TÉCNICO LISBOA

Brain Functional Connectivity in Schizophrenia

Disrupted network dynamics
revealed by functional magnetic resonance imaging

Miguel Lourenço Figueira Francisco Farinha

Thesis to obtain the Master of Science Degree in

Mathematics and Applications

Supervisors: Dr. Joana Ribeiro Barbosa Cabral
Prof. Dr. Maria da Conceição Esperança Amado

Examination Committee

Chairperson: Prof. Dr. António Manuel Pacheco Pires
Supervisor: Prof. Dr. Maria da Conceição Esperança Amado
Member of the Committee: Prof. Dr. Patrícia Margarida Piedade Figueiredo

November 2021

Acknowledgments

Em primeiro lugar, deixo o meu agradecimento à Doutora Joana Cabral que, apesar de não me conhecer, prontamente me deu a oportunidade de ser seu mestrando e que desde o início confiou em mim. Graças à sua orientação, pude descobrir que é nesta área que quero continuar a trabalhar. Obrigado por ter guiado o meu trabalho com rigor e dedicação, por ter acreditado nas minhas capacidades e valorizado a minha autonomia. Agradeço ainda toda a sua disponibilidade e o altruísmo com que partilhou o seu saber, que foram essenciais para o desenvolvimento desta dissertação. Espero que o futuro me traga a oportunidade de vir a colaborar novamente consigo.

À minha orientadora, Professora Conceição Amado, agradeço igualmente a disponibilidade que manifestou desde o início para me acompanhar e, em especial, a paciência e compreensão que sempre demonstrou mesmo nos momentos mais difíceis. Agradeço também todos os conselhos e sugestões que me deu, os quais contribuíram para enriquecer a dissertação que agora termino, que me proporcionou uma sólida e profícua aprendizagem.

Gostaria de agradecer ao Doutor Joaquim Reis por me ter orientado e me ter incentivado a procurar informação sobre esta área de estudo.

Agradeço a todos os meus amigos que me acompanharam ao longo do meu percurso e que me ajudaram a chegar até aqui. I would like to thank my friend Ben Harrison for always having time to help me, for always having kind and inspirational words for me and for helping me when I overthink.

Quero agradecer do fundo do meu coração aos meus pais e ao meu irmão. Convosco ao meu lado tudo faz sentido e tudo parece menos assustador. Vocês dão-me força para lutar e dar o meu melhor todos os dias. Obrigado por terem acreditado sempre em mim e por todo o vosso amor incondicional.

Abstract

Brain activity during rest has been demonstrated to evolve through a repertoire of functional connectivity (FC) patterns, whose alterations may provide biomarkers of schizophrenia - a psychotic disorder characterised by dysfunctional brain connectivity. In this study, differences between the dynamic exploration of resting-state networks using functional magnetic resonance imaging (fMRI) data from 71 schizophrenia patients and 74 healthy controls were investigated using a method focusing on the dominant fMRI signal phase coherence pattern at each time point. Through the lens of dynamical systems theory, brain activity in the form of temporal FC state trajectories was examined for intergroup differences by calculating the fractional occupancy, dwell time, limiting probability of each state and the transition probabilities between states. Results showed a medium to large effect size for a globally synchronised state which was found to have decreased fractional occupancy in schizophrenia. The fractional occupancy and limiting probability of FC states overlapping with canonical functional subsystems was identified to have increased in schizophrenia, with small to medium effect sizes. Medium to large effect size for the reduced probability of remaining in a global integrative state and for a number of altered state-to-state transition probabilities were revealed in schizophrenia. Finally, this study showed that using K-medoids clustering did not influence the observed intergroup differences - highlighting the utility of dynamical systems theory to better understand brain activity. Combined, these findings expose pronounced differences between schizophrenia patients and healthy controls - supporting and extending current knowledge regarding disrupted brain dynamics in schizophrenia.

Keywords

Resting-state fMRI, Dynamic functional connectivity, Functional networks, LEiDA, Schizophrenia

Resumo

A atividade cerebral em repouso evolui através de um repertório de padrões de conectividade funcional, cujas alterações poderão fornecer biomarcadores de esquizofrenia. Neste estudo foram investigadas diferenças entre a exploração dinâmica de redes de repouso em 71 pacientes esquizofrénicos e 74 controlos saudáveis, usando um método baseado no padrão de coerência de fases do sinal de Ressonância Magnética em cada instante de tempo. Na perspetiva da teoria de sistemas dinâmicos, foram examinadas diferenças, entre grupos, na atividade cerebral concebida como trajetórias entre estados, calculando a probabilidade de ocorrência, tempo de permanência, probabilidade limite de cada estado e as probabilidades de transição entre estados. Os resultados mostraram tamanho de efeito médio a elevado para um estado globalmente sincronizado, cuja probabilidade de ocorrência estava reduzida em esquizofrenia. A probabilidade de ocorrência e a probabilidade limite de estados relacionados com subsistemas funcionais canônicos foram identificadas como aumentadas em esquizofrenia, com tamanhos de efeito pequenos a médios. Em esquizofrenia, foi revelado tamanho de efeito elevado para a probabilidade reduzida de permanecer num estado global e ainda tamanhos de efeito médios para um conjunto de probabilidades de transição de estado alteradas. Por último, constatou-se que as diferenças observadas entre grupos não eram influenciadas pelo uso de agrupamento de K-medóides - salientando a maior utilidade da teoria de sistemas dinâmicos para a compreensão da atividade cerebral. Estes resultados, expondo diferenças demarcadas entre pacientes e controlos, reforçam e expandem o conhecimento atual sobre o dinamismo cerebral alterado na esquizofrenia.

Palavras Chave

Ressonância magnética funcional em repouso, Conectividade funcional dinâmica, Redes funcionais, LEiDA, Esquizofrenia

Contents

List of Figures

List of Tables

Acronyms

ARI	Adjusted Rand Index
ALFF	Amplitude of Low Frequency Fluctuations
ASD	Autism Spectrum Disorder
AAL90	Subset of 90 non-cerebellar Automated Anatomic Labeling brain areas
BP	Bipolar Disorder
BOLD	Blood Oxygenation Level Dependent
CLT	Central Limit Theorem
CDF	Cumulative Distribution Function
DMN	Default Mode Network
DTMC	Discrete Time Markov Chain
dFC	dynamic Functional Connectivity
EEG	Electroencephalogram
FD	Framewise Displacement
FC	Functional Connectivity
fMRI	Functional Magnetic Resonance Imaging
GSR	Global Signal Regression
GIG-ICA	Group Information Guided ICA
HC	Healthy Control
iid	independent and identically distributed
ICA	Independent Component Analysis
IVA	Independent Vector Analysis
LEiDA	Leading Eigenvector Dynamics Analysis

MRI	Magnetic Resonance Imaging
MEG	Magnetoencephalography
MNI	Montreal Neurological Institute
NaN	“not a number”
PC	Phase Coherence
PCC	Phase Coherence Connectivity
PS	Phase Synchronization
PCA	Principal Component Analysis
ROI	region of interest
ReHo	Regional Homogeneity
TR	Repetition Time
rs-fMRI	resting-state fMRI
RSN	Resting-State Network
SAD	Schizoaffective Disorder
SZ	Schizophrenia
SWA	Sliding-Window Analysis
std	standard deviation
sFC	static Functional Connectivity
TPM	Transition Probability Matrix
VI	Variation of Information

1

Introduction

Contents

1.1 Motivation

The use of Magnetic Resonance Imaging (MRI) data has increased significantly in recent years, resulting in a deeper understanding of the foundations of brain structure and function. In fact, the application of this technique within the clinical neuroimaging field has brought about significant breakthroughs, with the identification and validation of clinical biomarkers of diseases being one of the most relevant [1–4].

Schizophrenia, a severe neuropsychiatric disorder, has been the subject of an abundance of research in previous years. Notably, neural activity within this disorder has been studied both from an anatomical and functional perspective contributing with essential insights to the understanding of this disease [5–7]. Despite these efforts, the neuropathology of schizophrenia remains unknown, impeding the development of clinical diagnostics and effective treatments [7, 8].

A special focus regarding Functional Magnetic Resonance Imaging (fMRI) research lies on the examination of the dynamic fluctuations of brain connections, i.e., the way different brain areas communicate with each other. A plethora of promising methodologies to examine time-resolved fluctuations in connectivity have been proposed [9, 10]. The referred approaches were extensively applied to neuropsychiatric disorders, such as schizophrenia, due to their ability to detect transient changes in neural signalling which were hypothesised to underlie such disorders [11]. Nevertheless, inconsistent methodological considerations and statistical challenges for studying the dynamic fluctuations in brain activity in fMRI hinder the identification and validation of a universal biomarker for schizophrenia [10, 12, 13]. Furthermore, the clinical utility of such a marker, assuming one is found, is still questionable [3]. In this project, a novel methodology to study time-resolved neural activity will be applied to a large sample of patients with schizophrenia. This is to deal with the mentioned shortcomings and provide reliable and biologically meaningful biomarkers of schizophrenia.

1.2 Problem

Given input data collected from fMRI it is possible to characterise the spatiotemporal organisation of the functional interactions in both healthy and diseased brains. Initially, the brain's functional organisation was assessed over an entire scan, through the computation of average connectivity maps. Nevertheless, the clinical utility of such approaches is unclear [11]. As a result, recent years have seen a rapidly growing interest in studying time-resolved brain activity, which has been suggested to provide more sensitive and specific biomarkers of disease [9, 11, 13]. In fact, the time-resolved characterisation of the activity within the brain has enabled the detection of aberrant connectivity across regions, which is consistent with existing theories of psychiatric disorders that suggest impaired brain functionality [1, 14]. Notably, these approaches have propelled the identification of potentially promising biomarkers of several disorders, including schizophrenia [6, 7, 9, 11, 15].

A plethora of time-resolved techniques may be utilised to explore and characterise the temporal fluctuations of brain activity in schizophrenia patients. Nevertheless, the study of time-varying connectivity is not straightforward and several open questions remain under debate. On the one hand, multiple techniques and methodologies to extract and quantify time-resolved fluctuations in brain activity have been used in the literature, therefore hindering comparability between studies [9, 10, 13]. In addition, estimates of time-varying connectivity are strongly influenced by the preprocessing steps applied to the raw fMRI data [9, 10, 16, 17]. On the other hand, it is uncertain whether these approaches capture information of relevance regarding brain function, or simply resolve methodology-related artefacts [9, 10, 13, 18]. An additional important concern when assessing time-varying connectivity regards the omission or misuse of appropriate statistical testing of connectivity temporal variations [9, 10, 13]. Such procedures are paramount in order to guarantee that the observed connectivity temporal fluctuations are in fact caused by transient changes in neural activity, instead of simple artefacts or noise. Furthermore, it is critical that estimates of dynamic functional connections are reliable and sensitive enough to serve as robust biomarkers of disease. This is because they may illuminate significant intergroup differences that enable the detection and potential treatment of such diseased populations.

The aforementioned issues exacerbate the need to use approaches with a robust methodology and solid statistical analysis to facilitate the interpretability and comparability of relevant differences in brain function between healthy and diseased populations.

1.3 Objectives

Schizophrenia affects around 20 million people worldwide and, if left untreated, its symptoms can be persistent and disabling [19]. Hence, there is a great need for a better understanding of the pathology of this disease. Understanding whole-brain patterns of functional connections in patients with schizophrenia would be the first step towards discovering clinically relevant biomarkers that might have the potential to diagnose the disorder, to determine its prognosis, and to predict and monitor a patient's response to interventions.

The central aim of this MSc thesis is the detection of statistically significant differences between the neural activity observed in schizophrenia patients and healthy controls that could contribute with relevant biomarkers of this disease. To this effect, a recently proposed method is used to investigate whole-brain time-varying connectivity fluctuations observed in an open-source fMRI neuroimaging dataset from patients with schizophrenia and healthy controls [20].

This method, designated Leading Eigenvector Dynamics Analysis (LEiDA) [21], clusters the dominant functional connectivity patterns captured at each time point to define a repertoire of Blood Oxygenation Level Dependent (BOLD) phase coherence connectivity patterns which transiently emerge and dissolve

during neural activity - defining time courses of recurrent patterns of BOLD phase coherence. Firstly, this study assesses whether intergroup differences exist by characterising the estimated time courses of recurrent connectivity patterns using tools from dynamical systems theory. Specifically, differences between schizophrenia patients and healthy controls are investigated in terms of the probability of occurrence, the mean duration, the limiting probability of each pattern of connectivity and the one-step transition probabilities between patterns. Secondly, this research examines the quality and stability of the partitions resulting from the clustering procedure. This is because previous studies employing LEiDA have not closely examined the validity of the partitions of the data, tending to focus primarily on detecting differences between groups. Lastly, this project attempts to understand the influence of using an alternative clustering algorithm on the ability to differentiate patients with schizophrenia from healthy controls. This work hypothesised to find altered time-resolved functional connectivity in schizophrenia patients compared to healthy controls characterised by: (1) reduced excursions to a connectivity pattern possibly involved in the integration of segregated functional connections and (2) increased excursions to a number of functionally segregated connectivity patterns.

1.4 Outline

This project is organised into six chapters. The current chapter serves as an introductory overview of the problem under study, by stating its relevance and the objectives of this work. Chapter 2 provides a comprehensive literature review to provide context for the current research. This chapter starts by briefly describing the transverse workflow of the fMRI technique, followed by the discussion of analytical tools and theoretical concepts relevant for this study. Chapter 2 ends with the review of relevant past studies about time-resolved brain functional connectivity in schizophrenia measured by fMRI. Chapter 3 presents the materials and methods used to accomplish the objectives of this project. Firstly, the fMRI data and associated processing steps are outlined. Then a detailed description of the analytic approach used to capture the dynamic connectivity patterns is provided. This chapter ends by presenting the tools necessary to evaluate, compare and interpret the results. Chapter 4 outlines the findings obtained throughout this project. Results from the application of LEiDA to the neuroimaging data are provided in the first section. Chapter 4 ends by evaluating the influence of using a distinct clustering algorithm while conducting a LEiDA analysis. Chapter 5 provides the discussion and interpretation of the findings of the previous chapter. Chapter 6 reflects on the findings and contributions of this work. Limitations of this study, as well as areas of focus and recommendations for future work are also discussed.

2

Literature Review

Contents

This chapter aims at providing the reader with necessary background for the current research. The chapter starts by presenting a summary overview of schizophrenia in Section 2.1. Section 2.2 presents the fundamentals aspects of fMRI research. Section 2.3 surveys seminal findings of the resting-state fMRI studies. Relevant analytical methods to investigate resting-state brain activity are described in Section 2.4. Section 2.5 explains the theoretical background of two-sample location permutation tests. The chapter concludes by presenting the state of the art in resting-state schizophrenia research.

2.1 Schizophrenia

Schizophrenia is a chronic brain disorder characterised by disruption to thought processes, perceptions, emotional responsiveness and social interactions. This mental illness accounts for a disproportionately high healthcare burden [22]. Furthermore, patients with schizophrenia have an increased mortality and suicide [23] risk whilst also having a reduced quality of life [24].

The symptomatology of schizophrenia falls into three major categories. Positive symptoms include psychotic manifestations such as hallucinations, delusions and thought disorder, which are often the reason patients seek clinician's advice. However, the disorder is also associated with negative symptoms, such as amotivation and social withdrawal, and cognitive symptoms, which include deficits in working memory, executive function, and processing speed.

Schizophrenia has been described as a "disconnection syndrome" [25], referring to over and under-connection of neural circuits. In fact, brain-imaging studies suggest that schizophrenia is a disorder characterised by altered brain structure and function that may be related to clinical symptoms [2, 4, 6, 26, 27]. To date, no biomarker for schizophrenia diagnosis has been discovered and the condition is usually diagnosed after assessment by a specialist [28]. Consequently, one of the major goals in schizophrenia research is to identify underlying neural pathophysiological correlates of symptoms to yield a comprehensive understanding of this disease, thereby enabling the development of new treatment approaches, and establishment of personalised treatment planning. To this effect, neuroimaging tools and techniques may be particularly suitable schizophrenia-specific biomarkers, which may aid in developing accurate and reliable diagnostic platforms.

2.2 Functional Magnetic Resonance Imaging

Since its development, fMRI studies have become increasingly popular in neuroimaging research as they provide an unprecedented ability to safely and noninvasively image brain activity with very good spatial resolution. This neuroimaging method led to the expansion of a plethora of areas, such as clinical neuroimaging [1, 5, 6, 29] and cognitive neuroscience [29–35].

2.2.1 Basic principles

The technique of fMRI is most commonly performed using BOLD contrast imaging [36]. The increase in the magnitude of the BOLD signal provides an indirect measure of neural activity. This is triggered by a complex sequence of physiological events, which reflect changes in regional cerebral blood flow, volume and oxygenation. Briefly, exceeding the cerebral metabolic rate of oxygen leads to an increase in local neural activity - stimulating both higher energy consumption and increased blood flow. The local influx of oxygenated blood increases the oxy-to-deoxyhemoglobin ratio, which leads to an increase in the MRI signal compared to that of the surrounding tissue [31] (for an extensive discussion consult [37]).

Experimental designs used in fMRI studies may be either resting state or task-based. While task-based fMRI studies investigate brain response to various forms of stimuli and activity during task performance, resting-state fMRI (rs-fMRI) studies rely on the intrinsic spontaneous dynamics of the brain to detect and characterise activity within the brain [12, 30]. The goal of rs-fMRI is to measure the level of co-activation of the fMRI BOLD signals in different anatomical regions of the brain as an indicator of synchronous activity, i.e., interregional functional communication. Despite the challenges that rs-fMRI studies impose [12], studies of intrinsic brain dynamics and self-directed "resting" cognition, both in health and disease, are an indispensable tool to the neuroimaging community as they provide an important perspective on brain function.

2.2.2 Preprocessing and noise removal

In order to characterise or quantify brain activity using fMRI data, it is essential to differentiate the actual signal of interest from other noise-related fluctuations. Notably, one weakness of rs-fMRI lies in the complexity associated with noise/artefact identification and removal. The BOLD responses observed with rs-fMRI lack an *a priori* known task. As such, connectivity patterns are determined by measuring the temporal similarity of the BOLD time series using a given metric. Consequently, any non-neural activity-related process that is spatially or temporally correlated with the BOLD signals will also affect the derived measures of connectivity - introducing spurious results [12, 17]. These rs-fMRI confounds may be particularly problematic when comparing the brain's functional architecture between groups that display physiological or behavioural differences during "rest" [3, 16].

The analysis of rs-fMRI data has an intrinsically complex workflow, with quality control and preprocessing procedures playing a fundamental role in minimising the impact of artefacts and optimising data usability, reliability and reproducibility [12] (for a discussion on quality control and preprocessing consult [31]). Throughout rs-fMRI studies, spatial smoothing, temporal filtering and nuisance regression are commonly implemented preprocessing steps [12].

Spatial smoothing is a preprocessing technique which attempts to remove high-frequency signals. This technique involves the application of a filter to the data points, which replaces the signal at each voxel with a weighted average of that voxel's neighbours - resulting in blurred images [12, 31]. Whilst smoothing simultaneously increases the signal-to-noise ratio and the validity of the statistical tests while blurring any residual anatomical differences (decreases inter-subject variability), it also reduces the effective spatial resolution which may lead to the loss of meaningful local activations [12]. Smoothing typically involves convolving the functional images with a Gaussian kernel, often described by the full width of the kernel at half its maximum height - determining to which extent the data is smoothed [31].

Data from rs-fMRI studies contains noise, which may be observed at multiple frequency bands, depending on its source. Temporal filtering is performed in order to remove the effects of such noise confound signals with known or expected frequencies. This procedure hypothetically differentiates signal from noise frequency bands, resulting in increased signal-to-noise ratio [12, 38]. Research has shown that physiologically meaningful fMRI resting signals were predominantly "low-frequency" ($\sim<$ 0.1 Hz) spontaneous frequency fluctuations [16, 30, 33]. Furthermore, cardiac and respiratory physiological noise components typically reside in frequency bands above 0.1 Hz, where neural activity related signal is believed to be minimal [12, 16]. As a result, the standard strategy within investigations of rs-fMRI consists in applying a band-pass filter (typically, $0.01-0.08/0.1$ Hz). Nevertheless, it has been suggested that relevant neural activity may be detected at higher frequency ranges [39, 40].

Complementary to the aforementioned preprocessing procedures, a large body of research has been devoted to investigating cleanup methods to reduce the susceptibility of fMRI data to participant's physiological artefacts, such as head motion, cardiac and respiratory "noise", as well as vascular effects [17]. In fact, due to the indirect nature of fMRI, neural effects may be hard to distinguish from such sources of noise, thus hindering the interpretability of the results. In order to overcome the referred problem when conducting fMRI studies, a wide range of techniques have been proposed, such as time series-based strategies to remove nuisance regressors, that is, spurious fluctuations unlikely to be of neuronal origin [17, 41, 42], and the combination of fMRI with electrophysiological activity measurements [17], among others. Nevertheless, it is not currently possible to entirely isolate the true neural activity-related BOLD signals of interest from the confounds that may arise from many processes in the MRI environment [9, 10, 17]. Therefore, removing such confounds/noise will also inadvertently remove signals of interest [38]. Cleanup methods still require further refinements [16], which would facilitate more advanced applications to the healthy and diseased brains and, possibly, improve the reliability and interpretability of findings obtained from fMRI studies. However, despite the lack of consensus as to whether noise removal strategies should be applied and the fact that no "optimal" method has yet been identified, in recent years, a large number of studies have employed these techniques [41–43].

