Detection and Interactions of Dynamic Communities in Brain Functional Connectivity Imaging

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Biomedical Engineering

eingereicht von

Gladys Léger

Matrikelnummer 1428355

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Robert Sablatnig
Mitwirkung: Univ.Lektor Dipl.-Ing. Dr.techn. Georg Langs

Wien, 14. April 2016

_____   _____
       Gladys Léger   Robert Sablatnig

Detection and Interactions of Dynamic Communities in Brain Functional Connectivity Imaging

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Biomedical Engineering

by

Gladys Léger

Registration Number 1428355

to the Faculty of Informatics

at the Vienna University of Technology

Advisor: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Robert Sablatnig
Assistance: Univ.Lektor Dipl.-Ing. Dr.techn. Georg Langs

Vienna, 14th April, 2016

_____   _____
Gladys Léger   Robert Sablatnig

Erklärung zur Verfassung der Arbeit

Gladys Léger
15, rue des chênes, 72600 Mamers FRANCE

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 14. April 2016

—————————————————
Gladys Léger

Acknowledgements

Abstract

The way the brain connects its different areas to provide its unequaled level of efficiency is an aspect of the brain working that is currently not fully understood. A way to characterize it is to map the networks constituting functional connectivity owing to rest-functional Magnetic Resonance Imaging (rest-fMRI), a non-invasive imaging technique. Owing to connectivity metrics such as correlation or coherence calculated between rest-fMRI signals from different regions of the brain surface, Resting State Networks (RSNs) are constructed. Although functional networks share topological similarities with anatomical networks, they are not static as it has been assumed until 2010. The complexity of the problem is thus increased and methods are developed to extract dynamic properties from functional networks with a high enough time resolution. The main approach currently used to analyze dynamic functional connectivity identifies a finite set of connectivity states consisting of activity pattern reoccurring across time and subjects. However, the functional connectivity also varies spatially over time and the representation of dynamic functional connectivity as states is too restrictive.

The aim of this master thesis is to develop a new approach to represent dynamic functional connectivity as workable networks respecting both spatial and temporal variability. It is inspired by community evolution mining in social networks and text topics and proposes a richer alternative to connectivity states. The functional networks are thus considered as dynamic communities that interact across time. These interactions are characterized at different levels by events and their analysis provides insights in the functional organization of the brain. A clean and robust representation of the dynamic functional connectivity and their interactions across individuals is thus established and applied to a population of 200 subjects. Six different dynamic communities are thus identified across the population. They share similarities with static RSNs and they are repeatable across subsets of the population . The temporal characteristics of their activation allow to detect recurrent pattern in their co-occurrence and they are characterized by events whose significance can be evaluated.

Contents

Introduction

This chapter introduces the aims of the thesis. They are first justified by the current challenges of brain functional connectivity analysis which permits to clearly define the problems to be solved here. Then, the methodological approach constructed and tested in this thesis is briefly depicted. Finally, the structure of the report is exposed.

1.1 Motivation

The idea that brain functions can be localized and mapped appears in the 19th century [PMN11]. Franz Joseph Gall (1757-1828) used the skull shape to estimate the function localization and was thus the father of *phrenology* [PMN11]. Although the assumption that the skull and the brain function were linked appeared to be false, it has provided the starting point for brain function mapping [PMN11].

Moreover, networks in nature are everywhere, from the ecosystem to the cell level [Spo13]. Modeling the brain connectivity with networks seems therefore a spontaneous thought and it has became possible with the development of non-invasive neuroimaging modalities[Spo13]. An extensive mapping of the structural connectivity residing in brain anatomy has been thus realized and later, the identification of structured correlations in physiologically-based signals has ushered to the idea of a functional connectivity supported by the anatomical connectivity [Spo13]. This functional connectivity can be detected with several modalities such as ElectroEncephaloGraphy (EEG) or MagnetoEncephaloGraphy (MEG), but the Functional Magnetic Resonance Imaging (fMRI) is the modality of reference for identification of functional networks named Intrinsic Connectivity Networks (ICNs) [SZPBPH15]. Rest-fMRI (rest-fMRI) is particularly suited for this task because it detects the same networks as task-fMRI but with a shorter experiment duration and with less constraints [SZPBPH15]. The association of neuroscience with network science enables to study brain function as a complex system and to stop thinking of a brain as a organ responding to environmental stimulus while its functional activity is in part

ruled by its network activation and deactivation [Spo13]. The ICNs are made of nodes corresponding to voxels, group of voxels or Regions Of Interest (ROIs). They are linked by edges if their corresponding time series are statistically linked and thus revealed an interaction [HWA$^+$13].

Furthermore, it has been unsurprisingly demonstrated that functional networks are not static [CG10]&[MMH$^+$11]. Instead of using the whole experiment duration to establish the pattern of ICNs, new approaches have then arose to build dynamic Resting-State functional Networks (RSNs) [HWA$^+$13]. Most of them assume fixed spatial network components that form recurrent activity patterns considered as states [SVB$^+$13]&[CMPA14]. No existing representation permits to consider neither the spatial variation of the network by itself, nor the interactions between networks. However, other scientific fields using network modeling, such as sociology, have already developed techniques for analysis of complex systems which can be adapted to neurology ([TSFZ11], [CLT$^+$11], [VBAW15]).

1.2 Medical Applications and Relevance

Some neuropathologies stay unexplained and most of them are incurable, but RSNs can be utilized as biomarkers for their diagnosis and their comprehension [HWA$^+$13]. It is of key interest because they most of the time result in very serious degradation of life quality and are a real burden for the society [DO14]&[LZL$^+$15]. Moreover, our brain directs our life. We think, we feel, we learn, we dream through it and functional connectivity networks at rest have been demonstrated to be one technique to observe how the brain processes. Cognition tasks can for instance be detected through fMRI signals [SRR$^+$11] and a lot of work is realized to see how cognition relates to resting state networks properties (for instance: [KHS$^+$11]&[HSGJ$^+$12]). Emotions [EVDVSV11] and learning [BWP$^+$11] have also been linked to resting state networks fluctuations, as well as sleep [TL14], daydreaming [KD14] or consciousness [HCPP14].

1.2.1 Illustrative Examples: Schizophrenia, Bipolar Disorders and Epilepsy.

Bipolar disorders can be differentiated from Schizophrenia with rest functional networks [AKPC13]&[DAB$^+$14]. Schizophrenia and bipolar disorders are two complex diseases that are difficult to exactly diagnose [DO14]. Indeed both of them present psychotic symptoms and it leads to misdiagnosis that not only has disadvantageous economical consequences but also provokes human suffering due to the treatment inefficiency [DO14]. Figure 1.1 from Rashid et al. [RDPC14] provides an example of how these diseases can be differentiated. It highlights the differences of dynamic connectivity patterns between Healthy Controls (HC), Schizophrenia (SC) and Bipolar Disorders (BP) patients. Here states of connectivity are defined (see Section 2.4.1 for the definition of a state) and each state corresponds to a combination of components that are linked or not when the state is active. For instance, in state 4, it is visible that bipolar disorders impacts connectivity because there is a connection between a component from *temporal domain*

and a component from *parietal domain* for bipolar disorders patients, while it is absent for healthy patients. It is thus state specific and it can be used as marker. Moreover state 3 also presents interesting connection variations between schizophrenia RSNs and bipolar disorders RSNs. This kind of markers thus provides a promising tool for an early categorization of diseases [SZPBPH15].
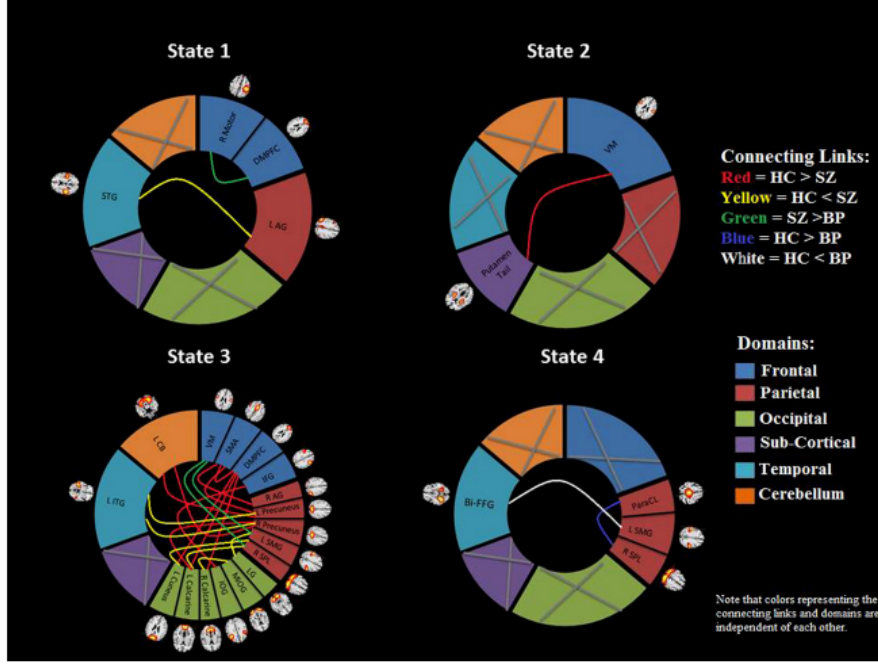
Figure 1.1: Visual summary of significant connectivity differences in four dynamic states. Each state represent an activation pattern on the brain surface that can be characterized by active subunits parts of different anatomical domains (frontal, parietal, occipital, subcortical, temporal and cerebellum). These subunits activation can be correlated or not and when they are, a colored link connects them if this connection is more present in one category of the subjects (HC: Healthy Control, BP: Bipolar Disorders, SZ: Schizophrenia). If a region shows no connectivity differences between the groups, it is marked with a gray cross. *from [RDPC14] with the permission by the authors.*

The graph properties of the networks are also used for characterization of diseases (see Section 2.2.3 for more details about graphs theoretic approaches in brain functional connectivity). Epilepsy is for example considered as a network disorder by Pedersen et al. [POWJ15]. By a graph analysis on static rest functional networks, they observe that epilepsy patients present more isolated network nodes and an higher network segregation. Their hypothesis to explain this observation is that it is a defense mechanism to prevent "instigation or spread of focal seizure" [POWJ15].

Other neuropathologies, such as Alzheimer disease ([HMP[+]13], [LOFB14], [BKP[+]15] for instance) or migraines [LZL[+]15], are also investigated through study of brain functional

connectivity.

1.2.2 Present state of development

Functional networks provide still tracks to be explored to understand and maybe contribute to prevent or at least reduce the effects of these diseases. Most of these possible applications are still at the research level, but static fMRI is already used for pre-surgical planning in order to identify hubs in the networks[SZPBPH15]. For instance it permits to shorten the intraoperative functional mapping and to prevent awake craniotomy needed to determine trajectories to access the pathological region and delimit the zone of resection [SZPBPH15].

It may become a diagnosis tool for brain-related disorders in the future but further studies are still required to assess the accuracy and the reliability of the functional dynamic networks extracted from rest-fMRI signals with low Signal-to-Noise Ratio, high variability across individuals and high sensitivity to environmental factors [LSS13]&[SZPBPH15].

1.3 Problem Statement and Aims

The main contribution of this thesis is the development of a robust representation of the dynamic RSNs on the whole brain surface. This representation enables the analysis of the interactions between RSNs. In contrast to other methods that consider only static networks or characterize the dynamic properties of RSNs by temporal chains of connectivity states (see Chapter 2), it integrates the spatial variability of localized RSNs as well as their temporal activation pattern (see Section 7.1).

Spatial overlap between the networks at time t is not allowed and it permits to know to which network a node belongs to. However, the temporal overlap is possible in order to observe which networks have a coordinated activity. Moreover, no prior assumptions concerning the regions to be integrated in the networks are made (see Chapter 5). The representation thus developed requires to circumvent noise problem to extract the maximum quantity of information from the fMRI signals (see Section 8.1).

In addition to an accurate identification of dynamic functional networks, tools and methods to detect interactions between networks are defined (see Section 5). The goal is to reveal, if it exists, one or several redundant interactions across subjects between identified networks.

To conclude, the aims of this thesis are:

- Detection of dynamic localized RSNs represented as Dynamic Communities (DCs),

- Development of algorithms for the analysis of dynamic networks interactions,

- Validation of the representation with a cohort of 200 subjects from the Human Connectome Project (HCP) [VEUA$^+$12] and detection of recurrent interactions .

1.4 Methodological Approach

The developed representation is based on Dynamic Communities (DCs) constructed as series of similar communities detected via the correlation between parcels of the brain. One DC corresponds to a dynamic RSN and if it is repeatable across the population, it is validated as a true dynamic RSN (see Section 7.1). The temporal overlap between the periods of activity of the DCs, latter named *co-occurrence*, is a first information concerning the possible interactions between the RSNs since two RSNs which co-exist during a long period are likely to interact (see Section 7.2). The interactions of DCs are characterized by *events* that are based on the starting and stopping of the activity of DCs and that are more or less significant (see Section 7.3).

After usual rest-fMRI pre-processing steps and parcellation of the brain surface, DCs are detected for each subject owing to these main steps:

1. Sliding Window Analysis to capture the correlation of the parcels at each time frame.

2. Community Detection to define a set of communities for each time frame.

3. Tracking and consolidation of DCs within subjects.

4. Identification of corresponding DCs across subjects on population level.

1.5 Structure of the Thesis

This thesis is structured as follows.

Chapter 2 provides first a brief overview of the physical and physiological basis of fMRI. Essential information to understand the significance of brain functional connectivity are then given before the presentation of the already implemented methods for the detection and analysis of RSNs, first as static network and then as dynamic networks.

Chapter 3 describes two methods: community detection and spectral clustering. They are part of the workflow and contribute to the construction of the representation.

Chapter 4 deals with the dataset used in this thesis and the way it is preprocessed to suppress the noise inherent to fMRI signals.

Chapter 5 focuses on methods used to build our new detection approach of dynamic RSNs, first at subject-level and then at population-level.

Chapter 6 consists in the description of three approaches developed for the validation of the results.

Chapter 7 presents the main experiments realized in order to validate our approach and define the best parameters for the representation. A description of the results goes along each experiment.

Chapter 8 discusses the results and the different choices made for the robustness of the representation.

Chapter 9 finally concludes the thesis with the major findings and the possible guidelines for future work.

State of the art

This chapter provides an overview of the current state of the art for brain dynamic function detection. The physical and physiological basics of fMRI are first briefly described in order to explain what is the meaning of the data used. In the second section, the main definitions and properties of brain function are given to construct the frame of the analysis. Then, the approaches already developed for identification of RSNs are presented to motivate the contributions of our method. Finally, the issues linked to these networks are discussed in order to demonstrate the utility of a new dynamic RSNs detection.

2.1 Functional Magnetic Resonance Imaging (fMRI)

fMRI signals are obtained voxel-wise from Magnetic Resonance Imaging (MRI) scans. It is therefore necessary to understand first the physical basics of image formation of MRI to be able to deal with fMRI. This imaging modality is a measure of a biological process which induces constraints that must be taken into account for every fMRI experiments [PMN11].

2.1.1 Nuclear Magnetic Resonance (NMR)

The NMR signal in medical imaging is based on the fact that a large part of the human body consists of protons ^1H [Lau73]. The proton is a paramagnetic atom as it has an unpaired electron and its spin I is thus non zero: $I = \frac{1}{2}$. Therefore its angular momentum $\vec{p} = \vec{I}.\hbar$ is also non zero, as well as its nuclear magnetic moment $\vec{\mu} = \gamma \vec{p}$. $\hbar$ is the Planck's constant and γ the gyromagnetic ratio which nucleus-specific [HSM04].

Without an external magnetic field, the spins of protons are randomly oriented and the vector sum of their magnetic moments, also named magnetization $\vec{M}$, is equal to zero. However, if an external magnetic field $\vec{B_0}$ is applied, the spin and consequently the magnetic moment, can have only two different orientations due to the rules of

quantum mechanics: parallel or anti-parallel to the direction of $\vec{B_0}$ [Str11]. Each of these orientations corresponds to an energy level. When an ensemble of spins without any interaction is considered, both energy levels are equally occupied and the total magnetization is equal to zero. However, if spins can interact with an energy reservoir and with each other, they tend to choose a state of minimal total energy and the lowest energy level is more occupied. A slightly higher number of spins are parallel to $\vec{B_0}$, the orientation with the lowest energy level, and the total magnetization $\vec{M}$ is no more equal to zero [Str11]&[HSM04]. Moreover, applying $\vec{B_0}$ also creates a torque on the magnetic moments and it causes a precession motion of $\vec{\mu}$ around $\vec{B_0}$-axis at the *Larmor frequency* $\omega_0 = \gamma B_0$, with B_0 the magnitude of the static magnetic field $\vec{B_0}$, in order to conserve the angular momentum [Str11]&[HSM04]. $\vec{M}$, the sum of all individual nuclear magnetic moments $\vec{\mu}$, is constant and aligned with $\vec{B_0}$. It is the equilibrium magnetization $\vec{M_0}$ and it has thus no precession motion in a constant external magnetic field. However, $\vec{M_0}$ is far too low to be directly measured (4.10^{-9} times B_0) and it needs to leave the equilibrium [Str11]&[HSM04]. It is realized by applying RadioFrequency (RF) magnetic field $\vec{B_1}(t)$ pulses resonating at the Larmor frequency ω_0. $\vec{M}$ is in a precession motion around an axis defined by the vector sum of static magnetic field $\vec{B_0}$ and time-varying RF magnetic field $\vec{B_1}$, as long as $\vec{B_1}$ is applied. The system is excited and increases its energy [Str11]&[HSM04]. In the rotating frame about $\vec{B_1}$-axis at Larmor frequency, $\vec{B_1}$ appears stationary. This is illustrated in Figure 2.1 where the rotating frame is (x', y', z') where x' corresponds to $\vec{B_1}$ direction and $\vec{B_1'}$ is the stationary RF field in this frame. If $\vec{B_1'}$ lies along the x'-axis, orthogonally to $\vec{B_0}$, then $\vec{M}$ precesses around x'-axis. The pulse-duration (2-3ms) must be just long enough to make this $90°$ rotation possible [Str11]&[HSM04].
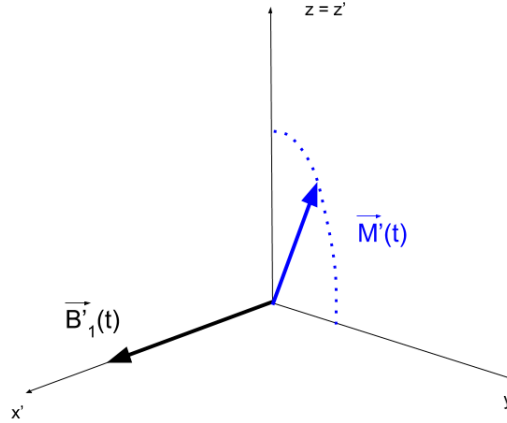

Figure 2.1: Magnetization during the RF pulse in the rotating frame (x',y',z') at Larmor frequency around z-axis

Then, when the RF field is switched off the relaxation of $\vec{M}$ from the (x, y) plane to the equilibrium state $\vec{M_0}$ aligned with z-axis is observed and this time, it is possible to measure

$\vec{M}$ by detecting the signal generated by its rotation into a coil. Indeed, by Faraday's law, a varying magnetic flux inside a coil induces voltage changes [Str11]&[HSM04]. Furthermore, it is necessary to spatially encode the data and it is realized by three sets of linear gradient coils in x-, y- and z-direction [Str11]&[HSM04]. First of all, a thin slice is "selected" owing to a linear magnetic field gradient in z-direction applied simultaneously with a RF-pulse shape determining the slice profile. It is now necessary to encode the 2D slice. Linear gradients are applied in x- and y-direction to realize respectively a frequency and phase encoding. Indeed, Larmor frequency is different for each row constructed by the gradient in x-direction and the precessions have different frequencies. For the phase encoding in y-direction, the phase is accumulated in the rotating frame because of the frequency encoding. It creates a new space named *k-space* which is the Fourier Transform of the image space for the reconstruction [Str11]&[HSM04]. Here, we have just given a brief overview of the basics of the techniques for image formation in 2D but it is enough to understand how to locate where does the information come from. However, this information we are interested in needs to be defined.

The relaxation of the magnetization can be decomposed into two mechanism: *spin-lattice* relaxation and *spin-spin* relaxation [HSM04]. Spin-lattice relaxation is characterized by its time constant T_1 and it corresponds to the return of magnetization longitudinal component M_z to the equilibrium. Its physical meaning is the transfer of the energy gain during RF-pulse from the spins to the energy reservoir, the lattice, mentioned above [HSM04]. Its temporal evolution after the RF field is switched off is illustrated on Figure 2.2.

$$M_z(t) = M_0 \left(1 - e^{-t/T_1} \right)$$
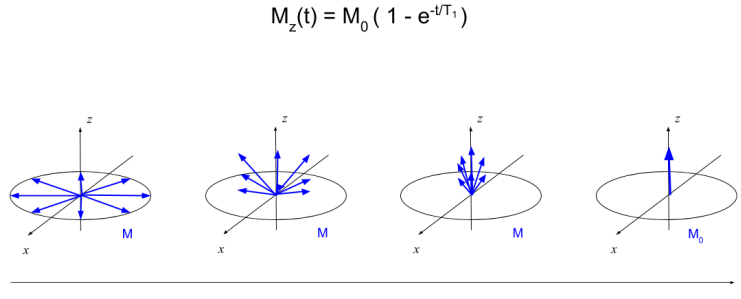
Figure 2.2: Evolution of the longitudinal magnetization M_z after several times T_2. M_0 is the magnitude of $\vec{M}$ at the equilibrium.

On the other hand, spin-spin relaxation stands for the interaction of spins between each other since each spin is a magnetic dipole and can slightly influence the magnetic field of its neighbors by the additional magnetic field it generates. It can thus also influence the precessional frequency [HSM04]. After $\vec{B}_1$ is switched off, it results in a progressive loss of phase between the rotating spins that is finally visible by an exponential decay of the magnetization in transverse plane M_{xy} with the time constant T_2 as illustrated by Figure 2.3.

$$M_{xy}(t) = M_{xy}(0) \ e^{-t/T_2}$$
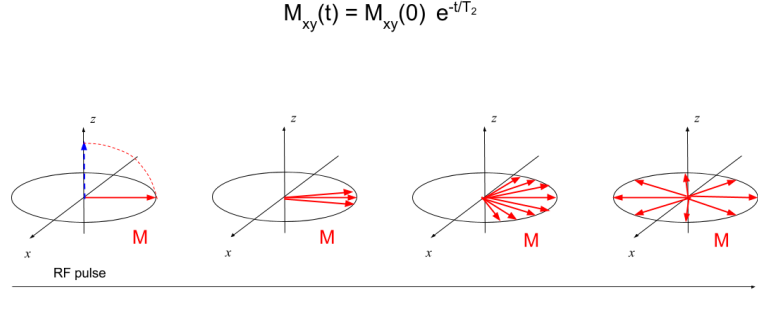
Figure 2.3: Evolution of the transverse component of the magnetization M_{xy} after the RF pulse.

The time constant T_2 is always lower than T_1 and both are highly dependent on the strength of $\vec{B_0}$ and on the tissue type which is excited [Str11]&[HSM04].

The RF-pulse is repeated at an interval named *Repetition Time (TR)* until enough data are acquired to generate an image [Str11]&[HSM04]. The choice of the TR is very important for the formation of the image since it gives more or less time to tissues to relax. If TR is short, tissues with short T_1 appear brighter than those whose T_1 are longer and therefore it emphasizes the difference of T_1. The *Echo Time (TE)* corresponds to the delay between the end of the excitation and the data acquisition [Str11]&[HSM04]. With a short TE, only the spins from tissues with the shortest T_2 have time to lose their phase coherence. These tissues appear darker than those with longer T_2 whose transverse magnetization has not enough time to decay. Therefore, the differences in T_2 between tissues are more visible. Moreover due to inhomogeneities in $\vec{B_0}$, spins lose their phase coherence quicker and we thus need to define an *effective transverse relaxation time $T_2^* < T_2$* [Str11]&[HSM04]:

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma \Delta B_0 \tag{2.1}$$

This formula is the basis of fMRI since *Blood Oxygen Level Dependent (BOLD)* signal induces magnetic field inhomogeneities also included in ΔB_0, the variation of static field magnitude [Str11]&[HSM04].

2.1.2 Blood Oxygen Level Dependent (BOLD) signal

fMRI has been developed in the early 1990s by Kwong and al (1992) [KBC+92]. It results from the observation that T_2^*-weighted magnetic resonance signals slightly increase (about 1%) in particular regions of the brain when the neuronal activity increases and experiments were inspired by the imaging of blood flow by Positron Emission Tomography (PET) [KBC+92]. Indeed, fMRI utilizes the physiological phenomenon that when a neuron is activated, it requires glucose and oxygen, and thus the region of the brain

where this neuron is located receives a larger blood flow. One particularity that still intrigues scientists and remains unexplained is that the amount of blood supplying the concerned neuron is higher than the one needed for a simple re oxygenation of the cell after energy consumption induced by the activation [PMN11]. This region presents thus a local excess of oxygen in veins and fMRI measures a signal depending on the variation of oxygenation, named BOLD signal. The variations of oxygen concentration is detectable by fMRI owing to the magnetic properties of hemoglobin. Indeed, hemoglobin bound to oxygen is diamagnetic since it presents no unpaired electron and no magnetic moment, while deoxyhemoglobin is paramagnetic due to an unpaired electron and a nonzero magnetic moment [Str11]. The variation of oxy/deoxyhemoglobin concentration creates thus inhomogeneities in the magnetic field due to the change of susceptibility and it influences the speed of phase decoupling of protons' spin. Thus the time T2* changes during activation, that is to say when the proportion of oxyhemoglobin is higher, and a T_2^*-weighted image provides information on the activation pattern [PMN11].

fMRI can thus detect neuronal activity that results in a surplus of oxygen linked to an increase of blood flow that is called hemodynamic response [PMN11]. Figure 2.4 illustrates the properties of the hemodynamic response that determine the interpretation of BOLD signals in fMRI. It has first to be noticed that the response to activation is rather slow since for a neuronal activity of few milliseconds, the maximum blood flow, that is to say the maximum of BOLD signal, is reached after 5 seconds. It forms a plateau as long as the neural activity is sustained and then it requires around 10 seconds to return to the baseline. Another aspect of BOLD signal demonstrated by Figure 2.4 is the high variability of response among subjects: change in BOLD signal can have different amplitudes and the time characteristics are also not the same among subjects. Moreover, different regions in the same subject may present temporal variability in their hemodynamic response, even if the delay before the maximum has a small variance [PMN11]. It can be explained by a difference in the vasculature and also by the task triggering the activation when different functional regions are compared [PMN11].

It can thus be tricky to analyze fMRI signals and preprocessing is highly necessary to obtain comparable results [PMN11]. Furthermore, the variation of BOLD amplitude is only about few percents, even if the blood flow is a very good indicator of neuron activation since it can locally increase about 60% during one task [HSM04]. However, blood flow is not the only physiological parameter that influences the BOLD signal. Indeed, the blood volume, determined by the current diameter of vessels, as well as the proportion of oxygen used will also determined the oxy/deoxyhemoglobin ratio and thus the BOLD signal, as illustrated in Figure 2.5. These three physiological components can also be influenced by diseases and other physiological parameters which explain the high variability of BOLD signals between individuals [PMN11].

2.1.3 Properties of fMRI

To obtain T_2^*-weighted imaging with short TR such as those of Human Connectome Project (HCP) data (see Section 4.1) and thus with high temporal resolution (sub second),
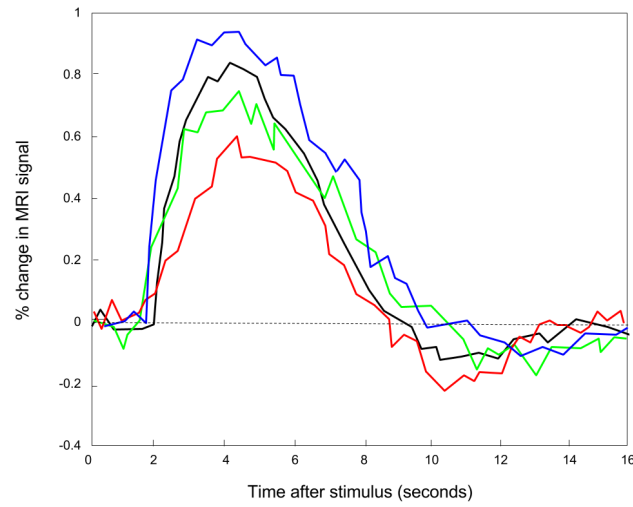
Figure 2.4: Hemodynamic responses to a 500ms stimulus sampled every 250ms. The four different plots are the hemodynamic response from different subjects. *figure following [PMN11].*
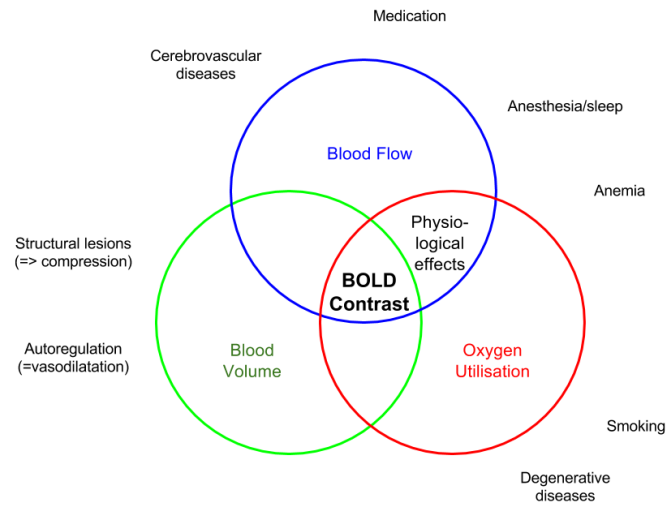

Figure 2.5: Representation of the different components of the BOLD signals and example of factors that can influence it. Inspired from `http://www.fil.ion.ucl.ac.uk/spm/course/slides10-zurich/Kerstin_BOLD.pdf`.

it is necessary to use Echo Planar Imaging (EPI) [ZFC$^+$14]&[SVB$^+$13]. On the contrary to classical data acquisition methods such as spin echo, EPI records several echoes of the RF-pulse per pulse, until the transversal magnetization has completely decayed. It permits to apply a different phase encoding to each echo and thus to acquire faster the whole volume of interest, since it is even possible to acquire in one TR the entire range of phase encoding steps, even if several pulse sequences are preferable for the image quality [PAMB$^+$01].

The spatial resolution of fMRI is limited by the size of the capillaries from where the BOLD signal is emitted [HSM04]. It is not able to resolve details at smaller scale than approximately 100 microns. However, a compromise between the spatial resolution and the scanning time is necessary since increasing the spatial resolution also increases the number of voxels to acquire. The choice of the spatial resolution depends on the aim of the experiment and for an entire brain acquisition, as in our case, few millimeters are enough [HSM04]&[SVB$^+$13].

2.1.4 Use of fMRI

It exists two type of fMRI experiments: rest-fMRI and task-fMRI [HSM04]. During rest-fMRI experiments, subjects receive the instructions to "lie still, think of nothing in particular, and not fall asleep" [SVB$^+$13]. The variability of what happens in the brain of all subjects is huge because nobody can think of nothing and rest experiments last from 5 to 15 min [PMN11]. However, as it is demonstrated in the following section, recurrent and meaningful information is extracted from resting-state.

On the other hand, task-fMRI requires the subjects to realize specific tasks at precise moments of the experiment [Str11]. A clear paradigm has to be constructed to ensure the reliability of the results. The set of possible tasks is infinite and task only requires a relative simplicity for its reproducibility and also a clear interpretation described by the paradigm. Thus, many academic fields use fMRI with success, such as cognitive neuroscience, psychology, psychiatry and radiology for example [HSM04].

fMRI is also used for clinical applications and especially for presurgical planning for brain tumor resections. It can for instance help the neurosurgeon to decide if the operation is too hazardous because the tumor is located in a key functional region, or also to choose the best approach to remove the tumor without additional tests during the operation [SZPBPH15]. Furthermore, fMRI is able to contribute to a personalized medicine approach by indicating how a tumor or a treatment has reorganized the functional connectivity[SZPBPH15].

2.1.5 Other modalities for detection of brain functional connectivity.

After fMRI, the main modalities that are used to characterize resting-state functional connectivity (see Section 2.2.1) are the ElectroEncephaloGraphy (EEG), the MagnetoEn-cephaloGraphy (MEG) and the Positron Emission Tomography (PET) [HWA$^+$13]. PET is historically the first modality that has detected functional connectivity [FFLF93].

The EEG measures the neural activity by recording electrical signals produced by the brain via scalp electrodes. The advantage of this technique is its high temporal resolution on the order of the millisecond, with a measurable bandwidth 1-100 Hz [BS09]. However, mainly because the electrical signal needs to go through several different resistive layers before reaching the electrodes, the spatial resolution is low, on the order of centimeter, while the fMRI provides a temporal resolution on the order of millimeter [BSR+15]. EEG and fMRI could therefore be complementary since a weakness of fMRI is its temporal resolution. Indeed, fMRI is restricted by the measurable frequency bandwidth 0.001-0.5 Hz [BS09]. The fMRI signal corresponds to an indirect measure of the brain activity via BOLD signal while in contrast EEG directly measures the electrophysiological activity of the brain. Several joint studies have been conducted to obtain both high spatial and temporal resolution ([BS09],[HWA+13],[CMPA14]). and it has been demonstrated that large scale EEG-detected RSNs (see Section 2.2.1) have "similar spatial patterns and correlated temporal dynamics to the networks independently derived from BOLD-fMRI data" [YDZ+15]. EEG can thus also be a validation tool for RSNs.

MEG is also a technique that measures electrophysiological signals. It records the extracranial biomagnetic field and detects its variations induced by the electrical neural activity. It thus also provides information to Functional Connectivity (FC), and on the contrary to EEG, it is not as much perturbed by changes in tissue conductivity and permittivity [HWA+13]. It has therefore an higher spatial resolution than EEG and conserves an higher temporal resolution than fMRI [HWA+13]. However, the spatial resolution stays low due to the indetermination of the inverse problem and it has also led to detect spurious correlations [dPDPS+10]&[HWH11]. On the other hand, it has permitted the early analysis of dynamic within- and across-network interactions, characterized at different frequencies and time scales [VLRM01]. However fMRI and MEG cannot be recorded simultaneously.

PET consists in injecting to the patient a positron emitting radiotracer which is involved in the biological process of interest. A measure of the positron emission permits to establish the distribution of the radiotracers in the body [SZPBPH15]. A common radiotracer is the FluoroDeoxyGlucose (FDG) which is considered the same way as glucose by blood owing to the high similarity between these two sugar [SZPBPH15]. It is very efficient to detect tumors because they need a lot of energy and therefore a lot of glucose. It is a mean to detect brain activity as we have seen that blood volume increases in active zone of the brain (see Section 2.1) and some "metabolic" networks revealed by this technique correspond to RSNs obtained via fMRI [DB12]. However, the spatial and temporal resolution of PET is lower than the one of fMRI or even MEG [DB12]. It is therefore not adapted for the analysis of dynamic RSNs.

2.2   Functional Brain Connectivity

The *Functional Connectivity (FC)* observed with fMRI enables modeling the interactions of brain regions and the properties progressively established owing to the networks it

forms [SVB$^+$13]. In this section, the different types of connectivity in brain are defined and the characteristics of FC are further discussed.

2.2.1 The notion of Functional Connectivity (FC) and other types of connectivity

The networks we want to detect in this thesis consist of nodes corresponding to locations on the brain surface (pixels, ROIs or identified components (see Section 2.3)). They are connected if there is a sufficiently high FC between them [SVB$^+$13]. FC is defined as "the temporal dependency of neuronal activation patterns of anatomically separated brain regions" [VdHHP10]. It evaluates to what extent distinct regions of the brain functionally communicate. It is differentiated from *effective connectivity* which only estimates direct connection and causality and which is not considered in this thesis [SVB$^+$13]. The different means to describe connectivity are explained in the Section 2.2.2.

FC is different from anatomical connectivity, also named *Structural Connectivity (SC)* [SVB$^+$13]. FC is revealed by metrics such as correlation between BOLD signals from different brain regions measured by fMRI. It only implies that activity from several regions are linked and these regions can be totally physically separated. Links between regions in SC are physical and can be detected by Diffusion Tensor Imaging for instance [PMN11].

There is an influence from cerebral anatomy to functional networks design [HSC$^+$09]& [VdHMKHP09] and its characteristics are still studied. Liao et al. [LYZ$^+$15] compare dynamic functional networks from rest fMRI with anatomical networks and demonstrate that structural connections limit the dynamicity of functional networks. Moreover, Shen et al.'s experiments [SHB$^+$15] on macaque shows that the stability of a functional networks is higher when the regions constituting this networks are anatomically linked. Liégeois et al. [LZP$^+$15] confirm these approach by demonstrating that dynamic functional networks with high modularity are primarily shaped by structural networks. These observations tend to confirm the hypothesis that SC would be a transitory structure for dynamic functional connectivity [LZP$^+$15].

The definition of networks also raises some issues. *ICNs* refer to the spatially static networks that are both detected during rest-fMRI and task-fMRI. A functional network detected at rest is named a Resting-State functional Network (RSN). ICNs are localized to gray matter regions [BDDS05]&[DLBDS$^+$06] and thus enable to restrict the analysis to the surface of the brain (see Section 4.1). The term has been introduced by Seeley et al. in [SMS$^+$07] in order to "avoid misconceptions evoked by resting-state networks" since these networks are obtained with a "resting" brain and prove by their existence that this brain is never really resting. Then, ICNs take on a new dimension when Smith et al. [SFM$^+$09] demonstrate that the functional modularity of the brain detected at rest can be extended to task experiments. The association of ICNs to task-fMRI finally permits to attribute functions to each ICN [LFE$^+$11]. These networks are even detected during sleep and anaesthesia [PDH13] and are consistent across subjects [YKS$^+$11] and also across some species [LHP$^+$13]. They are attributed to a function

according to their anatomical location and further task-experiments. The set of ICNs that is often used for the functional classification of networks is made of 7 networks ([YKS$^+$11], [LFE$^+$11], [ADP$^+$14], [ZFC$^+$14]): somatomotor, visual, auditory, ventral and dorsal attention, default-mode, and limbic networks. These networks can have different names in the literature and some of them are often considered as an aggregate of 2 networks. They are represented in Figure 2.6.

Connectivity structures detected at rest are reproducible during tasks. The fundamental difference between a rest and a task experiment lies in the fact that during rest, the cognitive activity is unconstrained [ADP$^+$14]. Therefore, during rest-fMRI the brain is involved in several types of mental activity which have an influence on functional connectivity organization [DDM$^+$10]. For instance, keeping the eyes closed or opened affects the spectrum of spontaneous activity and the functional connectivity pattern on specific brain regions ([MLPN$^+$08], [WEC10], [YLH$^+$09]). It can be remarked that EEG also detects this variation. Shirer et al. [SRR$^+$11] even classifies cognitive states with functional connectivity information. Therefore, rest-fMRI and task-fMRI are very similar, except that tasks during fMRI are not known. It is one of the difficulty to deal with in case of dynamic RSNs detection because all the subjects do not make the same cognitive tasks during rest experiments and the behavior of the detected networks can be totally different from one subject to another. However, a relatively high repeatability across the population is finally obtained, as demonstrated in Section 7.1.

2.2.2   Functional Connectivity (FC)

The previous paragraph defines what has to be detected, the FC, and what it represents. It is now necessary to model it to be able to measure it and form dynamic RSNs then.

Modeling the connectivity

FC observed with fMRI enables describing which distant brain regions interact with each other without apparent physical link. The FC between two data points of the brain can be estimated via different metrics such as correlation, coherence or mutual information between their respective time series [BS09]&[HWA$^+$13]. The most commonly used metric is the correlation as it directly provides the information if two regions are functionally connected. However, it must be kept in mind that a positive correlation can have different reasons, as presented on Figure 2.7 [PMN11]&[SVB$^+$13]. Indeed, if two regions A and B are correlated, it can either mean that A has a direct influence on B (and B on A respectively) and it is an effective connectivity, or that another region C is implied in this relation. It can result in an indirect influence if this region C is influenced by A and then influences B. Moreover, the region C can also directly influence both region A and B and it is then a shared influence [PMN11]. Therefore, if two data points have correlated time series, it does not mean the connection between them is direct. Another region can be involved in their interaction, but the latter can also be the result of variations in signal amplitude or noise level [Fri11]. The implicit assumption that correlation is equivalent to a direct connection has thus to be handled with care and has to be prohibited if the scale

Left

Right

■ Purple	(Visual)
■ Blue	(Somatomotor)
■ Green	(Dorsal Attention)
■ Violet	(Ventral Attention)
■ Cream	(Limbic)
■ Orange	(Frontoparietal)
■ Red	(Default)
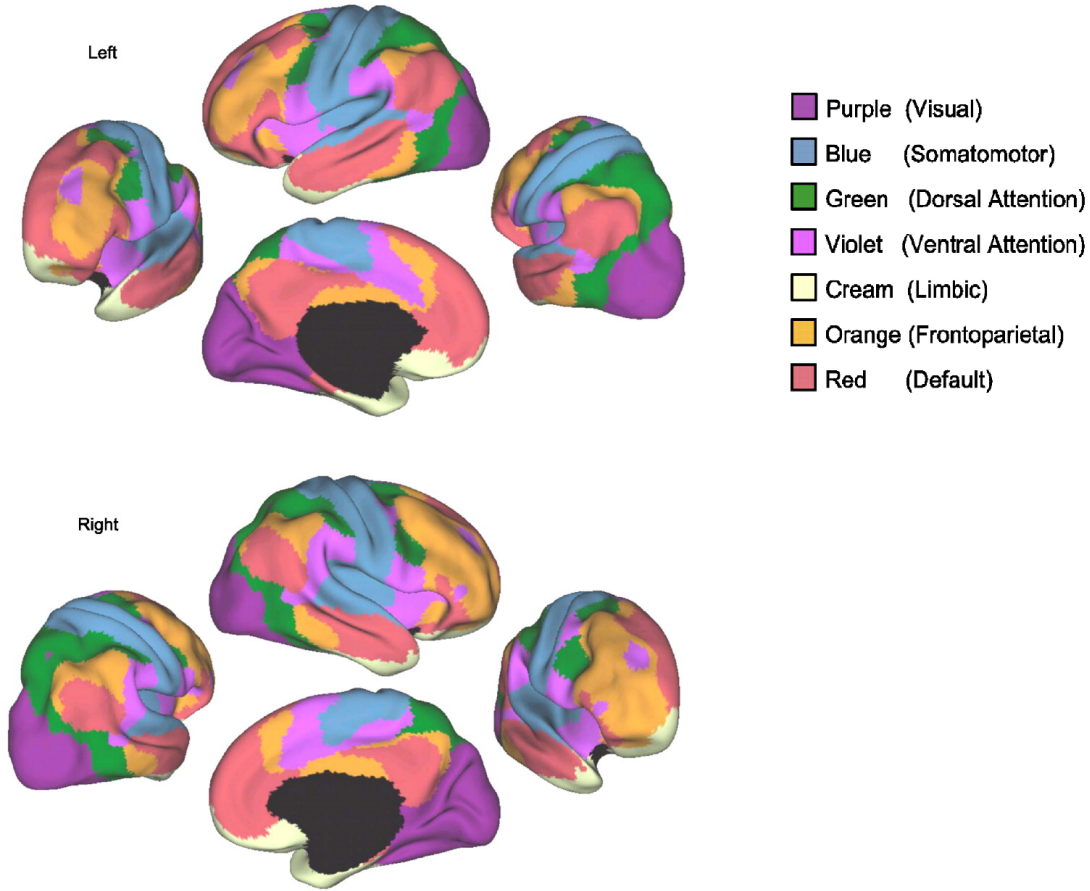
Figure 2.6: 7-networks partition of the human cerebral cortex based on rest-fMRI scans of 1 000 subjects. *from* [YKS+11]*with permission by the authors.*

of the sought networks is neural in order to identify the effective connections. However, it is commonly accepted for detection of functional networks [PMN11].

