Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Corso di Laurea Magistrale in Fisica

Simulated switching of the resting state functional connectivity in mouse brain using a real mesoscale connectome

Relatore:
Prof. Gastone Castellani

Presentata da:
Francesca Melozzi

Correlatori:
Dott. Christophe Bernard
Dott. Viktor Jirsa

Sessione III
Anno Accademico 2013-2014

Abstract

Capire come modellare l'attività del cervello a riposo, *resting state*, è il primo passo necessario per avvicinarsi a una reale comprensione della dinamica cerebrale.
Sperimentalmente si osserva che, quando il cervello non è soggetto a stimoli esterni, particolari reti di regioni cerebrali presentano un'attività neuronale superiore alla media. Nonostante gli sforzi dei ricercatori, non è ancora chiara la relazione che sussiste tra le connessioni strutturali e le connessioni funzionali del sistema cerebrale a riposo, organizzate nella matrice di connettività funzionale. Recenti studi sperimentali mostrano la natura non stazionaria della connettività funzionale in disaccordo con i modelli in letteratura.
Il modello implementato nella presente tesi per simulare l'evoluzione temporale del network permette di riprodurre il comportamento dinamico della connettività funzionale. Per la prima volta in questa tesi, secondo i lavori a noi noti, un modello di resting state è implementato nel cervello di un topo. Poco è noto, infatti, riguardo all'architettura funzionale su larga scala del cervello dei topi, nonostante il largo utilizzo di tale sistema nella modellizzazione dei disturbi neurologici. Le connessioni strutturali utilizzate per definire la topologia della rete neurale sono quelle ottenute dall'Allen Institute for Brain Science. Tale strumento fornisce una straordinaria opportunità per riprodurre simulazioni realistiche, poichè, come affermato nell'articolo che presenta tale lavoro, questo connettoma è il più esauriente disponibile, ad oggi, in ogni specie vertebrata. I parametri liberi del modello sono stati scelti in modo da inizializzare il sistema nel range dinamico ottimale per riprodurre il comportamento dinamico della connettività funzionale. Diverse considerazioni e misure sono state effettuate sul segnale BOLD simulato per meglio comprenderne la natura. L'accordo soddisfacente fra i centri funzionali calcolati nel network cerebrale simulato e quelli ottenuti tramite l'indagine sperimentale di Mechling et al., 2014 comprovano la bontà del modello e dei metodi utilizzati per analizzare il segnale simulato.

Abstract

Understanding how to model brain activity at rest, *resting state*, is the first necessary step towards a real comprehension of the cerebral dynamics.
Experimentally it is observed that, when the brain is not subject to external stimuli, characteristic networks of brain areas show neural activity greater than the mean. Despite despite the efforts of researchers, it is still unclear the link between structural connections and functional connections of the brain system at rest, organized in the functional connectivity matrix. Recent experimental studies show the non-stationary nature of the functional connectivity in disagreement with the models in literature.
The model used in the present thesis to simulate the network temporal evolution allow us to reproduce the dynamical behavior of the functional connectivity. For the first time in this thesis, at list from the literature we are aware of, a resting state model is implemented in the mouse brain. Actually little is known about large-scale functional architecture in mouse brain, despite the wide use of such system in the neurological disorders modeling. The structural connections used to define the network topology were obtained in the Allen Institute for Brain Science. The tool supplies an extraordinary chance to obtain realistic simulation since, as stated in the paper which presents the work, such connectome is the most comprehensive, available to date, in any vertebrate species. The free parameters of the implemented model were chosen in order to initialize the system in the optimal dynamical range to reproduce the non-stationary behavior of the functional connectivity. Several observations and measures were made on the simulated BOLD signals to better understand its nature. The agreement between the functional hubs calculated on the simulated brain network and the ones experimentally identified in the paper of Mechling et al., 2014, reveals the goodness of the implemented model and of the methods used to analyze the simulated signal.

Contents

Introduction

The human brain represents only the 2% of the body total mass but it consumes 20% of the body's energy, most of which is used to support spontaneous neural activity. Instead the task-related brain activity produces an increase in neuronal metabolism usually small ($< 5\%$) when compared with the large resting energy consumption [15]. Therefore, to understand how the brain operates, it is necessary to take into account the component that consumes most of the brain's energy: the *resting state* activity.
It was observed that such brain activity is not random but displays specific features, as the correlation in time of the BOLD signal between brain areas. Moreover the resting state activity it is not an human prerogative but it is observed also in non-human primate as rodents [24], [25].
Understanding how this activity emerges during rest is not a trivial problem since, in complex system like the brain, the collective outcome of system-wide dynamics is difficult to predict. Modeling studies [20], [18], [9] have demonstrated the importance of the interplay between the anatomical structure, the local neural dynamics and the noise in the emergence of resting state correlations described by the functional connectivity. Many models [18], [11], [13] operate at a working point close to the critical edge of instability; in this state the brain network is in a low firing stable state and it is continuously driven by the noise towards other configurations. Thus, according to such results, the resting state reflects the dynamical capabilities of the brain, which emphasizes the vital interplay of time and space.
Despite the progress in the discovery of potential mechanisms for explaining how a given anatomical structure produces dynamics at different temporal and spatial scales, the relationship between the anatomical connections in the brain and the resting-state functional connections is still not clear. One drawback in the empirical and simulated studies, is the assumption, often implicit, that the relationships between brain areas, shown by functional connectivity, are spatially and temporally stationary throughout the entire measurement period of BOLD signals. The groundlessness of such assumption was demonstrated by recently empirical studies as the one of Allen et al., 2014 [2], where

the authors observed that functional connectivity matrices computed over smaller time windows vary greatly during a scan session. In the work of Hansen et al., 2015 [19] the authors were able to reproduce the observed switching of the FC by using a model (*enhanced non-linearity mean-field model* - eMFM) that enriches the non-linearity of the dynamics of each brain region, in order to increase the number of attractors in the system.
Despite the large amount of mouse brain models used to understand human neurological disorders, little is known about the large-scale functional architecture of the mouse brain. This is the first study, at list from the literature we are aware of, where a resting state model is implemented in the whole mouse brain.
Moreover, unlike all cited works, where the anatomical wirings of the brain network are obtained with the DTI and/or DSI methods, in this work the structural connectivity matrix is an anterograde tracer mapping of the axonal projection in the mouse nervous system. The matrix is from the Allen Institute for Brain Science of Washington (USA) and it represents, to date, the most comprehensive connectivity map in any vertebrate species [28]. Using this careful biological realistic matrix, the resting state network is simulated with the eMFM in order to reproduce the functional connectivity switching. Several measures are performed on the simulated data to better understand the simulated system behavior. In particular, once defined the functional resting state networks, the hub regions of these networks are identified. The goodness of the comparison between such simulated functional hubs and the empirical ones as identified in the paper of Mechling et al., 2014 [25], give us confidence in our results and, more broadly, in the connectome based model approach.

This thesis is organized as follows. The first chapter introduces the resting state from the discovery to its possible, debated and fascinating, interpretations; moreover in the chapter the main principles of functional magnetic resonance imaging (fMRI) are summarized. The second chapter describes the concepts used in this work to model the brain activity. The neural population activity can be understood only by reducing the degrees of freedom of the system, and this is obtained by the mean field approximation; the most striking achievement in this regard is the reduction of a large population of spiking neurons to a distribution function describing their probabilistic evolution. In turn, this can be further reduced to a single variable describing the mean firing rate. This theoretical framework allows to understand the model used in the present work: the reduced Wong Wang model. The third chapter summarizes the state of the art of the resting state modeling. The fourth one illustrates the materials and methods used to simulate and to analyze the mouse brain resting state activity; it contains information about the structural connectivity matrix, the implementation of the computer simulations, the kind of analysis performed on the BOLD signals: the identification of the functional hubs and the decomposition. In the fifth chapter the results obtained are presented and discussed.

Part I

FRAMEWORK

Chapter 1

The resting state

The *resting state* is the spontaneous state of the brain i.e. the activity of the brain in absence of any external stimuli or other behaviorally salient events. Unlike the equilibrium state of an unperturbed noisy physical system, the spontaneous state of the brain does not show a trivial random activity, as was expected by the scientists until two decades ago. Actually there is a remarkable amount of coordinated activity taking place in the brain even in absence of any specific externally-driven task. Since the discovery of such striking behavior, interest has increased in trying to understand the dynamics of this resting condition, when we are supposedly not doing anything but our brain keeps active, organized and coordinated.

In this first chapter we provide a general outline of the resting state, starting from its history and ending with the current discussion on its interpretation; two sections deal respectively with the anatomy of the resting state network and its relevance for brain disease.

1.1 A brief history

The idea that the brain of humans exhibits ongoing activity regardless of the presence or absence of any observable behaviors is an ancient idea: as noted by Seneca in 60 A.D. in Epistulae morales ad Lucilium "The fact that the body is lying down is no reason for supposing that the mind is at peace. Rest is... far from restful".
Nevertheless the first clue that the brain activity persists during rest dates to the years following World War II from early studies of cerebral metabolism. Louis Sokoloff and colleagues (1948) noted that, although the human brain amounts to just 2% of the total

body mass, it consumes 20% of the body's energy; these measurements were made over brains in resting state. Seven years later the group adopted the same technique to ask whether cerebral metabolism changes globally when one goes from a quiet rest state to performing a challenging arithmetic problem. Surprisingly, metabolism remained constant; the local changes were too small (usually less than 5% compared with the resting energy consumption) to be detected by methods designed to measure the energy consumption of the brain as a whole.
Thus, the resting state is not truly a resting state at all.
Despite these considerations, for almost fifty years nobody studied the resting state: it was not clear how to undertake a study of intrinsic activity. This is why, nowadays, the vast majority of our knowledge about brain functioning comes from studies in which a task or stimulus is presented and the resulting neuronal activity changes are recorded as images. Neuroimaging practices, like the simple one explained in figure 1.1, were largely based on the assumption that ongoing activity is sufficiently random and then it can be averaged out in statistical analysis. Spontaneous brain activity was systematically overlooked.

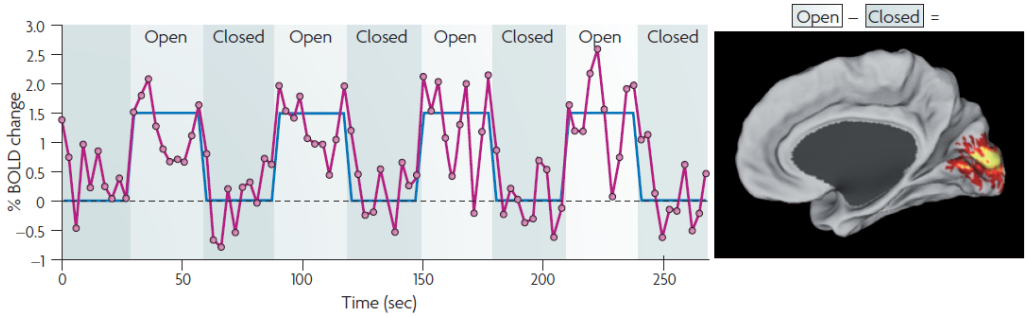


Figure 1.1: We can observe a change in the BOLD signal (magenta line) in a subject who is asked to perform a task as simple as closing and opening his eyes; by repeating the paradigm (blue line) several times, and subtracting the eyes closed condition from the eyes open condition, one can identify the BOLD signal intensity in the primary visual cortex (shown on the right). This successful experimental approach allows to build a topographical map of the brain related to functionality. However, an implicit assumption is taken: the spontaneous modulations of the BOLD signal not connected with the paradigm are regarded as *noise* and are minimized through averaging.
Source: Fox et al.,2007 [15] p.701

By the mid-1990s all the efforts of researchers were focused on understanding cognitive behavior; scans of resting state brain activity were often acquired across these studies

for mere control comparison, but researchers routinely began noticing that some brain regions showed more activity in resting state condition than during the execution of tasks. This progressively led to a radical change in the way of thinking the so-called *baseline state*, and would have shed new light on its mysteries.
The definitive recent event in the explication of the resting state came with the series of publication of Raichle, Gusnard and colleagues in 2001 [29]. In this study they individuate a set of brain regions, the Default brain Mode Network (DMN) as they call it, characterized by extraordinarily high metabolic rate during rest, and, on the other hand, by the greatest deactivation during externally imposed cognitive challenge. Their article made clear that DMN is to be studied as a fundamental neurobiological system, like the motor system or the visual system. It contains a set of interacting brain areas that are tightly functionally connected and distinct from other systems within the brain.
As pointed out in the study of Snyder et al., 2012 [32], Moreover convincing evidence of something equivalent to human DMN has been found in monkey and cat [32], rat [24] and mouse [25].

1.2 Method to investigate the resting state: fMRI

The main way used to reveal the manifestation of spontaneous neuronal activity is the fluctuation in the blood oxygen level dependent (BOLD) signal, revealed by the functional magnetic resonance imaging (fMRI). In the following the physical mechanism at the base of the magnetic resonance imaging and of the BOLD signal formation is briefly described.

1.2.1 Physical principles of MRI

The magnetic resonance imaging (MRI) is a non-invasive method to obtain images of the interior of structures as the human body. The method is based on the magnetic properties of materials composed of nuclei having a non-zero spin. Such nuclei, when placed in a magnetic field B_0, arrange themselves in the energetic levels according to the Boltzmann distribution and the total magnetization characterizes this order. After a perturbating pulse, which satisfies the resonance condition of the system, the magnetization realigns itself with B_0 through a characteristic time, in which nuclei make transitions to reestablish the equilibrium. The MRI detects the evolution of the system during the return to the equilibrium to obtain information about the system properties.

Let us consider an atomic nucleus with a non-zero total nuclear spin $\overrightarrow{I}$; the relation

between the magnetic moment μ and the spin is:

$$\overrightarrow{\mu} = \gamma\hbar\overrightarrow{I} \tag{1.1}$$

where γ is the magnetogyric ratio, which is typical for each nuclear isotope. Thus the component along the z direction of the magnetic moment is:

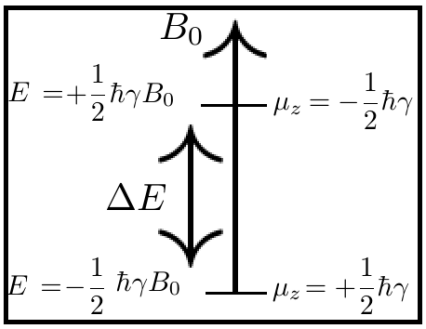
$$\mu_z = \gamma\hbar m \tag{1.2}$$

where m can take one of the $2I+1$ values in the interval $[-I, I]$.

As shown in the figure for the case of $I = {}^1/_2$, an homogeneous applied external magnetic field B_0 induces a splitting of the nuclear spin energy level:

$$\Delta E = \gamma\hbar B_0 \ . \tag{1.3}$$

Replacing the Planck-Einstein relation ($\Delta E = h\nu$) in the latter equation, it is simple to obtain the Larmor resonance frequency:



$$h\nu_0 = \frac{h}{2\pi}\gamma B_0 \qquad \Rightarrow \qquad \nu_0 = \frac{\gamma}{2\pi}B_0 \tag{1.4}$$

and the corresponding pulsation ω_0:

$$\omega_0 = \gamma B_0 \ . \tag{1.5}$$

We can not observe the motion of a single spin, but the collective motion of a set of N nuclei by means of the total magnetization $\overrightarrow{M} = N\langle\overrightarrow{\mu}\rangle$. The evolution in time of the magnetization $\overrightarrow{M}$ of a set of nuclei placed in the a magnetic field $\overrightarrow{B_0}$ is:

$$\frac{d\overrightarrow{M}}{dt} = \gamma\overrightarrow{M} \times \overrightarrow{B_0} \ . \tag{1.6}$$

The equation describes the precession of $\overrightarrow{M}$ around $\overrightarrow{B_0}$ at the angular velocity ω_0, only when $\overrightarrow{M}$ is not aligned with $\overrightarrow{B_0}$. At the equilibrium, the total magnetization of a paramagnetic material placed in a magnetic field $\overrightarrow{B_0}$, shares the same direction of $\overrightarrow{B_0}$ as stated by the Curie's law:

$$\overrightarrow{M_0} = C\frac{\overrightarrow{B_0}}{T} \tag{1.7}$$

where T is the absolute temperature and C is the Curie constant that depends on the material characteristics. In the following, for the sake of simplicity, $\overrightarrow{B_0}$, and thus $\overrightarrow{M_0}$, is considered aligned with the z axis. Applying a magnetic field $\overrightarrow{B_1}$ orthogonal to $\overrightarrow{B_0}$

and with frequency ν_0, the magnetization vector moves away from the z axis; the angle between z axis and the new position of the magnetization vector depends on the duration of radio frequency (RF) field $\overrightarrow{B_1}$ applied; in MRI the RF pulse is generated by a coil. At the end of the pulse application, the spin precession on the transverse plane induces an oscillatory electromotive force in the coil by electromagnetic induction, that originates a current in the probe. The detected signal is called *Free Induction Decay* (FID), which has an oscillating trend with exponential decaying, and it is originated by the photons in the radio-wave range emitted by the nuclei returning to the equilibrium.
After the RF pulse, the deterioration of the NMR signal is analyzed in terms of two separate processes, the longitudinal one and the trasverse one, each with their own time constants.

The longitudinal relaxation time T_1

The underlying process that leads the longitudinal component of the magnetization (along z) to reach its equilibrium value M_0, is the redistribution of nuclear spin populations according to the Boltzman distribution; such process takes place by energy exchanges between the nuclei and the surroundings.The longitudinal component of the magnetization decreases in time according to:

$$\frac{dM_z(t)}{dt} = -\frac{(M_z(t) - M_0)}{T_1} \tag{1.8}$$

and thus:

$$M_z(t) = M_z(0)e^{-t/T_1} + M_0(1 - e^{-t/T_1}) \ . \tag{1.9}$$

The transverse relaxation time T_2

The underlying process that leads the trasverse component of the magnetization to reach its equilibrium value, i.e. zero, is the decoherence of the transverse nuclear spin magnetization. Random fluctuations of the local magnetic field lead to random variations in the instantaneous NMR precession frequency of different spins. As a result, the initial phase coherence of the nuclear spins is lost and the total xy magnetization is null. The transverse component of the magnetization decays to zero in time according to:

$$\frac{dM_{xy}(t)}{dt} = -\frac{M_{xy}(t)}{T_2} \tag{1.10}$$

and thus:

$$M_{xy}(t) = M_{xy}(0)e^{-t/T_2} \tag{1.11}$$

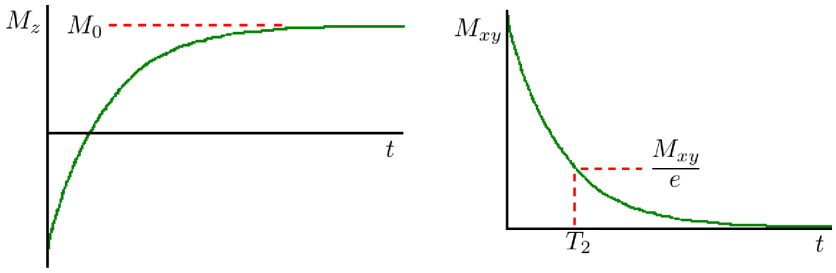


Figure 1.2: Evolution in time, after the RF pulse, of the longitudinal magnetization (left) and of the transverse magnetization (right) in the rotating frame.

1.2.2 Functional magnetic resonance imaging of the brain

Functional magnetic resonance imaging, fMRI, is a non-invasive way to assess the organ functionality using MRI signal changes. fMRI can be used to measure, indirectly, brain activity. The method is able to provide measurement of the consequence of neural activity in the energy metabolism of the brain; more precisely it is related to the level of oxygenation within the corresponding vasculature. Brain areas activation produces physiological changes in blood vessel resulting in a decreasing of deoxyhaemoglobin concentration. This effect is referred to as BOLD (blood oxygen level dependent) effect, that is the common used method for brain fMRI studies and manifests itself as a slight increase in the T_2 weighted MR images.
Let us examine the brain energy metabolism characteristics that lead to the BOLD signal formation. The activation of a brain region requires energy in form of oxygen (0_2) and glucose, which will be transformed in ATP. Since the brain does not store the glucose and the oxygen, such elements are transported, when required, by blood flows. The hemoglobin molecule (Hb) is able to transport four 0_2 units. The Hb molecule is composed by 4 units; at the center of each unit there is a Fe^{2+} atom, which is responsible of the link with the 0_2. The link or the separation between O_2 and Hb produces a change in the magnetic nature of the hemoglobin molecule. When Hb is not linked to 0_2, deoxyhemoglobin, the 4 Fe^{2e+} atoms are in the higher spin state ($S = 2$); consequently, Hb molecules are attracted by any externally applied magnetic field: Hb is paramagnetic. On the other hand, when 0_2 is bound to Hb (oxyhemoglobin, HbO_2) the two electrons of the Fe^{2+} are paired and, consequently, they are in the state of lower energy with a spin state $S = 0$: the $Hb0_2$ is diamagnetic. Since the deoxyhemoglobin is paramagnetic, it is able to reduce the NMR signal in the images weighted in T_2; indeed the rate of loss of proton spin phase coherence, measured through T_2, can be modulated by the presence of intravoxel deoxyhaemoglobin. Instead, the oxyhemoglobin, which is diamagnetic, does

not modify the NMR signal. During the neural activation of a brain area, there is an higher incoming blood flux with respect to the blood incoming flux during rest; in such area blood vessels expand and the brought-in oxygen is more than the oxygen consumed in burning glucose. Therefore, although paradoxical, in the activate brain region the concentration of oxygenated blood increases, and the concentration of deoxygenated blood decreases respect to the neighbour, non-active, brain areas. In conclusion, when a brain area is active it is possible to observe an increment in the NMR signal from that area with respect to the surrounding brain regions, due to the presence of an higher concentration of oxygenated blood. Such process is schematically shown in figure 1.3.
The change in the NMR signal from neuronal activity is called the hemodynamic response. The mathematical framework used in this work to model the BOLD signal is explained in appendix B.

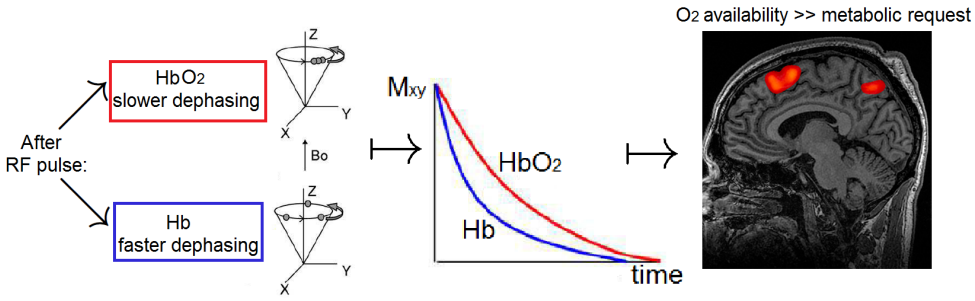
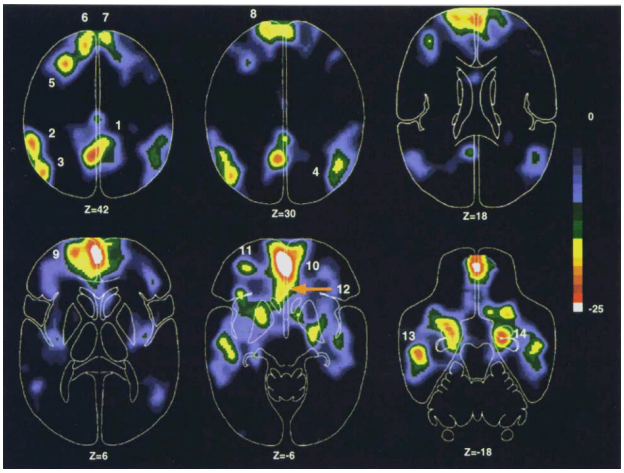


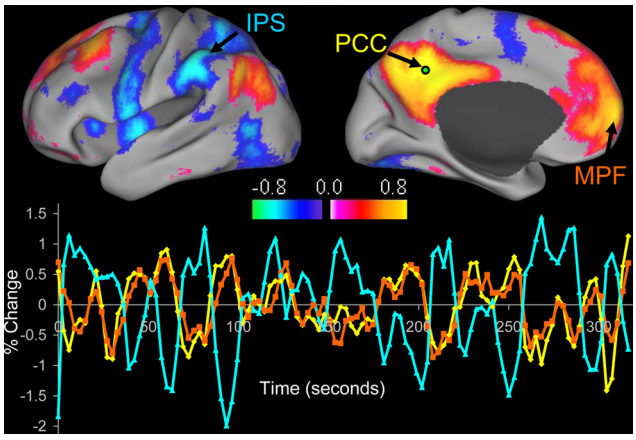
Figure 1.3: In figure is shown a schematically representation of the mechanism that lead to the BOLD signals formation.

1.3 Anatomy of the default mode network

The anatomy of the brain's default network has been characterized using multiple approaches. The DMN was originally identified with all those regions that showed a significant increase in neural activity during *passive* task states with respect to a wide range of *active* task states (Shulman et al., 1997 [31], figure 1.4a).
In more recent times, the preferred approach (and the one used in this work) moves its steps from the notion of *functional connectivity*: the temporal correlation in the recorded BOLD activity using fMRI data.
In practice, the resting state *Functional Connectivity* (FC) is a matrix whose each entry a_{ij} is the correlation (generally the Pearson correlation) in time between the intrinsic activity of the neural source i and the neural source j (figure 1.4b, see also section 3.1).



(a) *Source:* Shulman et al., 1997 [31] p. 651



(b) *Source:* Fox et al., 2005 [14] p.9674

Figure 1.4: Default mode network investigated with different techniques.
In figure 1.4a there are regions of the brain regularly observed to decrease their activity during attention demanding cognitive tasks. These data represent a meta-analysis of nine functional brain imaging studies performed with PET and analyzed by Shulman and colleagues; 132 individuals contributed to the data in these images. Such decreases appear to be largely task independent.
The different colors in the upper part of the figure 1.4b indicate regions of a single subject's brain that are correlated (positive values) and anticorrelated (negative values) during resting fixation in a functional MRI study by Fox et al., 2005 [14]. The BOLD flow in time of the posterior cingulate/precuneus region (PCC, yellow), almost overlaps with the flow of a region in the medial prefrontal cortex (MPF, orange) (*positive correlation*) and it is almost opposite to the flow of another region in the intraparietal sulcus (IPS, blue) (*negative correlation*).

Using the fMRI data and the FC analysis many neuroanatomical-systems have been proved to be coherent in their spontaneous activity including the somatomotor, the visual, auditory, hippocampus or episodic memory, language, dorsal attention and others (Fox et al., 2007 [15]). A meaningful finding is that regions with similar functionality -that is, regions that are similarly modulated by various task paradigm - tend to be correlated in the BOLD spontaneous activity. On the other hand regions with opposing functionality have been found to be negatively correlated in their spontaneous activity. Importantly, it was noted that resting state networks reflect the structure of the connections between brain regions. The information on the anatomical wiring of the brain are encoded in the *Structural Connectivity* (SC) or *anatomical connectivity* that is a matrix whose elements are a_{ij} are the weight of the connection between the region i and the region j. The SC is generally measured by the Diffusion Tensor Imaging (DTI) an MRI technique that exploits the restricted diffusion of water through myelinated nerve fibers linking connected areas (for more details see section 3.1). The DTI-derived structural connectivity can be compared to functional connectivity obtained by fMRI imaging. The comparison showed that although structural connectivity is a good predictor of functional connectivity - if there is a direct anatomical connection there is a functional connection - the opposite is not necessarily true. This work puts the stress on the importance to go over the anatomical connectivity in order to predict the realistic functional connectivity.

1.4 Resting state in brain disease

Most, if not all, physiological and psychiatric diseases have disrupted large-scale functional and/or structural properties [21]. Whether they are the cause or consequence of the disease is unclear, but it was observed that, in these case, brain neural populations exhibit significant changes in dynamic properties (for the definition of neural population see section 2.3); such fact may underlie many of the observed dysfunctions. Quantification of disrupted dynamics in neural populations may lead to a better understanding of the disorder, more targeted drug treatment, and eventually, diagnostic or prognostic indicators. As analyzed in the paper of Buckner et al., [6] there is a relationship between resting state network and mental disorder; the disorders where this link is most evident are autism, schizophrenia and Alzheimer's disease. Others for which hypotheses on the role of DMN have been advanced include depression, obsessional disorders, attention disorders and post-traumatic stress disorder.
Moreover, resting state networks take time to emerge in human beings, as it appears from experiments: it is not present in infant and start being shaped during childhood; changes could be observed in development and aging.

1.5 Interpreting the resting state

Numerous experiments have been performed to investigate the connections between the resting state network and cognition: the observation that the DMN is most active during passive cognitive state, when thought is directed towards internal channels (referred reflections, autobiographical memory, moral decision making, self judgments, daydreaming), encourages consideration of the possibility that the DMN is associated with spontaneous cognition.
Nevertheless, experimental evidences show that unconstrained cognition alone does not account for the greatest part of the intrinsic activity although it undoubtedly contributes a small increment. Indeed when subjects are asked to rest quiet in the scanner they perform mental tasks that result in changes in neural activity (cerebral metabolism studies performed by Sokoloff and colleagues, and discussed in first section); if these uncontrolled tasks are responsible for the coherence pattern observed, then the resting state studies may be no more interesting than controlled task-activation studies.
The *first* reason to reject the hypothesis of the unconstrained cognition as responsible for the DMN activation is that imposed tasks evoke response that are modest in magnitude in comparison to intrinsic activity; there is no reason to suppose that unconstrained thoughts are more energy demanding than constrained ones (Snyder et al., 2012 [32]).
The *second* reason is that the DMN persists during anesthesia and the early stages of sleep, states during which cognition is absent or at least very attenuated.
The relation between the DMN and the brain areas involved in self referential 'tasks' highlights a well know problem in neuroimaging, namely the problem of reverse inference: a certain cognitive operation activates a brain region but if the same brain region is shown in activity it does not necessarily mean that a cognitive process is occurring.
An interesting interpretation of the resting state network is the one proposed by Deco, Jirsa and McIntosh 2011 [10]; they link brain dynamics at rest to a constant inner state of exploration during which the brain makes prediction about the likely configuration for a given impeding input. They explained this concept with an effective metaphor in which they compared the brain at rest to a tennis player waiting for the service of his adversary; during the waiting the player is not static but continues to move with small lateral jumps to be able to react more effectively to the incoming service. The inner brain state exploration is driven by physiological characteristics of the brain such noise, delays in conductions, anatomical connections and intrinsic dynamics; these concepts will be deeper analyzed in this work.

Chapter 2

Modeling the brain

Biological systems are among the most complex systems in nature and the number of components of such systems is enormous; for example, the human brain contains 10^{11} neurons. Hence, in order to understand the brain dynamics, we need to make simplifications and to construct theoretical representations.
With reference to the spatial scale, the brain models could be divided as follows:

> **Micro**-scale models which take into account the way of exchanging information between the computing elements of the brain: the neurons.
>
> **Meso**-scale models which describe the dynamics of neural population.
>
> **Macro**-scale models which give information about the whole brain dynamics and about the interactions between large-scale neural systems, such as cortical regions.

The main goal of the chapter is to understand the reduced Wong Wang model, the macro-scale model used in this work. In order to reach this purpose, after a brief general description of the neuron, three neuron models are described: the leaky integrate-and-fire model, the Hodgkin-Huxley model and a simplification of the Hodgkin-Huxley model; the last one allows us to discuss the saddle node bifurcation. In the next sections the idea at the base of the large scale models are summarized: once defined the concept of neural mass and neural activity, the mean field approximation will be derived starting from the the Chapman-Kolmogorov equation.
Finally, once acquired the necessary tools, the reduced Wong Wang model is derived from the original paper of Wang 2002 [36].

2.1 The neuron

The basic unit of the nervous system is the neuron, an electrically excitable cell that processes and transmits information through electrical and chemical signals.
There are different kinds of neurons, and every neuron has many internal features; so, considering that even the smallest part of the simplest neuron is already a complex system, in the next sections the characteristics of the neuron, useful for the purpose of this work, are pointed out.

2.1.1 The structure

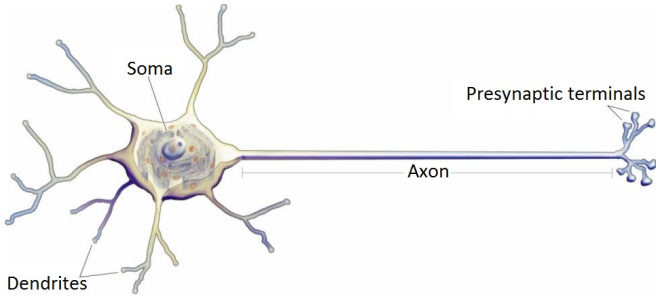


Figure 2.1: Schematic representation of a neuron with its principal components.

A "typical" neuron is divided into four distinct regions: the soma, the dendrites, the axon and the presynaptic terminals.
The *soma* is the cell body; it contains the nucleus and the hereditary material, which directs the operations of the neuron.
The *dendrites* are thin structures arising from the soma and extending for hundreds of micrometers; they branch multiple times, getting thinner at each branching and so giving rise to a complex "dendritic tree". This is where the input to the neuron occurs via the dendritic spine.
The *axon* is a tubular appendage of the soma. It is different from dendrites in structure and in function. Regarding the structure, it is longer (even up to 1 meter in human and even more in other species) and it maintains the same diameter as it extends. Also, unlike the dendrites, the axon doesn't receive signal, but it conducts the input away from the cell.
The swollen ends of the axon's branches are called *presynaptic terminals*; they transmit information from one neuron to another.