2.2.3 Brain parcellation

Brain parcellation is paramount for understanding brain organisation and function. It consists of defining distinct partitions within the brain, be they regions of interest (ROIs) or networks that comprise multiple discontinuous but closely interacting regions. Typically, BOLD time courses in each voxel in the brain are combined, according to a parcellation template, to extract the regional time courses for each of the areas or networks comprised in the selected parcellation scheme. To date, several brain parcellation templates have been proposed in the literature, with the optimal method yet to be discovered [44]. The Subset of 90 non-cerebellar Automated Anatomic Labeling brain areas (AAL90) parcellation divides the cortical and subcortical grey matter of the standard Montreal Neurological Institute (MNI) brain into 90 brain regions [45]. The procedure for parcellating fMRI data is summarised graphically in Figure 2.1.

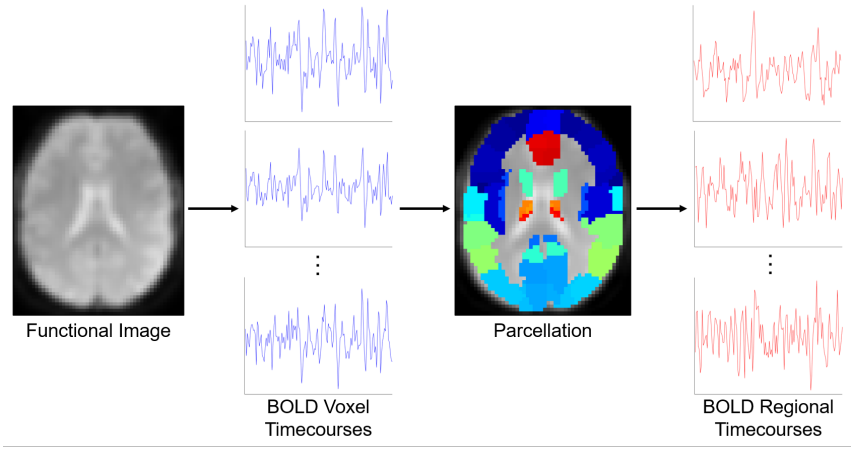


Figure 2.1: Axial slice of the brain obtained during "rest" at a given time point. BOLD time courses in each voxel of the brain are recorded and subsequently combined with a parcellation scheme (AAL90 was used) to recreate the (BOLD) regional time courses for each of the regions in the parcellation scheme.

2.3 Resting-state functional connectivity

The brain is characterised by a complex network of functional communications which enable the integration of information across different regions. The investigation of functional interactions in the human brain is of utmost importance, providing new valuable insights into the foundations of brain activity. In the context of functional neuroimaging, the characteristic temporal dependence of neural activity patterns of anatomically separated brain regions was termed Functional Connectivity (FC) [29]. FC reflects the level of functional communication between brain signals originating at different locations.

Early neuroimaging studies led to the recognition that spontaneous/intrinsic brain activity was a fundamental aspect of normal brain function. Resting-state FC, i.e., the spontaneous activity of the brain during "rest", was first described in the seminal discovery that estimated BOLD FC between the primary

motor cortex and other brain areas - revealing spatial FC patterns which mirrored those observed when executing a motor response [30]. This study showed brain regions could be synchronised in activity despite the absence of any task or stimulus [30]. Following this breakthrough, it became necessary to define a baseline - or default mode - of brain activity. This is because studies have shown that a given set of brain regions presented high activity levels during "rest" and low activity levels during task performance, and vice-versa - constituting a cohesive, task-negative default state of brain activity designated as Default Mode Network (DMN) [32]. The DMN has received special attention in investigations of cognitive misfunctioning in neurologic and psychiatric brain disorders [29, 46–48]. Subsequent neuroimaging studies of the brain at "rest" identified several other meaningful large-scale, task-positive, FC networks of correlated temporal patterns between anatomically separated brain regions. These FC networks were consistently observed across different subjects and corresponded to patterns of task-evoked activation. They define the so-called Resting-State Networks (RSNs) which constitute a discrete collection of anatomically separated, but functionally linked brain regions that display a high level of ongoing FC during "rest" [29, 33, 34]. Importantly, a wide range of studies have focused on examining possible functional dysconnectivity effects and/or altered activity within RSNs in neuropsychiatric disorders [4, 29, 47]. Figure 2.2 illustrates consistently reported RSNs across multiple studies as presented in [29].

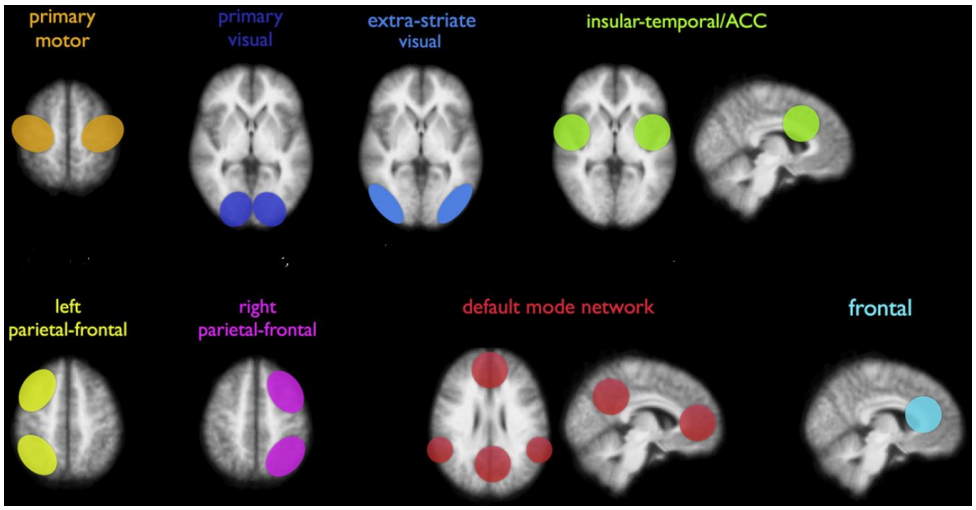


Figure 2.2: Consistent resting-state functional networks reported across different studies. The RSNs include the primary sensorimotor network, the primary visual and extra-striate visual network, a network consisiting of bilateral temporal/insular and anterior cingulate cortex regions, left and right lateralised networks consisting of superior parietal and superior frontal regions and the DMN consisting of precuneus, medial frontal, inferior parietal cortical regions and medial temporal lobe. Adapted from [29].

Motivated by prior investigations [30, 32–34], efforts were directed into the investigation of global functional organisation within the human brain, by mapping its functional connectivity. This resulted in reference FC maps which estimated the functional communications of the human cerebral cortex - providing information about the level of correlated dynamism across distinct brain regions [49]. Based

on rs-fMRI data collected from a large sample, the influential work of [49] provided a 7-network functional parcellation. The derived parcellation, presented in Figure 2.3, includes the Visual network, the Somatomotor network, the Dorsal and Ventral attention networks, the Limbic network, the Frontoparietal network and the Default network.

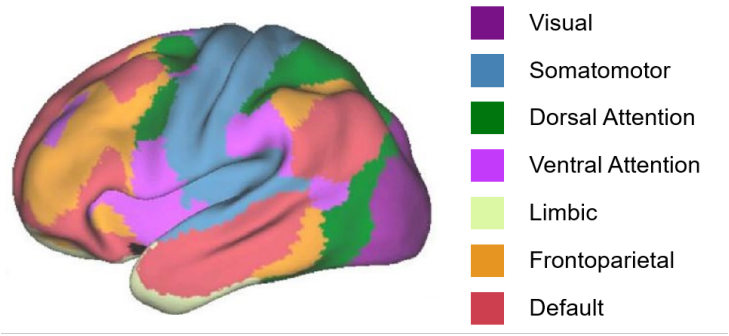


Figure 2.3: A coarse 7-network parcellation of the human cerebral cortex based on a large sample of rs-fMRI data. Adapted from [49].

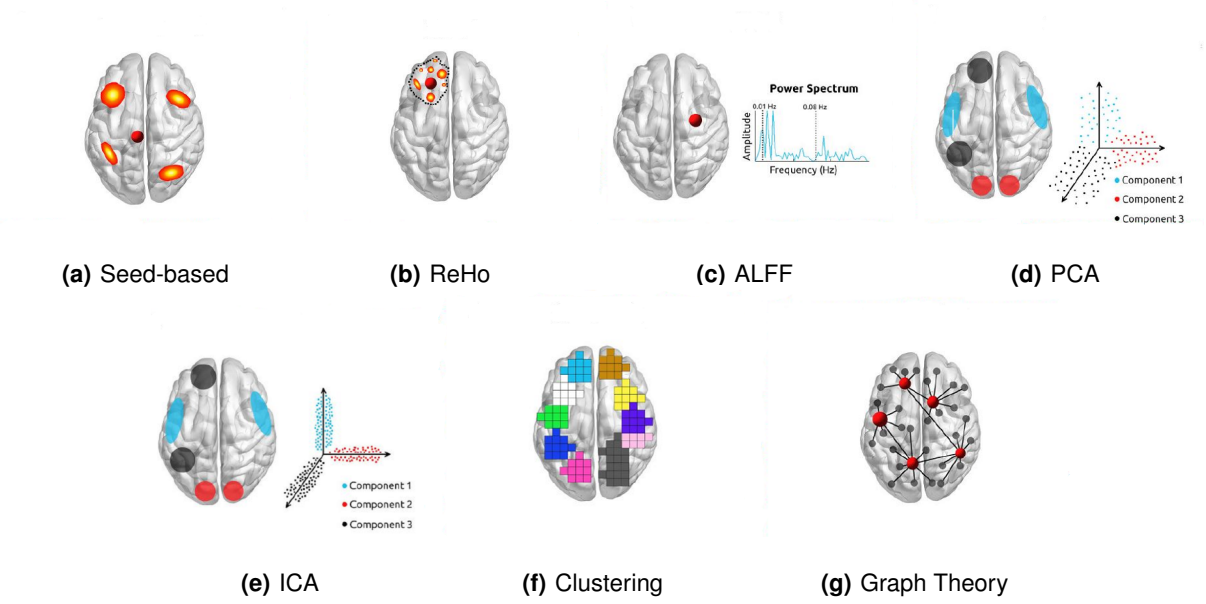
The authors suggested that these functionally coupled networks emphasise the need to adopt network approaches when exploring fluctuations in brain activity, since function within the brain is thought to emerge from the interactions among functionally interconnected brain areas [49]. Although research primarily focused on the role of the DMN in neurologic and psychiatric disorders, recent studies have started to examine whole-brain FC using a plethora of techniques - providing new insights in impaired brain communication and functional organisation in several disorders [4, 9, 10]. The following section moves on to describe commonly used analytical methods in rs-fMRI studies.

2.4 Resting-state functional connectivity analysis methods

Investigations of large-scale FC derived from rs-fMRI data have revealed valuable insights about the patterns of interregional neural interactions in both healthy and diseased brains [1, 4, 29, 46]. Such breakthroughs were obtained using a wide range of sophisticated analytic methods. Notably, the selection of the method to extract the relevant functional information plays a key role, since the method employed may greatly impact the results and, consequently, for instance, identify disease biomarkers [4, 50]. Traditionally, FC was assumed to be relatively static over the scan period (usually several minutes) and the majority of previous rs-fMRI studies applied static Functional Connectivity (sFC) approaches to assess brain functional organisation. Nonetheless, this assumption may be an oversimplification, as studies in humans demonstrated that measurements of BOLD signals during “rest” exhibited intrinsic spatiotemporal dynamic organisation [51]. Consequently, recent years have seen rapidly growing interest in studying time-resolved fluctuations in FC [13], often referred to as dynamic Functional Connectivity (dFC).

2.4.1 Static functional connectivity

Methods to extract relevant functional information from rs-fMRI data can be considered along a spectrum of temporal resolution. On one end, sFC methods assume that FC between regions is constant over an arbitrarily prolonged time window. These methods may focus on the local function of specific brain regions (functional segregation) or on the functional relationships between different brain areas (functional integration) [52, 53]. Functional segregation may be assessed utilising approaches such as Regional Homogeneity (ReHo) and Amplitude of Low Frequency Fluctuations (ALFF), while seed-based correlations, Principal Component Analysis (PCA), Independent Component Analysis (ICA), Clustering and Graph Theory may be used to evaluate functional integration [53]. Figure 2.4 presents the most common approaches for the analysis of rs-fMRI data.



(a) Seed-based **(b)** ReHo **(c)** ALFF **(d)** PCA

(e) ICA **(f)** Clustering **(g)** Graph Theory

Figure 2.4: Commonly used analysis methods in resting-state fMRI studies. Adapted from [12].

The first observation of resting-state FC from rs-fMRI data was accomplished using seed-based correlational analysis [30]. This technique is based on assessing the correlation among *a priori* defined ROIs or voxels, typically called seeds - resulting in FC maps [12, 50]. The FC maps provide information about with which regions the selected seed is functionally linked and to what extent [29].

Aiming to reflect regional FC (or co-activation), ReHo employs Kendall's coefficient of concordance to measure the similarity of the time series of a given voxel to those of its nearest neighbours (based on a pre-defined ROI) [54]. The ALFF [55] and its variant fractional ALFF [56], measure signal magnitude on a voxel-to-voxel basis and are usually implemented alongside ReHo, since these methods reflect different complementary properties of the BOLD signal [12].

The aforementioned model-driven strategies rely on prior knowledge to define sets of brain regions/voxels - limiting FC analysis to some specific regions/voxels. Such limitations were overcome by applying data-driven methods, such as PCA, ICA and clustering. PCA is concerned with explaining the variance-covariance structure of the data by finding a set of orthogonal axes (principal components) which reproduce the total or most of the data's variability [12]. ICA [57] is a widely used approach that has shown great promise in identifying network-based biomarkers of psychiatric disorders, such as schizophrenia [47, 58, 59]. Typically, ICA maps of fMRI studies are generated using spatial ICA which decomposes the BOLD signals into a linear combination of several maximally spatially independent components, of which some may be regarded as independent brain functional networks that represent synchronised BOLD activity [50, 53]. Clustering is another data-driven approach that groups brain voxels with similar connectivity (BOLD signal time courses) in the same cluster [12].

Functional brain networks have also been investigated through the lens of graph theory derived metrics, which have provided important insights to understand the architecture, development and evolution of brain networks [46, 52]. Under this framework, the brain is modelled as a network comprised of nodes (voxels or regions) and edges (connections between nodes, e.g., statistical dependencies among neuronal time series), which is commonly referred to as "Functional Connectome" [60]. This whole-brain functional network can be investigated at different levels of scale, and specific measures capture graph attributes at local (nodal) and global (network-wide) scales [46, 52]. These graph-based measures provide powerful features which integrate across the whole-brain - revealing useful biomarkers to characterise, identify and treat patients with psychiatric disorders such as schizophrenia [2, 5, 6, 27, 61].

The study of sFC, while undeniably informative, does not capture the spatiotemporal FC fluctuations in the human brain at behaviorally relevant timescales. Accordingly, on the other end of the spectrum, dFC-based methods, which are able to estimate time-resolved FC at each individual time point, were proposed. Remarkably, dFC-based approaches have been shown to provide more sensitive and specific markers of disease than sFC-based approaches [11, 50, 62]. Having introduced the general concepts of sFC, the following section will focus exclusively on the dynamical aspects of FC.

2.4.2 Dynamic functional connectivity

Recent functional neuroimaging studies have demonstrated that FC between different regions of the brain is not stable over time. In this perspective the term "Chronnectome" was employed to describe the time-varying, but reoccurring, patterns of functional communication among brain regions [11]. For instance, at "rest", the spatiotemporal variability of brain activity can be empirically observed by considering the spontaneous time-resolved activation and deactivation of a number of potential stable functional networks configurations, known as "FC states" or "FC patterns" [14, 21, 63–66]. Accordingly, dFC analysis provides valuable insights into the dynamic behaviour of the brain. Notably, a neural origin of mental

illnesses has been suggested by reports of altered resting-state brain activity - raising the possibility of utilising dFC approaches as disease biomarkers, by unveiling the intrinsic mechanisms leading to the diseased brain [9, 10, 14].

To date, several approaches have been proposed to assess and characterise the temporal variations in the spatiotemporal structure of BOLD signal fluctuations [9, 10, 67] and to test for its statistical significance [68]. However, the optimum analysis method is still uncertain and other unresolved issues still persist in this promising research area [13]. Moreover, despite its attractiveness to reflect time-varying fluctuations in FC, dFC approaches - like their stationary counterparts - are susceptible to noise and physiological artefacts [9, 10, 13].

The most common and straightforward approach to investigate the temporal evolution of FC is the Sliding-Window Analysis (SWA) [15, 65, 66]. This approach aims to capture pairwise connectivity from interregional synchrony [9, 10]. In this approach, a time window of fixed length (W) is selected, and within the temporal interval that it spans $(t - W, t + W)$, connectivity is computed between each pair of regional time courses using a given metric. The window is then shifted by a fixed step and the same calculations are repeated until the window spans the end part of the time courses. Figure 2.5 depicts an illustration of the computation of dFC using SWA.

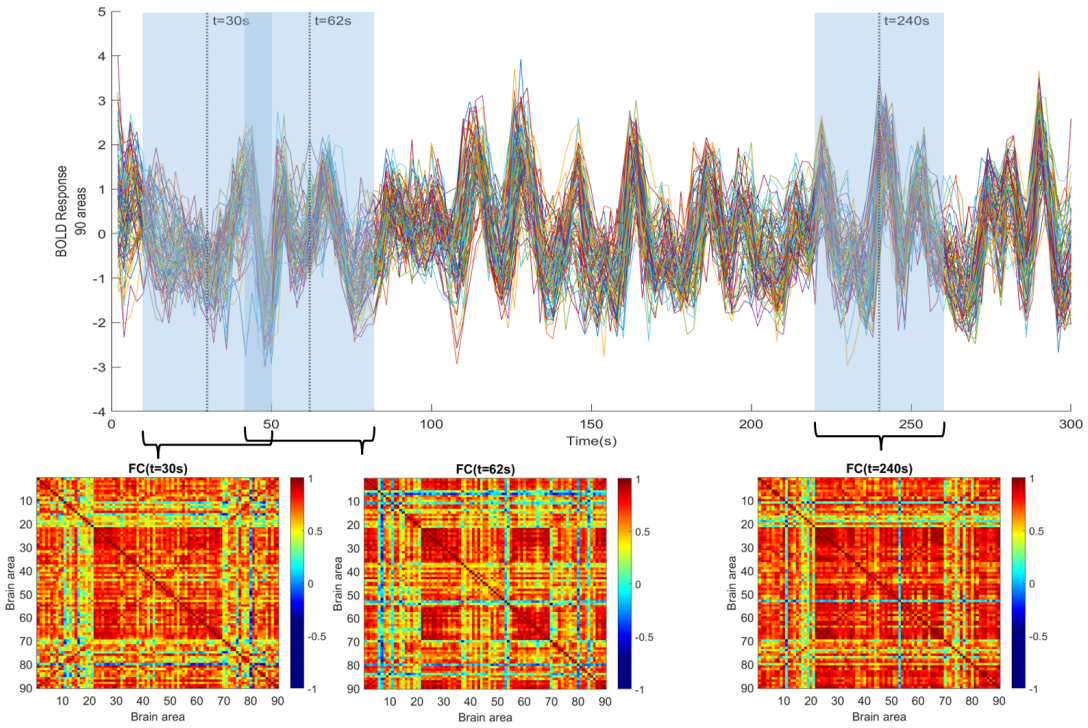


Figure 2.5: SWA using the Pearson correlation coefficient to extract dFC. BOLD signal measured at 90 distinct brain areas, according to the AAL90 parcellation scheme, from one healthy adult at "rest" (300 seconds, Repetition Time = 2 s). Example of correlation matrices $dFC(t)$ obtained for sliding-windows of 40 seconds centered at different time points t.

This iterative procedure outputs a three-dimensional time-varying tensor, with dimensions $N \times N \times T$, where N is the number of brain areas determined by the selected parcellation template and T the number of time windows considered, which describes the connectivity pattern of the brain during the examined temporal interval. The most frequently used metric in SWA is the Pearson correlation coefficient. Additionally, beyond this measure, higher order multivariate analyses have been experimented as well. In fact, by assessing FC in distinct time windows, sFC approaches can be easily extended to be time-resolved. Briefly, dFC can be evaluated by measuring FC among ROIs or voxels in a sliding-window, performing ICA (or Independent Vector Analysis (IVA)) on BOLD time courses in different windows, clustering time series of networks, evaluating time-resolved ReHo and computing time-varying graphs' metrics [10, 50]. Despite its popularity, the choice of a fixed window length limits the SWA to fluctuations in the frequency range below the window period, regardless of the true frequency content of the data [10]. Consequently, the choice of the ideal window length remains under debate [9, 10, 68].