Another metric of interest is the partial correlation. On the contrary to correlation, it gives access to a more accurate evaluation of direct connections but not to their directionalities [SVB+13]. In the case of the three nodes example, the partial correlation is calculated by taking each pair of nodes to regress out the third time series from the two first ones, and calculating then the correlation between the two nodes. For instance, if the connection between node A and node B is indirect as illustrated in the second column of Figure 2.7., regressing out C from A and B permits to eliminate the spurious connection between A and B. For more than three nodes, always two of them are selected and the others are regressed out of these two nodes [SVB+13]. Other methods can be used to obtain an even more exact representation of the networks. However, it depends on the level of connectivity it aims to model. Figure 2.8 from [SVB+13] shows the spectrum
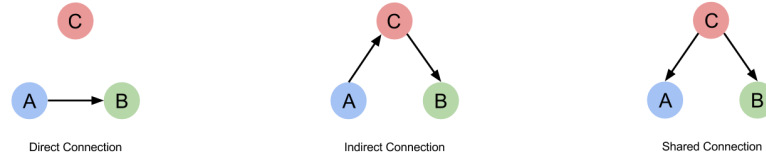
Figure 2.7: Different types of interactions between three nodes A, B, C that generate a correlation between A and B. The left panel depicts a direct interaction where correlation equal connection, the center panel illustrates an indirect influence from A to B via another node C and the left panel describes a shared influence from a common input region C to A and B.

of different network modeling approaches from neural-level brain to graph-theoretical networks. In this thesis, we are not interested in modeling the biophysical level at neuron scale and the fMRI is not adapted to this approach.
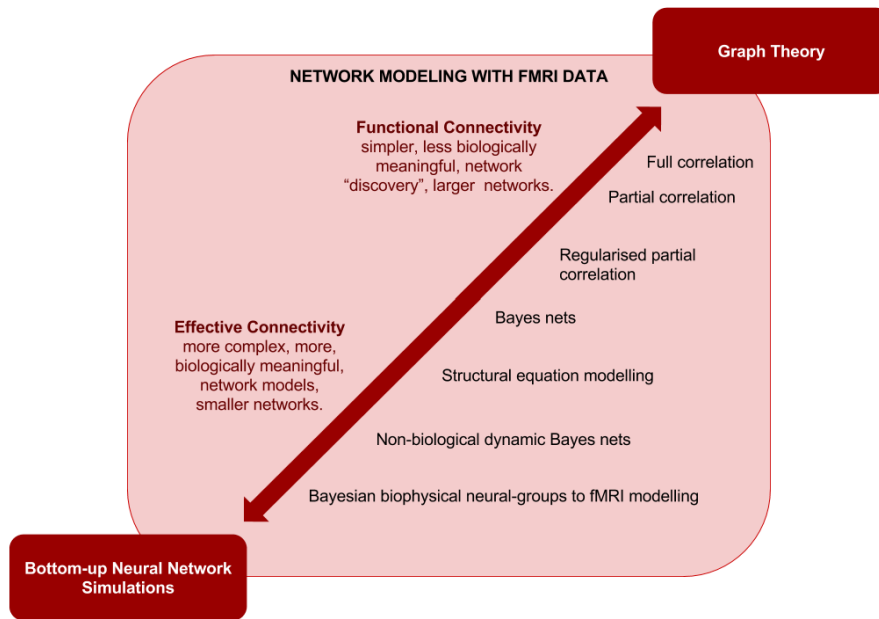
Figure 2.8: Different modeling-level of brain connectivity, from direct neuronal connection with effective connectivity to whole-brain Graph Theory with functional connectivity. *figure following [SVB+13].*

Therefore correlation, coherence or mutual information are the kind of metrics used to build functional networks based on Graph Theory [Fri11]. It provides less information than partial correlation for instance, but most of brain connections are anyway bidirectional and

even if correlation does not mean causation, correlation is a sign that a causation exists somewhere. If networks with higher spatial resolution are required, the computational cost increases and at the whole brain level, it becomes too high [SVB$^+$13]. As we are interested in the connectivity at the whole brain level, only metrics such as correlation, coherence or mutual information are discussed in this thesis.

Measuring Pairwise Connectivity: Metrics

Pearson's Correlation Coefficient   One popular metric to estimate the correlation is the *Pearson's correlation coefficient* r [HWA$^+$13]. The correlation between two time series (X, Y) is thus equal to their covariance $cov(X, Y)$ normalized by their respective variance σ_X and σ_Y, as described in Equation 2.2 where E is the expectation and μ_X and μ_Y their respective mean. It varies between -1 and $+1$ and $r = 1$ means a total positive correlation, $r = 0$ corresponds to no correlation and $r = -1$ is a total negative correlation.

$$r_{XY} = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X), (Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{2.2}$$

Spearman's Correlation Coefficient   Another correlation metric that is sometimes preferred to Pearson's correlation coefficient is the *Spearman's rank correlation coefficient* ρ [TF15]&[MMVDV15]. It also measures the statistical dependence between two variables, here two time series, but in a non-parametric way. It estimates how well the relationships between these two time series fits to a monotonic function, while Pearson's correlation coefficient refers to linear functions. Its value can be analyzed the same way as Pearson's correlation coefficient values. For two time series (X, Y) of same size n with respective raw scores X_i and Y_i, Spearman's rank correlation coefficient is defined as:

$$\rho_{XY} = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{2.3}$$

where d_i is the difference between rank x_i of X_i and y_i of Y_i.

Coherence   On the top of these two correlation coefficients, the *coherence* is also used to estimate if two brain regions are connected ([SMD04], [RAW15], [HD15], [YAMC15]). If their time series are ergodic - it means that their respective statistical properties can be obtained from a single, long enough, random sample of their time course - and if they are linearly related, coherence is a metric that estimates their causality with a value between zero, if they are not at all connected, and one, if one of the series can be perfectly predicted from the other. It differs from correlation coefficients since it requires a higher number of time points to be stable and it is therefore calculated over multiple time windows [RAW15]. It thus reduces the effect of noise but also the temporal resolution. The coherence $C_{X,Y}$ of two time series (X, Y) at frequency ω is computed as:

$$C_{XY}(\omega) = \frac{|G_{XY}(\omega)|^2}{G_{XX}(\omega)G_{YY}(\omega)} \tag{2.4}$$

19

with $G_{XY}(\omega)$ the cross-spectral density between X and Y and $G_{XX}(\omega)$ and $G_{YY}(\omega)$ the respective autospectral density of X and Y at frequency ω.

Mutual Information The *mutual Information I* is another metric that evaluates the FC. It permits to avoid the assumption of Gaussianity of the dependence structure between two time series (X, Y) of size n [CWBFF09]&[HPV$^+$11]. It is defined as:

$$I_{XY}(\omega) = \sum_{i=1}^{n}\sum_{j=1}^{n} p(X_i, Y_j) log \frac{p(X_i, Y_j)}{p(X_i)p(Y_j)} \tag{2.5}$$

with $p(X_i)$ the probability distribution function and $p(X_i, Y_j)$ the joint probability distribution function.

Other metrics are rarely used. For instance, Shine et al. [SKB$^+$15] have developed their own metric based on first order temporal derivatives of time series in order to increase the sensitivity of the SWA. The choice of the metric still depends on the aims of the experiments. Coherence is for instance more adapted than correlation when the frequency information has to be captured [YAMC15]. Pearson's correlation coefficient is a well-known statistical concept and it provides accurate results, at least as good as the other proposed metrics [HPV$^+$11]&[SKB$^+$15].

2.2.3 From Pair-wise Connectivity to Networks

Modeling brain connectivity with networks offers to use the rich set of methods in graph theory developed for multiple scientific fields [Spo13]. It is necessary to adapt the graph model to the information to be captured [SVB$^+$13]. In the following section, we review the graph theory concepts relevant for this thesis. In order to characterize the brain connectivity, the first studies typically used a small number of spatial maps built with the connectivity definition described above. Regions Of Interest (ROIs) were chosen and then, based on fMRI time series from other pixels of the brain, correlation spatial maps were built for each ROI [BZYHH95]. This approach is named a *seed based* approach and it is a model based approach since the "seeds" are fixed (see description of the method in Section 2.3.1). Later, brain connectivity has been considered as ensemble of networks, turning it into a model free approach using Graph Theory tools developed for numerous different scientific fields, such as biology or sociology [Spo13]. This is the most commonly used approach [Spo13]. It requires to define nodes, corresponding to functionally distinct brain regions, and edges to functionally link these nodes [SVB$^+$13].

One node can be a group of neighboring voxels resulting from a parcellation based on the similarity of their time series. Therefore, it is a clustering approach providing a hard parcellation without spatial overlap and with local nodes [dRVdH13]&[BJG$^+$13]. This is the approach we are using in this thesis. On the other hand Independent Component Analysis (ICA) on time series results in weighted spatial maps gathering distributed groups of points [Bec12]. Each map corresponds to one node of the global network and they can overlap with each other.

Once the nodes are defined, the edges are evaluated by comparing the time series of the different nodes consisting in the average of the pixels' time series contained in the corresponding node [SVB+13]. Most of the time, the directionality of the edges is not taken into account, except when models for effective connectivity are used, as explained in Section 2.2.2.

The networks and their analysis aim to answer questions such as: "How are the functional connections between brain regions organized? How efficiently can the brain integrate information between different sub-systems? And are there brain regions that have a specialized role in this efficient communication?" [VdHHP10]. To answer this kind of questions, graph theoretic approaches have been widely utilized [POWJ15]. Indeed, analyzing the connectivity structure of the human brain in the form of a network provides insights in its organization and function since the architecture of a network is directly impacted by its robustness, its performance to integrate information and to efficiently communicate [VdHHP10].

2.2.4 Graph Theoretic Approaches to Brain Network Analysis

Network Models

The most general categories of graphs are *regular, random* or *complex* networks [POWJ15]. A regular network shows high local connectivity with low global connectivity, while a random network is the opposite, a low local connectivity and a high global connectivity. As for the complex network, it displays a balance between local and global connectivity [POWJ15].

The graph metrics that are of interest to decide which of these three general models is the most fitted to model a network are [POWJ15]:

- The *local efficiency*: inverse of the shortest path length between connected nodes that are neighbors with the node of interest [POWJ15];

- The *clustering coefficient*: number of connections between the nearest neighbors of a node, that can be normalized by the maximum number of connections between these nodes for instance[SVB+13];

- The *modularity*: number of modules (highly connected groups of nodes that share few connections with other groups) inside the network [SVB+13];

- The *global efficiency*: average of the inverse of the shortest path lengths between all nodes [POWJ15];

- The *betweenness centrality*: proportion of shortest path length passing by the node [POWJ15].

Figure 2.9 illustrates these three models of networks by their respective typical adjacency matrix associated with their characteristic properties [POWJ15].
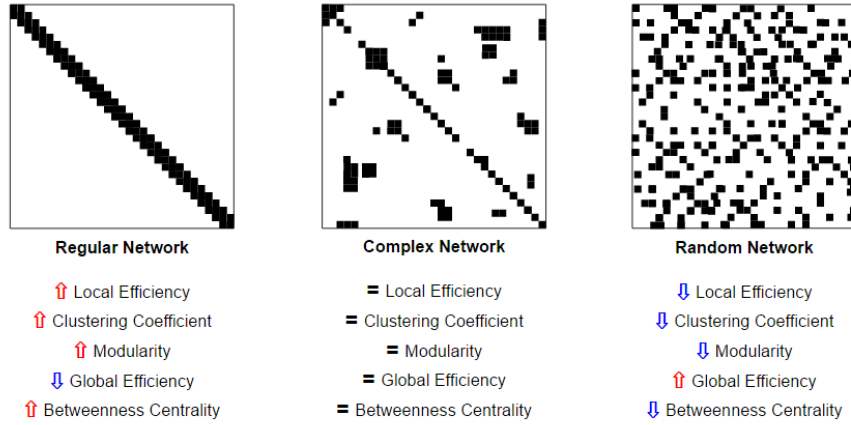
Figure 2.9: Adjacency matrices of three different network models (regular, random and complex networks) and comparison of their graph-theory metrics with complex model as reference. *Figure following [POWJ15].*

When brain connectivity is concerned, two models are of particular interest: the *small-world network* and the *scale-free network*, which are both complex networks [WC03]. Only three network parameters are required to determine if the studied network can be modeled by scale-free and/or small-world networks [Gri05]:

- The *connectivity distribution* $\mathbf{P}(\mathbf{k})$ which corresponds to the probability for any node to be connected to k other nodes,

- The *level of clustering* which is estimated by the clustering coefficient $\mathbf{C}$ defined above,

- The *average shortest path length* $\mathbf{L}$ between two nodes.

A *small-world network* shows a high level of clustering. It implies a high clustering coefficient $\mathbf{C}$ which means probable subgraphs within the global graph, and a short average node-to-node distance [WS98]. A *scale-free network* has a low average number of edges per node and yet a small number of highly connected nodes. The scale-free networks have thus a high global connectivity. Its connectivity distribution follows a power-law scaling: $\mathbf{P}(\mathbf{k}) \sim k^{-y}$ [BA99].

The combination of the properties from these two networks is recognized as a robust network architecture where information transfer and integration is processed with a high level of consistency and is therefore very adapted to model brain [VDHMP08]&[OHN$^+$14]. A large part of nodes are dedicated to local functioning owing to the small-worldness and fewer sets of nodes facilitate the global information processing [BS12].

Illustrations of use of graph metrics

If graph metrics such as **P**, **C** or **L** deviate from these complex model characteristics, considered as normal characteristics, toward random or regular network model characteristics, they can serve as biomarkers for disease detection for instance (see Section 1.2).

Centrality metrics enable to determine which nodes have an overwhelming influence on the rest of the graph and are therefore considered as *hubs*. A hub has a high *node degree* and a high *node betweenness*. It means that a lot of edges are connected to it and that a lot of shortest paths pass through this node. These particular nodes show for instance disconnections in neurodegenerative diseases and also in brain reorganization of comatose patients for instance [FRCA14]. Community structure analysis is also of key interest to comprehend brain function. It consists in detecting the more densely connected regions of the networks [FRCA14].

Limitations of graph theory

The reliability of graph analysis entirely depends on the correct network modeling at a lower level. The nodes have to be appropriately defined with a neurological base for the interpretation of the results [SVB^{+}13]. Moreover, non-directed graphs with the assumption that all connections are direct are used (see Section 2.2.2) and it has to be taken into account in the interpretation. Furthermore, graph metrics that are used are summarizing abstractions of the whole graph and no change in the metric does not mean no change in the graph since variations can be counterbalanced or averaged [SVB^{+}13]. A particular attention is needed for their interpretation since each graph metric can hold biases and the choice of the metric has to be adapted to the graph concerned [FRCA14].

2.2.5 From Static to Changing Graphs: Dynamicity

Until 2010 [CG10], all studies on resting-state functional networks have assumed that "the statistical interdependence of [fMRI] signals between distinct brain regions is constant throughout recording periods of task-free experiments" [HWA^{+}13]. It has provided remarkable advances in the field of neuroscience for the understanding of brain functions and yet it is based on an average of the spatio-temporal characteristics of these functions while dynamic aspects are likely to give further insight into the brain work [HWA^{+}13]. Recent research focuses on dynamic functional connectivity ([CG10], [ADP^{+}14], [LRVDV13], [AFPA14], [ZFC^{+}14], [KVDV15], [TF15], [YAMC15], [RAW15], [LLT^{+}16]), and successfully extends these dynamic properties to others species such as macaques or rats ([MMH^{+}11], [HWG^{+}13], [MHM^{+}16]). The dynamicity of brain function seems logical since brain must interpret, coordinate and respond in real time to internal and external stimulus, even at rest [HWA^{+}13].

The functional connectivity between two nodes can vary according to three different parameters: the strength of the connectivity measured by the absolute value of the

correlation coefficient, its sign and its existence since it can be lost if its strength is no more above a defined threshold as illustrated in Figure 2.10.
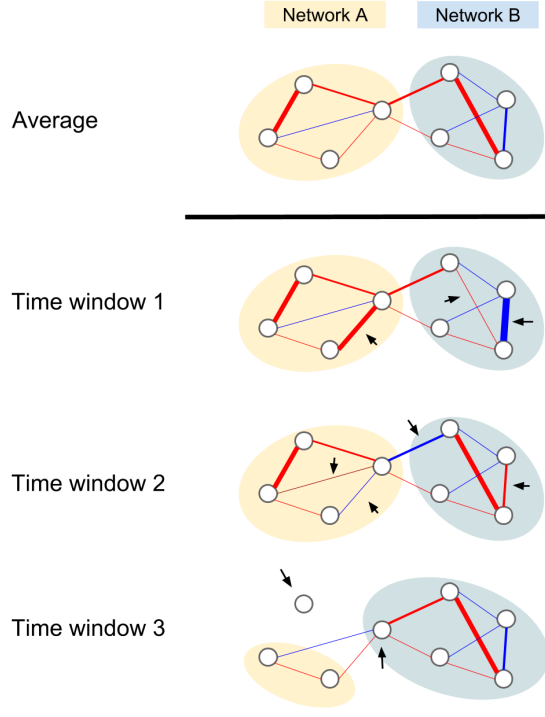
Figure 2.10: Illustration of time-varying properties of Functional Connectivity. The first row depicts an example of two averaged networks obtained when they are assumed to be static while row 2, 3 and 4 illustrate what such networks could look like at different time steps. The row 2 shows variations in the magnitude of the strength of the functional connectivity between two nodes, while the row 3 illustrates sign changes of strength and the row 4 represents lost or gain of connections depending on whether the connectivity strength is above or under a threshold. Such variations may lead to node membership changes as illustrated in time window 3 and the networks may have a different appearance at different time steps. Red edges represent positive connections, blue edges represents negative connections. *figure following [HWA+13].*

The formal definition for the *non-stationarity* of a time-series is that its mean and its covariance are not constant in time [HRGCB12]. The temporal variability of a metric alone does not mean it is non-stationary. In the context of functional connectivity, a difference must be made between true variability based on non-stationarity and variability generated by noise [HRGCB12].

FMRI signals are subject to numerous sources of non-stationary noise: non-neural noise such as noise generated by cardiac or respiratory activity, noise linked to the

measuring device, noise caused by the variations in BOLD signal mean and variance over time (see Section 2.5 for more details). All these noise sources can influence the dynamic functional connectivity metrics and induce dynamicity in a true stationary signal [HWA$^+$13]. In order to determine if observed dynamicity is generated by noise and has nothing to do with cognition, the changes in functional connectivity has been compared with other simultaneous measurements of neural or physiological processes such as EEG (see Section 2.1.5) and their reproducibility across populations has been established [HWA$^+$13]. Moreover studies which examine dynamic functional connectivity demonstrate a certain level of relation with cognition since it has even been proved that detected metastable states can be attributed to stages of consciousness [HCPP14].

The functional dynamic properties are also affected by diseases such as schizophrenia or epilepsy (see Section 1.2) and it is an additional argument for the relevancy of this dynamic fluctuations. Dynamic functional networks also share incontestable similarities with structural networks, as explained into Section 2.2.1.

It is important to know the time scale of interest for brain functional networks dynamicity to set the parameters of the methods designed to detect it. Van Dijk et al. demonstrate that the ICNs become stable within 4-5 minutes [VDHV$^+$10]. Therefore, according to the temporal resolution of fMRI, the functional connectivity is observed on the scale of the second. EEG and MEG have also been used to observe microstates on the scale of milliseconds with high accuracy [CMPA14].

2.3 Detecting Static Functional Networks

In this section we review the standard approaches to identify static Intrinsic Connectivity Networks (ICNs). They are important since many approaches for identification of dynamic RSNs are based on these methods [HWA$^+$13]. They also provide a simple illustration of the issues in FC detection.

2.3.1 Seed-Based Correlation Analysis

An obvious approach to study the functional connection of one region, defined by a pixel or a group of pixels, with other parts of the brain is to analyze the correlation between its time series and the other available time series from different brain region. It thus provides a functional connectivity map based on a particular region named "seed". A more sophisticated model than the one resulting from a linear correlation analysis is established by using a General Linear Model (GLM) if regressors of no interest are available [CSB10] (see Section 4.2 for a short description of a GLM)[CSB10] . The choice of the seed location is for instance determined by a functional activation map established during a previous task experiment [VdHHP10]. A region with a very high activity during the task is selected and then, with the resting state fMRI signals, the functional connectivity information of this region is extracted. The process of such a seed-based correlation analysis is described in Figure 2.11.
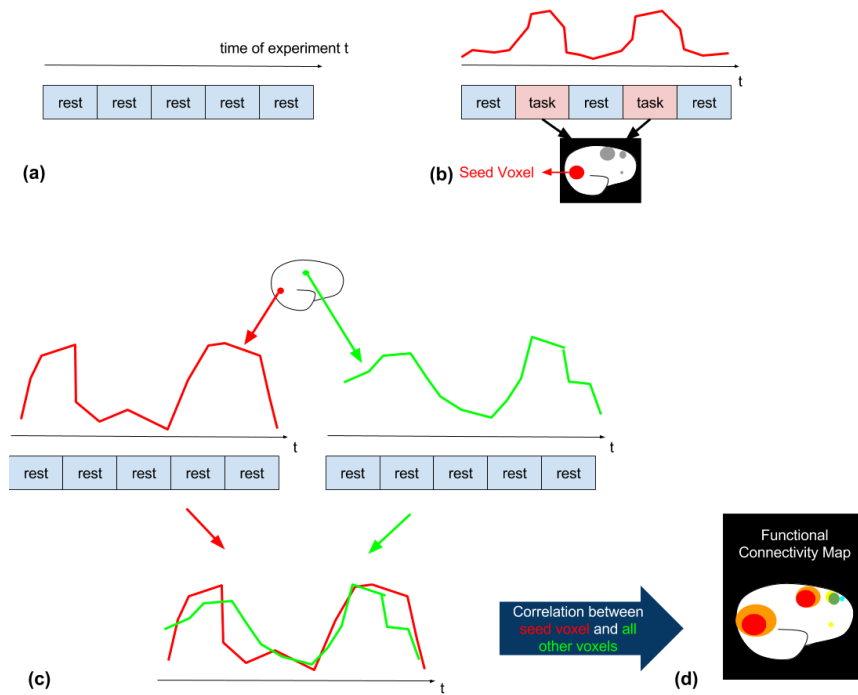
Figure 2.11: Different steps of a seed-based correlation analysis. (a) a resting-state fMRI experiment is measured. (b) from a task fMRI signal, a region of high activity is selected as region of interest, later used as seed voxel. (c) to study the functional connectivity between this seed region and other parts of the brain, the correlation between the seed voxel time serie from rest fMRI and the other time series corresponding to all other voxel is calculated. (d) a functional connectivity map is established from the correlation information. A high correlation between voxel j and the seed voxel means a high functional connectivity. *adapted from [VdHHP10].*

This approach has been the first one used for the evaluation of functional connectivity in 1995 by Biswal et al. [BZYHH95]. They has succeeded to demonstrate that left and right hemispheric regions of the primary structural motor network have correlated fMRI signals at rest and they have concluded that it is a sign for information exchange, or in other words, for FC. Several other studies have then confirmed this results for motor network and also highlights other connections of this anatomical network with primary visual network, auditory network and cognitive networks (for instance: [CHC+02], [DLSDS+05], [DRB+06]). Seed-based correlation analysis has also been used for clinical applications, to study the effect of aging for instance [AHSV+07] or to analyse physiological processes such as sleep [LPZN+09]. It is a very straightforward method that, although it results in a limited quantity of information concerning the FC, provides results easy to analyze on top of being uncomplicated to implement [BV07]. Indeed it gives a direct answer to the direct question about the connections of one ROI. Moreover, its reliability in the detection of static resting-state network has been demonstrated by Shehzad et al

[SKR$^+$09].

However, this method has major drawbacks. Using only few maps from the seed-based correlation analysis is an under-representation of the data and networks can be missed [CSB10]. Seed-based correlation analysis can also be influenced by structured spatial confounds such as other networks than the one which includes the seed, and also structured noise like residual head motion artifact [CSB10]. A temporal filtering reduces this negative influence but the residuals are always present in the detected networks. Furthermore seed size and location must be carefully defined since it influences the final networks [BAHS08].

It is possible to use all the available time series to establish the connectivity networks instead of using the univariate approach of the seed-based voxel analysis, and it is therefore preferred [CSB10]. Although comparing all the time series in a multivariate approach increases the complexity of the network interpretation, it also permits to obtain consistent functional maps that are not reachable with a seed-based approach since there are potentially as many spatial maps as possible seeds [CSB10].

2.3.2   Spatial Independent Component Analysis (ICA)

The ICA is a method that aims to represent non Gaussian multivariate data owing to linear combination of statistically independent vectors, or at least as independent as possible. It optimizes a measure of "non-Gaussianity" between the estimated components in order to obtain the "maximally statistically independent, non-Gaussian components" [CSB10]. A common example used to illustrate the idea of ICA is the situation where two people are speaking at the same time in a room, the cocktail party problem [HO00]. They are recorded by two microphones located at two different places and provide thus two different signals: $x_1(t)$ and $x_2(t)$. These signals are weighted sums of each acoustic signal emitted by the speakers s_1 and s_2, and the weights a_{ij} depend on the distance of the speaker j from the microphone i:

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) \tag{2.6}$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) \tag{2.7}$$

The aim of the ICA is to identify (s_1, s_2) from (x_1, x_2) without knowing weights (a_{ij}). It is possible by considering that s_1 and s_2 are statistically independent [HO00]. In a more general case, we want to separate n random variables $s = [s_1\ s_2 \ldots s_n]$ from a set of linear mixtures $x = [x_1\ x_2 \ldots x_n]$. The problem can be written as:

$$x = \mathbf{A}s \tag{2.8}$$

where $\mathbf{A} = (a_{ij})$ is the unknown mixing matrix. The ICA model is a generative model that identifies latent variables called independent components [HO00]. The decomposition into independent components is only possible if the components s_i are assumed to be statistically independent. It means that the value of s_i can not influence the value of

s_j and it is formally described by the equality between the joint probability density function $P(s_i, s_j)$ of s_i and s_j and the multiplication of their respective probability density functions $P(s_i)$ and $P(s_j)$. An additional assumption is also that the components do not have Gaussian distribution. To solve Equation 2.8, **A** is first estimated and then inverted to obtain s [HO00].

Application to RSNs

Spatial ICA is utilized in static functional connectivity analysis in order to extract independent spatial map to constitute nodes of a functional network. It means that, through ICA, the global resting state connectivity pattern can be described by a finite number of independent spatial source of resting state signals [VdHHP10]. The independence of the spatial map is not based on the spatial information but on the correlation of the signals. Therefore, spatial overlap is possible between spatial maps and it is not a hard parcellation as the one created by clustering [BDDS05]. Moreover, according to the quantity of information we are interested in, it is possible to choose a low or a high dimensional ICA. A low dimensional ICA, that is to say 10 to 30 components, is close to a seed-based analysis since each map generated by the ICA can be considered as one single seed-based map. Increasing the number of components up to several hundred of components gives access to more numerous and more detailed nodes [SVB$^+$13]. It thus increases the available quantity of information since we can for instance look which nodes inside two large-scale interacting networks are responsible from this interaction as illustrated on Figure 2.12 [SVB$^+$13].
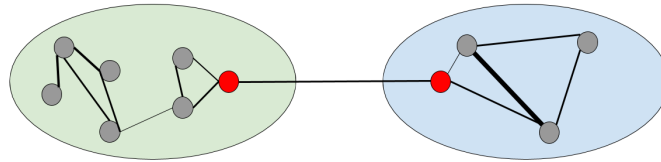
Figure 2.12: Two large scales networks interacting via the nodes in red.

One of the advantages of ICA compared to seed-based analysis is that ICA uses the whole-brain data without pre-established ROIs and is able to simultaneously compare activities from multiple distributed voxels. Consequently prior spatial assumptions are not necessary and ICA provides a richer description of the networks [SVB$^+$13]. Moreover, RSNs detected by ICA are less sensitive to noise than seed-based correlation analysis ([BMB08], [MBH$^+$09], [CSB10]). Indeed, structured noise effects can be separated from the signal by being sidelined into additional components that are then not considered as resting-state networks.

ICA has also drawbacks. First, ICA is a stochastic process and the components are built via iterative optimizations. It implies a run-to-run variability even on the same data but

software solution have been developed to establish an averaged solution from multiple ICA decompositions and circumvent this problem [ADP$^+$14]. Moreover, it has already been pointed out that it requires to specify a number of components. There are statistical tools to estimate the optimal model order but with a neurobiological point of view, it is impossible to consider there is a "best" number of components since the ICNs are hierarchically organized and this hierarchy is complex to capture [CSB10]. Every resting state networks can be separated in subnetworks and so on. Therefore, by choosing a number of components for the ICA, a complexity level in this hierarchy is also selected but the robustness of finest levels is restricted by the data quality, particularly the experiment duration [CSB10]. Furthermore, a large number of components representing subnetworks require more effort for the identification and classification [TFA$^+$08]. One of the main disadvantages of ICA compared to seed-based correlation analysis is the increased complexity for the interpretation of the results [CSB10].

Principal Component Analysis (PCA) is a possible alternative to ICA since they are both explorative methods [LRG$^+$13]. However, ICA is more adapted for fMRI signals analysis since the independence between components of ICA only implies that their time courses are not highly correlated, while for PCA, the decomposition is based on the orthogonality between the time courses and it can miss partially correlated signals [CSB10]. The ICA is therefore closer from biological models we have described in Section 2.2.2.

ICA has become one of the most popular method for the definition of RSNs and has provided interesting results, even for dynamic RSNs ([ADP$^+$14], [CA12], [HWA$^+$13], [RDPC14], [ZFC$^+$14]).

2.3.3 Clustering

Clustering aims to group data points with high level of similarity into a single group and data points with low level of similarity into different groups, so-called clusters [VdHHP10]. The metric used to measure the distance between the data points in terms of similarity is the correlation or any other measure described in Section 2.2.2 to estimate the connectivity.

On the contrary to ICA, clustering does not allow spatial overlap between networks and are patient specific [VdHHP10]. Indeed, to compare the results of clustering across subjects, additional seed-like processing steps are necessary. However, the obtained results are rather similar [VdHHP10].

Clustering algorithms have been used to build static functional networks, such as hierarchical [CHC$^+$02]&[SSC$^+$05], Laplacian [TDP06] and normalized cut clustering [VDHMP08]. However, this approach is no more used and ICA is preferred to clustering-based methods or PCA which is also another possibility for the detection of resting-state networks [CSB10].

2.4 From Static to Dynamic Networks

This section deals with the methods that expand network detection to dynamic networks that change in time. They can be split into two kinds of approaches [HWA+13]. The first one observes pairwise dynamic variations between regions by the mean of a Sliding Window Analysis (SWA) for instance. On the other hand, the second approach is multivariate and examines "global changing pattern of synchrony" via temporal ICA for example. All these approaches are explorative and are not based on neurobiological knowledge [HWA+13].

2.4.1 Sliding Window Analysis (SWA)

The Sliding Window Analysis (SWA) is currently the most commonly used approach in the literature to introduce the temporal dimension in functional connectivity ([SPK+10], [ADP+14], [LRVDV13], [ZFC+14], [SKB+15]). It consists in selecting points owing to a window that has a fixed length and in extracting then the functional connectivity information within this window. It is exactly the same approach as the one used for the detection of static characteristics except that only a part of the signals is examined and then, the window is shifted by a fixed number of time points. The length of this shift is important since it defines the temporal overlap between the windows and it takes its value between one and the total length of the window minus one. The SWA can be applied directly on preprocessed fMRI signals but also on spatially averaged signals resulting from a parcellation via ICA or clustering [HWA+13]. It is also possible to combine it with a seed-based correlation analysis [LRVDV13]. Moreover, as an alternative to one window size, results from different time window length can be integrated [HWA+13]&[TF15].

Tapered window

Instead of a rectangular window with uniform weights attributed to the points inside the window, a tapered window can be applied. It reduces the sensitivity to outliers and also the spurious fluctuations [ZFC+14]. Indeed, a "rectangular" window means that all the time points inside the window have the same weight in the correlation coefficient estimated at t. Therefore, if there is a single outlier in the value of correlation, all the sliding time window that contain this point are biased over their whole time length, provoking a false lecture of the data due to a brutal stop of the effect caused by the outlier [ZFC+14].

A tapered window attributes a structure of weights to the points it contains, reducing the duration of the outlier effect and smoothing its influence [PDMA12]. Different types of functions can be defined to taper the window. One type of function commonly used in financial studies for instance is exponential functions that give the highest weight to the last point [ZFC+14]. The window correlation at time t is therefore more specifically influenced by the last time point of the window defined therefore as the point t, while the weights of the $w - 1$ previous time points exponentially decrease with the distance to time t. Gaussian functions convoluted with rectangular windows are another class of functions

that is commonly used for rest fMRI signals are concerned [ADP$^+$14]&[YAMC15]. The main difference with exponential functions is the symmetry of Gaussian kernels which attributes the same weights to time points from the past and from the future of the time point t which corresponds to the point of highest weight on the middle of the window. Figure 2.13 describes these different windows and the corresponding notations.
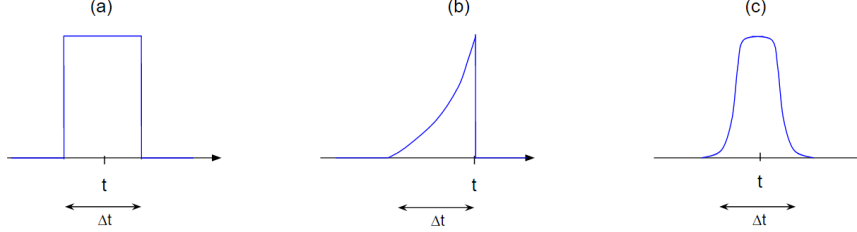
Figure 2.13: Examples of three different window types used in literature. (a) Rectangular window. (b) Exponential-tapered window. (c) Gaussian-tapered window.

The use of a tapered window has an impact on the calculation of Pearson's correlation coefficient [PDMA12]. From the definition of the window shape of size Δt, a structure of weights $\{w_1, \ldots, w_{\Delta t}\}$ are defined such as $\sum_{k=1}^{\Delta t} w_k = 1$ [PDMA12]. For a rectangular window, we have: $\forall\ k \in [1, \Delta t],\ w_k = \frac{1}{\Delta t}$. For a Gaussian-tapered or an exponential-tapered window, each weight corresponds to one point of the window as well.

Once the weights are defined, it is possible to compute the weighted Pearson's correlation coefficients r_{ij}^t in time window t of truncated signals $y_i = \{y_i(1), \ldots, y_i(\Delta t)\}$ and $y_j = \{y_j(1), \ldots, y_j(\Delta t)\}$ owing to the sample mean $\bar{y}_i^t$ and the corresponding variances σ_i^t and covariances σ_{ij}^t by applying the following equations [PDMA12]:

$$\bar{y}_i^t = \sum_{k=1}^{\Delta t} w_k y_i(k) \tag{2.9}$$

$$\sigma_i^t = \sqrt{\sum_{k=1}^{\Delta t} w_k (y_i(k) - \bar{y}_i^t)^2} \tag{2.10}$$

$$\sigma_{ij}^t = \sum_{k=1}^{\Delta t} w_k (y_i(k) - \bar{y}_i^t)(y_j(k) - \bar{y}_j^t) \tag{2.11}$$

$$r_{ij}^t = \frac{\sigma_{ij}^t}{\sigma_i^t \sigma_j^t} \tag{2.12}$$

Temporal variations of the functional connectivity metrics of interest, such as correlation coefficients in our case, are thus obtained. It is then possible to identify reoccurring patterns in the functional connectivity that are called states [HWA$^+$13] or to build networks inside each window and track them across time to build dynamic functional networks (see Section 5.3).

State Definition

Clustering methods can be used in order to detect reproducible correlation pattern in the temporal variation of functional connectivity. Clustering is applied among correlation matrices computed over time windows. Each row and each column of these correlation matrices represents a voxel, a ROI or a component from spatial ICA. Figure 2.14 adapted from [ADP+14] illustrates a complete process to detect connectivity states. Allen et al. [ADP+14] first perform a spatial ICA on the whole duration of rest-fMRI data to identify independent spatial maps that correspond to the static ICNs (see Section 2.3.2). Then, with a SWA, they compute a covariance matrix for each window and subsample them. They select the windows with the highest functional connectivity variance and name them "subject exemplars". Finally, they apply a k-means clustering algorithm to these sets of covariance matrices (one set per subject) in order to identify the redundant functional connectivity patterns and analyze their structure and frequency [ADP+14].

Clustering analysis provides satisfying results for the detection of states [HWA+13] but presents also shortcomings. Apart from the limitations introduced by the SWA, it is necessary to choose the adapted clustering algorithm and assumption, as well as appropriate parameters. Allen et al. [ADP+14] uses k-means clustering with L1 distance function and a number of clusters k=7, and perform the clustering only on selected time frames. They compare its results with those obtained with other algorithms such as hierarchical clustering, with other distance metrics such as correlation or Euclidean distance, and also test other number of clusters. They thus demonstrate that their results are consistent over these variations but further studies are necessary to study the sensitivity of states detection to clustering algorithms and parameters [HWA+13].

Furthermore, the following approaches are developed as an alternative for connectivity state detection. The clustering can for instance be based on the use of topological network metrics, such as modularity or community membership, as features ([BWP+11], [DBG12], [JVM+12], [KPCP12]). Indeed, graph theory is of key interest for network analysis and its metrics are extensively used to summarize static brain functional connectivity (see Sections 1.2 and 2.2.3). The temporal variability is introduced via multilayer network approaches that aim to quantify temporal variations owing to multilayer graph statistics or by directly observing the temporal fluctuations of graph metrics. In that last case, a weighted network is identified for each time window and the graph metrics are then calculated for each time window. Apart from state detection via clustering, it can be helpful to determine in which configuration (rich club, small worldness...) the brain spends more time for instance and compare this observation between diseased and healthy groups of patients [vdHKGS12]&[CMPA14]. Another approach to identify states is to detect change points in connectivity with model-based methods, therefore without using any clustering [CHA+12].
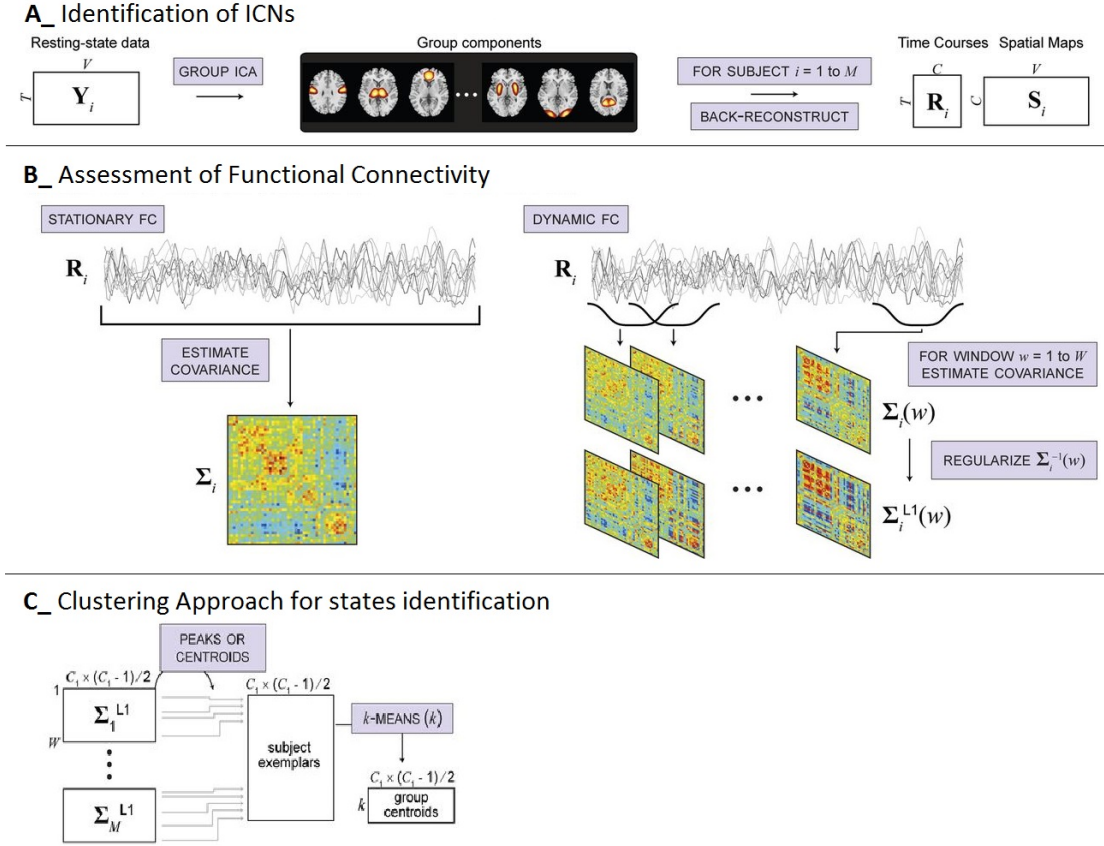
Figure 2.14: Steps for the states identification inside rest-fMRI signals. **A__** Static spatial ICA for the detection of spatial maps of interest. A group ICA is applied to the entire set of fMRI signals and independency between signal sources is maximized to obtained spatial maps whose corresponding averaged signal is then construct from original data by back-reconstruction. **B__** Functional connectivity analysis where a first evaluation of the covariance matrix is realized via stationary analysis and then the dynamic connectivity is extracted with a SWA. For each window, a covariance matrix is obtained and reorganized. **C__** First clustering on a selection of covariance matrices for each subject and then k-means to extract cluster centroids from the subject exemplars. The cluster centroids corresponds to a state. *adapted from* [ADP$^+$14]*, with permission by the authors.*

Limitations

SWA is a simple approach but it also has major shortcomings that have to be taken into account for the interpretation of the results. The most obvious issue in SWA is the choice of the window length w. As the frequency band of interest for dynamic analysis of fMRI is (0.01 - 0.1 Hz) (see Section 4.2) and because it is necessary to encompass several cycles of resting state fluctuations to obtain a robust estimation of functional connectivity [SVB+13], the window length w is included in the range (30 - 300 s) if the whole frequency spectrum of fluctuations is resolved. A too large time window does not detect fast connectivity variation by averaging the correlation or any other connectivity metric over too many time points and it is thus a low-pass filter behavior with a cutoff frequency of $1/w$ [TF15]. On the other hand, a too small window confers a high influence to high frequency components, acting like a high-pass filter with a cut-off frequency of $1/w$, and to noise too, since the number of points inside the window decreases and also because Signal-to-Noise Ratio of fMRI signals is particularly low at high frequency due to the low-pass properties of BOLD signal [HWA+13]. Shirer et al. [SRR+11] estimates that windows from 30 to 60 seconds permit to reliably detect cognitive states so it already reduces the range for the window length. However, Thompson et al. [TF15] demonstrate that the estimated functional connectivity is influenced by the window size because a window of length w is more sensitive to non-stationary phenomenon with a repeating frequency multiple of $1/w$. It is a problem for the accuracy of the connectivity analysis since Thompson et al. [TF15] also demonstrate that the fluctuations of within and between resting-state networks appear at different frequency and they report that the optimal window length for identification of visual network variations is different from the optimal window length that ensures the highest sensitivity for default mode network. Leornardi et al. [LVDV15] realize an extensive study to determine to what extent SWA generates spurious fluctuations of dynamic functional connectivity. They conclude that the common choice of a window length between 30 and 60 s is reasonable.