2.1.2 The electrical properties

The key of neural function is the synaptic signaling process which is partly electrical and partly chemical. The electrical aspect depends on properties of the neuron's membrane. Like all animal cells, the cell body of the neuron is enclosed by a plasma membrane, i.e. a bilayer of lipid molecules which acts as a powerful electrical insulator; in the membrane there are the *ion channels*, a types of integral membrane proteins that are capable of actively or passively moving specific types of ions from one side of the membrane to the other.
Since between the outside and the inside of the membrane there is a difference in charge, we can define the *membrane potential* as the difference between the potential inside the cell V_{in} and the potential outside the cell V_{out}:

$$V_m = V_{in} - V_{out} \, . \tag{2.1}$$

When a neuron is at rest there is an excess of positive charge on the outside with respect to the inside of the membrane; hence the *resting membrane potential* V_L is negative and it ranges from -60 mV to -70 mV.
The nerve cell peculiarity is the *electrical excitability*. In electrically excitable cell a sufficiently large depolarization of the membrane potential can evoke an action potential, during which the membrane potential changes quickly. Action potentials are generated by the activation of certain voltage-gated ion channels in consequence of a stimulation.

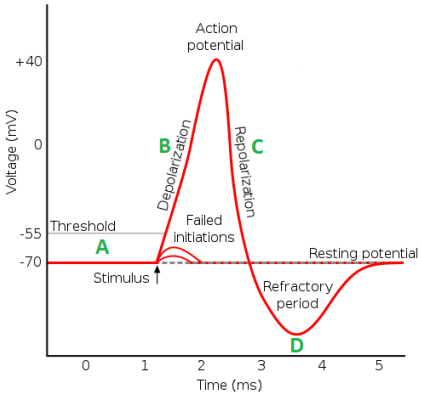


Figure 2.2: Time course of a neuron's action potential in response to a stimulus

The shape of the action potential is stereotyped and its time analysis clarifies the process. With reference to figure 2.2 in what follows the time intervals are described:

A. Before the stimulus, the neuron is at rest: the membrane potential is fixed at the resting membrane potential (V_L=[-70,-60] mV).

B. As effect of the stimulation, the Na^+ channels open and the Na^+ ions move into the axon, causing the depolarization. If the membrane reaches the threshold potential the action potential is generated.

C. Repolarization occurs when the K^+ channels open and K^+ ions move out of the axon. This creates a polarity change between the outside and the inside membrane. The impulse travels down the axon in one direction only, going towards the axon terminals where it signals to other neurons.

D. After the action potential has occurred, there is a transient negative shift, called the *refractory period* due to additional potassium currents. This is the mechanism that prevents an action potential from traveling back the way it just came; in this period of time an other action potential can not be generated.

Neurons have three passive electrical properties that are important to electrical signaling: intracelluar resistance along axons and dendrites, membrane capacitance and membrane resistance.

2.1.3 Synapses and synaptic receptors

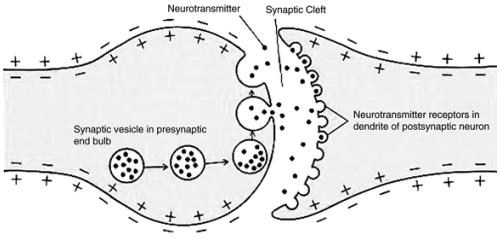


Figure 2.3: A schematic representation of the basical mode of operation of a chemical synapse.

The site where one neuron communicates with another is known as *synapse*. The part of the neuron that emits the signal is called the *presynaptic terminal*, and the part of the neuron that receives the signal is known as the *postsynaptic terminal*. There are two basic forms of synapses: chemical and electrical.
In order to understand the Wong Wang model utilized in this work and discussed in section 2.4, the mode of operation of the chemical synapses is briefly analyzed in a schematic way.

↪ Chemical synapse begins when an action potential reaches the presynaptic terminal.

↪ The action potential activates voltage-dependent Ca^{2+} channels, which allow calcium ions to enter the cell from the tissue fluids outside the neuron.

↪ In the presynaptic terminal the rise in the concentration of Ca^{2+} causes the migration of synaptic vesicles towards the presynaptic membrane.

↪ When a synaptic vesicle reaches the presynaptic membrane, both membranes fuse, and, consequently, the vesicle content is released in the space that separates the two neurons (the synaptic cleft).

↪ The released neurotransmitters diffuse across the synaptic cleft and bind to receptors in the postsynaptic membrane.

↪ The binding of neurotransmitters to their receptors on the postsynaptic cell membrane leads to the opening of the ion channels.

↪ The resulting ion flux modifies the conductance of the postsynaptic cell, thereby the transmission of the signal from one neuron to another is completed.

Synapses can be classified as excitatory or inhibitory. If the membrane potential of the postsynaptic cell increases, the postsynaptic neuron is driven towards its excited state, and the synapse is considered excitatory. On the other hand, inhibitory synapse is associated with a decrease in membrane potential, which drives the postsynaptic neuron towards its resting state membrane potential.

The major excitatory and inhibitory transmitters in the brain are Glutamate and GABA (-aminobutyric acid).
The three principal receptors are AMPA, NMDA and GABA.
The AMPA current activates and deactivates rapidly. In contrast, the NMDA current activates and deactivates slowly and it has a voltage dependence controlled by the extracellular magnesium concentration.
The GABA is the principal inhibitory neurotransmitter.

2.2 Dynamics of a single neuron

In this section are presented three models describing the dynamics of a single neuron: the *Leaky Integrate and Fire* model (LIF model), the *Hodgkin-Huxley* model and a modified version of the Hodgkin-Huxley model, which has a bifurcation diagram similar to the one of the reduced Wong Wang model (section 3.3.2). In this work the LIF model is treated

before the Hodgkin-Huxley model, although the LIf model is more recent and it is a simplification of the Hodgkin-Huxley model; it has been preferred following the sequence that make easier the understanding instead that following the chronological order.

2.2.1 The leaky integrate and fire model

Modeling a neuron as an electric circuit was firstly investigated by Lapicque over 100 years ago, and his model is known today as *integrate-and-fire* model or *single compartment* model. Lapicque's idea captures two of the most important aspects of neuronal excitability: the integration of the incoming signals and the generation of the spike once a certain threshold is exceeded.
This is obtained by considering the neuron an electric circuit as the one shown in figure 2.4 (without the resistor) and by describing the variation in time of the membrane potential with a single variable V:

$$C\frac{dV}{dt} = \frac{dQ}{dt}\,. \tag{2.2}$$

When an input current is applied, the membrane voltage increases with time until it reaches a constant threshold V_{th}; at this point a delta function spike occurs, the switch closes and shunts the capacitor that resets the voltage to its resting potential V_L.

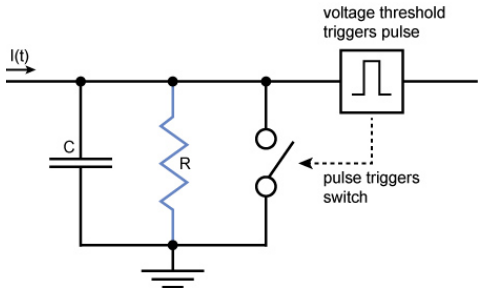


Figure 2.4: Integrate-and-fire circuit: the integrate-and-fire model consists of a capacitor, a threshold detector and a switch (without the resistor). Once the voltage reaches the threshold, the spike is fired and the switch is closed to shunt the capacitor. In the leaky version, a resistor is added in order to slowly drain the capacitor. This corresponds to leakage current through a membrane in a living cell.

The main shortcoming of the integrate and fire model is that it has no time-dependent memory. If the circuit receives a below-threshold signal, the voltage boosts forever until it fires again. This characteristic clearly does not reflect the observed neuronal behavior. In the *leaky integrate-and-fire model*, the memory problem is solved by adding a "leak"

term to the membrane potential, reflecting the diffusion of ions that occurs through the membrane when the cell does not reach the proper balance. In the previous electrical circuit, the LIF model is implemented adding a resistor in parallel with the capacitor (figure 2.4 with R); hence the equation 2.2 can be written as:

$$C\frac{dV(t)}{dt} = -g[V(t) - V_L] + \frac{dQ(t)}{dt} \tag{2.3}$$

where g is the conductance ($g = 1/R$).
Then a neural network of N_n spiking neurons can be modeled only adding to the previous equation the subscript i to denote the i-th neuron. Multiplying the equation by the specific membrane resistance R, and denoting with $\tau = C/R$ the membrane time constant, it is obtained:

$$\tau\frac{dV_i(t)}{dt} = -[V_i(t) - V_L] + RI_i(t) \tag{2.4}$$

where $I_i(t)$ is the total synaptic current flow into the cell i and it is the sum of the contributions of the δ-spikes produced by the presynaptic neurons. Considering N synapses going into a neuron i, and denoting with J_{ij} the efficacy of the synapse j, the total synaptic afferent current is given by:

$$RI_i(t) = \tau\sum_{j=1}^{N} J_{ij}\sum_{k}\delta(t - t_j^{(k)}) \tag{2.5}$$

where $t_j^{(k)}$ is the emission time of the k^{th} spike from the j^{th} presynaptic neuron.

2.2.2 The Hodgkin-Huxley model

The LIF model, described in the previous section, is a simplification, much more computationally tractable, of the Hodgkin-Huxley model.
In 1952, Alan Lloyd Hodgkin and Andrew Huxley developed their model to explain the ionic mechanisms that underlies the initiation and propagation of action potentials in the squid giant axon. For this work they received in 1963 the Nobel Prize in Medicine. A relevant explication of this model is given in the first chapter of the book of Jirsa and McIntosh [4] that will be briefly summarized in this section.
As in the LIF model, also in the Hodgkin-Huxley (HH) model the neuron is treated as an electrical element, but unlike the LIF model, here there are two voltage dependent electrical conductances associated with the voltage-gated ion sodium (Na^+) and potassium (K^+) and one linear conductance associated with the static chloride (Cl^-) channels.

Hence the circuit temporal evolution is given by:

$$C\frac{dV(t)}{dt} = I - g_{Na}(V(t))[V(t) - V_{Na}] - g_K(V(t))[V(t) - V_K] - g_L[V(t) - V_L] \tag{2.6}$$

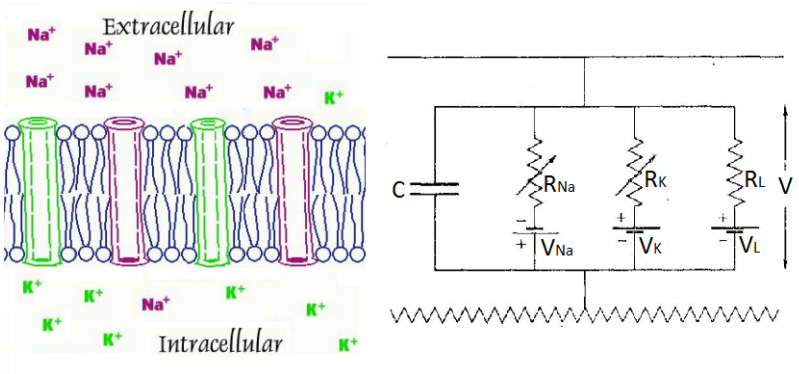


Figure 2.5: In the left part of the figure there is a schematic representation of the neuron membrane. When the Na^+ channels (magenta) open, the Na^+ ions flow inside the cell, depolarizing the neuron; instead when K^+ channels (green) open, K^+ ions flow outside the cell, hyperpolarizing the neuron.
In the right part of the figure there is the electric circuit used to model the neuron membrane: the Na^+ and K^+ channels are represented by two electric voltage dependent resistences, respectively R_{Na} and R_K; the linear resistence R_L represents the channels that are not voltage dependent.

where V_{ion} are the respective Nernst potentials[1]; the dependence on the voltage of the conductances can be transported as a dependence on the ion flows (which depend on the voltage):

$$g_{ion}(V) = \langle g_{ion} \rangle \, f_{ion}(V) \tag{2.8}$$

where $\langle g_{ion} \rangle$ represents the maximal conductance for the particular ion channel.
The Na and K ion flows reflects the state of the activation and inactivation channels, which, respectively, open and close as membrane voltage increases. These are given by:

$$\begin{aligned} f_{Na}(V) &= m(V)^M h(V)^H \\ f_K(V) &= n(V)^N \end{aligned} \tag{2.9}$$

where m and n are activation channels for Na and K, and h is the single inactivation channel for Na. The exponents are determined by the number of such classes of channel $M = 3$, $H = 1$ and $N = 4$. Hence equation 2.6 reflects the combined flow of all ion species as they are "pushed through" open channels according to the gradient between

[1] The expression of the Nernst potential is

$$V_{ion} = \frac{RT}{zF} ln \left(\frac{[\text{ion outside the cell}]}{[\text{ion inside the cell}]} \right) \tag{2.7}$$

the membrane and the Nernst potential.
To understand the process it is useful to notice that the Na and K channels act like a first order chemical reaction with two states; then this can be written as:

$$close \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} open \tag{2.10}$$

where the factors α and β are called the transition rate constants. α is the number of times per second that a channel in the closed state opens, while β is the number of times per second that a channel in the open state closes. All the gates within a particular class have the same value of α and β, but gates which belong to different classes may have different values of the transition rate constants: this gives the different classes their different properties.
Then for a whole generic population of gates (Na or K), it is possible to define the fraction of gates P that are in the open state (where P varies between 0 and 1) and consequently the fraction of channels *(1-P)* that are in the closed state. Therefore in a given time:

$$\begin{aligned} &Fraction\ of\ gates\ opening = \alpha(1-P) \\ &Fraction\ of\ gates\ closing = \beta P\ . \end{aligned} \tag{2.11}$$

The system is in the equilibrium state when the fraction of the gates in the open and in the closed state does not change; since the rate constants are voltage dependent, this situation occurs when the voltage is kept constant for a relatively long period of time. P reaches the steady state when the fraction of the gates open is equal to the fraction of the gates closed:

$$\alpha(1-P_\infty) = \beta P_\infty \quad \Rightarrow \quad P_\infty = \frac{\alpha}{\alpha+\beta}\ . \tag{2.12}$$

The rate at which P achieves a new value, following a change in voltage and consequently a change in α and β, is equal to the difference in the rate of channels closing and opening:

$$\frac{dP}{dt} = \alpha(1-P) - \beta P \tag{2.13}$$

The solution of the previous equation is:

$$P = P_\infty - (P_\infty - P)e^{-t/\tau} \quad with: \quad \tau = \frac{1}{\alpha+\beta}\ . \tag{2.14}$$

Putting in such equation the expression found for α and β by Hodgkin and Huxley, and remembering that m is the activation channel for Na, the kinetics of activation and inactivation of these channels are determined by:

$$\frac{dm(V)}{dt} = \frac{(m_\infty(V) - m(V))}{\tau_m}\ . \tag{2.15}$$

The kinetic of h (the inactivation channel for Na) and n (the activation channel for K) are of the same kind, although their rate constant τ are obviously distinct.
The form of the steady state configurations of ion channel populations as a function of the membrane potential is sigmoid shaped of the form:

$$m_\infty(V) = \frac{m_{max}}{1 + e^{(V_m - V)/\sigma}} \tag{2.16}$$

where V_m is the threshold potential for the ion channel and σ introduces the variance of this threshold.
The membrane dynamics discussed in this section is summarize in figure 2.6.

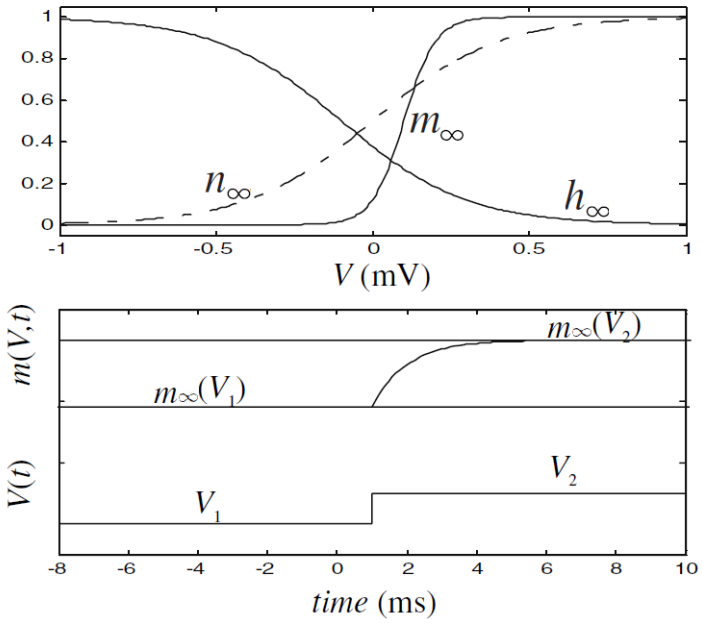


Figure 2.6: In the upper part of the figure is shown the sigmoid relationship between the membrane potential and the steady state conductances as stated by equation 2.16.
In the lower part of the figure it is displayed the changing in the transmembrane conductance according to the equation 2.15 following a discrete change in the transmembrane potential (lower line).
Source: Adapted from the book of Jirsa and McIntosh [4] p. 28

2.2.3 The saddle node bifurcation in a simplificated of the Hodgkin-Huxley model

The Hodgkin-Huxley model is a beautiful juncture of empirical and mathematical analysis that offers an explanation of neural firing and that captures quantitatively the complex shape of a neural depolarization. However, much of the qualitative behavior can be captured by good approximations of the model like the one explicated in the third chapter of the book of Izhikevich "Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting" [22], and summarized in this section. Furthermore the selected approximation permits to introduce the complex dynamics presented in the reduced Wong Wang model.

The Hodgkin-Huxley model can be reduced to one-dimensional system when all transmembrane conductances have fast kinetic. For the sake of illustration, let us consider a membrane having a leak current and a fast voltage-gated current, last one associated with the sodium channel:

$$C\frac{dV}{dt} = I - g_{Na}m(V)[V - V_{Na}] - g_L[V - V_L] \tag{2.17}$$

$$\frac{dm(V)}{dt} = \frac{(m_\infty(V) - m(V))}{\tau_m(V)}\,. \tag{2.18}$$

The assumption at the base of this approximation is that the gating kinetic (equation 2.18) is much faster than the voltage kinetic (equation 2.17), which means that the time constant $\tau_m(V)$ is very small ($\tau_m(V) << 1$) in the entire biophysical voltage range. Then it is possible to assume that fast sodium channels instantaneously reach their steady state (m_∞); therefore the two dimensional system (equation 2.17 and 2.18) can be reduced as follows:

$$C\frac{dV}{dt} = I \overbrace{-g_{Na}m_\infty(V)[V - V_{Na}]}^{\text{istantaneous I}} - g_L[V - V_L] \tag{2.19}$$

where the steady states m_∞ is given by:

$$m_\infty = \frac{m_{max}}{1 + e^{(V_m - V)/\sigma}}\,. \tag{2.20}$$

In order to make the next observations more clear, it is useful to introduce the experimental parameter values (table 2.1) obtained using whole-cell patch clamp [22].

Parameters		Values
Membrane capacitance	C	10 μF
Leak conductance	g_L	19 mS
Sodium conductance	g_{Na}	74 mS
Leak potential	V_L	-67 mV
Sodium potential	V_{Na}	60 mV

Table 2.1: Parameters used in the next calculations.
Source: book of Izikecih 2005 [22]

Attractors, stability and saddle node bifurcation

Before analyzing the complex dynamics of the previous system, let us define some basic concepts of continuous one-dimensional dynamical system, that can be written in the form:

$$\frac{dV}{dt} = f(V, I)\,. \tag{2.21}$$

The equilibrium points or attractors are the values of the state variables for which the state variable V does not change: $dV/dt = f(V, I) = 0$. A sufficient condition for the stability of an equilibrium point is that, in this point, the derivative of the function f with respect to V is negative (provided that the function is differentiable); instead positive slope of $f'(V_{eq})$ implies instability.

In general the one dimensional system 2.21, that has an equilibrium point $V = V_{sn}$ for some value of the parameter $I = I_{sn}$, is said to be at a *saddle-node bifurcation* when the following mathematical conditions are satisfied:

1. *Non-hyperbolicity*: the eigenvalue λ at V_{sn} is zero:

$$\lambda = \frac{\partial f(V, I_{sn})}{\partial V} = 0 \qquad (\text{at} V = V_{sn})\,. \tag{2.22}$$

2. *Non-degeneracy*: the second order derivative with respect to V at V_{sn} is non-zero:

$$\frac{\partial^2 f(V, I_{sn})}{\partial V^2} \neq 0 \qquad (\text{at} V = V_{sn})\,. \tag{2.23}$$

3. *Transversality* The function *f(V,I)* is non-degenerate with respect to the bifurcation parameter I:

$$\frac{\partial f(V_{sn}, I)}{\partial I} \neq 0 \qquad (\text{at} I = I_{sn})\,. \tag{2.24}$$

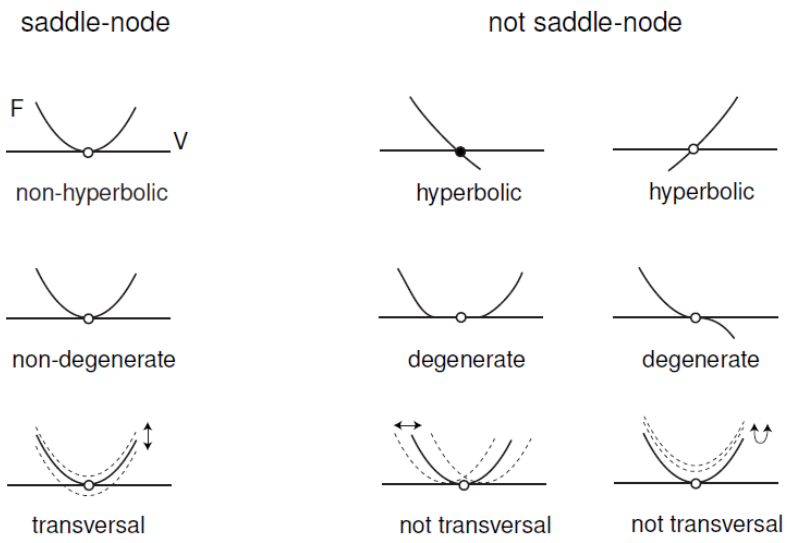


Figure 2.7: Geometrical meaning of the three conditions defining saddle node bifurcations.
Source: Izhikevich 2005 [22] p. 77

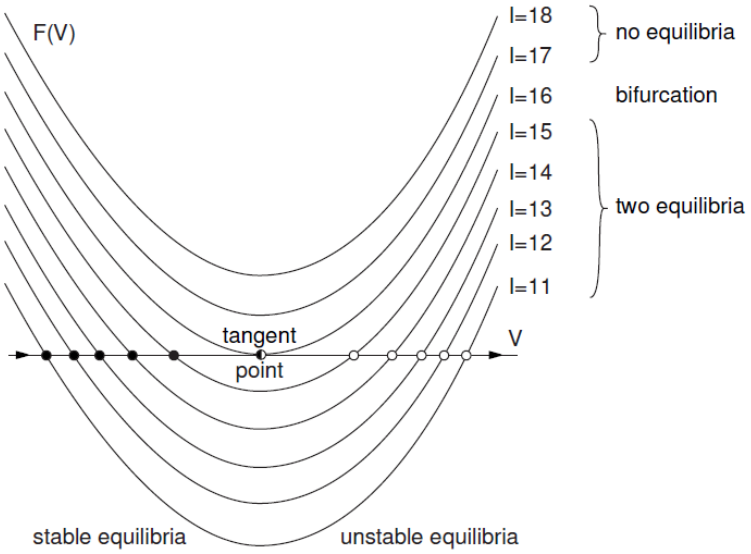


Figure 2.8: In the figure is displayed the saddle node bifurcation for the system 2.21: as the graph of *f(V)* is lifted up, the stable and unstable equilibria approach each other, coalesce at the tangent point, and then disappear.
Source: Izhikevich 2005 [22] p. 77

These conditions are illustrated in figure 2.7. Saddle node bifurcation consists in appearance or disappearance of a pair of equilibrium points as shown in figure 2.8. To draw the *bifurcation diagram*, it is necessary to determine the locations of the stable and unstable equilibria for each value of the parameter I and plot them in the (I, V) plane.

Saddle node bifurcation in the previous model

Let us studying the complex dynamics of the simplification of the Hodgkin-Huxley model (equation 2.19) when the bifurcation parameter is the injected dc-current. The equilibria are the zeros of the equation:

$$C\frac{dV}{dt} = I - \overbrace{[g_L(V - V_L) + g_{Na}m_\infty(V)(V - V_{Na})]}^{I_\infty(\mathrm{V})} . \tag{2.25}$$

In other words, the membrane potential V is at the equilibrium point if and only if the net membrane current $I - I_\infty(V)$ is zero.
In figure 2.9 there is the geometrical representation of this equation: any intersection between the steady-state I-V curve $I_\infty(V)$ and the injected dc curve (horizontal line) is a stable or unstable equilibrium; when I increases past 16 pA the bifurcation occurs.

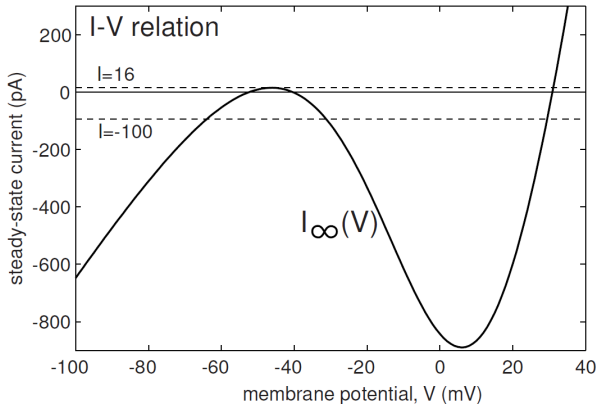


Figure 2.9: This diagram represents the steady-state I-V curve (I_∞) and the injected current I; it is obtained using the parameters values in table 2.1. The attractors are the intersections of the two curves.
Source: Izhikevich 2005 [22] p. 81

In figure 2.10 is represented the saddle node bifurcation diagram of the discussed model. The three branches of the S-shaped curve correspond to the rest, to the threshold and

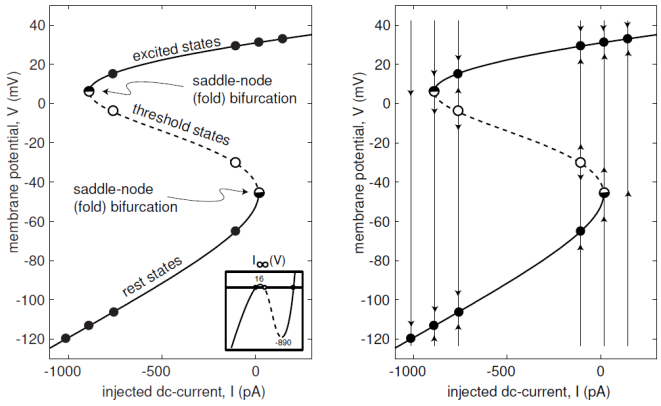


Figure 2.10: Saddle node bifurcation diagram of the discussed model obtained using the parameters in table 2.1. *Source:* Izhikevich 2005 [22] p. 82

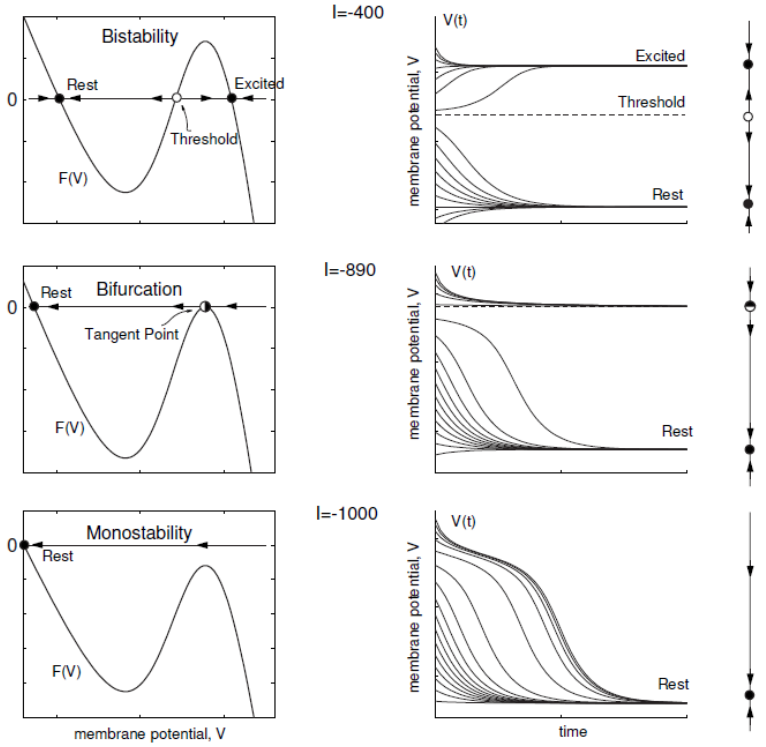


Figure 2.11: Bifurcation in the discussed model: when the injected current I is sufficiently small the excited state and the treshold state coalesce and disappear. *Source:* Izhikevich 2005 [22] p. 83

to the excited states of the model. Each point where the branches fold (minimum and maximum of $I_\infty(V)$) corresponds to a saddle-node bifurcation (I=16 pA and I=-890 pA). The first one (I=16 pA) corresponds to the disappearance of the rest state. The second one (I=-890 pA) corresponds to the disappearance of the excited state and it is illustrated in figure 2.11. The second occurs because I becomes so negative that the Na^+ inward current is no longer strong enough to balance the leak outward current and the negative injected current; the consequence is that the membrane can not maintain the depolarized state (excited state).
A neural circuit in the mammalian brain consists of many thousands of cells: its phase space is enormous. The dynamics of such systems can be very complex, exhibiting a wide gamut of spatiotemporal activity patterns. In these cases we use large scale models discussed in the next section, however the concept of bifurcations and attractors are still helpful for developing intuitions about attractor networks.

2.3 From small to large scale models

The goal of any large-scale description of neural dynamics is to reconstruct all relevant spatiotemporal dynamics of the neural system while preserving the mechanism which give rise to the observed dynamics. The assumption behind large scale models is that the neurocomputational units are more macroscopic than a single neuron. This assumption is justified by the fact that the key features of brain operations seem to emerge from the component interplay rather than being generated by each individual component.
Large scale models are the instruments to interpret the enormous data sets obtained from non-invasive brain imaging like functional magnetic resonance (*fMRI*) or electroencephalography (*EEG*).
Moreover large scale models are easier and less time-consuming to be solved computationally; actually today the the current technology allows to implement simulations of network with size significantly smaller than the amount of neurons we have in our brain.

2.3.1 Neural masses

A large scale model is composed of microscopic units which do not represent individual neurons, but rather complexes, also referred to as **neural masses**, capturing the non-reducible dynamics of a set of neurons [4]. These highly connected sets of neurons share common input and output pathways and functions. The activity of a neural mass (also known as **neural mass action**) is described by an m-dimensional vector variable at

location x and at time t

$$\psi(x,t) = \begin{bmatrix} \psi_1(x,t) \\ \psi_2(x,t) \\ \vdots \\ \psi_m(x,t) \end{bmatrix} . \tag{2.26}$$

The variable $\psi(x,t)$ is also referred as **neural population activity**. If the distance between neighboring neural masses is infinitesimally small, the physical space x is continuous and $\psi(x,t)$ is referred as **neural field**.
Since the neural mass action is generated by the N neurons within the neural mass, there is a mapping

$$\phi : Z(x,t) \mapsto \psi(x,t) \tag{2.27}$$

which unambiguously relates the N-dimensional neural activity (with N»m)
$Z(x,t) = \begin{bmatrix} Z_1(x,t) & Z_2(x,t) & ... & Z_N(x,t) \end{bmatrix}^T$ to the neural mass action ($\psi(x,t)$) [4].
The nature of this relation will be generally non trivial and involves the mean field reduction which will be discussed in the next section. A large scale model representation is successful when the large scale model simulation provides the same neural mass action $\psi(x,t+T)$ at a future time point $t+T$ as the simulation of the complete network dynamics using the microscopic neural activity $Z(x,t+T)$:

$$\phi : Z(x,t) \to \psi(x,t) \xrightarrow{\text{large scale network dynamics way}} \psi(x,t+T) \leftarrow Z(x,t+T) .$$

2.3.2 The mean field model

Mean field models attempt to model the dynamics of large, theoretically infinite, populations of neurons. Following the study of Deco et al., 2008 [8] we will derive the Fokker-Plank equation for neural dynamics that is specified in terms of spiking neurons; then, under the diffusion approximation, we will find the stationary solutions of the Fokker-Plank equation.