Brain activity constitutes a dynamical regime in which multiple spatial patterns of correlated activity across different brain regions, i.e., RSNs, transiently form and dissolute [14, 63, 64, 66, 69]. Aiming to capture this spatiotemporal dynamics of the phase interactions among rs-fMRI BOLD signals at a quasi-instantaneous level, Phase Synchronization (PS) or Phase Coherence (PC) was proposed [70]. As stated, the SWA had limitations associated to the window size - compromising the time resolution/reliability of the dFC estimates. The adoption of data-driven PC-based approaches attempts to solve this problem by assessing the complex time-varying FC [70] (although at the cost of greater susceptibility to high-frequency noise fluctuations [67]).

Recently, taking advantage of PC, Phase Coherence Connectivity (PCC), described in detail below, was proposed as a method to capture the amount of interregional BOLD signal synchrony at each time point through the computation of a phase coherence matrix [21, 69, 71, 72]. Alternative PC-based measures have been proposed, such as the Coherence Coefficient and Coherency (see [73] for a review).

To compute the pairwise phase synchrony between brain region n and every other brain region at each time point, first the BOLD phases at each time t, $\theta(n,t)$, are estimated using the Hilbert transform of each BOLD regional time course. Let $x(n,t)$ be a real-valued signal representing the BOLD regional time course of brain region n. To separate the signal's instantaneous amplitude from the instantaneous phase [70] it must be converted into its analytical representation, $x_a(n,t)$ - complex signal whose imaginary part is the Hilbert transform of the real part [74]. Thus, the analytic signal is written as:

$$x_a(n,t) = x(n,t) + j\mathrm{H}[x(n,t)] \tag{2.1}$$

where j is the imaginary unit and $\mathrm{H}[\cdot]$ is the Hilbert transform. Henceforth, it is assumed that the BOLD regional time course constitutes a narrowband signal, which can be expressed as the product of an amplitude-modulated low-pass signal, $a(n,t)$, and a high-frequency sinusoidal "carrier", $\theta(n,t)$, as:

$$x(n,t) = a(n,t)\cos(\theta(n,t)) \tag{2.2}$$

The Bedrosian's theorem states that the Hilbert transform of the product of a low-pass signal with a high-pass signal with non-overlapping spectra is given by the product of the low-pass signal and the Hilbert transform of the high-pass signal [74]. This can be written in mathematical terms as:

$$\mathrm{H}[f_{LP}(t)f_{HP}(t)] = f_{LP}(t)\mathrm{H}[f_{HP}(t)] \tag{2.3}$$

where $\mathrm{H}[\cdot]$ is the Hilbert transform, f_{LP} is the low-pass signal and f_{HP} is the high-pass signal. Application of the Bedrosian's theorem to the Hilbert transform of the narrowband signal $x(n,t)$ yields:

$$\mathrm{H}[x(n,t)] = \mathrm{H}[a(n,t)\cos(\theta(n,t))] = a(n,t)\mathrm{H}[\cos(\theta(n,t))] = a(n,t)\sin(\theta(n,t)) \tag{2.4}$$

Plugging in the results from Equation (2.2) and Equation (2.4) in Equation (2.1) and applying the Euler's formula, the analytic signal becomes:

$$x_a(n,t) = a(n,t)\cos(\theta(n,t)) + ja(n,t)\sin(\theta(n,t)) = a(n,t)e^{j\theta(n,t)} \tag{2.5}$$

where $a(n,t)$ is the instantaneous amplitude and $\theta(n,t)$ is the instantaneous phase of the BOLD signal of brain region n at time t. Upon calculating the phases (angles) of the BOLD signals over time, the PCC between brain areas n and p at time t, $dFC(n,p,t)$, is estimated using Equation (2.6) [21,69,71,72]:

$$dFC(n,p,t) = \cos(\theta(n,t) - \theta(p,t)) \tag{2.6}$$

From Equation (2.6) it follows that, at a given time t, two brain areas n and p with temporarily aligned BOLD signals will have a phase synchrony value $dFC(n,p,t)$ close to 1. Contrarily, if, at a given time t, the phases of the brain areas n and p are orthogonal, i.e., they are at right angles (90°) to each other, the phase synchrony value $dFC(n,p,t)$ will be close to 0. Since the measure is undirected, for each time $t \in \{1,\ldots,T\}$, where T is the total number of time points, $dFC(t)$ is a $N \times N$ symmetric matrix, where N is the total number of brain regions. Therefore, for a given subject the resulting dFC, characterised through PCC, is expressed as a three-dimensional tensor with size $N \times N \times T$.

It must be noted that, despite allowing for a higher temporal resolution, the computation of meaningful dynamic FC through PC requires dealing with narrowband signals - imposing restrictions on the analysed signals [70]. In previous studies, this caveat has been overcome by band-pass filtering the BOLD regional time courses within a specified narrowband frequency range [69–72]. However, the BOLD frequency band which best captures functionally relevant information is still uncertain [38–40]. Furthermore, performing intergroup comparisons within a given frequency band implies that the fluctu-

ations within the groups' dynamical repertoire of FC patterns are identical, which may not be the case. Specifically, in many brain disorders, if the underlying anatomical connections are damaged, the observed dynamical collection of FC states is reduced comparatively to healthy brain FC [26, 61, 63]. This supports the claim that band-pass filtering the BOLD signals may affect the time-resolved FC estimates and the statistical validation of the results [71].

Beyond the aforementioned methods, other relevant dFC-based approaches have been employed in studies of rs-fMRI (consult [9, 10, 67] for a complete review).

2.4.3 Estimation of dynamic functional connectivity states

After the estimation of whole-brain dFC, using either SWA, PCC or another approach, a number of methods have been employed in the literature to extract interpretable information from the dFC patterns. These methods include assessing graph theoretical measures and/or determining dynamic FC states [10]. The latter is of particular relevance to the present study and is, therefore, subsequently described in further detail. This approach consists of defining a collection of patterns of connectivity that repetitively occur over time and across subjects during resting-state acquisitions [10, 67]. The decomposition of the resting-state FC into a limited set of connectivity states relies on matrix factorisation techniques, such as clustering algorithms [65, 66] (other techniques described in [10]). Different techniques can produce distinct results, with each technique having particular advantages and disadvantages [11].

Within the possible clustering algorithms, the unsupervised K-means algorithm is one of the most widely used methods [9, 10]. This algorithm partitions a set of multivariate data into K clusters in which each observation belongs to the cluster with the nearest centroid (for a complete description of the method consult [75]). Typically, to compare dFC patterns over time, by providing the upper triangular elements of the symmetric phase coherence matrices $(dFC(t))$ of all subjects, the algorithm outputs K cluster centroids representing the resting-state FC states that describe the transition dynamics within a rs-fMRI scanning session [67]. In other words, according to the definition of the K-means algorithm, the FC states can be thought of as average patterns of connectivity that subjects tend to return to during the course of a resting-state acquisition [64, 76]. This algorithm has been successfully applied in the literature to characterise the switching dynamics of FC [9, 10, 65, 66]. Recently, as an alternative method, LEiDA, which considers only the leading eigenvector, $V_1(t)$, of each $dFC(t)$ matrix, was proposed [21]. Under the framework of this method, the K-means algorithm is applied to all leading eigenvectors V_1 across time points and subjects - yielding time courses of recurrent FC states, i.e., V_1 at each time point t is assigned to the cluster (FC state) that best represents the pattern of connectivity specified by $V_1(t)$ [21]. This method is explained in greater detail in Section 3.3.

Having obtained the different FC states, these may be characterised in distinct aspects, namely their spatial configuration, probability of occurrence and mean duration, among others [67, 76]. Notably, these

measures are thought to provide novel disease biomarkers [9,21,66,77,78], and appear to be useful for prediction [11,62] and disease characterisation [79].

2.5 Permutation tests for the two-sample location problem

Traditionally, in the functional neuroimaging field, the classical parametric statistical framework was applied to investigate group differences in whole-brain FC [31,80]. Under this framework, distributional assumptions about the data are required to identify whether the observed effects are statistically significant. Nevertheless, a large body of neuroimaging studies have resorted to non-parametric approaches based on permutation test theory to statistically differentiate healthy controls from diseased patients [21,77,78]. Permutation-based statistical testing is a conceptually simple method that relies on minimal assumptions and deals with the multiple-comparison problem [31,80,81]. Furthermore, in comparison to the parametric approach, the permutation method does not lead to excessive loss of power and provides a valid method for analysis when the assumptions required for a parametric analysis are not satisfied [80].

Here, the general theory of permutation tests for the two-sample location problem is outlined (for a detailed discussion consult [80, 81]). Let $\mathbf{X_1} = (X_{11}, \ldots, X_{1n_1}) \sim P_1$ denote the random sample, i.e., independent and identically distributed (iid) random variables, of population 1 from distribution P_1 with location μ_1 and $\mathbf{x_1} = (x_{11}, \ldots, x_{1n_1})^\intercal$ designate a realisation of $(X_{11}, \ldots, X_{1n_1})$. Let $\mathbf{X_2} = (X_{21}, \ldots, X_{2n_2}) \sim P_2$ denote the random sample of population 2 from distribution P_2 with location μ_2 and $\mathbf{x_2} = (x_{21}, \ldots, x_{2n_2})^\intercal$ designate a realisation of $(X_{21}, \ldots, X_{2n_2})$. Let the pooled random sample be designated by $\mathbf{Y} = (X_{11}, \ldots, X_{1n_1}, X_{21}, \ldots, X_{2n_2}) = (Y_1, \ldots, Y_{n_1}, Y_{n_1+1}, \ldots, Y_n)$ with a pooled sample size of $n = n_1 + n_2$. The hypothesis test concerns inferences about the difference in location of the two random samples and can be mathematically formulated as [81]:

$$H_0 : \mu_1 = \mu_2 \quad vs. \quad H_1 : \mu_1 \neq \mu_2 \tag{2.7}$$

The central assumption of permutation tests is that of exchangeability, i.e., under the null hypothesis, all permutations of the realisations of $\mathbf{Y}$ are equally likely [80, 81]. Let T_0^* designate an appropriate test statistic for the two-sample location problem under H_0. Then, let $\{T_{0,b}^*\}_{b=1}^{n!}$ designate the set of all possible permutation test statistics under H_0, i.e., $T_{0,b}^*$ is the test statistic corresponding to permutation b of the pooled sample $\mathbf{Y}$. Under the null hypothesis, the set of the values of $T_{0,b}^*$, $t_{0,b}^*$, for all possible permutations $(\{t_{0,b}^*\}_{b=1}^{n!})$ constitutes the permutation distribution [81]. Let t_0^* denote the observed value of the test statistic T_0^*, i.e., this value is computed for the original samples. The statistical procedure of exact permutation tests consists in permuting the group labels $n!$ times, each time recalculating the test statistic. In theory, after completing the previous computations, the exact two-sided p-value is consistently estimated by the number of times that $|t_{0,b}^*| \geq |t_0^*|$, with $b \in \{1, \ldots, n!\}$, is observed divided by the

total number of computed permutations [80, 81]. However, calculating the test statistic for all possible permutations is generally impractical for large sample sizes. Therefore, the p-value is usually estimated by taking a random sample of size B from all possible permutations [81, 82]. This approach is designated conditional Monte Carlo permutation tests, with the term conditional emphasising that this method is an ordinary Monte Carlo simulation carried out on the permutation sample space [81, 82]. These tests are also termed approximate permutation tests, since the permutation distribution is approximated by a subsample, leading to approximate p-values [81]. Given t_0^* and a random sample of B permutations of the pooled random sample $\mathbf{Y}$, the approximate two-sided p-value is estimated as [81, 82]:

$$\widehat{L}_{T^*} = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}_{\left\{|t_{0,b}^*| \geq |t_0^*|\right\}} \tag{2.8}$$

where $\mathbb{1}$ denotes the indicator function with $\mathbb{1}_{\left\{|t_{0,b}^*| \geq |t_0^*|\right\}} = 1$ when $|t_{0,b}^*| \geq |t_0^*|$ and $\mathbb{1}_{\left\{|t_{0,b}^*| \geq |t_0^*|\right\}} = 0$ otherwise. Notably, Monte Carlo permutation tests have been shown to provide minimal loss of power in comparison to the full permutation tests [80, 82].

2.6 State Of The Art

Over past decades, a large body of research has focused on investigating neural activity in schizophrenia during "rest". Most rs-fMRI studies employed a sFC approach - providing a significant wealth of knowledge about brain function in schizophrenia [6, 27, 61]. Nevertheless, their utility in clinical practice is unclear as sFC features may not provide enough sensitivity and specificity [11, 62]. Furthermore, sFC findings in schizophrenia lacked consistency across studies [5]. As such, a growing number of studies have investigated dFC to understand schizophrenia [15, 62, 79, 83–85]. Despite these efforts, no consistent clinically relevant biomarkers, have been found that can be used as a reliable and valid diagnostic tool or as prognostic tools for schizophrenia [7, 8].

Within the neuroimaging community, it is well-established that resting-state FC exhibits dynamic fluctuations of connectivity patterns over time [13, 51] - giving rise to a discrete repertoire of transient FC states [63, 64, 66, 69, 71]. Consequently, in this MSc project, only rs-fMRI studies applying dFC-based approaches to determine FC states were reviewed and included in the State Of The Art. Table A.1 presents a brief description of the methods and findings of the reviewed studies.

2.6.1 Resting-state dynamic functional connectivity in schizophrenia

In one of the earliest reports implementing dFC analysis to study schizophrenia, the SWA approach was combined with clustering to determine FC patterns and assess group differences between patients with this condition and Healthy Controls (HCs) [79]. Results from this study suggested that Schizophrenia

(SZ) patients, in general, spent significantly more time in relatively more sparsely connected FC states. Furthermore, results demonstrated hyperconnectivity between thalamus and sensorimotor networks, as well as hypoconnectivity within the sensory regions. A different study, employing IVA to extract spatial components from windowed correlation matrices, found that regarding the SZ group, significant spatial changes in network connectivity mainly occurred between the frontoparietal, cerebellar and temporal lobe regions [86]. Moreover, in the mentioned work, evidence suggested that SZ patients displayed more connectivity dynamics than HCs.

Schizophrenia research has also focused on connectivity dynamics within specific brain networks. In a combined ICA-graph theoretical approach, dFC in SZ patients within the DMN domain showed that, compared to HCs, patients spent less time in highly connected states and exhibited reduced connectivity strength, clustering coefficient, and both global and local efficiency [87]. These results reflected impairment in DMN subsystems, specifically in the posterior cingulate cortex and anterior medial prefrontal cortex, as well as reduced connectivity between dorsal medial prefrontal cortex and medial temporal lobe. This suggested that the FC patterns within the DMN may be harmed by schizophrenic disease condition. A subsequent study examining dFC within the DMN found a drop in the connectivity of the anterior cingulate cortex and increases in the connectivity between the precuneus and the posterior cingulate cortex of SZ patients [48]. Additionally, a potential relationship was identified between schizophrenia symptom severity and the dysregulation of the dynamical properties of DMN FC [48]. Lastly, an investigation of dFC within the visual sensory network in HCs and patients with SZ indicated a substantial disruption within the visual system in a shorter timescale in SZ patients, namely in the middle temporal gyrus, cuneus, calcarine gyrus and fusiform gyrus [88].

One possible application of the exploration of the temporal FC dynamics within the brain is related to the development of future preventive and therapeutic methods. It has been shown that the common transient functional disconnectivity observed in SZ patients also extends to their unaffected siblings - implying a potential risk for the healthy siblings of developing schizophrenia [89]. In a different study, using a novel Group Information Guided ICA (GIG-ICA) framework to estimate connectivity states, it has been established that individuals with clinical high-risk for psychosis displayed a number of disease-related connectivity abnormalities observed in SZ patients; hence reflecting the vulnerability of this population to develop psychosis [90]. Furthermore, targeting perceptual and cognitive deficits in SZ has been suggested to enhance the understanding of the disease's pathophysiology, thus promoting new treatment interventions [91].

Another possible use of dFC-based approaches lies in clinical diagnostics. Evidence concretely suggested that changes in connectivity dynamics could explain underlying similarities and differences in SZ and Bipolar Disorder (BP) patients [83]. In fact, features extracted from dFC analysis have been shown to provide more sensitive and disease-specific markers than sFC features - revealing connectivity

differences between SZ and BP patients sharing overlapping symptoms [62, 92]. An additional study, comparing dFC in patients with SZ and Autism Spectrum Disorder (ASD), concluded that SZ patients were characterised by a severe and pervasive pattern of temporal aberrations which was not observed both in ASD and HCs. This reinforced the clinical utility of dFC-based approaches to differentiate disorders with coinciding neuropathology [93].

Resting-state FC studies reporting dysconnectivity in schizophrenia have been overwhelmingly obtained in medicated subjects. Therefore, to isolate medication effects from intrinsic illness characteristics in schizophrenia, a prospective study to evaluate dFC changes in unmedicated SZ patients submitted to treatment with an antipsychotic medication concluded that such medications had no significant impact on the observed FC [94]. This reinforced, to some extent, previous research on dFC in schizophrenia.

The efficacy of using fMRI and Magnetoencephalography (MEG) individually, or potentially combined, has also been explored and has been suggested to shape the understanding of schizophrenia and other disorders more accurately, since each neuroimaging technique depicts the FC of patient groups in distinct ways [85].

The aforementioned investigations assumed that at each time point the observed dFC could be represented by a single FC pattern, allowing connectivity dynamics to be described as a process of moving from one to another of these summary FC states. Thus, to characterise the dFC features as temporal summary measures, a so-called 'meta-state approach' was proposed [84]. Results from this investigation showed that patients with SZ displayed less overall temporal dynamism as they exhibited fewer changes in overall connectivity patterns compared with HCs.

Overall, these studies indicated that patients with schizophrenia exhibited reduced dynamic connectivity. On the one hand, SZ patients visited less often and spent less time in states characterised by strong, large-scale brain connectivity than HCs. Furthermore, when SZ patients transitioned into the state of strongest connectivity, they switched states very rapidly. On the other hand, SZ patients visited more often and spent longer periods of time in FC states characterised by weak connectivity, and showed less interaction between FC states. Anatomically, group differences between HCs and SZ patients encompassed connectivity among DMN, temporal, and frontal components. Therefore, the temporal dynamics analysis seemed to provide robust and sensitive disease-specific markers by capturing the underlying intrinsic differences between the FC of HCs and SZ patients. In general, most research on dFC in schizophrenia has been carried out by first decomposing the imaging data using group spatial ICA, resulting in time courses of networks which were then used to estimate dFC through SWA. The statistical procedures employed to assess intergroup differences spanned a range of parametric and non-parametric statistical tests. Notably, non-parametric permutation tests were observed to be underused across the reviewed literature.

3

Materials & Methods

Contents

This chapter outlines the materials and methods utilised in this study. Section 3.1 characterises the analysed neuroimaging data. The rationale leading to not applying spatial smoothing, temporal filtering and nuisance regression strategies to the data is presented in Section 3.2. The method used to capture the dynamic functional connectivity and the tools used to characterise the time-evolving brain activity are detailed in Section 3.3. Section 3.4 outlines the statistical procedure used to assess intergroup differences. The comparison of connectivity patterns with canonical resting-state networks is described in Section 3.5. This chapter ends by presenting the tools used to validate clustering solutions.

3.1 Neuroimaging data acquisition and preprocessing

Neuroimaging data was obtained from the publicly available repository COBRE preprocessed with NIAK 0.17 - lightweight release [20]. This data constituted a derivative from the COBRE sample found in the International Neuroimaging Data-sharing Initiative (INDI). The collected neuroimaging data included preprocessed rs-fMRI data from 72 patients diagnosed with schizophrenia (58 males, age 18$-$65 years old) and 74 healthy controls (51 males, age 18$-$65 years old) [20]. The rs-fMRI data featured 150 echo planar imaging BOLD volumes obtained in 5 minutes, with Repetition Time (TR) = 2 s, echo time = 29 ms, acquisition matrix = 64$\times$64 mm^2, flip angle = 75° and voxel size = 3$\times$3$\times$4 mm^3 [20].

Inspection of the BOLD data from each participant was performed using MATLAB R2019b [95]. This resulted in the exclusion of one participant whose data did not include all 150 BOLD volumes. Therefore, the final dataset used in the subsequent analysis included 71 patients with schizophrenia (80.28% males) and 74 healthy controls (68.92% males). Both groups had the same age range of 18-65 years old. The schizophrenia group had a mean age of 38.14 years old (standard deviation (std) of 13.99) and the healthy control group had a mean age of 35.82 years old (std of 11.58). The Framewise Displacement (FD), as defined in [96], provided a quantitative indication of each participant's head motion during the scanning period. The schizophrenia group had a mean FD of 0.4758 mm (std of 0.2768) and the healthy control group had a mean FD of 0.3177 mm (std of 0.1534).