Another approach consists in applying wavelet transform to focus on particular frequency ranges (see paragraph below on time-frequency analysis). The Wavelet Transform Coherence (WTC) is for instance utilized to access to instantaneous correlation coefficients at different frequency bands owing to a window length varying according to the frequency content of the time series ([CG10], [HWG+13], [YAMC15]). However, the integration of the results for different frequency bands is complex and there is no relevant summarizing measures yet [LVDV15]. Furthermore, Lindquist and al. [LXNC14] take inspiration from the field of finance and apply multivariate volatility models to fMRI signals to study their functional connectivity in a parametric way, without any choice of frequency band of interest. They demonstrate their method to be less sensitive to noise than SWA but further studies are necessary to establish the reliability of their approach. To conclude, there are leads for alternatives to SWA and its window length problem. However, an informed and cautious interpretation of the results from SWA provides recurrent well-defined dynamic networks in an intuitive way that has been extensively studied [HWA+13]. Therefore, SWA is currently the best established method for a direct

analysis of resting-state dynamic functional connectivity [LVDV15].

2.4.2 Time-Frequency Analysis

A time frequency analysis can be applied to rest-fMRI time series in order to estimate the coherence and phase lag, also named time shift, between each couple of time series. It utilizes WTC to realize a multi-resolution time-frequency coherence analysis [TC98]. It provides information dependent at once on time and frequency without selection of time window length since the effective size of the analysis window is the scale of the wavelet and it is adapted to the frequency components of the signal [HWA+13]. First, shorter windows are used for fast fluctuations (high frequencies) and then, the window length progressively increases to capture lower frequency variations. This exploratory analysis provides a high quantity of information such as the frequency which has the highest influence on coherence in a particular region, or the temporal variations of magnitude and phase relationships between two nodes within a given frequency band. It is an asset but also a drawback when it must be extended to multiple subjects or even to multiple brain regions[HWA+13]. It is then necessary to decrease the dimensionality of the information.

Chang and Glover [CG10] for instance design a time-averaged coherence profile. Frequency bands can also be selected and the overall fluctuation of variability is estimated via standard deviation or mean-squared successive differences [HWA+13]. Yaesoubi and al. [YAMC15] are the first to extend this technique to a whole-brain analysis. They first decompose the resting-state data into independent components via spatial ICA and apply a complex wavelet transform to obtain information from both time and frequency domains. Then, they apply a clustering to identify the different functional networks and detect the recurring patterns of connectivity corresponding to states in both temporal and frequency domains. It directly provides the temporal fluctuation and frequency and phase profiles of each recurrent functional connectivity states. They extract states that are present in many frequency bands as well as states occurring only in narrow ranges of frequency. It demonstrates the frequency domain also hold information that can be used as a complement to the dynamic connectivity information extracted from the temporal domain [YAMC15].

2.4.3 Temporal Independent Component Analysis

A temporal ICA identifies "distinct temporal functional modes" from the rest-fMRI signals [SVB+13]. These modes are the extracted components from the temporal variability of the nodes and serve as functional networks. ICA allows spatial overlapping of its component [SVB+13]. However, the spatial or temporal ICA both provides models with constant spatial weight over time and thus are not able to indicate the possible time-varying correlation strength between nodes [SVB+13]. Moreover, temporal ICA components are obligatory independent over time. Therefore, in the case of one node shared by two mutually exclusive functional networks, if this node is particularly active for only one function at time t, the temporal ICA considers the networks as temporally

independent and the correlation is equal to zero, while a negative correlation has to be detected [SVB$^+$13].

Importance of Networks spatial overlap

The main difference between SWA and ICA is based on the possibility of overlapping networks. Smith et al. [SMM$^+$12] demonstrates that for instance the default mode network can be decomposed into several overlapping subnetworks. They also applied SWA to original time series and to times series reconstructed from the temporal functional modes obtained by temporal ICA. The variability shows a difference of 25% between the two SWA models and they assume that this difference is due to the overlap of the networks. Indeed, it must be considered that some regions of the brain can have different roles by being involved in different functional networks. However these networks can also variate over time and interact, possibilities that temporal ICA cannot deal with [HWA$^+$13]&[SVB$^+$13].

Independent Vector Analysis (IVA)

Independent Vector Analysis (IVA) examines spatial coupling and therefore allows dynamic changes in the spatial patterns while ICA establishes only the time-varying functional connectivity between spatially fixed networks [CMPA14]. The principle of IVA is to extend ICA to multiple datasets. These dataset can be generated via SWA, as illustrated in Figure 2.15. Instead of applying the ICA to the whole dataset, ICA is applied to subsets where obtained components, also named sources, are still mutually time-independent but in assuming that each component is dependent on at most one component from another subset. Then a joint source separation is performed between the subsets to take into account this dependence [KLL06]&[AFPA14]. Therefore, the independence of the sources inside temporal subsets are maximized and their dependence across subsets is also captured. A high dependence of sources across the dataset means that the corresponding networks are approximately static while a high variability characterize dynamic networks. Figure 2.15 describes the different steps of the IVA that finally describes spatiotemporal variability of functional connectivity via a data-driven approach.

2.4.4 Co-Activation Patterns (CAPs)

Previous studies remark that in specific individual time frames, the networks obtained via methods such as seed-based correlation analysis or spatial ICA, are similar to ICNs obtained on the whole time course [HWA$^+$13]. Therefore, it means that within the individual time frame where the signal amplitude of the nodes of the concerned ICN is high, one of the detected networks has a high spatial similarity with this ICN. ICNs can thus be identified from a small number of time frames from the resting state fMRI signals [LD13].
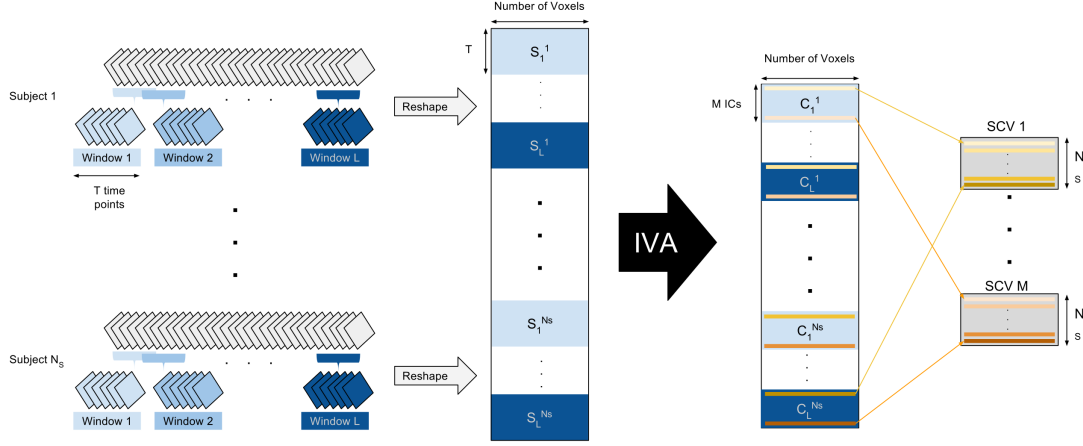
Figure 2.15: Steps of Independent Vector Analysis. A SWA is first applied independently to N_S subjects and the resulting data are reorganized into subsets C_i^j containing data from window i in subject j. L is the number of windows with a fixed length T. These subsets are the input of a temporal IVA decomposition that identify M Independent Components (ICs) for each window and they are contained into C_i^j. These components are then reorganized via a joint source separation into M spatial maps of dynamic connectivity named Source Component Vectors (SCVs). *figure following [CMPA14].*

Based on this observation, Liu et al [LD13] apply a clustering algorithm on selected individual BOLD volumes registered during a rest-fMRI experiment and use the spatial similarity as distance between them. The BOLD volumes are selected if they present a seed region with a high signal intensity. The resulting cluster centroids are named Co-Activation Patterns (CAPs) and they depend on the selected seed. These CAPs intend to "characterize a set of representative instantaneous configurations of BOLD activity" [HWA+13]. Liu and al. identify 30 different CAPs but they observe that their connectivity patterns are very different from those obtained with spatial ICA. Moreover, they remark that the static networks obtained with a seed-based correlation analysis performed over a given time interval are actually a sum of CAPs. Therefore, the temporal variations observed in SWA can be considered as the relative occurrence of CAPs between the consecutive time windows [LD13]. It is a new approach that demonstrates the existence of different level in the time-varying connectivity [HWA+13]. The observation of this reoccurring single-volume snapshots of BOLD activity also raises the issue of whether consecutive sequences of BOLD activity are also recurrent. Majeed and al. [MMH+11] develop an algorithm to reliably answer this question and they detect similar patterns across a scan that are reproducible across subjects and that for instance involve well-known regions of the default-mode network. These identified patterns are responsible from 25 to 50 % of the variance in the low frequency BOLD time courses [HWA+13].

It is now established that rest functional connectivity can be decomposed spatially and temporally into recurrent patterns. Karahanoğlu et al. [KVDV15] also demonstrate

that temporal components not only reoccurred into cycles as Majeed et al. [MMH$^+$11] assume it, but also present temporal overlaps, as well as spatial patterns. They use recent methods of signal processing to temporally deconvolve fMRI signals on their entire spectrum. Indeed the bandpass filtering is made unnecessary by the prior analysis of hemodynamic response function and they thus obtain denoised fMRI signals. With the deconvolved signals, they identify "transient" time points that are characterized by high variation of the signal intensity (activation and de-activation) and observe the activity spatial map at this particular moments. To study the consistency of these spatial maps, they apply a clustering on them and identify 20 CAPs which spatially overlap. Furthermore, by back-projection, they reconstruct their time lines and observe temporal overlapping between specific resting-state networks that also have a behavioral significance [KVDV15].

2.5   Limitations and Challenges

There are numerous approaches that aim to represent FC, from a static to a dynamic representation [CMPA14]. Each one have specific drawbacks but all of them are also affected by the major limitation of fMRI, that is to say noise [HWA$^+$13].

2.5.1   Noise

The Signal-to-Noise Ratio of fMRI signal is low and the partition of the signal for dynamic functional analysis still decreases it since less data points are used [HWA$^+$13]. Moreover, the noise is hard to estimate because its strength varies across the experiment and this variation can also appear as spurious dynamic function information during analysis [HWA$^+$13]. It is the reason why preprocessing steps are essential, particularly for dynamic connectivity analysis. However, all the noise cannot be removed in spite of numerous approaches developed to reduce it (for instance [BL07], [CCG09], [PBS$^+$12], [KIE$^+$12]). Some methods integrate the denoising to their application. The ICA, temporal or spatial, can for instance increase its dimensionality to enable to select only the most reliable components and consider the others as noise [CSB10]&[ADP$^+$14]. Karanoglu and al. also developed an approach based on recent signal-processing techniques where they need to deconvolve the fMRI signals and it results into denoised signals [KVDV15].

2.5.2   Weaknesses of the current representation of dynamicity

Currently, there are two different kinds of representations for brain dynamic function: states and components. Indeed, CAPs are actually spatial and/or temporal components of a certain type of states (see Section 2.4.4). There are also different kinds of states defined according to different properties (temporal independence, connectivity patterns. etc.).

Definitions of Regions of Interest

Prior to the detection of functional connectivity, the nodes of the networks need to be defined. Different methods are described in the previous sections to establish a spatial parcellation and they most of the time rely on static networks. For dynamic studies, spatial ICA is the dominating method [HWA$^+$13] and components are identified from the whole fMRI duration and may not recover the entire brain surface. It is therefore interesting to eliminate this spatial constraint based on static information. Moreover, using ICA means that the networks are made of independent components and it is particularly inappropriate to study interactions between them [LLT$^+$16].

Another problem raised by the use of a supervised parcellation is the assumption that ROIs are reproducible across subjects while they can be subject specific [LLT$^+$16]. This observation underlines the problem of the level of analysis: the link between individual and population is not fully established by neuroscience. The dynamic networks are first studied at the subject level and then translated to the population level to test their robustness and obtain exploitable results. However, a high quantity of information is neglected in the process [LLT$^+$16].

Spatial Overlap of the networks

One question that has also to be addressed when the dynamic functional networks are identified is how to deal with networks' spatial overlap [SMM$^+$12]. Indeed, one node can be correlated at time t with other nodes of a network A, while at time $t + 1$ it is correlated with nodes from network B. This spatial overlap asks the question of whether the correlation variations inside a network is due to nodes shared between several overlapping networks and that change their membership, or due to internal connections being non-stationary. Both of these possible explanations of connectivity changes are neurobiologically relevant but it is currently not possible to determine which one is dominant in the dynamic behavior of resting-state dynamic functional networks [SVB$^+$13].

Use of the available data

Defining states and components require assumptions. Indeed, they are defined through their independence or their temporal reproducibility (see Section 2.4) while static networks are only based on connectivity measure, such as correlation. Furthermore, the construction of the states and components is often based on a finite number of time points or spatial components where the connectivity is particularly strong or dynamic ([ADP$^+$14], [LD13], [ZFC$^+$14], [KVDV15]). The information put aside is indeed more likely to result from noise but it can yet hold a high quantity of information. Therefore, these approaches are too restrictive.

2.6 Contribution of this thesis

Several of the problems exposed in the previous part are addressed by the detection method developed in this thesis.

Concerning noise, our approach applies the conventional preprocessing steps in order to remove the maximum of noise before identifying functional networks. At different steps of the workflow, constraints are applied to prevent the noisy networks to be detected, by for instance requiring a minimum number of nodes or a minimum time of activity. The transition from the subject level to the population level acts also as a filter by requiring a high number of recurrence among subjects since a network corresponding to noise is less likely to be reproducible across subjects. Moreover, several validation steps are performed to assess the effect of the remaining noise.

The parcellation used in this thesis is a geometric parcellation. It is not based on ROIs and therefore, no assumption on the subject variability is made. A geometrical parcellation prevents the spatial overlap between the nodes of the networks but it has the advantage to take into account the whole brain surface. Moreover, as they are not supervised, they do not require assumption based on static information or on respective independency as for ICA. It is therefore more adapted to the detection of interaction [LLT+16]. No prior assumptions are made and the approach described in the following paragraphs only use the correlation information and identify dynamic networks very closed from the static ICNs. The approach used here is very similar to a repetition of the method described in [YKS+11] for the identification of static ICNs inside successive windows selecting a small number of time points. Furthermore, no time point of interest are selected and the functional networks are constructed with the complete time course of the experiment. The quantity of information taking into account for the detection of dynamic RSNs is consequently increased but the cost is that a higher amount of noise is including this information.

The problem of the spatial overlap between networks is addressed by allowing a node to change its membership. Current representations cannot deal with the spatial overlap this way since they are not building real dynamic networks with node membership. In our case, time-varying spatial maps is the solution to assess changes of node membership without influencing the detection of within networks fluctuations. Concerning the interactions between networks, exchange of nodes would reveals a spatial and temporal interaction. Currently, the interactions between states and components are limited transition probabilities for instance [ADP+14].

The representation we propose here is similar to the concept of component representation since it identifies individual networks that resemble static ICNs and allows their co-occurrence as well as their spatial overlap over time. However, it is more flexible and requires the fewest assumptions on what should the dynamic networks look like without noise. By taking into account all the available data, it is more influenced by noise but it also accesses to more information and built easy to handle networks that creates new outlooks, such as new biomarkers for neuropathologies (see Section 1.2).

2.7 Summary

The Functional Connectivity (FC) is mainly detected via fMRI. This imaging modality detects variation of blood flow oxygenation near activated cells. Several metrics, such as the Pearson's correlation coefficient, are available to extract this connectivity from a set of signals attributed to different locations on the brain surface. Each location represents a node. They are thus pairwise related and it is possible to represent the functional connectivity information as a graph.

This graph can be static or a temporal dimension can also be introduced. In that case, the dynamic FC is analyzed. There are different possible representations of the dynamic FC but they rely on static assumption, such as a parcellation based on static independence, and they are not considering all the available data, restricting either the spatial or the temporal dimension. At the end of the day, the dynamic characteristics of the FC are represented by states or components of connectivity and their interactions are not considered.

General Methods for the construction of Dynamic Communities (DCs)

The previous chapter gives an overview of what has already been done for the analysis of the functional connectivity. In this thesis, a new approach is developed and before explaining it in more details, it is necessary to present few techniques used for its construction. This chapter thus describes a method for the community detection as well as the principles of spectral clustering, two techniques used in Chapter 5.

3.1 Community Detection based on Stability

A community detection aims to separate the data from a graph into relevant groups. It differs from clustering because existing communities are unknown. They can have totally unequal size and density, and their number is also unknown. Moreover, they often have a hierarchy [For10]. Communities are identified by detecting groups of nodes with a higher connection density internally than with the rest of nodes. It thus relies on connection inhomogeneities to partition the network.

According to this definition of community, an obvious optimization criterion for the community detection is the *modularity*. It is defined as the sum of the difference for each community between the actual fraction of links inside this community and the fraction of links inside this same community if links would be randomly placed. However, the modularity optimization is not adapted to handle the hierarchical properties of networks because it has a resolution limit that lead to miss the detection of small communities [FB07]. It is thus necessary to define a quality criterion adapted to multi-resolution

networks and small communities. Le Martelot and Hankin [LMH13] demonstrates that stability can be one of them .

3.1.1   Concept of Stability Measure

One way to conceptualize the stability approach is to visualize edges of the graph as vessels filled of water and whose width depends on their weight. If a droplet of ink is put on a node, the ink will diffuse through the vessels and we can observe how it diffuses over time, as in Figure 3.1.
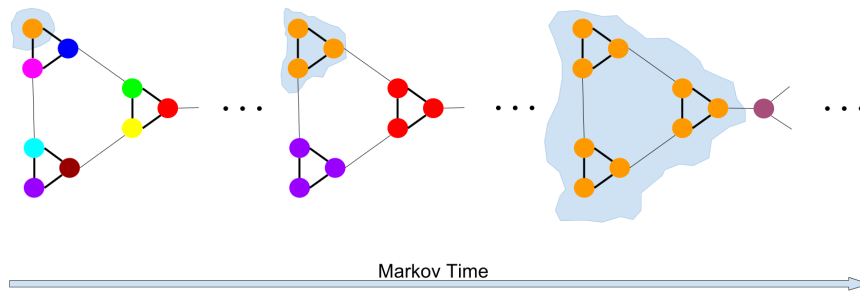

Figure 3.1: Diffusion dynamic on the graph at different time steps. Nodes with different colors mean they are not members of the same community. The blue stain represents the diffusion at the represented time step. *adapted from: `http://wwwf.imperial.ac.uk/~mpbara/Partition_Stability/`*

The droplet of ink can be deposed at each node and its diffusion can be observed for each of them. What we want to see is if the ink is trapped in a particular region during an abnormally long time. Such an observation would reveal the existence of communities at that particular time. The advantage of this method here is that we can continue to observe the diffusion behavior and remark several time period where the partition realized by the ink is stable in time. It gives the different time scales where different relevant communities appears in the graph, and the more we wait, the coarser the communities are.

3.1.2   Stability definition

Stability definition considers the whole network as a Markov Chain [LMH13]. It means that each node a_i is a state and the weight of each edge between node a_i and node a_j represents the transition probability between a_i and a_j. With n nodes and m edges (all the nodes are not linked so all the transitions are not possible), an adjacency matrix $\mathbf{A} = (A_{ij})$ of size $n \times n$ can be built with A_{ij} equal to the weight between node i and j. In our case, the graph is undirected but this definition acknowledges also directed graphs [LMH13]. A vector $\mathbf{d}$ of size n is also defined to store the degree of each node, that is to say the number of edges connected to it, and $\mathbf{D} = diag(\mathbf{d})$. The chain distribution

is the stationary distribution $\pi = \frac{\mathbf{d}}{2m}$ and we define the diagonal matrix $\mathbf{\Pi} = diag(\pi)$. A stationary distribution means that we are assuming that in the long run, all states are equally occupied [DYB10]. It is also named a random walk and thus, we can say that the probability of leaving a state at the next Markov time is equally distributed between outgoing edges. It is thus possible to define the probability vector $\mathbf{p_t}$ at Markov time t_M such as [LMH13]:

$$\mathbf{p}_{t_M+1} = \mathbf{p}_{t_M}\mathbf{D}^{-1}\mathbf{A} \equiv \mathbf{p}_{t_M}\mathbf{M} \tag{3.1}$$

and $\mathbf{M} = \mathbf{D}^{-1}\mathbf{A}$ is the transition matrix of size $n \times n$. By considering a community partition into c communities, we also define the $n \times c$ 0-1 matrix $\mathbf{H} = (H_{ij})$ providing for each node its community. If $H_{ij} = 1$, then the node i belongs to community j. If a label α_i is attributed to each node, the process results in a random variable $(\mathbf{X}_t)_{t \in \mathbb{N}}$ which consists of a sequence of α_i. If the community partition defined by $\mathbf{H}$ is relevant over a given time scale, then transitions between nodes of a same community are more likely than transitions between nodes of different communities for such a time span [DYB10]. The quantification of this kind of phenomenon is realized via the autocovariance of the observable defined as [DYB10]:

$$\mathbf{cov}[\mathbf{X}_{t_M}, \mathbf{X}_{t_M+\tau}] = \mathbf{E}[\mathbf{X}_{t_M}\mathbf{X}_{t_M+\tau}]\mathbf{E}[\mathbf{X}_{t_M}]^2 \tag{3.2}$$

where $\mathbf{E}$ is the expectation. The less transitions between different communities happen, the longer are the periods where $\mathbf{X}_{t_M}$ and $\mathbf{X}_{t_M+\tau}$ are correlated and the higher is the autocovariance. If the autocovariance is high, the community partition is relevant and it is stable. To study the stability as a function of time, the evolution of the autocovariance has to be established as a function of the lag τ [DYB10]. With the previously defined notations, the clustered autocovariance matrix at Markov time t_M of size $c \times c$ from the transition matrix at time t_M $\mathbf{M_{t_M}}$ is defined as [LMH13]:

$$\mathbf{R}_{t_M} = \mathbf{H}^T(\mathbf{\Pi}\,\mathbf{M}^{t_M} - \pi^T\pi)\mathbf{H} \tag{3.3}$$

$(\mathbf{R}_{t_M})_{ij}$ is equal to the probability of starting the Markov chain in a community i and arriving in a community j after t_M time steps, minus the probability that two random walkers are located in i and j in a stationary situation [DYB10]. Therefore, the answer of the question to what extent the Markov chain stays in the same community during t_M Markov times is located on the diagonal of $\mathbf{R}_{t_M}$. The stability $Q_S^{t_M}$ for a random path of length t_M is then equal to the trace of the auto-covariance matrix [DYB10]. As global stability definition, we have:

$$Q_S = \min_{0 \leq t_M \leq \tau_M} trace(\mathbf{R}_{t_M}) \tag{3.4}$$

where τ_M is the upper bond of Markov time that has been used.

The algorithm used in this thesis to perform the community detection based on this stability definition is described in Section 5.3.

3.2 Spectral Clustering

Another method of interest used in Chapter 5 is the *spectral clustering*. It uses the spectral information contained in the graph to represent the data in a space where the data points can be attributed to clusters by a classical clustering algorithm [VL07].

3.2.1 Graph definition

The first step of the spectral clustering is to build a graph. An appropriate similarity function needs to be defined and a common function is the Gaussian function [VL07]. In our case, we already have one, as explained in Section 5.5.2. A similarity matrix $\mathbf{S}$ is thus build and it is used to construct a similarity graph [VL07]. The larger the similarity is, the lower is the weight. It is now possible to formulate the clustering as the partition of the graph into subgraphs where the edges within the groups have low weights and the edges between the groups have high weights [JGGF16].

The similarity graph $\mathbf{G}$ is described as $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ where $\mathbf{V} = \{v_1, \ldots, v_n\}$ is the set of vertices and $\mathbf{E}$ are the undirected weighted edges which are constructed according to the selected type of graph. The weights of each possible edge are stored into $\mathbf{W} = (w_{ij})$. If two vertices i and j are not connected, $w_{ij} = 0$.

Transforming a set of vertices $\mathbf{V} = \{v_1, \ldots, v_n\}$ with associated pairwise similarities S_{ij} into a graph can be performed by several means [VL07]. In this thesis, we are just describing the one we choose to use, the undirected *k-Nearest Neighbors (kNN)* graph. It consists in connecting the vertex v_i only to its kNN, that is to say the k vertices which are the most similar to v_i. No direction is attributed to the edges, consequently an edge between v_i and v_j either means that v_j is one of the kNN of v_i, or that v_i is one of the kNN of v_j, or even that v_i and v_j are mutual kNN [VL07]. Then the weights w_{ij} attributed to these newly formed edges are inversely proportional to their similarity S_{ij}. The degree of a vertex v_i is defined as $d_i = \sum_{j=1}^{n} w_{ij})$ and the degree matrix $\mathbf{D}$ is a diagonal matrix with $D_{ii} = d_i$.

3.2.2 Graph Laplacian

Spectral clustering is based on the *graph Laplacian matrices*. There are different types of graph Laplacian and here we use the *normalized symmetric graph Laplacian* $\mathbf{L}_{sym}$ defined as [VL07] (see Section 8.3 for the explanations of the choices made here):

$$\mathbf{L}_{sym} := \mathbf{D}^{-\frac{1}{2}} \, \mathbf{L} \, \mathbf{D}^{\frac{1}{2}} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \, \mathbf{W} \, \mathbf{D}^{\frac{1}{2}} \tag{3.5}$$

where $\mathbf{L}$ is the unnormalized graph Laplacian:

$$\mathbf{L} := \mathbf{D} - \mathbf{W} \tag{3.6}$$

The important properties of $\mathbf{L}_{sym}$ are:

1. $\forall\ \mathbf{f} \in \mathbb{R}^n$,

$$\mathbf{f}^T \mathbf{L}_{sym} \mathbf{f} = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} \left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \right)^2 \tag{3.7}$$

2. λ is an eigenvalue of $\mathbf{L}_{sym}$ with eigenvector $\mathbf{D}^{\frac{1}{2}}\mathbf{u}$ only if λ and $\mathbf{u}$ solve the generalized eigenvalue problem : $\mathbf{Lu} = \lambda \mathbf{Du}$.

3. 0 is an eigenvalue of $\mathbf{L}_{sym}$ with eigenvectors $\mathbf{D}^{\frac{1}{2}}\mathbb{1}$, with $\mathbb{1}$ the constant one vector.

4. $\mathbf{L}_{sym}$ is positive semi-definite and have n non-negative real-valued eigenvalues $\lambda_1 \leq \ldots \leq \lambda_n$.

These properties are demonstrated in Luxburg Tutorial [VL07].

The key concept of spectral clustering is that the multiplicity of the Eigenvalue 0 of $\mathbf{L}_{sym}$ is equal to the number of connected components $A_1, \ldots, A_k$ in the similarity graph $\mathbf{G}$, and the Eigenspace of $\lambda = 0$ corresponds to $\mathbf{D}^{\frac{1}{2}}\mathbb{1}_A$ where $\mathbb{1}_A$ is the indicator vector defined as $\mathbb{1}_A = (f_1 \cdots f_n)^T \in \mathbb{R}^n$ *with* $f_i = 1$ *if* $v_i \in A$ *and* $f_i = 0$ *otherwise* [VL07]. A connected component A_i is defined as a group of vertices that are connected together. For more simplicity, we explain why this property is important via the unnormalized graph Laplacian $\mathbf{L}$ which can be linked then to $\mathbf{L}_{sym}$ [VL07]. If $\mathbf{f}$ is an Eigenvector of $\mathbf{L}$ corresponding to Eigenvalue $\lambda = 0$, then :

$$\mathbf{f}\mathbf{L}\mathbf{f}^T = 0 \Leftrightarrow \sum_{i,j=1}^{n} w_{ij} \left(f_i - f_j \right)^2 = 0 \tag{3.8}$$

As w_{ij} is positive or equal to 0, all the terms in the sum needs to be equal to 0. Therefore if there are a vertex i and a vertex j which are connected, then $w_{ij} \neq 0$ and $f_i = f_j$. If there is only one connected component, for all connected vertices (i, j), $f_i = f_j$ and the Eigenvector $\mathbf{f}$ is equal to $\mathbb{1}$. It means that all vertices with at least one connection are connected together with identical weights. If we consider the normalized graph Laplacian $\mathbf{L}_{sym}$, then all connected nodes are also connected together but with weights depending on their degree and the Eigenvector $\mathbf{f}_{sym}$ of $\mathbf{L}_{sym}$ respects $\mathbf{D}^{\frac{1}{2}}\mathbf{f}_{sym} = 1$ [VL07]. If the number of connected components k is different from 0, it is possible to write $\mathbf{W}$ as a block diagram if the vertices are ordered such as connected vertices are gathered. The resulting unnormalized graph Laplacian $\mathbf{L}$ has also a block shape and each block can be considered as an individual graph Laplacian. As a property of block matrices, the Eigenvalues of the entire matrix are also those from the block matrices, as well as the corresponding Eigenvectors filled with zeros at position of other block matrices. Therefore, as each block is a graph Laplacian of a connected graph, each one has 0 as Eigenvalue of multiplicity 1 with the corresponding Eigenvector whose components are equal to 1 for rows of the considered block and 0 otherwise. To conclude, the similarity graph $\mathbf{G}$ has as many connected components $A_1, \ldots, A_k$ as the multiplicity of Eigenvalue 0 for $\mathbf{L}$ or $\mathbf{L}_{sym}$ [VL07].

3.2.3 Rearrangement of the data

There are several algorithms for spectral clustering according to the graph Laplacian that is used. In our case, we use a normalized spectral clustering developed by Jordan and Weiss with the symmetric normalized graph Laplacian [NJW⁺02]. The similarity matrix $\mathbf{S}$ as well as the number of clusters to build k are provided as input. The first step is to build the similarity graph $\mathbf{G}$ which is here an undirected kNN graph. Then, the symmetric normalized Laplacian $\mathbf{L_{sym}}$ is calculated and its first k Eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u_k}$ corresponding to the k smallest Eigenvalues are computed. The rows of the matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_k]$ are normalized to norm 1 and result in the matrix $\mathbf{T} = (t_{ij})$ with:

$$t_{ij} = \frac{U_{ij}}{\sqrt{\sum_k U_{ik}^2}} \tag{3.9}$$

A k-mean clustering is finally applied to the rows of $\mathbf{T}$. Each row corresponds to one point while normalized Eigenvectors acts as variables. Therefore t_{ij} is equivalent to an estimation of the membership of point i in the connected component A_j. As output, clusters $C_1, \ldots, C_k$ are obtained [VL07].

3.2.4 k-means Clustering

The k-means clustering algorithm is here only briefly described in the way we use it since it is not the main point of the consolidation step (see Section 5.5.2). K-means clustering is also called Lloyd's algorithm and it is an iterative, data-partitioning algorithm. It aims to attribute each observation to only one of the k-clusters that are defined by a key observation, named centroid. The number of cluster k is an input of the algorithm. It is initialized by a random definition of k cluster centroids among the observation. According to these initial cluster centroids, the final partition changes. Then the distances between every observation and each cluster centroid are calculated. For every observation, a vector contains its associated values for a set of variables. Every distance metric can be used, such as squared Euclidean distance, cosine or the sum of absolute difference, but only one must be selected. In the case of a Batch update, each observations is then attributed to the cluster with the closest centroid. Once the initial clusters formed, the averaged of the points in each cluster is computed to produce k new cluster centroid locations. The distances between observations and the centroids are once again calculated and the observations assigned to cluster. The operation is repeated until the cluster assignment becomes stable [FHT01]. However, one constraint of k-means clustering is the choice of k and there are several methods to select the most appropriate [VL07].

3.2.5 Selected metrics for the choice of the number of clusters k

Silhouette Coefficient

The *Silhouette coefficient* permits to directly estimate the accuracy of a partition by combining a measure of cohesion and a measure of separation [Rou87]. The cohesion a of

a partition corresponds to the sum of the weights of all the edges between the points of the same cluster. On the other hand, the cluster separation b is the sum of the weights of all the edges linking a point from one cluster to another cluster. The Silhouette Coefficient s is calculated vertex per vertex and for a vertex x into a cluster $\mathcal{D}_\alpha^{*l}$, it is equal to:

$$s(x) = \begin{cases} 1 - \frac{a(x)}{b(x)} \text{ if } a(x) < b(x) \\ \frac{b(x)}{a(x)} - 1 \text{ otherwise} \end{cases} \tag{3.10}$$

with:

$$a(x) = \sum_{y \in \mathcal{D}_\alpha^{*l}} w_{xy} \tag{3.11}$$

and:

$$b(x) = \min_{\mathcal{D}_\alpha^{*l'} \in \mathcal{D}_\alpha^*} \sum_{y \in \mathcal{D}_\alpha^{*l'}} w_{xy} \tag{3.12}$$

The Silhouette coefficient for an entire cluster is equal to the averaged silhouette coefficient of the point of the cluster [Rou87]. It varies between -1 and 1. The aim of the optimization is to obtain a cluster coefficient as close as possible from 1 since it means that the average of $\frac{a}{b}$ is low and therefore that the cohesion a tends to be low and the separation b tends to be high. A low value of a means high similarities between the points of a same cluster so the cohesion is high, while a high value of b means that the two most similar clusters have yet a low similarity, so they are well separated. For our optimization procedure, we calculate the average cluster Silhouette coefficient for values of cluster number in a range of [7,13] (see Section 7.4.3 for justification). A fixed number of clusters for all subjects may not be extracted since all subjects may not activate the same RSNs during the experiment. Therefore, the optimization process is run independently for each subject and reiterated 50 times per number of clusters to obtain a consistent number of clusters. The results of the optimization procedure are described in Section 7.4.3.

Eigengap Heuristic

Luxburg et al. [VL07] recommend to use *Eigengap heuristic* to determine k. It is indeed a natural approach since spectral clustering is based on Eigenvalues from the graph Laplacian and it has been demonstrated above that the multiplicity of the Eigenvalue 0 is equal to the number of connected components in the graph. In real dataset, the number of completely disconnected components is lower than k, but the k lower Eigenvalues are still close to zero if the graph is structured in k possible clusters. Therefore, to optimize the number of clusters k, the Eigengap heuristic approach consists in choosing k such as the ordered Eigenvalues $\lambda_1, \ldots, \lambda_k$ are very small and λ_{k+1} is significantly larger [VL07]. The largest Eigengap designates the transition between λ_k and λ_{k+1} and thus the optimal k is identified.

3.3 Summary

In this chapter, the general methods and tools utilized for the accurate detection of the dynamic RSNs are explained. First, a community detection based on the concept of stability has been depicted. This method enables to perform a partition of the space in identifying more or less homogeneous groups of nodes. Second, the spectral clustering and the way it reorganizes the space to enhance the separation between clusters is explained.

All the aspects of these methods are not dealt with because they are described according to the choices made for their implementation in Chapter 5. For more details concerning these methods, the literature cited in this chapter provides the basics of each techniques. Moreover, the Section 8.3 justifies the choices made for the spectral clustering.

<div align="right">

CHAPTER 4

</div>

Data

The dataset that forms the basis of this thesis consists of 200 subjects. It corresponds to rest-fMRI scans provided by the Human Connectome Project (HCP) [VEUA$^+$12]. It is essential to know the parameters that have been utilized to record them in order to choose appropriate parameters for the methods implemented then. Moreover, preprocessing is required to eliminate as much noise as possible.

4.1 Dataset: Human Connectome Project (HCP)

The HCP responds to a request from the American National Institute of Health (NIH) emitted in 2009 and has been launched by two consortia in 2010, one made of Washington University, University of Minnesota and Oxford University ("WU-Minn" HCP Consortium), and the other one created by Massachusetts General Hospital and University of California Los Angeles. The aim of this project is to "characterize human brain connectivity and function in a population of 1200 healthy adults and to enable detailed comparisons between brain circuits, behavior, and genetics at the level of individual subjects" [VEUA$^+$12]. The second consortium is interested in leading-edge structural connectomics measured by diffusion MRI, while the WU-Minn HCP Consortium generates and gives free access to both functional and structural brain connectivity data from a very large cohort of more than 1000 subjects. Concerning rest-fMRI, approximately 300 subjects are available for the moment and 200 subjects are selected as basis for this thesis, since it is a sufficiently high number to ensure the statistical significance of the analysis ([ZFC$^+$14], [KVDV15] & [LLT$^+$16]).

Subject cohort

HCP data are acquired on adults twins and their non-twin siblings. Studying twins and their siblings enables to take into account the genetic factor and the influence of

environment [SVB$^+$13]. However this aspect is not considered in this thesis and we use 200 random subjects. Moreover, subjects are between 22 and 35 years old.

Experiments

For each subject, two pairs of 15-min sequences of whole-brain resting-state fMRI are acquired with a $3T$ scanner at different days. The spatial resolution is $2 \times 2 \times 2\ mm^3$ and a temporal resolution of 0.7 seconds is obtained owing to an EPI acceleration factor of 8 (see Section 2.1). Moreover, scans at higher frequency and therefore with a higher temporal resolution are available for subsets of the cohort, as well MEG data [SVB$^+$13]. In addition to resting-state fMRI sequences, task fMRI data are also provided by the HCP. They are not used in this thesis since the aim is to detect dynamic RSNs so the paradigms are not described here.

4.2 Preprocessing

MRI spatial distortions and signal loss are corrected by the HCP consortium and the alignment between subjects is also performed at high resolution ($0.7mm$) [SVB$^+$13]. It also provides a gray matter surface mesh obtained from the original 3D fMRI matrices. The surface mesh is used here because the information concerning FC is located in gray matter [HWA$^+$13] and it thus permits to reduce the amount of data to be stored by conserving only the data of interest into a simpler representation [VEUA$^+$12]&[SVB$^+$13].

Three main operations are still necessary before processing the signals :

- applying movement regressors,

- filtering the signals to only obtain the frequency band of interest,

- parcellating the brain surface in order to decrease the computational cost.

4.2.1 Head Motion Correction

Head motion corrections must always be applied to fMRI signals [PMN11]. Indeed, all the subjects irremediably move their head, due to swallowing for instance, and head movements have an overwhelming influence on fMRI data [PMN11]. Two main effects of head motion can be distinguished. On the one hand, the displacement of the head can result in a mislocation of the recorded time series on the brain surface. As the whole brain is involved in this effect, it is also referred as bulk motion [PMN11]. The basic motion correction techniques uses a single reference image to realign the time series. This effect is particularly visible at the edges of the brain where the motion has the biggest influence since the concerned pixels can contain brain tissue at time t and then no more brain tissue at time $t + 1$. On the other hand, the MRI signal itself can be affected by head movement and this effect is called the spin history effect. It is produced by the displacement of protons from one slice to another. Indeed, the MRI scanner selectively

excites slices for the 3D reconstruction of the signal and the response of the protons is analyzed according to their excitation (see Section 2.1). Therefore, if the proton is no more in its slice of excitation, it corrupts the reconstruction of the signal. At the end of the day, it can result in large alteration of the intensity in a single slice or in a finite set of slices [PMN11]. Standard motion corrections do not handle this effect but exploratory methods such as ICA can be used to leave out the components of the signals corresponding to this effect into additional components [PMN11].