The population density approach

Each neuron can have several attributes like the membrane potential V, the current I and the time T since the last action potential. Each attribute induces a dimension in the phase plane of the neuron; if neurons have all the three attributes the state of each neuron correspond to a point $\nu = \{V, I, T\} \in \Re^3$ in the phase space. It is possible to define the neural ensemble probability density $\rho(\nu, t)$ that evolves according to the state of each

neuron. In this approach, individual integrate and fire neurons are grouped together into populations of statistically similar neurons. In general neuron with the same *V(t)*, at a given time t, have a different history because of random fluctuations in the input current. The key assumption in the population density approach is that the input currents, that arrive to the same population neurons, are uncorrelated. Then, neurons that belong to the same population and that have the same *V(t)* are indistinguishable.
It is possible to define the fraction of neurons that at time t have a membrane potential *V(t)* in the range $[\nu, \nu + d\nu]$ using the notion of probability density function:

$$\rho(\nu, t)d\nu = Prob\{V(t) \in [\nu, \nu + d\nu]\} . \tag{2.28}$$

The evolution of the population density is given by the Chapman-Kolmogorov equation:

$$\rho(\nu, t + dt) = \int_{-\infty}^{+\infty} \rho(\nu - \epsilon, t)\rho(\epsilon|\nu - \epsilon)d\epsilon \tag{2.29}$$

where $\rho(\epsilon|\nu - \epsilon)$ is the conditional probability that the membrane potential starting from the value $(\nu - \epsilon)$ at time t, increases is value of ϵ in an infinitesimal interval of time dt:

$$\rho(\epsilon|\nu - \epsilon) = Prob\{V(t + dt) = \nu|V(t) = \nu - \epsilon\} . \tag{2.30}$$

The Chapman Kolmogorv equation can be written in a differential form by performing a Taylor expansion in $\rho(\nu', t)\rho(\epsilon|\nu')$ around $\nu' = \nu$ (assuming that $\rho(\nu', t)$ and$\rho(\epsilon|\nu')$ are infinitely many time differentiable in ν'):

$$\rho(\nu', t)\rho(\epsilon|\nu') = \sum_{k=0}^{\infty} \frac{(-\epsilon)^k}{k!} \frac{\partial^k}{\partial \nu'^k} [\rho(\nu', t)\rho(\epsilon|\nu')]_{\nu'=\nu} . \tag{2.31}$$

Inserting this expansion in equation 2.29 we obtain:

$$\begin{aligned}
\rho(\nu, t + dt) = &\int_{-\infty}^{+\infty} \left(\sum_{k=0}^{\infty} \frac{(-\epsilon)^k}{k!} \frac{\partial^k}{\partial \nu^k} \rho(\nu, t)\rho(\epsilon|\nu) \right) d\epsilon \\
=&\rho(\nu, t) \int_{-\infty}^{+\infty} \rho(\epsilon|\nu)d\epsilon - \frac{\partial}{\partial \nu} \left[\rho(\nu|t) \left(\int_{-\infty}^{+\infty} \epsilon\rho(\epsilon|\nu)d\epsilon \right) \right] + \\
&+ \frac{1}{2} \frac{\partial^2}{\partial \nu^2} \left[\rho(\nu, t) \left(\int_{-\infty}^{+\infty} \epsilon^2\rho(\epsilon|\nu)d\epsilon \right) \right] + ... = \\
=&\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\partial^k}{\partial \nu^k} [\rho(\nu, t)\langle \epsilon^k \rangle_\nu] .
\end{aligned} \tag{2.32}$$

Finally, taking the limit for dt that tends to zero:

$$\frac{\partial \rho(\nu, t)}{\partial t} = \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \frac{\partial^k}{\partial \nu^k} \left[\rho(\nu, t) \lim_{dt \to 0} \frac{1}{dt} \langle \epsilon^k \rangle_\nu \right] . \tag{2.33}$$

This equation expresses the time evolution of the population density in a differential form and it is known as the Kramers-Moyal expansion of the original integral Chapman Kolmogorov equation (eq. 2.29).

The diffusion approximation

In order to understand the dynamics of the population density obtained in equation 2.33 we need to know the expression for the moments $\langle \epsilon^k \rangle_\nu$, or, in other words, the average of the membrane voltage variation ϵ with respect to the conditional probability $\rho(\epsilon|\nu)$ at a given ν during the interval dt. These calculations can be obtained using the mean-field approximation. As mentioned above, the history of the currents that arrive on each neuron is uncorrelated, therefore each current have the same statistic. Assuming the ergodicity for all neurons in a population, the mean field approach allows to discard the index that identifies each single neuron; the infinitesimal change in the membrane potential of all neurons can be written as [2]

$$dV(t) = \langle J \rangle_J N r(t) dt - \frac{V(t) - V_L}{\tau} dt \tag{2.34}$$

where N is the number of neurons, $\langle J \rangle_J$ is the average of the synaptic weights in the population, and *r(t)* is the mean population firing rate defined as:

$$r(t) = \lim_{dt \to 0} \frac{n_{spikes}(t, t+dt)}{N dt} \,. \tag{2.35}$$

The moments of the infinitesimal depolarization $\epsilon = dV(t)$ can be calculated using the equation 2.34. The first two moments in the Kramers-Moyal expansion (eq. 2.33) are called respectively the drift and the diffusion coefficients:

$$M^{(1)} = \lim_{dt \to 0} \frac{1}{dt} \langle \epsilon \rangle_\nu = \langle J \rangle_J N r(t) - \frac{\nu - V_L}{\tau} = \frac{\mu(t)}{\tau} - \frac{\nu - V_L}{\tau} \tag{2.36}$$

$$\begin{aligned} M^{(2)} &= \lim_{dt \to 0} \frac{1}{dt} \langle \epsilon^2 \rangle_\nu = \lim_{dt \to 0} \frac{1}{dt} \left[\langle J \rangle_J N r(t) dt - \frac{V(t) - V_L}{\tau} dt \right]^2_\nu = \\ &= \lim_{dt \to 0} \left[\langle J^2 \rangle_J N^2 \left(\lim_{dt \to 0} \frac{n_{spikes}(t, t+dt)}{N dt} \right)^2 dt \right] + \\ &\quad + \lim_{dt \to 0} \left[\frac{(\nu - V_L)^2}{\tau^2} dt \right] - \lim_{dt \to 0} \left[2 \frac{\langle J \rangle_J N r(t) (\nu - V_L)}{\tau} dt \right] = \\ &= \langle J^2 \rangle_J N r(t) = \frac{\sigma(t)^2}{\tau} \,. \end{aligned} \tag{2.37}$$

In general for $k > 1$:

$$M^{(k)} = \lim_{dt \to 0} \frac{1}{dt} \langle \epsilon^k \rangle_\nu = \langle J^k \rangle_J N r(t) \,. \tag{2.38}$$

[2] Taking into account the expressions obtained for the LIF model (section 2.2.1):
$\tau \frac{dV_i(t)}{dt} = -[V_i(t) - V_L] + RI_i(t)$
$RI_i(t) = \tau \sum_{j=1}^N J_{ij} \sum_k \delta(t - t_j^{(k)})$
$\Rightarrow \quad dV_i(t) = \sum_{j=1}^N J_{ij} \sum_k \delta(t - t_j^{(k)}) dt - \frac{V_i(t) - V_L}{\tau} dt$

The diffusion approximation is exact in the limit of infinitely large network ($N \to \infty$) and when the synaptic weights scale appropriately with network size: $J \to 0$ but $NJ^2 \to cost.$ In case of exact approximation, all the moments higher than two become negligible in relation to the drift (μ) and to the diffusion (σ^2) coefficients.
In the Kramers-Moyal expansion (equation 2.33), under the diffusion approximation, it is possible to omit all higher orders $k > 2$; the above-mentioned equation becomes the Fokker-Plank equation:

$$\frac{\partial \rho(\nu,t)}{\partial t} = \frac{1}{2\tau}\sigma^2(t)\frac{\partial^2 \rho(\nu,t)}{\partial \nu^2} + \frac{\partial}{\partial \nu}\left[\left(\frac{\nu - V_L - \mu(t)}{\tau}\right)\rho(\nu,t)\right] . \tag{2.39}$$

When the drift is linear and the diffusion coefficient $\sigma^2(t)$ is constant, the Fokker-Plank equation express an *Ornstein-Uhlenbeck* process, that is a stochastic process that tends to drift towards its long-term mean; for this reason it is called mean-reverting. This process describes the temporal evolution of the membrane potential when the input afferent currents is given by:

$$RI(t) = \mu(t) + \sigma\sqrt{\tau}w(t) \tag{2.40}$$

where *w(t)* is white noise. This equation allows to interpret the sum of many Poisson process in equation 2.5 as a normal random variable with mean $\mu(t)$ and variance σ^2.

The mean field approach

The mean field approximation consists in finding the Fokker-Plank equation (equation 2.39) stationary solutions that represent the stationary solutions of the original leaky and integrate neural system under the diffusion approximation.
We can use the Fokker-Plank equation to define the flux of probability F:

$$\frac{\partial \rho(\nu,t)}{\partial t} = -\frac{\partial F(\nu,t)}{\partial \nu} \equiv -\frac{\partial}{\partial \nu}\left[-\frac{\nu - V_L - \mu(t)}{\tau}\rho(\nu,t) - \frac{\sigma^2(t)}{2\tau}\frac{\partial \rho(\nu,t)}{\partial \nu}\right] . \tag{2.41}$$

At the threshold potential, the probability current has to give the average firing rate r of the population. Then, if V_{th} is the neural firing threshold, the stationary solution should satisfy the following boundary condition:

$$\begin{aligned} \rho(V_{th},t) &= 0 \\ \frac{\partial \rho(V_{th},t)}{\partial \nu} &= -\frac{2r\tau}{\sigma^2} \ . \end{aligned} \tag{2.42}$$

Additionally, for $\nu \to -\infty$, the probability density has to vanish fast enough to be integrable:

$$\begin{aligned} \lim_{\nu \to -\infty} \rho(\nu,t) &= 0 \\ \lim_{\nu \to -\infty} \nu\rho(\nu,t) &= 0 \, . \end{aligned} \tag{2.43}$$

In order to make the model realistic, we need to take in account the refractory period of a neuron t_{ref} that is the time that occurs after an action potential; during this time the neuron is not able to generate another action potential (section 2.1.2). This means that the probability mass leaving the threshold at time t has to be re-injected at the rest potential at time $t+t_{ref}$; then Fokker-Plank equation (eq. 2.41) becomes:

$$\frac{\partial \rho(\nu,t)}{\partial t} = -\frac{\partial}{\partial \nu}[F(\nu,t) + r(t-t_{ref})\Theta(\nu - V_{reset})] \tag{2.44}$$

where Θ is the Heaviside function. Solving this equation with the boundary conditions, and taking into account the normalization of the mass probability, the population transfer function ϕ of Ricciardi is obtained:

$$r = \left[t_{ref} + \tau\sqrt{\pi} \int_{\frac{V_{reset}-V_L-\mu}{\sigma}}^{\frac{V_{th}-V_L-\mu}{\sigma}} e^{x^2}(1+erf(x))dx \right]^{-1} = \phi(\mu,\sigma)\,. \tag{2.45}$$

The stationary dynamics of each population can be described by the population transfer function of Ricciardi ϕ, which provides the average population rate as a function of the average input current. The result found in equation 2.45 can be generalized for more than one population of neurons whose input currents share the same statistical properties and fire spikes independently at the same rate:

$$r_i = \phi(\mu_i,\sigma_i)\,. \tag{2.46}$$

To solve equation 2.46 for all the neural population i, the differential equation below can be integrated while it describes the approximate dynamics of the system which has a fixed point solutions corresponding to equation 2.46:

$$\tau_i \frac{dr_i}{dt} = -r_i + \phi(\mu_i,\sigma_i)\,. \tag{2.47}$$

The mean field approach ensures that this dynamics will converge to a stationary attractor that is consistent with the steady state required.
In the next section an extended mean field framework will be analyzed. This model is consistent with the LIF model and with the realistic synaptic equations that consider both the fast and slow glutamatergic excitatory synaptic dynamics (AMPA and NMDA) and the dynamics of GABA inhibitory synapses.

2.4 The reduced Wong Wang model

The section concerns the model used in this work: the *reduced Wong Wang* model. The model is obtained through successive approximation from the work of Wang (2002) [36]

that will be described in the first part of the section. Next, will be treated the model reduction by Wong and Wang (2006) [37], and Deco et al.(2013) [13].

2.4.1 The biophysically realistic cortical network model

The network model in Wang 2002 [36] (inspired to the work of Brunel and Wang 2001 [5]) was constructed in order to understand the maintenance of an item in working memory, which is achieved through persistent activity in selective neural assemblies of the cortex. This persistent activity of neurons group over a delay period was observed in experiments with monkeys. The monkeys perform a paradigm of visual motion discrimination like the one of Funahashi [17] that will be briefly explained. The monkeys are trained to fix their eyes on a central position during a sudden presentation of cue. The cues are presented in one of eight different locations separated by 45° angles. Once the cue is given, there is a delay period until a 'go' signal which indicates that the monkey has to focus its eyes on the previous location of the cue. The task requires the monkey to retain the location in its memory during the delay period. These experiments show that some neurons in the dorsolateral prefrontal cortex are active during the delay period. Thus, the delayed activity of these neurons indicates the location the eyes must be directed towards after the delay period.
Starting from these experimental evidences Wang tried to investigate the basic cellular and synaptic mechanisms in a perceptual decision-making process. In Wang model the populations of neurons are linked through recurrent excitatory and inhibitory connections in order to produce an attractor dynamics that amplify the difference between conflicting inputs and that generate a binary choice.

The organization of the network

The network represents a cortical module. It is composed of three populations of pyramidal cells (N_E, 80%) and one populations of interneurons (N_I, 20%). One of the excitatory population is sensitive to a leftward motion of the stimulus and another to the rightward. The interneurons and the remaining pyramidal cells are non-selective to any of the motion stimuli (figure 2.12). Each neuron receives:

C_E excitatory synaptic contacts from pyramidal cells;

C_I inhibitory synaptic contacts from interneuron. Then the inhibitory population acted not only to globally inhibit the excitatory cells, but also to inhibit itself.

C_{ext} excitatory synaptic contacts from outside the network. These connections send to the network all the information (stimuli) received from the outside world, as well

as background noise due to spontaneous activity outside the module.

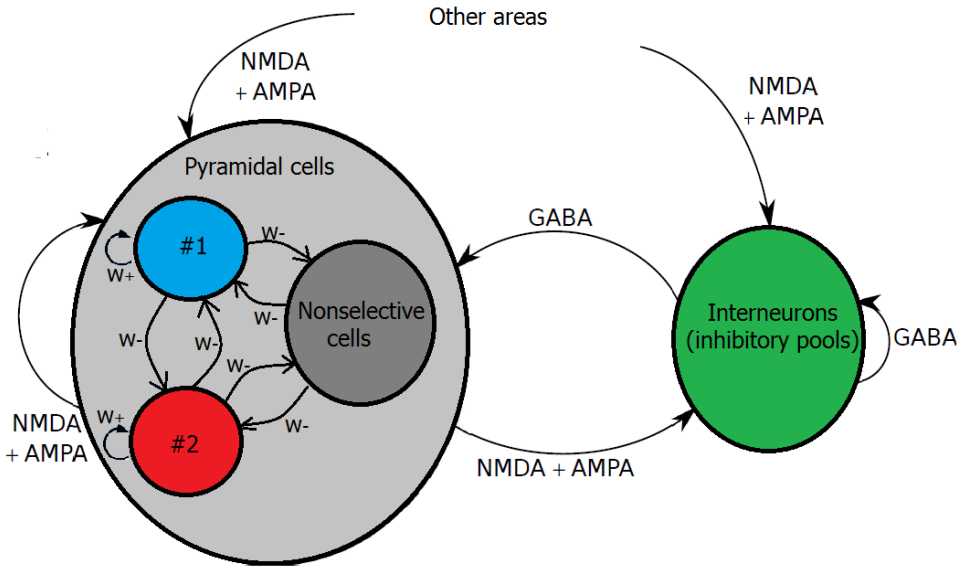


Figure 2.12: Schematic description of the model. There are three pyramidal population of cells: one non selective to the stimuli (in gray), and the are two groups (blue and red) selective to one of the two external stimuli (#1 and #2). Pyramidal cells send connections to other pyramidal cells and to interneurons through AMPA and NMDA synapses. Interneuron cells send connections to itself and to pyramidal cells through GABA synapses. Cells within a selective population have relatively stronger connections (modulated by $w_+ > 1$), while connections outside the selective population are relatively weaker (modulated by $w < 1$).
Source: Adapted from Brunel and Wang 2001 p. 67 [5]

Neurons

Each neuron is described by the leaky and integrate model (section 2.2.1):

$$C_m \frac{dV_i(t)}{dt} = -g_m \left(V_i(t) - V_L\right) - I_{syn}(t) \tag{2.48}$$

where $I_{syn}(t)$ is the total synaptic current flowing into cell i, V_L is the leak of the resting potential of the cell in absence of external input, C_m is the membrane capacitance and g_m is the conductance membrane leak. When $V(t)$ reaches V_{th} the cell generates an action potential in the form of a spike (a delta function). After that, the cell stays in a short absolute refractory period, where the membrane potential is clamped at V_{reset}. All the

values of these parameters for pyramidal cells and interneurons are shown in the table 2.2.

Synapses

The total synaptic current come from both external (ext) inputs and recurrent (rec) synaptic connections. The recurrent inputs include both excitatory and inhibitory currents. The excitatory inputs consist of AMPA ($I_{rec,AMPA}$) and NMDA receptor-mediated synapses ($I_{rec,NMDA}$), while the inhibitory input is a GABA receptor-mediated ($I_{rec,GABA}$). Thus the total synaptic current is given by:

$$I_{syn}(t) = I_{ext,AMPA}(t) + I_{rec,AMPA}(t) + I_{rec,NMDA}(t) + I_{rec,GABA}(t) \tag{2.49}$$

where:

$$I_{ext,AMPA}(t) = g_{ext,AMPA}(V_i(t) - V_E) \sum_{j=1}^{C_{ext}} s_j^{ext,AMPA}(t) \tag{2.50}$$

$$I_{rec,AMPA,}(t) = g_{rec,AMPA}(V_i(t) - V_E) \sum_{j=1}^{C_E} w_{ij} s_j^{rec,AMPA}(t) \tag{2.51}$$

$$I_{rec,NMDA}(t) = \frac{g_{NMDA}(V_i(t) - V_E)}{1 + \gamma e^{-\beta V_i(t)}} \sum_{j=1}^{C_E} w_{ij} s_j^{NMDA}(t) \tag{2.52}$$

$$I_{rec,GABA}(t) = g_{GABA}(V_i(t) - V_I) \sum_{j=1}^{C_I} s_j^{GABA}(t)\,. \tag{2.53}$$

Here g_i is the synaptic conductances, s_i the synaptic gating variable (fraction of open channels), and V_E and V_I are respectively the excitatory and inhibitory reversal potentials (parameters in table 2.2). The sum over j represents a sum over the synapses formed by presynaptic neurons j . Observing the equation 2.52 it can be noted that the NMDA currents are voltage dependent and they are modulated by intracellular magnesium concentrations ($\gamma = [Mg2+]/3.57$).

The dimensionless parameter w_{ij} is the synaptic weight between nodes i and nodes j, and it is defined as follows:

- Excitatory synapses, which connect cells inside the selective population, are chosen to be relatively stronger than those which connect cells outside the selective population. This particular structure is similar to the Hebbian rule where cells that tend to fire together form stronger synapses. In this case $w = w_+ > 1$

Parameters		Excitatory neurons	Inhibitory neurons
Membrane capacitance	C_m	0.5 nF	0.2 nF
Leak conductance	g_m	25 nS	20 nS
Resting potential	V_L	-70 mV	-70 mV
Firing threshold	V_{th}	-50 mV	-50 mV
Reset potential	V_{reset}	-55 mV	-55 mV
Refractory period	τ_{ref}	2 ms	1 ms
AMPA external conductance	$g_{ext,AMPA}$	2.496 nS	1.944 nS
AMPA recurrent conductance	$g_{rec,AMPA}$	0.104 nS	0.081 nS
NMDA recurrent conductance	g_{NMDA}	0.327 nS	0.258 nS
GABA recurrent conductance	g_{GABA}	0.4375	3.4055 nS

Synaptic parameters		Values
Excitatory potential	V_E	0 mV
Inhibitory potential	V_I	-70 mV
Decay time of AMPA	τ_{AMPA}	2 ms
Rise time of NMDA	$\tau_{NMDA,rise}$	2 ms
Decay time of NMDA	$\tau_{NMDA,decay}$	100 ms
Decay time of GABA	τ_{GABA}	10 ms
	α	0.5 kHz
	β	0.062
	γ	0.28

Table 2.2: Neural and synaptic parameters used in the model. These values are from Brunel and Wang 2001 [5]

- Instead $w = w_- < 1$ for the synapses between two different selective populations, and for synapses between the nonselective population and the selective ones.

- $w = 1$ for all the other connections.

The gating variables $s_j^i(t)$ are the fractions of open channels of neurons and their evolution

is given by:

$$\frac{ds_j^{AMPA,ext}(t)}{dt} = -\frac{s_j^{AMPA,ext}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k) \tag{2.54}$$

$$\frac{ds_j^{AMPA,rec}(t)}{dt} = -\frac{s_j^{AMPA,rec}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k) \tag{2.55}$$

$$\frac{ds_j^{NMDA,rec}(t)}{dt} = -\frac{s_j^{NMDA,rec}(t)}{\tau_{AMPA}} + \alpha x_j(t)(1 - s_j^{NMDA,rec}(t)) \tag{2.56}$$

$$\frac{dx_j^{NMDA,rec}(t)}{dt} = -\frac{x_j^{NMDA,rec}(t)}{\tau_{NMDA,rise}} + \sum_k \delta(t - t_j^k) \tag{2.57}$$

$$\frac{ds_j^{GABA}(t)}{dt} = -\frac{s_j^{GABA}(t)}{\tau_{GABA}} + \sum_k \delta(t - t_j^k)\,. \tag{2.58}$$

The sum of delta functions denotes the sum of spikes generated from presynaptic neurons j at times t_j^k. $\tau_{NMDA,rise}$ and $\tau_{NMDA,decay}$ are the rise and decay times for the NMDA synapses, and τ_{AMPA} and τ_{GABA} are the decay times for AMPA and GABA synapses (note that the rise times of both AMPA and GABA synaptic currents are neglected because they are short (< 1 ms)).

2.4.2 Model reduction

Thanks to its biological realism, the Wang model explained in the previous section is able to replicate some of the psychophysical and physiological observations obtained during experiments of the visual discrimination reaction task on primates [37]. The weakness of the model is its complexity: it consists of thousands of spiking neurons that interact with each other in a highly non linear manner; then it is difficult to deeply analyze and understand how it works. This was the reason which motivated Kong-Fatt Wong and Xiao-Jing Wang to construct a reduced version of the Wang model through a mean-field approach. Their model was presented in 2006 in the paper "A Recurrent Network Mechanism of Time Integration in Perceptual Decisions" [37]. Through consecutive approximations, they reduced the original model composed of 7200 dynamical equations to a simplified model of only 2 variables: the *reduced Wong Wang model*. This model, despite of its conciseness, is able to reproduce much of the behavior of the original spiking neuron model of Wang. The main idea of the reduction is that all the neurons in one population are considered to be homogeneous, while each population in the spiking neural network is different; then the input of a neuron can be treated as a Gaussian process. As already mentioned, in this view, the mean activity of a large homogeneous population can be modeled by the activity of a representative unit. Taking into account that the population

firing rates depend on the input currents which in turn depend on the firing rates, it is clear that the firing rate of the representative unit must be self-consistent in order to represent the population firing rate.
In the next sections the necessary steps to arrive to this model are presented, following the paper of Wong and Wang 2006 [37].

Mean Field approximation

As it was observed in section 2.3.2 (eq. 2.45), the firing rate r of a population of leaky integrate and fire neurons, receiving a noisy input current, is given by the transfer function ϕ of Ricciardi:

$$r = \phi(\mu) = \left[t_{ref} + \tau\sqrt{\pi} \int_{\frac{V_{reset}-V_L-\mu}{\sigma}}^{\frac{V_{th}-V_L-\mu}{\sigma}} e^{x^2}(1+erf(x))dx \right]^{-1} \tag{2.59}$$

where τ is the membrane time constant, V_{th} and V_{reset} are respectively the spiking threshold and the reset value of the membrane voltage. σ is the membrane potential standard deviation, t_{ref} is the refractory period, V_L is the leak potential, μ is the drift coefficient obtained in equation 2.36 and defined as:

$$\mu = \langle J \rangle_J NQ(t)\tau = \frac{I_{syn}}{g_L} \tag{2.60}$$

where I_{syn} is the total input current.
Defining $V_{ss} = V_L + I_{syn}/g_L$ equation 2.59 becomes:

$$r = \phi(I_{syn}) = \left[t_{ref} + \tau\sqrt{\pi} \int_{\frac{V_{reset}-V_{ss}}{\sigma}}^{\frac{V_{th}-V_{ss}}{\sigma}} e^{x^2}(1+erf(x))dx \right]^{-1} . \tag{2.61}$$

In the work of Wong and Wang (2006), instead of using this formula, they used a simplified input output function:

$$\phi(I_{syn}) = \frac{c_{E,I} I_{syn} - I_{E,I}}{1 - e^{-g_{E,I}(c_{E,I} I_{syn} - I_{E,I})}} \tag{2.62}$$

where E and I are respectively labels for pyramidal cell and interneuron. $c_{E,I}$ is the gain factor and $g_{E,I}$ is the noise factor.
The values of these parameters are obtained by fitting the model to the Ricciardi transfer function ϕ of a LIF neuron whose external stimulus is driven by AMPA receptor mediated by external Gaussian noise (figure 2.14). As stated in section 2.3.2, using the mean field approximation, it is possible to describe the temporal dynamics of the system with equation 2.47; in effect this dynamics system will converge to the steady state described by equation 2.59, and for the observations made above, also to the steady state required in

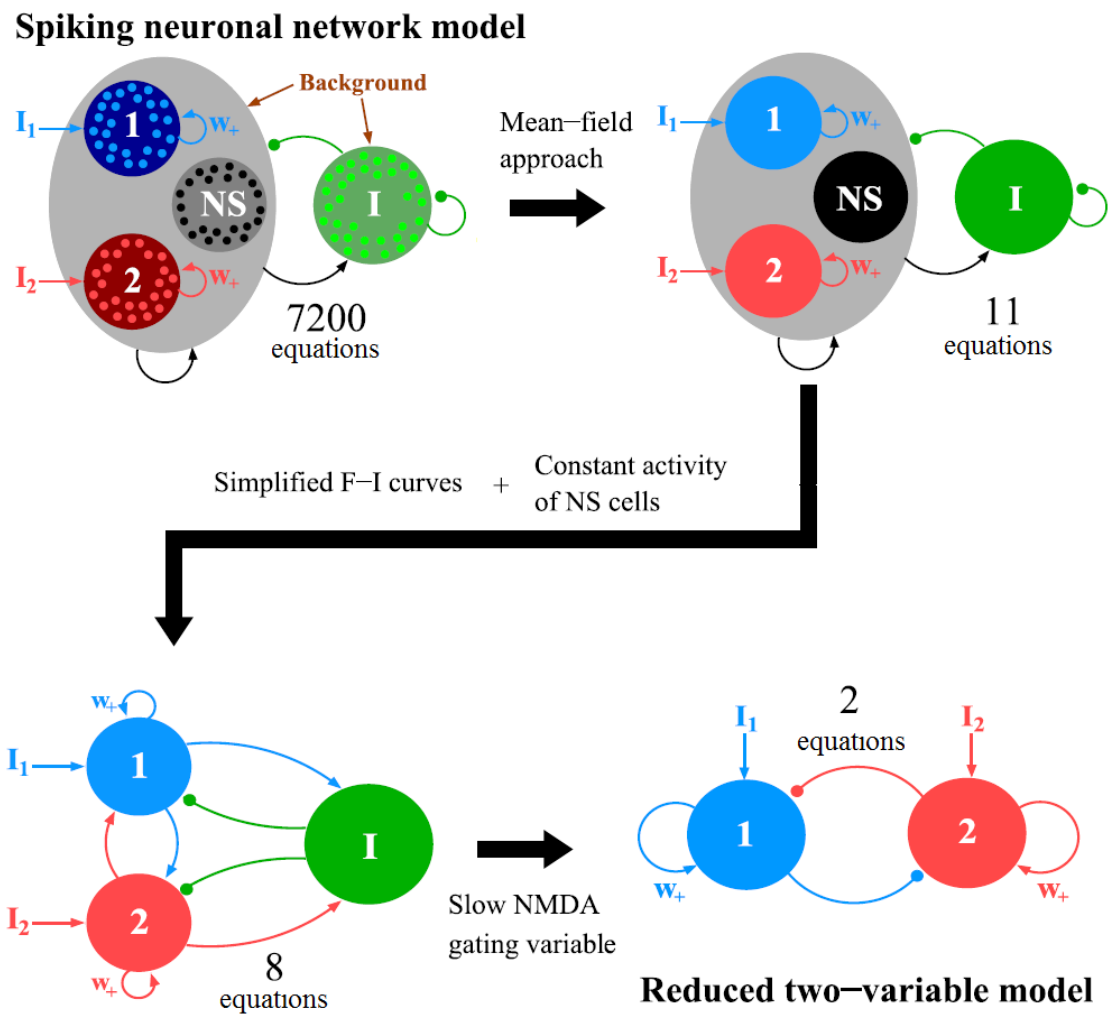


Figure 2.13: Schematic representation of the approximation required to arrive to the reduced Wong Wang model starting from the Wang model.
NS and I denote respectively the nonselective excitatory (black) and inhibitory (green) pools of cells. Arrows indicate excitatory connections and circles inhibitory connections. I_i is the input from external stimulus to i-th selective neural populations. Brown arrows indicate background noisy inputs. w_+ denotes enhanced excitatory connections within each selective neural pool. The numbers on the right display the total number of dynamical equations involved in the model.
Source: Adapted from Wong and Wang 2006 [37] p. 1316

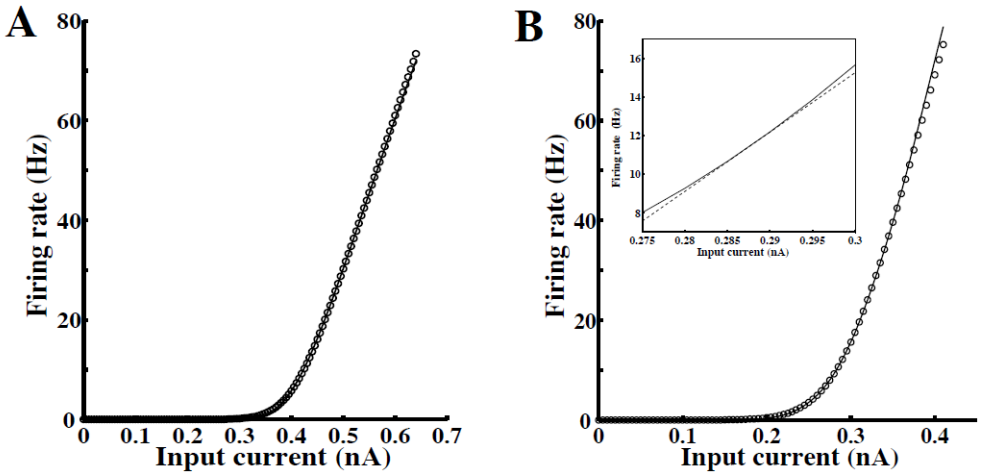


Figure 2.14: Firing rate as function of the synaptic input current of a pyramidal cell (A) and of an interneuron(B). Bold lines are plotted using the Ricciardi formula (eq. 2.61) and circles are plotted using the simplified equation 2.62. In the inset there is a close up of the input-output relation of an interneuron; dashed line is obtained fitting this relation with the linear function (eq. 2.70)
Source: Wong and Wang (2006) [37], Supplementary Material, p. 4

this case and described by equation 2.62. Then the mean firing rate dynamics is described by four equations:

$$\begin{cases} \tau_r \frac{dr_i}{dt} = -r_i + \phi(I_{syn,i}) & \text{with } i = 1,2,3 \\ \tau_r \frac{dr_I}{dt} = -r_I + \phi(I_{syn,I}) \end{cases} \tag{2.63}$$

where i labels the two selective excitatory populations and the non-selective one, and I labels the inhibitory population. The dynamics of the output synaptic NMDA gating variable S (equation 2.56 in the Wang model) is described by:

$$\frac{dS_{NMDA,i}}{dt} = -\frac{S_{NMDA,i}}{\tau_{NMDA}} + (1 - S_{NMDA,i})F(\psi(r_i)) \quad \text{with } i = 1,2,3 \tag{2.64}$$

with:

$$F(\psi_i) = \frac{\psi_i}{\tau_{NMDA}(1 - \psi_i)} \tag{2.65}$$

where ψ_i is the steady state of S_i. $F(\psi_i)$ can be calculated analytically when the input spike train is Poissonian (Brunel and Wang (2001) [5]). The result is:

$$\psi = \frac{\gamma \cdot r \cdot \tau_{NMDA}}{1 + \gamma \cdot r \cdot \tau_{NMDA}} \tag{2.66}$$

where $\gamma = 0.641$. Replacing the expression found in equation 2.65 it results $F(\psi) = \gamma r$; consequently the dynamics of the NMDA gating variable (eq. 2.68) becomes:

$$\frac{dS_{NMDA,i}}{dt} = -\frac{S_{NMDA,i}}{\tau_{NMDA}} + (1 - S_{NMDA,i})\gamma r_i \,. \tag{2.67}$$

Then the dynamics of the gating variables (equation 2.55, 2.56 and 2.58 in the Wang model) is:

$$\begin{cases} \frac{dS_{AMPA,i}}{dt} = -\frac{S_{AMPA,i}}{\tau_{AMPA}} + r_i \\ \frac{dS_{NMDA,i}}{dt} = -\frac{S_{NMDA,i}}{\tau_{NMDA}} + (1 - S_{NMDA,i})\gamma r_i \quad \text{with } i = 1,2,3 \\ \frac{dS_{GABA}}{dt} = -\frac{S_{GABA}}{\tau_{GABA}} + r_I \end{cases} \tag{2.68}$$

Summarizing, with the mean field approximation, it is possible to reduce the original model of Wang (2002) to a system of 11 variables i.e. the mean firing rates and the synaptic gating variables of four different neural population: two selective and one non selective excitatory populations, and the inhibitory population. The dynamics of these variables are described by the equation in the system 2.63 (the mean firing rates) and in the system 2.68 (the gating variables NMDA, AMPA and GABA).