To assess whether the mean age was different between groups, the two-sided Wilcoxon Rank-Sum test with Bonferroni correction provided in R software version 4.0.4 [97] was used, since the Shapiro-Wilk test suggested non-normality of the data. No significant difference was detected between the mean age of the groups by the two-sided Wilcoxon Rank-Sum test $(p = 0.4253)$. The same test was used to compare the groups' mean FD. The two-sided Wilcoxon Rank-Sum test detected a significant intergroup difference in the group mean FD $(p < 0.001)$. Specifically, on average, the BOLD signals of schizophrenia patients were characterised by larger amounts of head motion.

The data preprocessing as well as packaging was implemented by Pierre Bellec, CRIUGM, Department of Computer Science and Operations Research, University of Montreal, 2016 [20]. According to

the author, the datasets were analysed using the NeuroImaging Analysis Kit (NIAK) version 0.17, under CentOS version 6.3 with Octave version 4.0.2 and the Minc toolkit version 0.3.18. Each fMRI dataset was corrected for inter-slice difference in acquisition time and the parameters of a rigid-body motion were estimated for each time frame. Rigid-body motion was estimated within as well as between runs, using the median volume of the first run as a target. The median volume of one selected fMRI run for each subject was coregistered with a T1 individual scan using Minctracc, which was itself non-linearly transformed to the MNI template using the CIVET pipeline. The MNI symmetric template was generated from the ICBM152 sample of 152 young adults, after 40 iterations of non-linear coregistration. The rigid-body transform, fMRI-to-T1 transform and T1-to-stereotaxic transform were all combined, and the functional volumes were resampled in the MNI space at a 6 mm isotropic resolution.

In this study, the neuroimaging data was not subject to any further preprocessing procedures. According to the description provided by the author, no spatial smoothing and temporal filtering were implemented to the neuroimaging data [20]. In addition, along with the data, this repository provided a number of confounding variables which were estimated from the BOLD time series of each subject. Importantly, according to the author, no confound/noise removal techniques had been applied to the BOLD time series [20]. Therefore, the neuroimaging data contained the neural activity captured in the BOLD signals, as well as the participant's physiological artefacts.

3.2 Remarks on preprocessing

Typically, the preprocessing pipeline implemented in rs-fMRI studies includes spatial smoothing, temporal filtering and noise removal [12, 17]. In this project, contrary to the literature reviewed in Section 2.6, the aforementioned preprocessing steps were not applied to the neuroimaging data. This section presents a thorough discussion of the rationale leading to this "disruptive" decision.

3.2.1 Spatial smoothing

Previous studies have demonstrated that smoothing artificially shifts and combines activations from adjacent functionally and anatomically distinct brain regions - leading to less reliable detection of local activations [98]. Within investigations of whole-brain FC applying a ROI-based framework, spatial smoothing averages voxel signals both within and across the boundaries of ROIs - blurring the boundaries among ROIs and increasing the similarity among brain regions with respect to their time series. Nonetheless, the mentioned increase is limited in range by the kernel width, as neighbouring voxels will be significantly affected whereas the smoothing effect on distant voxels may be negligible or non-existent. Consequently, smoothing results in artefacts that depend on the parameters of the applied smoothing process - affecting differences in functional organisation between subject groups [99].

Despite its common use, spatially smoothing the rs-fMRI data was considered unsuitable given the objectives of this study. In this project, dFC is assessed by computing and for all participants, a phase coherence matrix at each time point. This matrix represents the phase synchronisation between any pair of ROIs, as described in Equation (2.6). Therefore, with the application of spatial smoothing, it was assumed that the phase relationship between neighbouring ROIs could be affected more significantly than the phase relationship between anatomically distant ROIs. Specifically, by implementing smoothing to the rs-fMRI data, the instantaneous phase of adjacent ROIs could become more similar - resulting in increased phase coherence values. Consequently, the detection of meaningful RSNs could be hindered as the captured FC patterns could become more segregated and less integrated. In other words, connectivity within local clusters (similar phase coherence) would be increased whereas inter-cluster connectivity would be weaker, thus "masking" the observed functional architecture of the brain in both the control and diseased groups. In fact, smoothing would raise the question as to whether the detected differences were due to actual variations in brain functional organisation or just a by-product of this preprocessing procedure. Finally, this study relies on parcellating the brain to investigate dFC. To some extent, this corresponds to a level of smoothing proportional to the size of the ROIs of the selected parcellation template.

3.2.2 Temporal filtering

As stated in Section 2.2.2, BOLD fMRI signal fluctuations used to study FC during "rest" are typically temporally filtered in a low-frequency band ($0.01-0.1$ Hz), predominantly to avoid the influence of physiological noises (respiration and cardiac signal). However, studies have provided evidence supporting the persistence of meaningful connectivity patterns at higher frequency bands - challenging the assumption that rs-fMRI BOLD signals are simple "low-frequency" spontaneous signal fluctuations [39, 40, 76]. Accordingly, this preprocessing procedure could result in increased BOLD signal variance being removed from the dataset at random [38]. Additionally, the implementation of narrow filters limits the BOLD signals to ultra-slow fluctuations. This has been shown to eliminate the unique signature of BOLD time courses [76] and encourage unreliable correlations among unrelated ROIs [16].

Here, no temporal filtering was applied to the resting-state BOLD signals. Consequently, the whole-frequency spectrum within the BOLD activity was considered, since it could conceal intrinsic connectivity patterns differences between schizophrenia patients and healthy controls. In addition, conducting comparisons between whole-brain FC in patients with schizophrenia and healthy controls within a specified frequency spectrum could have led to spurious results. Temporally filtering the rs-fMRI data implied that the frequency band of interest, where all relevant BOLD activity occurs, was known *a priori*. However, at the time of writing, it was unknown whether such a frequency band exists. Furthermore, more than one frequency band of interest within a given group and across participants could have been observed.

3.2.3 Nuisance regression

Estimates of dFC may be susceptible to confounds induced by hardware-related artefacts, subject movements, and other noise sources. Such artefacts can be mistaken for neural effects and are particularly problematic in studies of intergroup differences [16]. Additionally, it has been suggested that nuisance effects might account for a significant portion of the fluctuations in dFC estimates [9, 10]. Numerous cleanup methods have been proposed to isolate and remove these confounds - removing signals thought to be of non-neural origin from the BOLD time courses [16, 17, 42]. Typically, in rs-fMRI studies, nuisance regression entails regressing motion-related artefacts (consult [13, 17, 100] for complete discussions on the effects of motion-related confounds) and the global signal (see [13, 17, 101] for complete discussions on the effects of global signal regression). Nevertheless, understanding the influence of these cleanup strategies on rs-fMRI studies is still under active debate [102].

Motion regression was not performed in this study. It must be noted that this choice may result in biased estimates of dFC, since a substantial portion of the observed resting-state FC dynamics may be attributable to head motion [100]. A consequent body of evidence provided the scientific and methodological reasoning for not performing motion regression. Firstly, simple motion correction does not completely remove the influence of motion-related fluctuations from estimates of FC [43, 96, 102]. This issue could be addressed by applying "scrubbing" [96], i.e., censor complete volumes with excessive motion. However, this option was not viable given the data used in this project. On the one hand, if the data from all subjects after censoring was used for analysis, the large number of within-subject censored volumes could potentially interrupt the temporal autocorrelation structure of the data [42] - resulting in a considerable loss of signal variance of the data. On the other hand, following the recommendations proposed in [43], only 22 subjects had a number of uncensored volumes considered suitable for further analysis, thus resulting in extensive loss of data. Secondly, using excessive motion regressors was assumed to be detrimental to the extent that interesting BOLD signal could be removed from the rs-fMRI data [103]. Thirdly, the effects of motion regression in dFC estimates have been investigated mostly in studies employing the SWA approach [9, 10, 42]. Accordingly, the effects of motion regression on dFC estimates obtained using a PC-based approach [21], used in this project, are yet to be addressed. This approach may prove to be more robust to motion-related artefacts, as PC estimates the dFC at a quasi-instantaneous level by only considering the difference in phases between brain regions at each time point. Finally, evidence suggests that motion regression may extract BOLD signal variance that may be decomposed into functionally meaningful networks [103].

Here, the application of Global Signal Regression (GSR) to the resting-state BOLD signals was considered inappropriate. Without an accepted gold standard, the use of GSR has been a topic of controversy throughout rs-fMRI literature [101, 104]. In fact, although the global signal (mean BOLD signal computed across all voxels in the brain) reflects non-neural confounds in the data, it also likely includes

a neural component [101]. Furthermore, the detection of biologically meaningful RSNs could be dramatically altered by this procedure, potentially biasing group differences in estimates of FC. Specifically, in this project, the application of GSR was found to hinder the detection of a FC pattern corresponding to a global pattern of BOLD phase coherence. This was in line with previous studies employing a similar methodology to study dFC [21, 76–78]. Finally, the interpretation of FC results after GSR is complex and, therefore, it has been recommended that this rs-fMRI cleanup method should be avoided [17, 102].

In general, there is still little understanding of the effects of nuisance regression on dFC estimates. Research conclusions are inherently dichotomous as to whether this procedure should be employed or not in rs-fMRI studies [16, 103]. On the one hand, although it is not currently possible to remove all confounds, application of noise removal techniques may increase confidence in studies' results and interpretations [16, 17, 105]. On the other hand, there is great variability in the effectiveness of distinct cleanup methods, with no optimal method identified at the time of writing [17, 41–43]. Furthermore, evidence suggests that any set of nuisance regressors may extract BOLD signal variance that can be related to functional networks [17, 103] and, if unfiltered, may even reintroduce nuisance-related variation [105]. Additional studies have concluded that the difference between dFC estimates obtained prior and post nuisance regression were relatively small - questioning the usefulness of such strategies [102]. The modelling of confound effects in rs-fMRI studies is still an area of active investigation [17, 102] and its influence on dFC estimates obtained through PC remains largely unknown. Considering the reviewed studies, none of the confound regression strategies available in the literature were applied.

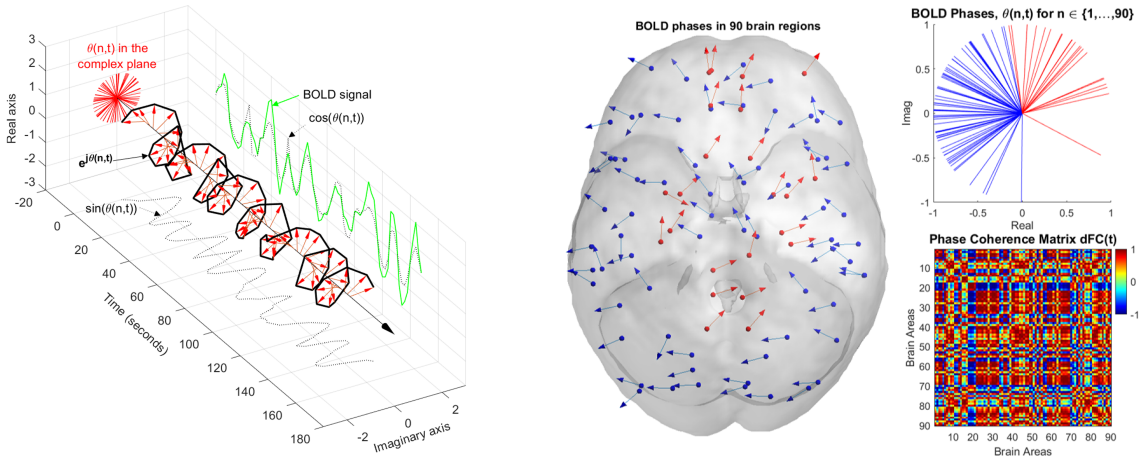
3.3 Dynamic functional connectivity analysis

In this project, rather than employing time-averaged approaches to investigate FC, the analysis presented in subsequent chapters leveraged on methods which enabled the estimation of time-resolved fluctuations in FC. This analysis used MATLAB R2019b [95], the Statistics and Machine Learning Toolbox™ and the Econometrics Toolbox™. The steps to investigate dFC are described below.

3.3.1 Computation of dynamic functional connectivity

While there is no gold standard for ROI selection, following the methodology in [21, 76–78], the canonical AAL90 template was used to parcellate the entire brain of each participant into 90 cortical and sub-cortical non-cerebellar regions [45]. This template provided a clear parcellation and an explicit description of whole-brain regions [92]. Accordingly, for each region in the brain, at each time point (TR), the BOLD signals were averaged over all voxels belonging to that brain area to compute the regional BOLD time courses. For each subject, this procedure resulted in a $N \times T$ BOLD dataset, where $N = 90$ is the number of brain areas and $T = 150$ is the total number of recording frames in each scan (volumes).

For each subject in the dataset, the PCC method [69–71], described in Section 2.4.2, was used to compute a phase coherence matrix with dimension 90×90 at each time point (TR). Each entry in this matrix represented the amount of BOLD signal synchrony between two brain areas defined by the AAL90 parcellation scheme at a given time point. To compute the phase coherence between each pair of AAL90 regions, first the instantaneous phase of the BOLD signals across all brain regions $n \in \{1, \ldots, N\}$ for each time $t \in \{2, \ldots, T-1\}$, $\theta(n,t)$, were estimated by computing the Hilbert transform of their BOLD regional time courses. It must be noted that throughout the entirety of this analysis the first and last TR of each fMRI scanning session were excluded due to possible signal distortions induced by the Hilbert transform following the procedure of [76]. Figure 3.1(a) illustrates the phase of the analytic representation of the BOLD signal for a given brain region, n, at each TR.



(a) Plot of the complex BOLD phase of one brain region n at each TR. It is depicted as $e^{j\theta(n,t)}$, with $sin(\theta(n,t))$ representing the imaginary part of the analytic phase and $cos(\theta(n,t))$ representing its real part (black dotted lines). The real part of the analytic phase, $cos(\theta(n,t))$, captures the oscillatory behaviour of the original BOLD signal (green). Red arrows represent the Hilbert phases at each TR.

(b) Representation of the BOLD phases in all $N = 90$ brain regions in both the cortical space and in the complex plane at a single time point t. Phase coherence matrix, $dFC(t)$, at the same time t, where 1 (red) indicates full coherence and -1 (blue) indicates areas are in anti-phase.

Figure 3.1: Summary of the method for computing a dFC matrix from rs-fMRI data.

Then, to obtain a whole-brain pattern of BOLD phase synchrony, the phase coherence between brain areas n and p at each time t, $dFC(n,p,t)$, was estimated using Equation (3.1):

$$dFC(n,p,t) = cos(\theta(n,t) - \theta(p,t)), \quad n,p \in \{1,\ldots,N\}, \quad t \in \{2,\ldots,T-1\} \tag{3.1}$$

where phase coherence values range between -1 (area n has a BOLD phase shift of 180° in relation to area p at time t) and 1 (areas n and p have synchronised BOLD signals at time t). This computation was repeated for all pairwise combinations of brain regions (n,p), with $n,p \in \{1,\ldots,90\}$, at each time

point (volume) t, with $t \in \{2, \ldots, 149\}$, and for all subjects. The resulting dFC for each subject was therefore a three-dimensional tensor with dimension $N \times N \times T'$, where $T' = 148$, i.e., 148 $dFC_{90\times90}(t)$ matrices were computed. Importantly, from Equation (3.1), it follows that phase coherence matrices are symmetric at each time point. Figure 3.1(b) depicts, at a single time point t, the BOLD phases in all 90 brain regions both in the cortical space and in the complex plane. The colouring of brain area n was determined by the sign of the n^{th} element of the leading eigenvector of the $dFC(t)$ matrix at time t (explained below). The $dFC(t)$ matrix displayed in Figure 3.1(b) captured the BOLD phase synchrony between each pair of brain areas at a single time point t (volume).

3.3.2 Functional connectivity leading eigenvector

Prior studies investigating dFC, typically analysed the symmetric $N \times N$ $dFC(t)$ matrices at each time point by considering the vectors found in the upper triangular elements of the matrices ($N(N-1)/2$ elements) [9,10,66,71]. This approach could result in the analysis of large amounts of multi-dimensional dFC data - hindering the estimation of recurrent patterns of connectivity. Alternatively, the method LEiDA, considering only the leading eigenvector, $V_1(t)$, of each $dFC(t)$ matrix, was proposed to capture the temporal evolution of dFC with reduced dimensionality [21]. This study uses the mentioned method to characterise the temporal evolution of the dFC in healthy controls and schizophrenia patients.

In detail, the leading eigenvector, $V_1(t)$, is a $N \times 1$ column vector that captures the dominant connectivity pattern of BOLD phase coherence at each time t, i.e., $V_1(t)$ represents the main orientation of the BOLD phases over all brain areas [21,76–78]. It must be noted that the dominant FC pattern of the dFC matrix at time point t can also be reconstructed back into matrix format by computing the $(N \times N)$ outer product $V_1(t)V_1^{\intercal}(t)$. Notably, considering only the leading eigenvector, $V_1(t)$, was found to considerably reduce the dimensionality of the classical approaches from $N(N-1)/2$ to N, while still explaining most of the variability observed in the spatiotemporal fluctuations of the connectivity patterns [21]. Additionally, LEiDA has shown increased robustness to high-frequency noise, improved temporal resolution and improved signal-to-noise ratio in the analysis of the time-evolving structure of resting-state fluctuations compared to the traditional approaches [21].

Under this framework, for each time t, the associated leading eigenvector, $V_1(t)$, partitions the $N = 90$ brain areas into two modules/communities by separating the positive elements of $V_1(t)$ from the negative elements of $V_1(t)$ [21, 106], as suggested by the projection of the BOLD phases into the cortical space of Figure 3.1(b). In other words, at time point t, a given brain area $n \in N$ is assigned to either one of the two modules according to the sign of the n^{th} row of $V_1(t)$. When all elements of $V_1(t)$ have the same sign, the BOLD phases between brain regions are coherent, which is indicative of a global mode of phase coherence governing all BOLD signals [21, 76–78]. This implies that all brain regions belong to the same community. Contrarily, if the elements of $V_1(t)$ have different signs (i.e., positive

and negative), the connectivity pattern between brain regions is not coherent. As a result, each ROI will be assigned to one of the two communities according to their BOLD phase relationship [21, 77, 78] (see cortical representation in Figure 3.1(b)). Additionally, the absolute value of each element in the eigenvector weighs the contribution of each brain area to the assigned community [21, 106].

Given that if $V_1(t)$ is a leading eigenvector, so is $-V_1(t)$, following the methodology in [76–78], it was ensured that most of the elements in $V_1(t)$ had negative values. This is because by assigning positive values to the brain areas whose BOLD phases did not follow the global mode, meaningful functional brain networks (i.e., canonical RSNs) were distinctly detected [76–78].

To illustrate the preceding considerations, Figure 3.2 presents the vector, matrix and cortical space representations of the leading eigenvector of the phase coherence matrix presented in Figure 3.1(b).

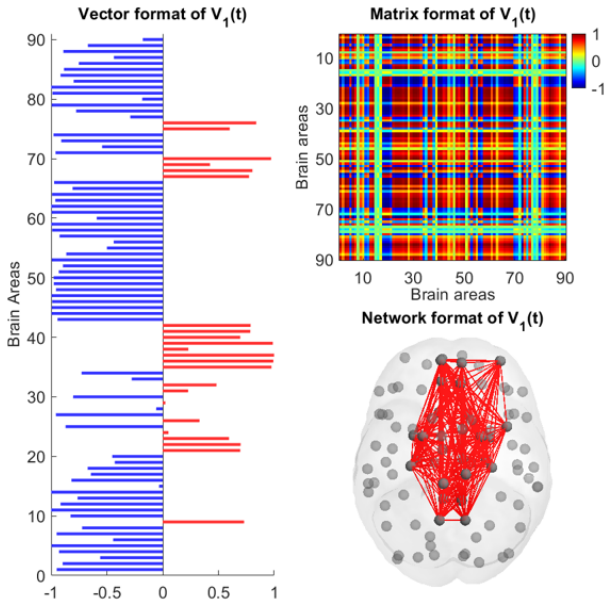


Figure 3.2: Vector, matrix and network representations of the leading eigenvector, $V_1(t)$, of the phase coherence matrix, $dFC(t)$, computed at time point t depicted in Figure 3.1(b).

To confirm the robustness of this approach, i.e., that reducing the dimensionality of the data did not lead to a considerable loss of information, the proportion of the variance in phase coherence explained by the leading eigenvector of each $dFC(t)$ at all time points was assessed across all subjects [21, 78]. As shown in Appendix A.2, the leading eigenvector consistently accounted for more than 50% of the variance in phase coherence at all time points and for all subjects - capturing most of the variability in the phase coherence matrices computed across the recording frames in each scan. Additionally, the two-sided Wilcoxon Rank-Sum test did not identify differences in the mean proportion of variance explained by the leading eigenvector at all time points between the control and the schizophrenia groups, as suggested in Figure A.1(b). Therefore, the health condition of each participant did not impair the ability of the leading eigenvector to represent the dominant connectivity pattern at a given time point.

3.3.3 Estimation of functional connectivity states

Upon computing the leading eigenvector, $V_1(t)$, of the phase coherence matrix, $dFC(t)$, for each recording frame (volume), the next step in the analysis was to characterise the evolution of the dFC over time by identifying recurrent FC patterns in the data, i.e., FC states that defined a set of repeating patterns of connectivity transiently emerging and dissolving through time [21].