The aim of the head motion correction is to realign the fMRI images at different time series. It is thus necessary to estimate the motion that has caused the misalignment of these images and it is done owing to movement regressors [PMN11]. It assumes that head motion can be approximated by a rigid body transformation: it can be translated or rotated along three axis but it can not be deformed. At least 6 regressors are needed, one for each translation and one for each rotation, but for a better accuracy, their first derivatives are also used [PMN11].

The set of recorded signals is considered as a matrix $\mathbf{Y}$ that can be defined such as :

$$\mathbf{Y} = \mathbf{g}\ \mathbf{b} + \boldsymbol{bold} \tag{4.1}$$

with $\mathbf{g}$ the matrix of the movement regressors, $\mathbf{b}$ the part of the signals linked to head movement and $\boldsymbol{bold}$ the part of the signal holding the functional information that has to be extracted and which is independent from the motion. It is a Generalised Linear Model (GLM). It enables to regress noise by estimating $\mathbf{b}$ by $\hat{\mathbf{b}}$ via an ordinary least squares estimation :

$$\hat{\mathbf{b}} = (\mathbf{g}^T\mathbf{g})^{-1}\ \mathbf{g}^T\ \mathbf{Y} \tag{4.2}$$

Then the BOLD signal component $\boldsymbol{bold}$ corresponding to the residual in this GLM is easily approximated by :

$$\boldsymbol{bold} = \mathbf{Y} - \mathbf{g}\hat{\mathbf{b}} \tag{4.3}$$

The head motion is thus corrected by processing only $\mathbf{bold}$.

4.2.2 Band-pass Filtering

One source of noise that has not been preprocessed yet is the "physiological noise" [HWA+13]. This noise can be produced by cardiac pulsation and respiration that provokes shifts in the main magnetic field $\mathbf{B_0}$ due to body motion. Moreover, the changes of respiratory and cardiac rate have also an influence on BOLD contrast and it is particularly important to eliminate it for rest fMRI analysis because they are located in the low frequency domain we are interested in ([BMB08], [CG09], [DIH99], [SvGdZ+07]). Indeed, noise linked to cardiac activity occurs around 0.1-0.5 Hz while respiratory activity produces noise between 0.6 and 1.2 Hz [BS09].

Moreover, the frequency band of interest selected for functional connectivity analysis in resting-state is usually [0.01,0.1]Hz [SVB+13]. The inferior cutoff frequency permits to remove the slowest temporal drifts in the data [SVB+13] while the superior one select
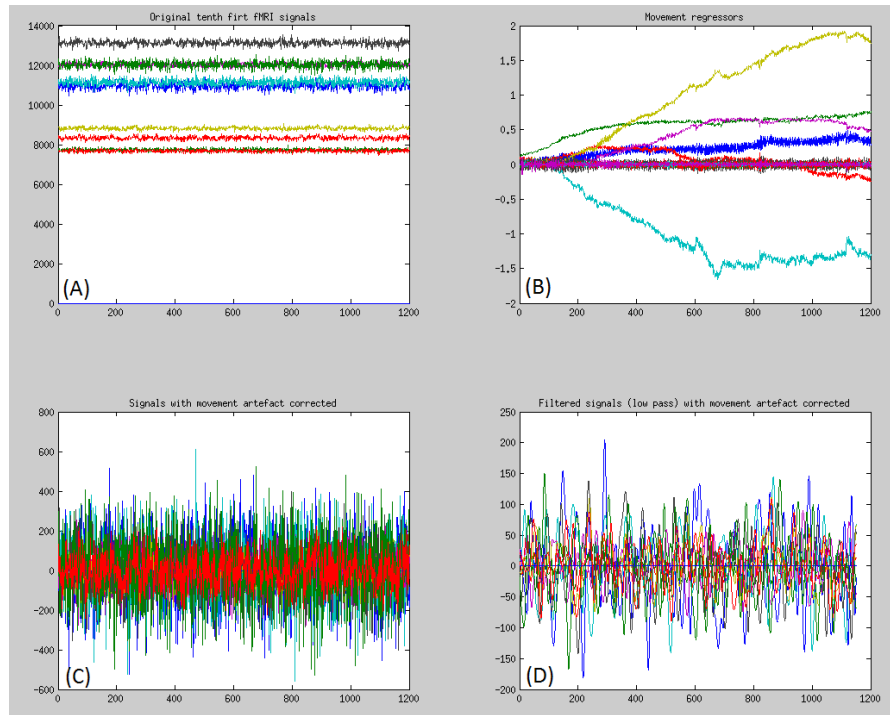
Figure 4.1: (A) Set of 10 initial non corrected fMRI signals from 10 different time points, HCP data. (B) Movement regressors used for the head motion correction. (C) Set of the same 10 fMRI signals corrected for head motion. (D) Set of the same 10 fMRI signals corrected for head motion and after band-pass filtering.

only low frequency domain since, for instance, correlation between auditory, visual and somatomotor regions has been detected in this domain [CSB10]&[VdHHP10]. Thus, we apply a band-pass filter to select only this frequency domain because it allows to select the information of interest and to eliminate most of physiological noise.

Figure 4.2: Brain surface parcellated into 800 regions, 400 on each hemisphere.

4.2.3 Parcellation

The HCP data provides 64984 spatial points on the brain surface corresponding to voxel of gray matter in the MRI scan. However, for our analysis, this number is too high and would requ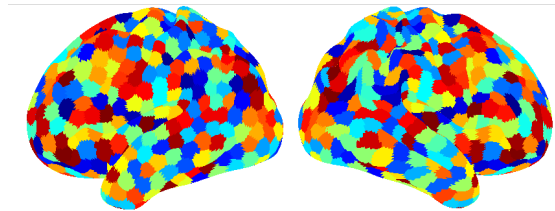ire a too expensive computational cost. Therefore, the brain surface is reorganized into 800 parcels [BJG⁺13] where the fMRI signals from each voxel included in one parcel are averaged to provide only one signal per parcel and simplify the analysis. It also permits to improve the Signal-to-Noise ratio. On the other hand, it decreases the spatial resolution, one of the advantage of the fMRI compared to EEG for example, but it remains relatively high. Figure 4.2 illustrates the parcellation realized on the brain surface.

4.3 Summary

The approach developed in this thesis is tested on 200 subjects from the HCP that aims to map the brain structural and functional connectivity [VEUA⁺12]. They are already partly preprocessed but additional corrections are necessary. The head motion is corrected via a GLM and a specific frequency range is selected. Moreover, the brain cortex is parcellated to decrease the computational cost and it increases also the Signal-to-Noise Ratio since the resulting signals, input of our workflow, are the averaged signals of the voxels included in this parcel.

Methodology

The methods defined in this chapter have been used to create a robust representation of dynamic functional networks via Dynamic Communities (DCs). They have been selected and designed to fit the constraints required by the nature of the data and the aims defined in the first chapter. Most of them are based on graph theory techniques and their implementation are partly inspired by their use in social networks.

5.1 Notation and Overview

The approach developed in this thesis is different from those explained in Section 2.4 and is summarized in Figure 5.1. The dynamic RSNs are considered as communities and as they can change over time, they are called *Dynamic Communities (DCs)*.

Before preprocessing, the brain surface from one subject α initially has $N_0 = 64984$ nodes and their corresponding signals made of T samples are represented here by $\mathbf{S}_\alpha$. It is then meshed with $N = 800$ nodes and the communities are made of these nodes. Each of the N nodes has a signal contained into $\mathcal{S}_\alpha^*$, the set of preprocessed signals. Before extending the analysis to the whole population containing $N_s = 200$ subjects, it is necessary to process each subject individually.

The first part of the approach is the detection of communities in each time point at subject level. A Sliding-Window Analysis (SWA) first acquires the correlation between signals from $\mathcal{S}_\alpha^*$ inside a window of w time points shifted time point per time point. The correlation information is thus extracted for $T - w$ time points and stored into $\mathcal{C}_\alpha = \left\{ \mathbf{C}_\alpha^1 \dots \mathbf{C}_\alpha^{T-w} \right\}$. Each $\mathbf{C}_\alpha^t$ is a correlation matrix for the N nodes. From each $\mathbf{C}_\alpha^t$, a community detection is then performed to obtain a set of static communities inside each window t represented by the vector $\mathbf{C}_\alpha^{*t} = (C_{i,\alpha}^{*t})$. $C_{i,\alpha}^{*t}$ is the label of the community to which node i is affected. For the subject α, the entire set of communities from all the time points is accessible via $\mathcal{C}_\alpha^* = \left\{ \mathbf{C}_\alpha^{*1} \dots \mathbf{C}_\alpha^{*T-w} \right\}$.

With $\mathcal{C}_\alpha^*$, it is possible to construct the Subject-level Dynamic Communities (sDCs). A *community tracking* is performed to link the communities from $\mathbf{C}_\alpha^{*t}$ with the communities from $\mathbf{C}_\alpha^{*t+1}$ $\forall t \in [1, T - w - 1]$. For the subject α, n_α sDCs are first detected and stored into $\mathcal{D}_\alpha$. The matrix $\mathbf{D}_\alpha = (D_{itj,\alpha})$ represented on Figure 5.1 is the front-tracer matrix whose coefficient $D_{itj,\alpha}$ is equal to 1 if the node i is in sDC j at time t and 0 if not. $\mathcal{D}_\alpha$ also contains properties of interest of the sDCs. The n_α^* post-processed sDCs and their properties are stored into $\mathcal{D}_\alpha^*$.

Then, when the sDCs are available for each subject, the Population-level Dynamic Communities (pDCs)) can be constructed by identifying the recurrent sDCs in the population. The pDCs are represented by $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_n\}$ with $\mathcal{P}_m = \left\{ \mathcal{D}_{\alpha_1}^{*\ l_1}, \ldots, \mathcal{D}_{\alpha_{nm}}^{*\ l_{n_m}} \right\}, \forall\ m \in [1, n]$.
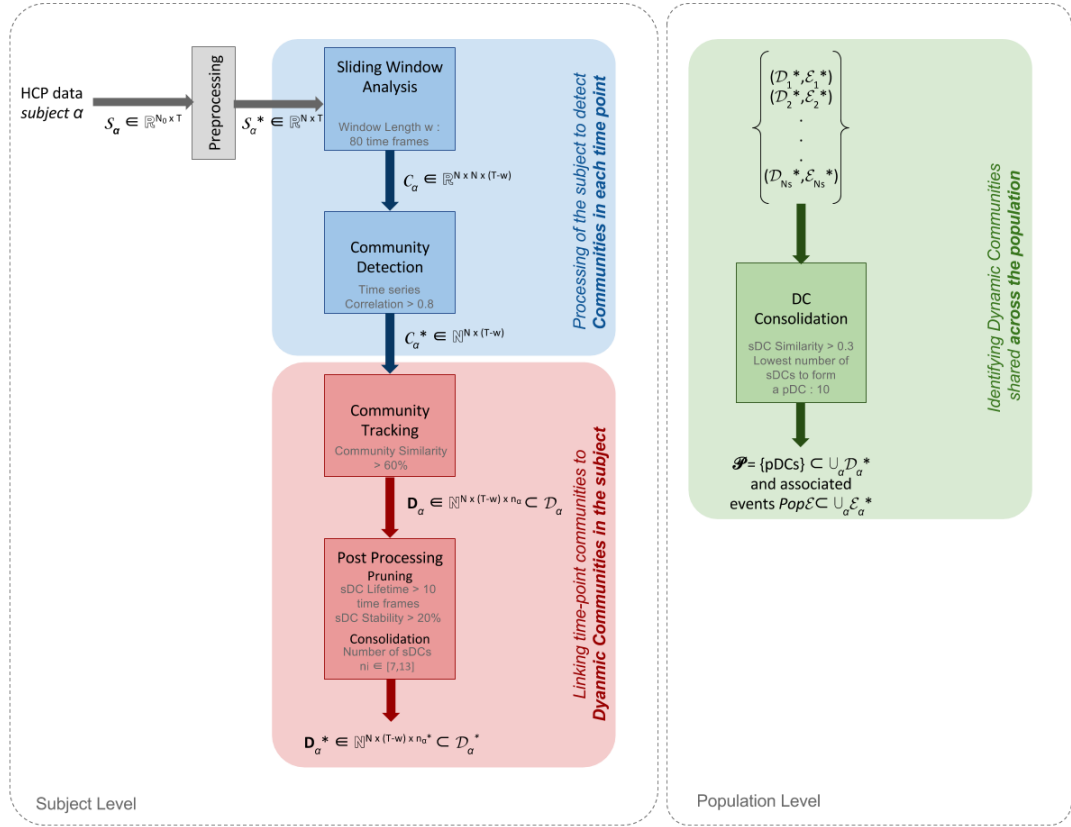

Figure 5.1: General Work-flow. For more details on each step, see the corresponding paragraph. The parameters indicated in the blocks are key parameters that illustrate the compromises made to reach accurate sDCs and pDCs with their associated events.

5.2 Sliding Window Analysis (SWA)

The Sliding Window Analysis (SWA) is applied on the preprocessed signal from individual subjects. In the previous section, the set of N signals for a subject α was represented by $\mathcal{S}_\alpha^*$. Each signal has T time points. As we have seen in Section 2.4.1, SWA consists in sliding a window of fixed length w and in calculating the correlation at a time t between each pair of signals owing to the w time points contained inside the window. Here $w = 80$ is selected. As the TR of the HCP data is 0.7 seconds [VEUA+12], it corresponds to a duration of $80 \times 0.7 = 56$ seconds and we are therefore in the range described in Section 2.4. This window is a Gaussian tapered window constructed by the convolution of a rectangular window with a width $\frac{w}{2}$ and Gaussian kernel made of $\frac{w}{2}$ points with a standard deviation equal to $\frac{1}{\frac{3}{11}w}$ [ADP+14]. This window is represented on Figure 5.2.
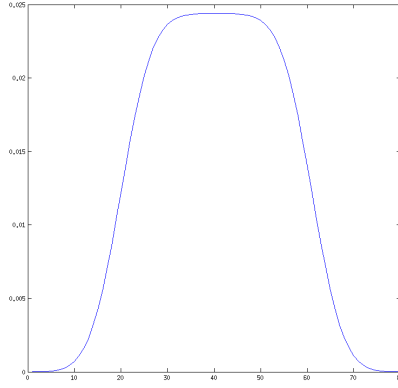

Figure 5.2: Shape of the window used for the SWA.

The correlation is measured through the tapered Pearson's correlation coefficient r (see Equation 2.12) calculated between all pairs of signals structured by the window.

Output

As output of the SWA, we obtain a symmetric correlation matrix between all the parcels and at each time $t \in [1, T - w]$ and with the notation defined in Section 5.1, it is represented by $\mathcal{C}_\alpha = \left\{ \mathbf{C}_\alpha^1 \dots \mathbf{C}_\alpha^{T-w} \right\}$. It permits to build a graph for each window, that is to say $T - w$ graphs, as explained in the following section. Figure 5.3 shows the input and the output of this SWA with the number of parcels N equal to 6.

5.3 Community Detection at each time point

At each time t, a correlation matrix $\mathbf{C}_\alpha^t$ contains the Pearson's correlation coefficient for each couple of signals. From this correlation matrix, an adjacency matrix $\mathbf{A}_t$ can be calculated to construct an undirected graph $\mathbf{G}_t = (\mathbf{V}, \mathbf{E}_t)$. The vertices $\mathbf{V}$ are parcels on brain surface and edges $\mathbf{E}_t$ connects nodes whose correlation at time t is positive and
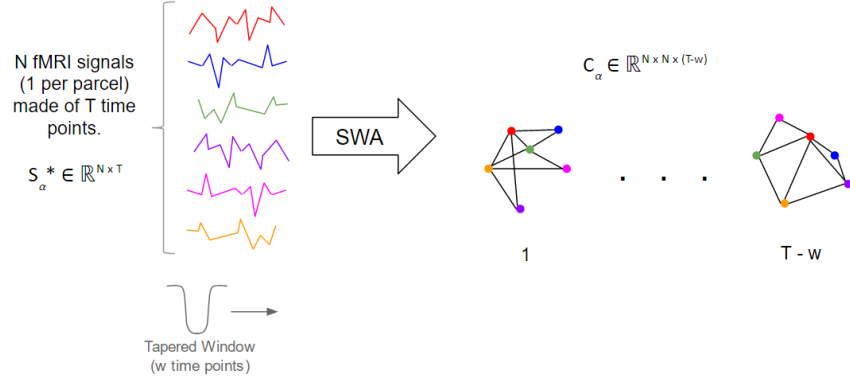
Figure 5.3: Input and output of the SWA. To simplify the visualization, the number of parcels N considered here is equal to 6. T is the number of time points in the fMRI signals and w is the window length. The window is shifted time point per time point and is a Gaussian tapered window. S_α^* is the matrix where the pre-processed fMRI signals are stored and C_α is the three dimensional matrices containing the correlation matrices for each window. The construction of the graphs represented here in an arbitrary way is based on these correlation matrices.

above a fixed threshold equal to 0.8. Edges are weighted according to the correlation between the two nodes they connect: the more they are correlated, the closer they are.

5.3.1 Motivations

For each of this graph, a community detection based on stability is performed as defined in Section 3.1. It is convenient to use community detection rather than a classical clustering algorithm because it can access hierarchical information, but above all it requires no assumption on the separation of the data space [LMH13]. They can have totally different size and density and there is no need to specify a number of communities to detect [For10].

The resulting communities are the active RSNs at time t, considered as static inside the window centered on t. The only constraint is a high stability within the graph. They may have totally different properties and size and this structure is therefore less influenced by the noise since it can be integrated to a community without changing the overall separation structure.

5.3.2 Implementation via an Optimization algorithm

Le Martelot and Hankin [LMH13] have developed a simple greedy algorithm for community detection based on stability optimization. Each parcel of the brain is one node and the edges between nodes are weighted via the correlation coefficient of the fMRI signals from this two nodes. The community detection algorithm is applied independently at each time window. Only one Markov time is used because it is not necessary to obtain multi-scale

communities and it is preferable to identify finer communities at this moment of the procedure since further steps are then planed to consolidate them if necessary. Although the multi-resolution properties of the stability optimization is finally not used, the choice of this method instead of the modularity optimization is still relevant because it provides higher resolution communities [FB07] and as future work, it could be interesting to go further in the analysis of the networks to detect larger communities. The algorithm presented is a simplification of the one from Le Martelot and Hankin [LMH13] since it uses only one Markov time. The complete version is available in [LMH13].

The initialization step consists in partitioning the graph $\mathbf{G}_t$ such as each initial community contains one node. The partition is represented by $\mathbf{H}$ and it can be visualized as a set of vectors $\mathbf{H}_i$ of size N, the number of nodes. Each vector $\mathbf{H}_i$ defines a community at the considered time t by attributing 0 or 1 to a node and $\sum_i \mathbf{H} = \mathbb{1}$.

The global stability Q_S of partition $\mathbf{H}$ is calculated and stored as the best known stability. Then, we try for each single node n to merge its community with each neighboring community and calculate the stability $Q_{s,temp}$. If at least one merge of its community with another community permits to increase the stability, the merge producing the highest stability Q_{S*} is acted and we look at the next node. If not, nothing happens and we look at the next node as well. When all nodes have been inspected and if at least one merge has occurred, we pass to the community level and calculate for each single community $\mathbf{C_i}$ the stability if it is merged to a neighboring community $\mathbf{C_j}$. With the same principle used at node level, if one combination with this single community increases the stability, it is realized and we inspect the next community. If not, nothing happens and we examine the next communities. When all communities have been analyzed, if no community merge has been made, the detection ends. If at least one community merge happens, the procedure starts again at the node level and is repeated until merge does not happen anymore. This algorithm is explained by the pseudo code 5.1. The output of the algorithm, $\mathbf{H}$, is the final partition of the graph at time t.

Output

Finally, a set of communities is obtained for each window and a community i can be represented by a footprint $\mathbf{H}_i$ on the brain surface since the nodes of the communities are parcels of brain surface. Only the communities with more than 40 nodes are conserved and the other are rejected as noise. All the nodes are finally not attributed to a community. For each time t, a set of labels is attributed to each community and a vector $\mathbf{C}_\alpha^{*t}$ of size N permits to read to which community the node is part of. The community membership for each node and for each time window is stored into $\mathcal{C}_\alpha^* \in \mathbb{N}^{N \times (T-w)}$ for subject α. Figure 5.11 permits to visualize the input and the output of the community detection. It is now possible to construct DCs.

Algorithm 5.1: Community Detection via Stability Optimization, adapted from [LMH13]

Input: Graph $\mathbf{G}_t = (\mathbf{V}, \mathbf{E}_t)$
Output: Partition $\mathbf{H}$

1 Initialize the community partition $\mathbf{H}$ by creating 1 different community per node of $\mathbf{G}_t$;

2 Calculate the stability Q_S of community partition $\mathbf{H}$;

3 while *at least 1 community merge with another community* **do**

4   **while** *at least 1 node change its community membership* **do**

5   **for** *all nodes n in the graph* **do**

6 $Q_S^* \leftarrow -\infty$;

7   **for** *all community neighbors $\mathbf{H}_j$ of the community $\mathbf{H}_i \mid n \in \mathbf{H}$* **do**

8 $\mathbf{H}^{temp} \leftarrow \mathbf{H}$;

9   Merge $\mathbf{H}_j$ and $\mathbf{H}_i$ in $\mathbf{H}^{temp}$;

10   Calculate the stability Q_S^{temp} of $\mathbf{H}^{temp}$;

11   **if** $Q_S^{temp} > Q_S^*$ **then**

12 $Q_S^* \leftarrow Q_S^{temp}$;

13 $\mathbf{H}_* \leftarrow \mathbf{H}_j$;

14   **end**

15   **if** $Q_S^* > Q_S$ **then**

16   Merge $\mathbf{H}_*$ and $\mathbf{H}_i$ in $\mathbf{H}$;

17   Calculate the stability Q_S of community partition $\mathbf{H}$;

18   **end**

19   **end**

20   **end**

21   **end**

22   **for** *all communities $\mathbf{H}_i$ of $\mathbf{H}$* **do**

23 $Q_S^* \leftarrow -\infty$;

24   **for** *all neighbouring communities $\mathbf{H}_j$ of $\mathbf{H}_i$* **do**

25 $\mathbf{H}^{temp} \leftarrow \mathbf{H}$;

26   Merge $\mathbf{H}_i$ and $\mathbf{H}_j$ in $\mathbf{H}^{temp}$;

27   Calculate the stability Q_S^{temp} of $\mathbf{H}^{temp}$;

28   **if** $Q_S^{temp} > Q_S^*$ **then**

29 $Q_S^* \leftarrow Q_S^{temp}$;

30 $\mathbf{H}_* \leftarrow \mathbf{H}_j$;

31   **end**

32   **end**

33   **if** $Q_S^* > Q_S$ **then**

34   Merge $\mathbf{H}_*$ and $\mathbf{H}_i$ in $\mathbf{H}$;

35   Calculate the stability Q_S of community partition $\mathbf{H}$;

36   **end**

37   **end**

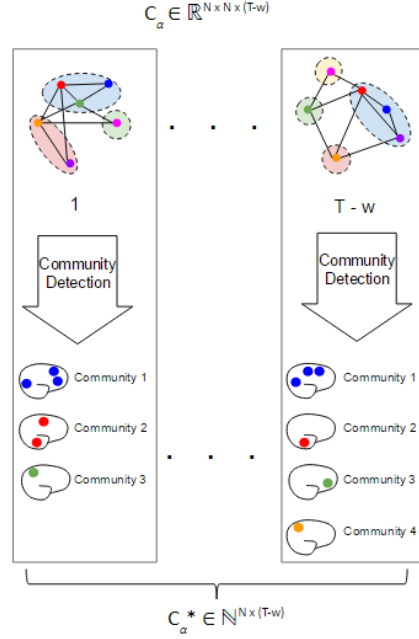**38 end**

39 return $\mathbf{H}$;

Figure 5.4: Input and output of the Community Detection. The colors of the nodes are not related to the colors of communities. It is an illustration example with $N = 6$ parcels.

5.4 Community Tracking

The communities defined by $\mathcal{C}_\alpha^*$ for subject α are equivalent to static communities detected inside each window of w time points. The community tracking, based on the similarity of the community footprints, temporally links them to form the subject-level Dynamic Communities (sDCs).

5.4.1 Problem Definition

For each time window $t \leq T$ and for a subject α, a set of static communities $\mathcal{C}_\alpha^{*t} = \left\{ C_{\alpha,1}^t, C_{\alpha,2}^t, \ldots, C_{\alpha,n_t}^t \right\}$, where n_t is the number of communities identified at time t, is now available via $\mathcal{C}_\alpha^* = \left\{ \mathbf{C}_\alpha^{*1} \ldots \mathbf{C}_\alpha^{*T-w} \right\}$. Indeed, the nodes labeled by $i \in [1; n_t]$ in $\mathbf{C}_\alpha^{*t}$ are members of $C_{\alpha,i}^t$. To introduce the temporal dimension, it is necessary to link these communities across time to form DCs $\mathcal{D}_\alpha$. One DC $\mathcal{D}_\alpha^l$, where $l \leq n_\alpha$ with n_α the final number of sDCs for subject α, is a sequence of static communities ordered by time: $\mathcal{D}_\alpha^l = \left\{ C_{\alpha,i_1}^{t_1}, C_{\alpha,i_2}^{t_2}, \ldots, C_{\alpha,i_l}^{t_l} \right\}$, with $t_1 < t_2 < \ldots < t_l$ and $\forall j < T$, $i_j < n_j$. This definition implies that $\mathcal{D}_\alpha^l$ does not have a community C_{α,i_j}^t at each time step t and that it cannot present more than one community C_{α,i_j}^t per time step t. Figure 5.5 illustrates the input and the output of this step named here tracking.

A community $C_{\alpha,i}^{t_j}$ will be attributed to the DC $\mathcal{D}_\alpha^l$ if its footprint on the brain surface
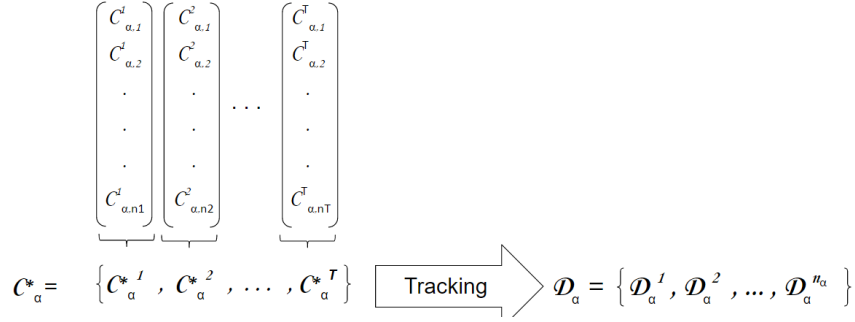
Figure 5.5: Input and output of the community tracking.

is similar enough to the footprint of the previous community in time $C_{\alpha,i_{j-1}}^{t_{j-1}}$ of $\mathcal{D}_\alpha^l$ and if no other community $C_{\alpha,i'}^{t_j}$ is more similar. The community tracking is based on the similarity because the footprint of DCs have to be as spatially stable as possible.

5.4.2 Similarity Definition

It is thus necessary to define a similarity function to build the DCs by tracking communities. One commonly used similarity metric in community mining is the Jaccard ([GDC10],[VBAW15]) defined as :

$$J(C_{\alpha,i_j}^{t_j}, C_{\alpha,i_{j-1}}^{t_{j-1}}) = \frac{\left| C_{\alpha,i_j}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\left| C_{\alpha,i_j}^{t_j} \cup C_{\alpha,i_{j-1}}^{t_{j-1}} \right|} \tag{5.1}$$

Takaffoli and al. [TSFZ11] also uses another similarity function :

$$sim(C_{\alpha,i_j}^{t_j}, C_{\alpha,i_{j-1}}^{t_{j-1}}) = \begin{cases} \frac{\left| C_{\alpha,i_j}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\max\left(\left| C_{\alpha,i_j}^{t_j} \right|, \left| C_{\alpha,i_{j-1}}^{t_{j-1}} \right| \right)} & \text{if } \frac{\left| C_{\alpha,i_j}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\max\left(\left| C_{\alpha,i_j}^{t_j} \right|, \left| C_{\alpha,i_{j-1}}^{t_{j-1}} \right| \right)} \geq k \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

where they introduce a similarity threshold k whose optimization is described in the Section 5.4.4. It implies that $C_{\alpha,i_j}^{t_j}$ and $C_{\alpha,i_{j-1}}^{t_j}$ cannot be similar if they do not share at least a proportion of k parcels in their respective footprint. The second definition of similarity is selected for the community tracking and this choice is explained in more details in Section 8.2. It is now possible to use community mining tools developed for social networks, and in this field, the tracking of DCs can be characterized by events.

5.4.3 Events Definition

The tracking is performed time point per time point. At $t = 0$, there is no DC. At t=1, each community detected in the first time window initializes a DC. Then, a DC can

survive only if in the following time windows there is at least one community which is similar enough to its last community. If such a community exists at time t, a *survival* happens at time t, but if there is no such community it is the *death* of the DC. Moreover, it can happen that several communities have a high similarity with the same DC at the same time step without being correlated and it provokes a *split* of the DC. On the contrary, if a unique community matches several non-correlated DCs, it is a *merge*. If a community has no match with any of the DC, it is a *birth*. These events are illustrated in Figure 5.6.
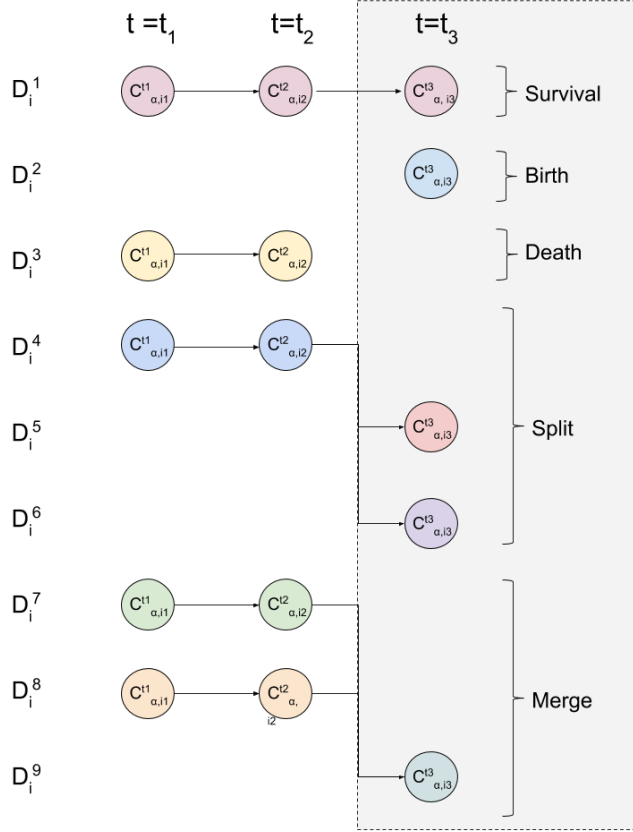

Figure 5.6: Example of DC partition illustrating the different types of events : survival, birth, death, split, merge. Each color represents one DC.

In the context of the brain dynamic RSNs represented by these DCs, *birth/death/survival* simply mean that the activation of the concerned RSN starts/stops/continues. However, in the case of a *split* or a *merge*, it reveals a high level of interaction between the RSNs involved in these events. Indeed, if for instance a DC A splits into DC B and DC C, it means that actually B and C share a high proportion of their parcels with A and are likely to merge again later to form one unique RSN or to merge with other DCs with a recurrent pattern. It is only an example but it is something interesting to observe if it
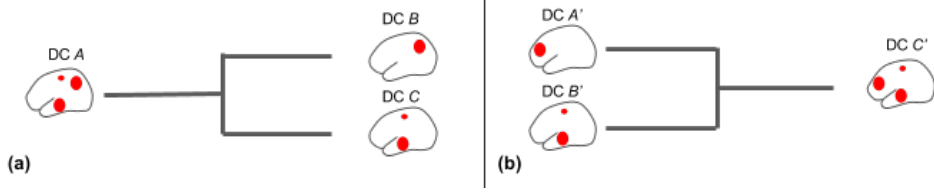
Figure 5.7: Illustration of community events on network footprints. (a) Split of DC A into DC B and DC C. (b) Merge of DC A' and DC B' into DC C'.

exists. Figure 5.7 depicts the translation of splits and merges from DCs to the concrete footprints on the brain surface.

Mathematically, the events are defined this way :

- *Birth*: $\mathcal{D}_\alpha^l$ is born at time t_j if:

$$\forall\, t_m < t_j,\ \left\{ sim(C_{\alpha,i_m}^{t_m}, C_{\alpha,i_j}^{t_j}) \geq k \right\} = \emptyset \tag{5.3}$$

- *Death*: $\mathcal{D}_\alpha^l = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\}$ dies at time $t_j = t_{j-1} + 1$ if:

$$\forall\, t_m \geq t_j,\ \left\{ sim(C_{\alpha,i_j}^{t_j}, C_{\alpha,i_{j-1}}^{t_{j-1}}) \geq k \right\} = \emptyset \tag{5.4}$$

- *Survival*: $\mathcal{D}_\alpha^l = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\}$ survives at time $t_j > t_{j-1}$ if $\exists\,!\ C_{\alpha,i_j}^{t_j}$ *where*:

$$sim(C_{\alpha,i_j}^{t_j}, C_{\alpha,i_{j-1}}^{t_{j-1}}) \geq k \tag{5.5}$$

- *Split*: $\mathcal{D}_\alpha^l = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\}$ splits at time $t_j > t_{j-1}$ if $\exists\, \mathcal{C} = \left\{ C_{\alpha,1}^{t_j}, \ldots, C_{\alpha,\eta}^{t_j} \right\}$ with $\eta \geq 2$ *where* :

$$\begin{cases} \forall\, C_{\alpha,i}^{t_j} \in \mathcal{C}, \ \dfrac{\left| C_{\alpha,i}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\left| C_{\alpha,i}^{t_j} \right|} \geq k. \\[2em] \dfrac{\left| (C_{\alpha,1}^{t_j} \cup \ldots \cup C_{\alpha,\eta}^{t_j}) \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\left| C_{\alpha,i_{j-1}}^{t_{j-1}} \right|} \geq k. \end{cases} \tag{5.6}$$

- *Merge*: A set of DCs $\mathcal{D} = \left\{ \mathcal{D}_\alpha^{l_1}, \ldots, \mathcal{D}_\alpha^{l_\eta} \right\}$ with $\eta \geq 2$ merges at time t_j if $\exists\, C_{\alpha,i}^{t_j}$ *where* :

$$\begin{cases} \forall\, \mathcal{D}_\alpha^\lambda = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\} \in \mathcal{D}, \ \dfrac{\left| C_{\alpha,i}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{\left| C_{\alpha,i_{j-1}}^{t_{j-1}} \right|} \geq k. \\[2em] \dfrac{\left| (C_{\alpha,1}^{t_{j-1}^{l_1}} \cup \ldots \cup C_{\alpha,\eta}^{t_{j-1}^{l_\eta}}) \cap C_{\alpha,i}^{t_j} \right|}{\left| C_{\alpha,i}^{t_j} \right|} \geq k. \end{cases} \tag{5.7}$$

The definition of *split* means that a DC A splits into communities $\{A_1, \ldots, A_\eta\}$ if at least k proportion of the parcels of communities A_i, $\forall i \in [1, \eta]$, are in A. The second condition of the definition prevents the effects of nodes from A leaving the network. Indeed, all the brain parcels are not attributed to a community at each time step due to a constraint on their size during the community detection and if the size of A_i is relatively small compared to the size of A, the first condition is easily fulfilled. The definition of *merge* requires from community A', which results from the merge of communities $\left\{A'_1, \ldots, A'_\eta\right\}$, that it shares at least k proportion of its parcels with community A'_i, $\forall i \in [1, \eta]$. The second condition of the definition is to ensure that most of the nodes of A' were already in the networks inside communities $\{A_1, \ldots, A_\eta\}$ before the merge because if community A'_i contains only few parcels, its parcel proportion k represents a low number of common parcels with A'.

These definitions are adapted from definitions developed by Takaffoli et al. [TSFZ11] for community mining in social networks. The notations are however different in order to be closer to the ones used in our algorithm. Moreover, some modifications have been realized in order to adapt definitions from [TSFZ11] to the specificity of our data. Indeed, after a DC splits, it dies and the resulting DCs are born, while after a set of dynamic communities merges, these DCs die and the resulting dynamic community is born. A *split* or a *merge* cannot be associated with a *survival*. These modifications are necessary in the algorithm to prevent the generation of spurious events as those illustrated in Figure 5.8(a1)&(a2) and Figure 5.8(b1)&(b2). The deaths of the involved DCs are particularly necessary after a *merge*. Indeed, if a DC is not active during several time steps, it is not considered as dead to allow temporal gap in the activity. Therefore, DCs that have merged can continue to merge with their resulting DCs at each time step until the initial merged community is no more active (see illustration on Figure 5.8(b1)&(b2)).

When no *split* or *merge* happens to a DC, its *death* is never declared. It is necessary in order to allow temporal gaps in its activity. It would be possible to check if a DC is dead or not at time t with this requirement, but it would be necessary to look at each time step $t' > t$ if it survives, splits or merges and it would imply a very high computational cost.

Alternatives for the definition of events and consequently of the DCs have been tested but the selected approach provides the most robust DCs with a relatively high number of *splits* and *merges*. Its main drawback is the formation of a high number of DCs that are considered as noise due to their short life time and that present a high redundancy. The definitions selected here therefore enhance the need for the post-processing steps explained in Sections 5.5.1 and 5.5.2.

5.4.4 Optimization of the similarity threshold k

The choice of the similarity threshold k for the community tracking is of key-interest since it determines the event detection as well as the stability of the detected dynamic communities. Indeed, the higher the similarity threshold is, the more stable are the

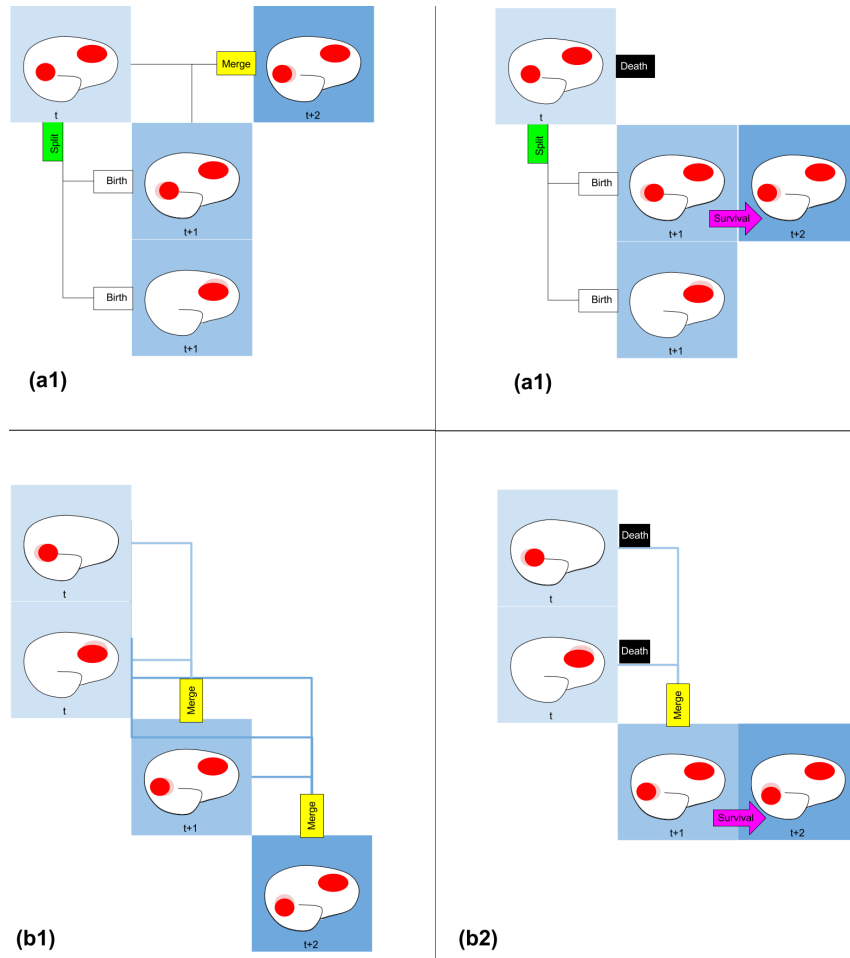Figure 5.8: 2 scenarios to illustrate the necessity of the *death* of the involved DCs after a *split* (a) and after a *merge* (b). (a1) Spurious *merge* detected after a *split* at t+1.(a2) Solution to prevent this effect : after a *split*, the concerned DC dies.(b1) Spurious cascade of *merges* after a true *merge* at t+1. (b2) Solution to prevent this effect : after a *merge*, the DCs that have merged die.

identified dynamic communities. Moreover, if the similarity threshold is too high, the number of *birth* increases because *survival, split* and *merge* happen only for higher similarities. Therefore, the number of DCs is a criterion for the choice of k, even if the consolidation step is able to solve the issue of a too high number of DCs. The other criterion used by Takaffoli et al. [TSFZ11] is the number of *mutual topics*. A mutual topic is defined in the context of text mining but it can also be applied to brain networks by using brain surface parcels instead of keywords. They assume that a DC which survives during several time frames is more likely to conserve the same parcels, or when text is concerned, to be about the same topic. Thus the average number of mutual topic between all the DCs obtained must be as high as possible because long living DC are assumed to be more robust. As a minimum of stability between the consecutive communities of the DCs is required due to the noise, the similarity threshold is tested in the range of [0.4,0.8]. Indeed, the higher is k, the higher is the stability. If several maxima appear into this range, the choice of k must be decided according to the expected stability of DCs [TSFZ11]. k must not be too high because the community tracking also detects events and a too high k results in no event of high level of interaction. The number of events is therefore another criterion for the choice of k and counterbalances the need of a minimal stability. On the other hand, it is also important to optimize k in keeping in mind that it must isolate the spurious connectivity resulting from noise into short-lifetime DCs which can be eliminated via the pruning step. It is therefore another argument for a rather high k.

For our experiments the optimized k is equal to 0.6 (see result of the optimization in Section 7.4.2).