Constant activity of non-selective cells

Under a wide range of conditions, the firing rate of the nonselective population changes only by a modest amount [37]. Consequently we can assume that the firing rates of the non-selective cells is constant; this observation reduces the system of four population to three population: the inhibitory population and two excitatory populations.
Then the dynamics of the network is given by the following 8 dynamical equations:

$$\begin{cases} \tau_r \frac{dr_i}{dt} = -r_i + \phi(I_{syn,i}) \\ \tau_r \frac{dr_I}{dt} = -r_I + \phi(I_{syn,I}) \\ \frac{dS_{AMPA,i}}{dt} = -\frac{S_{AMPA,i}}{\tau_{AMPA}} + r_i \qquad \text{with } i = 1,2 \\ \frac{dS_{NMDA,i}}{dt} = -\frac{S_{NMDA,i}}{\tau_{NMDA}} + (1 - S_{NMDA,i})\gamma r_i \\ \frac{dS_{GABA}}{dt} = -\frac{S_{GABA}}{\tau_{GABA}} + r_I \end{cases} \tag{2.69}$$

Linearization of the transfer function of the inhibitory population

The mean firing rate of the inhibitory population typically is in the range of $8 \div 15$ Hz. As can be noticed in the close up of figure 2.14, within this range the input-output relation

is almost linear, and fitted by:

$$\phi(I_{syn}) = \frac{1}{g_2}(c_I I_{syn} - I_I) + r_0 \tag{2.70}$$

where g_2 and $r_0 = 11.5$ Hz.

Slow dynamics of NMDA gating variable

The synaptic gating variable of the NMDA receptors has a much longer decay time constant ($\tau_{NMDA} = 100$ ms) than AMPA ($\tau_A MPA = 2$ ms) and GABA ($\tau_{GABA} = 10$ ms) receptors. Furthermore it is possible to assume that the NMDA gating variable (S_{NMDA}) dominates the time evolution of the system since it reaches its steady state much slower than the GABA and AMPA gating variables.
Then, if we assume that the GABA and AMPA gating variables reach their steady states instantaneously, the S_{GABA} and S_{AMPA} become proportional to the average firing rate of presynaptic cells:

$$\begin{cases} \dfrac{dS_{AMPA,i}}{dt} = 0 \Rightarrow \quad S_{AMPA,i}(t) = \tau_{AMPA} r_i(t) = \tau_{AMPA}\phi_i(t) \\ \dfrac{dS_{GABA}}{dt} = 0 \Rightarrow \qquad S_{GABA}(t) = \tau_{GABA} r_I(t) = \tau_{GABA}\phi_I(t) \end{cases} \tag{2.71}$$

In conclusion, the network model is reduced into a two-variable system described by the following dynamical equations:

$$\begin{cases} \dfrac{dS_1}{dt} = -\dfrac{dS_1}{dt} + (1 - S_1)\gamma r_1 \\ \dfrac{dS_2}{dt} = -\dfrac{dS_2}{dt} + (1 - S_2)\gamma r_2 \end{cases} \tag{2.72}$$

where *1* and *2* label the two different excitatory populations, S and τ_S denote respectively S_{NMDA} and τ_{NMDA}.
The firing rate r is given by the simplified version of the Ricciardi transfer function (eq. 2.62) and depends on the total synaptic current $I_{syn,i}$ that is given by:

$$\begin{cases} I_{syn,1} = J_{N,11}S_1 - J_{N,12}S_2 + J_{A,11}r_1 - J_{A,12}r_2 + I_0 + I_1 + I_{noise,1} \\ I_{syn,2} = J_{N,22}S_2 - J_{N,21}S_1 + J_{A,22}r_2 - J_{A,21}r_1 + I_0 + I_2 + I_{noise,2} \end{cases} \tag{2.73}$$

where I_i is the stimulus to the i-th excitatory population, $I_{noise,i}$ is the noise term and I_0 is the mean external input common to both populations. The coefficients $J_{N,ij}$ and $J_{A,ij}$ are the coupling constants from population i to population j mediated respectively by $NMDA$ and $AMPA$ receptors.

Since the terms $J_{x,ij}$ have negative sign, the connectivity between the two populations is negative. This is because the inhibitory population I receives stimulus from both excitatory populations, and then its output (proportional to input, because of the linearization of the transfer function) is of the form:

$$J_{N,I\to E}J_{E\to I}(S_1 + S_2) \tag{2.74}$$

$$J_{A,I\to E}J_{E\to I}(r_1 + r_2)\,. \tag{2.75}$$

Then, for example, in $I_{syn,1}$ the S_2-dependent term associated with the NMDA gating variable is:

$$(J_{N,2\to 2} - J_{N,I\to 2}J_{N,2\to I})S_2 \doteq -J_{N,12}S_2\,. \tag{2.76}$$

In order to further simplify the problem it is necessary to find an explicit expression for the firing rate; indeed the firing rate expression $r_i = \phi(I_{syn,i})$ depends on $I_{syn,i}$ that in turns depends on S_1, S_2, r_1, r_2. Once defined the following new variables:

$$x_1 = J_{N,11}S_1 1 - J_{N,12}S_2 + I_0 + I_1 + I_{noise,1} \tag{2.77}$$

$$x_2 = J_{N,22}S_2 - J_{N,21}S_1 + I_0 + I_2 + I_{noise,2}\,. \tag{2.78}$$

Wong and Wang, in the Supplementary Materials of their article [37], found an equivalent response function H in terms of them:

$$\begin{aligned} r_1 &= H_1(x_1, x_2) = \\ &= \frac{a(J_{A,11})x_1 - f_A(J_{A,12}, x_2) - b(J_{A,11})}{1 - exp\{-d(J_{A,11})[a(J_{A,11})x_1 - f_A(J_{A,12}, x_2) - b(J_{A,11})]\}} \end{aligned} \tag{2.79}$$

$$\begin{aligned} r_2 &= H_2(x_2, x_1) = \\ &= \frac{a(J_{A,22})x_2 - f_A(J_{A,21}, x_1) - b(J_{A,22})}{1 - exp\{-d(J_{A,22})[a(J_{A,22})x_2 - f_A(J_{A,21}, x_1) - b(J_{A,22})]\}} \end{aligned} \tag{2.80}$$

where the parameters a, b, c, d and the function f are chosen to fit the numerical solutions:

$$a(J_{A,11}) = 239400J_{A,11} + 270 \qquad [(VnC)^-1] \tag{2.81}$$

$$b(J_{A,11}) = 97000J_{A,11} + 108 \qquad [Hz] \tag{2.82}$$

$$c(J_{A,11}) = -30J_{A,11} + 0.154 \qquad [s] \tag{2.83}$$

$$f_A(J_{A,12}, x_2) = J_{A,12}(-27x_2 + 106)\theta(x_2 - 0.4) \qquad [Hz] \tag{2.84}$$

where $\theta(x)$ is 0 if x is negative, and 1 if $x \geq 0$.
Finally, combining these results with the dynamical equations in the system 2.72, Wong and Wang obtained the two dynamical equations able to describe the system:

$$\begin{cases} \dfrac{dS_1}{dt} = -\dfrac{S_1}{\tau_S} + (1 - S_1)\gamma H(x_1, x_2) \\ \dfrac{dS_2}{dt} = -\dfrac{S_2}{\tau_S} + (1 - S_2)\gamma H(x_2, x_1) \end{cases} \tag{2.85}$$

2.4.3 Single neural population reduction

Two years ago in the paper "Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations." of Deco et al., (2013) [13] (see also subsection 3.2.3), the reduced Wong Wang model was used in resting state modeling with an additional simplification: instead of considering two excitatory populations, they considered only one neural population.

This is obtained by neglecting the contributions of the AMPA receptors to local recurrent excitation, i.e. neglecting the terms $J_{A,ij}$. This means that the firing rates of the two populations are "dis-entangled" since the equations defining the response function H_i (eqs. 2.79 and 2.80) depend only on x_i:

$$r_1 = H_1(x_1) = \frac{ax_1 - b}{1 - exp[-d(ax_1 - b)]} \tag{2.86}$$

$$r_2 = H_2(x_2) = \frac{ax_2 - b}{1 - exp[-d(ax_2 - b)]}\,. \tag{2.87}$$

In conclusion the mean activity of a neural population composed by excitatory and inhibitory leaky integrate and fire neurons, interconnected via NMDA synapses, can be expressed by a single one variable: S.
Then the whole brain dynamics of the network composed by these kind of neural populations is described by the following set of coupled nonlinear stochastic differential equations:

$$\frac{dS_i(t)}{dt} = -\frac{S_i}{\tau_s} + (1 - S_i)\gamma H(x_i) + \sigma\eta_i(t) \tag{2.88}$$

$$H(x_i) = \frac{ax_i - b}{1 - exp(-d(ax_i - b))} \tag{2.89}$$

$$x_i = wJ_N S_i + J_N G \sum_j C_{ij} S_j + I_0 \tag{2.90}$$

where, summarizing, S_i, $H(x_i)$ and x_i represent respectively the NMDA average synaptic gating variable, the collective population firing rate and the total synaptic input to a cortical region i. The parameters and their values are shown in the third column of the table 4.1 in section 4.2.

Chapter 3

Modeling the resting state

The aim of the chapter is to discuss the state of the theoretical framework used to understand the dynamics of the resting state.
As discussed in the first chapter, spontaneous brain activity during wakeful rest is not random but displays specific features, as the slow fluctuating spatiotemporal patterns in BOLD signals. Understand how this activity emerges during rest is not a trivial problem and it has enthralled many scientists.
Once defined two concepts essential in this field (the *structural* and the *functional* connectivity), the theoretical analysis about the mechanism of the resting state brain fluctuations, that are interesting for the purpose of the present thesis, are summarized.
The first papers analyzed concern a series of models of resting state dynamics, in macaque and human cortices, that use different implementation for node dynamics and interaction terms: Honey et al., 2007 [20] ,Ghosh et al., 2008 [18], Deco et al., 2009 [9], Deco and Jirsa, 2012 [11] and Deco et al., 2013 [13]. These models have achieved a large consistent set of results: the dependence of functional connectivity from the time on which it is calculated and the key role in the resting state modeling of the anatomical structure, of the noise and of the time delay. In these works it is stated that the resting state oscillations emerge from noise-induced transient fluctuations around the stable equilibrium state of the brain network.
Despite that, one year ago, Messé et al., [26] showed that a simple linear model was able to reproduce a prediction in the functional connectivity similar to the one obtained with these models.
Moreover empirical evidences show that changes both in strength and direction of functional connections appear at different time scales [2]. These observations shed light on the necessity of a new metric: the Functional Connectivity Dynamics. The chapter ends

with the analysis of the paper of Hansen et al., 2015 [19], where the authors were able to simulate the functional connectivity switching by enhancing the dynamics repertoire of the reduced Wong Wang model. The functional connectivity states, defined in the paper, are successfully connected to the default mode networks. The ideas and the strategy used in the 2015 paper of Hansen et al., are the starting point of this work.

3.1 Structural and functional connectivity

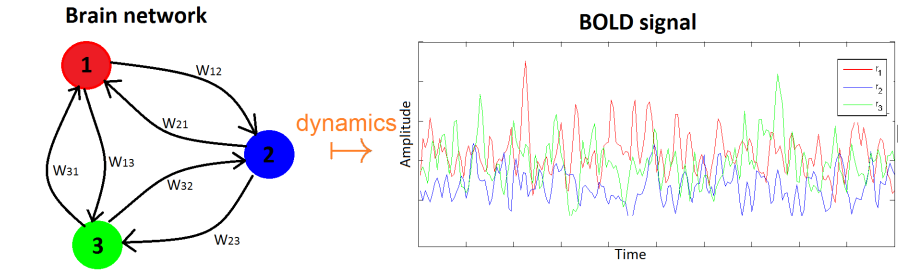
Before starting to analyze the resting state salient models it is necessary to define the empirical data on which they are applied to generate predictions (*the structural connectivity*) and then to discuss the measure used to compare the results obtained (the *functional connectivity*).

The structural connectivity

Structural Connectivity (SC) is the set of physical and structural connections linking the elements, neurons or bigger structures, within the network. Anatomical connections range from local circuits to large-scale network of inter-regional pathways.
The principal technique used to generate this kind of data is the Diffusion Tensor Imaging (DTI); this is an MRI based technique that was developed in the last decades and which tracks the diffusion of water molecules, that in the brain is maximally oriented along the axonal fibers. DTI can only track one direction per voxel and, for this reason, methods which improve this limitation have been introduced; for example Diffusion Spectrum Imaging (DSI), together with computational algorithms (tractography), is able to infer different orientations within a single voxel. The state of the art technique consists in using the photon tomography to map point-to-point connectivity between all anatomical regions in the animal brain. This innovative method produces the better data obtained until now in neuroscience and in this thesis innovative data of this kind are used to simulate the resting state dynamics (details in section 4.1).
These two techniques are compared and deeper explained in appendix A.

The functional connectivity

To compare the predicting power of the different models existing in the literature, some kind of measures have to be performed both on the simulated data and on the empirical one. One can be interested in the actual temporal evolution or in the mean measures, one can look for the presence of oscillations in particular frequency ranges or calculate



$$SC = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}; \; FC = \begin{bmatrix} corr[r_1(t), r_1(t)] & corr[r_1(t), r_2(t)] & corr[r_1(t), r_3(t)] \\ corr[r_2(t), r_1(t)] & corr[r_2(t), r_2(t)] & corr[r_2(t), r_3(t)] \\ corr[r_3(t), r_1(t)] & corr[r_3(t), r_2(t)] & corr[r_3(t), r_3(t)] \end{bmatrix};$$

Figure 3.1: Explicative cartoon of meaning of the structural and the functional connectivity in a network composed of three neural populations

the amount of mutual information or of transfer entropy. The range of possibilities is not narrow, but as simple linear correlation (for example covariance or Pearson correlation coefficients) proved to be informative on some characteristics of the network [11], it has become a common measure for the Functional Connectivity (FC).
Measuring the correlation of the time series, the functional connectivity is able to capture patterns of deviations from statistical independence between neuronal units. Functional connectivity is highly time-dependent and does not make any explicit reference to causal effects or to underlying structural model [33].
The resting state brain activity has been investigated with multiple procedures such as positron emission tomography (PET), magnetoencephalography (MEG), and electroencephalography (EEG), but the dominant approach is presently functional magnetic resonance imaging (fMRI). Resting-state fMRI is a non-invasive method in which the FC and other properties of Blood Oxygen Level Dependent (BOLD) signals are examined from scans acquired with no explicit task.

3.2 The RSNs and the dynamics repertoire

In general the resting state models have three principal building blocks (figure 3.2).
The *first building block* is the anatomical connectivity between brain areas (the structural connectivity defined in the previous section); the models analyzed in the present section use structural connectivity from human and macaque.
The *second building block* is the time delays of the signal between brain area due to finite

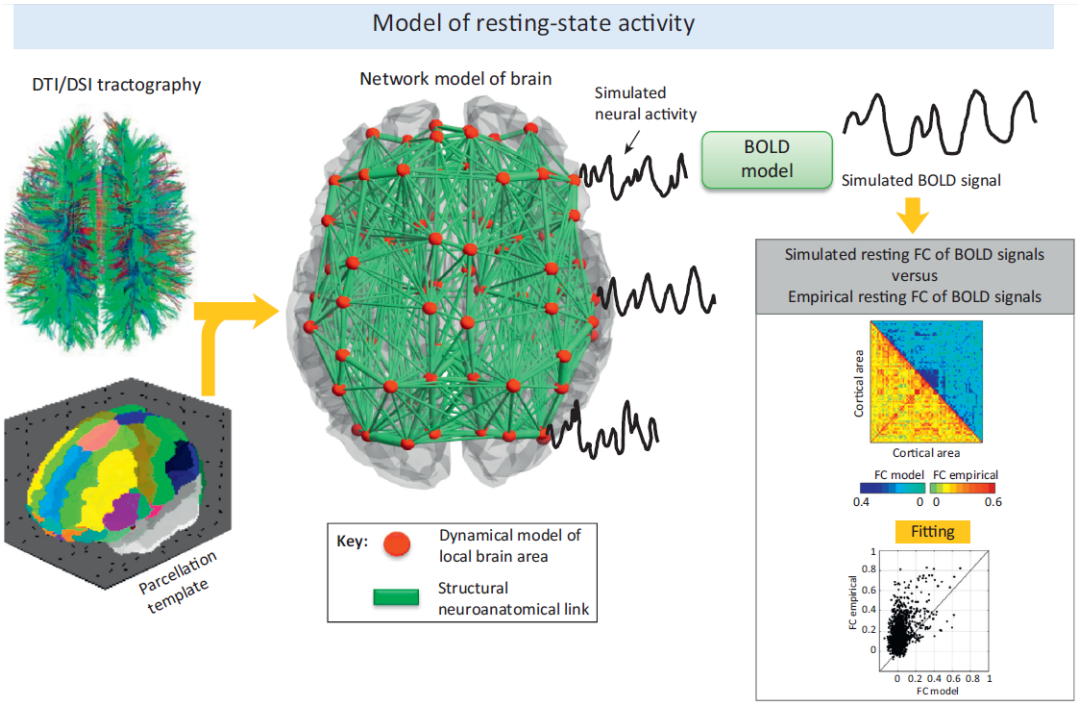


Figure 3.2: In this scheme the principal steps used in neuroscience to model the resting state activity are summarized. By applying the model to the empirical structural connectivity (in this case obtained with DTI or DSI) the dynamics of the brain is simulated; then, once the BOLD signals from the simulated brain activity are calculated, the simulated FC, i.e. the correlation in time of the time-series, is achieved. Finally comparing the simulated FC with the empirical FC (in this case obtained with fMRI) it is possible to establish the goodness of the model in reproducing the dynamics of the brain.
Source: Deco et al., 2013 [12] p. 270

nature of the speed propagation along nerve fibers. The paper of Ghosh et al., 2008 [18] and the paper of Deco et al., 2009 [9] state that the space-time structure of the couplings shapes the emergent network dynamics. Instead in the paper of Deco and Jirsa 2012 [11] the time delays are not used in modeling the RSN with spiking neurons because of absence of oscillations.

The *third building block* is the dynamics of the neural populations in the brain i.e. the dynamics of the network nodes. Some models considered simple oscillatory dynamics, as in the work of Ghosh et al., 2008 [18] and Deco et al., 2009 [9]; other models considered chaotic dynamics, as the first model analyzed in the next pages (Honey et al., 2007 [20]); finally, Deco and Jirsa 2012 [11] and Deco et al., 2013 [13], used the reduced Wong Wang

model in order to model a brain area as a realistic biological network with excitatory and inhibitory populations of spiking neurons coupled throw NMDA synaptic dynamics.
A common characteristic of all models is that the emergence of the resting state networks occurs only if the system is at the edge of instability.

3.2.1 Neural populations as chaotic oscillators

In the paper of 2007, Honey et al., [20] investigate functional connectivity patterns at multiple time scale; they simulate the resting state neural activity of a macaque using a large-scale connectome.
The units of the model describe interconnected excitatory and inhibitory neurons whose behaviors are determined by the membrane channels activity. They used the macaque structural connectivity to determine connections between the network nodes. Non linear instabilities generate activity in the system; the oscillations of the system are spontaneous and self organizing and they generate transient synchronous activations between the nodes i.e. brain regions. They calculated the correlation both on brain activity lasting 4 minutes and on brain activity lasting 30 s. They observed that the functional connectivity, calculated over long time series, reflects the anatomical connections; instead at smaller time windows, shorter-lived patterns of functional connectivity, that were not predicted by the anatomy, emerged. This observation shed light on the temporal dependence of the FC.
In order to compare the dynamics simulated with the empirical BOLD data, the oscillations for each node were converted into a BOLD signals using the Balloon - Windkessel haemodynamic model (Appendix B). The FC is computed as the correlation over BOLD time series, and the authors individuate the DMNs as the brain region correlated and anti-correlated. In conclusion, in this model, resting-state activity patterns result from cluster synchronization between nodes, and different clusters correspond to different DMNs.

The interesting result of this paper is that the functional connectivity depends on the time scale over which the correlation is calculated.
The principal two lacks of this model are that it does not include the noise and the delay in the propagation of the signals between brain regions [10].

3.2.2 The importance of noise and time delay in resting state fluctuations

Unlike the discussed paper of Honey et al., (2007) [20], in the work of Ghosh et al., (2008)[18] and in the work of Deco et al., (2009) [9], not only they introduce the noise and the time delay in modeling the resting state, but they also put the stress on the fundamental relevance of these factors.

In the paper of Ghosh et al., (2008) [18], the dynamics of each node of the network were modeled as oscillators with a constant value of noise. Specifically they associate at each node i.e., at each neural population, two state variables u and v; therefore the evolution in time of the network, composed of N nodes, is given by the following system:

$$\begin{cases} \dfrac{du_i}{dt} = g(u_i, v_i) - G\displaystyle\sum_{j=1}^{N} W_{ij} u_j(t - \Delta t_{ij}) + n_u(t) \\ \dfrac{dv_i}{dt} = h(u_i, v_i) + n_v(t) \end{cases} \tag{3.1}$$

where n_u and n_v represent gaussian noise and Δt_{ij} is the time delay of the signal between node i and j. W_{ij} is the weight of the anatomical connection between node i and node j and G is the coupling strength i.e. a parameter which scales all connection strengths without altering the connection topology of the structural connectivity matrix. The functions g and h represent the intrinsic dynamics of the node, that ,in this case, are:

$$\begin{cases} g(u_i, v_i) = \tau(v_i + \gamma u_i - \dfrac{u_i^3}{3}) \\ h(u_i, v_i) = \dfrac{1}{\tau}(u_i \alpha + b v_i) \end{cases} \tag{3.2}$$

the values of the parameter are in table 3.1.
In the paper the time delay Δt_{ij} between two coupled brain area i and j is defined as the

Parameters	Values
α	1.05
β	0.2
γ	1
τ	1.25

Table 3.1: Parameters used in the paper of Ghosh et al., (2008) [18] to implement the model for simulating activity of the brain during rest.

ratio between the Euclidean distance between the two nodes defined in the three dimensional space, and the propagation velocity of the impulse along the connecting fibers; this velocity is approximate to a constant. Using this model, the authors perform simulations with the purpose of finding the critical boundary between stable and unstable regime of the network in function of coupling strength and of transmission delay. The resulting stability diagram is shown in figure 3.3 where the degree of instability is quantified by the real part of the eigenvalue ($Re[\lambda]$) following a linear stability analysis of the network equilibrium points.

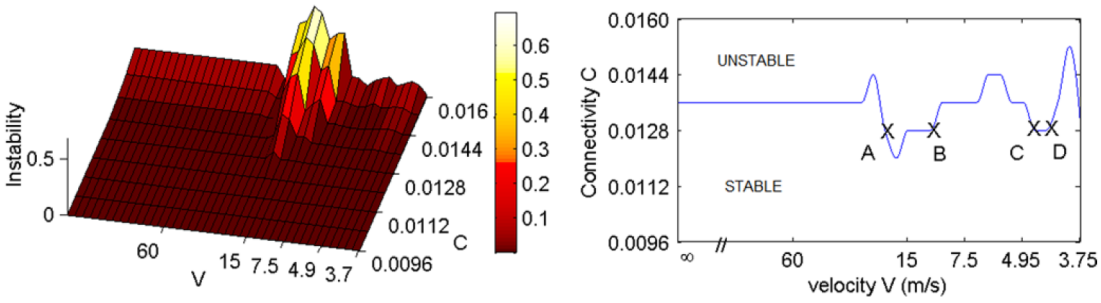


Figure 3.3: In the upper part of the figure the degree of instability, equivalent to Re[λ], is plotted as a function of connection strength and of propagation velocity. In the lower part of the figure the critical boundary, equivalent to Re[λ] = 0, is plotted as a contour line separating unstable and stable regions. The intervals $[A; B]$ and $[C; D]$ indicate biological relevant velocity ranges.
Source: Ghosh et al.,2008 [18] p.7

Adding the noise, the behavior of the system in the different regimes is studied. When the system is in the the stable regime, far from the critical boundary, it is observed that all oscillations are strongly damped; in the unstable regime, the system displays high amplitude oscillatory behavior, which resembles pathological, as epileptic, activity. Instead when the system is close to the critical boundary, it is possible to observe the emergence of the characteristic oscillatory behavior of the resting state ($\sim 10Hz$).
The ongoing interplay between noise drive and oscillatory return leads to the exploration of the brain's dynamics repertoire. A effective way to visualize this stochastic mechanism is the one used by Deco et al., (2011) [10] and shown in figure 3.4.

Implementing the present model with a real anatomy connectivity and with a biological realistic propagation velocity, the BOLD signals are computed; in this way realistic resting state networks are founded.
The work of Ghosh et al., demonstrate the importance of the spatial (anatomical connectivity) and temporal (time delay) structure to construct realistic simulation of the resting

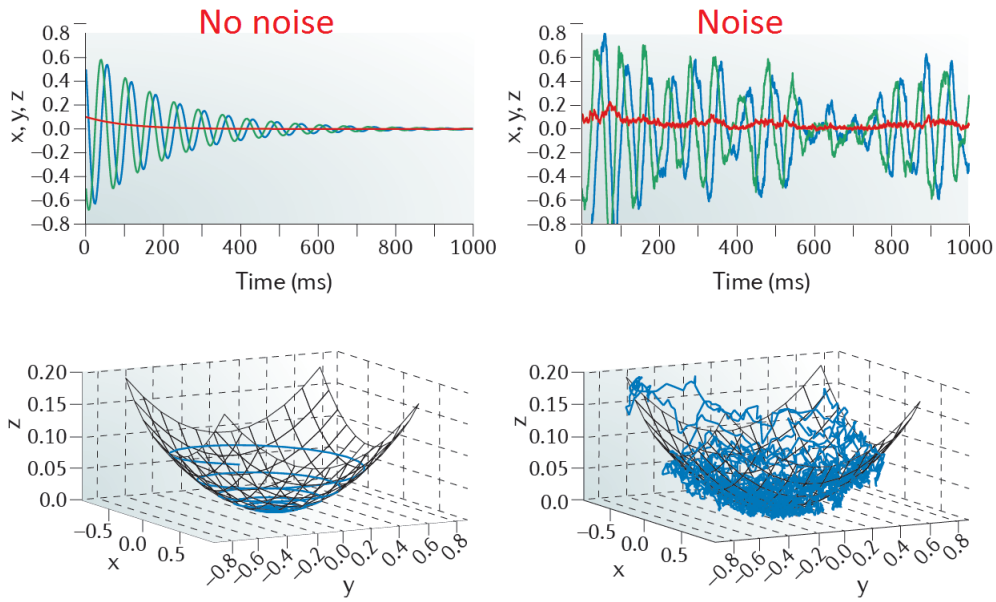


Figure 3.4: In this figure are shown the the time series (upper part) and the phase plane (lower part) of a three dimensional system with an equilibrium point in order to illustrate the stochastic mechanism that leads to the resting state oscillations in the Ghosh model [18]. In the absence of noise (left part of the figure) the system approach to the stable attractor by spiraling down a paraboloid; the corresponding time series of two of the three system variables display a damped oscillation (in green and blue), the third one (in red) relaxes to zero in a non-oscillatory way. In the presence of noise (right part of the figure) the system explores the neighborhood of the equilibrium point. Each excursion further away from the equilibrium is followed by an oscillatory return along the paraboloid. This behavior is revealed in the time series with intermittent, fast neurophysiological oscillations.
Source: Deco et al.,2011 [10] p.50

state. When the system is closed to the instability, it is able to explore its dynamical repertoire, i.e. a number of potential functional network configurations: the default mode networks.
The Ghosh et al. results were confirmed by the paper of Deco et al., 2009 [9], where it is underlined that the synchronization between brain area in the resting state strongly depends of three factors: the neuroanatomical connectivity structure, the delays in the transmission of information between different brain nodes and the noisy fluctuations.
In summary, the most important finding of these two works is that spontaneous ongoing activity is built up by multistable attractors, each one related to different specific tasks or

stimulations; the resting state activity fluctuates through transitions between attractor states by reason of unstructured input: the noise [10].

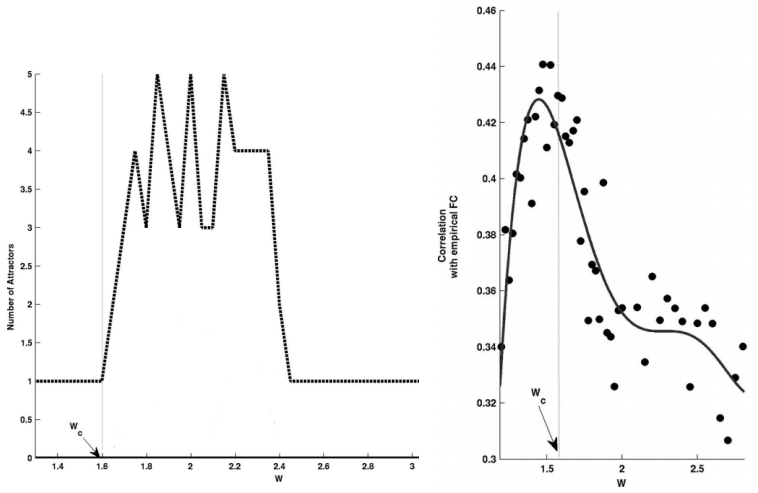
3.2.3 Resting state modeling with the reduced Wong Wang model

Starting from the previous consideration, i.e. the resting state fluctuations emerge if the system is at the edge of instability, Deco and Jirsa in their paper of 2012 [11] decide to use a realistic model to explain the dynamics of each brain area: the reduced Wong Wang model. The stationary fix points of the models as function of the coupling strength is studied; the attractor landscape of the system (shown in figure 3.5a) is obtained by simulating the deterministic evolution of the model with 1000 different initial conditions for each value of the coupling strength; in this way three different regimes are founded:

1. for small value of the coupling strength ($W < W_{-}$) only one stable attractor exists. This attractor corresponds to the spontaneous ground state of the system where all neurons are firing at low level of activity.
2. For increasing value of the coupling strength there are more than one attractor; this leads to a multistable regime of the system. These attractors correspond to high firing activity in particular brain area.
3. For very large value of the coupling strength only one attractor is stable: this corresponds to the state in which all the excitatory neurons are highly activated in the brain (epileptic state).

The authors simulated the BOLD signals and, calculating the correlation matrix of the BOLD activity between all brain areas, they obtained the simulated FC. Plotting the correlation between the empirical FC and the simulated FC as function of the coupling strength (figure 3.5b), they were able to state that the correlation is maximum for values of the coupling strength near the bifurcation as shown in figure 3.5a. In conclusion, using a biological realistic model, Deco and Jirsa were able to prove that the best fit between empirical and simulated data is obtained when the brain network is at the critical point. Default mode networks emerge as structured noise fluctuations around the stable low activity state induced by the presence of a latent "ghost" multistable attractors at the edge of the bifurcation.

Stated the effectiveness of the Wong Wang model in predicting the resting state fluctuations, Deco et al., in the paper of 2013 reduced this model to only one variable as already discussed in section 2.4.3. As in the previous paper, they evaluated the number of the system stationary states running deterministic simulations of the model with 1000 different initial conditions. They plotted the maximum firing rate activity over all cortical



(a) (b)

Figure 3.5: (a) Attractor landscape of the system in function of the coupling strength. In the graph (b) dots represent the correlation between the empirical and the simulated FC. *Source:* Deco and Jirsa 2012 [11] p. 3371

areas as function of the coupling strength (figure 3.6 B) for both the reduced model and the original one.

The value of the external input current (I_0, equation 2.90) of the reduced model was tuned in order to get a structure of the stationary attractors of the global system similar to the one obtained by using the complete Wong Wang model. In this way the bifurcation diagram (figure 3.6 B) displays three different regimes as found in the previous work (Deco and Jirsa 2012 [11]). For small values of the coupling strength only one stable state exists, characterized by a low firing activity in all cortical areas. For a critical value of the coupling strength, it emerges a first bifurcation and new multistable states of high activity appear while the state of low activity is still stable. For even larger values of coupling strength the spontaneous state becomes unstable.

The correlation, obtained in the paper, between the empirical functional connectivity and the simulated functional connectivity with the reduced model, is shown in figure 3.6 A. The maximal correlation occurs at the edge of the second bifurcation where the low firing rate state loses its stability and, therefore, the noise drives the system in the exploration of the other attractors shaped by the underlying anatomy.