In previous research using LEiDA, a number of FC states were detected by applying the K-means algorithm to the leading eigenvectors from all subjects [21,77,78]. These studies aimed to obtain the set of FC states that differed most significantly between groups, often overlooking the quality and validity of clustering results. Here, as stated in Section 1.3, the objectives were threefold. On the one hand, through the application of the statistical procedure explained in Section 3.4, this work aimed to detect statistically significant differences in the FC patterns observed in schizophrenia patients and healthy controls, independently of the number of FC states considered for analysis. On the other hand, this project intended to explore the quality and validity of the obtained clustering results, as well as the influence of using the K-medoids algorithm instead of the K-means algorithm to estimate FC states.

Firstly, to achieve this, the conventional LEiDA clustering analysis was performed [21,77,78]. Specifically, the K-means algorithm was applied to the dataset of all leading eigenvectors computed at the set of volumes $\{2, \ldots, 149\}$ across all 145 participants, which corresponded to a combined total of $148 \times 145 = 21460$ leading eigenvectors. The K-means clustering algorithm was run with a value of K from 2 to 20, i.e., dividing the set of leading eigenvectors into $K = \{2, 3, \ldots, 20\}$ clusters. The cosine distance was used as the distance metric for minimisation as in previous studies [21,76–78]. Aiming to minimise the chances of getting trapped in a local minima, the K-means algorithm was run 1000 times, considering in each iteration new initial cluster centroid positions selected according to the K-means++ algorithm [107].

The K-medoids clustering method searches for objects among the set of all observations which are representative of the clusters (medoids[1]) and assigns each object to the nearest medoid to build the clusters (for a review of the K-medoids algorithm consult [75]). The influence of using this algorithm to conduct a LEiDA analysis was investigated through its application to the combined total of 21460 leading eigenvectors from all participants. The algorithm was ran with a value of K from 2 to 20. Similarly, the cosine distance was selected as the distance metric for minimisation and the algorithm was run 1000 times, considering in each iteration new initial cluster medoid positions selected according to the K-means++ algorithm [107]. The algorithm to find medoids and other input arguments were set to the default parameter specifications.

[1] A medoid is defined as the object of a cluster whose average dissimilarity to all the objects of that cluster is minimal, that is, it is a most centrally located point in the cluster [75].

3.3.4 Characterisation of functional connectivity state temporal trajectories

Independently of the clustering algorithm, the LEiDA clustering procedure implemented in this study outputs one distinct clustering solution for each value of K clusters. In detail, each clustering solution contains K clusters $\mathbf{C} = \{C_1, \ldots, C_K\}$, with $K \in \{2, \ldots, 20\}$. This corresponds to decomposing the N-dimensional space of pooled leading eigenvectors into a K-dimensional state-based space. Let V_{C_α} be the vector of dimension $N \times 1$ representing the coordinates of the centroid/medoid of cluster C_α, where $\alpha \in \{1, \ldots, K\}$. Then, for a given clustering solution, each V_{C_α} represents a recurrent FC pattern or FC state, i.e., a pattern of BOLD phase coherence [21]. These patterns of connectivity are commonly characterised by the activation of a small subset of ROIs which display coherent BOLD signal - forming functional brain networks. It must be noted that functional networks were assumed to be represented by the subset of ROIs with positive sign in V_{C_α}, since these defined a group of brain areas detaching from the global mode of BOLD phases. The contribution of brain area $n \in \{1, \ldots, 90\}$ from the AAL90 parcellation to a given functional network is given by the n^{th} row of V_{C_α} [21,106]. Importantly, throughout this work, the FC state represented by the centroid/medoid V_{C_α} was referred to as simply FC state α. Furthermore, each estimated FC state α can be represented back into a $N \times N$ matrix by computing the outer product $V_{C_\alpha} V_{C_\alpha}^{\intercal}$, with positive values between all elements with the same sign in V_{C_α} and negative values between elements of different signs.

For each clustering solution, the set of estimated K FC states was subsequently utilised to obtain, for each participant, time courses of FC states. This was accomplished by representing each $V_1(t)$ at time point t by the FC state (centroid/medoid) of the cluster to which it was assigned by the clustering algorithm (this corresponded to the output vector of cluster assignments). Specifically, following the conceptual framework proposed by [76], resting-state BOLD time series were assumed to temporally evolve through a finite state trajectory of recurrent patterns of BOLD phase coherence. Following this rationale, each clustering solution with K FC states was assumed to define a finite state space $S = \{1, \ldots, K\}$. Furthermore, for a clustering solution with K clusters, the cluster (FC state) to which $V_1(t)$ was assigned at each time t, denoted by V_t, was assumed to define a stochastic process, $\{V_t : t \in \{2, \ldots, 149\}\}$, with associated state space as defined above. Consequently, considering the Markov property [108] holds, the temporal trajectory of FC states was assumed to define a time-homogeneous Discrete Time Markov Chain (DTMC) [76]. Importantly, it must be noted that, although brain activity is an uninterrupted process, the limited fMRI scanning period implied the state trajectories were temporally limited. This resulted in a number of DTMCs not spanning the entire state space, i.e., under this framework, despite assuming all FC states within the state space existed, some FC states were not observed. Moreover, the time courses of FC states were defined in a discrete sense, which could have been an oversimplification.

Figure 3.3 graphically summarises the estimation and characterisation of the temporal trajectories of FC states for a representative clustering solution considering $K = 7$ clusters.

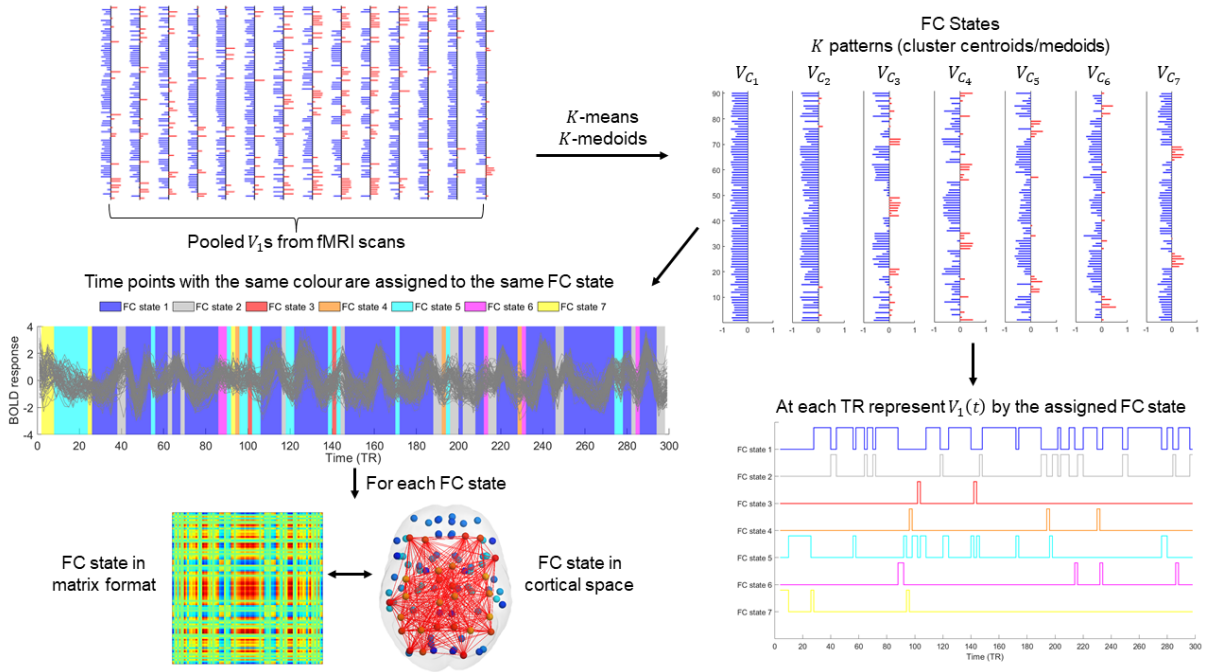


Figure 3.3: Graphical illustration of the estimation and characterisation of the temporal trajectories of recurrent FC states obtained by using LEiDA. The pooled leading eigenvectors are clustered into 7 FC states, which are used to define a time course of FC states for a given participant.

A number of descriptive properties (measures) of the underlying temporal trajectories of FC states observed in healthy controls and schizophrenia patients were considered in this study. Specifically, the fractional occupancy and dwell time were estimated for each individual FC state. The one-step Transition Probability Matrix (TPM) which represents the one-step probability of switching between FC states was also computed. Furthermore, the limiting behaviour of the time courses of FC states was also investigated by estimating the limiting distribution of the corresponding DTMCs. These measures provided insights into the stability of BOLD phase coherence states and facilitated the characterisation of the dynamical properties of the connectivity patterns emerging spontaneously and consistently during "rest" [76]. Notably, these properties were shown to provide relevant insights on aberrant dynamic brain network connectivity in previous LEiDA analysis [21, 77, 78].

3.3.4.A Fractional occupancy

The fractional occupancy (probability of occurrence) of a FC state α represents the proportion of times V_t was assigned to cluster C_α throughout the duration of a scan [76]. The fractional occupancy of FC state α for the scanning session of participant s, $P_\alpha^{(s)}$, is calculated (estimation) as follows:

$$P_\alpha^{(s)} = \frac{1}{T'} \sum_{t=1}^{T'} \mathbb{1}_{\left\{V_t^{(s)} \in C_\alpha\right\}} \quad , \quad \alpha \in \{1, \ldots, K\} \tag{3.2}$$

where $T' = 148$ is the number of time points (first and last volume of each scan were excluded), $\mathbb{1}$ is the indicator function and $V_t^{(s)}$ is the FC state to which $V_1(t)$ was assigned at time t. For each clustering solution with $K = \{2, \ldots, 20\}$ clusters, the fractional occupancy was estimated for each of the K FC states separately for each fMRI scan. Importantly, the restricted fMRI scanning time window and low temporal resolution resulted in FC states having a null probability of occurrence throughout the analysis.

3.3.4.B Dwell time

The dwell time (mean duration) of a FC state represents the mean number of consecutive epochs spent in that state throughout the duration of a scan [76]. The dwell time of FC state α measured for participant s, $DT_\alpha^{(s)}$, is defined (estimation) as:

$$DT_\alpha^{(s)} = \frac{1}{k_\alpha^{(s)}} \sum_{d_\alpha=1}^{k_\alpha^{(s)}} R_{d_\alpha} \quad , \quad \alpha \in \{1, \ldots, K\} \tag{3.3}$$

where $k_\alpha^{(s)}$ is the number of consecutive periods in which $V_t^{(s)}$ was assigned to FC state α (cluster C_α) and R_{d_α} is the duration of each of the $k_\alpha^{(s)}$ periods. For each clustering solution with $K = \{2, \ldots, 20\}$ clusters, the dwell time was estimated for each of the K FC states separately for each fMRI scan. It must be pointed out that unobserved FC states (null fractional occupancy) had a mean duration of zero.

3.3.4.C One-step transition probability matrix

According to [76], the probability of being in FC state α at time t and transition to FC state β at time $t+1$, considering a clustering solution with state space $S = \{1, \ldots, K\}$ (K clusters), can be calculated (estimated) by the following expression:

$$P_{\alpha\beta}^{(s)} = \frac{1}{T'-1} \sum_{t=1}^{T'-1} \mathbb{1}_{\left\{V_t^{(s)} \in C_\alpha \, , \, V_{t+1}^{(s)} \in C_\beta\right\}} \quad , \quad \alpha, \beta \in \{1, \ldots, K\} \tag{3.4}$$

From Equation (3.4), for a clustering solution with K FC states, it follows that the TPM of the fMRI scan of participant s, $\mathbf{P}^{(s)}$, is defined (estimation) as:

$$\mathbf{P}^{(s)} = \left[P_{\alpha\beta}^{(s)}\right]_{\alpha,\beta \in \{1,\ldots,K\}} = P\left[V_{t+1}^{(s)} \in C_\beta \mid V_t^{(s)} \in C_\alpha\right] = \frac{P_{\alpha\beta}^{(s)}}{P_\alpha^{(s)}} \quad , \quad \alpha, \beta \in \{1, \ldots, K\} \tag{3.5}$$

For the tentative optimal clustering solution, a TPM was estimated separately for each DTMC representing a fMRI scan. Furthermore, for the DTMCs defined by the optimal clustering solution, the average TPM for the control and schizophrenia groups was obtained by averaging the TPMs of all participants from each group.

It must be noted that for any FC state with an associated null probability of occurrence, the probability of transitioning from that state to any other state is zero according to Equation (3.4). Consequently, according to Equation (3.5), it follows that for any unobserved FC state j the row j of the TPM will have K undetermined elements ("not a number" (NaN)), i.e., the one-step transition probabilities given by $\left[P_{j\beta}^{(s)}\right]_{\beta\in\{1,\ldots,K\}}$ will be undefined since $\frac{P_{j\beta}^{(s)}}{P_{j}^{(s)}} = \frac{0}{0}$ for all $\beta \in \{1,\ldots,K\}$. This implied that not all estimated TPM were right stochastic matrices [108]. Importantly, in this study, since unobserved FC states were considered to occur at some point within the state trajectories of brain activity, all estimated TPMs were considered for subsequent computations and analysis, with the exception of all NaN values. Furthermore, the NaN values were not replaced by zeros, since this was observed to modify the distribution of one-step transition probabilities - biasing group comparisons of these quantities.

3.3.4.D Limiting probability

The stationary distribution of the DTMCs was calculated to study the limiting probabilities of the FC states defining the temporal state trajectories. The estimation of this property of the DTMCs aimed to bypass the restrictions imposed by the temporal windows of the fMRI scanning sessions.

To compute this quantity, all considered DTMCs were irreducible [108], with finite state space $S = \{1,\ldots,K\}$, with $K = \{2,\ldots,20\}$. Hence, all states were positive recurrent [108]. Furthermore, all considered DTMCs were aperiodic [108]. Therefore, for every participant, s, with an associated irreducible, positive recurrent and aperiodic DTMC, it followed that:

$$\lim_{t\to\infty}\left(P_{\alpha\beta}^{(s)}\right)^{t} = \pi_{\beta}^{(s)} > 0 \quad , \quad \alpha,\beta \in \{1,\ldots,K\} \tag{3.6}$$

where the row vector $\boldsymbol{\pi}^{(s)} = \left[\pi_{\beta}^{(s)}\right]_{\beta\in S}$, with dimension $1\times|S|$, was the unique stationary distribution of the DTMC of participant s (see [108] for theorem) given (estimation) by:

$$\boldsymbol{\pi}^{(s)} = \mathbf{1}\times\left(\mathbf{I}-\mathbf{P}^{(s)}+\mathbf{ONE}\right)^{-1} \tag{3.7}$$

where $\mathbf{1}$ is a $1\times|S|$ vector of ones, $\mathbf{I}$ is the identity matrix with rank $|S|$, $\mathbf{P}^{(s)}$ is the TPM of subject s and $\mathbf{ONE}$ is a $|S|\times|S|$ matrix all of whose entries are one. For a given FC state β, π_{β} (element β of the row vector $\boldsymbol{\pi}$) was the measure to be used to perform intergroup comparisons. Importantly, since only aperiodic DTMCs were considered, π_{β} can be understood as the limiting probability that the DTMC is in FC state β and as the long-run fraction of time the DTMC spends in FC state β. It must be noted that intergroup comparisons between the estimated stationary distributions were not performed in this study.

For the tentative optimal clustering solution, the limiting probability of each FC state was estimated separately for each fMRI scan. Due to the inclusion criteria imposed on the DTMCs defined by the tentative optimal state trajectories, i.e., irreducibility and aperiodicity, only 37 and 46 DTMCs from the

control and schizophrenia groups, respectively, were considered for further analysis. This corresponded to excluding all DTMCs with an associated TPM with at least one row of zeros, i.e., with at least one unobserved FC state.

3.4 Intergroup comparisons

In this work, intergroup comparisons between estimates of the properties of the temporal trajectories of FC states were performed through the application of Monte Carlo permutation tests by adapting the statistical procedure used in previous LEiDA analysis [21, 77, 78]. Permutation tests were selected over parametric approaches as they did not impose stringent assumptions, did not lead to excessive loss of inferential efficiency and have been extensively applied within the neuroimaging community [80, 81, 109].

Here, Monte Carlo permutation tests were used to assess intergroup differences because the dataset considered for analysis contained 145 participants - leading to an extremely large number of possible permutations of the group labels. Furthermore, the appropriateness of the application of Monte Carlo permutation tests in the context of this analysis relied in the assumption of exchangeability of the estimated data. In fact, the assumption of independent subjects implied exchangeability, since any intergroup differences were assumed to relate to variations in the time-resolved FC, irrespective of the effect of other variables. Under the null hypothesis that there were no actual differences between healthy controls and schizophrenia patients, the group membership became arbitrary. Consequently, due to the large sample size and given that the assumption of exchangeability was satisfied, the suitability of Monte Carlo permutation tests was guaranteed.

The general framework for the hypothesis tests associated with the properties presented in Sections 3.3.4.A to 3.3.4.D is subsequently described. Each of the aforementioned measures was designated as measure m to improve readability. For each measure, consider the random sample $\mathbf{X}_1^m = (X_{11}^m, \ldots, X_{1n_{\text{HC}}}^m)$ of size n_{HC} from a distribution with mean μ_{HC}^m from the control group, where X_{1i}^m is the value of measure m for the i^{th} chosen healthy control. Similarly, for each measure, consider the random sample $\mathbf{X}_2^m = (X_{21}^m, \ldots, X_{2n_{\text{SZ}}}^m)$ of size n_{SZ} from a distribution with mean μ_{SZ}^m from the schizophrenia group, where X_{2i}^m is the value of measure m for the i^{th} chosen schizophrenia patient. Here, both random samples were statistically independent. For any measure m, the null hypothesis (H_0) to be tested states there is no difference between the mean of measure m in the control group and the schizophrenia group, whereas the alternative hypothesis (H_1) states there is a difference between the mean of measure m between groups. Mathematically, this can be written as:

$$H_0 : \mu_{\text{HC}}^m = \mu_{\text{SZ}}^m \quad vs. \quad H_1 : \mu_{\text{HC}}^m \neq \mu_{\text{SZ}}^m \tag{3.8}$$

For each clustering solution, the realisations of the previously defined random samples were given

by the estimates of measure m computed from the temporal trajectories of FC states of the control and schizophrenia groups, i.e., $\mathbf{x}_1^m = (x_{11}^m, \ldots, x_{1n_{\text{HC}}}^m)^\intercal$ and $\mathbf{x}_2^m = (x_{21}^m, \ldots, x_{2n_{\text{SZ}}}^m)^\intercal$, respectively.

Given the high computational power available, Monte Carlo permutation tests were performed using $B = 100000$ permutations to produce an accurate approximate estimation of the permutation distribution. However, studies have suggested that 10000 permutations suffice for permutation testing at a 1% significance level [82]. The remainder of this section describes the statistical procedure used to perform intergroup comparisons for the estimates of the measures described in Sections 3.3.4.A to 3.3.4.D. It should be noted that the same procedure was applied to all clustering solutions detected by the K-means and K-medoids algorithms.

With the exception of dwell time, the remaining measures were probabilities. This implied that the assumptions required to perform a two-sample t-test or a two-sample Welch's t-test on the estimates of the fractional occupancy, transition probabilities and limiting probabilities were not satisfied. In other words, the distribution of the test statistic could not be assumed to follow a t-distribution. This prompted the use of Monte Carlo permutation tests to approximately generate the distribution of the mentioned test statistic. It must be noted that despite not representing a probability, intergroup comparisons of the estimates of the dwell time were also conducted using Monte Carlo permutation tests. The Levene's test was used to assess the homogeneity of the variances of the populations from which the group-samples were drawn. The threshold to reject the hypothesis of homogeneity of variance was fixed at a 5% significance level, i.e., the hypothesis of homocedasticity was rejected if $p_{\text{Levene's test}} < 0.05$ and not rejected otherwise. The results from the Levene's tests were then utilised to determine the statistic to be employed on the Monte Carlo permutation tests. The two statistics used throughout the Monte Carlo permutation tests (under H_0) were given by Equation (3.9):

$$T_0^* = \begin{cases} \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_{\text{HC}}} + \frac{S_2^2}{n_{\text{SZ}}}}}, & p_{\text{Levene's test}} < 0.05 \\ \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_{\text{HC}}-1)S_1^2 + (n_{\text{SZ}}-1)S_2^2}{n_{\text{HC}}+n_{\text{SZ}}-2} \cdot \left(\frac{1}{n_{\text{HC}}} + \frac{1}{n_{\text{SZ}}}\right)}}, & \text{otherwise} \end{cases} \tag{3.9}$$

where $\bar{X}_1$ and $\bar{X}_2$ are the random sample means, S_1 and S_2 are the random sample standard deviations and n_{HC} and n_{SZ} are the sample sizes for the control and schizophrenia groups, respectively. Only one statistic from Equation (3.9) was used depending on the relation between the variances of the group data samples. The statistic from Equation (3.9) used to perform the statistical test was subsequently used to obtain the value of the statistic under each of the B permutations of the sample data. Once the statistic was selected, for each permutation, the group labels were randomly permuted and $t_{0,b}^*$, with $b \in \{1, \ldots, B\}$, was obtained by plugging in the values of the permuted sample data. Under the null hypothesis, the set of values $\{t_{0,b}^*\}_{b=1}^B$ was used to estimate the permutation distribution. This produced

an approximation of the distribution of the statistic (T_0^*) employed throughout the statistical permutations of the sample data. After calculating the B permutation statistics, the permutation distribution was estimated by applying a normal kernel smoothing function to the set of permutation statistics ($\{t_{0,b}^*\}_{b=1}^{B}$). This resulted in an empirical Cumulative Distribution Function (CDF), which was observed to converge towards the CDF of a normal distribution, as expected from the Central Limit Theorem (CLT) [110]. To obtain an approximate p-value, linear interpolation was then used to calculate the value of the empirical CDF at the point $|t_0^*|$. Finally, the two-sided p-value was estimated by plugging in the value of the empirical CDF evaluated at the point $|t_0^*|$ into the expression $p = 2 \cdot (1 - \text{CDF}(|t_0^*|))$.