5.4.5 Algorithm

Simplified case

To illustrate the approach used in the developed algorithm for community tracking, the definitions given above are simplified. A simple manner to consider this problem is to use a similarity matrix. If we consider that all communities C_i, $\forall i \in \mathbb{N}$ have the same size N_C, we have :

$$\forall (i,j) \in \mathbb{N}^2, \ sim^*(C_i, C_j) = \frac{|C_i \cap C_j|}{|max(C_i, C_j)|} = \frac{|C_i \cap C_j|}{N_C} \tag{5.8}$$

This definition of the similarity sim^* is adapted from the similarity function sim defined in Equation 5.2. The difference between these two functions is only that the similarity threshold is no more considered in sim^*.

Then, if we consider only the first condition in *split* and *merge* definition, events can be defined as :

- *Split*: $\mathcal{D}_\alpha^l = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\}$ splits at time $t_j > t_{j-1}$ if $\exists \, \mathcal{C} = \left\{ C_{\alpha,1}^{t_j}, \ldots, C_{\alpha,\eta}^{t_j} \right\}$ with $\eta \geq 2$ *where* :

$$\forall \, C_{\alpha,i}^{t_j} \in \mathcal{C}, \frac{\left| C_{\alpha,i}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{N_C} \geq k. \tag{5.9}$$

- *Merge*: A set of DCs $\mathcal{D} = \left\{ \mathcal{D}_\alpha^{l_1}, \ldots, \mathcal{D}_\alpha^{l_\eta} \right\}$ with $\eta \geq 2$ merges at time t_j if $\exists \, C_{\alpha,i}^{t_j}$ *where* :

$$\forall \, \mathcal{D}_\alpha^\lambda = \left\{ C_{\alpha,i_1}^{t_1}, \ldots, C_{\alpha,i_{j-1}}^{t_{j-1}} \right\} \in \mathcal{D}, \frac{\left| C_{\alpha,i}^{t_j} \cap C_{\alpha,i_{j-1}}^{t_{j-1}} \right|}{N_C} \geq k. \tag{5.10}$$

In this particular case, the similarity matrix $\mathbf{S}^{t_j} = (S_{a,b}^{t_j})$ established at a time step t_j is defined as $S_{a,b}^{t_j} = sim^*(C_{i,a}^{t_j}, C_{i,b}^{t_{j-1}})$. This matrix is turned into a binary matrix $\mathbf{S}^{*t_j} = \{\mathbf{S}^{t_j} \geq k\}$ and a similarity matrix such the one represented on Figure 5.9 is obtained.
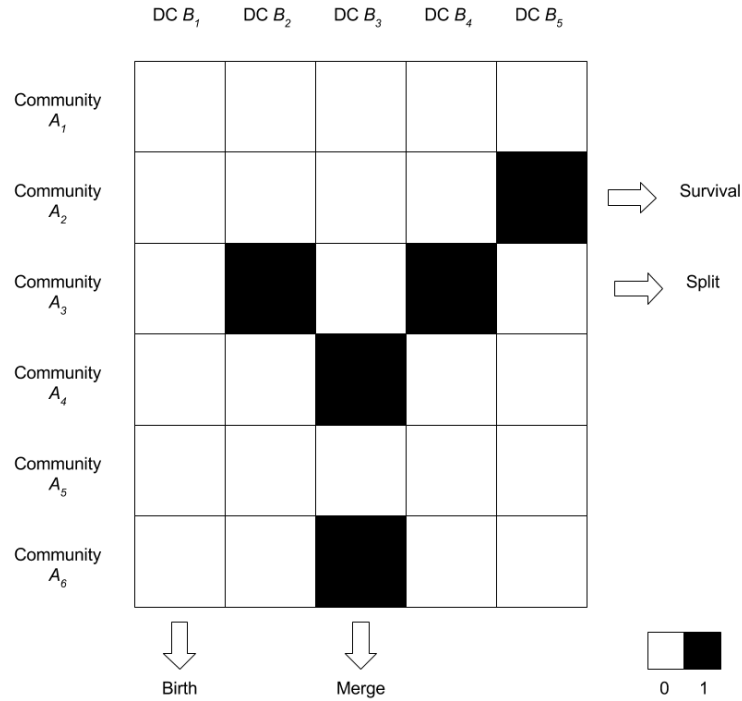

Figure 5.9: Binary Similarity matrix representing the similarity between the community of the current time window and the DCs constructed until the previous time window. The coefficient of the matrix is different from zero if the similarity if higher than k, the similarity threshold. Each event situation is represented here.

It is the idea illustrated by Figure 5.9 that is used in the tracking algorithm. If in the row of the community A_i, no DC matches, it is a *birth*. If at least two DCs match, it is a *split*. If one DC match and if it is the only matching of the DC, it is a *survival*, but if the DC has more than one matching, it is a *merge*. In case of mute DC at time step t, nothing has to be done as it has been decided that the dynamic community stay available even with temporal gap.

Real case

All the dynamic communities do not have the same size. Three different "similarity" matrices are necessary : $\mathbf{SS}^{t_j} = (SS_{a,b}^{t_j})$ to detect the splits, $\mathbf{SM}^{t_j} = (SM_{a,b}^{t_j})$ to detect the merges and $\mathbf{S}^{t_j} = (S_{a,b}^{t_j})$ to detect survival and birth. Their respective definitions are given in equation (5.11). n_{t_j} is the size of $\mathcal{C}_\alpha^{*t_j}$, that is to say the number of static communities identified at time t_j, and n_{D_α} is the size of $\mathcal{D}_\alpha = \left\{ \mathcal{D}_\alpha^1, \dots, \mathcal{D}_\alpha^{N_{D_\alpha}} \right\}$ at t_{j-1}, that is to say the number of already born DCs at time t_{j-1}, in subject α.

$$\forall (a,b) \in \mathbb{N}^{n_{t_j} \times n_{D_i}}, \begin{cases} SS_{a,b}^{t_j} = \dfrac{\left| C_{i,a}^{t_j} \cap C_{\alpha,i_{j-1}}^{b,t_{j-1}} \right|}{\left| C_{i,a}^{t_j} \right|} \\[2ex] SM_{a,b}^{t_j} = \dfrac{\left| C_{i,a}^{t_j} \cap C_{\alpha,i_{j-1}}^{b,t_{j-1}} \right|}{\left| C_{\alpha,i_{j-1}}^{b,t_{j-1}} \right|} \\[2ex] S_{a,b}^{t_j} = \dfrac{\left| C_{i,a}^{t_j} \cap C_{\alpha,i_{j-1}}^{b,t_{j-1}} \right|}{max\left(\left| C_{i,a}^{t_j} \right|, \left| C_{\alpha,i_{j-1}}^{b,t_{j-1}} \right| \right)} \end{cases} \tag{5.11}$$

A first scan of communities is made with matrix $\mathbf{SS}^{t_j}$ to detect *splits* (see Figure 5.10(a)), the second with $\mathbf{SM}^{t_j}$ to detect *merges* (see Figure 5.10(b)), and a final one with $\mathbf{S}^{t_j}$ to look at *survivals* and *births* (see Figure 5.10(c)). These different similarity matrices are defined for specific events and only the concerned events are detected via the analysis of each matrix. Before declaring a *split* or a *merge*, the second condition of each definition is verified (see equations (5.6) and (5.7)) on the potential *splits* and *merges* respectively detected with $\mathbf{SS}^{t_j}$ and $\mathbf{SM}^{t_j}$.

Output

After this tracking step, a set of n_α sDCs is available for the subject α. Each sDC is characterized by its dynamic footprint corresponding to the footprints of the communities part of this sDC, and by the time line of its activation. When the sDC is not active at t, it has no corresponding community in the time window t. These data are represented by a front-tracer matrix $\mathbf{D}_\alpha = (D_{itj,\alpha})$ of size $N \times (T - w) \times n_\alpha$. $D_{itj,\alpha} = 1$ if the node i is in the community j in the window t and $D_{itj,\alpha} = 0$ if not. This matrix is stored into a structure $\mathcal{D}_\alpha$ with the characteristics of the *splits* and *merges* and also with properties of the sDCs, such as the label of their constitutive community $C_{\alpha,i}^t$ for each t where it is active and the similarity between the corresponding consecutive footprints.
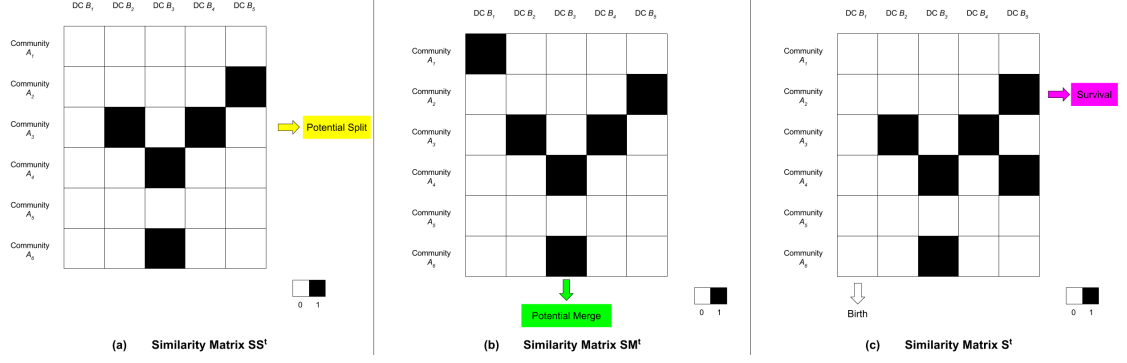
Figure 5.10: Example of different similarity matrices used for the community tracking at time t. (a) Matrix for the detection of the splits $\mathbf{SS}^t$. (a) Matrix for the detection of the merges $\mathbf{SM}^t$. (a) Matrix for the detection of the survival and birth $\mathbf{S}^t$.
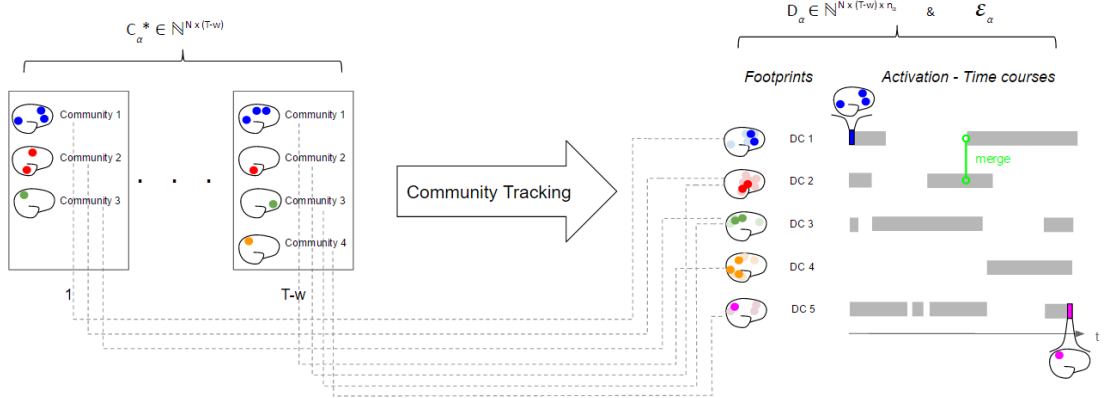

Figure 5.11: Input and output of the Community Tracking for the subject α. The choice of colors is arbitrary and two communities of the same color are not necessary members of the same dynamic community. The dotted lines link the communities to the DC they are part of. It is an illustration example.

The latter is for instance used to calculate the stability of a sDC and it also holds the information of the lifetime of a sDC, two parameters of interest for the next step. Figure 5.11 summarize the process performed with this community tracking step. The DCs require then post-processing.

5.5 Post-processing of the Subject-level Dynamic Communities (sDCs)

After tracking, when we observe the resulting time lines and footprints contains into $\mathcal{D}_\alpha$ for the identified sDCs into subject α, a maority of the sDCs displays one of these three characteristics :

- a very short lifetime,

- a low stability,

- a high redundancy.

The two first observations are characteristics for noisy dynamic communities which have to be removed by *pruning*. The high redundancy among the remaining set of dynamic communities requires a consolidation step to link the fragments of sDCs that should not have been detected as different sDCs during the tracking for different reasons.

5.5.1 Pruning

In this pruning operation, we assume that a dynamic community with a too short lifetime and/or a too low stability corresponds to noise. Indeed, the preprocessing realized before the SWA is not able to remove all the noise from the data. These properties are calculated with the data from $\mathcal{D}_\alpha$.

The stability of a dynamic community is defined as the mean of the similarity between all its constitutive footprints and described to what extent it is well defined on the brain surface. The similarity definition is the same as the one used for the community tracking and as at each time frame, a different footprint is attached to the dynamic community, the averaged similarity is obtained by dividing the sum of the similarity between each time frame by the number of alive time points. A dynamic community is considered as noise if its stability is under a defined threshold or if its lifetime is under another threshold. The pruning on the footprint stability restricts the spatial variability of a dynamic community.

Concerning the pruning on lifetime, the temporal resolution is reduced since an assumption on the lifetime of a dynamic community is made. It has an impact for the neurobiological interpretation of the sDCs, but it is necessary due to noise.

The thresholds for the pruning have been arbitrary fixed according to the observation of the time lines that permits to obtain an estimation of a "low" lifetime as well as through

the observation of the stability values for different DCs. With these observations, the stability threshold is equal to 0.2 and the lifetime threshold is equal to 10. The choice of the thresholds has also been influenced by the quantity of information we need concerning the events. The threshold are as small as possible to preserve as many splits and merges as possible.

The sDCs that are pruned are simply removed from $\mathcal{D}_\alpha$ and the front-tracer matrix $\mathbf{D}_\alpha$ is rearranged. The *splits* and *merges* which involve these sDCs are also erased if the remaining sDCs are not sufficient to ensure their integrity. For a *split*, the initiating sDC has to survive the pruning and at least two resulting sDCs have to survive as well. For a *merge*, at least two initiating sDCs and the resulting sDC have to remain. This pruning step is performed before the consolidation because if the noise is consolidated with true DCs, the latter are blurred.

5.5.2   Consolidation

The consolidation consists in linking DC fragments that have been identified as different sDCs during the tracking. After the pruning used to delete the sDCs resulting from noise, it is possible to realize this operation on $\mathcal{D}_\alpha$.

Motivation

There are several reasons that can explain that two DCs are declared different during the tracking although they should be considered as the same DC. First of all, the events definitions imply that after a *split* or a *merge* at time t, the concerned DCs die at time $t+1$ (see Figure 5.12(a), left panel). Therefore, no more community can be integrated to the sequence of this DC for $t' > t$. However, it is highly probable that one of the DCs resulting from a split is highly similar to the DC that has been *split* and when a *merge* is concerned, one of the DC that has merged may also be very similar to the resulting DC (see Figure 5.12(a), right panel). These highly similar DCs are then two fragments of the same DC which need to be consolidated. Even though DCs resulting from events are not fragments of the initial DCs, the latter are prevented to later reappear and it can provoke the birth of a new DC after several time frames which is just a fragment but that must be labeled differently due to the events (see Figure 5.12(b), left panel). These two fragments have then to be associated in the consolidation step (see Figure 5.12(b), right panel).

Furthermore, tracking is made step by step and a DC is characterized by the footprint of its last living frame. Therefore, if this frame is noisy (see Figure 5.13(c)) or slightly shifted (see Figure 5.13(b)), the corresponding DC may not survive whereas if previous footprints were available, it would have survived, as illustrated in Figure 5.13.

The aim of the consolidation is thus to decrease the redundancy generated by these two phenomena in the set of DCs.
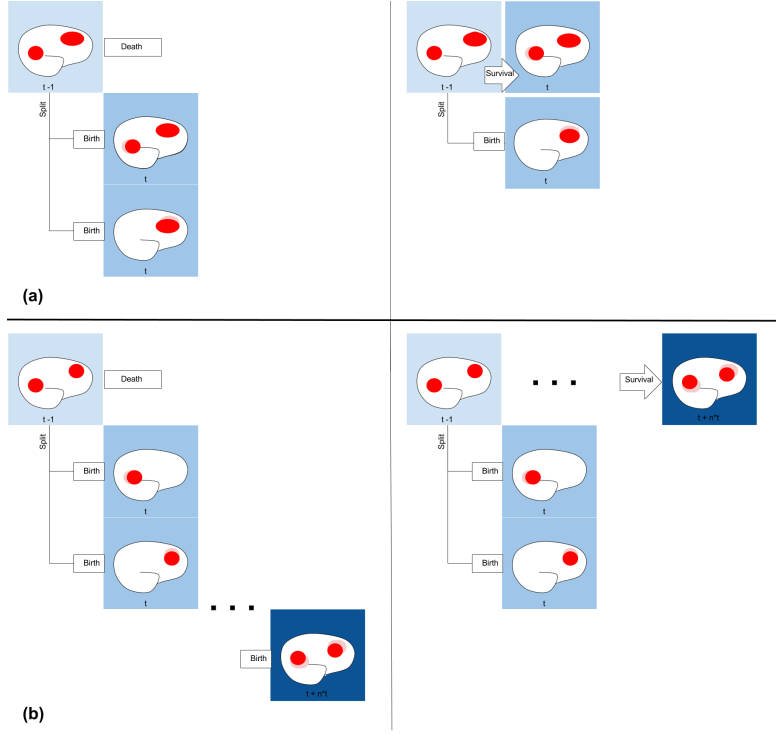
Figure 5.12: Illustration of two cases where the consolidation is necessary after an event. The DCs are represented by their footprint in red and the shadows for $t' \geq t$ represents the footprint of the initial DCs from $t - 1$. **(a)** Left panel: before the consolidation, a split results into 2 new DCs and the first one is very similar to the DC that has split. Right panel: after the consolidation. **(b)** Left panel: before the consolidation, after the split two DCs are born but none is similar enough to the initial DC to be consolidated with it. After several time frames represented by dots, a new DCs is born highly similar to the initial DC and they must be consolidated. Right panel: after the consolidation.

Spectral Clustering

To realize this consolidation step, a spectral clustering based on the similarity between the averaged footprints of each DCs is performed. It means that the averaged footprints (see definition in the next paragraph) of the DCs included in the same cluster are similar while averaged footprints from different clusters are dissimilar. The advantage of spectral clustering here is that we do not need to define a similarity threshold for the DCs to be in the same group. However, we need to fix the number of resulting clusters (see Section 7.4.3). Each group obtained from clustering is then considered as a unique DC.

Averaged Footprint A DC has potentially a different footprint at each time frame it is alive. It has a set of footprints that are similar enough to be attributed to the same DC but that also displays differences which can be explained by noise or by small spatial
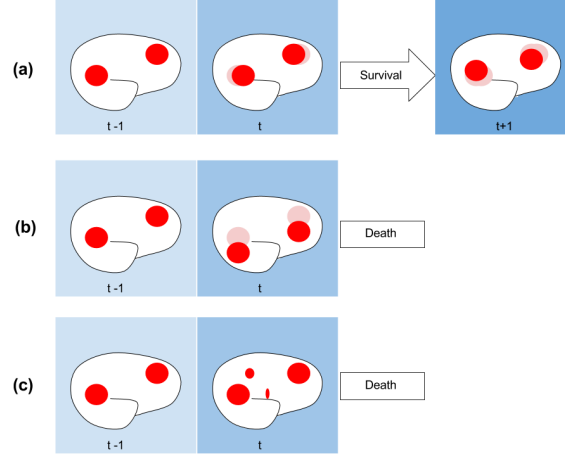
Figure 5.13: Examples to illustrate the effect of noise on DC survival. (a) Ideal sequence. (b) Sequence cut off due to a shift of the footprint at time t. (c) Sequence cut off due to noisy footprint at time t.

fluctuations inherent to the dynamic characteristics of functional networks. However, to compare the DCs in order to implement the consolidation, it is necessary to have for each DC a single footprint. In order to embrace the whole dynamic footprint with its spatial variability, all the static footprints of one DC are simply summed and the sum is then normalized by the number of time frames when DC is alive. For each DC, an averaged footprint is thus generated. Each parcel that has been active at least once in the lifetime of the considered DC has a coefficient different from 0 and the most recurrent parcels have a coefficient close to 1 in this footprint. The averaged footprint of sDC i in subject α is represented by $\bar{\mathbf{D}}_\alpha^i \in \mathbb{R}^N$.

Similarity definition   During the community tracking, a similarity function has already been defined according to the definition introduced by Takafolli et al. [TSFZ11] (see Section 5.4). This definition is used to build communities and we have preferred it to Jaccard for the community tracking (see Section 8.2 for the explanation). However, Jaccard is here more appropriate to the consolidation since it is a harder metric (see Section 8.2).

Moreover, in order to reduce the influence of noise appearing in the footprint $\bar{\mathbf{D}}_\alpha^i$ in the shape of low coefficient nodes, only the nodes that are present in at least half of the footprints of the DCs are selected and therefore nodes whose coefficient in $\bar{\mathbf{D}}_\alpha^i$ is higher than 0.5. The counterpart of this selection is that we are also losing a part of the time-varying footprint but the consolidation is based on its skeleton since it aims to gather fragments of the same DC at different time points. Therefore, the similarity matrix coefficient S_{ij} corresponding to the Jaccard between the averaged footprint of

sDC i and the averaged footprint of sDC j in subject α can thus be defined as [GDC10]:

$$S_{ij} = \frac{\bar{D}_\alpha^i \cap \bar{D}_\alpha^j}{\bar{D}_\alpha^i \cup \bar{D}_\alpha^j} \qquad (5.12)$$

Graph definition The similarity matrix $\mathbf{S}$ defined above can be utilized to construct an undirected similarity graph as described in Section 3.2. The similarity graph $\mathbf{G}$ is described as $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ where $\mathbf{V} = \{v_1, \ldots, v_n\}$ is the set of vertices here corresponding to the DCs and $\mathbf{E}$ are the undirected weighted edges which are constructed according to the selected type of graph. The weights of each possible edges are stored into $\mathbf{W} = (w_{ij})$. If two vertices i and j are not connected, $w_{ij} = 0$. Moreover, the larger the similarity is, the lower is the weight.

Implementation To perform the spectral clustering, the approach used is described in Section 3.2. Instead of imposing a fixed number of cluster for the k-means clustering, we have decided to add an optimization step to evaluate which number of final sDCs has be identified in $\mathcal{D}_i$. An automatic clustering method could have been used [JGGF16] but as the literature already provides estimation of the number of RSNs (see Section 2.2.1), an optimization based on simple metrics with a fixed range for this number is an appropriate choice. The selection of this metric, the Silhouette coefficient, is explained in Section 7.4.3.

The Algorithm 5.2 adapted from [NJW$^+$02] describes the procedure adopted. The number of clusters can vary between $k_{min} = 7$ and $k_{max} = 13$ and is optimized for each subject owing to the calculation of the Silhouette coefficient on several iteration of the spectral clustering with different k (see Section 3.2.5). Therefore, all subjects do not have the same number of final DCs but it is normal since a rest-fMRI sequence lasts approximately 10 minutes and all the dynamic RSNs may not be active during this time.

Algorithm 5.2: Spectral clustering for consolidation of sDCs

    Input: Similarity matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$, frame $[k_{min}, k_{max}]$ for the number of cluster k
    Output: Clusters $A_1, \ldots, A_k$

1 Construct the similarity graph $\mathbf{G}$ as kNN graph from $\mathbf{S}$. $\mathbf{W}$ is the weighted adjacency matrix of $\mathbf{G}$.;

2 Compute the symmetric normalized graph Laplacian $\mathbf{L}_{sym}$;

3 Find the best k in $[k_{min}, k_{max}]$ for the subject;

4 Calculate the first k eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u_k}$ of $\mathbf{L_{sym}}$ and store them as columns of matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_k] \in \mathbb{R}^{n \times k}$;

5 Normalize rows of $\mathbf{U}$ to norm 1 in order to obtain $\mathbf{T} = [\mathbf{y}_1 \cdots \mathbf{y}_n]^T \in \mathbb{R}^{n \times k}$;

6 Cluster the points $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ into clusters $C_1, \ldots, C_k$ with the k-means algorithm;

7 **return** *Clusters $A_1, \ldots, A_k$;*

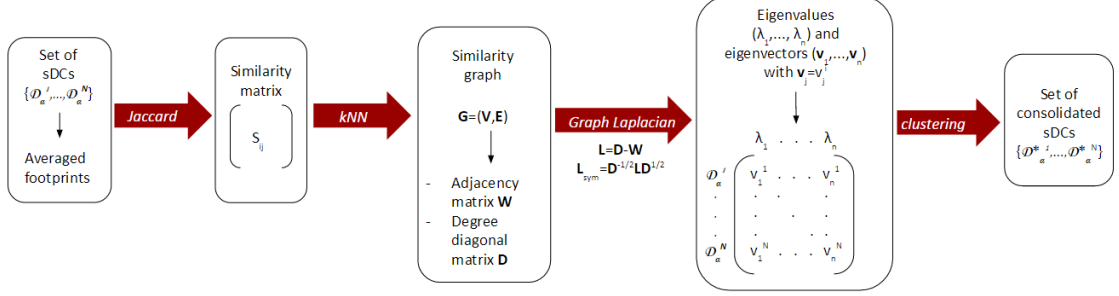The process of the spectral clustering is reminded and summarized in Figure 5.14.

Figure 5.14: Steps of the spectral clustering of the N sDCs $\left\{\mathcal{D}_\alpha^1 \ldots \mathcal{D}_\alpha^N\right\}$ from subject α, as described in Section 3.2.

Output As output, the new structure $\mathcal{D}_\alpha^*$ is similar to $\mathcal{D}_\alpha$ except that is contains the data for the post-processed sDCs $\mathcal{D}_\alpha^{*i}$ of subject α. The consolidation agglomerates the sDCs $\mathcal{D}_\alpha^i$ surviving to the pruning into a smaller number of sDCs $\mathcal{D}_\alpha^{*i}$ according to their similarity. It remains between 7 and 13 sDCs $\mathcal{D}_\alpha^{*i}$ for each subject α. The detection of population-level DCs (pDCs) is now possible.

5.6 Detection of population-level Dynamic Communities (pDCs)

We have now a set of sDCs $\mathcal{D}_i^*$ with their associated events for each subject i. In order to check the consistency of our approach and also to detect recurrent patterns in DCs interactions, we need to expand from subject level to population level. All the steps described above are applied to $N_S = 200$ subjects which form our population. It is now possible to look for DCs that are present in several subjects, as summarized in Figure 5.15.

The detection of pDCs is based on the similarity of their averaged footprint. The latter is constructed in an analog way as in the consolidation step (see Section 5.5.2), but this time for the consolidated sDCs. Then, like in the consolidation step, a similarity matrix **S** is calculated to compare the whole dynamic communities from every subject and the similarity function is also the same as in consolidation at subject-level. Particular constrains prevent to use once again spectral clustering to parcellate the set of dynamic communities into sub-groups. First of all, all the sDCs are not necessarily attributed to a pDC. Indeed, in spite of the de-noising steps, some of the detected sDCs are only noise or are at least affected by artifacts. Then, it is necessary to differentiate the DCs identified in the same subject: two sDCs from the same subject cannot be associated to the same pDCs. It could happen if clustering would be the selected solution because

Figure 5.15: Input and output of the Population Tracking. The colors are arbitrary. The number of sDCs and pDCs, as well as the footprints represented here are not meaningful, they are a simplification of the real case to facilitate the visualization. N_S is the number of subjects constituting the population.

despite the consolidation, few subjects still have redundant sDC footprint. To respect these constraints, a "manual" clustering is thus required.

A similarity graph $\mathbf{G}_{Pop} = (\mathbf{V}_{Pop}, \mathbf{E}_{Pop})$ is constructed with sDCs as vertices in $\mathbf{V}_{Pop}$. The edges are constructed with the thresholded similarity matrix defined above and thus, two sDCs are linked if their similarity is above the defined threshold. It is an unweighted graph. The similarity threshold is arbitrary fixed to 0.3, that is to say that the footprints of dynamic communities must at least display an overlap of 30% to be connected. It is a rather low threshold but the variability between subject connectivity networks can be high [HWA$^+$13]. Several tests of thresholds have however permitted to select this threshold as a compromise between stability across subject and number of subjects it involves. Moreover, in order to respect the second constraint that prevent the association of two sDCs to the same pDC, all the links between the dynamic communities from the same subject are removed.

The algorithm described in this paragraph corresponds to the pseudo-code 5.3. The degree of each sDC is calculated and sDCs are sorted in a descending order according to their degree. The sDC with the highest degree is selected and all the other dynamic communities that are connected to it and that are not in the same subject are matched

79

with it to form the first pDC. Their respective properties are merged and they are removed from the graph to construct a new graph. The degrees are recalculated in this new graph and the remaining sDCs are reorganized in a descending order with regard to their degree in this new graph. Once again, the sDC with the highest degree is selected as cluster centroid and the same steps are iterated until the highest degree of the current graph reaches the lowest acceptable degree for a center that we have fixed to 10 after several tests. Thus, the set of pDCs are constructed and it is described by $\mathcal{P}$.

Output

A set of pDCs that represent the consistent dynamic functional networks identified among subjects is finally constructed. The properties and events of their constitutive sDCs are as well integrated to them, and it is now possible to realize a population-level analysis. All the data concerning the population level, such as the constitutive sDCs of each pDC or their averaged footprint, are stored into a structure $\mathcal{P}$.

5.7 Definition and Analysis of the Interactions

The representation developed in this thesis that result into $\mathcal{P}$ access to two levels of interactions defined by events. The higher level of interaction is characterized by *splits* and *merges* whose definitions are described in Section 5.4. It is a spatial and temporal interaction since it is revealed by an exchange of node at a particular time point. They are inherent to community tracking. Another set of events needs to be defined in order to characterize the lower level of interaction only based on temporal information. These events are detected at subject level, after the community tracking and the post-processing. Then, methods to independently measure the occurrence of the events and their accuracy as representation of interactions are required.

5.7.1 Definition of the events

Definitions of interactions based on temporal and spatial information: higher level of interaction.

The events of higher interaction are the *splits* and *merges* defined in Section 5.4. Their definition is yet too restrictive and they are not enough to survive at the population level. Consequently other markers are needed to analyze the interactions between the sDCs and then between the pDCs. A lower level of interactions is thus characterized.

Definitions of interactions based on temporal information: lower level of interaction.

Events characterizing a lower-level of interaction between DCs are likely to occur more frequently than *splits* and *merges*. For two sDCs A and B, they are defined as:

- $death - death(A, B) = 1$ if A dies at t and B dies between $t - \tau$ and $t + \tau$.

Algorithm 5.3: Population-level DC tracking

    Input: Sets of sDCs $\left\{ \mathcal{D}_1^*, \ldots, \mathcal{D}_{N_s}^* \right\}$, events $\mathcal{E}$, similarity threshold k_{Pop}, lowest
           degree for center d_{min}
    Output: pDCs partition $\mathcal{P} = \{\mathcal{P}_1, \ldots, \mathcal{P}_n\}$, events $\mathcal{E}_{Pop}$

1 Calculate the similarity matrix $\mathbf{S} = (S_{ij})$ between all the sDCs in $\left\{ \mathcal{D}_1^*, \ldots, \mathcal{D}_{N_s}^* \right\}$;

2 $\mathbf{W} = (W_{ij}) \leftarrow \mathbf{S}$;

3 $W_{ii} \leftarrow 0$;

4 **if** $W_{ij} < k_{Pop}*$ **then**

5 | $W_{ij} \leftarrow 0$;

6 **end**

7 **if** $W_{ij} \neq 0$ *for two sDCs of the same subject* **then**

8 | $W_{ij} \leftarrow 0$;

9 **end**

10 Construct the similarity graph $\mathbf{G}_{Pop}$ with $\mathbf{W}$ as adjacency matrix;

11 Compute the degree d_i of each node;

12 $N_s \leftarrow 0$;

13 **while** $max(d_i) > d_{min}$ **do**

14 | $N_s \leftarrow N_s + 1$;

15 | Define the node with the highest degree $max(d_i)$ as a new cluster centroid.;

16 | Attribute all its direct neighbors to the new cluster $\mathcal{P}_{N_s}$ and merge their data.;

17 | **for** *all nodes i in the cluster P_{N_s}* **do**

18 | | $S_{i.} \leftarrow 0$;

19 | | $S_{.i} \leftarrow 0$;

20 | **end**

21 | $\mathbf{W} \leftarrow \mathbf{S}$;

22 **end**

23 **return** $\{\mathcal{P}_1, \ldots, \mathcal{P}_n\}$ *and the corresponding events $\mathcal{E}_{Pop}$;*

- $birth - birth(A, B) = 1$ if A is born at t and B is born between $t - \tau$ and $t + \tau$.

- $death - birth(A, B) = 1$ if A dies at t and B is born between $t - \tau$ and $t + \tau$.

- $death - survival(A, B) = 1$ if A dies at t and B is active between $t - \tau$ and $t + \tau$.

- $birth - survival(A, B) = 1$ if A is born at t and B is active between $t - \tau$ and $t + \tau$.

Two parameters are of key interest in these definitions. The first one is τ which corresponds to the time delay tolerated for the simultaneity of death and/or birth. Indeed, when the DC A is born/dead, the DC B must be born/dead less than τ time frames before or after the birth/death of DC A. The observation of the timelines of the DCs and the evolution of their correlation permits to arbitrary fix τ at 5. This τ is necessary due to the nature of SWA that average the information and introduce an uncertainty in

time. The second parameter that must be adequately defined is the number of inactive time frames required to declare a DC as dead. It is not an explicit of the definition this issue must be fixed. Indeed, in the definition of community tracking (see Section 5.4), temporal gaps are allowed in DCs to enable the DC to survive after as many time frame as necessary. A DC can thus have several episodes of life and death during the scan. A larger tolerance to temporal gap, that is to say a larger required number of inactive time frames to declare a DC as dead, decreases the number of death and birth and therefore of additional events. However, in counterpart it increases the reliability of the events because a long death is more likely to be a true death. In the results presented in Section 5.4, this minimal number of dead time points is 20. Other numbers have been tested but 20 is a good compromise and it is rather consistent with the lifetime threshold equal to 10 selected for the pruning step.

5.7.2   Interpretation of the events

Higher level of interaction

Splits and *merges* reveals a high level of interaction since they link two networks that have been a unique network during the time course. These two networks share thus a high proportion of their nodes and their time courses should be intricate. A network that splits or merges can then be related to different functionalities during its lifetime. Such an observation would for instance give an insight on how the brain is organized to switch its function and it could serve as disease biomarker if it happens that some diseases affect this organization.

Lower level of interaction

The analysis of *death-death, birth-birth, death-birth, death-survival* and *birth-survival* events across subjects permits for instance to observe which pDCs stop often their activity together when it is a recurrent *death-death* event, or which ones activate at the same time according to *birth-birth* events, are they the same or not. With the *death-birth* events, if the death of pDC is often simultaneously with the birth of another pDC, it could for instance means that switching off one pDC implies switching on another. Mutual influences of functional networks could thus be analyzed. For instance, it could give hints about the information integration by the brain by determining pattern of activation through events, and it could also be interesting and easy to access biomarkers.

5.7.3   Methods for event analysis

Increasing the accuracy of the low interaction events

In order to increase the reliability of the death and birth as a complement to the parameters described above, data from initial steps are re-used. During the community tracking, a DC is born at time t if the parcels constituting its footprints at time t are correlated. For the definition of events, a death at time t is declared as a death only if

the correlation of the footprint at time t has decreased at least of 70% after the necessary time to be declared as dead (here 20 time points). For a birth at time t, the upcoming being born footprint, that is to say the footprint of the DC at time t, must have a 70% lower correlation at t minus the required time necessary to declare the previous episode of life as dead. These constraints on the correlation permits to ensure a real birth and a real death of the DC and therefore an accurate additional event detection.

From the subject level to the population level

The events are first identified at the subject level and characterize thus the interaction of sDCs. However, to obtain a statistically relevant and robust estimation of the interactions between the dynamic functional networks, it is necessary to transfer these events at the population level. Therefore, it reduces the number of events taken into account in the analysis since they must involve at least two sDCs that are components of a pDC (see Section 5.6). The other events that happen between sDCs that are not linked to a pDC are not considered in the analysis.

Event occurrence

An event is considered as relevant if it appears in more than one subject. Indeed, it increases the probability that it is not a spurious event and if it characterizes an interaction, its repeatability across subjects demonstrates this interaction exists. The aim is therefore to identify the top-occurring events that happen in several subjects.

Matrices of occurrence for *death-death, birth-birth, death-birth, death-survival* and *birth-survival* events are used to visualize the couple of pDCs that are involved in the top-occurring events. Indeed, events characterizing the lower level of interaction happens between two pDCs i and j. The coefficient of one of this matrix in the row i and column j corresponds to the number of times the death/birth of pDC i occurs at the same time as the death/birth/survival of pDC j. For the *death-death* and *birth-birth* events, these occurrence matrices are symmetric by definition. Indeed, if the pDC i dies/is born at time t and the pDC j dies/is born at time $t' \in [t-5, t+5]$, it implies that the pDC i dies/is born at time $t \in [t'-5, t'+5]$ since it is equivalent to write $t \in [t-10, t+10]$. On the other hand, the occurrence-matrix for *death-survival, death-survival* and *birth-survival* are not necessarily symmetric.

For the higher level of interaction, the events are less numerous and sparse. Therefore, a simple extraction of the top-occurring events is enough.

One issue for the analysis of events is the dependency of their number of occurrence regarding the occurrence, and therefore the co-occurrence, of the DCs it involves. Indeed, if a DC lives longer than the other DCs, it has more birth and death events and is more likely to survive when others die or are born. Therefore, its amount of additional event it is involved in is higher. Moreover, two DCs with long lifetimes may have more temporal overlap and may also share an higher number of events than the other couples of DCs. To evaluate the significance of the events and interpret them, it is thus necessary to

study the relation between occurrence and co-occurrence, as well as the relation between occurrence and events. The matrices representing the event occurrence can also be normalized to decrease the influence of these two factors, the occurrence of pDCs and their co-occurrence. To suppress the first one, the occurrence of an event involving a pDC A and a pDC B is divided by the sum of occurrences of pDC A and pDC B. The influence of the co-occurrence on event is eliminated by dividing the number of events involving pDC A and pDC B by the co-occurrence of pDC A and pDC B. Nevertheless, these normalizations must be applied with caution because they are also susceptible to favor outliers. Indeed, even though an event appears one time in one subject, its weight can be increased by the normalization in regard to the other events if it involves pDCs with low number of occurrences or co-occurrences. It will thus appear as significant while it is not. Therefore, a further analysis of the influence of occurrence on co-occurrence and on number of events is required and is performed in Chapter 7.

5.8   Summary

This Chapter has presented the different steps of the workflow developed in this thesis, from the preprocessing to the final characterization of the interactions between the detected dynamic RSNs.

The final output of the workflow are pDCs. Each pDCs is characterized by its constitutive sDCs occurring in individual subjects. The sDCs change their footprints over time with a relative stability and they have defined periods of activity. During a period of activity, at each time step it is equivalent to a community of nodes on the brain surface with an individual footprint. The set of individual footprints finally forms a dynamic footprint and associated with the timelines, they characterize a sDCs. The properties of a pDC depends on its sDCs.

Owing to the occurrence of sDCs and their respective timelines, the global co-occurrence of the pDCs can be established. The significance of this co-occurrence needs to be evaluated to estimate if it holds information or if it is just a random co-occurrence.

Moreover, events are expanded to the population-level. If they are numerous enough, they are analyzed to identify recurrent patterns. As for the co-occurrence, the estimation of their significance is necessary.

Our methodology provides thus dynamic RSNs represented by pDCs and it allows to examine interactions between them. Few methods need to be define for the validation of the results before their presentation.

Methods for the Validation of the Results

Several methods are developed to validate the results. The main validation is performed by testing the repeatability by dividing the population into subsets but it is also possible to analyze specific properties and verify if the initial requirements are still met at the population level.

6.1 Properties of interest for the validation of the Population-level Dynamic Communities (pDCs)

To validate the identified pDCs, some simple characteristics require need to be examined to evaluate the robustness of the identified dynamic networks.

6.1.1 Repeatability across subjects

The number of subjects where the pDC is identified is an estimation of the consistency of the pDC across population. Indeed, the more it is present among subjects, the more its existence as a population wide dynamic functional network is demonstrated. The minimal number of subjects necessary to declare a set of sDCs as a pDC is determined by the lowest degree for center d_{min} (see Section 5.6). Here it is fixed to 50. The numbers of subjects where the pDC is present can be lower than the lowest degree for center because the node corresponding to the cluster center of the pDC can be linked to several sDCs from the same subject and only one of them is finally attributed to the cluster. This observation also demonstrates the necessity to decrease as much as possible the redundancy among sDCs since it obligatory generate a redundancy in pDCs, a level where the interactions are also analyzed.

6.1.2   Occurrence within the population

The occurrence of a pDC, that is to say the number of time frames it is active across the whole population, is linked to the number of subjects where this pDCs is active. However, the latter does not entirely determine the lifetime of the DC since it is possible that subjects present high difference of activation duration between the pDCs. Thus, a pDC can be present in a lot of subjects but for very tiny moments. Furthermore, the occurrence of a pDC is a metric of interest since it is also related to the number of events the pDC is involved. Indeed, if a pDC is active in a higher number of time frames, it may increases the probability for higher number of events since there are more time frames for them to happen.

6.1.3   Stability of the dynamic footprint

Another metric of interest is the stability of the pDCs (see Section 5.5.1 about pruning for the definition of stability of a DC). We previously explain that we introduce the spatial variation of the pDCs across time but that a relatively high stability must be preserved to enable the differentiation of the pDCs and prevent their agglomeration into an inappropriately low number of pDCs. Therefore, the final pDCs must present a relatively high similarity to testify they have been constructed in an appropriate way.

Therefore, these three properties must be evaluated to estimate if the pDCs identified are appropriate representatives of the functional connectivity at a population level. The occurrence is of particular interest because it can more or less influence the detection of interactions between pDCs.