An explication of the meaning of the state of the network brain when it is at the edge of bifurcation is presented in another work of Deco et al., 2013 [12]. The authors observed that, at the point of criticality, neuronal networks optimize several aspects of information

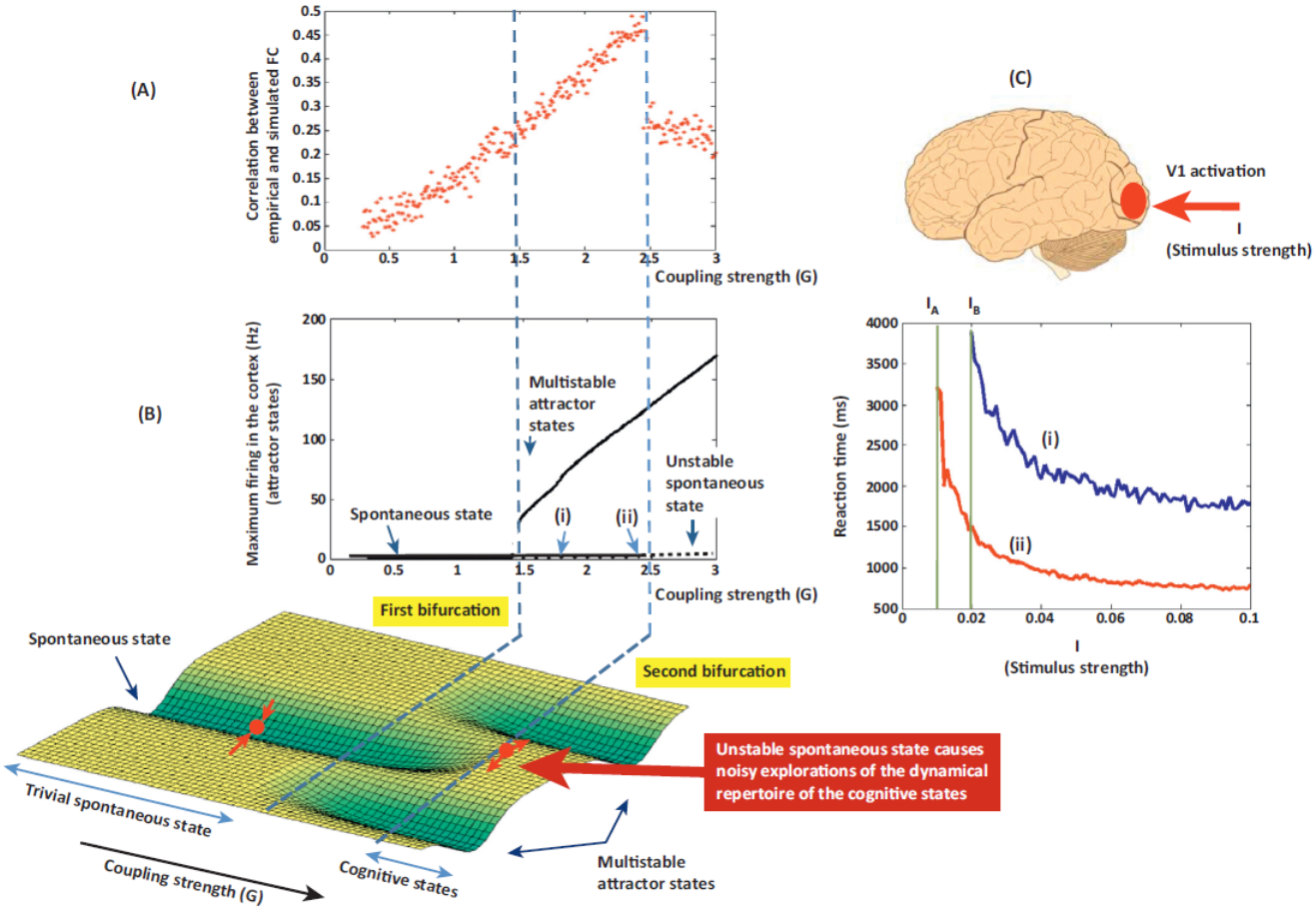


Figure 3.6: Correlation between the empirical and the simulated functional connectivity with the reduced model is plotted as function of the coupling strength (part **A** of the figure).
In the part **B** of the figure the maximum firing rates activity in function of the coupling strength is shown; in the graph the three different regimes discussed can be observed.
In the part **C** of the figure it is plotted the time required to the visual network to converge to the appropriate attractor after a stimulation in area V1. The system is implemented with the characteristic of two different working points: one at the edge of the second bifurcation (point *(ii)*, red line) and the other one further away (point *(i)*, blue line).
Source: Deco et al.,2013 [12] p. 271

processing. They simulated the stimulation of the visual network by the application of an external bias in the V1 area; for different strengths of the external stimulus I, they measured the time that the system required to converge to the appropriate network at-

tractor. In figure 3.6 C it is shown the reaction time in function of the strength of the stimulus when the system is implemented with the characteristics of two different working points (figure 3.6 B): one at the edge of the second bifurcation (in red) and the other further away (in blue). The optimal working point is the one at the brink of the second bifurcation; in this point the system is maximally sensitive to external stimulation and able to respond quickly.

3.3 Non-stationarity of the Functional Connectivity

Nevertheless the biological realism of the models discussed in the previous section, Messé et al., in 2014 paper [26], demonstrate that the accord between empirical and predicted FC of these models is good as the accord between empirical and predicted FC with a simple linear model.
Empirical evidences, as the one contained in the paper of Allen et al., 2014 [2], show that the FC calculated upon smaller time series is significantly different from the one calculated upon longer time series.
These observations conduct at the paper on which is based the work of this thesis: the paper of Hansen et al., 2015 [19].

In what follows the three mentioned papers will be summarized.

3.3.1 Evidence of functional connectivity dynamics

Up to this point, when treating FC, it is implicitly assumed that the relationships between brain areas are constant throughout the length of recording; instead resting state networks have a rich spatiotemporal signature. As stated for the first time in the paper of Honey et al., 2007 [20] (see section 3.2.1), the calculation of the correlation across different length of time leads different results. Across a long time window, distinct functional networks, related to the underlying anatomy, emerge; however at shorter time steps, subnetworks emerge and dissolve as the full network evolves (figure 3.7). Consequently quantifying changes in FC metrics over time may provide greater insight into fundamental properties of brain networks.
The temporal dependency of the functional connectivity has been empirically validated in fMRI studies as the one of Allen et al., 2014 [2]. In this work resting state data from a large sample (n=405) of human are used and for the analysis the sliding window approach is used. In this approach a time window of fixed length is selected, and data points within that window are used to calculate FC. The window is then shifted in time by a fixed number of data that define the amount of overlap between successive windows (figure 3.7). By

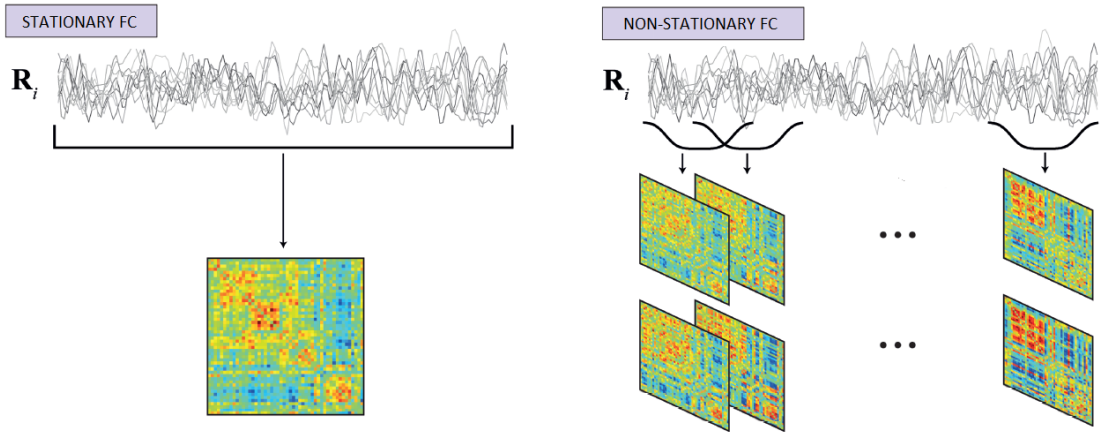


Figure 3.7: Schematic representation of sliding windows analysis.
Source: Adapted from Allen et al.,2014 [2] p. 3

the analysis of the windowed FC matrices, the authors not only were able to demonstrate the dynamic behavior of the functional connectivity, but, performing clustering analysis, they individuate recurring short-term connectivity patterns which they called *FC states*. A part of the individuated FC states diverge strongly from connectivity patterns individuated by global FC.
In conclusion FC, when estimated over long time series (called stationary FC), breaks down into a variety of correlation patterns that can be observed only if the estimate is done over short time windows. Since the correlation over long time series leads, for the definition of the Pearson correlation, to a loss of information about the dynamics evolution of the system, the stationary functional connectivity is mainly related to the underlying anatomy. On the other hand, once FC is estimated over short time windows, it mostly reflects recurrent transitory patterns that aggregate when the FC estimate is done on a whole session.
An other interesting result is obtained in the paper of Messé et al., 2014 [26]. The authors investigate the relative contributions in the emergence of the empirical FC of anatomical connections, stationary and non-stationary dynamics. In order to do that, the predicting power of increasing complexity models is systematically compared. The predicting power is calculated as Pearson correlation between the simulated and the empirical FC. The simplest model considered is a spatial autoregressive model (*SAR*, as called in the paper) which is explicitly associated with a stationary FC; more complex models are expected to generate dynamics that are compatible with the non-stationarity of the FC. Despite of the expectation, the authors demonstrate that all models tend to perform similarly, irrespective of model complexity (figure 3.8). The best performing model in most cases

is the SAR model, which, for its definition, is connected to the stationary part of the FC, although a significant part of the variance in the empirical data remains unexplained (the 63% according to the paper results). This large remaining variance must contain the non-stationary behavior of the FC as evidenced in literature and that, the models discussed until know, are not able to capture.

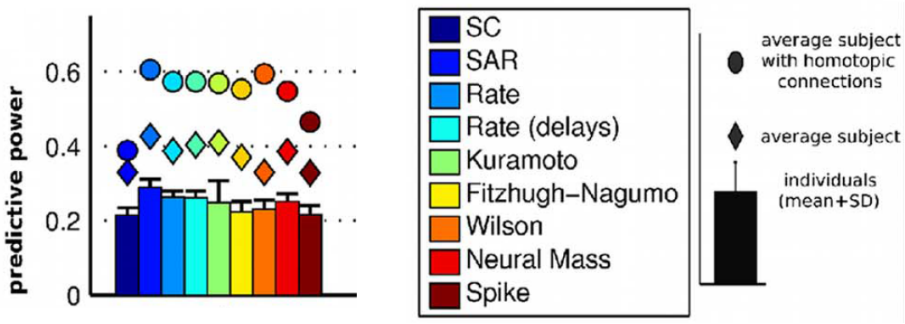


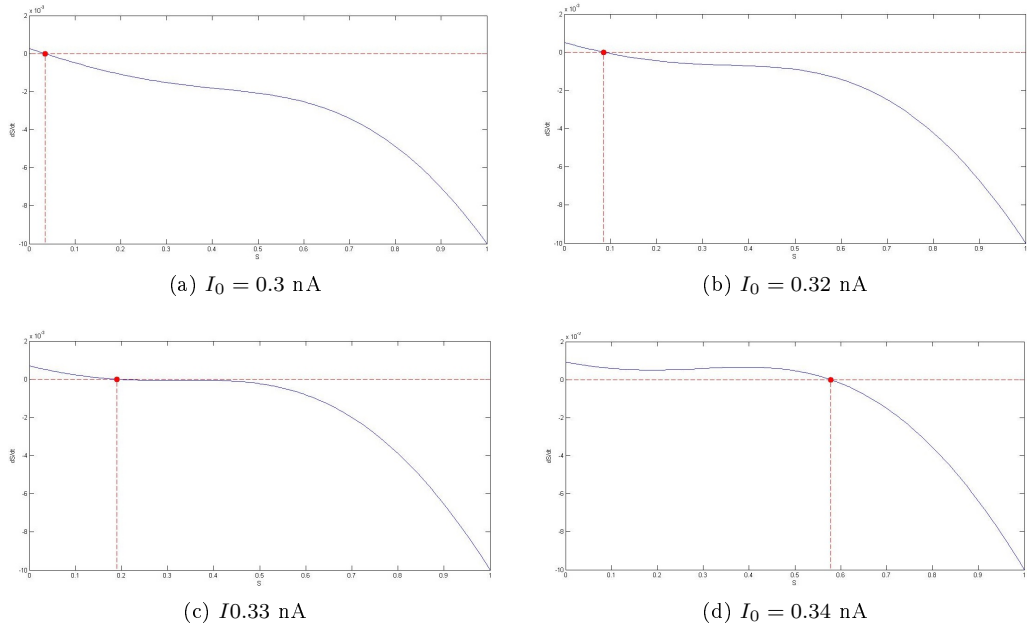
Figure 3.8: Predictive power of the model analyzed in the study of Messé et al.,. Bar charts represent means and associated standard deviations of the individual predictive powers. The diamonds represent the predictive powers for average subject using the original SC obtained by fiber tractography based on diffusion weighted imaging (DWI). The circles represent the predictive powers for average subject using the SC with addition of connections across cerebral hemispheres that are usually estimated improperly using DWI (homotopic connections).
Source: Adapted from Messé et al., 2014 [26] p.3

3.3.2 Exploring the dynamics of the FC

The work of Hansen, Battaglia, Spiegler, Deco and Jirsa "Functional connectivity dynamics: Modeling the switching behavior of the resting state", published in the current year, represents a new stance, necessary at this time, to study the resting state dynamics. The authors were able to simulate the transition between states at rest.
The structural connectivity matrix, used to reproduce the brain activity, represents the connections between 66 anatomical cortical regions of human brain obtained by the diffusion spectrum imaging.
In order to model the resting state activity the authors considered two models; the first one is the reduced Wong Wang model, with the reduction performed in Deco et al., 2013 [13], and, following the authors label, in the next this model will be referred as MFM (Mean Field Model).

In the paper the dynamics of a single node (setting the coupling strength to zero G=0 in equation 2.90) is studied analyzing the relationship between NMDA current related variable S and its time-derivative dS/dt as it is shown in figure 3.9 and 3.10. Remembering that the values of S for which dS/dt is equal to zero correspond to fixed points of neural activation, it is possible to notice that a variation of the local excitability I_0 and the local excitatory recurrence w leads to different regimes.
Fixing the local excitatory recurrence at 0.9, the analysis of the phase portrait of the model (figure 3.9) reveals that there is only one stable equilibrium for different, realistically biological values of the local excitability I_0; the stable point corresponds to low activity regime (figures 3.9a, 3.9b, 3.9c), and to high activity regime (figure 3.9d).
Interestingly, changing the value of the recurrent connections w from 0.9 to 1.0, it is possible to notice that when I_0 is in the range between 0.32 and 0.34 nA there are two new equilibria, one stable and one unstable (figure 3.10c); this situation corresponds to a bistable configuration with two fixed point (high and low activation). For other values of I_0 there is only one stable equilibrium (figure 3.10a, 3.10c and 3.10d).

In conclusion, when each brain region is considered isolated from other brain regions ($G = 0$), the increase of the recurrent connections w causes the dynamics of each brain region to change from monostable to bistable for certain values of the external input I_0. Therefore, in order to increase the dynamic repertoire of the mean field system, the authors decided to model the resting state with the *enhanced non-linearity mean-field model* (eMFM); such model shares with the MFM the same dynamical equations but with a slightly modification of the parameters.
To understand how a change in the coupling strength G produces a change in the dynamics of the system, for both models 1000 deterministic simulations, using random initial conditions, were run for each value of G in the range of $[0 : 3.25]$ with increments of $\Delta = 0.05$. Effectively, in absence of noise, the system converges to fixed point of activity in which each brain region displays a steady level of activity. The final stable patterns of the simulations of each model are correlated with two measurements from graph theory calculated on the structural connectivity matrix: the in-strength and the s-core. The in-strength of a brain region i represents the density of ingoing white matter connections connected to region i; the in-strength is a descriptor of local topology. The s-core of a network is the maximal connected subnetwork in which all nodes have at least a value of strength of s or higher; the nodes of the s-core capture the global aspect of the network topology.
In figure 3.11 are shown the graphs obtained for the MFM and the eMFM correlating the stable patterns with the in-strength (blue dots) and with the s-core (red dots). Observing



(a) $I_0 = 0.3$ nA

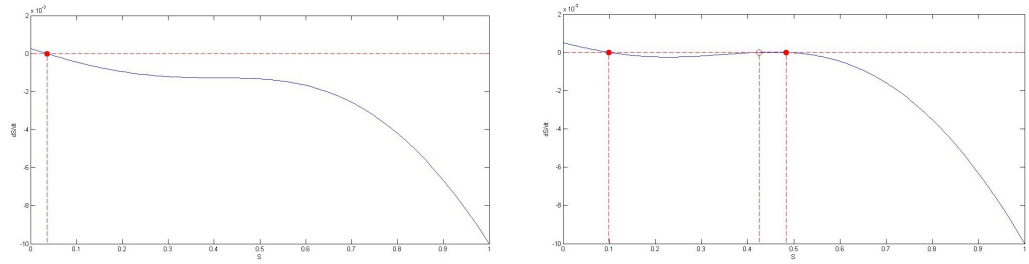
(b) $I_0 = 0.32$ nA

(c) $I0.33$ nA

(d) $I_0 = 0.34$ nA

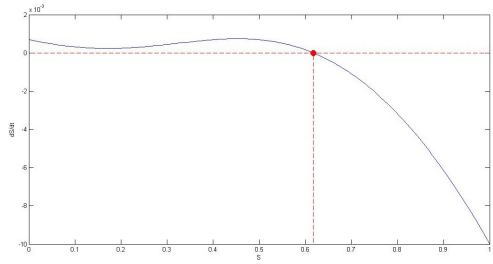
Figure 3.9: In the graphs is shown the neural activation dynamics at the level of a single brain region when the local excitatory recurrence w is fixed at 0.9. Varying the value of the local excitability I_0 there is only a stable fixed point (filled dots) that corresponds to low activation regime in figure 3.9a,3.9b, 3.9c, and to high activation regime in figure 3.9d

the graph it is possible to notice that in the range $G_{c-} < G < G_{c+}$ the dynamic repertoire of the MFM is enriched, as already notice in Deco and Jirsa 2012 [11] and in Deco et al., 2013 [13]; however all the sampled fixed points are strongly correlated with the local SC topology (in-strength) or with the global SC topology (s-core). This evidence confirms the results obtained in the study of Messé et al., 2014 [26] : the brain dynamics simulated with the MFM reflects the underlying anatomical structure. In the eMFM correlation graph (right part of figure 3.11) additional classes of fixed points emerge which has poor correlations with both the local and the global SC topology. As expected, the local multistability of the eMFM (figure 3.10b) reinforces the global emergence of multi-stability; moreover the presence of these new classes of attractors mirrors a non-trivial interplay between local and collective dynamics that can not be explained by the SC alone. This result represents the step forward respect to the models analyzed by Messé et al., 2014 [26]. The goodness of the eMFM in reproducing the resting state activity is stressed by the fact that with such a model the authors were able to observe the switching in FC. In

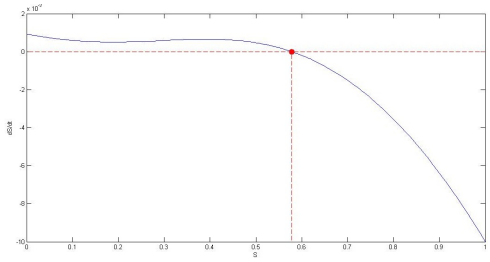


(a) $I_0 = 0.3$ nA

(b) $I_0 = 0.32$ nA



(c) $I0.33$ nA



(d) $I_0 = 0.4$ nA

Figure 3.10: The graphs show the neural activation dynamics at the level of a single brain region when the local excitatory recurrence w is fixed at 1. For the value of the local excitability I_0 equal to 0.3 nA there is only a stable fixed point (filled dots) that correspond to a low activation regime (fig 3.10a); at the value of I_0 equal to 0.32 nA there is also an unstable fixed point (empty dots) that leads the system in a bistable condition between high and low activity regimes (fig 3.10b). For values of I_0 greater than 0.32 nA the isolated brain area exhibits a stable high activation regime (a stable high activity fixed point).

order to catch the FC switching, the *Functional Connectivity Dynamics* (FCD) matrix is constructed as explained in the following (an explanatory cartoon of the procedure is shown in figure 4.3 in the next chapter). The BOLD signal time-series is divided in windows of 60 seconds duration, centered at time t, with an overlap of 58 seconds between windows. The functional connectivity $FC(t_i)$ of the BOLD time-series of each time windows is calculated. Finally the Pearson correlation is computed between the components of the upper triangular part of each FC(t_i) arranged in a vector. Such correlations are organized in a matrix: the functional connectivity dynamics matrix, the FCD, which is the new metric able to catch the rich structure of the FC non stationarity behavior.

The coupling strength and noise value have an important role in shaping FC dynamics

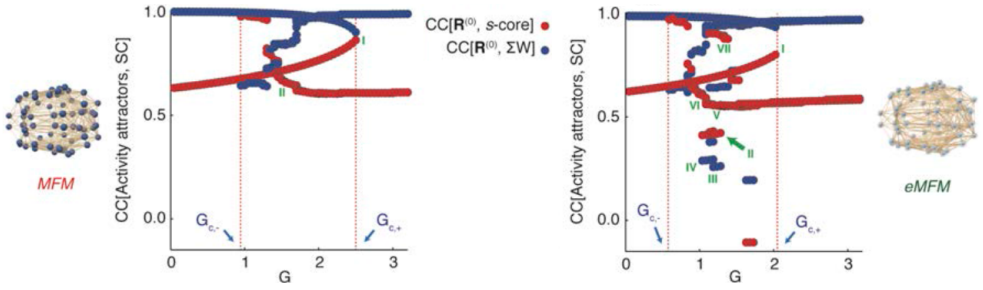


Figure 3.11: The graphs represent the correlation between the deterministic attractor of the neural activity (MFM, left and eMFM, right) and the total in-strength of each brain region (blue dots) and the s-core (red dots) of the SC matrix. The fixed points of activity for each model were found by running 1000 deterministic simulations lasting 15 s for every value of the coupling strength G. The correlation points accumulate to form a discrete set of different branches; this indicates that different families of attractors exist with distinctive relations to local and global topology. G_{c^-} and G_{c+} indicate the bifurcation points. It is interesting to notice that in the eMFM correlation graph there are values of correlation not present in the MFM correlation graph. The Roman numerals indicate the working point used as initial conditions for simulations.
Source: Hansen et al., 2015 [19] Supplementary data, figure S5

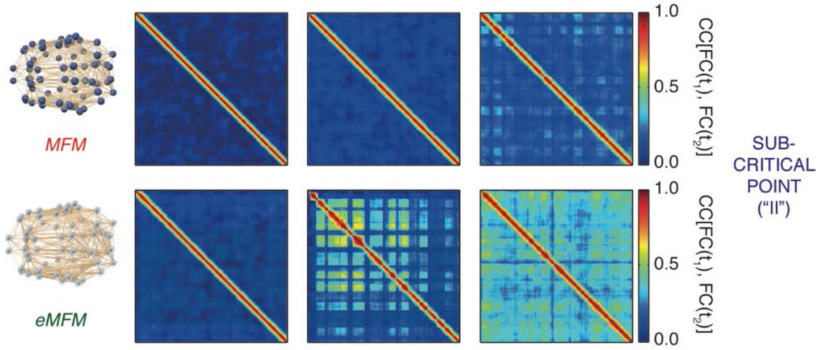


Figure 3.12: FCD matrices for different noise levels σ and for the coupling strength value II with reference to figure 3.11. In both models the FC switching occurs at high noise level, but the FCD generated with the eMFM displays a richer structure with respect to the one obtained with the MFM, mirroring its bistable regime at the brain region level.
Source: Hansen et al.,2015 [19] Supplementary data, figure S3.

for both eMFM and MFM. Indeed, observing figure 3.12, it is possible to notice that also MFM gives rise to structured non-stationary patterns, when G is selected at the

center of the range $[G_{c-}; G_{c+}]$ and the noise is increased until $\sigma = 0.01$. However this patterns are always less prominent that for eMFM; the only significant correlations are in the main diagonal, which means that only functional connectivity states close in time are correlated. In both model the FC switching is suppressed in presence of low level of noise ($\sigma = 0.001$) and when the coupling strength is set near the critical points (G_{c-} or G_{c+}). Such behavior is due, in both models, to a poor dynamical repertoire: the dynamics of the network is settled in a fixed point. At the contrary, when the system is maintained out-of-equilibrium by higher noise levels, the trajectories of the dynamics of the brain system explore different subspaces; the trajectories either converge back towards the fixed point from which they originated or diverge towards a different attractors.
Observing the empirical and the eMFM FCD in figure 3.13, it is possible to notice the

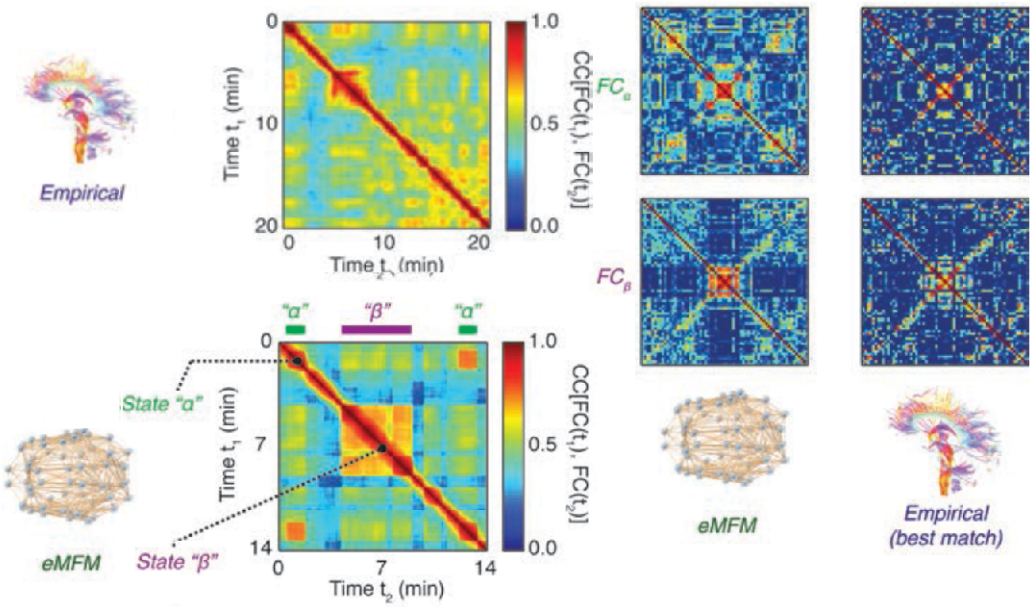


Figure 3.13: In the left part of the figure the empirical and the eMFM FCD matrices are shown; α and β indicate the epoch of stability of the FC. In the right part of the figure are shown the FC states of the epoch of stability and the best match with the FC(t) found in the empirical data set.
Source: Hansen et al.,2015 [19] p. 530

presence of red, square-shaped blocks occurring along the diagonal that are the signs of the presence of FC(t)s closely correlated during epochs lasting several minutes. Using the usnusperving clustering method (K-means) the boundaries between blocks are precisely defined; following the work of Allen et al., 2014 [2], the FC(t) clusters are called FC

states. The two epochs found in the work, α and β, are shown in the right part of the figure 3.13 where are also displayed the closest matches found in the empirical data-set. Finally in the paper the authors compared the simulated FC states with the resting state network in literature.
In conclusion the work of Hansen et al., 2015 shows that the previous large scale model (MFM) does not reproduce the switching between the FC states observed in the empirical data (Allen et al., 2014 [2]). Instead a slight enhancement of the non-linearity of the network nodes leads the system to enrich its repertoire of attractors and consequently the simulated FC to display non-stationary switching similar to empirical resting state recordings.

Part II

ORIGINAL WORK AND RESULTS

Chapter 4

Materials and methods

In this chapter materials and methods used to individuate the functional connections in mouse brain resting state activity simulation are discussed.
Firstly the matrix used to define the anatomical connections between the brain regions is presented; such matrix is an anterograde tracer mapping of axonal projections across the mouse central nervous system and it is obtained in the Allen Institute for Brain Science [28]. Such structural connectivity matrix represents, to date, the most comprehensive connectivity map in any vertebrate species. Consequently, this instrument is an extraordinary opportunity for understanding the mammalian brain, and specifically for simulating the resting state with a careful biological realistic connectivity structure. The first section briefly analyzes the matrix using graph theory measures.
Section 4.2 describes the mathematical frameworks used for modeling each local brain area, i.e. each node in the global brain network. Such models are the (*i*) Wong Wang model with the further reduction performed by Deco et al., 2013 [13] (the *Mean Field Model*, MFM) and (*ii*) the *enhanced non-linearity Mean-Field Model* (eMFM). The eMFM shares with the MFM the same dynamical equations, but with the modification of a parameter value which causes, at the brain areas level, the introduction of bistability between high and low firing rate states; in the global network such behavior is reflected in presence of attractors which don't have any trivial relationship with the local and global network topology (as shown in the paper of Hansen et al., 2015 [19] discussed in section 3.3.2).
Then the characteristics of the computer simulations are briefly summarized.
The method used to calculate the functional connectivity matrix and the functional connectivity dynamics matrix starting from the simulated BOLD signals, is illustrated in section 4.4.
After that, the approach employed to individuate the functional brain hubs is explained.

In the same section are also summarized the results obtained in the paper of Mechling et al., 2014 [25] experimenting resting state fMRI on mice; this article is of interest since the experimentally functional hubs found by the authors are compared with the ones detected in the simulation under discussion.
Finally a strategy to reduce the dimensionality of the simulated brain activity is presented.

4.1 The structural connectivity

In this work the state-of-the-art mouse brain anatomical connectivity obtained in the Allen Institute for Brain Science of Washington (98103 USA) is used. Such structural connectivity is an anterograde tracer mapping of axonal projections across the mouse central nervous system and it is, to date, the most comprehensive connectivity map in any vertebrate species [28].
The method used to map the brain connections is the anterograde tracer method. Such method permits to trace the axonal projections. This goal is obtained by the injection of observable tracer molecules which are absorbed locally by the soma of various neurons and transported to the axon terminals; in this way it is possible to detect the tracer migration and consequently to trace axonal projections from their source, soma, to their point of termination, synapse. In particular, to obtain the Allen connectivity matrix, as explained in the study of Oh et al., 2014 [28], the axonal projections were mapped by injecting in adult male C57Bl/6J mice the recombinant adeno-associated virus which expresses the enhanced green fluorescent protein (EGFP) anterograde tracer. The green fluorescent protein exhibits bright green fluorescence when exposed to light in range from blue to ultraviolet, and it was isolated for the first time in the jellyfish Aequorea victoria. The tracer migration signal was detected with a serial two-photon tomography system with voxel resolution of $100\,\mu$m x$100\,\mu$m x$100\,\mu$m. Repeating systematically this approach, and after several data processing and quality control, the whole brain connectivity matrix was obtained. The matrix, plotted in figure 4.1, contains the connection strengths between 426 brain regions.
The connectivity matrix reveals that the mouse brain exhibits features of both small-world networks and scale-free networks [28], i.e. the presence of hubs and the presence of clusters of nodes; the cluster of nodes are subsets of nodes that have more connections inside than outside the subset.
Unlike DTI and DSI methods, which force the connectivity matrix to be symmetric, the tracer method catches the asymmetry of the connectivity matrix and consequently, when used as base for brain simulations, allows to reproduce more realistic behaviors of the network (details about the comparison of such methods are in appendix A). The asym-

metry of the anatomical structure is revealed by the difference between the in-strength and the out-strength of each brain region, i.e. respectively the sum of the incoming and the outcoming connection weights (figure 4.2). Both the in-strength and the out-strength distributions have the characteristics of log-normal distributions with a small number of strong connections and a large number of weak connections [28].

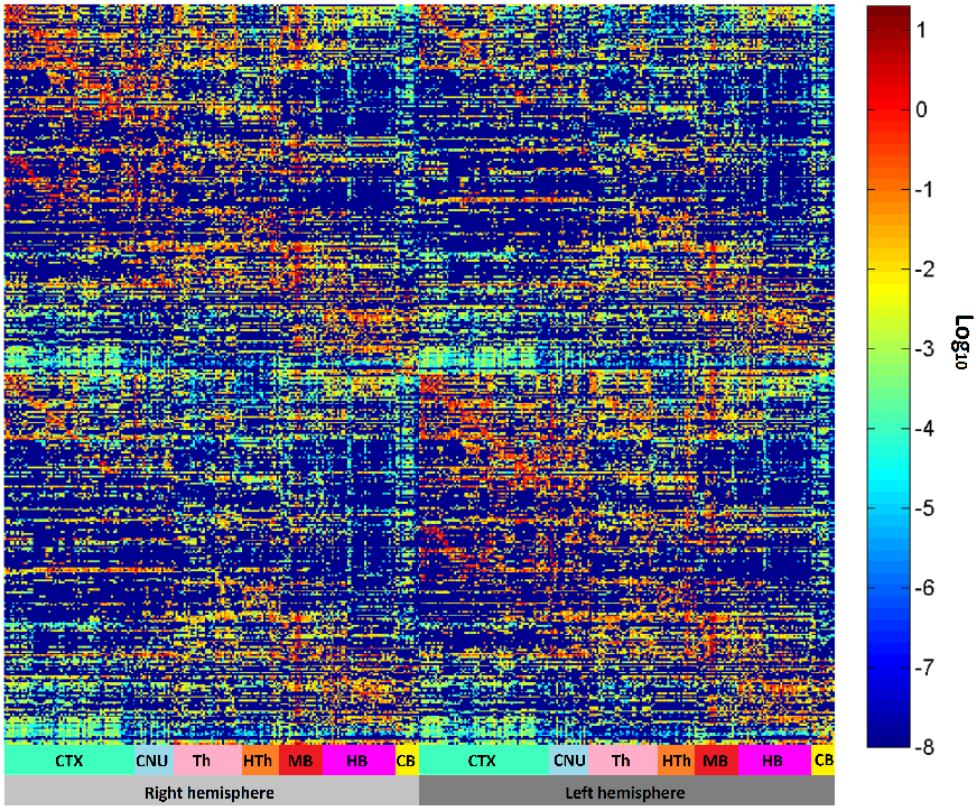


Figure 4.1: The color map represents the structural connectivity matrix by encoding the connections strengths by base ten logarithmic scale (for identifying the brain region see the abbreviation list).