It must be noted that the statistical procedure employed in this work to assess intergroup differences differed from the one applied in prior studies employing LEiDA [21,77,78]. Specifically, a different statistic was used according to the homocedasticity of the populations from which the samples were drawn.

Importantly, the variability observed for each measure was assumed to be independent of the effect of other variables, such as age, gender and head motion. Here, the effect of these variables on the estimates of the measures considered for analysis was not investigated. Section 6.2 discusses this limitation and provides recommendations to study the mentioned effects in future research.

3.5 Comparison to reference resting-state functional networks

As explained, for each of the considered algorithms used to estimate FC states, the data was partitioned into K clusters, with $K \in \{2, \ldots, 20\}$. For each clustering solution, the functional relevance of the estimated FC states was investigated by assessing whether there was a significant spatial overlap between the cluster centroids/medoids which represented each FC state and any of the seven reference functional networks proposed by [49] (cortical representation of the seven networks depicted in Figure 2.3). This was accomplished by following the procedure used by [76, 78]. Briefly, the seven RSNs defined by [49] were transformed into seven non-overlapping vectors with dimension 1×90, where each entry of the vectors corresponded to each of the regions defined by the AAL90 parcellation template. For each of the seven functional networks, the n^{th} element of its vector was obtained by calculating the number of voxels from brain region n of the AAL90 parcellation that were assigned to that particular network [76].

Finally, the Pearson correlation coefficient was used to assess the spatial overlap between these seven reference RSNs and the centroids/medoids V_{C_α}, with $\alpha \in \{1, \ldots, K\}$, corresponding to the FC states estimated for each clustering solution with K clusters. As in previous work, all the negative values of the centroids'/medoids' vectors were set to zero, to consider only areas that were thought to define relevant functional networks [76, 78]. The associated p-values were evaluated to determine whether the spatial overlap was significant. Specifically, for each clustering solution with K FC states, the overlap between a given FC state and any of the seven reference RSNs was considered significant if $p < 0.05/K$.

3.6 Validation of clustering solutions

In this study, cluster validation techniques were used to assess the quality of the results produced by the clustering algorithms, i.e., to determine whether the generated partitions of the data captured the actual group-structure of the data. Here, internal and external clustering validation techniques, as well as a method to evaluate clustering stability were utilised.

3.6.1 Unsupervised cluster validation criteria

In this study, the class labels of the leading eigenvectors obtained for each TR were unknown. Consequently, given the assignment of leading eigenvectors to a given cluster formed an unsupervised classification task, the quality of clustering structures was evaluated using internal validation measures. Namely, the average Silhouette coefficient [111] and the Dunn's index [112] were measured for each clustering solution outputted by the clustering algorithms, with the highest measured value determining the optimal number of clusters. Despite their broad applicability, the performance of these measures has been shown to be affected by various characteristics of the data [75, 113]. Therefore, internal validation criteria were only used as an accessory tool to select an optimal clustering solution.

3.6.2 External validation using clustering agreement measures

Here, clustering agreement measures based on paired agreement and mutual information were utilised to compare the clustering outputs from distinct clustering algorithms. These indices can be written based on a $k \times r$ contingency table [114] of all possible overlaps between each pair of clusters in U and V (two different partitions), where the entry n_{ij} shows the intersection of cluster U_i and V_j ($n_{ij} = |U_i \cap V_j|$), $n_{i.} = \sum_j n_{ij}$ and $n_{.j} = \sum_i n_{ij}$ represent the table's row and column totals, respectively, with $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, r\}$, and $n = \sum_{i=1}^{k} \sum_{j=1}^{r} n_{ij}$ represents the number of observations.

The Adjusted Rand Index (ARI) [114] quantifies the similarity between partitions based on paired comparisons considering a correction for agreement by chance. This measure has a null value when the agreement between partitions does not exceed the agreement by chance.

$$ARI(U,V) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{r}\binom{n_{ij}}{2} - \left[\sum_{i=1}^{k}\binom{n_{i.}}{2}\sum_{j=1}^{r}\binom{n_{.j}}{2}\right]/\binom{n}{2}}{\frac{1}{2}\left[\sum_{i=1}^{k}\binom{n_{i.}}{2} + \sum_{j=1}^{r}\binom{n_{.j}}{2}\right] - \left[\sum_{i=1}^{k}\binom{n_{i.}}{2}\sum_{j=1}^{r}\binom{n_{.j}}{2}\right]/\binom{n}{2}} \tag{3.10}$$

The Variation of Information (VI) [115] measures the amount of information that is lost or gained in changing between partitions U and V and ranges from 0 (overlapping partitions) to $2\log(\max(k, r))$.

$$VI(U,V) = \sum_{i=1}^{k}\sum_{j=1}^{r}\frac{n_{ij}}{n}\log\left(\frac{n_{i.}n_{.j}/n^2}{n_{ij}^2/n^2}\right) \tag{3.11}$$

3.6.3 Clustering stability evaluated by K-fold cross-validation

To assess the stability of a clustering solution a method based on a K-fold cross-validation procedure, used in supervised learning, was adapted and applied to the pooled leading eigenvectors [116, 117]. This procedure is described in Table 3.1.

Table 3.1: Clustering K-fold cross-validation procedure used to evaluate stability. Adapted from [117].

Step	Action	Output
1	For each fold $k \in K$, split the data sample into training and test samples	Training and test samples
2	Cluster the training sample	Clusters in the training sample (partition P_1)
3	Apply a Nearest Centroid classifier to the test dataset	Classes in the test dataset (class set P_2)
4	Cluster the test dataset using the clustering algorithm of step 2	Clusters in the test sample (cluster set P_3)
5	Obtain a contingency table between partitions P_2 and P_3 of the test sample	Contingency table with the true class (P_2) and the predicted cluster (P_3)
6	Calculate clustering agreement measures	Indicator of stability of clustering solution

In this study, the data was randomly divided into 10 disjoint folds, each of which with approximately the same number of observations. Performing 10-fold cross-validation was preferred to simple cross-validation to increase the number of observations used as validation data. As observed in Table 3.1, for each training-test split, a Nearest Centroid Classifier was considered in step 3. This classifier assigned an observation (a leading eigenvector) to the cluster whose respective centroid/medoid was nearest to the mentioned observation. The clustering agreement measures used to evaluate the stability of a clustering solution were the ARI and the VI. Given that, for all 10 folds, the partitions of the test sample obtained by applying the classifier (P_2) and the clustering algorithm (P_3) contained the same number of clusters, the fraction of instances correctly assigned was computed by summing the values in the diagonal of the contingency table and dividing by the total number of observations (percent agreement) - measuring the level of association between the class set P_2 and the cluster set P_3.

4

Results

Contents

This chapter presents the results derived from the application of LEiDA to investigate differences in dFC between schizophrenia patients and healthy controls. Section 4.1 presents the findings obtained from performing a LEiDA analysis using the clustering solutions detected by the K-means algorithm. Lastly, the influence of using K-medoids instead of K-means to study dFC is investigated in Section 4.2.

4.1 Exploration of dynamic functional connectivity detected using the K-means algorithm

The K-means clustering algorithm was applied to the leading eigenvectors of all participants computed at each time point to detect FC states. The application of the K-means algorithm yielded 19 distinct clustering solutions (partition models), each considering a number of clusters, K, ranging from 2 to 20. For each clustering solution, the cluster centroids (FC states) served as the prototype of the obtained clusters and were assumed to represent averaged network configurations that consistently recurred across fMRI scans [76]. As stated, for each participant, given that each observation in the dataset represented a time point, the vector of cluster assignments represented a time course of FC states.

Here, Sections 4.1.1 and 4.1.2 outline the differences between groups found for the fractional occupancy and dwell time calculated for the detected FC states respectively. The overlap between the detected FC states and reference functional networks is assessed in Section 4.1.3. Section 4.1.4 evaluates the validity of clustering solutions through internal and external criteria. In Section 4.1.5 a tentative optimal clustering solution is selected and characterised. Its stability is examined in Section 4.1.6. Finally, the intergroup differences identified in the dynamical properties of the state trajectories produced by the optimal clustering solution are presented in Sections 4.1.7 and 4.1.8.

4.1.1 Fractional occupancy of functional connectivity states

The collection of clustering solutions was investigated to search for FC configurations whose mean fractional occupancy most significantly and consistently differed between schizophrenia patients and healthy controls. Figure 4.1 presents, for each clustering solution, K two-sided p-values obtained from evaluating whether the mean probability of occurrence of a FC state differed between schizophrenia patients and healthy controls. It must be noted that for a partition model with K FC states K hypothesis tests were performed. Consequently, instead of rejecting the null hypothesis at a significance level of $\alpha_1 = 0.05$, two corrected significance thresholds were considered. The first corrected significance threshold consisted in a Bonferroni corrected significance level of $\alpha_2 = 0.05/K$ - accounting for the increased probability of false positives. Additionally, a conservative significance threshold of $\alpha_3 = 0.05/\sum_{K=2}^{20} K$ was considered - correcting for the total number of computed comparisons.

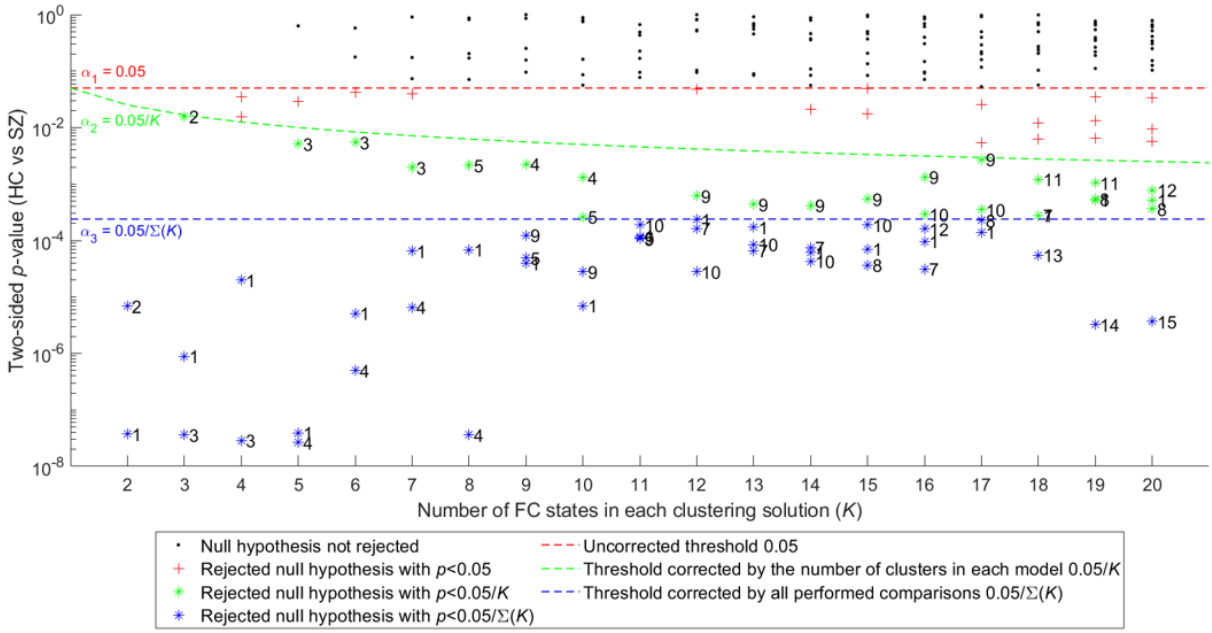


Figure 4.1: Two-sided p-values obtained for the intergroup comparisons of the mean fractional occupancy of each FC state for each partition model. FC states (clusters) are ranked according to their estimated probability of occurrence, where cluster 1 consisted of the largest number of objects and cluster K consisted of the least number of objects.

As depicted in Figure 4.1, for partitions of the dataset into a larger number of clusters, there was an increase in the number of FC states for which no significant differences in the mean probability of occurrence were detected. Furthermore, on average, as shown in Figure 4.1, the number of FC states for which the null hypothesis was rejected at a 5% significance level but did not survive the corrected significance thresholds was reduced $(\alpha_2 < p < \alpha_1)$. Notably, across all partition models, the mean fractional occupancy differed significantly between groups in at least two FC states ($p < \alpha_2$, two-tailed tests), as evidenced in Figure 4.1. In general, as observed in Figure 4.1, the clustering algorithm consistently produced clusters with cluster centroids, i.e., FC states, with an associated two-sided p-value that fell below the conservative significance threshold α_3.

As shown in Figure 4.1, the statistical procedure identified significant differences in the mean fractional occupancy of FC state 1 (centroid of the cluster with the largest number of assigned leading eigenvectors) between groups ($p < \alpha_2$, two-tailed tests) for all clustering solutions. In fact, the null hypothesis of no intergroup differences in the mean probability of occurrence of FC state 1 was rejected at a significance level of α_3 (two-tailed tests) for the clustering solutions with K ranging between 2 and 17 clusters. For each partition model, inspection of the cluster centroid associated with FC state 1 revealed this recurrent connectivity pattern represented a global mode of BOLD phase coherence (all elements of the cluster centroid had the same sign). Hence, FC state 1 was referred to as the Global Mode. Notably, the cluster centroids representing the Global Mode detected for each clustering solution

were highly correlated (minimum pairwise Pearson correlation coefficient of 0.706) - indicating that all FC states 1 represented similar forms of global BOLD phase coherence. Additionally, across all partition models, further non-global FC states were characterised by a significant difference in the mean fractional occupancy observed for the control and schizophrenia groups ($p < \alpha_2$, two-tailed tests). Importantly, visual inspection of the FC states with associated two-sided p-value below the significance threshold α_3 revealed these states represented varying forms of the same underlying patterns of connectivity. Specifically, regarding these states, it was noticed that connectivity patterns detected for lower values of K could be obtained by combining the fine-grained patterns of connectivity identified in partition models with larger values of K - evidencing the dependence among the hypothesis tests performed across partition models.

As stated, the mean probability of occurrence of a variety of FC states obtained across the range of clustering solutions was identified as significantly different between groups. To understand whether the mean fractional occupancy of the detected FC states was increased or decreased in schizophrenia patients in relation to healthy controls, one-tailed hypothesis tests were performed based on the inspection of the estimates of the group mean fractional occupancy presented in Figure 4.2.

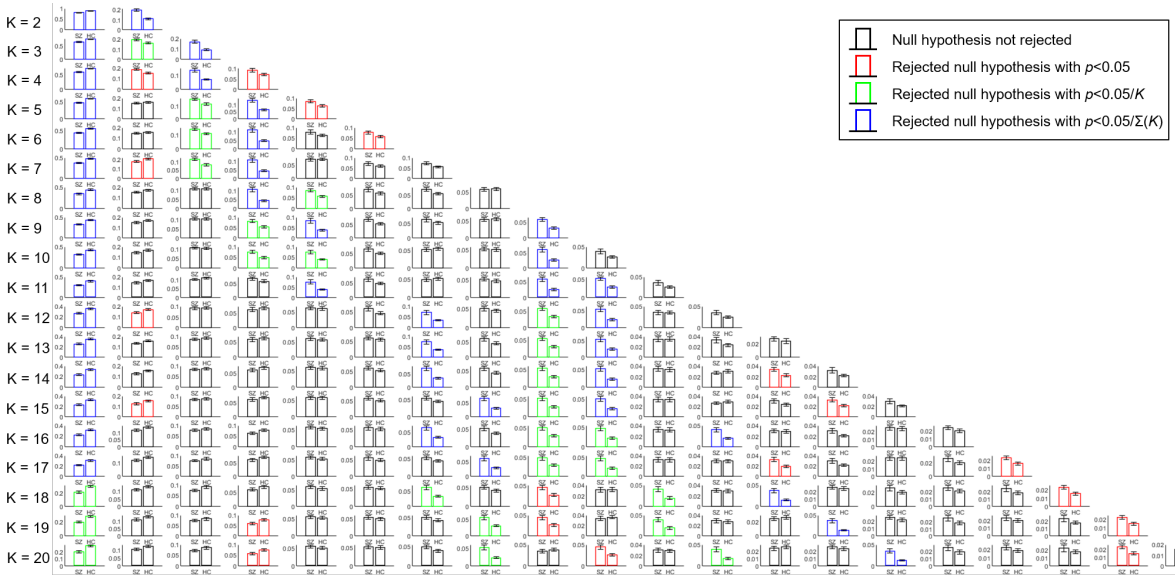


Figure 4.2: Barplot of the estimated mean fractional occupancy with associated standard error of each FC state for each group. For each FC state, the colour of the bars indicates whether the null hypothesis of no intergroup differences in the mean probability of occurrence was rejected (two-tailed tests). Black bars indicate the null hypothesis was not rejected at a 5% significance level. Red, green and blue bars indicate the null hypothesis was rejected at a 0.05, α_2 and α_3 significance level, respectively. The standard error of each bar was calculated as the standard deviation of the sample data divided by the square root of the sample size.

Analysis of Figure 4.2 suggested that the estimated mean fractional occupancy of FC state 1 was decreased in patients with schizophrenia compared to healthy controls irrespective of the clustering

solution. This was corroborated by the one-sided p-values computed from testing whether the mean probability of occurrence of this state was decreased in schizophrenia patients compared to healthy controls across all clustering solutions ($p < \alpha_3$ for $K \in \{2, \ldots, 18\}$, one-tailed tests). Furthermore, inspection of Figure 4.2 revealed that all non-global FC states with an associated two-sided p-value below the significance threshold α_3 (Figure 4.1) were characterised by a higher estimated mean probability of occurrence in the schizophrenia group compared to the control group. In fact, the one-sided p-values obtained from testing whether the mean fractional occupancy of each of these states was increased in patients with schizophrenia compared to healthy controls confirmed the previous findings ($p < \alpha_3$, one-tailed tests).

4.1.2 Dwell time of functional connectivity states

As previously described, the collection of clustering solutions revealed connectivity patterns which were consistently characterised by a mean fractional occupancy that differed significantly between groups. With respect to intergroup differences in the mean dwell time of the detected FC states, the K two-sided p-values calculated for each clustering solution are displayed in Figure 4.3.

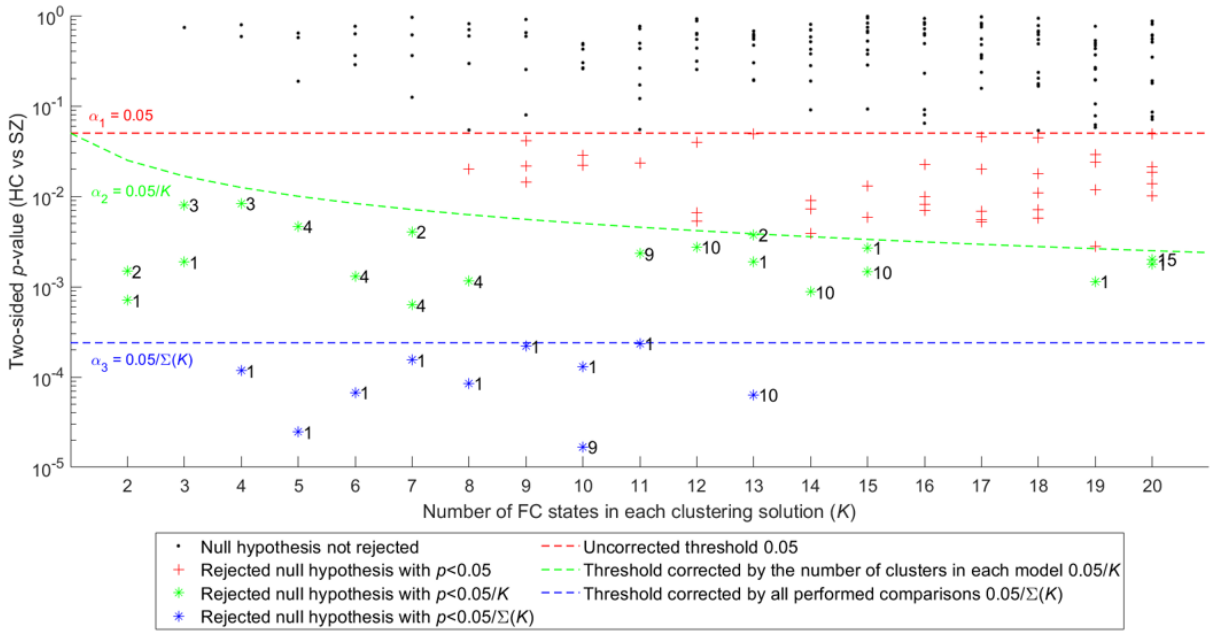


Figure 4.3: Two-sided p-values obtained from testing the null hypothesis of no difference between the mean dwell time of each FC state obtained for the control and schizophrenia groups. FC states (clusters) are ranked according to their estimated fractional occupancy.