6.2   Freezing the Occurrence within the population

In order to establish to what extent the occurrence determines the co-occurrence or the number of event, an experiment where the occurrence of pDCs is fixed at the population level is designed. The co-occurrence and the number of events can thus be analyzed independently from the occurrence. It is realized by generating sets of synthetic subjects via the process illustrated by Figure 6.1, here with only 3 subjects. To form one population of 200 synthetic subjects, the sDCs constituting one pDC are randomly re-attributed to subjects presenting this pDC in the real population, and it is performed for each pDC. This random process is repeated to form several populations of synthetic subjects. The number of possible combinations is very high and in the following experiments, the basic number of synthetic population is equal to 1000. The total occurrence of each pDC is thus conserved across population via this operation and the occurrence is thus fixed.
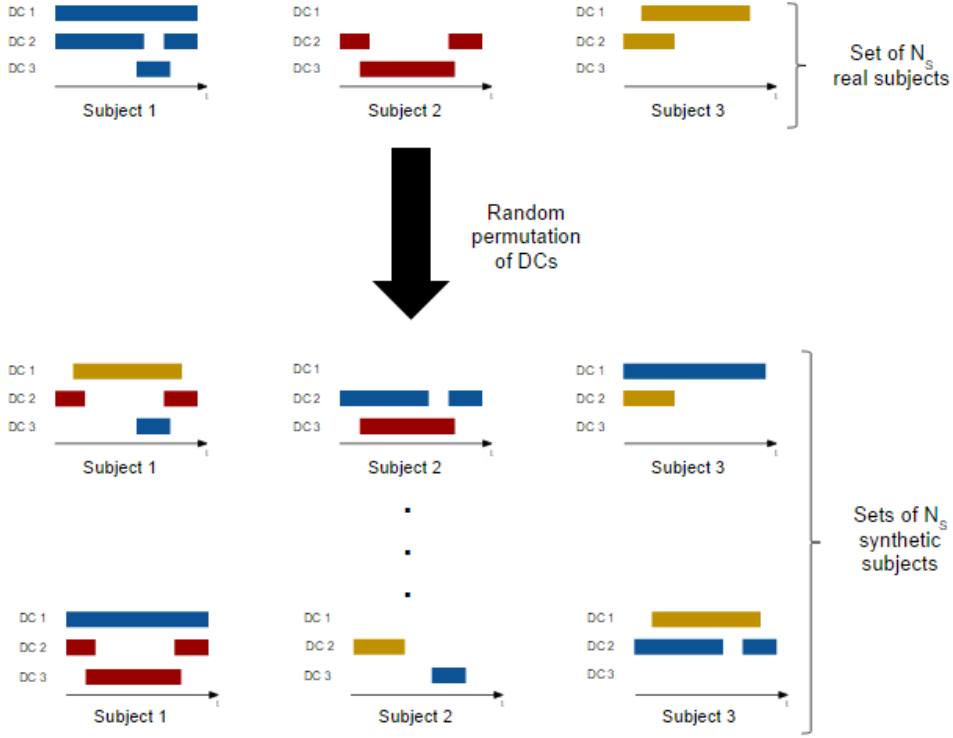
Figure 6.1: Illustration of the composition of sets of synthetic subjects from $N_S = 3$ real subjects.

6.3 Analysis of the correlation of the Subject-level Dynamic Communities (sDCs)

The identification of the pDCs is based on the detection of relevant sDCs. This detection is performed by an analysis of the correlation between signals from parcels of the brain surface. With time courses of the sDCs and their varying footprint, it is possible to estimate which brain parcels were considered as correlated during the community tracking at each time t for the initial fixed window size w_0, here equal to 80 time points. Figure 6.2(a) provides a support to illustrate the way the correlation is reconstructed from the sDCs. By applying the steps of the approach described along this chapter on rest-fMRI signals, the footprint of a sDC is identified at each time step it is active. Therefore, after the post-processing steps, it is possible to reconstruct the evolution of the correlation inside the final footprint of the sDC since for each parcel, the fMRI signal is still available. The relevancy of these sDCs is thus estimated.

The correlation can be reconstructed for different footprints of interest of the sDC. First of all, the correlation of a footprint is the averaged correlation between the parcels of the footprint and it is named the correlation strength. To have the more exact information

concerning the correlation strength of the sDC, the correlation strength must be calculated for the "dynamic" footprint that may vary at each time t. However, when the sDC is not active at time t, no footprint corresponds to this sDC at time t and the "dynamic" correlation is equal to zero. In order to obtain information about what happens to the parcels of the sDC during a period of death, it is possible to calculate the correlation of the footprint from the last-living frame, the one that would have to match another community footprint if the DC survives. The correlation strength of the next-living frame, the one where the footprint of the DC corresponds to this community footprint that has been match to the previous last-living footprint for the re-birth, is also relevant to access the behavior of the sDC when it is not active. The association of this three correlation strengths provides a complete visualization of the dynamic evolution of the correlation inside a sDC and testify their accuracy. Furthermore, it is possible to plot for every time point the correlation strength of the averaged footprint. It allows to state its significance as a static representation of the dynamic networks.

The access to the dynamic footprints and their corresponding signals also enables to see the influence of the size reduction of the window on the correlation strength of an entire footprint. The developed approach is the same as the one previously described to reconstruct the correlation from the footprints of the sDCs, except that the correlation is calculated on a smaller number of time points, as explained in Figure 6.2(b). A window is available at each time point since it is shifted point per point and the tapered window gives the highest weight to the correlation of its center point. Therefore, the footprint at time t is constructed owing to the correlation measured in the window $[t - \frac{w_0}{2}; t + \frac{w_0}{2} - 1]$. To observe the influence of the window size, the correlation strength is calculated the same way as in Figure 7.12(a) but instead of $w_0 = 80$ with a shift of 1 time point, it is calculated with for $w < w_0$ with a shift of 1 time point as well. The correlation strength derived from sDC footprint at time t is then calculated in the window $[t - \frac{w}{2}; t + \frac{w}{2} - 1]$. It is difficult to take a number of time points w lower than 50 because, as explained in Section 2.4, the range of interest for the dynamic information is included in $[30,300]$ seconds and with $TR = 0.7$, a window containing 50 time points has a length of $0.5 \times 0.7 = 0.35$ seconds. Therefore, it is already very close from 30 seconds.

6.4 Summary

The three methods described in this chapter enable to validate the results. The properties defined in Section 6.1 establish to what extent the detected pDCs are robust and reliable to represent the FC. In Section 6.2, the method developed to fix the occurrence of pDCs inside a population enables to solve the problem of dependence on occurrence of the relation between the different pDCs. Finally, the Section 6.3 allows to estimate the influence of the length of the window during SWA at subject level in using again the correlation information. It also permits to visualize the output of the community detection and the community tracking from the point of view of the correlation, the fundamental metric of our representation.
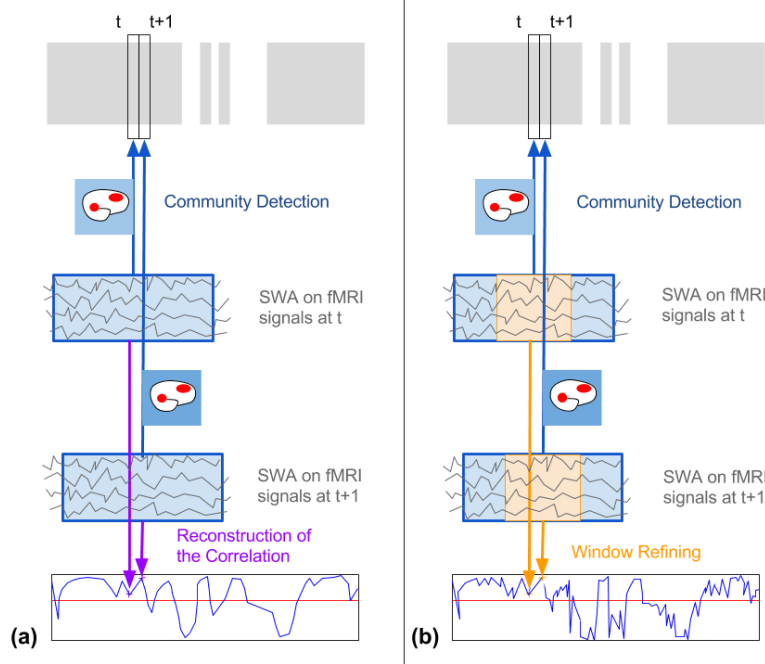
Figure 6.2: Methods for correlation reconstruction. The process is illustrated here for only two consecutive time frames. The first row represents the time course of one DC. For each time frame, the SWA and the following steps have extracted a footprint of the DC from the signal via a window here represented in blue on second and third rows. (a) The fourth row represents the correlation strength of the dynamic footprints reconstructed at each time t with the blue windows from the signals of each parcel of the footprint at time t. (b) Window refinement: the fourth row represents the correlation strength of the dynamic footprints reconstructed at each time t with the red windows.

CHAPTER

Experiments and Results

This chapter presents the results obtained with our representation. First the existence of pDCs is demonstrated and their properties are depicted. Then their co-occurrences are examined as a first insight into the relationship they can share and this aspect is next characterized by the analysis of different types of events occurring between pDCs. Finally, the relevancy of our approach, parameters and results is validated through different experiments.

7.1 Identification of Population-level Dynamic Communities (pDCs)

To determine if it exists pDCs in the functional connectivity, repeatability and other particular properties described in Section 6.1 must be testified for the identified DCs. This is why sDCs are tracked across the population. Several choices of parameters and approaches, on top of additional experiments for the validation of the latter, finally permits to obtain 6 different pDCs from a set of 200 subjects.

7.1.1 Footprints of the pDCs

The averaged footprints of the 6 pDCs thus detected are represented on Figure 7.1.

It is possible to compare these pDCs to static RSNs to approximate their function and also to demonstrate that the detected pDCs are not random distributions of parcels on the brain surface. Dice coefficients, a standard measure of result similarity [BM10], are calculated between the static RSNs identified by Yeo et al. in [YKS+11] and our pDCs are calculated. The Dice similarity coefficient d for two segmentations A and B is defined as [ZWB+04]:
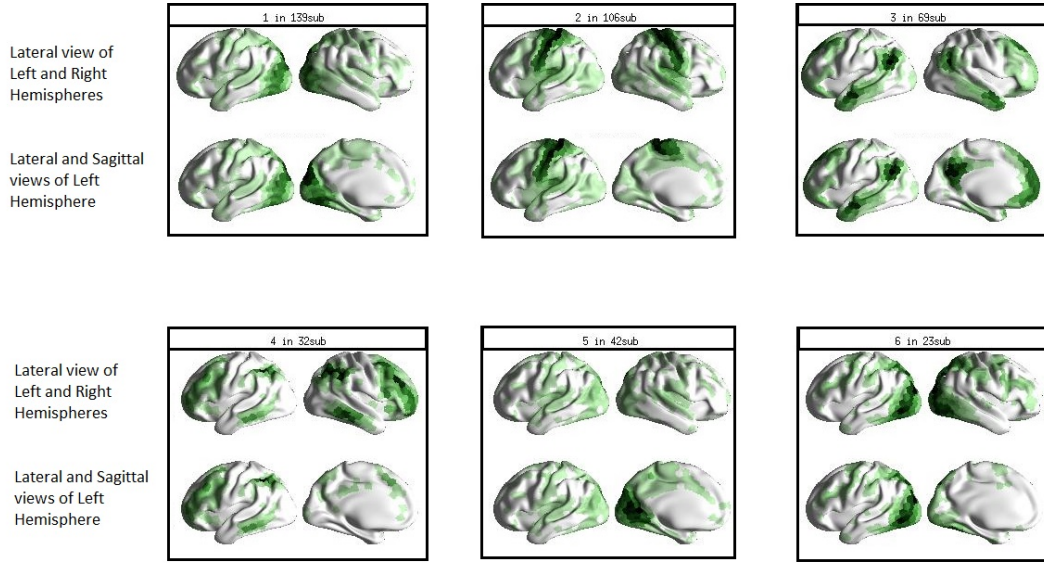
$$d(A, B) = 2\frac{A \cap B}{A \cup B} \tag{7.1}$$

Figure 7.1: pDCs detected in a set of 200 subjects. The intensity of the parcels on the brain surface are dependent on the number of subjects in which the parcels is member of the concerned pDC. The darker the parcel is, the more frequently it is present in the averaged footprint of sDCs constituting the pDC. Two views are proposed here : the lateral view of the pDC averaged footprint on both hemispheres and the lateral and sagittal view of the same footprint only on the left hemisphere. pDCs are labeled from 1 to 6 and the number of subjects in which they appear are specified.

It is a commonly used metric in medical imaging to evaluate spatial overlap and the reproducibility of a segmentation [ZWB+04]. The similarity matrix between the set of pDCs and the set of RSNs from Yeo et al. [YKS+11] is represented by Figure 7.2. For comparison of the footprints, the averaged footprint of the pDCs is simplified by considering only the parcels that are present in more than 10 subjects in order to compare a sharp enough footprint but in conserving the subject variability. Indeed, if the averaged footprint is made of all the parcels that are present at least one time in the sDCs constituting a sDC, it covers a very large proportion of the brain due to noise and subject variability. Moreover, it has to be noticed that on the contrary to Yeo et al.'s parcellation, the spatial overlap between pDCs is allowed. With the observation of the averaged footprints, a minimum of 10 subjects is a good compromise.

The similarity matrix permits to attribute a pDC to one or several static RSNs. A pDC remains however different from a static RSN since it has dynamic properties: it activates and deactivates and it slightly changes its footprint across time. The link between the averaged footprint of the pDC and the RSN permits to attribute a possible function to the pDC and also validate our approach that reveals dynamic RSNs represented as DCs similar to the static RSNs. Thus, by combining Dice coefficient of Figure 7.2 and
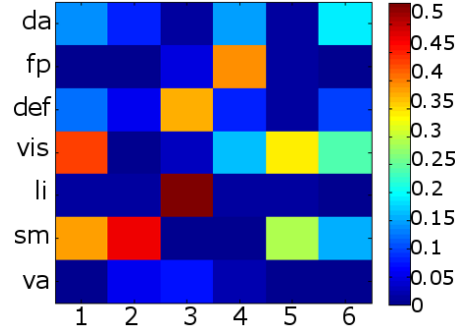
Figure 7.2: Similarity matrix gathering the Dice coefficients between the pDCs and static RSNs. The static RSNs are : dorsal attention (da), frontoparietal (fp), default (df), visual (vis), limbic (li), somatomotor (sm) and ventral attention (va). The pDCs labels corresponds to those defined by Figure 7.1.

information about parcel occurrence in pDCs from 7.1, it is possible to state that :

- pDC1 present in 139 subjects is similar to the visual RSN.

- pDC2 present in 106 subjects is similar to the somatomotor RSN.

- pDC3 present in 69 subjects is similar to the limbic and default RSNs.

- pDC4 present in 32 subjects is similar to the frontoparietal RSN.

It is more difficult to associate pDCs 5 and 6 to RSNs because the Dice coefficients are lower. Both of them have their highest similarity with the visual network.

7.1.2 Properties of the pDCs

The properties of interest serve as evaluation of the accuracy of pDCs and are also helpful for the analysis of their interactions, as explained in Section 6.1. The number of subjects where each pDC is identified, the stability of each pDC and the number of time frames within the population where each pDC is active are summarized in Table 7.1.

In the previous paragraph, it has already been demonstrated that the footprints of the pDCs 1, 2, 3 and 4 are more accurately defined than those of pDCs 5 and 6. They particularly match the static RSNs while the pDCs 5 and 6 cannot be attributed to one RSN and share similarity with other pDCs, particularly the first one. The first row of Table 7.1 shows that they are also present in less subjects than pDCs 1, 2 and 3. However, it does not mean it does not hold information since they are present in more than 10% of subjects after the post-processing. Moreover, the pDC 4 is present in less subjects than the pDC 5 but it has been detected first because it has an higher degree in the similarity graph. Therefore, the significance of the number of sDCs in a pDC must be

93

Properties	pDC1	pDC2	pDC3	pDC4	pDC5	pDC6
Number of sDCs	139	106	69	32	42	23
Alive time	35865	26742	16932	7623	8009	4480
Stability	0.8434	0.8342	0.8304	0.8239	0.8280	0.8384

Table 7.1: Properties of interest of individual pDCs.

considered in association with the label of the pDC that also access the weight of the pDC regarding the complete set of sDCs, without consideration for the membership to a subject. The data for the occurrence of the pDCs in the second row of the Table 7.1 are here correlated to the number of sDCs constituting each pDC in the first row.

The last row of the Table 7.1 presents the averaged stability value of each pDC across sDCs being part of it. The values are rather similar for each pDC and are relatively high, particularly compared to the stability threshold defined for pruning (0.2). It demonstrates that the expansion of sDCs into pDCs does not decrease the stability and that the final pDCs are stable.

It is therefore possible to conclude that these evaluations of the properties of pDCs confirm that the identified pDCs meet the requirements to represent the dynamic functional connectivity. To check if they are representative of the population, their repeatability is interrogated.
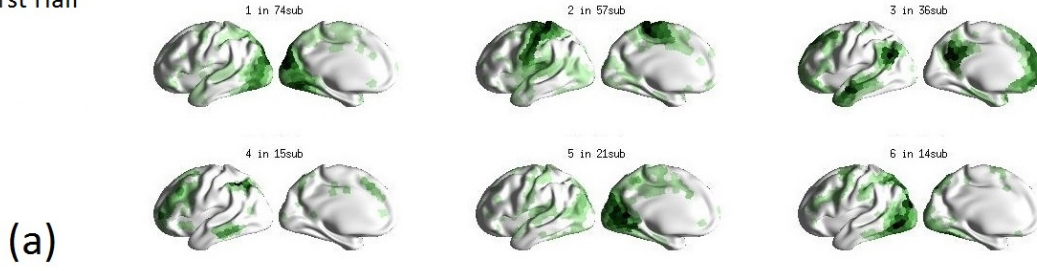
7.1.3 Repeatability of the pDCs

In order to test the repeatability of the results, the set of subjects is separated into 2 random sets of 100 subjects and the previously described experiments are independently applied to them with the same parameters. The repeatability of the experiments is thus quantitatively estimated via cross correlation.

The subjects where the pDCs are present are known and therefore, their characteristics can be directly evaluated within subsets of the population without the repetition of the population-level tracking. However, owing to preliminary tests realized first on 20 and then 100 subjects for the construction of the workflow, this tracking has been performed independently on small subsets of subjects. From a qualitative point of view, the results obtained with these smaller populations are as similar as possible from the results with 200 subjects. Indeed, the footprints corresponds to the same RSNs but they are present in a lower number of subjects.

The functional connectivity measured in this thesis is mainly represented by the footprints of the pDCs. To evaluate their repeatability, the averaged footprints are recalculated over each subset of subjects and they are represented in Figure 7.3. It is already possible to compare the number of subjects the pDCs are part of. Although it is not an exactly uniform repartition, this split of the population used for the repeatability analysis is appropriate because the pDCs are almost equally distributed among the two subsets.

First Half

1 in 74sub   2 in 57sub   3 in 36sub

4 in 15sub   5 in 21sub   6 in 14sub

(a)

Second Half

1 in 65sub   2 in 49sub   3 in 33sub
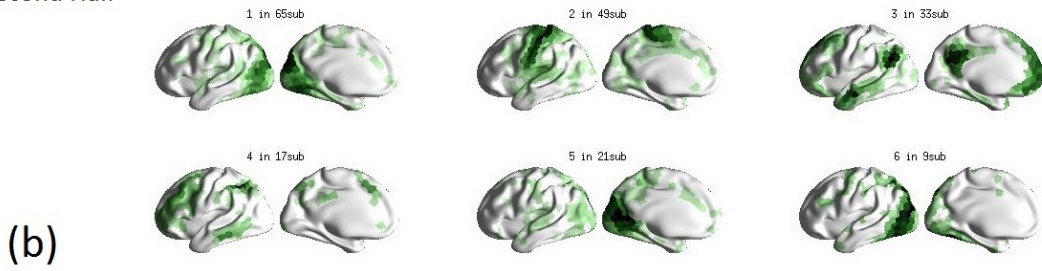
4 in 17sub   5 in 21sub   6 in 9sub

(b)

Figure 7.3: Averaged footprints for 2 subsets of the population. (a) First half of the population. (b) Second half of the population

Moreover, in order to compare the footprints between the two subsets, the Dice coefficients between averaged footprints of each pairs of pDCs from different subsets are calculated, without selection of the parcels. They are represented in Figure 7.4.
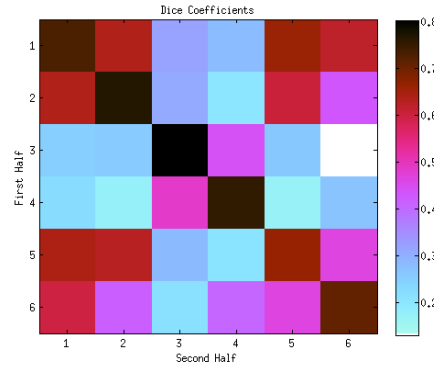

Figure 7.4: Dice coefficients between the pDCs of the first half of the population and the pDCs of the second half of the population. All the parcels of the averaged footprints are used.

The mean of the Dice coefficients between corresponding pDCs in the two subsets, that is to say the mean of the diagonal of the matrix represented in Figure 7.4, is equal to 0.7373 with a minimum dice coefficient for pDC 6 equal to 0.6622. If the Dice coefficients are

calculated with pDCs averaged footprints made of parcels present in at least 5 subjects, the mean is equal to 0.8603 with a minimum dice coefficient for pDC 6 equal to 0.7749. Therefore, it confirms the repeatability of the pDC footprints between the different subsets because when the core of the footprints are addressed, the similarity is still higher and the weakest pDC is the one with the lowest number of subjects, that is to say the less significant.

Moreover, the number of time frames where the respective pDCs are active, their occurrence, is also repeatable. The correlation between the occurrence of each pDC in each subset is calculated and the Pearson's correlation coefficient is equal to 0.98769 with a p-value equal to 0.0002265. The repeatability of the occurrence is consequently demonstrated.

Therefore we can conclude that the pDCs identified with our methodological approach are repeatable and therefore robust. A major part of the aims defined in Section 1 are thus validated.

7.2 Co-occurrence of the Population-level Dynamic Communities (pDCs)

In the previous section, it has been demonstrated that robust and repeatable pDCs exist in brain functional connectivity. They are active during finite periods of time and some pDCs are more often active within the population than others. It is possible that different pDCs have temporally overlapping periods of occurrence that indicate a *co-occurrence.* The analysis of this co-occurrence is interesting in the sense that the pDCs that co-occur may interact and be components of a state of activation. On the other hand, pDCs that never co-occur can be considered as complementary.

7.2.1 Co-occurrence patterns

It is therefore necessary to detect the couples of pDCs that often co-occur. There are several points of view to define the co-occurrence: it can be independently calculated in each subject as the number of time frames where both pDCs are active and then averaged with the number of subjects, or it can be considered as the absolute number of time frames where the couple co-occurs in the complete population. The second definition is selected because the whole analysis is conducted at the population level and the first definition is more influenced by possible spurious co-occurrence happening in single subjects. The co-occurrence is therefore characterized by the total number of time frames where both pDCs are simultaneously active.

Figure 7.5(a) is the upper triangular matrix of co-occurrence for each couple of pDCs. The color of its coefficient in row i and column j corresponds to the number of frames where DC i and DC j are active according to the color bar on the right. The pairs of pDCs with the darkest colors are remarkable co-occurrence patterns. Figure 7.5(b) is a bar plot representing the occurrence, that is to say the number of time frames where

each pDC is active and it corresponds to the data of Table 7.1. If we compare the top-co-occurring couples of pDCs with the top-occurring pDCs, it is obvious that they are corresponding. Indeed, the three couples with the highest number of shared time frames are DC1-DC2, DC2-DC3 and DC1-DC3. The significance of the co-occurrence can therefore be questioned, but it is first necessary to verify if this co-occurrence configuration is repeatable.
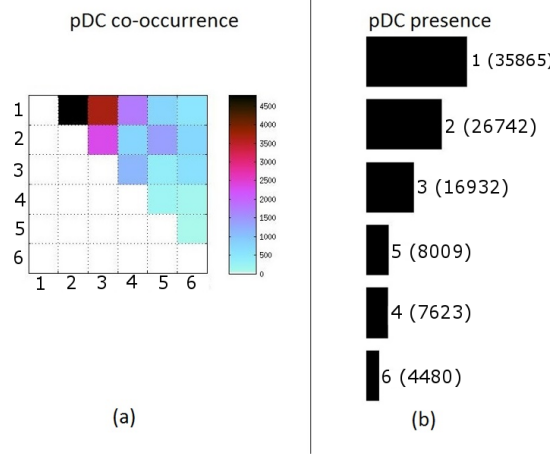

Figure 7.5: (a) Co-occurrence matrix whose coefficients visualized by a color scale are the number of time frames where the concerned couple of pDCs co-occurs. (b) Bar plot representing the number of time frames across the population where each pDC is active.

7.2.2 Repeatability of the co-occurrence

The repeatability of the co-occurrence is demonstrated in the same manner as the repeatability of the pDCs in Section 7.1. The population is split into two subsets of 100 subjects and the co-occurrence of each pair of pDCs is computed within each subset. The co-occurrence matrices are compared and the correlation between the subsets is calculated. The results are presented in Figure 7.6.

The Pearson's correlation coefficient is equal to 0.9751 with a p-value equal to 8.00910^{-24} that demonstrates the significance of this correlation. In both subset the three top-co-occurring patterns are again DC1-DC2, DC1-DC3 and DC2-DC3. Therefore, it is possible to conclude that the co-occurrence is repeatable and it is now necessary to establish to what extent it is significant.

7.2.3 Significance of the co-occurrence

In order to know if the co-occurrence reveals something else than the occurrence of the pDCs, it is interesting to study the influence of occurrence of pDCs on co-occurrence. The occurrence is thus frozen as explained in Section 6.2 and 1000 synthetic populations are
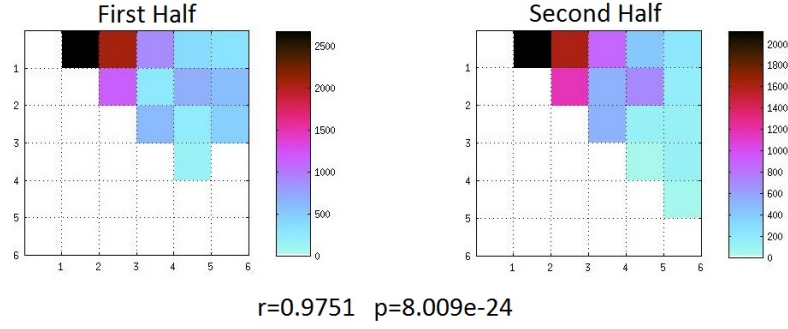
Figure 7.6: Upper triangular matrices for the two subset of the population. r is the Pearson's correlation coefficient between the two sets of co-occurrence in each subset and p is the corresponding p-value.

generated. It is thus possible to compare the distribution of the co-occurrence at subject level for each pair of pDCs between the 200 original subjects and $1000 * 200 = 200000$ synthetic subjects.

The comparison is realized owing to statistical tests. First of all, the distributions of the co-occurrence of pDCs pairs are not normal according to one-sample Kolmogorov-Smirnov test and data visualization, so it is not possible to apply a two-samples t-test to compare the co-occurrence distributions between the real subjects and the synthetic subjects. Therefore we use a two-sample Kolmogorov-Smirnov test which tests the null hypothesis H_0 that the distribution of co-occurrence in subjects from synthetic populations and real subjects comes from the same continuous distribution with a level of significance equal to 0.05. H_0 is rejected for two couples of pDCs:

- pDC 1 and pDC 4 with a p-value equal to 0.020161.

- pDC 2 and pDC 4 with a p-value equal to 0.023369.

The multiple statistic correction with a q-value equal to 0.3 conserves these two pairs of pDCs whose co-occurrence is independent from the occurrence. The q-value is however relatively high and it means that the probability that these two results are actually false positive is equal to 30% and it has to be taken into account.

The rejection of H_0 for these two pairs of pDCs means that their co-occurrence is not a random association and is therefore not just an effect of the occurrence, as well as its event. The box-plots of the distribution of real and synthetic subjects are represented in Figure 7.7. It shows that the co-occurrence of pDC 1 and pDC 4 happens globally more often than random while the co-occurrence of pDC 2 and pDC 4 tends to happen less often than random.
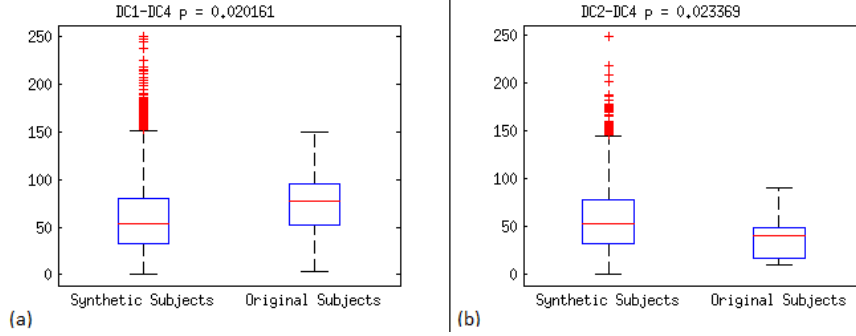
Figure 7.7: Box-plots of the distributions of number of frames where a couple of pDCs co-occurs in one subject. The ordinate represents the number of frames. (a) Left box-plot: Co-occurrence of pDC 1 and pDC 4 in 200 000 synthetic subjects. Right box-plot: Co-occurrence of pDC 1 and pDC 4 in the 200 original subjects. (b) Left box-plot: Co-occurrence of pDC 2 and pDC 4 in 10 000 synthetic subjects. Right box-plot: Co-occurrence of pDC 2 and pDC 4 in the 200 real subjects.

7.3 Analysis of Events at Population Level

Among the events detected with the sDCs at the subject level, a part of them survives at the population level if they involve sDCs associated to a pDC. The detection of the events is performed at the subject level but their analysis is made at the population level to increase its statistical weight. The different steps of the construction of pDCs from the fMRI signals are designed to detect as many events as possible in order to obtain the highest amount of data and by this way, to increase the accuracy of the analysis. Indeed, it is necessary to extend the dataset because noise remains in the pDCs and it is therefore also present in the interactions identified through events of high and low interaction levels.

7.3.1 High Level of Interaction

In spite of all the effort implemented to conserve *splits* and *merges* that characterize spatial and temporal interaction between two or more DCs, the pruning step after the community tracking where they are identified suppress a large part of them. It is the cost for the denoising essential to obtain stable and consistent pDCs. The low number of remaining events can either mean that most of *splits* and *merges* results from noise and does not really happen in the brain, or that the pruning is too hard and suppress too much information. Several thresholds have been tested for the minimal lifetime and stability but it is difficult to choose a lower threshold than 10 time frames and 0.20 for stability if the priority is the noise suppression. Therefore, among the 200 subjects, only 10 *splits* remains at the population level and only one appears in two different subjects. The others occurs one time in a unique subject. Therefore, these high level of interaction events have a too low statistical weight to produce a reliable model for the interaction of
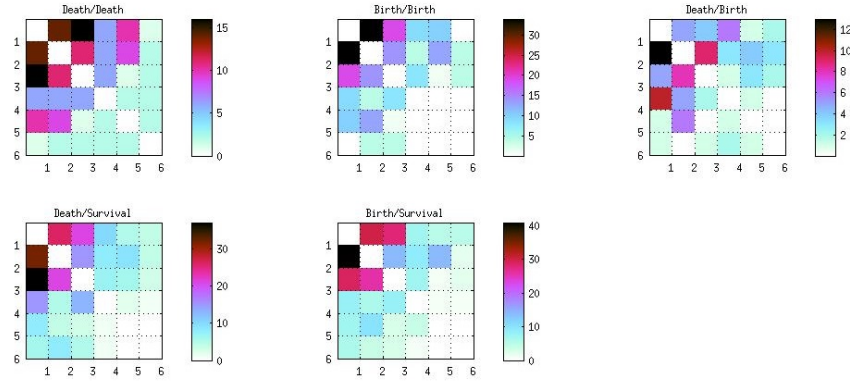
Figure 7.8: Event matrices without scaling for *death-death*, *birth-birth*, *death-birth*, *death-survival* and *birth-survival* events. The coefficients of the matrices represents the number of occurrences. Rows and columns corresponds to pDCs.

pDCs.

7.3.2   Low Level of Interaction

The co-occurrence can also be considered as the lowest level of interaction and events used to represent the low level of interaction, only based on temporal information, are at the end more or less a characterization of meaningful co-occurrence. Indeed, it highlights for instance the pDCs that often co-occur with a similar pattern via *death-death* and *birth-birth* events and those which are complementary via *death-birth* events. At this lower level of interaction, the events are more numerous and some of them are recurrent across subjects. The matrices of occurrence described in Section 5.7.3 are represented in Figure 7.8.

The *death-death* and *birth-birth* events are symmetric by definition, but it is visible that the other matrices are almost symmetric too. A possible interpretation is the influence of occurrence and co-occurrence of DCs on the number of events they are involved in. Indeed the co-occurrence is a symmetric properties since if DC *A* occurs at the same time as DC *B*, it also means that DC *B* occurs at the same time as DC *A* and it has already be explained that occurrence more less influences the co-occurrence of particular couples of pDCs.

Furthermore, the matrices of Figure 7.8 underline that for each type of event, some of them occurs more frequently. Therefore, re-occurring patterns exist at this level of interaction. The top-occurring events often involve top-occurring pDCs 1, 2 and 3 but, for instance, the second top-occurring event for *death-birth* events happens between pDC 1 and pDC 4. It is particularly interesting because the co-occurrence of this couple of pDCs is demonstrated as significant in Section 7.2.3. Therefore, it is an example of significant event that can be of interest for neuroscientists. However, the repeatability of the events must be established in order to validate their occurrence.

Events	r	p
Death-Death	0.73019	4.3164e-7
Birth-Birth	0.87819	1.9772e-12
Death-Birth	0.49947	1.9265e-3
Death-Survival	0.83758	1.8962e-10
Birth-Survival	0.8494	5.7801e-11

Table 7.2: Correlation between the number of events in two subsets.

7.3.3 Repeatability of the events

The population is once again separated into two subsets to perform a test of repeatability on the lower level of interaction. Indeed, there are not enough events characterizing the higher level of interaction to process a relevant analysis of their repeatability.

For each of the two subsets of the population, the different kinds of events are counted. The matrices of occurrence are build and the correlation between those corresponding to the same type of events are calculated. The results are summarized in the Table 7.2.

It demonstrates that we obtain a high repeatability between the two subsets of the population since the computed correlation is close from 1. The *death-birth* events have a lower correlation value but it is also the type of event with the lower number of occurrence and it is therefore the most affected by the reduction of the number of subjects in the population. This experiment validates the events occurrence. Their independence regarding the occurrence of the pDCs can be tested to evaluate their significance.

7.3.4 Evaluation of the significance of the events

To eliminate the influence of the occurrence on the number of events, the simplest approach consists in normalizing, for all couple of pDCs (A, B), the number of events between pDC A and pDC B by the sum of the respective occurrence of A and B. The normalized event matrices thus generated are depicted in Figure 7.9. The events are more distributed among pDCs than in Figure 7.8 and some of them tends to show a higher recurrence, such as *death-birth* between pDC 1 and pDC 4 or the *death-survival* between pDC 4 and pDC 3. To validate this approach, the repeatability is analyzed the same way as in Section 7.3.3. The correlation and p-values thus obtained are summarized in Table 7.3.

When we compare Table 7.2 and Table 7.3, the correlation is globally lower when the number of events is normalized by the occurrence. It can be interpreted in different ways, such as the non-repeatability of the events and therefore their non-significance, or it can also indicate that this normalization is not appropriate. Indeed, it favors the events with a small number of occurrences and therefore, an event that happens few times and in few subjects, but with pDCs of short lifetime, has finally a relatively high weight in the related event matrix compared to the other events. It can reflect the true configuration
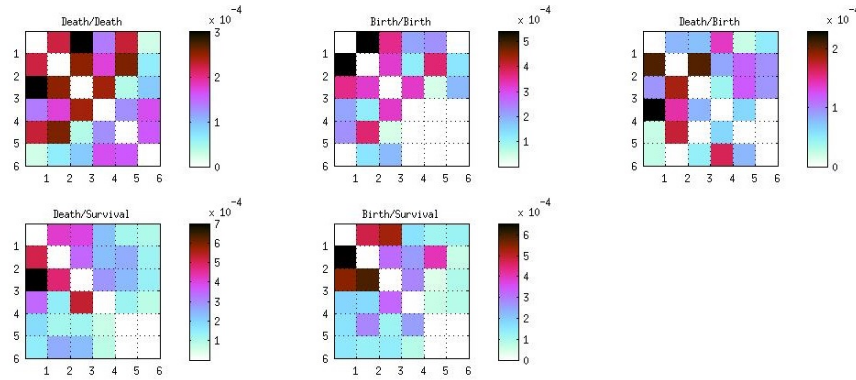
Figure 7.9: Event matrices normalized with occurrence for *death-death, birth-birth, death-birth, death-survival* and *birth-survival* events. The coefficients of the matrices represents the normalized number of occurrences. Rows and columns corresponds to pDCs.

Events	r	p
Death-Death	0.24857	0.14378
Birth-Birth	0.83152	3.3662e-10
Death-Birth	0.20703	0.22569
Death-Survival	0.62945	3.9108e-5
Birth-Survival	0.66994	7.8544e-6

Table 7.3: Correlation between the normalized number of events in two subsets.

of the brain activity but it is necessary to further explore the links between occurrence and event to establish the significance of the latter.

Number of events in synthetic populations with fixed occurrences

Sets of synthetic subjects with fixed occurrence are generated according to the method described in Section 6.2 to provide an evaluation of the significance of the events, independently from the occurrence of the pDCs they involve. These synthetic subjects are considered as 1000 different populations and the number of events are calculated across each population. For each possible event, it thus creates a distribution of their number, with a mean and a standard deviation. The z-score of the number of events for each event in the original population is calculated to evaluate to what extent the number of events is affected by the randomization of the sDCs. The results are represented in Figure 7.10 with matrices of z-scores.

It is visible in this figure that the number of positive and negative z-scores is balanced, and that some events have a particularly high z-score. With a two-tailed test and a maximum p-value equal to 0.05, that is to say a level of significance equal to 0.05, the significant z-scores are those whose absolute value is higher than 1.96. They are highlighted in
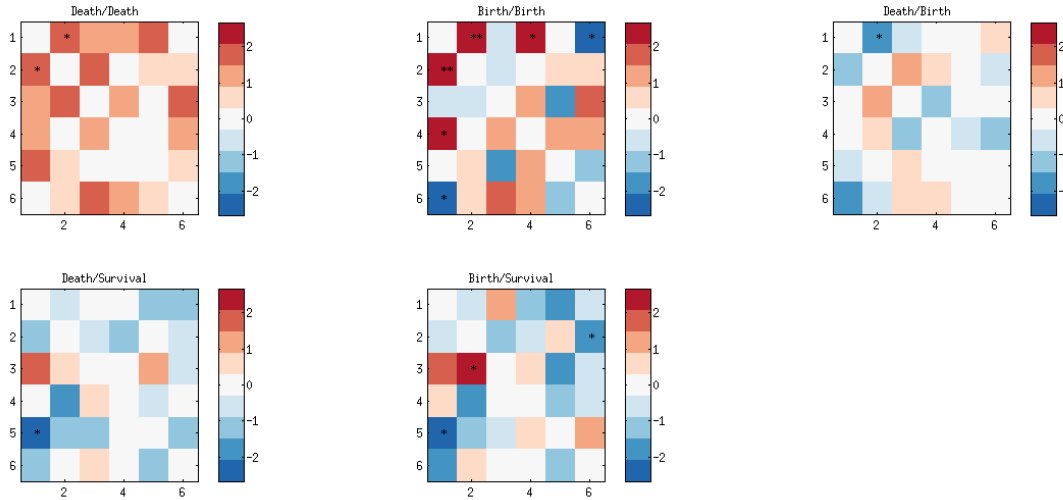
Figure 7.10: Matrices of z-scores between the distribution of number of events in synthetic populations and the number of event in the original population. An event with one star means that the z-statistic reject the null hypothesis and the event with two stars is the one surviving to the multiple statistic correction with a q-value of 0.05.

Figure 7.10 by stars and summarized with their z-score and corresponding footprint in Figure 7.11. Moreover, if a multiple statistic correction is applied at a q-value of 0.05, it remains one event labeled with two stars in Figure 7.10. This event is the simultaneous birth of pDC 1 and birth of pDC 2. It has a positive score and it happens consequently more often than random.

This experiment demonstrates that a certain number of events are more significant, admittedly with more or less confidence, but it demonstrates that they are holding information. It would be now interesting to observe for instance how their occurrence is influenced by neuropathologies.

7.4 Validation of the parameters

The results described above meet the requirements defined in Section 1. Indeed, a robust representation of the dynamic functional networks is constructed and a selection of events characterizes interactions in a significant way. Further experiments are however helpful to justify the choices made for the construction of this new representation that uses fMRI signals with low Signal-to-Noise Ratio and cannot supported by a ground truth because none exists.

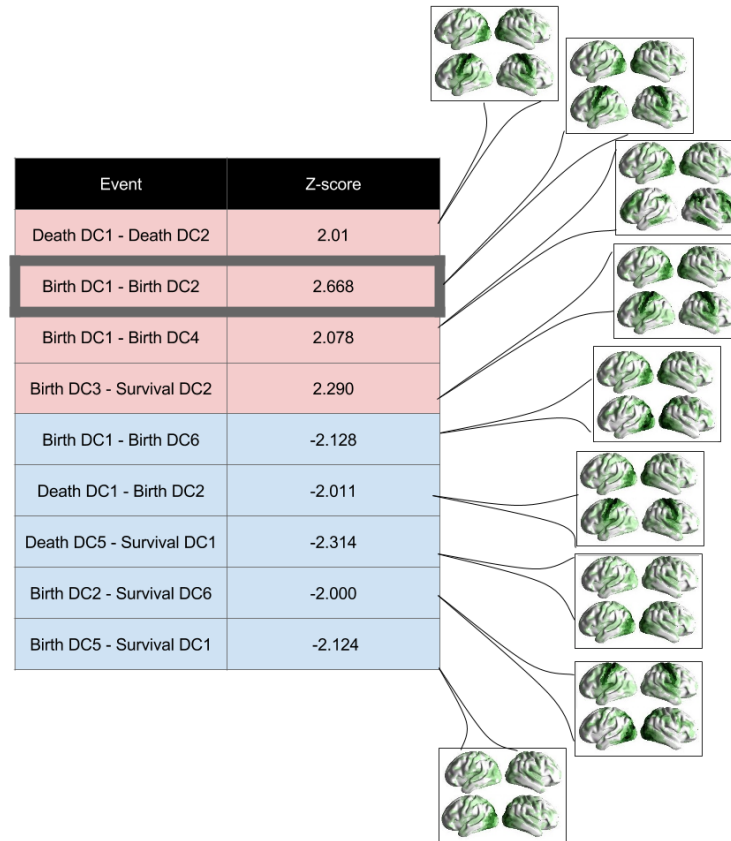| Event	Z-score
Death DC1 - Death DC2	2.01
Birth DC1 - Birth DC2	2.668
Birth DC1 - Birth DC4	2.078
Birth DC3 - Survival DC2	2.290
Birth DC1 - Birth DC6	-2.128
Death DC1 - Birth DC2	-2.011
Death DC5 - Survival DC1	-2.314
Birth DC2 - Survival DC6	-2.000
Birth DC5 - Survival DC1	-2.124

Figure 7.11: Table of the events declared as significant by a statistical test on their occurrence in synthetic subjects and in real subjects. The concerned events are indicated in the first row and their related z-score in the second row. The footprints of the involved pDCs are represented on the right of the table. The framed event is the one that survives after the multiple statistic correction with q-value equal to 0.05.