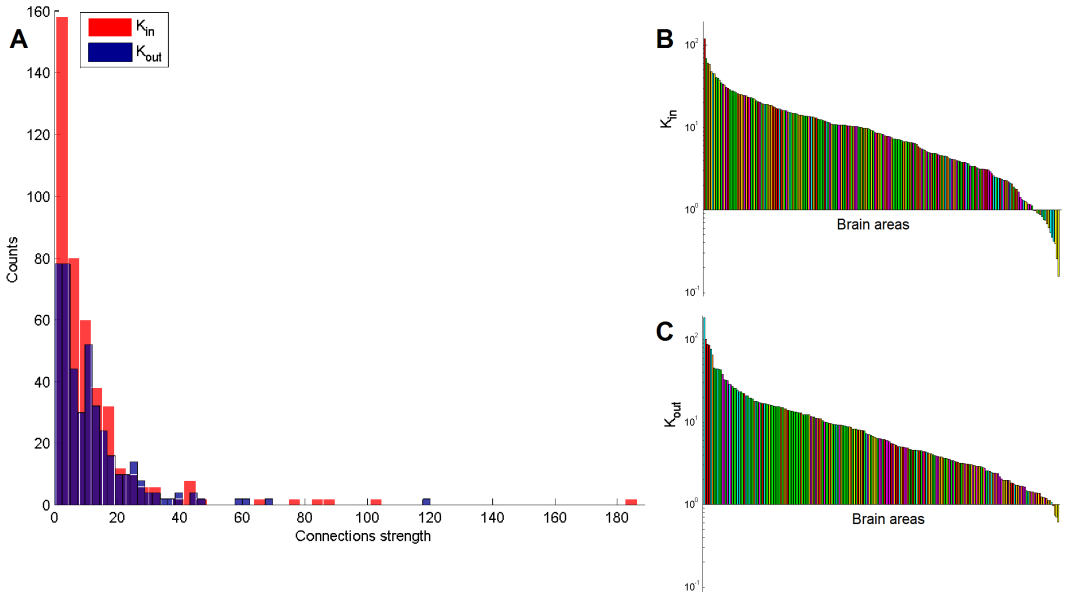


Figure 4.2: (**A**) Distribution of the in-strength (red) and out-strength (blue) connections of the network nodes. Both distributions have a log-normal shape i.e., long tailed distributions with a small number of strong connections and a large number of weak connections. In-strenght (**B**) and out-stength (**C**) grading of the anatomical connections relative to each brain area in logarithmic scale. Different colors label different groups of brain regions (such as cerebral cortex) with reference to figure 4.1.

4.2 The brain models

The mean activity of each brain region, composing the mouse brain network, is described by the reduced Wong Wang model. In this approach, deeply analyzed in section 2.4, the dynamics of a brain region is given by the whole dynamics of excitatory and inhibitory populations of LIF neurons interconnected via NMDA synapses. In this work we take in account this model with the further reduction performed in Deco et al., 2013 [13], analyzed in section 2.4.3. In this framework, the dynamics of the output synaptic NMDA gating variable S of a local brain area i is strictly bound to the collective firing rate H_i,

and is given by the following coupled equations:

$$\frac{dS_i(t)}{dt} = -\frac{S_i}{\tau_s} + (1 - S_i)\gamma H_i + \sigma\eta_i(t) \tag{4.1}$$

$$H_i = \frac{ax_i - b}{1 - exp(-d(ax_i - b))} \tag{4.2}$$

$$x_i = wJ_N S_i + J_N G \sum_j C_{ij} S_j + I_0 \tag{4.3}$$

where x_i is the synaptic input to the i-th region. γ is a kinetic parameter fixed to 0.641, τ_s is the NMDA decay time constant and its value is 100 ms; a,b and d are the parameters of the input and output function H and are respectively equal to 270 nC^{-1}, 108 Hz, 0.154 s. G is the coupling strength i.e., a scalar parameter which scales all the connection strengths C_{ij} without altering the connection topology of the anatomical connectivity matrix; the value of G, together with the value of the noise amplitude σ of the normally distributed stochastic variable η_i, are tuned to find the optimal network regime as explained in the next section. $J_N = 0.2609$ nA scales the synaptic input current and I_0 is the external input current sett to 0.3 nA. Finally w is the local excitatory recurrence.
The difference in the two models implemented in this work (i.e. the *Mean Field Model* (MFM) and the *enhanced non-linearity Mean-Field Model* (eMFM)) lies in the value of the parameter w: in the MFM $w = 0.9$, in the eMFM $w = 1$. Indeed a modification of such parameter is able to change the kind of model since it is able to modify the dynamics at the level of a single brain area and consequently to modify the dynamics of the global brain network. This behavior was noted in the work of Hansen et al., 2015 [19], as summarized in section 3.3.2. When $w = 0.9$ the analysis of the phase portrait (figure 3.9) reveals that there is only one stable equilibrium (MFM). On the other hand, when w is set to 1, as in the eMFM (figure 3.10), the system is characterized by two stable and one unstable equilibrium points; this configuration leads the system in a bistable condition. Such behavior at the level of a single brain area has an impact on the global network by introducing attractors without any trivial relation with the structural connectivity (figure 3.11); these observations, made in Hansen et al., 2015 [19], are deeply treated in section 3.3.2.
The values of the models parameters used to implement simulations are summarized in table 4.1.

4.3 The simulations

All the simulations are performed in Matlab. According with the works discussed in chapter 3, to simulate the resting state activity it is necessary that the system is in an unstable

Parameters		Values	
Name	Abbr.	MFM	eMFM
Kinetic parameter	γ	0.641	0.641
NMDA decay time constant	τ_s	0.1 s	0.1 s
Parameters of the input-output function	a	270 $(nC)^{-1}$	270 $(nC)^{-1}$
	b	108 Hz	108 Hz
	d	0.154 s	0.154 s
Local excitatory recurrence strength	w	0.9	1.0
Synaptic coupling	J_N	0.2609 nA	0.2609 nA
External input	I_0	0.3 nA	0.3 nA
Coupling strength	G	variable	variable
Noise amplitude	σ	variable	variable

Table 4.1: In the table the values of the parameters used to implement the brain models, MFM and eMFM, are summarized. These values are taken from Deco et al., 2013 [13], p. 11242, except for the value of w in the eMFM model that is fixed in accord with the work of Hansen et al., 2015 [19]

regime. Thus, to identify all possible stationary states, 200 deterministic simulations were run, each lasting 15 s, with different random initial conditions (100 simulations with random initial conditions in the range $[0; 0.2]$ in order to initialize the system in the low firing rate regime; 100 simulations with random initial conditions in the range $[0.8; 1.0]$ in order to initialize the system in the high firing rate regime). This procedure was repeated varying the value of the coupling strength G of the models (equation 4.3) in the range $[0; 0.1]$ with an increment of $\Delta = 2 \cdot 10^{-3}$. The deterministic Euler integration was used with a fixed integration step of 0.1 ms.
Once found the optimal values for the coupling strength, stochastic simulations, for both models, were run. The stochastic simulations last 20 min and were performed using the stochastic Euler integration with a fixed integration step of 0.1 ms. Simulated BOLD signal was obtained by converting the simulated neural activity using the Balloon-Windkessel method; details about this model are in appendix B. The BOLD time-series were downsampled to 1 s.

4.4 Functional connections

Functional connections in the simulated time-series are explored from a spatial point of view and a temporal point of view by using, respectively, the functional connectivity (FC)

and the functional connectivity dynamics (FCD).
The ij element of the FC matrix is calculated as the Pearson correlation between the BOLD signal of the brain region i and the BOLD signal of the brain region j.
To estimate the FCD, the entire BOLD time-series was divided in time windows of a fixed length (2 min) and with an overlap of 118 s; the data points within each window, centered at the time t_i, were used to calculate FC(t_i).

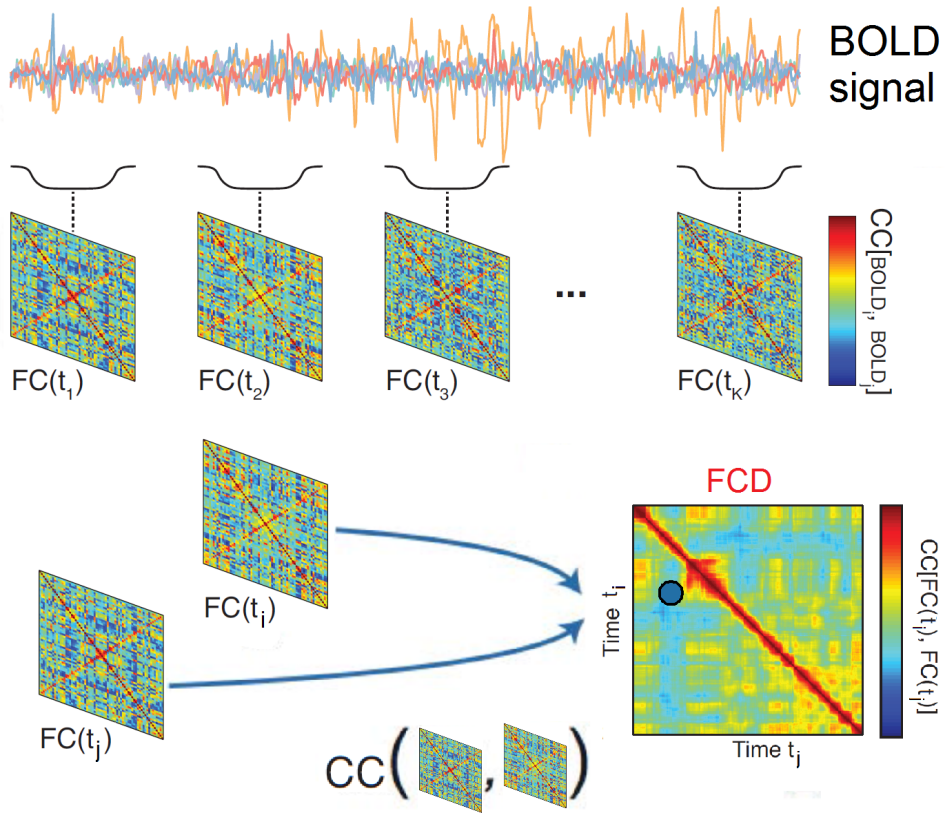


Figure 4.3: The image provides an explanatory cartoon of the FCD matrix construction. After the division of BOLD signals in time windows, the FC(t_i) of each time window is computed. The correlation values between the triangular part of the FC(t_i)s, arranged in a vector, are calculated and organized in the FCD matrix.
Source: adapted from Hansen et al.,2015 [19] p. 528

The ij element of the FCD matrix is calculated as the Pearson correlation between the upper triangular part of the $FC(t_i)$ matrix arranged as a vector and the upper triangular part of the $FC(t_j)$ matrix arranged as a vector.
The FCD matrix allows us to identify the epochs of stable FC configurations as blocks

of elevated inter-$FC(t)$ correlation. These blocks are organized around the FCD matrix diagonal [19]. For each epoch the corresponding FC matrix was calculated as the mean value of the $FC(t_i)$, with t_i belonging to that epoch.

4.5 The resting state functional cores

The simulated data are compared with the empirical data obtained in the paper of Mechling et al., 2014 [25]. Performing resting state fMRI experiments on mice and using graph theory analysis, the authors have individuated 5 functional modules; furthermore the hub regions for each modules are identified. In what follows the methods and the results of the article are summarized. In the last part of the section the approach used to individuate functional hubs in our simulated data is explained.

4.5.1 The experimental functional hubs in literature

The article "Fine-grained mapping of mouse brain functional connectivity with resting state fMRI" by Mechling et al., 2014 carefully describes the more functionally active regions during the resting state in the mouse brain.
The fMRI experiments were executed on 13 female mice anesthetized with isoflurane. The imaging was performed with a 7 T animal scanner; 12 slices of the mouse brain were acquired using single shot gradient Echo EPI. The acquisition protocol (lasting 15 minutes) was run twice during the same session to check for stability and reproducibility of the connectivity patterns.
The Independent Component Analysis (ICA) was performed on the BOLD signal in order to divide it into additive components generated by several sources, i.e. brain regions; with this approach the brain is parceled in 92 areas. In figure 4.4 is shown the adjacency matrix, averaged across all animals, obtained by calculating the partial correlation coefficient between each pair of region. Graph theory method was used to divide in cluster the functional network defined by the weighted PC matrix. For each division in M modules the modularity Q was calculated as [30]:

$$Q = \sum_{u \in M} \left[e_{uu} - \left(\sum_{v \in M} e_{uv} \right)^2 \right] \tag{4.4}$$

where e_{uv} is the proportion of all links that connect nodes in module u with nodes in module v. The best partition was individuated as the one for which the modularity Q is maximized. The best modular structure is composed of $M = 5$ functional modules as shown in figure 4.5.

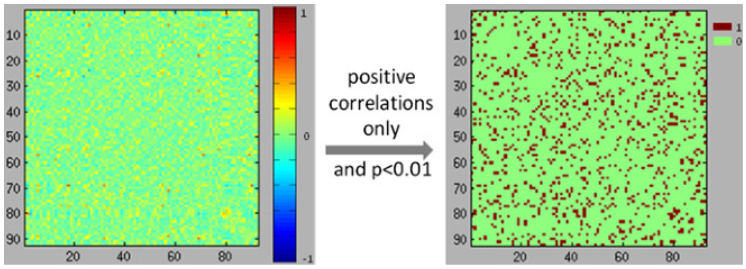


Figure 4.4: The matrices summarize the fMRI resting state data obtained in the study of Mechling et al., 2014. The left matrix is the weighted indirect partial correlation matrix obtained by correlating the BOLD signals (15 min.) of the 96 brain areas found with the ICA. The right matrix is obtained from the left matrix replacing all the positive entries by 1, and all the negative entries by 0.
Source: Mechling et al.,2014 [25] p. 208

Furthermore, in the paper, for each node of the functional network the normalized con-

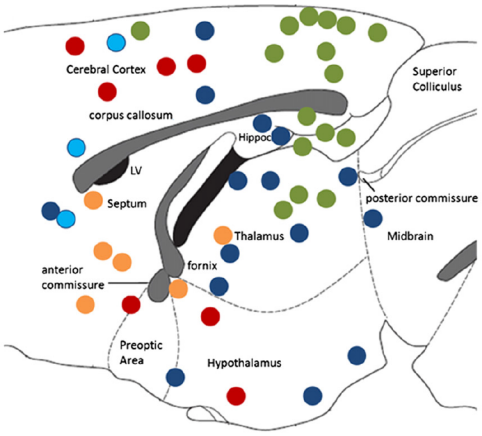


Figure 4.5: It is schematically shown, in the sagittal plane of the mouse brain, the modular organization of the resting state FC obtained as the division which maximizes the modularity. Different colors label different modules.
Source: Mechling et al.,2014 [25] p. 210

nection strength and the regional connection diversity were calculated; these quantities estimate the centrality of a node in a network. The measures will be briefly defined according to the paper of Rubinov and Sporns 2011 [30]. Let us consider a network of n nodes where the weight of the connection between node i and node j is denoted by $w_{ij}^{\pm}$ (in the case of positive connection $w_{ij}^{+} =]0, 1]$ and $w_{ij}^{-} = 0$, in the case of negative

connection $w_{ij}^{+} = 0$ and $w_{ij}^{-} =]0, 1]$) and where there is a modularity partition M which divides the network in m modules; in this case the normalized connection strength $s_i^{'\pm}$ and the regional connection diversity $h_i^{\pm}$ of the i-th node are respectively defined as:

$$s_i^{'\pm} = \frac{1}{n-1} s_i^{\pm} \ , \qquad \text{with} \qquad s_i^{\pm} = \sum_j w_{ij}^{\pm} \tag{4.5}$$

$$h_i^{\pm} = -\frac{1}{log\, m} \sum_{u \in M} p_i^{\pm}(u) log\, p_i^{\pm}(u) \ , \qquad \text{with} \qquad p_i^{\pm}(u) = \frac{s_i^{\pm}(u)}{s_i^{\pm}} \tag{4.6}$$

where $s_i^{\pm}(u)$ is the total weight of connections of node i to all nodes in the module u. As we can notice in equation 4.6, the topological diversity of a node in a network is quantified as the Shannon entropy of the strength of node i within module u. The strength and the diversity are calculated in the paper of Mechling et al., 2014 to distinguish the importance of each node in the functional network; in figure 4.6 is shown the ranking of the mouse brain node according to the highest strength and diversity.

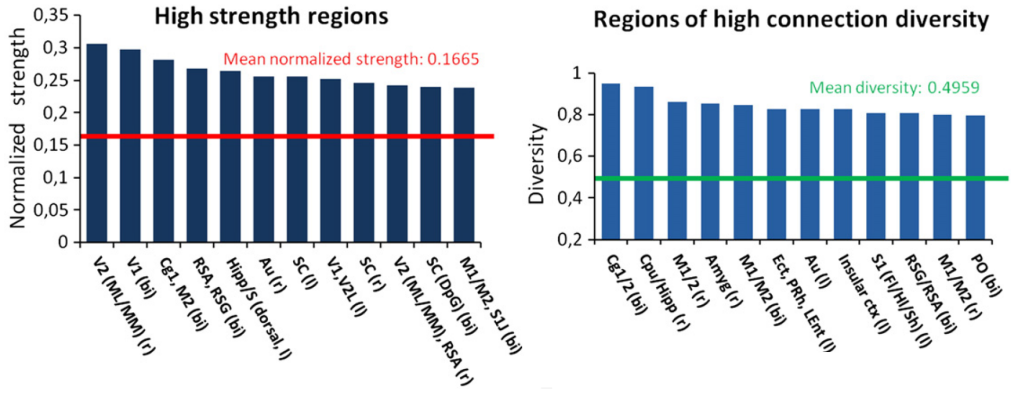


Figure 4.6: The right and the left graph represent respectively the ranking of the mouse brain area with the higher normalized strength and with the higher diversity. The authors define as functional hubs the regions that have both the value of the strength and the value of the diversity over the mean. The brain regions can be identified following the abbreviation list.
Source: Mechling et al.,2014 [25] p. 211

The authors define functional hub regions as the network nodes that have simultaneously strength and diversity over the average; at each module is associated one or more hub regions; these results are summarized in table 4.2.

Module	Hub regions
1. *Motor system*	Primary motor cortex Secondary motor cortex
2. *Somatosensory system*	Primary somatosensory area
3. *Limbic system*	Retrosplenial granular cortex Retrosplenial agranular cortex
4. *Visual system*	Visual cortex Dorsal lateral geniculate nuclei Pretectum
5. *Autonomic system*	Hypothalamus

Table 4.2: The table summarizes the results obtained in the paper of Mechling et al., 2014 [25]. In the first column there are the 5 modules in which the empirical resting state FC matrix results partitioned according to the maximization algorithm of the modularity (equation 4.4). The corresponding hub regions are individuated as the network nodes that have the strength value and the diversity value greater than the mean value.

4.5.2 The simulated hubs

The functional connectivity matrix of each epoch defines a functional network; for each functional network we identify the hub regions with an approach analogous to the one used in graph theory for defining the eigenvector centrality of a network node [27]. Let us assume the functional centrality $\phi^{(i)}$ of a brain region i as the sum of the functional centrality of the other brain regions weighted to the functional connection strength fc_{ij}:

$$\phi^{(i)} = \frac{1}{\lambda} \sum_{j=1}^{N} fc_{ij} \phi^{(j)} \tag{4.7}$$

where λ is a constant. Defining the vector $|\phi\rangle$ as the column vector whose components are the functional centrality of each network region, we can rewrite the previous equation in the matricial form:

$$|\phi\rangle = \frac{1}{\lambda} FC\,|\phi\rangle \quad . \tag{4.8}$$

It is simple to notice that $|\phi\rangle$ is the eigenvector of the functional connectivity matrix associated with the eigenvalue λ. Since FC is a real symmetric matrix (thus diagonalizable), we can decompose it as:

$$FC = \Phi \Lambda \Phi^T = \begin{bmatrix} |\phi_1\rangle & \cdots & |\phi_N\rangle \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \begin{bmatrix} \langle\phi_1| \\ \vdots \\ \langle\phi_N| \end{bmatrix} = \sum_{i=1}^{N} \lambda_i |\phi_i\rangle \langle\phi_i| \quad . \tag{4.9}$$

It follows that the magnitude of the eigenvalue gives a measure of the role of the corresponding eigenvector in reproducing the original matrix.
Taking into account all these observations, we identify the functional hub regions of the mouse brain as the regions with the largest eigenvector components associated with the three largest eigenvalues of the FC matrix.

4.6 Dimensionality reduction

To better understand the nature of the simulated time-series (the entire one, and the epoch's one), it is interesting to study their main components through a decomposition. An useful approach to reach this goal is the principal component analysis (PCA).
Let us consider the eigenvector $|v\rangle$ of the covariance matrix of the signal $|\psi(t)\rangle$, where $|\psi(t)\rangle$ is the original time-series once subtracted its mean value. The set of N eigenvectors, with N number of the total brain areas, forms an orthonormal base. Thus, since $\sum_{i=1}^{N} |v_i\rangle \langle v_i|$ is equal to the identity matrix, the time-series with null mean value, $|\psi(t)\rangle$, can be written as:

$$|\psi(t)\rangle = \begin{bmatrix} \psi_1(t) \\ \psi_2(t) \\ \vdots \\ \psi_N(t) \end{bmatrix} = \sum_{i=1}^{N} |v_i\rangle \langle v_i|\psi(t)\rangle = \underbrace{\sum_{i=1}^{M<N} |v_i\rangle \langle v_i|\psi(t)\rangle}_{|T(t)\rangle} + \underbrace{\sum_{i=M+1}^{N} |v_i\rangle \langle v_i|\psi(t)\rangle}_{|R(t)\rangle} \quad (4.10)$$

where the residual term $|R(t)\rangle$ is the part of the data that does not lie in the space spanned by the $M < N$ vectors $|v\rangle$. If we estimate $|\psi(t)\rangle$ as the projection on the first M vectors $|v\rangle$, the dimensionality of the problem is reduced, and consequently we lose information; it is possible to express the goodness of fit g between the lower dimensional data and the original one as explained in Banerjee et al., 2008 [3]:

$$g = 1 - \frac{E[\langle R(t)|R(t)\rangle]}{E[\langle \psi(t)|\psi(t)\rangle]} \quad (4.11)$$

where $E[x]$ indicates the expectation value of x.
In what follows we show that, when the time-series has a null mean value, the choice of $\{|v\rangle\}$ as the first M eigenvectors of the covariance matrix leads to have a goodness of fit equal to the sum of the corresponding M eigenvalues.

$$
\begin{aligned}
g &= 1 - \frac{E[\langle R(t)|R(t)\rangle]}{E[\langle \psi(t)|\psi(t)\rangle]} = \\
&= \frac{E[\langle \psi(t)|\psi(t)\rangle] - E[(\langle \psi(t)| - \langle T(t)|)(|\psi(t)\rangle - |T(t)\rangle)]}{E[\langle \psi(t)|\psi(t)\rangle]} = \\
&= \frac{E[2\,\langle \psi(t)|T(t)\rangle - \langle T(t)|T(t)\rangle]}{E[\langle \psi(t)|\psi(t)\rangle]} = \\
&= \frac{E[2\,\langle \psi(t)|\sum_{i=1}^{M}|v_i\rangle\,\langle v_i|\psi(t)\rangle - \sum_{i=1}^{M}\sum_{j=1}^{M}\langle \psi(t)|v_i\rangle\,\langle v_i|v_j\rangle\,\langle v_j|\psi(t)\rangle]}{E[\langle \psi(t)|\psi(t)\rangle]} = \\
&= \frac{E[\sum_{i=1}^{M}\langle \psi(t)|v_i\rangle\,\langle v_i|\psi(t)\rangle]}{E[\langle \psi(t)|\psi(t)\rangle]} = \sum_{i=1}^{M}\langle v_i|\,\frac{E[|\psi(t)\rangle\,\langle \psi(t)|]}{E[\langle \psi(t)|\psi(t)\rangle]}\,|v_i\rangle = \sum_{i=1}^{M}\langle v_i|C|v_i\rangle \quad .
\end{aligned}
\tag{4.12}
$$

Let us consider the covariance matrix of the time-series:

$$
Cov = E[|\psi(t)\rangle\,\langle \psi(t)|] - E[|\psi(t)\rangle]E[\langle \psi(t)|] \quad . \tag{4.13}
$$

When the mean value of the time-series is zero, the expression of the covariance becomes:

$$
Cov_0 = E[|\psi(t)\rangle\,\langle \psi(t)|] \quad . \tag{4.14}
$$

This is the expression found for the C matrix in equation 4.12 except for the multiplied scalar factor $(E[\langle \psi(t)|\psi(t)\rangle])^{-1}$. It is simple to prove that a matrix, X, and the same matrix multiplied by a scalar, $X' = yX$, share the same eigenvectors; moreover the following relation between the eigenvalues holds: $\lambda_i(X) = y \cdot \lambda_i(X')$. Consequently, if we normalize both sets of eigenvalues between 0 and 1, there will be no difference in the eigenvalues ($\overline{\lambda_i}(X) = \overline{\lambda_i}(yX)$).
In conclusion we have:

$$
g = \sum_{i=1}^{M}\overline{\lambda_i}(Cov_0) \quad . \tag{4.15}
$$

Chapter 5

Results and discussion

In the chapter the results obtained with the materials and methods explained in the previous chapter are described and discussed.

5.1 The optimal range

Firstly, for both models (MFM and eMFM) the stationary states has been calculated in function of the control parameter G, i.e. the coupling strength (equation 4.3). This was obtained running 200 deterministic simulations for each value of $[0; 0.1]$ with an increment $\Delta G = 2 \cdot 10^{-3}$. Deterministic simulations were run for 15 s in order to achieve the full equilibrium: the state of the network at the end of the simulation is stationary. In particular we initialize the system in the low activity state and then in the high activity state using random initial conditions belonging, respectively, to the range $0 \leq S_0 \leq 0.2$, and to the range $0.8 \leq S_0 \leq 1$. The diagrams in figure 5.2 show the maximum value of the firing rate activity (H in equation 4.2) over all brain areas for all possible stationary states. The resulting diagrams reveal that, in both models, for small values of the coupling strength ($G \leq 0.04$ in the MFM, and $G \leq 0.024$ in the eMFM) the firing rate in all brain areas reaches values below 50 Hz independently of the state in which the system is initialized; this means that the system always converges to a stable state characterized by a low firing rate activity in all brain areas. In the low range of the coupling strength G it exists only one stable state. For a critical value of the coupling strength G_c the diagrams show that the maximum value of the firing rate depends on the initial condition used to simulate the model; this means that a bifurcation occurs at G_c and, for $G > G_c$ at least two stable states exist one characterized by low and the other one by high firing rate

activity. The graphs show that the bifurcation occurs for values of G greater in MFM than in eMFM. Moreover, the MFM graph displays that the state of low activity, for values of G greater than 0.06 becomes unstable since the system converges to the high activity state also when initialized to low values. A similar behavior happens in eMFM for values of the coupling strength $G > 0.05$ when the state of low activity becomes unstable but, however, systems with different initial conditions converge to different stable states. These observations are consistent with the expectations: the dynamical repertoire of eMFM is richer than that of MFM. In the paper of Deco et al., 2013 [13], discussed in section 3.2.3, the authors obtain, with the same method, the bifurcation diagram using the MFM and simulating the dynamics with an human anatomical connectivity achieved with DSI technique. Their result, shown in figure 5.1, is consistent with the one of this thesis.
According to the works described in chapter 3, the resting state activity is simulated in the range where the dynamics is multistable, i.e. the models are implemented with values of the coupling strength belonging to the range $]0.04; 0.06]$ for the MFM and to the range]0.0240.058] for the eMFM (since a firing rate greater than 250 Hz is not biologically possible).

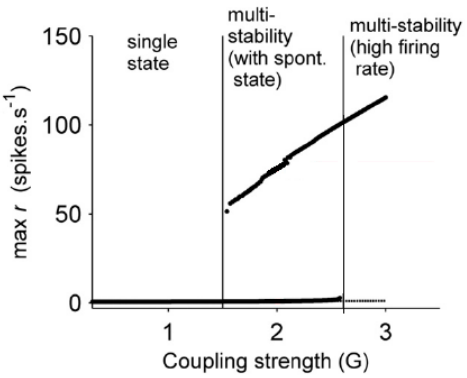


Figure 5.1: Bifurcation diagram obtained by Deco et al., 2013 [13] for the MFM using human SC. Each point represents the maximum firing rate (H in equation 4.2) among all nodes when the system is in the steady state. The graph is obtained running 1000 deterministic simulations, lasting 15 s, for each value of G. This graph is analogous to the one obtained in this work (figure 5.2, upper part).
Source: Deco et al., 2013 [13], p. 11245

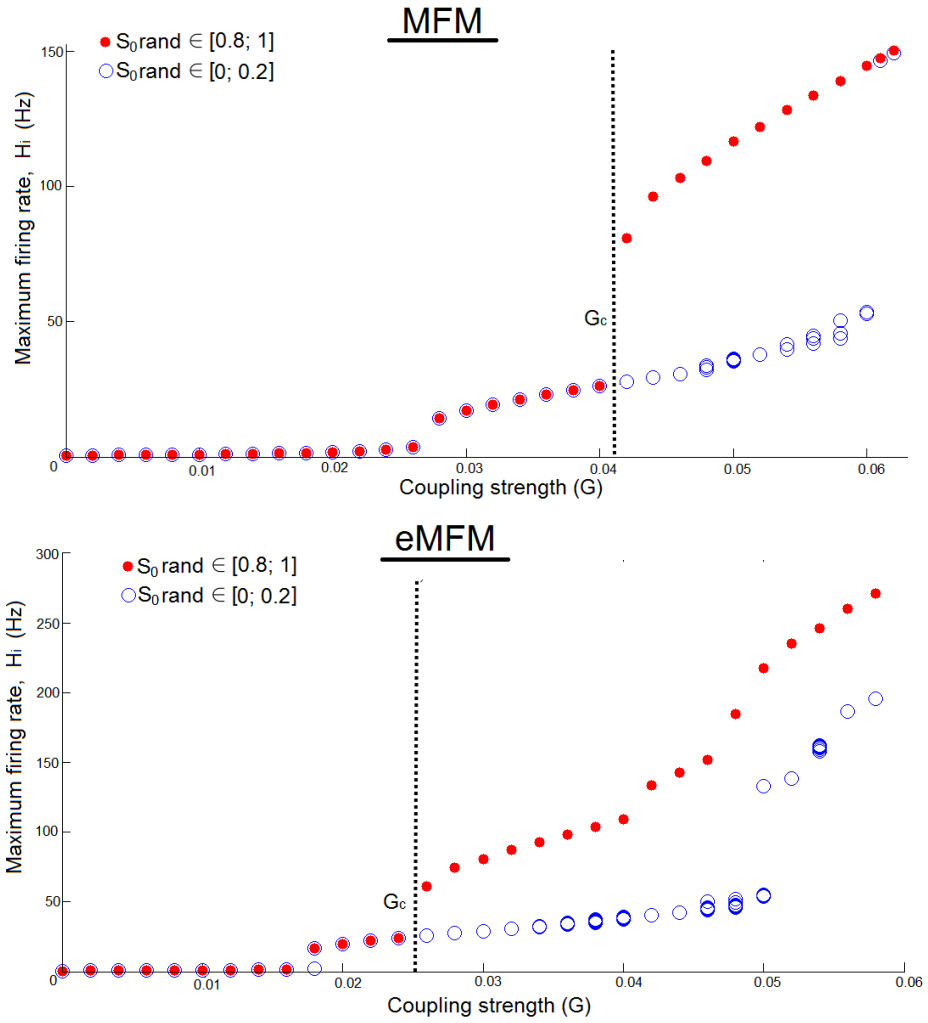


Figure 5.2: Each point represents the maximum firing activity rate (H in equation 4.2) among all nodes when the system is in the steady state. The equilibrium is achieved running 200 deterministic simulations, for each value of G. Each simulation lasts 15 s and it is implemented with random initial conditions belonging to the low range (blue dots) and to the high range (red dots). The model used in the top graph is MFM, while the model used in the bottom graph is eMFM. For low values of the coupling strength ($G < G_c$) the system converges to a single stable state of low activity. Increasing the value of the coupling strength ($G > G_c$) in addition to the low activity state the high activity state appears.