Inspection of Figure 4.3 revealed that the number of FC states with an associated two-sided p-value that fell below the significance thresholds α_2 and α_3 was reduced in comparison to the findings of Figure 4.1. As shown in Figure 4.3, significant intergroup differences in the mean dwell time of FC state

1 were detected across the majority of the clustering solutions ($p < \alpha_2$, two-tailed tests). According to Figure A.2, the estimated mean dwell time of this state was decreased in schizophrenia patients compared to healthy controls. The one-tailed hypothesis tests confirmed the observed decrease was significant for a group of clustering solutions ($p < \alpha_3$ for $K \in \{4, \dots, 11\}$). Conversely, as depicted in Figure A.2, the estimated mean dwell time of a reduced number of non-global FC states appeared to be increased in the schizophrenia group compared to the control group. One-tailed hypothesis tests confirmed this increase to be significant at a significance level of α_3 for only two FC states. Notably, the Pearson correlation coefficient between the two mentioned states was found to be approximately 0.996 - reinforcing the fact that significant intergroup differences were consistently detected in analogous connectivity patterns irrespective of the clustering solution.

4.1.3 Overlap with reference functional networks

Intergroup differences in the mean fractional occupancy and in the mean dwell time were consistently detected in a number of varying forms of the same FC patterns. In Figure 4.4, for all partition models, the cluster centroids are projected into the cortical space and color-coded according to the reference RSN defined by [49] with which they most significantly overlapped. This aimed to elucidate the correlation among the mentioned states and assess their functional relevance.

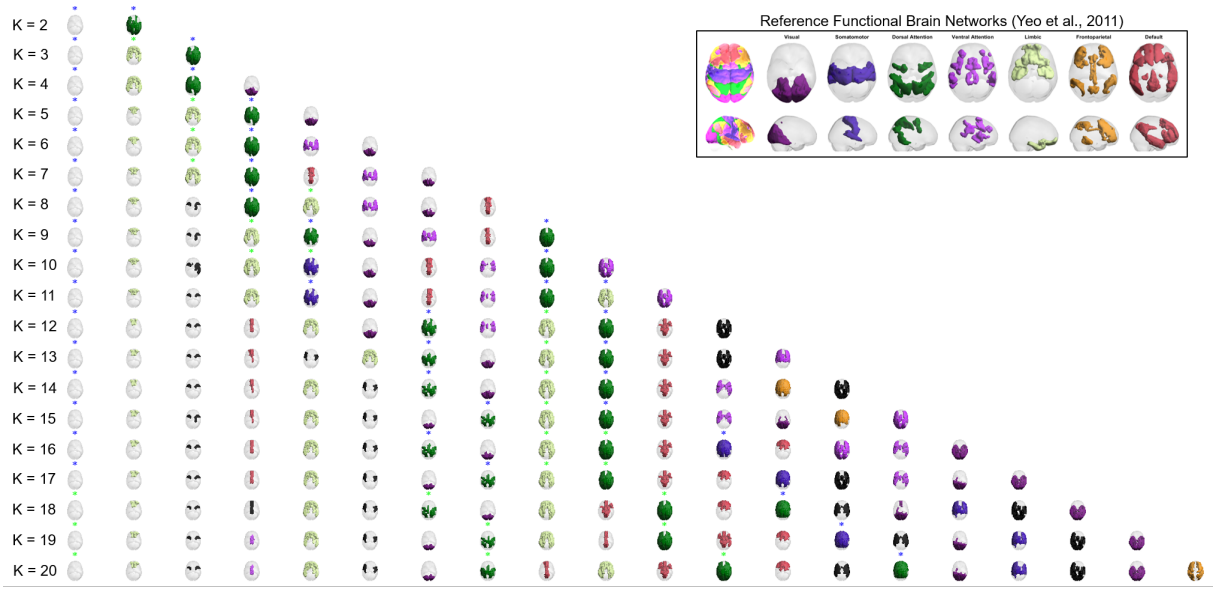


Figure 4.4: Representation of the centroids obtained for each clustering solution in cortical space. The rendered ROIs correspond to positive elements in the vectors of the centroids. Brain regions are coloured according to the reference RSNs defined by [49] whose p-value obtained from computing the Pearson correlation coefficient was lowest (with $p < 0.05/K$). Centroids not significantly overlapping with any of the reference RSNs are coloured in black. For each FC state, title asterisks indicate whether significant intergroup differences in the mean fractional occupancy were detected. Green and blue asterisks indicate $p < \alpha_2$ and $p < \alpha_3$ (two-tailed tests), respectively.

Firstly, as shown in Figure 4.4, FC state 1 was represented as a transparent brain across all partition models. This is because this state represented a global mode of BOLD phase coherence, with no activation of any particular subset of functionally coupled ROIs. In line with the findings from Figure 4.1, for the partition models with K ranging between 2 and 17, the title of each cortical representation of the Global Mode reiterated that significant differences in the mean fractional occupancy between groups were identified at a significance level of α_3 (two-tailed tests).

Secondly, concerning the non-global FC states, most of the cluster centroids were found to significantly overlap with a reference functional network defined in [49], as evidenced in Figure 4.4. It must be noted that each non-global FC state could have significantly overlapped with more than one reference functional network, i.e., the observed connectivity patterns could have encompassed the activation of multiple functionally relevant subsets of ROIs. As shown in Figure 4.4, centroids detected in distinct clustering solutions significantly overlapped with the same canonical RSNs. Furthermore, visual inspection of the corresponding cortical rendering revealed approximately similar patterns of connectivity were captured across these states and partition models. This result reinforced the finding that different partition models produced varying forms of the same underlying FC patterns.

Thirdly, the non-global FC states with a mean fractional occupancy identified as significantly increased in schizophrenia patients ($p < \alpha_2$, one-tailed tests) were found to repeatedly overlap with the same reference RSNs. Specifically, as depicted in Figure 4.4, the mentioned states shared significant spatial similarities with the Somatomotor network, Dorsal Attention network and Limbic network $(p < 0.05/K)$ - indicating that the mean fractional occupancy of networks with functional relevance closely related to that of the Somatomotor, Dorsal Attention and Limbic RSNs was significantly increased in schizophrenia patients. In fact, for each group of non-global states with identified significant intergroup differences in the group mean fractional occupancy ($p < \alpha_2$, two-tailed tests) and overlapping with the Somatomotor, Dorsal Attention and Limbic networks, computation of the Pearson correlation coefficient between the cluster centroids revealed the lowest correlation coefficient for each group was 0.870, 0.594 and 0.858, respectively. The relatively lower correlation coefficient obtained for the group of cluster centroids overlapping with the Dorsal Attention network was due to the fact that, for partition models with more than 11 clusters, at least two FC states significantly overlapped with this reference RSN, with each state being characterised by the activation of different subsets of ROIs, as observed in Figure 4.4.

Finally, together with Figure 4.3, inspection of Figure 4.4 showed that the non-global FC states characterised by a significantly increased mean dwell time in schizophrenia patients compared to healthy controls ($p < \alpha_2$, one-tailed tests) significantly overlapped with the Dorsal Attention network and the Limbic network.

Overall, the K-means algorithm provided clustering solutions which suggested that FC state 1 recurred significantly less ($p < \alpha_3$, one-tailed tests) and lasted significantly reduced consecutive periods

of time ($p < \alpha_2$, one-tailed tests) in schizophrenia patients compared to healthy controls. Additionally, across partition models, the mean fractional occupancy and the mean dwell time of the non-global FC states, possibly with functional activity analogous to that of the Somatomotor, Dorsal Attention and Limbic RSNs, was found to be significantly increased in schizophrenia patients compared to healthy controls ($p < \alpha_2$, one-tailed tests). These findings are graphically summarised in Table 4.1.

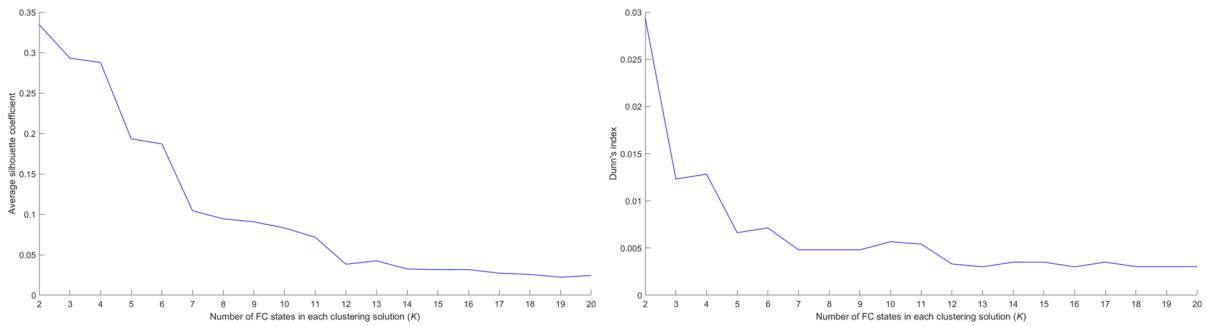
Table 4.1: Graphical summary of the mean fractional occupancy and of the mean dwell time in schizophrenia patients compared to healthy controls observed across partition models detected by the K-means algorithm. The non-global FC states are represented by the reference RSN defined by [49] with which they most significantly overlapped. Upward/downward arrows indicate the mean fractional occupancy and/or the mean dwell time were significantly increased/decreased in the schizophrenia group in relation to the control group across partition models ($p < \alpha_2$, one-tailed tests). Crosses indicate the null hypothesis of no intergroup differences in the mean of each measure were not rejected at a significance level of α_2.

		Reference functional networks defined by [49]						
	Global Mode	Visual	Somat.[a]	Dorsal Attention	Ventral Attention	Limbic	Front.[b]	Default
Fractional Occupancy	⇓	×	⇑	⇑	×	⇑	×	×
Dwell Time	⇓	×	×	⇑	×	⇑	×	×

[a] Somatomotor; [b] Frontoparietal;

4.1.4 Internal and external validation of K-means clustering solutions

Figure 4.5 depicts the average Silhouette coefficient and the Dunn's index utilised to evaluate the quality of the clustering solutions produced by the K-means algorithm.



(a) Average Silhouette coefficient. **(b)** Dunn's index.

Figure 4.5: Internal validation of K-means clustering solutions.

As observed in Figure 4.5, the maximum value in both criteria was obtained for clustering solutions with a low number of FC states, which were of limited interest for the present study. Contrarily, for

clustering solutions with K larger than 12, both criteria remained relatively constant at low values - indicating such partitions were also of limited interest. Notably, for clustering solutions with a number of clusters between 7 and 11, the average Silhouette coefficient decreased smoothly and the Dunn's index remained relatively constant, as shown, respectively, in Figures 4.5(a) and 4.5(b) - suggesting that these partition models were of potential interest for further analysis.

Clustering solutions detected by the K-means algorithm using as input the pooled leading eigenvectors from all participants were compared with clustering solutions detected by the K-means algorithm using only the leading eigenvectors from the control group as input. This comparison revealed number matching FC states were moderately correlated. Remarkably, the FC states from each of the clustering approaches with a mean fractional occupancy identified as significantly different between groups ($p < \alpha_3$, two-tailed tests) were found to be highly intercorrelated. Furthermore, results from investigating the functional relevance of the FC states detected by using only the leading eigenvectors from the control group were identical to the findings reported in Table 4.1. Lastly, the ARI and the VI computed between clustering solutions with the same number of clusters detected by each approach indicated relatively similar partitions of the data were obtained across the range of K.

4.1.5 Selection of the optimal clustering solution

For the subsequent analysis, a tentative optimal clustering solution was selected. In this study, the partition model with 11 FC states was selected as the optimal clustering solution. The rationale leading to this decision was based upon the ability to identify a collection of FC states with properties that significantly differed between groups and the quality and stability of the actual partition of the data.

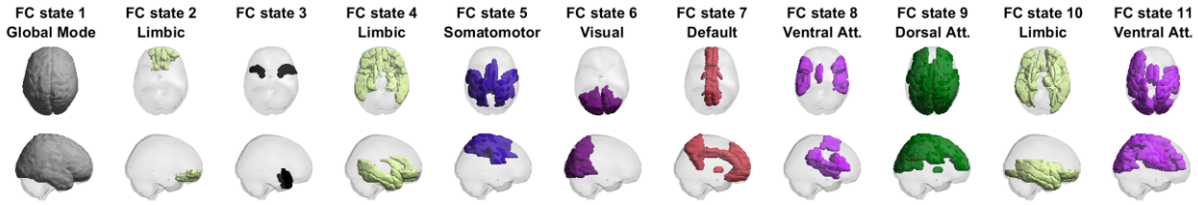
Firstly, visual inspection of Figure 4.1 revealed this partition model provided an optimal separation between the FC states with mean fractional occupancy that did and did not differ between groups - reducing the likelihood of considering false negatives and false positives in subsequent analysis. Specifically, the null hypothesis of no intergroup differences in the mean fractional occupancy of 4 FC states was rejected at a significance level of α_3 (two-tailed tests), while the same hypothesis was not rejected for the remaining 7 FC states ($p > 0.05$, two-tailed tests). Furthermore, as shown in Figure 4.3, the mean dwell time of 2 FC states was identified as significantly different between groups ($p < \alpha_2$, two-tailed tests). It must be noted that the two-sided p-value associated with FC state 9 did not fall below the α_3 significance level, as observed in Figure 4.3. Nonetheless, this FC state had a Pearson correlation coefficient of 0.999 with FC state 9 of the partition model with $K = 10$ clusters whose associated two-sided p-value fell below the α_3 significance level.

Secondly, the FC states with significantly different intergroup mean fractional occupancy and significantly different intergroup mean dwell time revealed a broad spectrum of functional relevance. In fact, the functional relevance of these states was found to possibly relate to that of the Somatomotor, Dorsal

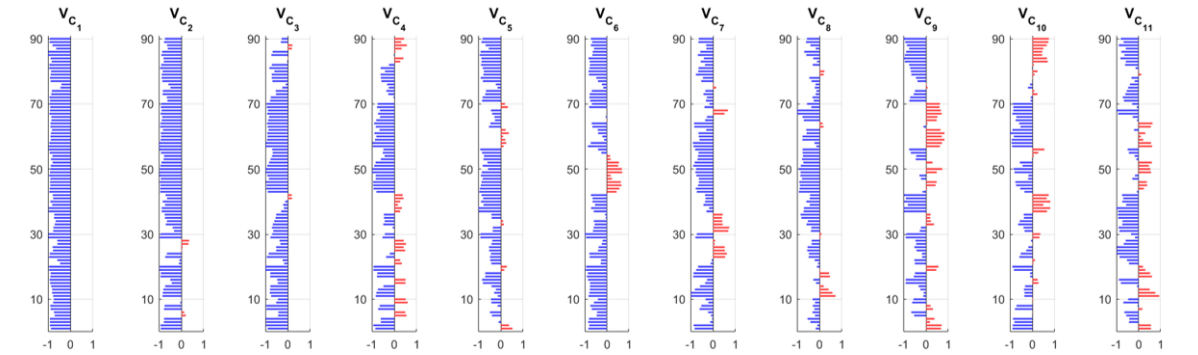
Attention and Limbic networks defined by [49], as demonstrated in Figure 4.4.

Finally, the quality of the structure of the clustering solution with 11 FC states was considered to lay within the range of acceptable values according to the evaluated internal validation criteria, as evidenced in Figure 4.5. In detail, the average Silhouette coefficient and the Dunn's index were relatively similar to the values observed for the clustering solutions with K ranging between 7 and 10 clusters. Additionally, considering 11 FC states was speculated to better enable the detection of specific and fine-grained functional networks with potentially altered activity in schizophrenia patients.

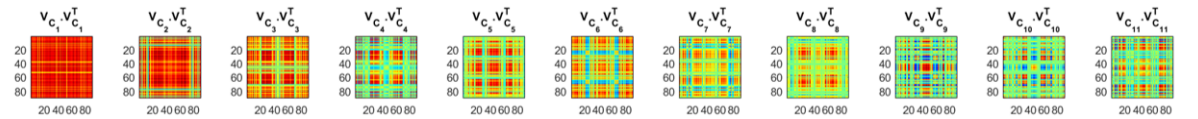
The collection of 11 BOLD phase coherence patterns represented the functional activation of different subsets of ROIs, as depicted in Figure 4.6.



(a) Cortical space representation of each FC state.



(b) Vector representation of each FC state.



(c) Matrix representation of each FC state.

Figure 4.6: Collection of 11 FC states obtained from the selected optimal clustering solution. The FC states are arranged (left-to-right) according to decreasing estimated fractional occupancy.

For the selected partition model, the mean fractional occupancy of the Global Mode was significantly decreased in schizophrenia patients compared to healthy controls ($p < \alpha_3$; Hedge's $g = 0.694$, medium to large effect size). The estimated mean occurrence probability of this state was $29.8 \pm 19.8\%$ and $39.8 \pm 14.9\%$ (mean $\pm$ std) for the schizophrenia and control groups respectively. The mean fractional

occupancy of FC states 5, 9 and 10 was significantly increased in schizophrenia patients compared to healthy controls ($p < \alpha_3$; Hedge's $g = \{0.611, 0.630, 0.629\}$, respectively, medium to large effect size). The estimated mean fractional occupancy for FC states 5, 9 and 10 was $7.61 \pm 7.78\%$, $5.67 \pm 6.34\%$ and $5.14 \pm 4.09\%$ for the schizophrenia group and $4.02 \pm 3.15\%$, $2.52 \pm 3.24\%$ and $2.94 \pm 2.81\%$ for the control group, respectively (mean $\pm$ std). These results are graphically summarised in Figure 4.7.

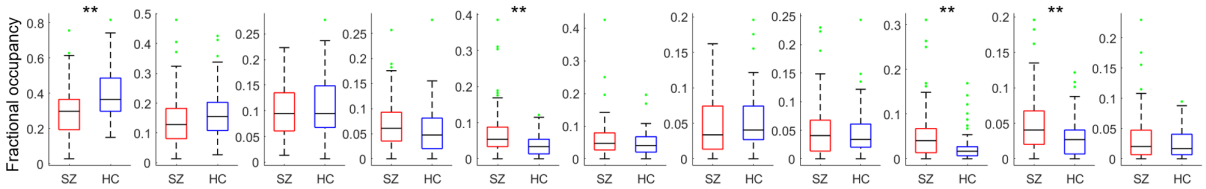


Figure 4.7: Boxplot of the fractional occupancy values for FC state 1 to 11 (left-to-right) for the schizophrenia and control groups. Asterisks indicate significant intergroup differences ($p < \alpha_3$, one-tailed tests). Green points represent outliers, according to the Tukey criterion.

The mean dwell time of the Global Mode was significantly decreased in schizophrenia patients compared to healthy controls ($p < \alpha_3$; Hedge's $g = 0.618$, medium to large effect size). The estimated mean dwell time of this state was $4.986 \pm 2.306\,s$ and $6.574 \pm 2.803\,s$ (mean $\pm$ std) for schizophrenia patients and healthy controls respectively. The mean dwell time of FC state 9 was significantly increased in schizophrenia patients compared to healthy controls ($p < \alpha_2$; Hedge's $g = 0.519$, medium to large effect size). The estimated mean dwell time of FC state 9 was $2.287 \pm 0.763\,s$ for the schizophrenia group and $1.830 \pm 0.977\,s$ for the control group (mean $\pm$ std). The boxplots of Figure 4.8 represent the dwell time values for each of the 11 FC states.

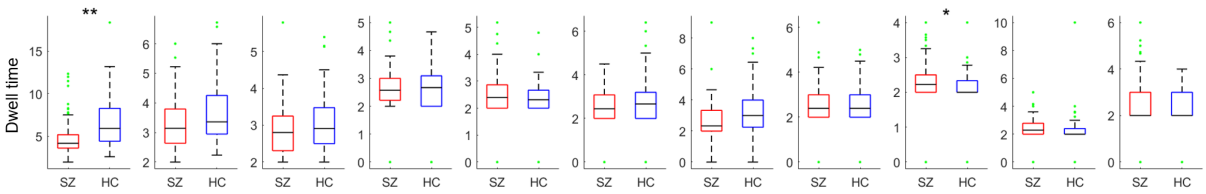


Figure 4.8: Boxplot of the dwell time values for FC state 1 to 11 (left-to-right) for the schizophrenia and control groups. Single and double asterisks indicate significant intergroup differences with $p < \alpha_2$ and $p < \alpha_3$ (one-tailed tests), respectively. Green points represent outliers, according to the Tukey criterion.