7.4.1 Post-refining of the Window Size to study the Consistency of Community detection

These experiments use the technique described in Section 6.3 to have a closer look to the correlation inside the final footprints of the sDCs.

Validation of the community tracking and importance of the time varying footprint

Firstly, reconstructing the time-varying correlation of the footprints when they are active enables to validate the community detection. In the graph used for the community detection, two vertices are connected if their correlation is higher than a correlation cutoff of 0.8, that is to say if they are highly correlated. However, all the points of one community do not necessarily have a pairwise correlation higher than 0.8 because all pairs of node in one community are not necessarily connected. The community detection algorithm aims to ensure a high correlation between parcels of the community and it is possible to check how high is this correlation with the first part of this experiment. Figure 7.12(a) represents with gray blocks the time course of each sDC for one subject with the associated averaged correlation, also named here correlation strength, within their dynamic footprints in blue. If the sDC is not active at time t, the sDC has no dynamic footprint at time t and the "dynamic" correlation is equal to zero. To visualize the information during a period of death, the correlation of the footprint from the last-living frame, the one that would have to match another community footprint if the DC survives, is plotted in green and the correlation of the next-living frame, the one where the footprint of the DC corresponds to this community footprint that has been matched to the previous last-living footprint for the re-birth, is plotted in magenta. Moreover, from observation on several subjects, a correlation threshold for the activation of the sDCis identified. Even though, it is an arbitrary value, it provides a better visualization for the comparison of the activity peaks. It is approximately equal to 0.6 and it is here represented by a red horizontal line. Figure 7.12(a) demonstrates that the parcels from sDCs footprints are globally more correlated when sDC is active. Indeed, the last-living and next-living footprint correlation respectively decreases and increases during the inactivity periods while the dynamic footprints have a higher correlation during activity. Therefore, the community detection results in reliable pDCs after the community tracking.

The same tool is used to visualize the correlation of the "static" footprint of the sDCs, that is to say their averaged footprint which is used to build the similarity matrix in consolidation and tracking across the population. Figure 7.12(b) represents for the same subject as Figure 7.12(a) the evolution of their correlation in parallel with their time course. It highlights the importance of the time-varying footprint since the static footprint is far more smoother and globally higher than the dynamic footprint during the period of activity. However, their evolution are similar and it is possible to observe that the last-living and next-living footprints share correlation variations with the static footprints,

even if both of last-living and next-living footprints correlation are not correlated together most of the time.

To conclude, this experiment demonstrates that the averaged sDC footprint is not an accurate representation for the dynamic properties of the pDC. However, the dynamic community detection is validated here since the communities it extracts from the data conserve their accuracy after the community tracking and the post-processing, as it is demonstrated in Figure 7.12(a).

Post-refining of the window size

The selection of a window length for the SWA is one of the weakness of our representation (see Section 2.4.1). In order to verify the influence of the window size on our results without repeating the SWA which is computationally expensive (45 minutes per subject), it is possible to use the data from the already constructed sDCs to observe the correlation fluctuations at a smaller time scale, as explained in Section 6.3. A refinement to 50 time points may present high peaks when the correlation strength obtained with 80 time points is subtracted to it. A peak can for instance be an outlier whose influence had been smoothed by the use of more time points, but it can also be a significant peak of activation that has been missed due to the same smoothing.

Figure 7.12(d) represents the absolute difference of correlation strengths between a reconstruction with 80 time points and 50 time points. It does not exist major differences that could have revealed missed activation. Indeed, when we compare Figures 7.12(c) and 7.12(a) that respectively represent the correlation strength for a window length of 50 and a window length of 80, it is visible that the variations of the three different correlation strengths are globally the same. However, it is also remarkable that they are sharper for 50 time points.

The significance of this experiment is limited since it is built on results obtained with a larger time window but if major differences would have existed between dynamic networks detected via a SWA with a window length of 80 time points and dynamic networks from a SWA with a window length of 50 time points, they would have yet been revealed by this experiment. The constructed sDCs and the associated pDCs are therefore still relevant at a finer temporal scale and the use of 80 time points is thus validated.

7.4.2 Similarity Threshold for the sDCs Tracking

In order to select the most accurate similarity threshold for community tracking at subject level, we observe how the number of mutual topics (Figure 7.13(a)), the number of DCs (Figure 7.13(b)) and the number of events (Figure 7.13(c) and (d)) evolve when k varies between 0.4 and 0.8. These elements are of particular interest as explained in Section 5.4.4 where the concept of mutual topics is also defined.

The results of these experiments are presented on Figure 7.13. The community tracking is tested for $k = \{0.4, 0.5, 0.6, 0.7, 0.8\}$ on 10 different subjects, and the mean of the four
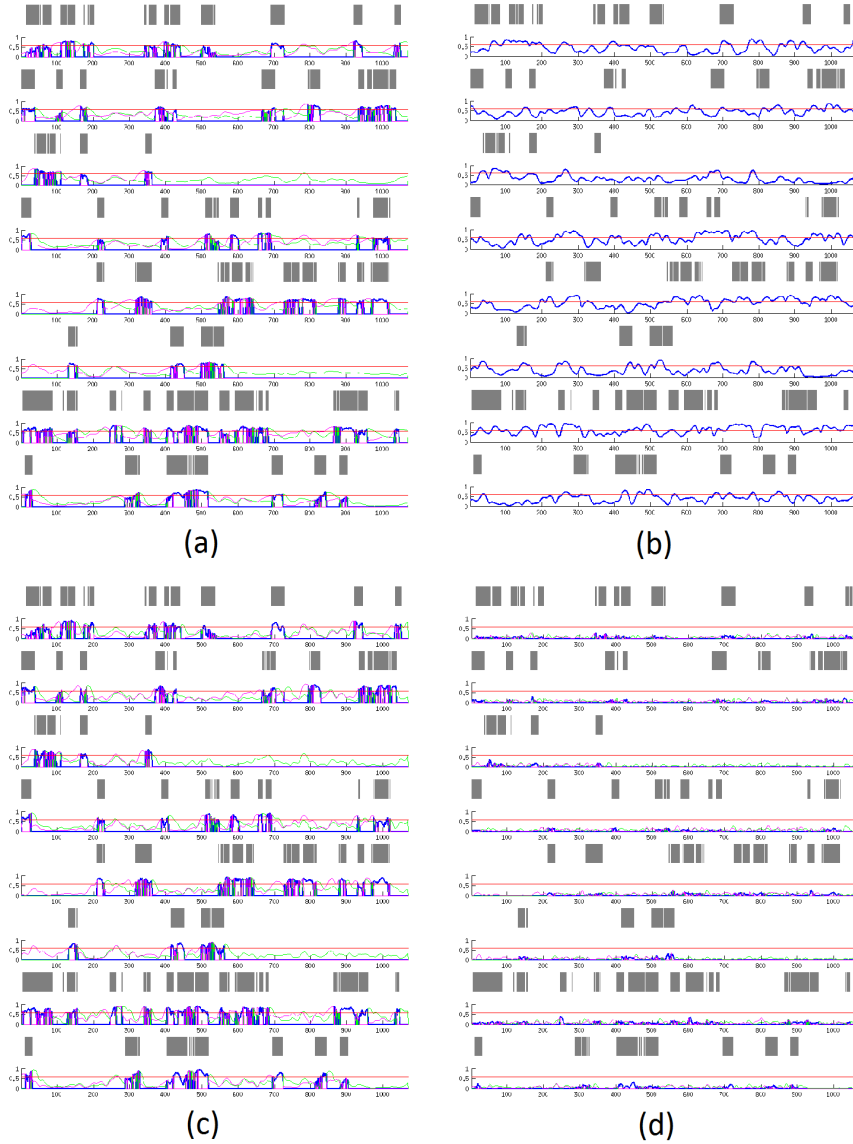
(a)

(b)

(c)

(d)

Figure 7.12: Time courses of pDCs represented by gray blocks aligned with the correlation of dynamic pDC footprints for one subject. The subject has 8 pDCs and therefore 8 sets of plots. The tapered correlation strengths are plotted between 0 and 1 for each time window. The x-axis is the time axis where one unit of time corresponds to one time frame since windows are shifted with a step of one time frame. The red line is arbitrary located at a correlation of 0.6 and aims to better visualize the correlation variations. (a) Dynamic correlation strengths reconstructed from the dynamic footprints with $w_0 = 80$. The blue plot is the tapered correlation strength for the footprint of the pDC at time t and it is different from 0 only when the pDC is active. The green plot corresponds to the correlation coefficient of the last alive footprint before a death of the pDC, while the magenta plot is the correlation coefficient for the upcoming being born DC footprint after a death. These two plots are different from 0 when the pDC is not active. (b) Correlation strength reconstructed from the constant averaged footprint with $w_0 = 80$. The blue plot represents this correlation strength. (c) Dynamic correlation strengths reconstructed from the dynamic footprints with $w_0 = 50$. (d) Absolute differences between the reconstructed correlation strengths of dynamic footprints at $w_0 = 80$ and $w'_0 = 50$.

107

criteria of interest is plotted. The number of mutual topics corresponds to the number of vertices that are present in the same communities during two consecutive time frames. It is obvious that $k = 0.6$ provides a higher number of mutual topics than the other values of k, except for $k = 0.4$. As the stability is a requirement of our representation, k=0.6 is favored. The number of DCs logically increases with k but has its minimum at k=0.5 and it is similar for $k = 0.4$ and $k = 0.6$. Concerning the number of events, it decreases when k increases because the number of birth increases. Indeed, a community produces a *split* or a *merge* if it shares at least a proportion k of its parcels with several communities. When k increases, the requirement are higher and instead of *split* and *merge*, it is birth of new DCs that happens.
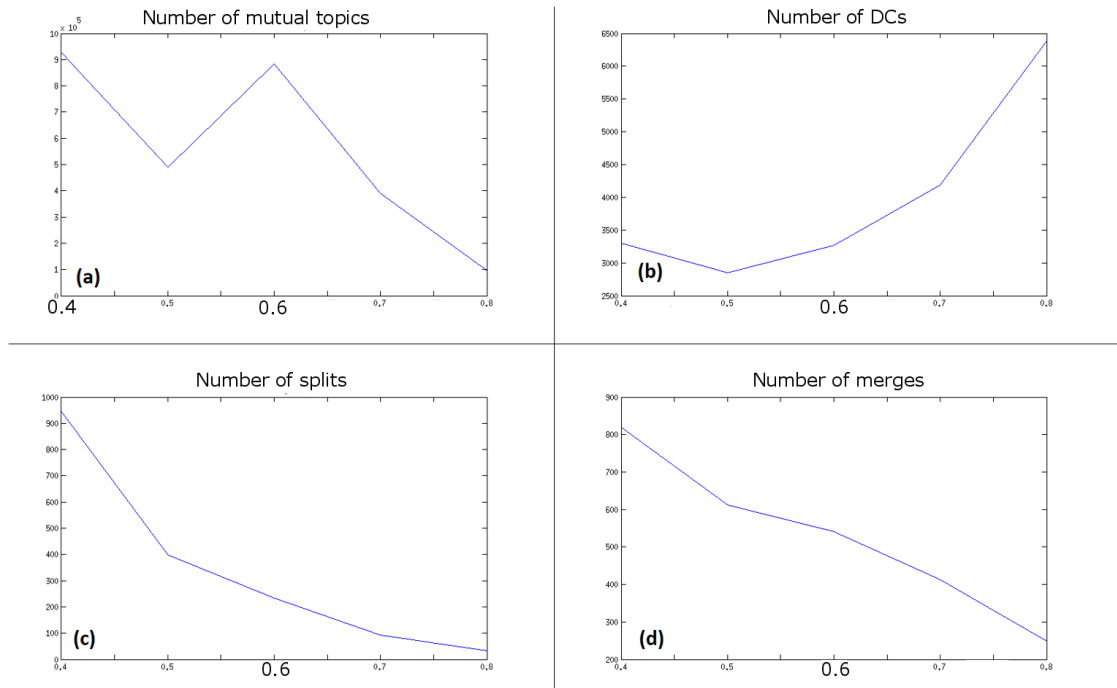

Figure 7.13: Illustration of the effect of the similarity threshold k. The community tracking is applied on 10 subjects with $k = \{0.4, 0.5, 0.6, 0.7, 0.8\}$. (a) Number of mutual topics in 10 subjects. (b) Number of DCs. (c) Number of splits. (d) Number of merges.

To conclude, from this experiment on 10 subjects, the most appropriated similarity threshold k for the community tracking is equal to 0.6. This value enables a high stability of the resulting sDCs and also the detection of events in a reliable way.

7.4.3 Determination of the number of Subject-level Dynamic Communities (sDCs)

The choice of a clustering algorithm for the consolidation requires the choice of a number of clusters that corresponds here to the number of sDCs detected within a subject.

From the literature [YKS$^+$11], there are several networks partitions but the dominant trend is to split the brain from 7 ICNs up to 17 ICNs where most of the networks are actually sub-networks of the 7 ICNs ([ADP$^+$14], [TF15], [LSGVDV14], [YAMC15]). As mentioned in Section 5.5.2, different approaches are implemented for the determination of the number of DCs.

Selecting an order of magnitude for the number of sDCs

As a first try, the number of cluster is supposed to be closer from 17 because a higher number of clusters permits to obtain a higher quantity of information concerning the interactions. However, the redundancy and the inaccuracy of some resulting sDCs (see Figure 7.14) urges us to decrease the range of interest. Results obtained with a fixed number of 8 clusters (see Figure 7.15) are compared with results for 18 clusters (see Figure 7.14). These values, 8 and 18, were inspired from calculation of Silhouette coefficient on a small sample of subjects for different number of clusters. It finally appears that the sDCs are less redundant for a lower number of clusters and therefore for a harder consolidation. Increasing the number of clusters increases only the redundancy and not the number of effective sDCs. The footprints for $k = 8$ depicted in Figure 7.15 are relatively well defined and some of them correspond to well-known RSNs such as the default-mode networks or the motor network (see Section 2.2.1 for an overview of the static networks).
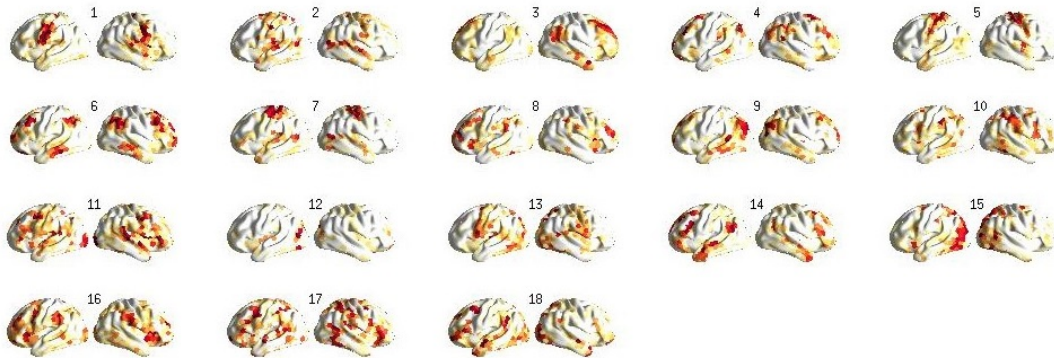

Figure 7.14: Averaged footprints of 18 consolidated sDCs for one subject. DC 5 and 7 are very similar. DCs 3, 10, 11, 12, 17 are for instance probably not real sDCs because they are either non-symmetric on both hemisphere (3) or very distributed (11, 17) or not enough stable (10, 12).

Consistency of the number of sDCs per subject

In order to check if the number of clusters can be the same for all subjects, the clustering algorithm is applied with numbers of clusters from 5 to 25 across 20 subjects and with 50 iterations. The mean cluster Silhouette coefficient is then calculated for each number
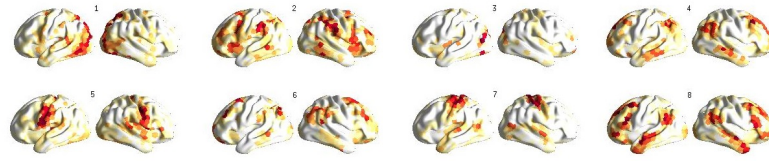
Figure 7.15: Averaged footprints of 8 consolidated sDCs for the same subject as in Figure 7.14. DC 1 is similar to the RSN of dorsal attention, DC 2 is similar to ventral attention, DC 5 is part of the motor network, DC 7 is also part of motor network, DC 8 is similar to default-mode network. The others do not look like known RSNs [YKS+11].

of clusters (see Section 3.2 for explanation on Silhouette coefficient). The aim of this experiment is to determine if a local maximum of the subject Silhouette coefficient exists or not. The plot of Figure 7.16 demonstrates that no local maximum is within this range and the smallest number of clusters, 5, gives the highest Silhouette coefficient.
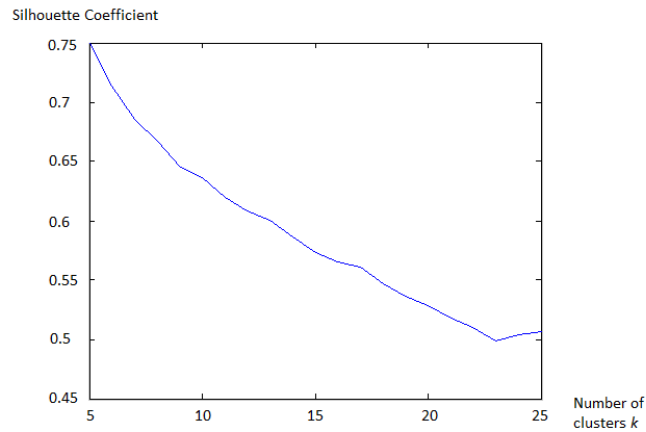

Figure 7.16: Averaged subject Silhouette coefficient for 50 iterations on 20 subjects as function of number of clusters $k \in \{5 \dots 25\}$

However, when individual subject Silhouette coefficients are observed, a high variability between the plots of the subject Silhouette coefficients is obvious and the lowest number of clusters corresponds not always to the largest Silhouette coefficient, as illustrated on Figure 7.17.

Therefore, 5 is not an optimal number of clusters for all subjects and restricting the number of sDCs to 5 also decreases the quantity of information about pDCs interactions then. Indeed, the number of clusters also influences the number of events of low and high interaction, because a lower number of DCs reduces the possible numbers of actors for these events. Although it increases the computational cost, it is thus necessary to perform an individual optimization step for each subject to determine its appropriate
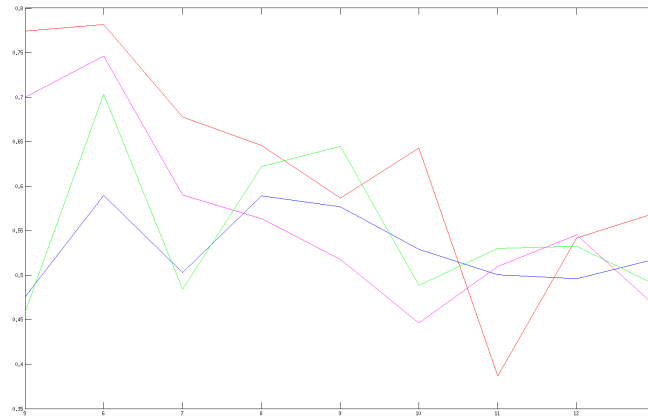
Figure 7.17: 4 different subject silhouette coefficients as function of number of clusters $k \in \{5 \ldots 13\}$

number of clusters for the consolidation. Several optimization procedures have been tested.

Individual optimization of the number of sDCs per subject

The first optimization method, that also permits to draw the previous conclusion, is the maximization of the Silhouette coefficient in a specific range of number of clusters $k \in \{7 \ldots 13\}$. For each subject, the spectral clustering is applied 50 times per value of k and the k corresponding to the maximum of the silhouette coefficient is selected as the optimized number of clusters k. The final pDCs obtained with this optimization procedure performed on 100 subjects is presented on Figure 7.19.

The second optimization algorithm uses Eigengap heuristic (see Section 5.5.2) to estimate the number of connected components in the graph built for the consolidation and the range for k is also $\{7 \ldots 13\}$. Figure 7.20 shows the resulting pDCs footprints with exactly the same parameters as Figure 7.19.

Finally, we replace clustering by community detection based on stability, a method already used for the community tracking. It does not require any fixed range and determines thus the number of communities in the pruned sDCs at the first hierarchical level (only 1 Markov time is used). It confirms the use of a range closer to 7 than to 17 for the optimization of the number of clusters since for 100 subjects, the number of clusters varies between 3 and 8. Figure 7.21 corresponds to the final pDCs, always with the same parameters.

Figure 7.18 permits to compare the three optimization processes tested by displaying the distribution of the difference of cluster numbers for each subject according to the method selected. It is obvious that the silhouette coefficient and the Eigengap heuristic methods provide higher number of DCs than the community detection. Indeed, the use

of community detection permits to consolidate the sDCs to form very stable DCs but they are at the end too stable and even close from being static when the time lines are considered. Moreover, reducing the number of sDCs to very low values such as 3 implies that a lot of information is lost since the temporal overlap of the consolidated DCs is obligatory high. Therefore, the community detection is not appropriated for the consolidation because it is a too hard process. Figure 7.21 supports this choice because with equivalent parameters, only three pDCs are detected and one of them is not meaningful, while consolidation using clustering techniques detects at least 6 pDCs.
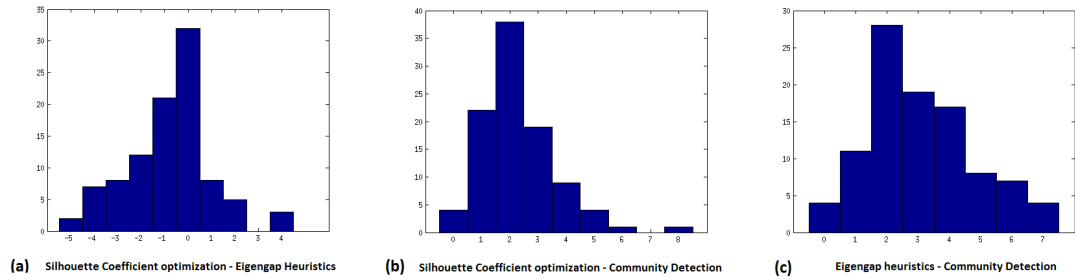

Figure 7.18: Histograms representing the distribution of the differences between the number of sDCs obtained with 2 different techniques. (a) Number of sDCs obtained via optimization of the silhouette coefficient minus number of sDCs obtained via Eigengap heuristics.(b) Number of sDCs obtained via Silhouette coefficient optimization minus number of sDCs obtained via Community Detection.(c) number of clusters obtained via Eigengap Heuristic minus number of sDCs obtained via Community Detection.

The Silhouette coefficient and the Eigengap heuristic provide more similar results since the distribution of differences of their number of sDCs on Figure 7.18(a) is centered on 0. Eigengap heuristic tends to detect globally more sDCs than Silhouette coefficient. The final number of pDCs is finally higher with Eigengap heuristic optimization, as illustrated in Figure 7.20, although three of its pDCs are present in less than 10 subjects, while only one is in less than 10 subjects for Silhouette coefficient optimization. Silhouette coefficient and Eigengap heuristic optimizations provide thus similar results. After several tests with 200 subjects, the Silhouette coefficient is finally preferred to determine the number of sDCs after the consolidation because the number of subjects where the pDCs are present is globally higher for procedures using Silhouette coefficient instead of Eigengap heuristic. Moreover, it is a compromise between community detection and Eigengap heuristics.

To conclude, the number of sDCs k is optimized between 7 and 13 owing to the mean Silhouette coefficient estimated via 50 iterations per value of k of the spectral clustering. It provides then approximately 6 pDCs according to the selected parameters for the population-level DC tracking. To reduce the redundancy of the detected networks and obtain stable pDCs, a maximum number of clusters equal to 13 is selected. However, because a too hard consolidation leads to loss of dynamic information and abusive merge of DCs, a minimum cluster number of 7 is also set. Moreover, each subject can have a

different number of clusters and it is neurophysiologically relevant since during the fMRI experiment, all the DCs may be not active. The minimum is yet 7, the number of static ICNs, because sub-networks can also be identified and some sDCs might correspond to noise which can be thus isolated. Furthermore, this range has been confirmed by several tests with other ranges and the observation of the resulting footprints.
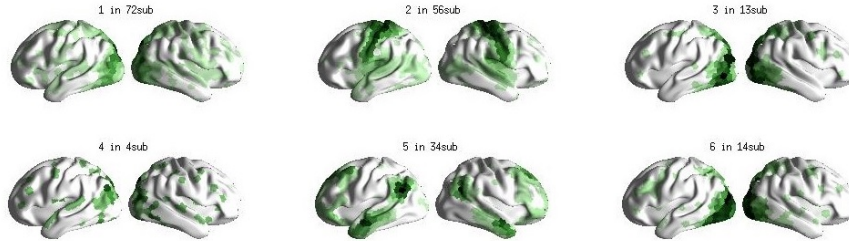

Figure 7.19: pDCs obtained after consolidation with Silhouette Coefficient optimization for 100 subjects. The similarity threshold between the sDCs is 30% and the lowest degree for center is equal to 30 (see explanations of this parameters in 5.6). pDCs 1, 3 and 4 are similar to dorsal attention and visual networks, pDC 2 is similar to motor network, pDC 5 is similar to default-mode network. pDC 4 is only present in 4 subjects and does not lot looks like any known ICN [YKS$^+$11]

Figure 7.20: pDCs obtained after consolidation with Eigengap heuristics for 100 subjects. The similarity threshold between the sDCs is 30% and the lowest degree for center is equal to 30. pDCs 1 and 2 are similar to dorsal attention and visual networks, pDCs 3 and 7 are similar to motor networks, pDC 5 is similar to default-mode network, pDC 6 is similar to a part of the motor network. The pDC 4 is present in only 3 and cannot be identified to one of the 7 ICNs described in [YKS$^+$11].

7.5   summary

The key parameters of the three main steps of our approach, the community detection, the community tracking and the consolidation at subject level, are consequently validated by the experiments. Added to the individual validation of the different results, it permits to conclude that our approach is reliable.
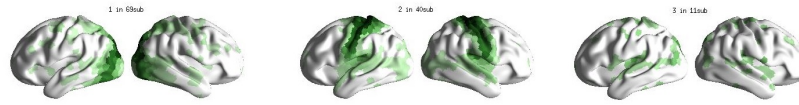
Figure 7.21: pDCs obtained after consolidation with Community detection based on stability for 100 subjects. The similarity threshold between the sDCs is 30% and the lowest degree for center is equal to 30. pDC 1 is similar to dorsal attention network, pDC 2 is similar to motor network and DC 3 does not look like any known ICN [YKS$^+$11].

The three main elements highlighted in this chapter are the pDCs that correspond to the components of the dynamic FC, their co-occurrence and their interactions. For each of them, its repeatability across two subsets of the population is tested. A particularly interesting point concerning the pDCs is their similarity with specific static RSNs. Moreover, it has been demonstrated that our approach enables to detect significant co-occurrence of couple of pDCs and events for the analysis of interactions and it is sufficient to state that interactions exist between pDCs.

CHAPTER 8

Discussion

In this chapter, the results are first discussed with regards to the aims of the thesis and the related limitations are highlighted. The main contribution of the thesis is a reliable and robust representation of the dynamic RSNs via spatial and temporal components respecting specific requirements. In addition, we define methods for the analysis of their interactions.

Two key choices made for the methodology are also explained in this Chapter before the discussion of the limitations of the overall methodology examined step by step.

8.1   Results

The results are validated through different experiments described in the Chapter 7. The dynamic functional connectivity is represented by 6 different pDCs inside a population of 200 subjects (see Figure 7.1). 4 of them are highly similar to static ICNs (see Figure 7.2) and they are all reproducible (see Figure 7.3). Moreover, two couples of pDCs have a particularly significant co-occurrence that increases the significance of their events (see Section 7.2.3). The significance of the events is also evaluated independently from the occurrence and some of them happen more often or less often than random (see Section 7.3). The pDCs act thus as components of the dynamic brain functional connectivity, being sometimes simultaneously active and exchanging nodes at different time steps. They present a spatial variation as well as activations and de-activations that permits to study their interactions. The initial requirements defined in Chapter 1 are thus respected, even if the interactions are only identified at the temporal level (see Section 7.3.1).

8.1.1   Loss of information

The functional dynamic connectivity cannot be analyzed in all the subjects since one subject presents a reduced number of pDCs and sometimes none. The histogram depicted

by Figure 8.1 illustrates that approximately 30% of the population has one or none of its sDC related to a pDC. It is normal that all the pDCs are not represented into one subject since they may be not all active during the time of the experiment. However, it is not possible to have none of the functional network active within a subject and it is necessary to accept that the pDCs detected with our approach do not characterize the complete functional connectivity. It is the cost to obtain robust pDC. A part of the information is sacrificed to ensure the reliability of the remaining part but it is still accessible and it is therefore possible to extend the pDCs to more numerous subjects if less defined footprints are acceptable. Working at the population level is necessary at this preliminary stage of the representation development because it needs to be validated, as well as the tools it provides, to be then applied at the subject-level.
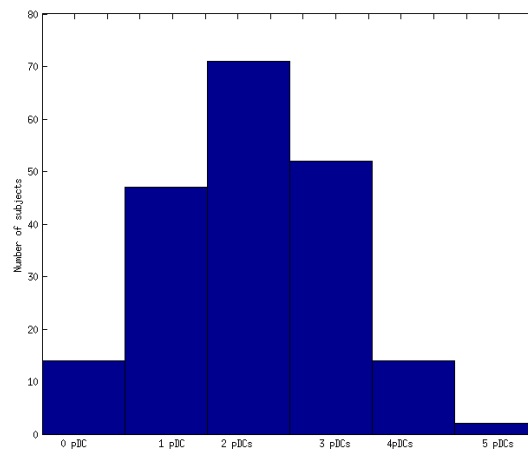

Figure 8.1: Distribution of the number of pDCs per subject. The maximum number of pDCs detected in 1 subject is 5 while some subjects present no pDC at all. This distribution is made of 200 subjects.

Furthermore, in order to obtain consistent pDCs, a consolidation step is performed and also creates a loss of information. Its parameters have been designed in order to decrease the redundancy between the sDCs (see Section 7.4.3) but it also results in a compression of the sDC. With Figure 8.2, it is possible to remark that it indeed reduces efficiently the redundancy: when the footprints of the pDCs, constructed with parcels which are in at least 5 constitutive sDCs, are compared between each other via their Dice coefficients, it is visible that the redundancy is globally eliminated within subjects. Indeed, the Dice coefficients are low, except between the pairs pDC1-pDC5 and pDC1-pDC6. However, it has already been pointed out that pDC5 and pDC6 are less reliable than the other pDCs and the properties of these two DCs must be considered with care (see Section 7.1). The low values of Dice coefficients also confirm that the probability of *splits* and *merges* between these pDCs is very low since it requires sharing of nodes at different time points. Most of the *splits* of *merges* are suppressed during the pruning on lifetime. It could either mean that they happens at very short time scale (less than ten time frames, threshold

116

used during the pruning) or that they do not happen in our representation despite of the optimization of the parameters along the steps of construction of pDCs, or also that the pruning is too hard and it is an additional source for the loss of information. It can be remarked that loss of information happens most of the time during processes that aim to eliminate noise. Denoising is indeed one of the biggest challenge when fMRI data are used (See Section 2.5), and if the efficiency of the pre-processing steps is increased, a higher quantity of significant information could be used.
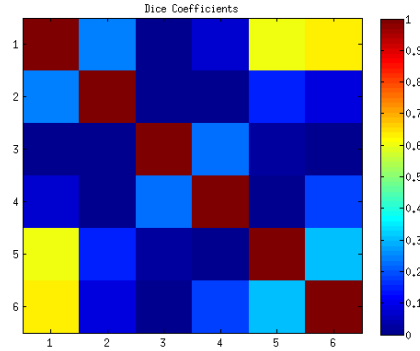

Figure 8.2: Dice coefficients between the pDCs.

8.1.2 Interpretation of the detected interaction

The events are defined according to death and birth of a DC and therefore according to a notion of "activation" which is used several time here. In our representation, a network is "active" at time t if it is formed at time t, that is to say if its parcels have correlated signals at time t. It is therefore assumed that when the brain needs one of its function, regions constituting the concerned networks become linked to form this network. However, correlation of fMRI signals does not mean activation [VdHHP10]. The referred concept of activity here is thus the activity of the network. The translation to the activity of a function must be handled with care for the interpretation of the dynamic characteristics of the networks.

Several experiments are performed to determine to what extent the interaction identified by this representation can be trusted as true interaction and not as a random combination of *death*, *birth* and *survival* (see Section 7.2.3 and 7.3.4). Their repeatability is first established between two subsets that ensure similar footprints for the pDCs (see Figure 7.3 and Section 7.3.3). The other experiments uses a fixed occurrence by generating synthetic subjects (see Section 7.3.4). All the possible combinations of pDCs are not generated and it is a problem for the evaluation of the event significance if the number of synthetic populations is lower than 200 according to experiments, since some of the events detected as significant at one iteration are no more significant at the next iteration. A set of events is however stable across iterations for a number of 1 000 populations and it is sufficient to validate our approach (see Figure 7.11).

The co-occurrence is not directly defined in this thesis as an interaction. It has already been explained that the events are a sort of characteristic behavior of the co-occurrence (see Section 7.2). The co-occurrence is influenced by the occurrence but for two couples of pDCs it happens more or less often than random in a significant way (see Figure 7.7). On the other hand, a set of events are not influenced by the occurrence and most of them do not happen between the two couples of pDCs mentioned previously (see Figure 7.11). Our representation thus provides different ways to study the interaction of the dynamic functional networks it constructs.

8.2   Definitions of Similarity

Two different similarity functions are utilized in this thesis to compare two communities. Equations 8.1&8.2 remind below their respective definition with simple notations. A and B are two groups of points which can for instance be communities as in community tracking or footprints on the brain surface as in consolidation or pDCs detection.

$$sim(A, B) = \frac{|A \cap B|}{\max(|A|, |B|)} \tag{8.1}$$

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{8.2}$$

The first definition of similarity is used for the community tracking while the second one, the Jaccard, is utilized for the consolidation and the pDC detection.

If two communities A and B share a very low number of parcels, the similarity is close to zero because $|A \cap B| \to 0$. The Jaccard $J(A, B)$ is closer to zero than the similarity function $sim(A, B)$ because $|A \cup B| \geq \max(|A|, |B|)$. On the other hand, if A and B share almost all their parcels, the similarity is close to one. In that case, $\max(|A|, |B|)$ is still lower than $|A \cup B|$ and then the similarity function sim is closer to one than the Jaccard. Therefore, the Jaccard is a harder similarity function than sim. The difference between these two similarity functions decreases when the difference between the size of A and the size of B increases because then $|A \cup B| \xrightarrow[|A| \to \infty, |B| \to 0]{} \max(|A|, |B|)$. Moreover, function sim is globally higher than Jaccard because $|A \cup B| \geq \max(|A|, |B|)$. It is illustrated in Figure 8.3 which represents the two similarity functions evolution for a toy example. Two communities represented by two vectors of size 50 filled with 0 and 1 shares the number of vertices indicated on the x-axis and always contain the same number of vertices.

As the similarity function sim is generally higher than the Jaccard, its tolerance to noise in the footprints would be higher too and less spurious DCs are created in the community detection. Moreover, to increase the stability, the communities associated to the same sDC must have approximately the same size and it is therefore particularly relevant to
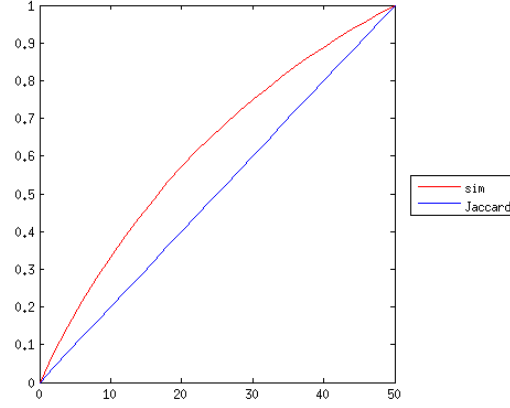
Figure 8.3: Illustration of the evolution of Jaccard and *sim* function when two groups of points A and B increase the number of points they share x represented on the x-axis. $A \cup B = 50$ is constant, $A \cap B = x$ and $|A| = |B| = 25 + x$

use *sim*. However, for the consolidation, the tolerance to noise has already been allowed in the community tracking and the loss of information must be restricted by a hard enough similarity function. It permits a compromise between a too permissive and a too hard consolidation. For the population-level DC detection, the similarity threshold is already low and this step is delicate because it is the one that determines which events remain at the population level. They need to be as accurate as possible and a harder metric like the Jaccard is more appropriate.

Moreover, an interpretation of their difference is that *sim* takes one of the community as reference while Jaccard used both communities. It is relevant for tracking to compare one community to another because this process aims to link one specific community at time t to another at time $t + 1$. On the other hand, consolidation and pDCs detection look for ensemble of communities.

8.3   Choices for the Consolidation of sDCs: Spectral Clustering

The consolidation is performed owing to a spectral clustering, as explained in Section 5.5.2. It is a sensitive step designed to make the sDCs workable for their translation at the population level and it could be realized in several different ways.

8.3.1   Choice of the method

For the consolidation step, we could have used a different way to link the fragments of sDCs that are separated by the community tracking. The first try was to perform a kind of "manual" consolidation also based on the similarity of the averaged footprint of the sDC, in a similar way to pDCs tracking. The advantage of this method is its

119

ability to introduce the constraint of temporal overlap: two fragments cannot be linked if they temporally overlap. A tolerance of five time frames is accepted, in analogy with the tolerance for the definition of events. However, a similarity threshold must be fixed and if it is too high, the redundancy is still high and the sDCs are not consolidated. On the other hand, if this similarity coefficient is too low, all non-overlapping sDCs are linked in a non-accurate way. Methods to find an optimal similarity threshold have been developed but none of them provides an appropriate result. These methods aim to decrease the similarity between the consolidated sDCs and to increase their respective stability (see definition in Section 5.5.1). Obviously, increasing the similarity threshold provides a softer consolidation with more sDCs and therefore higher stability within sDCs and also higher similarity between sDCs. As decreasing the similarity threshold would have the contrary effect, an optimal similarity should exist and be detected by the joint analysis of stability and similarity, but none has been found. Because the similarity and stability metrics are very similar since the stability is actually the averaged similarity of the constitutive footprints of the sDCs, with a slightly different definition of similarity (see Section 5.4), the problem described above is similar to a partition problem which can be solved with a clustering algorithm. Using clustering induces to delete the constraint on the temporal overlap but it is finally not so important because this constraint cannot be satisfied without spatial redundancy in the sDCs as described above. This redundancy is not acceptable because two sDCs with the same footprints represent the same dynamic RSN, and yet the events that permit to study the interactions between RSNs would be attributed to two different sDCs. It is not accurate and it blurs the data.

Therefore a spectral clustering is selected to perform the consolidation. The spectral clustering reduces the dimensionality of the data and is particularly adapted here since we already have a similarity metric. It is therefore less computationally expensive. However, it still raises the issue of parameters and metrics choices, and more particularly of the number of clusters (see Section 7.4.3).

8.3.2   Choice of spectral clustering parameters

A spectral clustering is based on the spectral information to perform a partition of the considered graph in a more appropriated space. There are different steps to access this information and at each step, the choice of a method and the associated parameters must be made. Section 7.4.3 already illustrates the process to define the number of final clusters and it has required several experimental tests to fulfill the requirements of the functional specification of the representation.

Similarity Graph

The first step of the spectral clustering is to build the similarity graph from the similarity matrix. The most common similarity graph are ϵ-neighborhood, k-Nearest Neighbors (kNN), mutual kNN and fully connected graph [VL07].

The ϵ-neighborhood graph consists in linking nodes whose distance is below ϵ, and it is usually an unweighted graph since all edges have more or less the same scale. The difficulty with this kind of graph is the choice of ϵ and the optimal epsilon depends on the distance scale inside the clusters. Therefore, ϵ-neighborhood graph are not adapted to detect clusters with different distance scales [VL07]. The kNN graph described in Section 5.5.2 does not have the problem of ϵ-neighborhood graph with distance scale since it is not based on a distance threshold but on the number of neighbors. A node will be linked to its k closest neighbors even if they are relatively far. However, it implies that it requires a relatively high density within clusters to be identified and a large inter-clusters distances to produce an accurate parcellation [VL07]. The mutual kNN graph is similar to the kNN graph but instead of linking nodes i and j if at least one of this two nodes is one kNN of the other, both nodes must be one of the kNN of the other. The weight of the edges is inversely proportional to the similarity between the two nodes it links. The mutual kNN graph is a compromise between kNN graph and ϵ-neighborhood graph and is appropriate for the detection of clusters with different densities [VL07]. The fully connected graph simply links two nodes if their similarity is positive and therefore, the returned adjacency matrix is often non-sparse [VL07].

Luxburg et al. [VL07] finally recommend the kNN graph as first choice and it is the choice made in this thesis since the consolidation step is performed on different subjects which can present very different similarity matrices and an adaptable graph is required.