5.2 The functional connectivity dynamics

The BOLD signals, for both models, are obtained by running stochastic simulations and implementing the brain system in different configurations; in particular the network behavior is studied for different values of the coupling strength G (belonging to the optimal range individuated in the previous section) and for different noise levels (changing σ in equation 4.1). Then, for each situation, the FCD is calculated as explained in section 4.4. Between all the simulations performed, the most representatives are selected and are shown in figure 5.3; this will be discussed in what follows. Specifically we have chosen the BOLD signals and the related FCD obtained using $G = 0.058$ in the MFM and $G = 0.053$ in the eMFM; in both cases the levels of noise used to perturb the system are: $\sigma = 10^{-3}$, $\sigma = 3 \cdot 10^{-3}$ and $\sigma = 4 \cdot 10^{-3}$.
For the weakest level of noise, the network nodes, simulated with the MFM, display only high or low level of neural activity. There is no switching between the two states, i.e. each node has a constant level of neural activity for all duration of the simulated signal, and, consequently, its BOLD signal displays a constant amplitude, in spite of noise perturbations; this can be noticed in the time-series plotted in figure 5.3. Such behavior is reflected in the FCD matrix which does not contain any clue of FC switching. Simulating the BOLD signals with the same level of noise (the weakest one selected) and the eMFM, it can be noticed that some nodes jump from one state to the other one after $t = 15$ min. However this switching behavior is not enough intensive to be considered a clue of the system multistable dynamics, as also indicated by the absence of any significant FC(t_i) correlations in the related FCD matrix.
Thus we can deduce that in both models in the presence of low levels of noise the dynamics of the network settles into a fixed point, as expected.
Increasing the level of noise in the MFM means only pushing the system into a chaotic regime as in the time-series obtained with $\sigma = 2{\cdot}10^{-3}$ and $\sigma = 3{\cdot}10^{-3}$. In such graphs the BOLD signals of certain brain areas display oscillations between high and low amplitude values, i.e. the network nodes oscillate between high and low activity state. Such network behavior does not mean that the system is exploring different attractors, but it only reflects the effect of the high noise perturbation. The absence of any constructed pattern is indicated also in the FCD matrices which does not contain significant correlations.
On the other hand, increasing the noise level in the eMFM simulations means that the system is allowed to explore several stable states. As we can notice from the BOLD signals obtained in these cases, the nodes jump between high activity and low activity state not in a chaotic way, as it happens in the MFM, but in a structured way: the trajectories of the

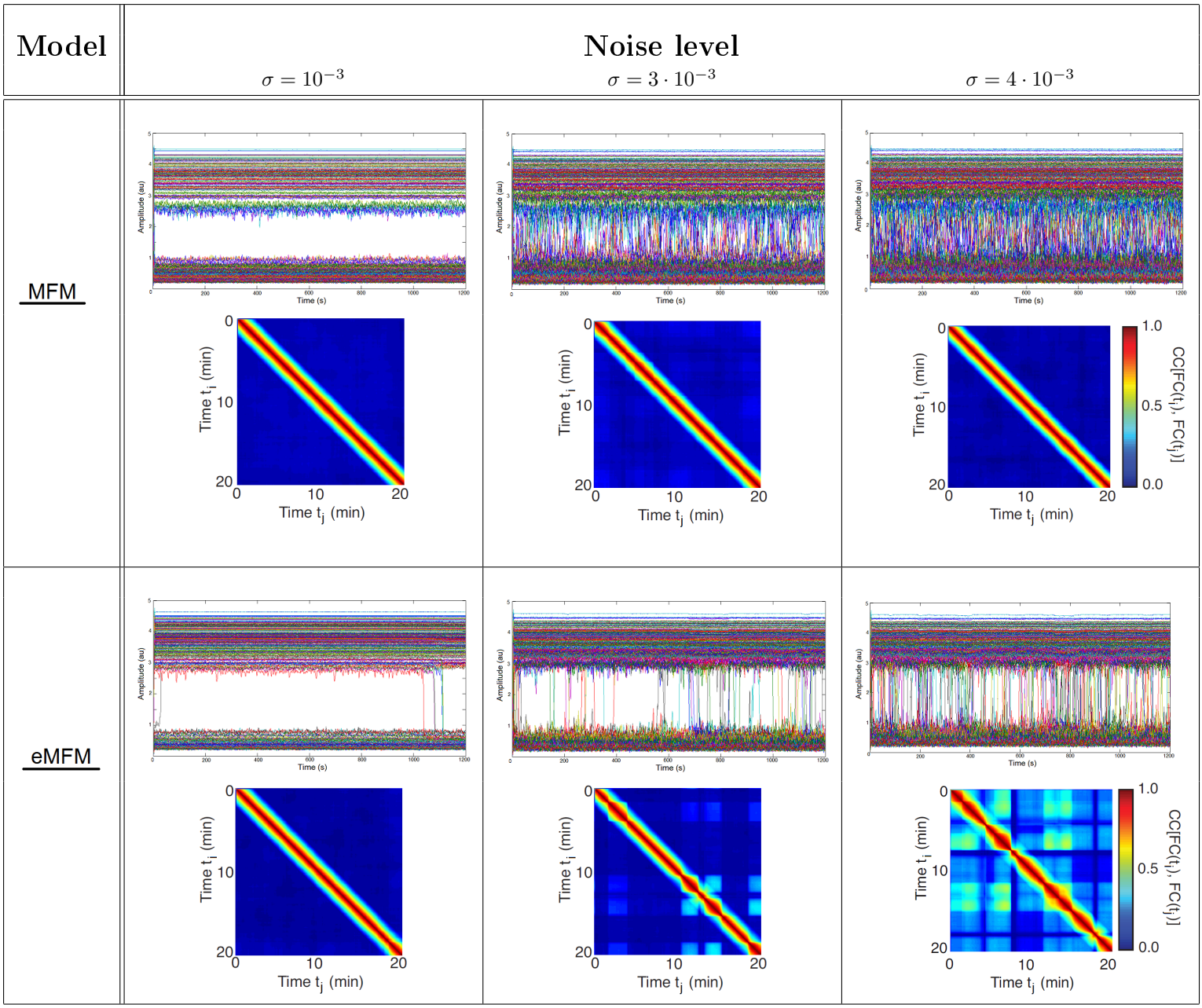


Figure 5.3: 20 minutes of simulated BOLD signals and relative FCD matrices obtained implementing the MFM with $G = 0.058$ (first row) and the eMFM with $G = 0.053$ (second row), and perturbing the system with different noise levels (columns, weaker to stronger noise from left to right).

dynamics of the system explore different subspaces. The trajectories either converge back towards the fixed point from which they originated or diverge towards a different attractor. The dynamics behavior of the system give rise to structured dynamics itinerancy as can be seen by the emergence of stability pattern in the FCD matrices obtained in these noise regimes.

We can conclude that the MFM is not able to reproduce the switching between the functional connectivity states observed in the empirical BOLD data. Introducing bistability at the level of single brain area, i.e. implementing the eMFM, it is possible to observe the emergence of functional connectivity patterns when the system is maintained out-of-equilibrium by appropriate noise levels.

The results are consistent with the work of Hansen et al., 2015 [19].

Between all the simulations performed, the analysis as explained in the method chapter (individuation of epoch of stability, identification of functional hubs and decomposition), is carried out on one of them that has the following characteristics:

- ▷ model: eMFM,
- ▷ coupling strength: $G = 0.052$,
- ▷ noise level: $\sigma = 3.6 \cdot 10^{-3}$.

Unless otherwise specified, from now on, the general terms time-series, BOLD signals, FCD matrix will refer to the result of *this* simulation. The BOLD signals and the FCD are respectively shown in figure 5.4 A and 5.4 B. The shape of the time-series inform us that the the brain network is exploring a rich dynamical repertoire. Thus, the functional patterns evolve in time as shown by the FCD matrix. One can notice that there are blocks around the diagonal of the FCD matrix that correspond to time intervals during which the FC(t_i) matrices are strongly correlated ($\gg 0.5$ Pearson correlation); we call these temporal lengths *epochs of stability* since they correspond to periods of functional stability patterns. In the FCD matrix we individuate 3 epochs of stability indicated with black lines in figure 5.4 B. Figure 5.4 C displays the FC matrices calculated over the identified epochs of stability as explained in the chapter of methods.

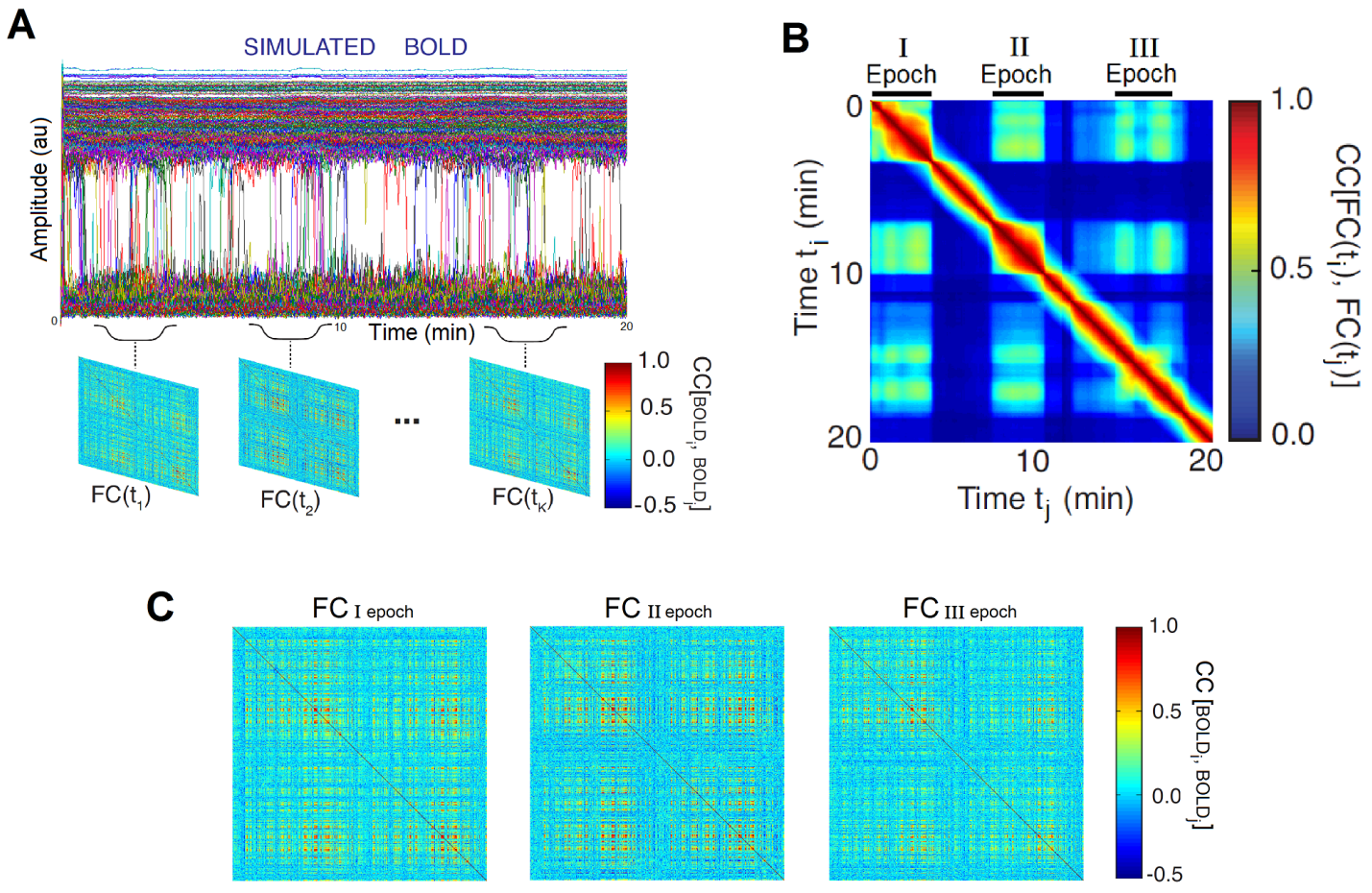


Figure 5.4: A. The simulated BOLD time-series of the resting state activity. The stochastic simulation was performed implementing the eMFM with the coupling strength G equal to 0.052 and the noise level σ equal to $3.6 \cdot 10^{-3}$; the values of the other parameters of the model are in table 4.1.

B. The FCD matrix obtained from the concerned time-series. Black lines label the individuated epochs of stability of the functional connectivity patterns.

C. The FC matrices corresponding to the epochs of stability.

5.3 The functional hubs

Before starting next discussion it is useful to provide the reader with a tool to facilitate the orientation in the mouse brain; some of the mouse brain areas visualized in the 2D coronal view in figure 5.5 can be easily associated with their names using the table under the figure. Such view will be used to plot the components of the eigenvector of the functional and structural connectivity matrices.

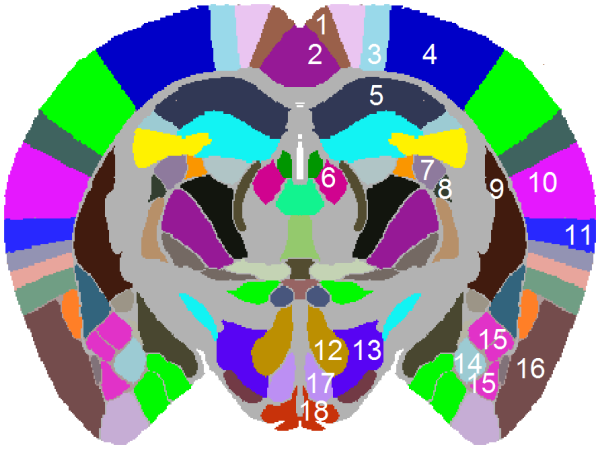


Figure 5.5: This figure supplies an usefull guide for the nomenclature of the coronal slice of the mouse brain. Different colors indicate different brain areas according to the parcellation of the SC matrix used here (the gray color labels regions which are not contained in the SC matrix). The table below contains the name of the brain areas labeled with the numbers in figure.

1	Restrosplenial area, dorsal part	10	Primary auditory areas
2	Retrosplenial area, ventral part	11	Temporal association areas
3	Primary motor area	12	Posterior hypothalamic nucleus
4	Primary somatosensory area, trunk	13	Lateral hypothalamic area
5	Field CA1	14	Basomedial amygdalar nucleus
6	Lateral habenula	15	Basolateral amygdalar nucleus
7	Dorsal part of the lateral geniculate complex	16	Piriform area
8	Ventral part of the lateral geniculate complex	17	Dorsomedial nucleus of the hypothalamus
9	Caudoputamen	18	Bed nuclei of the stria terminalis

Firstly the functional networks are defined as the networks whose adjacency matrix is the FC matrix computed on the identified epochs (see previous section). Then the hub regions of each functional network are individuated; as already discussed in the previous chapter, the functional hubs of the ith epoch are defined as the brain regions linked to the largest eigenvector components associated with the first 3 biggest eigenvalue of the FC matrix computed over the ith epoch. Such eigenvector components are plotted over the coronal view of the mouse brain as shown in figure 5.6.

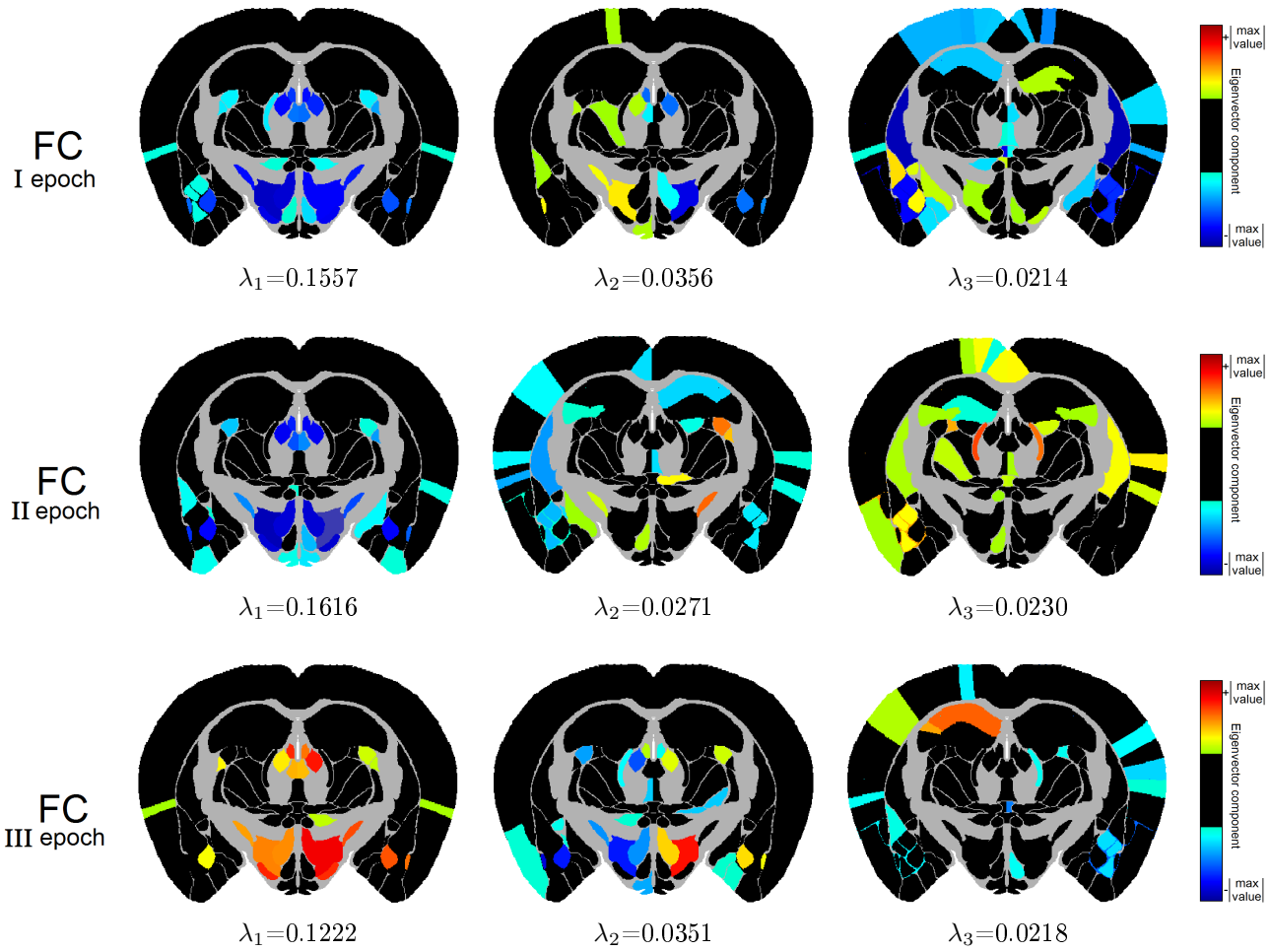


Figure 5.6: Eigenvectors components mapping in the 2D coronal view of the mouse brain. Images in the same rows represent the plotting of the eigenvectors components of FC belonging to the same epoch. The images organized in the first, second and third column are referred respectively to the first the second and the third biggest eigenvalue of the matrices (the normalized magnitude of the eigenvalues is under each figure). The scale used changes for each eigenvector: the scale is symmetrical stretched until the biggest eigenvector component, in absolute value; such scale permits to efficiently visualize the relative difference between the eigenvectors. According to our definition (see method section), the colored areas are the hub regions of the brain network defined by the FC matrices in left part of the figure; the importance of each hub regions is proportional to the corresponding eigenvalue.

We decide to use a color map that associates black color with all the values close to zero, so that only the eigenvector components with the highest absolute value are highlighted. The scale used change for each eigenvector: the scale is symmetrically stretched until the biggest eigenvector component, in absolute value; this choice is motivated by the fact that we are interested in visualizing the relative difference between the eigenvectors since any multiple of an eigenvector is also an eigenvector with the same eigenvalue.
The functional hubs obtained are compared with the anatomical connectivity used in the simulations, and with the functional hubs observed experimentally in literature and specifically with the ones obtained in the paper of Mechling et al., 2014 [25] summarized in section 4.5.1.

5.3.1 Comparison with the structural hubs

We define the structural hubs with the same methods used to define the functional hubs. The eigenvectors corresponding to the first three highest eigenvalues of the structural connectivity matrix are plotted in figure 5.7.

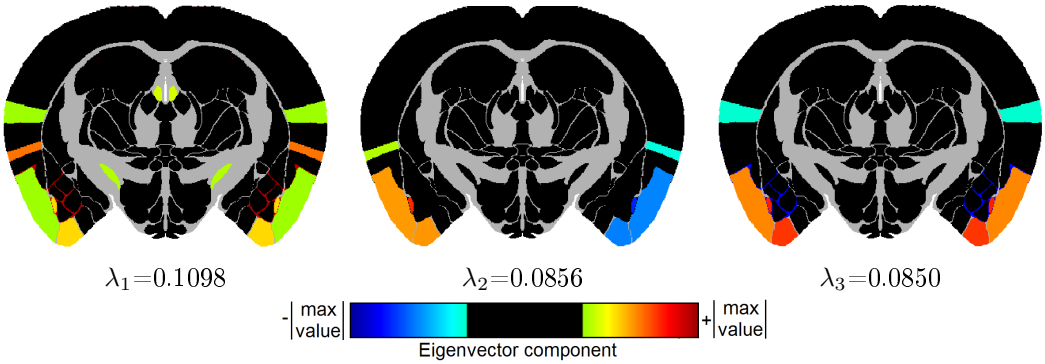


Figure 5.7: The figure represents the mapping, on the mouse brain coronal view, of the eigenvector components corresponding to the three highest eigenvalues of the structural connectivity matrix used to simulate the resting state activity. The normalized value of the eigenvalue is indicated under the plot of the relative eigenvector. According to our definition, the couloured areas are the structural hub regions; the importance of each hub regions is proportional to the corresponding eigenvalue.

The structural hubs and the functional hubs are clearly different. Thus, the hub regions we have detected in our simulation cannot be explained with the underlying anatomy

only, but they are the result of the interplay of the topology and of the dynamics of the brain system.

5.3.2 Comparison with the experimental functional hubs in literature

As discussed in section 4.5.1, in the work of Mechling et al., 2014 [25] the authors divided the functional connectivity matrix, obtained performing resting state fMRI acquisition on mice, in 5 functional modules. Using graph theory tools, one or more hub regions are associated at each module. The functional hubs obtained in the paper are plotted in the 2D coronal view of the mouse brain in figure 5.8. In the plot and in the following analysis we neglect the asymmetry in the functional hubs found in the work of Mechling et al., 2014 [25]; in order to be consistent we neglect also the asymmetry of the functional hubs found in the simulation.

The hypothalamus and the geniculate complex
Observing the results in figure 5.6 it is possible to notice that the hypothalamus, the ventral and the dorsal part of the geniculate complex are associated, in each epoch, with highest components of the eigenvectors of FC. Thus they are classified as functional hub regions of the functional networks arranged during the simulated resting state activity. This result is consistent with the work of Mechling et al., 2014 [25] where the authors identify the hypothalamus and the geniculate complex as hub regions of the autonomic system (violet area in figure 5.8) and of the visual system (yellow area in figure 5.8) respectively.
The retrosplenial cortex
The retrosplenial cortex (RSC), divided in granular and agranular part, is the the main hub of the human resting state network (together with the posterior cingulate cortex) [6], [29], [35]. The RSC is associated with cognitive functions including episodic memory, navigation, imagination of future events. These cognitive tasks are the ones that characterize the functional resting state network as discussed in the first chapter of this work. It was noticed that the RSC is a component of the resting state network also in the rat brains [24]. In the study of Mechling et al., 2014 [25] the authors individuate in the retrosplenial cortex the hub region of the limbic system (the blue area in figure 5.8). The results plotted in figure5.6 and achieved in this work show that the dorsal part of the RSC is associated with one of the highest eigenvector components of the third eigenvalue of the functional connectivity matrix, when calculated on the first epoch; in the second epoch the ventral part of the RSC is associated with a significant eigenvector component of the second eigenvalue of FC; the ventral and dorsal part of the RSC are associated also

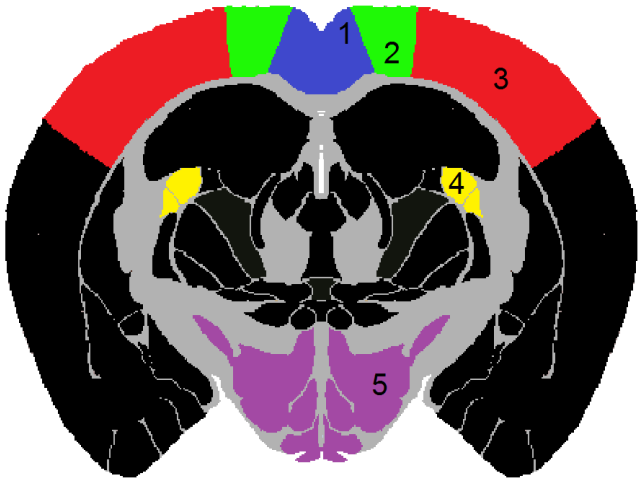


Figure 5.8: The different colors in the coronal view of the mouse brain label the functional hub regions associated with different modules according to the study of Mechling et al., 2014 [25] as explained in the table below. We have neglected the hemisphere asymmetry that is present in the paper's data.

Num	Module	Hub regions
1	*Limbic system*	Retrosplenial granular cortex Retrosplenial agranular cortex
2	*Motor system*	Primary motor cortex Secondary motor cortex
3	*Somatosensory system*	Primary somatosensory area
4	*Visual system*	Visual cortex Dorsal lateral geniculate nuclei Pretectum
5	*Autonomic system*	Hypothalamus

with important eigenvector components of the third eigenvalue of FC, when calculated over the second epoch. Thus in our results, the RSC is composed of core regions of the network defined by the functional connectivity of the first and of the second epoch of stability.

The amygdala

The other brain structure that in all the epochs plays an important functional role is the amygdala. In particular, as shown in figure 5.6, the basomedial and basolateral amygdala nucleus, during each epoch, are associated with high components of the eigenvector of all the considered eigenvalue. The amygdala is a functional core in our simulated resting state activity. Such area is composed of structurally and functionally distinct nuclei that contribute to the processing of emotion through interactions with other subcortical and cortical structures; it is an important part of the limbic system. The amygdala is a component of the human default mode network [6]. In the work of Lu et al., 2012 [24], using experimental data, the authors pointed out that all the structures, involved in the

rat resting state networks, have direct or indirect connections with limbic structures as hippocampus and amygdala. In the study of Mechling et al., 2014 [25] the amygdala is not classified as a functional hub region. However in the grading of brain regions that has high connection diversity (left graph in figure 4.6) the amygdala has an high ranking: the fourth. Since the hub regions in the paper are defined as the regions that have simultaneously high strength and diversity (respectively equation 4.5 and 4.6), the amygdala is not included in this definition because it has not an high strength. On the other hand when considering the binary FC matrix (right part of the figure 4.4) the amygdala shows a number of connections above the mean value: it is the sixth most functional connected brain region. All these observations underline the significant functional role of the amygdala in the experimental mouse data; the same importance can be found in our results.

5.4 The decomposition of the BOLD signals

The complete BOLD time-series and the ones corresponding to each epoch are decomposed as explained in the methods section. The decompositions of the complete (20 min) BOLD signals and of the BOLD signals of each epoch are shown in figure 5.9 and in figure 5.10 respectively. Each figure displays the original BOLD time-series, after subtracting its mean value, the projected time-series and the four components along the four eigenvectors associated with the four biggest eigenvalues of the covariance matrix of the BOLD signals.
The temporal evolution of the 4 principal components of the first and of the second epoch are very similar. Let us assume, in first approximation, that the principal components of the first and of the second epoch are equal. Following the approximation, the two epochs represent the length of time during which the network nodes have the same covariance matrix, up to multiplicative factors, and then the evolution of the network nodes is similar. We can say that a particular functional resting state network is arranged during the first epoch and then it is decomposed. The same network configuration emerges again during the second epoch.

As already mentioned in the chapter of method, the operation of projection onto the space spanned by the eigenvectors of the four largest eigenvalues implies a loss of information. The loss of information is less significant for the epochs than for the entire time-series as can be seen comparing the goodness of fit that is obtained for the epochs with respect to the entire time-series. Since the goodness of fit is closely related to the covariance matrix of the time-series itself (as shown in the methods section), this means that the normalized values of the first four eigenvalues of the covariance matrix, calculated over

the entire BOLD signals, are less significant than the ones of the covariance matrices calculated over the epochs.
In order to understand the meaning of this observation let us consider the limit case in which all the eigenvalues have the same magnitude and thus all the components $\langle v_i|\psi\rangle$ are equal. It comes out that the covariance matrix is the identity matrix, since it is real and symmetric [1]. A covariance matrix equal to the identity matrix is typical of a BOLD signals where all the dimensions (i.e. all brain areas) are statistically independent. The previous statements, once translated in our case, lead to the consideration that the network nodes in the entire BOLD signals behave more uncorrelated that the network nodes in the signal belonging to each epoch. This is a clue of the fact that each epoch defines length of time during which stable brain network configurations emerge; such functional networks are not evident when all the entire signal length is considered.

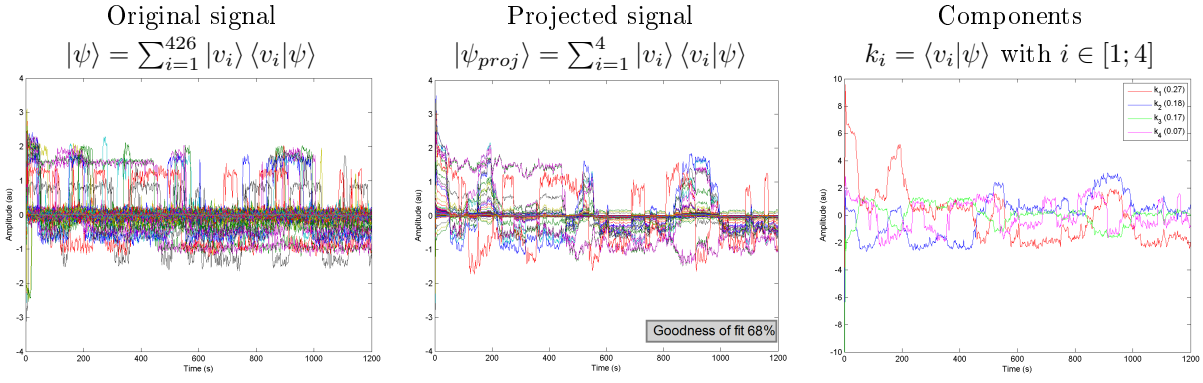


Figure 5.9: In the first column the simulated BOLD signals, after subtracting its mean value, is shown; in the second column the BOLD signals projected on the first 4 eigenvector of the covariance matrix, are illustrated, as discussed in the method section; in the last column the time evolutions of the 4 components k_i are displayed.

[1] Let us consider the matrix C that is real and symmetric (thus diagonalizable). If all its eigenvalues are equal to one, then there exists a matrix V such that:

$$C = VIV^{-1} = VV^{-1} = I \tag{5.1}$$

where I is the identity matrix.

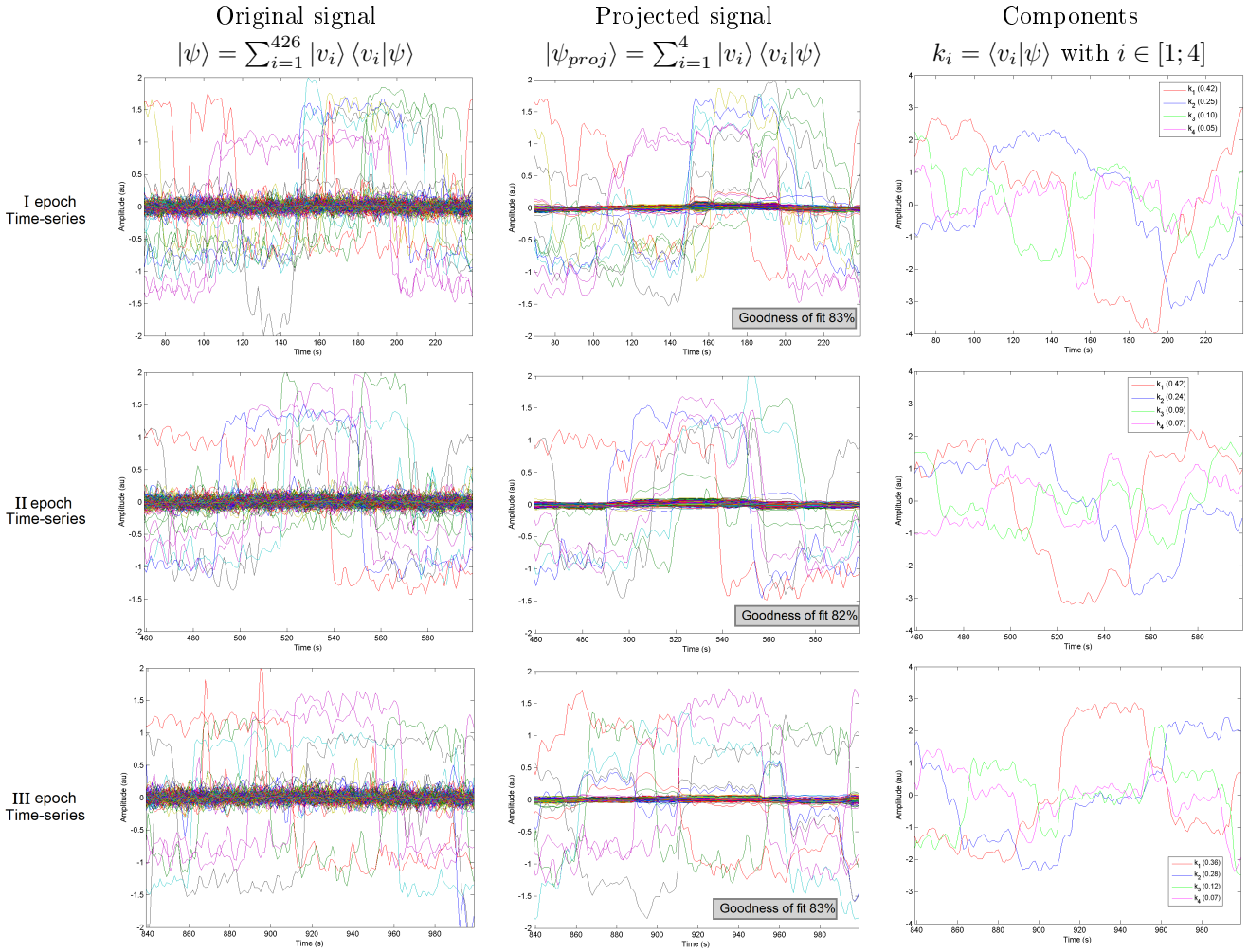


Figure 5.10: The graphs in the same row are related with the same epoch. In the first column the simulated BOLD signals, after subtracting their mean value, are shown; in the second column the BOLD signals projected on the first 4 eigenvector of the covariance matrix, are illustrated, as discussed in the method section; in the last column the time evolutions of the 4 components k_i are displayed.

Conclusion

The increasing amount of efforts in studying the spontaneous brain activity is motivated by the promise that such research carries with itself: help to shed light on large-scale brain dynamics. Models to predict this intrinsic activity were conceived and tested against resting-state data and they have added pieces to our understanding of the dynamics underlying brain functioning.

However, whatever is the dynamics rule, such activity is to a certain degree constrained by structure, i.e. the anatomical wirings of brain regions that define the topology of the network used in the model. The structural connectivity, used in typical resting state simulation, is obtained with DTI or DSI. These methods have several lacks such as low spatial resolution ($\sim 6mm$), ability to give information only about the connection's course and not on the connection's direction, ability to investigate only white matter and not brain structures. The structural matrix, used in the present work, is obtained with the anterograde tracing method, that is able to overcome all the previous limitations: it is applicable to any brain structure and it permits to label the fiber origin and destination with microscopic resolution. For these reasons, the connectome used in this work is, to date, the most comprehensive in any vertebrate species [28]. Such state-of-the-art tool, obtained in the Allen Institute for Brain Science, contains the connections weight between 426 brain regions of the mouse brain.

The simulation of the BOLD resting state activity in a mouse brain represents the second original aspect of this thesis. The non-invasive recording of resting state activity in the human brain has been tied to several neurological and psychiatric conditions, including, among others, Alzheimer's disease, autism spectrum disorders, and attention deficit hyperactivity disorder. However, key limitations exist in human-only studies, and the ability to use animal models greatly advances our understanding. Despite the large amount of human brain disorders that can be modeled in mice, little is known about the large-scale functional architecture measured by functional brain imaging: the mouse brain activity in resting state is explored experimentally in few articles and, as far as we know, no work has explored it from a theoretical point of view until the present one. Here we have used a

resting state model to explore how the large-scale functional connections emerge in mouse brain.

A measure often used to evaluate the goodness of resting state models, against empirical data from BOLD signals, is the linear Pearson correlation coefficient computed among each pair of brain areas over long time-series (of the order of 10 minutes); such correlations are usually organized in the so-called functional connectivity matrix (FC). Unfortunately this method does not give justice to the brain complexity since resting-state brain activity is not stationary. In fact the empirical study [2] showed that FC, when computed over smaller time windows, varies greatly during a scan session. As shown in the study of Hansen et al., 2015 [19], the previous large-scale computational models of the human brain, such as the mean-field model (MFM) [11], [13], did not reproduce the switching between functional connectivity states observed in empirical BOLD data. This is because the parameters of these models were chosen to fit empirical functional connectivity matrices constructed by using the entire time series of BOLD signals. Consequently, the studies focused only on those regions of the parameters space in which the dynamics of the MFM correlate closely with the empirical functional connectivity. Instead a slight enhancement in the non-linearity of the network nodes leads the system to enrich its repertoire of attractors and consequently FC, stochastically simulated, displays non-stationary switching similar to empirical resting state recordings. The model used in Hansen et al., 2015 [19], the so-called *enhanced mean field model* (eMFM), is the one used in this work that gave us the opportunity to reproduce the dynamical behavior of the functional connections in mouse brain; instead, simulating the dynamics with the MFM, no FC patterns emerge. In order to evaluate the evolution in time of the functional connections we calculated the functional connectivity dynamics (FCD) matrix [19]. Such quantity allow us to individuate the length of time, *epoch*, in which the functional connectivity is stable. The FC matrices calculated over these epochs are the so-called functional connectivity states since they define a functional configuration of the brain network that persists for several minutes.

Then we calculated the hub regions of such functional network as the brain regions associated with the greater, in absolute value, components of the eigenvectors corresponding to the three largest eigenvalues of the FC states. These functional cores are compared with the experimental ones in literature [25], [24], [35]. The comparison leads to positive results: our simulation was able to catch the central role of the hypothalamus, of the geniculate complex, of the retrosplenial cortex and of the amygdala as experimentally observed in mouse resting state networks. The goodness of the comparison gives us confidence in the method used to simulate the resting state activity and in the analysis performed on the BOLD signals. Finally the possibility of restricting the dynamics to a small sub-space is explored with the aim to reduce the degrees of freedom of the complex

network and to obtain a more tractable system that allows further insights on otherwise masked characteristic of the network. In particular, the complete time-series and the ones calculated over the epochs of stability, were projected on the eigenvectors associated with the 4 largest eigenvalues of the covariance matrix of the respective BOLD signals. The projection onto such four dimensional space is able to reproduce around 80% of the signal belonging to the epochs and around 68% of the entire signal (20 min). As shown in this work, the goodness of the signal projection is quantified by the normalized magnitude of the eigenvalues of the covariance matrix of the signal. This means that during the complete 20 minutes of BOLD signals the network nodes behave in a more uncorrelated way compared to what happens during the BOLD signals belonging to each epoch. Such consideration confirms our approach: the epochs are time intervals during which stable brain network configurations emerge.

The rightness of our results follows from their ability to reproduce the dynamics behavior of the functional connectivity and from their agreement with the empirical evidences.

Abbreviation list of brain areas

Abbreviation	Brain area
Amyg	Amygdala
Au	Auditory cortices
CB	Cerebellum
Cg	Cingulate cortex
CNU	Cerebral nuclei
CPu	Caudate putamen
CTX	Cerebral cortex
DpG	Deep gray layer of the superior colliculus
Ect	Ectorhinal cortex
Ent	Entorhinal cortex
HB	Hindbrain
Hipp	Hippocampus
Hth	Hypothalamus
HypoTh	Hypothalamus
M1	Primary motor cortex
M2	Secondary motor cortex
MB	Midbrain
MM	Mediomedial area
ML	mediolateral area
PRh	Perirhinal cortex
PO	Preoptic area
RSA	Retrosplenial agranular cortex
RSC	Retrosplenial cortex
RSG	Retrosplenial granular cortex
S	Subiculum
S1	Primary somatosensory cortex
S1BF	Primary somatosensory cortex barrel field

S1Fl/Hl	primary somatosensory cortex, front/hind legs
S2	Secondary somatosensory cortex
SC	Superior colliculus
Th	Thalamus
V1	Primary visual cortex
V2	Secondary visual cortex

Appendix A

Anatomical measures of brain connectivity

In brain modeling, the anatomy of brain wiring, organized in the so-called structural connectivity matrix (section 3.1), is considered as an established and permanent body of knowledge which supposedly provides a secure basis for building neural network models. Then the goodness of the brain topology used in the modeling is strictly bound to the methods used to measure the structural connectivity matrix.
The most popular method used, to date, for measuring the structural connectivity is the DTI; while in this work the method used to construct the matrix is the anterograde tracing. Both methods will be described and then compared following the work of Kötter [23] and the one of Alexander and Lobaugh [1] in *Handbook of Brain Connectivity*.

Diffusion Tensor Imaging

Magnetic resonance imaging (MRI) methods provide information on white matter tissue. White matter is composed of myelinated axons and glial cells. Myelin is an electrically insulating material that forms a layer, the myelin sheath, usually around the axon of a neuron in order to increase the conduction speed of the action potential. Either demyelination, myelin degradation, or poor myelin development will impede the efficiency of action potentials and affect neural connectivity. The glia ("brain glue") are non-neural cells and are the supporting cells of the nervous system. Imaging methods that can characterize the properties of this complex tissue matrix may be valuable for investigating the

influence of tissue substrates on neural connectivity.
The MRI methods provide indirect measurement of tissue architecture.
The Diffusion Tensor Imaging (DTI) tracks the diffusion of water molecules in white matter. The Einstein diffusion equation:

$$\langle \Delta r^2 \rangle = 2nD\Delta t \tag{A.1}$$

states that the mean-squared displacement $\langle \Delta r^2 \rangle$ from diffusion is proportional to the diffusivity D over the diffusion time Δt; the displacement is scaled by spatial dimensionality n ($n = 3$ in biological tissue).
In fibrous tissue, such as white matter tracts in the brain, water diffusion is less restricted in the direction parallel to the fiber orientation; on the other hand water diffusion is highly restricted in the direction perpendicular to the fibers. Diffusion imaging experiments, in order to characterize such diffusion anisotropy, use measurement of parallel ($D_{\parallel}$) and perpendicular ($D_{\perp}$) diffusion components.
The diffusion tensor is a model of water diffusion, which assumes that the diffusion is described by a 3D, multivariate normal distribution:

$$P(\Delta \overrightarrow{r}, \Delta t) = \frac{1}{\sqrt{(4\pi\Delta t)^3 |D|}} \; e^{\frac{-\Delta \overrightarrow{r}^T D^{-1} \Delta \overrightarrow{r}}{4\Delta t}} \tag{A.2}$$

where Δr is the displacement vector, Δt is the diffusion time, and D is the diffusion tensor:

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \; . \tag{A.3}$$

The eigenvalues (λ_1, λ_2, λ_3) and the corresponding eigenvectors ($\overrightarrow{\varepsilon}_1$, $\overrightarrow{\varepsilon}_2$, $\overrightarrow{\varepsilon}_3$) of the diffusion tensor D describe, respectively, the relative amplitudes of diffusion and the directions of the principle diffusion axes (figure A.1). In the central nervous system, water diffusion is typically anisotropic in white matter regions ($\lambda_1 \neq \lambda_2 \neq \lambda_3$) and isotropic in both gray matter and cerebrospinal fluid ($\lambda_1 = \lambda_2 = \lambda_3$). The direction of white matter trajectories is estimate by assuming that the largest diffusion eigenvector is parallel to the tract orientation.

The motion of water molecules in white matter may cause the signal intensity to decrease in MRI. The diffusion MR imaging acquisition, typically, is performed using two gradient pulses with equal magnitude and duration (figure A.2). The first gradient pulse dephases the magnetization across the sample (or voxel in imaging); and the second pulse rephases the magnetization. For stationary (non-diffusing) molecules, the phases induced by both gradient pulses will completely cancel, the magnetization will be maximally coherent, and there will be no signal attenuation from diffusion. In the case of coherent flow in the

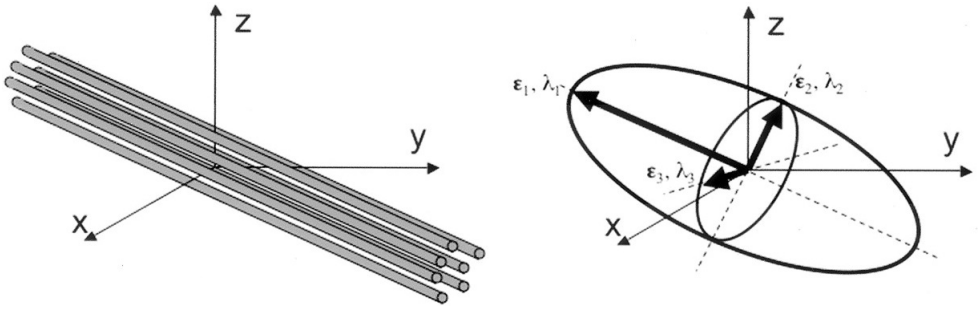


Figure A.1: Schematic representations of diffusion displacement distributions for the diffusion tensor. In the left part of the image a schematic representation of the white matter fibers (gray tubes) in the space is shown. In the right part of the image the common visual representation of the diffusion tensor is shown: an ellipsoid with the principal axes aligned with the eigenvectors and with the axes lengths proportional to the eigenvalues. The direction of greatest diffusivity is generally assumed to be parallel to the direction of white matter. When the eigenvalues are equal the diffusion is istropic, otherwise the diffusion is anisotropic.

direction of the applied gradient, the bulk motion will cause the signal phase to change by different amounts for each pulse so that there will be a net phase difference $\Delta\phi$, which is proportional to the velocity, v, to the gradient pulses amplitude, G, to the duration, δ, to the spacing between the pulses, Δ and to the gyromagnetic ratio γ:

$$\Delta\phi = \gamma v G \delta \Delta \ . \tag{A.4}$$

In the presence of diffusion gradients, each diffusing molecule will accumulate a different

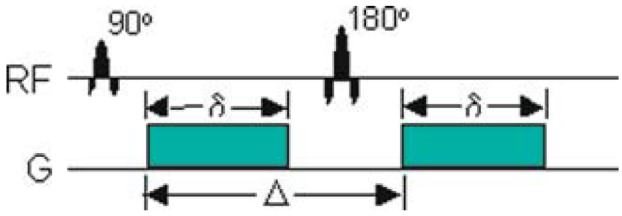


Figure A.2: Spin echo pulse sequence scheme for pulsed-gradient diffusion weighting.

amount of phase. The diffusion-weighted signal is created by summing the magnetization from all water molecules in a voxel. The phase dispersion from diffusion will cause destructive interference,and consequently signal attenuation. For simple isotropic Gaussian diffusion, the signal attenuation for the diffusion gradient pulses, figure A.2, is described by;

$$S = S_0 e^{-bD_{app}} \tag{A.5}$$

where S is the diffusion-weighted signal, S_0 is the signal without any diffusion weighting gradients, D is the apparent diffusion coefficient, and b is the diffusion weighting described by the properties of the pulse pair:

$$b = (\gamma G \delta)^2 (\Delta - \delta/3) \ . \tag{A.6}$$

Thanks to image processing, it is possible to calculate the diffusion tensor from the apparent diffusion coefficient.

Anterograde tracing

Anterograde tracing method allows to trace axonal projections from their source (soma) to their point of termination (synapse). The general approach comprises a vast range of substances with the common features that they are taken up by neurons and that they are spread along their projections, where they can be visualize. To reach such goal the tracer is a molecules with two main characteristics:

(*i*) the capability to be absorbed locally by the soma of neurons and transported to the axon terminals, or the capability to be absorbed by axons and transported to the soma of the neuron;

(*ii*) in order to be visualized by imaging technique the tracer has to be radioactive or fluorescent; for such reason the tracing one is an invasive methods.

Tract tracing	Diffusion-weighted imaging
microscopic resolution	low spatial resolution (several mm)
any species (limited by ethics)	applicable to large brains (humans)
fibers quantifiable	surrogate myelin measures
long history and known validity	unclear validity
invasive animal experiment	non-invasive, applicable to humans
few injections per brain	entire brain imaged
fiber origin + destination	shows fiber course, no direction
applicable to any brain structure	limited to white matter

Table A.1: Comparison between the characteristic features of tract tracing and the characteristic features of diffusion-weighted imaging.
Source: Kötter 2007 [23]

Appendix B

Modeling the BOLD signal

The BOLD signal is an indirect measure of the neural activity as discussed in section 1.2.2. The transformation from neural activity to measured BOLD signal is not trivial and, in the present study, is done following the mathematical approach described in the paper of Friston et al., 2000 [16].
The model used to describe the hemodynamic response detected with fMRI technique during brain activation, is based on the work of Buxton et al., 1998 [7] and its extension by Friston et al., 2000 [16]. As schematically shown in figure B.1, such model consists of three parts:

1. the link between the neural activity and regional cerebral blood flow (rCBF) (Friston et al., 2000 [16]);
2. the link between the rCBF and the changes in the blood volume (ν) and the link between the rCBF and the deoxyhemoglobin content (q) (the Balloon model by Buxton et al., 1998 [7]);
3. the link between ν and q and the BOLD signal change (Buxton et al., 1998 [7]).

For the sake of clarity, let us examine the three steps starting from the last one.

The Ballon model

Based on theoretical and empirical results, Buxton et al., 1998 [7] derived the following equation for BOLD signal change (ΔS) during activation, relative to the resting signal

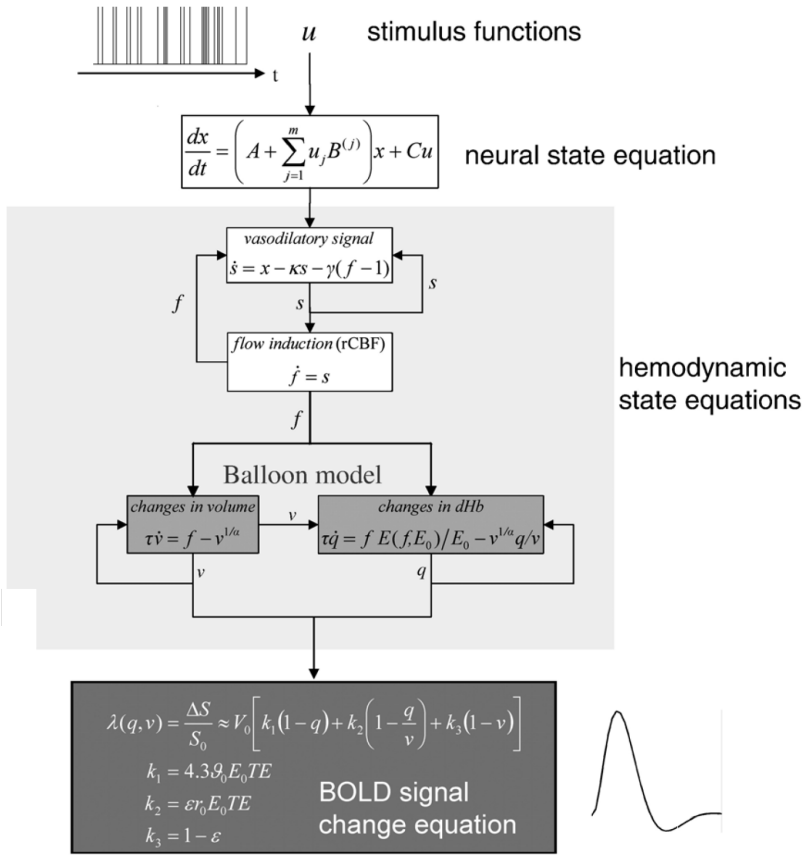


Figure B.1: Schematic representation of the hemodynamic model used in this work. *Source:* Stephan et al., 2007 [34] p. 388

S_0 (lower part of figure B.1):

$$\begin{aligned} \lambda(q,\nu) &= \frac{\Delta S}{S_0} \approx V_0 \left[k_1(1-q) + k_2(1-\frac{q}{\nu} + k_3(1-\nu) \right] \\ k_1 &= 7E_0 \\ k_2 &= 2 \\ k_3 &= 2E_0 - 0.2 \end{aligned} \tag{B.1}$$

where, as already mentioned, q and ν are respectively the venus blood volume and deoxyhemoglobin content (both normalized at their rest values). V_0 is the resting venous blood volume fraction and E_0 is the oxygen extraction fraction at rest. The coefficients k_1,k_2, and k_3 were estimated by assuming a magnetic field of 1.5 T and an Echo Time (TE) of 40 ms; these parameters are the ones used to implement the Matlab code in this

work.
The previous equation (equation B.1) requires the knowledge of the blood volume ν evolution and of the deoxyhemoglobin content q evolution; such request is satisfied by the Ballon model. The main idea of this model is that the increase in blood flow inflates the venous as a *balloon*, and consequently the deoxygenated blood is diluted and expelled at greater rate; the reduction of the deoxygenated blood causes an increase in NMR signal (see section 1.2.2). The description is based on two main assumptions: (*i*) the vessels react to an increase in inflowing blood like an inflating balloon, and (*ii*) oxygen extraction is strictly coupled to blood flow.
The first assumption (*i*) allows to describe the change in vessel volume ν as the difference between the incoming (f_{in}) and outcoming (f_{out}) blood flow:

$$\tau \frac{d\nu}{dt} = f_{in}(t) - f_{out}(t) \tag{B.2}$$

where τ is a time constant that indicates the average time that blood takes to traverse the vessel, and it corresponds to the ratio between the resting blood volume, V_0, and the resting blood flow, F_0 ($\tau = V_0/F_0$). In the discussed framework, the outflow is function of the volume ν; this function models the balloon-like capacity of the vessel to expel blood at greater rate when distended. Defining the parameter α as the resistance of the venous balloon (the vessel stiffness), the outflow has the following form:

$$f_{out} = \nu^{\frac{1}{\alpha}} \ . \tag{B.3}$$

The second model assumption (*ii*), i.e. the link between the oxygen extraction and the blood flow, is the instrument to determine the equation for the deoxyhemoglobin content q. The change of the deoxyhemoglobin content dq/dt is the difference between the delivery of deoxyhemoglobin into the venous compartment and the one expelled. The expelled deoxyhemoglobin is simply the outflow f_{out} times the concentration q/v . The delivered deoxyhemoglobin into the vessel corresponds to the blood inflow f_{in} times the oxygen fraction extracted from the inflow $E(f_{in}, E_0)$, normalized to the oxygen fraction extracted at rest E_0. That is:

$$\tau \frac{dq}{dt} = f_{in}(t) \frac{E(f_{in}, E_0)}{E_0} - f_{out}(\nu) \frac{q(t)}{\nu(t)} \ . \tag{B.4}$$

The fraction of oxygen extracted from the inflowing blood can be approximated as:

$$E(f_{in}, E_0) = 1 - (1 - E_0)^{f_{in}} \ . \tag{B.5}$$

In conclusion the Balloon model determines the following state equations for ν and q :

$$\begin{cases} \tau \dfrac{d\nu}{dt} = f_{in}(t) - \nu(t)^{\frac{1}{\alpha}} \\ \tau \dfrac{dq}{dt} = f_{in}(t) \dfrac{1 - (1 - E_0)^{f_{in}}}{E_0} - \nu(t)^{\frac{1}{\alpha}} \dfrac{q(t)}{\nu(t)} \ . \end{cases} \tag{B.6}$$

In the present work we took the values of the parameters of the oxygen extraction fraction at rest (E_0), the mean transit time (τ) and the stiffness exponent (α) from the work of Stephan et al., 2007 [34].
Observing the resulting system B.6, it is possible to notice that the only thing required to specify the evolution of q and ν, and thus of the BOLD signal, is the incoming blood flow f_{in}. The issue of how the blood flow depends on neural activity is discussed in the paper of Friston et al., 2000 [16] and summarized in the following section.

The regional cerebral blood flow (rCBF)

The regional cerebral blood flow (rCBF) is the bloody supply needed to satisfy the metabolic requirements of a specific brain region at any given time.
Blood flow and synaptic activity are linearly related, as appears by experimental evidences. Thus, Friston et al., 2000 [16] have chosen the most simple model to dynamical relation between the incoming flow f_{in} and the flow inducing signal s:

$$\frac{df_{in}}{dt} = s \ . \tag{B.7}$$

The signal s is assumed to be generated by the neural activity $u(t)$. In particular, in the paper of Friston et al., 2000 [16], the authors consider the behavior of s as homologous to a very damped oscillator: changes in neural activity $u(t)$ evoke an exponentially decaying vasodilatatory signal that is subject to feedback-regulation by the flow it induces. That is:

$$\frac{ds}{dt} = \epsilon u(t) - \frac{s}{\tau_s} - \frac{f_{in} - 1}{\tau_f} \tag{B.8}$$

where ϵ represents the effectiveness of the neural activity in producing an increment in the signal, τ_s is the time decaying constant of the signal, and τ_f is the time constant for autoregulatory feedback for blood flow.

Bibliography

[1] Alexander A. L., and Lobaugh N. J. (2007). "Insights into Brain Connectivity Using Quantitative MRI Measure of White Matter". In: Jirsa Viktor K and Anthony R McIntosh "Handbook of brain connectivity" Vol. 1. Springer, pp. 221-271.

[2] Allen E., Damaraju E., Plis S. M., Erhardt E., Eichele T., and Calhoun V. D. (2014) "Tracking whole-brain connectivity dynamics in the resting state". In: *Cerebral Cortex*, 24 (3), pp. 663-676.

[3] Banerjee A., Tognoli E., Assisi C. G., Kelso J. A. S., and Jirsa V. K. (2008). "Mode level cognitive subtraction (MLCS) quantifies spatiotemporal reorganization in large-scale brain topographies". In: *NeuroImage*, 42 (2), pp. 663-674.

[4] Breakspear M. and Jirsa V. K. (2007). "Neural Dynamics and Brain Connectivity". In: Jirsa Viktor K and Anthony R McIntosh "Handbook of brain connectivity" Vol. 1. Springer, pp. 3-59.

[5] Brunel N. and Wang X. J. (2001). "Effects of Neuromodulation in a Cortical Network Model of ObjectWorking Memory Dominated by Recurrent Inhibition". In: *Journal of Computational Neuroscience*, 11, pp. 63-85.

[6] Buckner R. L., Andrews-Hanna J. R., and Schacter D. L. (2008). "The brain's default network". In: *Annals of the New York Academy of Sciences*, 1124 (1), pp. 25-32.

[7] Buxton R. B., Uludag K. and Frank L. R. (1998). "Dynamics of blood flow and oxygenation changes during brain activity: the Balloon model". In: *Magnetic Resonance in Medicine*, 39 (6), pp. 855-864.

[8] Deco G., Jirsa V. K., Robinson P. A., Breakspear M. and Friston K. (2008). "The Dynamic Brain: From Spiking Neurons to Neural Masses and Cortical Fields". In: *Plos Computational Biology*, 4 (8), e1000092.

[9] Deco G., Jirsa V. K., McIntosh A. R., Sporns O., and Kötter R. (2009). "Key role of coupling, delay, and noise in resting brain fluctuations". In: *Proceedings of the National Academy of Sciences*, 106, pp. 10302-10307.

[10] Deco G., Jirsa V. K. and McIntosh A. R. (2011). "Emerging concepts for the dynamical organization of resting-state activity in the brain". In: *Nature Reviews Neuroscience*, 12 (1), pp. 43-56.

[11] Deco G. and Jirsa V. K. (2012). "Ongoing cortical activity at rest: criticality, multistability and ghost attractors". In: *Journal of Neuroscience*, 32 (10), pp. 3366-3375.

[12] Deco G., Jirsa V. K. and McIntosh A. R. (2013). "Resting brains never rest: computational insights into potential cognitive architectures". In: *Trends in Neurosciences*, 36 (5), pp. 268-274.

[13] Deco G., Ponce-Alvarez A., Mantini D., Romani G. L., Hagmann P., and Corbetta M. (2013). "Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations". In: *Journal of Neuroscience*, 33 (27), pp. 11239-11252.

[14] Fox M. D., Snyder A. Z., Vincent J. L., Corbetta M., Van Essen D. C., and Raichle M. E. (2005). "The human brain is intrinsically organized into dynamic, anticorrelated functional networks". In: *Proceedings of the National Academy of Sciences*, 102, pp. 9673-9678.

[15] Fox M. D. and Raichle M. E. (2007). "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging". In: *Nature Reviews Neuroscience*, 8 (9), pp.700-711

[16] Friston K., Mechelli A., Turner R., and Price C. (2000). "Nonlinear responses in fMRI: The balloon model, volterra kernels, and other hemodynamics". In: *NeuroImage*, 12 (4), pp. 466-477.

[17] Funahashi S., Bruce C.J., and Rakic P. S. G. (1989) "Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex". In: *Journal of Neurophysiology*, 61, pp.331-349.

[18] Ghosh A., Rho Y., McIntosh A. R., Kötter R. and Jirsa V. K. (2008). "Noise during rest enables the exploration of the brain's dynamic repertoire". In: *PLoS Computational Biology*, 4, e1000196.

[19] Hansen E. C. A., Battaglia D., Spiegler A., Deco G. and Jirsa V. K. (2015). "Functional connectivity dynamics: Modeling the switching behavior of the resting state". In: *Neuroimage*, 105, pp. 525-535.

[20] Honey C. J., Kötter R., Breakspear M. and Sporns O. (2007). "Network structure of cerebral cortex shapes functional connectivity on multiple time scales". In: *Proceedings of the National Academy of Sciences*, 104, pp. 10240-10245.

[21] Hutchison R. M., Womelsdorf T., Allen E. A., Bandettini P. A., Calhoun V. D., Corbetta M., Della Penna S., Duyn J. H., Glover G. H., Gonzalez-Castillo J., Handwerker D. A., Keilholz S., Kiviniemi V., Leopold D. A., de Pasquale F., Sporns O., Walter M., and Changi C. (2013). "Dynamic functional connectivity: Promise, issues, and interpretations". In: *NeuroImage*, 80, pp 360-378.

[22] Izhikevich E. M. (2005). "Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting". In: MIT press, pp. 55-87.

[23] Kötter R. (2007). "Anatomical Concepts of Brain Connectivity". In: Jirsa Viktor K and Anthony R McIntosh "Handbook of brain connectivity" Vol. 1. Springer, pp. 221-271.

[24] Lu H., Zou Q., Gu H., Raichle M. E., Stein E. A., and Yang Y. (2012). "Rat brains also have a default model network". In: *Proceedings of the National Academy of Sciences*, 109, pp. 3979-3984.

[25] Mechling A. E., Hbner N. S., Lee H. L., Henning J., Elverfeldt D., and Harsan L. A. (2014). "Fine-grained mapping of mouse brain functional connectivity with resting-state fMRI". In: *NeuroImage*, 96, pp. 203-215.

[26] Messé A., Rudrauf D., Benali H. and Marrelec G. (2014). "Relating Structure and Function in the Human Brain: Relative Contributions of Anatomy, Stationary Dynamics, and Non- stationarities". In: *PLoS Computational Biology*, 10, e1003530.

[27] Newman M. E. J. (2008). "Mathematics of networks". In: The New Palgrave Encyclopedia of Economics, 2nd edition. Palgrave Macmillan, Basingstoke.

[28] Oh S.W. , Harris J. A., Ng L., Winslow B., Cain N., Mihalas S., Wang Q., Lau C., Kuan L., Henry A. M., Mortrud M. T., Ouellette B., Nguyen T. N., Sorensen S. A., Slaughterbeck C. R., Wakeman W., Li Y., Feng D., Ho A., Nicholas E., Hirokawa K. E., Bohn P., Joines K. M., Peng H., Hawrylycz M. J., Phillips J. W., Hohmann J. G., Wohnoutka P., Gerfen C. R., Koch C., Bernard A., Dang C., Jones A.R., and Zeng H. (2014). "A mesoscale connectome of the mouse of the brain". In: *Nature*, 508(7495), pp. 207-214.

[29] Raichle M. E., MacLeod A. M., Snyder A. Z., Powers W. J., Gusnard D. A., and Shulman G. L. (2001). "A default mode of brain function". In: *Proceedings of the National Academy of Sciences*, 98 (2), pp. 676-682.

[30] Rubinov M. and Sporns O. (2011). "Weight-conserving characterization of complex functional brain networks". In: *NeuroImage*, 56, pp. 2068-2079.

[31] Shulman G.L., Fiez J.A., Corbetta M., Buckner R.L., Miezin F.M., Raichle M.E., and Petersen S.E. (1997). "Common blood flow changes across visual tasks: II. Decreases in cerebral cortex". In: *Journal of Cognitive Neuroscience*, 9, pp. 648-663.

[32] Snyder A.Z. and Raichle M.E. (2012). "A brief history of the resting state: The Washington University perspective". In: *NeuroImage*, 62, pp. 902- 910.

[33] Sporns O. and Tononi G. (2007). "Structural Determinants of Functional Brain Dynamics". In: Jirsa Viktor K and Anthony R McIntosh "Handbook of brain connectivity." Vol. 1. Springer, pp 117-147.

[34] Stephan K. E., Weiskopf N., Drysdale P. M., Robinson P. A., and Friston K. J. (2007). "Comparing hemodynamic models with DCM". In: *NeuroImage*, 38 (3), pp. 387-401.

[35] Vann S. D., Aggleton J. P. and Maguire E. A. (2009). "What does the retrosplenial cortex do?". In: *Nature Reviews Neuroscience*, 10 (11), pp. 792-803.

[36] Wang X.J. (2002). "Probabilistic decision making by slow reverberation in cortical circuits". In: *Neuron* 36, pp.955-968.

[37] Wong K.F. and Wang X.J. (2006). "A recurrent network mechanism of time integration in perceptual decision". In: *Journal of Neuroscience*, 26, pp. 1314-1328.