Choice of k for the k-Nearest Neighbors graph

The choice of k for the kNN graph influences the number of connected components in the graph. If k increases, the number of connected components decreases. This number of connected components must be lower than the number of clusters to be detected, otherwise the clustering algorithm trivially returns the connected components as clusters [VL07]. In that case, it is then important to be sure that the connected components are really the clusters of interest but it is preferable to have a connected graph to take advantage of spectral clustering [VL07]. For large graph, when number of nodes $n \to +\infty$, it has for instance been demonstrated that k should be on the order of $log(n)$ [BCQY97]. However, n is here only on the order of 10^2. We test different values of k for different subjects with an optimized number of clusters for the spectral clustering (description of the optimization procedure in Section 7.4.3) and it appears that this asymptotic result $log(n)$ provides rather appropriate footprints. Moreover, a test is performed in our algorithm with a Matlab function using the Tarjan's algorithm to check if the graph that is clustered is connected or not.

Choice of graph Laplacian

The graph Laplacian used for the consolidation step in this thesis is a normalized graph Laplacian with a symmetric matrix. Alternatives are unnormalized graph and non-symmetric normalized graph closer from a random-walk. The choice of graph Laplacian is not decisive if the degrees of the similarity graph are not highly distributed

[VL07]. However in our graph, the degrees are roughly included between 5 and 20 for approximately 100 nodes.

An unnormalized graph Laplacian does not directly enable to maximize the similarity within the clusters, while normalized graph Laplacians enable to fulfill both objectives of clustering algorithm: a low similarity between clusters (high separation) and a high similarity within clusters (high cohesion). Therefore the unnormalized graph Laplacian must be avoided [VL07]. The difference between a symmetric and non-symmetric normalized graph Laplacian is that the non-symmetric normalized graph Laplacian does not normalize the rows of the matrix containing the Eigenvectors of the normalized Laplacian graph [VL07] (see definitions of both graphs in Section 3.2). It is a small difference but the symmetric normalized graph is more appropriate to accurately cluster graphs when the cohesion within clusters varies significantly across clusters [NJW+02]. It is our case here and the symmetric normalized graph is therefore the selected graph Laplacian for the consolidation of sDCs.

Choice of clustering algorithm

The final clustering step is made very simple by the use of a graph Laplacian and it is indeed the aim of the spectral clustering [NJW+02]&[VL07]. Therefore, the choice of the clustering algorithm is not determinant for the quality of the consolidation and a simple clustering algorithm is appropriated to extract the final clusters that form the sets of sDCs for each subjects. A k-mean algorithm is thus used in this thesis.

8.4 Limitations

At each step of the construction, approximations are made.

During the preprocessing, a band-pass filtering is applied on the fMRI signals in order to reduce noise (see Section 4.2) and the frequency spectrum is thus reduced. Moreover, a parcellation of the brain surface is then realized to decrease the computational cost by averaging signals from several voxels into one parcel signal. This two operations are highly necessary because the noise is a real burden for fMRI (see Section 2.5) and the SWA needs already approximately 45 minutes per subject to run with a reduced number of nodes. Spatial and temporal resolution are consequently decreased during preprocessing.

Then, the SWA reduces the temporal resolution again due to noise. Even if almost every time point of the fMRI scan gets a set of static communities, the latter are established by considering 80 time points. It does not mean that the temporal resolution of 0,7 seconds turns into $0, 7 * 80 = 56$ seconds owing to the temporal overlap of the window, but it can no longer be considered as the TR of the fMRI scan. The size of the window cannot be reduced because several time points are necessary to increase the Signal-to-Noise Ratio of signals and it has been demonstrated in Section 7.4.1 that a reduction of the window length until its minimum does not significantly influence the community detection then.

In addition to the window length, a possible limitation of our approach is the use of correlation as selected metric to measure the functional connectivity. For instance, mutual information or coherence are based on higher order statistics [ADP+14] and could be an interesting alternative to correlation. More work is necessary to quantify the influence of the metrics used to establish connectivity.

During the community detection based on the information extracted by the SWA, the community has to contain at least 40 parcels. It increases the consistency of the future sDCs but small networks are neglected and considered as noise. An arbitrary minimum size of networks is thus defined and corresponds to one twentieth of the nodes, permitting still a large variation in the number of nodes in each community. On the other hand, the hierarchical properties of the community detection based on stability are not fully exploited since only one Markov time is implemented. The spatial scale used here is therefore the finest available one, but exploring larger scales could give additional information on the structure of the dynamic connectivity.

The community tracking generates additional sDCs due to the compulsory death of DC involved in events, but the consolidation step is implemented to reduce this effect. Community tracking also limits the temporal variability of the spatial footprints with the similarity threshold. This variability needs to be limited, otherwise every DC would almost always find a community to survive and a relative stability is one of our prerequisite. The similarity threshold is thus optimized to maximize the stability and it does not necessarily corresponds to the highest similarity threshold since an high similarity threshold generate more sDCs and less events. A similarity threshold closer from 0.4 (see Figure 7.13) might be a solution to access the higher level of interaction because the number of *splits* and *merges* would be higher but to the detriment of the robustness of the sDCs. At the end of the day, the experiment performed in Section 7.4.1 permits to validate this community tracking since at the population level, the dynamic communities it identified are still robust.

The pruning on lifetime and stability requires the assumption that sDCs with short lifetime and low stability corresponds to noise (see Section 5.5.1). Pruning all sDCs with short lifetime reduces the temporal resolution of the methods, while pruning sDCs with low stability is an additional restriction of the spatial variability of the footprints. The thresholds for lifetime and stability cannot be realistically decreased or it would be equivalent to the suppression of this pruning step, that is yet necessary because the community tracking return too many sDCs with high redundancy (more than 100).

The consolidation is maybe the most sensitive part of the workflow because it sets the final number of sDCs and compresses them while the output of the community tracking should not need this operation. It agglomerates sDCs in generalizing the information of the cluster centroid to all the sDCs of the cluster when conflict of interest appears due to temporal overlap. A high quantity of information is therefore lost during this steps and parameters of the spectral clustering must be carefully chosen (see Section 7.4.3). The current number of sDCs per subject is included between 7 and 13 to be closer from a 7-networks parcellation as presented in Section 2.2.1, but extracting smaller

network may provide a more exact insight in the brain organization. However, with the presented architecture of the representation, it is not possible since increasing the number of clusters just increases the redundancy. In addition to the loss of information this step requires, the consolidation is also based on the averaged footprints of the pDCs. The results of Section 7.4.1 concerning the averaged footprints demonstrate that it must be handled with care since it does not take into account the dynamic properties of the networks they represent. Indeed, the consolidation is only based on static information. The Section 7.4.1 also demonstrates that the dynamic footprints of the sDCs have high internal correlation strengths, in spite of this consolidation step. Moreover, the stability of the final pDCs is also high (see Section 7.1) and it demonstrates that the consolidation is not too hard.

Concerning the detection of events characterizing the temporal interactions, their number is also partly determined by arbitrary parameters. As a reminder, the death of two communities can be declared as spontaneous even if in our representation it happens at 5 time frames of difference. It can thus creates additional events that should not be taken into account but it is necessary due to the incertitude of the time resolution. On the other hand, a death is considered as a death only if the sDC is inactive during at least 20 time frames. Therefore, death and birth can be missed and additional events are not detected whereas they should have been. By setting existence conditions on the correlation of the dynamic footprint around death and birth, the number of events is still decreased in order to conserve only the significant events. It is finally a hard selection of the events but it is justified since only few of them can be considered as significant, as demonstrated into Section 7.3.

Finally, the population sDCs tracking also provokes a non-negligible loss of information because the pDCs it detects are not present in all subjects (see Figure 8.1). The interactions have consequently less opportunities to be detected and if the subject-level permits to build the population-level, the contrary is not true. A decision must be taken concerning the sDCs that are not attributed to a pDC during this operation. Here, they are not taken into account in the final representation because they are questioning the robustness of the representation but less restrictive parameters can be used to include them in the analysis of the functional connectivity. However, if the artifacts are more efficiently suppressed, this issue is suppressed as well.

8.5 Summary

The first aim of this thesis is the construction of a robust representation of the dynamic functional connectivity as networks. The choices of parameters and methods generate approximations and loss of information. They mainly aim to remove noise from the data in order to obtain only reliable pDCs, stable enough, and related to the RSNs of the literature. The constraints thus imposed by the low Signal-to-Noise Ratio may require to be softened and independently examined to allow more flexibility in the representation. Indeed, even if the Signal-to-Noise Ratio happens to be improved in fMRI signals, some

factors such as the between-subject variability irremediably complicates the population-wise analysis of the functional connectivity. However, the dynamic networks provided by our approach already open new possibilities that can be investigated (see Section 9.2).

<div align="right">

CHAPTER 9

</div>

Conclusion and Future Work

9.1 Major Findings of this thesis

In this thesis, a new approach is developed to detect and analyze the functional connectivity in neuroscience. The aim is to provide a richer representation than states or component of connectivity. It uses the complete time course of experiments as well as the complete surface of the brain. It has been applied to real subjects to demonstrate its robustness and its significance. Moreover, its characteristics enables to develop methods to establish the possible interactions between functionally separated regions of the brain.

9.1.1 A new and robust representation of the dynamic Functional Connectivity (FC)

The dynamic FC is represented for each subject by a set of sDCs. They are dynamic because they have periods of activity, during which they are formed, and their spatial footprint changes over time in conserving a minimum stability in order to conserve a relative consistency across time frames. Moreover, periods of activity from different sDCs can overlap and they are therefore spatial and temporal components of the dynamic FC.

It is necessary to extend these sDCs at the population level in order to test their repeatability and also to take into account the subject variability. Thus, for 200 subjects, 6 different pDCs are identified and validated. Over the whole experiment and the whole population, they spatially overlap but at the subject level and during individual time frames, they are spatially separated. The main concern during their construction is to suppress noise and it is therefore a limitation. A second concern is to access interaction information.

9.1.2 An insight in dynamic interactions between components of dynamic Functional Connectivity (FC)

The interactions defined in this thesis can be classified in two different levels of interaction. The higher level of interaction is based on the dynamic sharing of nodes between the DCs and is therefore related to both spatial and temporal characteristics of the DCs. The lower level of interaction only rely on temporal characteristics but it is only the most accessible level. The interactions are characterized by events and their are analyzed at the population level in order to increase the reliability of the analysis.

The lower level of interaction is related to the temporal overlap of the DCs, the co-occurrence, which is particularly significant for specific couples of DCs. The significance at this level of interaction is established with both the repeatability across subsets of the population and the independence regarding the occurrence. A set of events is as well detected as significant and their recurrence can also be estimated owing to the methods developed here.

9.2 Future work

A revision of the priorities for the construction of the sDCs and pDCs may provide interesting results concerning the effect of noise. The pruning step can for instance be suppressed to access the higher level of interaction and estimate its significance. Moreover, during the detection of the pDCs in the set of sDCs, it is possible to urge each subject to participate to each pDCs. It is probable that the resulting spatial identity of the pDCs represented by their footprint on the brain surface would not be sharp but the significance of the interactions can also be affected.

A relevant test for the detected pDCs would be to observe their behavior in subjects performing tasks during task-fMRI. For instance, the HCP provides task-fMRI scans that can be used since the preprocessing is the same. During a motor task, it would be expected that the pDCs associated to the motor RSN has higher occurrence. The interactions may be as well related to the execution of a task. Different motor tasks can also enable to detect components of the motor dynamic RSN.

Finally, our approach can be utilized to compare the pDCs of patients suffering from neuropathologies and the pDCs of controls. Interaction events may be clear biomarkers permitting to realize a classification, as well as dynamic graph metrics available with our representation.

Acronyms

BOLD Blood Oxygen Level Dependent. 10–14, 34, 36, 37, 51, 123

CAP Co-Activation Pattern. 37, 38

DC Dynamic Community. xi, 4, 5, 41, 42, 44, 46, 48, 55, 59, 61–76, 79–81, 84, 86, 89, 90, 94, 97, 98, 104, 106–110, 112, 114–116, 121, 125, 126, 128, 129, 131

EEG ElectroEncephaloGraphy. 1, 14, 16, 25, 52

FC Functional Connectivity. 14–16, 20, 25, 26

fMRI Functional Magnetic Resonance Imaging. 1, 2, 4, 5, 7, 10, 11, 13–16, 18, 20, 23–26, 30–38, 50–52, 58, 85, 97, 101, 110, 115, 120, 122, 124, 125

GLM General Linear Model. 25

HCP Human Connectome Project. 5, 49

ICA Independent Component Analysis. 20, 27–30, 32, 33, 35–38, 40, 124

ICN Intrinsic Connectivity Network. 1, 2, 15, 16, 25, 29, 32, 36, 40, 106, 110–113, 129

IVA Independent Vector Analysis. 36, 37, 125

kNN k-Nearest Neighbors. 44, 46, 75, 118, 119

MEG MagnetoEncephaloGraphy. 1, 14, 50

MRI Magnetic Resonance Imaging. 7, 49, 50

PCA Principal Component Analysis. 29

pDC Population-level Dynamic Community. xii, 56, 76–85, 89–105, 108–117, 121, 122, 125–130

PET Positron Emission Tomography. 10, 13, 14

rest-fMRI rest-fMRI. 1, 4, 5, 17, 49, 75, 123

RF RadioFrequency. 8–10, 13, 123

ROI Region Of Interest. 2, 15, 20, 26, 28, 32, 39, 40

RSN Resting-State functional Network. 2–5, 7, 14, 16, 25, 28, 29, 40, 47, 48, 55, 58, 63, 75, 82, 89–92, 107, 108, 118, 122, 127, 129

SC Structural Connectivity. 15

sDC Subject-level Dynamic Community. xii, 56, 61, 69, 71, 72, 74–79, 82–87, 89–92, 97, 100, 103, 104, 106–114, 116–118, 120–122, 125–127, 129

SWA Sliding Window Analysis. 30, 32–34, 36, 37, 57, 58, 79, 104, 120, 121, 124, 125

TE Echo Time. 10

TR Repetition Time. 10, 11, 13, 57, 120

WTC Wavelet Transform Coherence. 34, 35

Bibliography

[ADP+14]   Elena A. Allen, Eswar Damaraju, Sergey M. Plis, Erik B. Erhardt, Tom Eichele, and Vince D. Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

[AFPA14]   M. Anderson, G.-S. Fu, R. Phlypo, and T. Adali. Independent vector analysis: Identification conditions and performance bounds. *IEEE Transactions on Signal Processing,*, 62(17):4399–4410, 2014.

[AHSV+07]   J. R. Andrews-Hanna, A. Z. Snyder, J. L. Vincent, C. Lustig, D. Head, M. E. Raichle, and R. L. Buckner. Disruption of large-scale brain systems in advanced aging. *Neuron*, 56(5):924–935, 2007.

[AKPC13]   M. R. Arbabshirani, K. A. Kiehl, G. D. Pearlson, and V. D. Calhoun. Classification of schizophrenia patients based on resting-state functional network connectivity. *Front. Neurosci*, 7(133):10–3389, 2013.

[BA99]   A. L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[BAHS08]   R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter. The brain's default network. *Annals of the New York Academy of Sciences*, 1124(1):1–38, 2008.

[BCQY97]   M. R. Brito, E. L. Chavez, A. J. Quiroz, and J. E. Yukich. Connectivity of the mutual k-nearest-neighbor graph in clustering and outlier detection. *Statistics & Probability Letters*, 35(1):33–42, 1997.

[BDDS05]   C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1457):1001–1013, 2005.

[Bec12]   C. F. Beckmann. Modelling with independent components. *Neuroimage*, 62(2):891–901, 2012.

[BJG+13]   T. Blumensath, S. Jbabdi, M. F. Glasser, D. C. Van Essen, K. Ugurbil, T. E. J. Behrens, and S. M. Smith. Spatially constrained hierarchical parcellation of the brain with resting-state fmri. *Neuroimage*, 76:313–324, 2013.

[BKP+15]   J. Blautzik, D. Keeser, M. Paolini, V. Kirsch, A. Berman, U. Coates, M. Reiser, S. J. Teipel, and T. Meindl. Functional connectivity increase in the default-mode network of patients with alzheimer's disease after long-term treatment with galantamine. *European Neuropsychopharmacology*, 26(3):602–613, 2015.

[BL07]   E. B. Beall and M. J. Lowe. Isolating physiologic noise sources with independently determined spatial measures. *Neuroimage*, 37(4):1286–1300, 2007.

[BM10]   C. M. Bennett and M. B. Miller. How reliable are the results from functional magnetic resonance imaging? *Annals of the New York Academy of Sciences*, 1191(1):133–155, 2010.

[BMB08]   R. M. Birn, K. Murphy, and P. A. Bandettini. The effect of respiration variations on independent component analysis results of resting state functional connectivity. *Human brain mapping*, 29(7):740–750, 2008.

[BS09]   E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[BS12]   E. Bullmore and O. Sporns. The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5):336–349, 2012.

[BSR+15]   B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq, and F. Vidal. Spatial and temporal resolutions of eeg: Is it really black and white? a scalp current density view. *International Journal of Psychophysiology*, 97(3):210–220, 2015.

[BV07]   R. L. Buckner and J. L. Vincent. Unrest at rest: default activity and spontaneous network correlations. *Neuroimage*, 37(4):1091–1096, 2007.

[BWP+11]   D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton. Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108(18):7641–7646, 2011.

[BZYHH95]   B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

[CA12]   V. D. Calhoun and T. Adali. Multisubject independent component analysis of fmri: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE Reviews in Biomedical Engineering*, 5:60–73, 2012.

[CCG09]   C. Chang, J. P. Cunningham, and G. H. Glover. Influence of heart rate on the bold signal: the cardiac response function. *Neuroimage*, 44(3):857–869, 2009.

[CG09]   C. Chang and G. H. Glover. Relationship between respiration, end-tidal co 2, and bold signals in resting-state fmri. *Neuroimage*, 47(4):1381–1393, 2009.

[CG10]   C. Chang and G. H. Glover. Time–frequency dynamics of resting-state brain connectivity measured with fmri. *Neuroimage*, 50(1):81–98, 2010.

[CHA+12]   I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, and M. A. Lindquist. Dynamic connectivity regression: determining state-related changes in brain connectivity. *Neuroimage*, 61(4):907–920, 2012.

[CHC+02]   D. Cordes, V. Haughton, J. D. Carew, K. Arfanakis, and K. Maravilla. Hierarchical clustering to measure connectivity in fmri resting-state data. *Magnetic resonance imaging*, 20(4):305–317, 2002.

[CLT+11]   W. Cui, S. Liu, L. Tan, C. Shi, Y. Song, Z. J. Gao, H. Qu, and X. Tong. Textflow: Towards better understanding of evolving topics in text. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2412–2421, 2011.

[CMPA14]   V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı. The chronnectome: time-varying connectivity networks as the next frontier in fmri data discovery. *Neuron*, 84(2):262–274, 2014.

[CSB10]   D. M. Cole, S. M. Smith, and C. F. Beckmann. Advances and pitfalls in the analysis and interpretation of resting-state fmri data. *Frontiers in systems neuroscience*, 4(8), 2010.

[CWBFF09] B. Chai, D. Walther, D. Beck, and L. Fei-Fei. Exploring functional connectivities of the human brain using multivariate information analysis. In *Advances in neural information processing systems*, pages 270–278, 2009.

[DAB+14]   E. Damaraju, E. A. Allen, A. Belger, J. M. Ford, S. McEwen, D. H. Mathalon, B. A. Mueller, G. D. Pearlson, S. G. Potkin, A. Preda, et al. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5:298–308, 2014.

[DB12]   X. Di and B. B. Biswal. Metabolic brain covariant networks as revealed by fdg-pet with reference to resting-state fmri networks. *Brain connectivity*, 2(5):275–283, 2012.

[DBG12]   K. W. Doron, D. S. Bassett, and M. S. Gazzaniga. Dynamic network structure of interhemispheric coordination. *Proceedings of the National Academy of Sciences*, 109(46):18661–18668, 2012.

[DDM⁺10]   P. Delamillieure, G. Doucet, B. Mazoyer, M. R. Turbelin, N. Delcroix, E. Mellet, L. Zago, F. Crivello, L. Petit, N. Tzourio-Mazoyer, and M. Joliot. The resting state questionnaire: An introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain Research Bulletin*, 81(6):565 – 573, 2010.

[DIH99]   M. S. Dagli, J. E. Ingeholm, and J. V. Haxby. Localization of cardiac-induced signal change in fmri. *Neuroimage*, 9(4):407–415, 1999.

[DLBDS⁺06] M. De Luca, C.F. Beckmann, N. De Stefano, P.M. Matthews, and S.M. Smith. fmri resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*, 29(4):1359 – 1367, 2006.

[DLSDS⁺05] M. De Luca, S. Smith, N. De Stefano, A. Federico, and P. M. Matthews. Blood oxygenation level dependent contrast resting state networks are relevant to functional activity in the neocortical sensorimotor system. *Experimental brain research*, 167(4):587–594, 2005.

[DO14]   M. DiLuca and J. Olesen. The cost of brain diseases: a burden or a challenge? *Neuron*, 82(6):1205–1208, 2014.

[dPDPS⁺10] F. de Pasquale, S. Della Penna, A. Z. Snyder, C. Lewis, D. Mantini, L. Marzetti, P. Belardinelli, L. Ciancetta, V. Pizzella, G. L. Romani, et al. Temporal dynamics of spontaneous meg activity in brain networks. *Proceedings of the National Academy of Sciences*, 107(13):6040–6045, 2010.

[DRB⁺06]   J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann. Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, 103(37):13848–13853, 2006.

[dRVdH13]   M. A. de Reus and M. P. Van den Heuvel. The parcellation-based connectome: limitations and extensions. *Neuroimage*, 80:397–404, 2013.

[DYB10]   J.-C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *Proceedings of the National Academy of Sciences*, 107(29):12755–12760, 2010.

134

[EVDVSV11] H. Eryilmaz, D. Van De Ville, S. Schwartz, and P. Vuilleumier. Impact of transient emotions on functional connectivity during subsequent resting state: a wavelet correlation approach. *Neuroimage*, 54(3):2481–2491, 2011.

[FB07] S. Fortunato and M. Barthélemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.

[FFLF93] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. J. Frackowiak. Functional connectivity: The principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism*, 13(1):5–14, 1993.

[FHT01] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.

[For10] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.

[FRCA14] F. Fallani, J. Richiardi, M. Chavez, and S. Achard. Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1653), 2014.

[Fri11] K. J. Friston. Functional and effective connectivity: a review. *Brain connectivity*, 1(1):13–36, 2011.

[GDC10] D. Greene, D. Doyle, and P. Cunningham. Tracking the evolution of communities in dynamic social networks. In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 176–183, Aug 2010.

[Gri05] M. G. Grigorov. Global properties of biological networks. *Drug discovery today*, 10(5):365–372, 2005.

[HCPP14] A. E. Hudson, D. P. Calderon, D. W. Pfaff, and A. Proekt. Recovery of consciousness is mediated by a network of discrete metastable activity states. *Proceedings of the National Academy of Sciences*, 111(25):9283–9288, 2014.

[HD15] H. Huang and M. Ding. Linking functional connectivity and structural connectivity quantitatively: A comparison of methods. *Brain connectivity*, 6(2):99–108, 2015.

[HMP+13] K. Hahn, N. Myers, S. Prigarin, K. Rodenacker, A. Kurz, H. Förstl, C. Zimmer, A. M. Wohlschläger, and C. Sorg. Selectively and progressively disrupted structural connectivity of functional brain networks

in alzheimer's disease—revealed by a novel framework to analyze edge distributions of networks detecting disruptions with strong statistical evidence. *Neuroimage*, 81:96–109, 2013.

[HO00]   A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.

[HPV$^+$11]   J. Hlinka, M. Paluš, M. Vejmelka, D. Mantini, and M. Corbetta. Functional connectivity in resting-state fmri: is linear correlation sufficient? *Neuroimage*, 54(3):2218–2225, 2011.

[HRGCB12]   D. A. Handwerker, V. Roopchansingh, J. Gonzalez-Castillo, and P. A. Bandettini. Periodic changes in fmri connectivity. *Neuroimage*, 63(3):1712–1719, 2012.

[HSC$^+$09]   C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040, 2009.

[HSGJ$^+$12]   L. Heine, A. Soddu, F. A. Gomez Jaramillo, A. Vanhaudenhuyse, L. Tshibanda, M. Thonnard, V. Charland-Verville, M. Kirsch, S. Laureys, and A. Demertzi. Resting state networks and consciousness alterations of multiple resting state network connectivity in physiological, pharmacological and pathological consciousness states. *Frontiers in Psychology*, 3(3):295, 2012.

[HSM04]   S. A. Huettel, A. W. Song, and G. McCarthy. *Functional magnetic resonance imaging*, volume 1. Sinauer Associates Sunderland, 2004.

[HWA$^+$13]   R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80:360–378, 2013.

[HWG$^+$13]   R. M. Hutchison, T. Womelsdorf, J. S. Gati, S. Everling, and R. S. Menon. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human brain mapping*, 34(9):2154–2177, 2013.

[HWH11]   O. Hauk, D. G. Wakeman, and R. Henson. Comparison of noise-normalized minimum norm estimates for meg analysis using multiple resolution metrics. *Neuroimage*, 54(3):1966–1974, 2011.

[JGGF16]   A. José-García and W. Gómez-Flores. Automatic clustering using nature-inspired metaheuristics: A survey. *Applied Soft Computing*, 41:192–213, 2016.

[JVM$^+$12]   D. T. Jones, P. Vemuri, M. C. Murphy, J. L. Gunter, M. L. Senjem, M. M. Machulda, S. A. Przybelski, B. E. Gregg, K. Kantarci, D. S. Knopman, et al. Non-stationarity in the "resting brain's" modular architecture. *PloS ONE*, 7(6):1–15, 2012.

[KBC$^+$92]   K. K. Kwong, J. W. Belliveau, D. A. Chesler, I. E. Goldberg, R. M. Weisskoff, P. Poncelet, D. N. Kennedy, B. E. Hoppel, Cohen. M. S., R. Turner, H. M. Cheng, T. J. Brady, and B. R. Rosen. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences USA*, 89:5675–5679, 1992.

[KD14]   A. Kucyi and K. D. Davis. Dynamic functional connectivity of the default mode network tracks daydreaming. *Neuroimage*, 100:471–480, 2014.

[KHS$^+$11]   M. G. Kitzbichler, R. N. A. Henson, M. L. Smith, P. J. Nathan, and E. T. Bullmore. Cognitive effort drives workspace configuration of human brain functional networks. *The Journal of Neuroscience*, 31(22):8259–8270, 2011.

[KIE$^+$12]   P. Kundu, S. J. Inati, J. W. Evans, W.-M. Luh, and P. A. Bandettini. Differentiating bold and non-bold signals in fmri time series using multi-echo epi. *Neuroimage*, 60(3):1759–1770, 2012.

[KLL06]   T. Kim, I. Lee, and T.-W. Lee. Independent vector analysis: definition and algorithms. In *Fortieth Asilomar Conference on Signals, Systems and Computers (ACSSC'06)*, pages 1393–1396. IEEE, 2006.

[KPCP12]   J. Kinnison, S. Padmala, J. M. Choi, and L. Pessoa. Network analysis reveals increased integration during emotional and motivational processing. *The Journal of Neuroscience*, 32(24):8361–8372, 2012.

[KVDV15]   F. I. Karahanoğlu and D. Van De Ville. Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks. *Nature Communications*, 6, 2015.

[Lau73]   P. C. Lauterbur. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature*, 242(5394):190–191, March 1973.

[LD13]   X. Liu and J. H. Duyn. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11):4392–4397, 2013.

[LFE$^+$11]   A. R. Laird, P. M. Fox, S. B. Eickhoff, J. A. Turner, K. L. Ray, D. R. McKay, D. C. Glahn, C. F. Beckmann, S. M. Smith, and P. T. Fox.

Behavioral interpretations of intrinsic connectivity networks. *Journal of cognitive neuroscience*, 23(12):4022–4037, 2011.

[LHP+13] L. Li, X. Hu, T. M. Preuss, M. F. Glasser, F. W. Damen, Y. Qiu, and J. Rilling. Mapping putative hubs in human, chimpanzee and rhesus macaque connectomes via diffusion tractography. *NeuroImage*, 80:462 – 474, 2013.

[LLT+16] Y. B. Lee, J. Lee, S. Tak, K. Lee, D. L. Na, S. W. Seo, Y. Jeong, J. C. Ye, et al. Sparse spm: Group sparse-dictionary learning in spm framework for resting-state functional connectivity mri analysis. *NeuroImage*, 125:1032–1045, 2016.

[LMH13] Erwan Le Martelot and Chris Hankin. Multi-scale community detection using stability optimisation. *International Journal of Web Based Communities*, 9(3):323–348, 2013.

[LOFB14] E. R. Lowther, J. T. O'Brien, M. J. Firbank, and A. M. Blamire. Lewy body compared with alzheimer dementia is associated with decreased functional connectivity in resting state networks. *Psychiatry Research: Neuroimaging*, 223(3):192–201, 2014.

[LPZN+09] L. J. Larson-Prior, J. M. Zempel, T. S. Nolan, F. W. Prior, A. Z. Snyder, and M. E. Raichle. Cortical network functional connectivity in the descent to sleep. *Proceedings of the National Academy of Sciences*, 106(11):4489–4494, 2009.

[LRG+13] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J. M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville. Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *NeuroImage*, 83:937–950, 2013.

[LRVDV13] N. Leonardi, J. Richiardi, and D. Van De Ville. Functional connectivity eigennetworks reveal different brain dynamics in multiple sclerosis patients. In *2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI)*, pages 528–531, 2013.

[LSGVDV14] N. Leonardi, W. R. Shirer, M. D. Greicius, and D. Van De Ville. Disentangling dynamic networks: separated and joint expressions of functional connectivity patterns in time. *Human brain mapping*, 35(12):5984–5995, 2014.

[LSS13] M.H. Lee, C.D. Smyser, and J.S. Shimony. Resting-state fmri: A review of methods and clinical applications. *American Journal of Neuroradiology*, 34(10):1866–1872, 2013.

[LVDV15]   N. Leonardi and D. Van De Ville. On spurious and real fluctuations of dynamic functional connectivity during rest. *Neuroimage*, 104:430–436, 2015.

[LXNC14]   M. A. Lindquist, Y. Xu, M. B. Nebel, and B. S. Caffo. Evaluating dynamic bivariate correlations in resting-state fmri: A comparison study and a new approach. *Neuroimage*, 101:531–546, 2014.

[LYZ+15]   X. Liao, L. Yuan, T. Zhao, Z. Dai, N. Shu, M. Xia, Y. Yang, A. Evans, and Y. He. Spontaneous functional network dynamics and associated structural substrates in the human brain. *Frontiers in Human Neuroscience*, 9, 2015.

[LZL+15]   J. Liu, L. Zhao, F. Lei, Y. Zhang, K. Yuan, Q. Gong, F. Liang, and J. Tian. Disrupted resting-state functional connectivity and its changing trend in migraine suffers. *Human brain mapping*, 36(5):1892–1907, 2015.

[LZP+15]   R. Liégeois, E. Ziegler, C. Phillips, P. Geurts, F. Gómez, M. Bahri, B. T. T. Yeo, A. Soddu, A. Vanhaudenhuyse, S. Laureys, and R. Sepulchre. Cerebral functional connectivity periodically (de)synchronizes with anatomical constraints. *Brain Structure and Function*, pages 1–13, 2015.

[MBH+09]   K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage*, 44(3):893–905, 2009.

[MHM+16]   A. Medda, L. Hoffmann, M. Magnuson, G. Thompson, W. J. Pan, and S. Keilholz. Wavelet-based clustering of resting state mri data in the rat. *Magnetic resonance imaging*, 34(1):35–43, 2016.

[MLPN+08]   M. McAvoy, L. Larson-Prior, T. S. Nolan, S. N. Vaishnavi, M. E. Raichle, and G. d'Avossa. Resting states affect spontaneous bold oscillations in sensory and paralimbic cortex. *Journal of Neurophysiology*, 100(2):922–931, 2008.

[MMH+11]   W. Majeed, M. Magnuson, W. Hasenkamp, H. Schwarb, E. H. Schumacher, L. Barsalou, and S. D. Keilholz. Spatiotemporal dynamics of low frequency bold fluctuations in rats and humans. *Neuroimage*, 54(2):1140–1150, 2011.

[MMVDV15]   D.-E. Meskaldji, S. Morgenthaler, and D. Van De Ville. New measures of brain functional connectivity by temporal analysis of extreme events. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 26–29, April 2015.

[NJW⁺02]   A. Y. Ng, M. I. Jordan, Y. Weiss, et al. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 2:849–856, 2002.

[OHN⁺14]   S. W. Oh, J. A. Harris, L. Ng, B. Winslow, N. Cain, S. Mihalas, Q. Wang, C. Lau, L. Kuan, A. M. Henry, et al. A mesoscale connectome of the mouse brain. *Nature*, 508(7495):207–214, 2014.

[PAMB⁺01]   M. Poustchi-Amin, S. A. Mirowitz, J. J. Brown, R. C. McKinstry, and T. Li. Principles and applications of echo-planar imaging: A review for the general radiologist. *RadioGraphics*, 21(3):767–779, 2001. PMID: 11353123.

[PBS⁺12]   J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen. Spurious but systematic correlations in functional connectivity mri networks arise from subject motion. *Neuroimage*, 59(3):2142–2154, 2012.

[PDH13]   D. Picchioni, J. H. Duyn, and S. G. Horovitz. Sleep and the functional connectome. *Neuroimage*, 80:387–396, 2013.

[PDMA12]   F. Pozzi, T. Di Matteo, and T. Aste. Exponential smoothing weighted correlations. *The European Physical Journal B*, 85(6):1–21, 2012.

[PMN11]   R. A. Poldrack, J. A. Mumford, and T. E. Nichols. *Handbook of functional MRI data analysis.* Cambridge University Press, 2011.

[POWJ15]   M. Pedersen, A. H. Omidvarnia, J. M. Walz, and G. D. Jackson. Increased segregation of brain networks in focal epilepsy: an fmri graph theory finding. *NeuroImage: Clinical*, 8:536–542, 2015.

[RAW15]   L. F. Robinson, L. Y. Atlas, and T. D. Wager. Dynamic functional connectivity using state-based dynamic community structure: method and application to opioid analgesia. *NeuroImage*, 108:274–291, 2015.

[RDPC14]   B. Rashid, E. Damaraju, G. D. Pearlson, and V. D. Calhoun. Dynamic connectivity states estimated from resting fmri identify differences among schizophrenia, bipolar disorder, and healthy control subjects. *Frontiers in Human Neuroscience*, 8, 2014.

[Rou87]   P. J.. Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.

[SFM⁺09]   S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann. Correspondence of the brain's functional architecture during

activation and rest. *Proceedings of the National Academy of Sciences*, 106(31):13040–13045, 2009.

[SHB⁺15]   K. Shen, R. M. Hutchison, G. Bezgin, S. Everling, and A. R. McIntosh. Network structure shapes spontaneous functional connectivity dynamics. *The Journal of Neuroscience*, 35(14):5579–5588, 2015.

[SKB⁺15]   J. M. Shine, O. Koyejo, P. T. Bell, K. J. Gorgolewski, M. Gilat, and R. A. Poldrack. Estimation of dynamic functional connectivity using multiplication of temporal derivatives. *NeuroImage*, 122:399–407, 2015.

[SKR⁺09]   Z. Shehzad, A. M. C. Kelly, P. T. Reiss, D. G. Gee, K. Gotimer, L. Q. Uddin, S. H. Lee, D. S. Margulies, A. K. Roy, B. B. Biswal, et al. The resting brain: unconstrained yet reliable. *Cerebral Cortex*, 19(10):2209–2229, 2009.

[SMD04]   F. T. Sun, L. M. Miller, and M. D'Esposito. Measuring interregional functional connectivity using coherence and partial coherence analyses of fmri data. *Neuroimage*, 21(2):647–658, 2004.

[SMM⁺12]   S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, et al. Temporally-independent functional modes of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 109(8):3131–3136, 2012.

[SMS⁺07]   W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *The Journal of neuroscience*, 27(9):2349–2356, 2007.

[SPK⁺10]   Ü. Sakoğlu, G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun. A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 23(5-6):351–366, 2010.

[Spo13]   O. Sporns. Structure and function of complex brain networks. *Dialogues Clin Neurosci*, 15(3):247–262, 2013.

[SRR⁺11]   W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, 22(1):158–165, 2011.

[SSC⁺05]   R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. D. Bullmore. Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral cortex*, 15(9):1332–1342, 2005.

[Str11]   P. W. Stroman. *Essentials of functional MRI*. CRC Press, 2011.

[SVB+13]   S.M. Smith, D. Vidaurre, C.F. Beckmann, M.F. Glasser, M. Jenkinson, K.L. Miller, T.E. Nichols, E.C. Robinson, G. Salimi-Khorshidi, M.W. Woolrich, D.M. Barch, K. Uurbil, and D.C. Van Essen. Functional connectomics from resting-state fmri. *Trends in Cognitive Sciences*, 17(12):666 – 682, 2013. Special Issue: The Connectome.

[SvGdZ+07]   K. Shmueli, P. van Gelderen, J. A. de Zwart, S. G. Horovitz, M. Fukunaga, J. M. Jansma, and J. H. Duyn. Low-frequency fluctuations in the cardiac rate as a source of variance in the resting-state fmri bold signal. *Neuroimage*, 38(2):306–320, 2007.

[SZPBPH15]   N. Shevzov-Zebrun, N. M. Petrovich Brennan, K. K. Peck, and A. I. Holodny. Chapter 3 - advanced functional imaging: fmri, pet, and meg. In A. J. Golby, editor, *Image-Guided Neurosurgery*, pages 63 – 89. Academic Press, Boston, 2015.

[TC98]   C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1):61–78, 1998.

[TDP06]   B. Thirion, S. Dodel, and J. B.. Poline. Detection of signal synchronizations in resting-state fmri datasets. *Neuroimage*, 29(1):321–327, 2006.

[TF15]   W. H. Thompson and P. Fransson. The frequency dimension of fmri dynamic connectivity: Network connectivity, functional hubs and integration in the resting brain. *Neuroimage*, 121:227–242, 2015.

[TFA+08]   J. Tohka, K. Foerde, A. R. Aron, S. M. Tom, A. W. Toga, and R. A. Poldrack. Automatic independent component labeling for artifact removal in fmri. *Neuroimage*, 39(3):1227–1245, 2008.

[TL14]   E. Tagliazucchi and H. Laufs. Decoding wakefulness levels from typical fmri resting-state data reveals reliable drifts between wakefulness and sleep. *Neuron*, 82(3):695–708, 2014.

[TSFZ11]   M. Takaffoli, F. Sangi, J. Fagnan, and O.R. Zaïane. Community evolution mining in dynamic social networks. *Procedia - Social and Behavioral Sciences*, 22:49 – 58, 2011.

[VBAW15]   C. Vehlow, F. Beck, P. Auwärter, and D. Weiskopf. Visualizing the evolution of communities in dynamic graphs. *Computer Graphics Forum*, 34(1):277–288, 2015.

[VdHHP10]   M. P. Van den Heuvel and H. E. Hulshoff Pol. Exploring the brain network: A review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20(8):519 – 534, 2010.

[vdHKGS12]   M. P. van den Heuvel, R. S. Kahn, J. Goñi, and O. Sporns. High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, 109(28):11372–11377, 2012.

[VdHMKHP09]   M. P. Van den Heuvel, R. C. W. Mandl, R. S. Kahn, and Hilleke E. Hulshoff P. Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping*, 30(10):3127–3141, 2009.

[VDHMP08]   M. Van Den Heuvel, R. Mandl, and H. H. Pol. Normalized cut group clustering of resting-state fmri data. *PloS ONE*, 3(4):1–11, 2008.

[VDHV+10]   K. R. A. Van Dijk, T. Hedden, A. Venkataraman, K. C. Evans, S. W. Lazar, and R. L. Buckner. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of Neurophysiology*, 103(1):297–321, 2010.

[VEUA+12]   D.C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T.E.J. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S.W. Curtiss, S. Della Penna, D. Feinberg, N. Glasser, M.F.and Harel, A.C. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S.E. Petersen, F. Prior, B.L. Schlaggar, S.M. Smith, A.Z. Snyder, J. Xu, and E. Yacoub. The human connectome project: A data acquisition perspective. *NeuroImage*, 62(4):2222 – 2231, 2012.

[VL07]   U. Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.

[VLRM01]   F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: phase synchronization and large-scale integration. *Nature reviews neuroscience*, 2(4):229–239, 2001.

[WC03]   X. F. Wang and G. Chen. Complex networks: small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, 3(1):6–20, 2003.

[WEC10]   L. Wu, T. Eichele, and V. D. Calhoun. Reactivity of hemodynamic responses and functional connectivity to different states of alpha synchrony: A concurrent eeg-fmri study. *NeuroImage*, 52(4):1252 – 1260, 2010.

[WS98]   D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *Nature*, 393(6684):440–442, 1998.

[YAMC15]   M. Yaesoubi, E. A. Allen, R. L. Miller, and V. D. Calhoun. Dynamic coherence analysis of resting fmri data to jointly capture state-based phase, frequency, and time-domain information. *Neuroimage*, 120:133–142, 2015.

[YDZ+15]   H. Yuan, L. Ding, M. Zhu, V. Zotev, R. Phillips, and J. Bodurka. Reconstructing large-scale brain resting-state networks from high-resolution eeg: Spatial and temporal comparisons with fmri. *Brain connectivity*, 2015.

[YKS+11]   B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 106(3):1125–1165, 2011.

[YLH+09]   C. Yan, D. Liu, Y. He, Q. Zou, C. Zhu, X. Zuo, X. Long, and Y. Zang. Spontaneous brain activity in the default mode network is sensitive to different resting-state conditions with limited cognitive load. *PLoS ONE*, 4(5):1–11, 05 2009.

[ZFC+14]   A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear. Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences*, 111(28):10341–10346, 2014.

[ZWB+04]   K. H. Zou, S. K. Warfield, A. Bharatha, C. M. C. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells, F. A. Jolesz, and R. Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index 1: Scientific reports. *Academic Radiology*, 11(2):178–189, 2004.