4.1.6 Examination of the stability of the optimal clustering solution

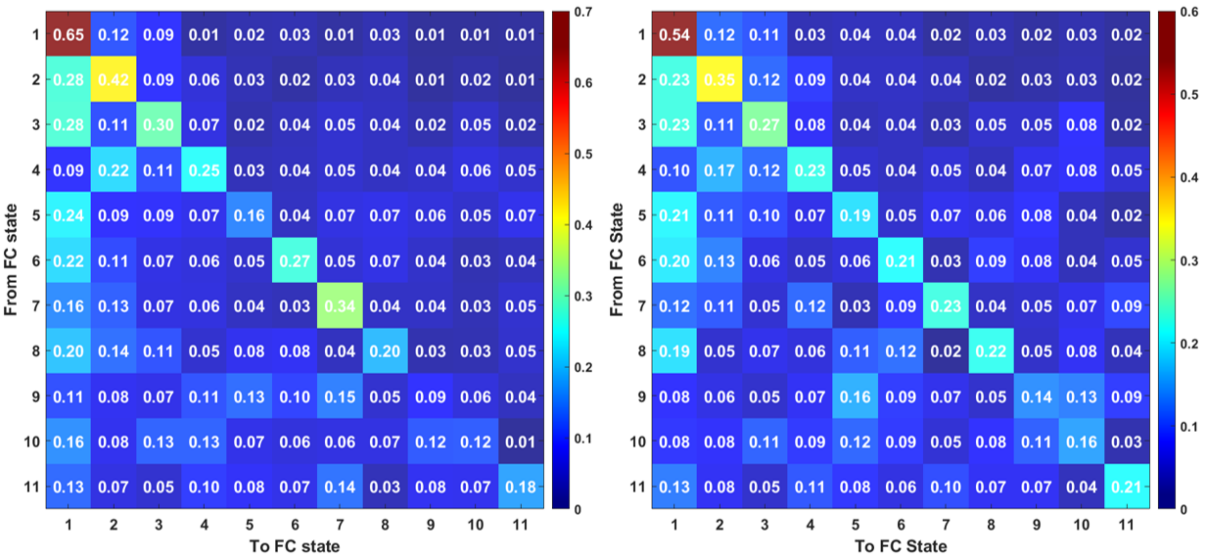
The percent agreement, ARI and VI obtained for each fold of the 10-fold cross-validation procedure described in Table 3.1 are provided in Table 4.2. The results suggested respectively, good levels of association and paired agreement between partitions of the test sample and that the amount of information that was lost in changing from the class set P_2 to the cluster set P_3 of the test sample was relatively low. Consequently, the optimal clustering solution was considered valid and appropriate for further analysis.

Table 4.2: Results from the 10-fold cross-validation stability analysis of the optimal clustering solution.

	Cross-validation fold number									
Indices	1	2	3	4	5	6	7	8	9	10
Percent agreement	0.642	0.536	0.583	0.688	0.529	0.564	0.602	0.698	0.547	0.523
Adjusted Rand	0.692	0.712	0.717	0.757	0.696	0.678	0.749	0.718	0.773	0.709
Variation of information	1.379	1.504	1.349	1.237	1.396	1.402	1.291	1.229	1.181	1.388

4.1.7 Transition probability matrix of the optimal state trajectories

With respect to the optimal clustering solution with 11 FC states, for all participants, the individual DTMC defined by the temporal trajectory through the finite state space $S' = \{1, \ldots, 11\}$, was characterised by its estimated one-step TPM. The estimated average TPM for the control group, $\mathbf{P}^{(\mathrm{HC})}$, and the schizophrenia group, $\mathbf{P}^{(\mathrm{SZ})}$, is depicted in Figures 4.9(a) and 4.9(b), respectively.



(a) Estimated average TPM of the control group.

(b) Estimated average TPM of the schizophrenia group.

Figure 4.9: Estimated group average one-step TPM. Each element of $\mathbf{P}^{(\mathrm{HC})}$ and $\mathbf{P}^{(\mathrm{SZ})}$ represents the estimated average one-step probability of switching from state α to state β with $\alpha, \beta \in S'$.

As shown in Figure 4.9, the estimated group average TPMs were aperiodic and asymmetric - reinforcing the transient emergence and dissolution of distinct connectivity patterns. The highest estimated group average transition probabilities were observed in both Figures 4.9(a) and 4.9(b) along the diagonal

(representing the probability of remaining in the same FC state) and the first column (representing the probability of transitioning back to FC state 1) - indicating that the dynamical system of state trajectories presented a higher propensity to maintain a given pattern of connectivity (stay locked within a given FC state) and to return back to a state of global BOLD phase coherence. Furthermore, the FC state transitions of the form $\alpha \to \alpha$ $(\alpha \in S')$, which indicate the relative stability of each FC state [76], suggested that FC state 1 displayed the highest stability (with $65 \pm 14.2\%$ and $54 \pm 15.9\%$ (mean $\pm$ std) estimated mean probability of remaining in this state for the control and schizophrenia groups, respectively), while the stability of the remaining non-global FC states was relatively lower.

For each one-step probability of transitioning from state α to state β $(\alpha \to \beta,\ \alpha, \beta \in S')$, a two-sided p-value was obtained by testing whether its group mean differed between groups. The state-to-state transitions that were significantly affected in the schizophrenia group compared to the control group are presented in Figure 4.10.

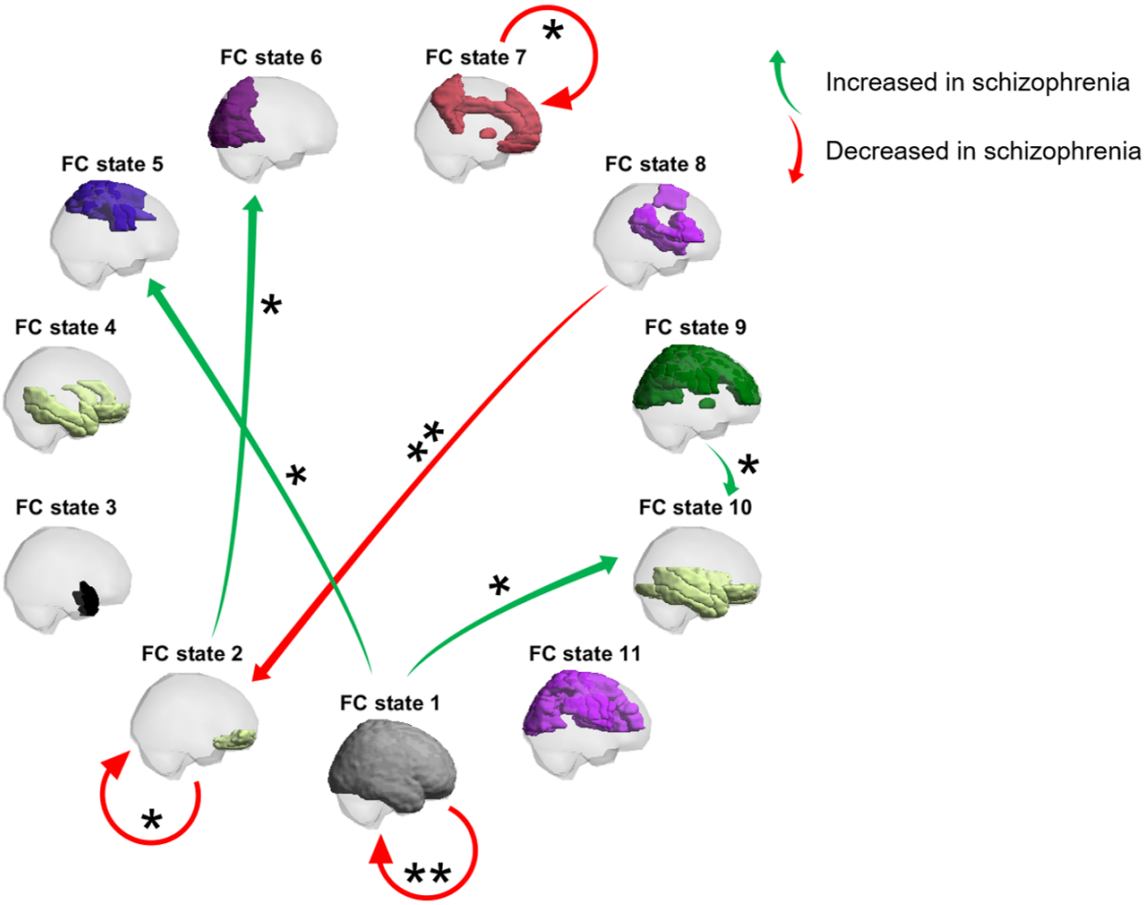


Figure 4.10: Transition diagram of the state-to-state transitions significantly altered in schizophrenia patients compared to healthy controls. Each arrow represents a mean transition probability that was significantly increased (green) or significantly decreased (red) in schizophrenia patients compared to healthy controls. Single and double asterisks indicate significant intergroup differences with $p < 0.05/11$ and $p < 0.05/(11 \times 11)$ (one-tailed tests), respectively.

On the one hand, as shown in Figure 4.10, the mean probability of remaining in a state of global BOLD phase coherence (FC state 1) was significantly reduced in schizophrenia patients compared to healthy controls ($p < 0.05/(11 \times 11)$, one-tailed test; Hedge's $g = 0.726$, medium to large effect size). Additionally, the mean probability of transitioning from FC state 8 to FC state 2 was also found to be significantly reduced in schizophrenia patients ($p < 0.05/(11 \times 11)$, one-tailed test; Hedge's $g = 0.568$, medium to large effect size). Lastly, the mean probability of remaining in FC states 2 and 7 was significantly reduced in the schizophrenia group compared to the control group ($p < 0.05/11$, one-tailed tests; Hedge's $g = \{0.449, 0.515\}$, respectively, small to medium effect size). On the other hand, the mean probability of transitioning from the Global Mode to FC states 5 and 10 was significantly increased in schizophrenia patients compared to healthy controls ($p < 0.05/11$, one-tailed tests; Hedge's $g = \{0.452, 0.513\}$, respectively, small to medium effect size). Furthermore, the mean probability of switching from FC state 9 to FC state 10 was identified to be significantly increased in the schizophrenia group compared to the control group ($p < 0.05/11$, one-tailed test; Hedge's $g = 0.461$, small to medium effect size). Finally, the mean probability of switching from FC state 2 to FC state 6 was found to be significantly increased in schizophrenia patients compared to healthy controls ($p < 0.05/11$, one-tailed test; Hedge's $g = 0.466$, small to medium effect size).

4.1.8 Limiting probability of the optimal functional connectivity states

The estimates of the mean row vector representing the stationary distribution for the group of 37 healthy controls, $\boldsymbol{\pi}^{(\text{HC})} = \left[\pi_{\beta}^{(\text{HC})}\right]_{\beta \in S'}$, and the group of 46 schizophrenia patients, $\boldsymbol{\pi}^{(\text{SZ})} = \left[\pi_{\beta}^{(\text{SZ})}\right]_{\beta \in S'}$, with irreducible and aperiodic DTMCs, and countable state space $S' = \{1, \ldots, 11\}$, are given by:

$$\begin{aligned} \boldsymbol{\pi}^{(\text{HC})} &= \begin{bmatrix} 0.309 & 0.157 & 0.112 & 0.070 & \boxed{0.052} & 0.059 & 0.063 & 0.057 & 0.039 & \boxed{0.043} & 0.038 \end{bmatrix} \\ \boldsymbol{\pi}^{(\text{SZ})} &= \begin{bmatrix} 0.272 & 0.135 & 0.109 & 0.066 & \boxed{0.077} & 0.060 & 0.054 & 0.064 & 0.056 & \boxed{0.058} & 0.049 \end{bmatrix} \end{aligned} \tag{4.1}$$

where the highlighted values refer to estimates of the group mean limiting probabilities for which the null hypothesis of no intergroup differences was rejected at a 5% significance level. Interestingly, the null hypothesis of no intergroup differences in the mean limiting probabilities was never rejected at the corrected significance value $0.05/11$. Inspection of Equation (4.1) indicated that the estimated long-run proportion of TRs spent in FC state 1 was 0.309 ± 0.104 for the group of 37 healthy controls and 0.272 ± 0.111 for the group of 46 schizophrenia patients (mean $\pm$ std). Surprisingly, no intergroup differences were found in the group mean limiting probability of this state (two-tailed test; Hedge's $g = 0.342$, small to medium effect size). The group mean limiting probability of FC states 5 and 10 was identified as significantly increased in the schizophrenia group compared to the control group ($p < 0.05$, one-tailed tests; Hedge's $g = \{0.464, 0.449\}$, respectively, small to medium effect size).

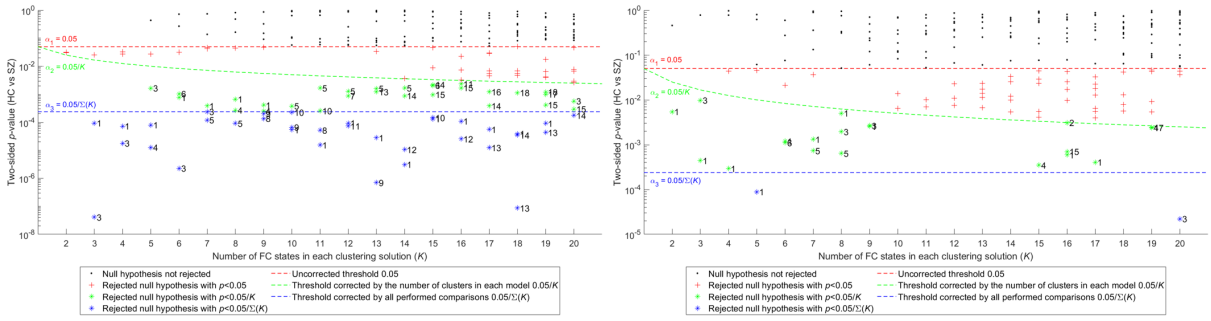
4.2 Influence of using the K-medoids algorithm instead of the K-means algorithm

The leading eigenvectors from all participants were used as input to the K-medoids clustering algorithm. The algorithm considered values of K that ranged between 2 and 20 - producing 19 distinct partition models consisting of K clusters and respective cluster medoids. For each clustering solution, the cluster medoids (FC states) served as the prototype of the obtained clusters and were assumed to be the most representative network configurations among groups of consistently recurring connectivity patterns.

Here, the clustering solutions detected by the K-medoids algorithm are analysed in Section 4.2.1. Specifically, the fractional occupancy, the dwell time and the functional relevance of the detected FC states are investigated. Lastly, the clustering solutions outputted by the K-medoids algorithm and the K-means algorithm are compared in Section 4.2.2.

4.2.1 Investigation of clustering solutions detected by the K-medoids algorithm

The K two-sided p-values obtained from testing whether the mean fractional occupancy and the mean dwell time differed between healthy controls and patients with schizophrenia are depicted in Figure 4.11.



(a) Fractional occupancy.

(b) Dwell Time.

Figure 4.11: Two-sided p-values obtained from the intergroup comparisons between the mean fractional occupancy and the mean dwell time of each FC state for each partition model detected by the K-medoids algorithm. FC states (clusters) are ranked according to their estimated fractional occupancy.

Analysis of Figure 4.11 revealed that the application of the K-medoids algorithm enabled the detection of FC states with a mean fractional occupancy and a mean dwell time which consistently and significantly differed between groups, in line with the results from Sections 4.1.1 and 4.1.2. Across the range of partition models detected by the K-medoids algorithm FC state 1, which represented a global mode of BOLD phase coherence, was repeatedly identified as having a significantly different mean fractional occupancy in the schizophrenia group compared to the control group ($p < \alpha_2$, two-tailed tests), as observed in Figure 4.11(a). Significant intergroup differences in the mean fractional occupancy were

also identified in a number of non-global FC states ($p < \alpha_2$, two-tailed tests). With respect to the non-global states with an associated two-sided p-value below the significance threshold α_3, the connectivity patterns detected by both the K-means and K-medoids algorithms were found to be highly correlated - representing the phase coherence among the BOLD signals of similar subsets of ROIs. In agreement with previous findings, intergroup differences in the mean dwell time were only detected in a reduced number of FC states, as evidenced in Figure 4.11(b).

In detail, across partition models and FC states detected by the K-medoids algorithm, the increase or decrease of the estimates of the group mean fractional occupancy and of the estimates of the group mean dwell time in the schizophrenia group relative to the control group were approximately analogous to the findings derived using the K-means algorithm, as observed in Figures A.3(a) and A.3(b). In fact, the mean fractional occupancy and the mean dwell time of the Global Mode were identified to be significantly decreased in the schizophrenia group compared to the control group across a number of clustering solutions ($p < \alpha_2$, one-tailed tests). Finally, the mean fractional occupancy and the mean dwell time were found to be significantly increased in schizophrenia patients compared to healthy controls at a significance level of α_2 across multiple non-global FC states (one-tailed tests), as shown in Figures A.3(a) and A.3(b).

The overlap between the cluster medoids (FC states) detected by the K-medoids algorithm and the reference functional networks defined by [49] is depicted in Figure A.4 - revealing the detected FC patterns displayed the activation of subsets of ROIs with possibly relevant functional relevance. The combination of the findings from Figures A.3 and A.4 is summarised in Table 4.3.

Table 4.3: Graphical summary of the mean fractional occupancy and of the mean dwell time in schizophrenia patients compared to healthy controls observed across partition models detected by the K-medoids algorithm. The non-global FC states are represented by the reference RSN defined by [49] with which they most significantly overlapped. Upward/downward arrows indicate the mean fractional occupancy and/or the mean dwell time were significantly increased/decreased in the schizophrenia group in relation to the control group across partition models ($p < \alpha_2$, one-tailed tests). Crosses indicate the null hypothesis of no intergroup differences in the mean of each measure were not rejected at a significance level of α_2.

		Reference functional networks defined by [49]						
	Global Mode	Visual	Somat.[a]	Dorsal Attention	Ventral Attention	Limbic	Front.[b]	Default
Fractional Occupancy	⇓	×	⇑	⇑	×	⇑	no overlap	×
Dwell Time	⇓	×	×	⇑	×	⇑	no overlap	×

[a] Somatomotor; [b] Frontoparietal;

Inspection of Table 4.3 suggested that the connectivity patterns detected by the K-medoids algorithm for which significant intergroup differences in the mean fractional occupancy and the mean dwell time were identified represented FC configurations with functional relevance closely related to that of the

patterns of connectivity detected by the K-means algorithm whose mean fractional occupancy and mean dwell time also differed significantly between groups.

4.2.2 Comparison of partition models detected by K-means and K-medoids

The clustering solutions detected by the K-medoids algorithm were compared to the partition models previously obtained by the K-means algorithm. For each of the 19 clustering solutions, the correlation coefficient between the centroids and the medoids with the same label detected by each of the algorithms was found to be relatively low. In detail, the ARI and the VI showed that the clustering solutions with the same number of FC states detected by each of the clustering algorithms exhibited relatively low paired agreement and that the amount of information that was lost from changing between clustering solutions was relatively high. The values calculated for each of the external clustering agreement measures are provided in Table 4.4 for the clustering solutions with K ranging between 5 and 14 clusters.

Table 4.4: Comparison between clustering solutions with the same number of clusters detected by the K-means and K-medoids algorithms.

	Clustering solution with K FC states									
Indices	5	6	7	8	9	10	11	12	13	14
Adjusted Rand	0.684	0.541	0.581	0.552	0.564	0.467	0.446	0.406	0.385	0.388
Variation of information	1.054	1.499	1.345	1.594	1.642	1.926	2.039	2.206	2.372	2.421

The results of Table 4.4 suggested that the clustering algorithms influenced the outputted clustering solutions. Furthermore, closer inspection of Table 4.4 showed the ARI decreased and the VI increased for successively increasing values of K - indicating that the partitions of the pooled leading eigenvectors obtained from each of the clustering algorithms were increasingly dissimilar. This could suggest that the detected centroids and the detected medoids, i.e., the FC states, represented the activation of progressively different subsets of ROIs. Nevertheless, for each partition model with K clusters, the FC states detected by each of the clustering algorithms whose mean fractional occupancy and mean dwell time differed significantly between groups ($p < \alpha_2$, two-tailed tests) were found to be highly correlated - representing approximately identical connectivity patterns. Therefore, despite providing distinct partition models, both the K-means algorithm and the K-medoids algorithm were found to effectively identify similar FC patterns with properties that significantly and consistently differed between groups.

5

Discussion

This MSc project investigated differences in resting-state dFC in schizophrenia patients compared to healthy controls through the application of an approach focusing solely on the dominant FC pattern captured by the leading eigenvector of BOLD phase coherence matrices [21]. Under this framework, brain activity was conceived as a DTMC which represents the discrete temporal evolution of resting-state connectivity patterns through a finite state space consisting of whole-brain BOLD phase coherence configurations. By quantitatively estimating the fractional occupancy, dwell time and limiting probability of each FC state, as well as the transition probabilities between states, it was possible to characterise the time-resolved functional connectivity in both schizophrenia patients and healthy controls. Supported by the analysis of clustering results, the permutation-based statistical procedure facilitated the identification of recurrent connectivity patterns with properties distinctly and consistently altered in schizophrenia.

In agreement with previous studies [21, 76–78], two distinct groups of BOLD phase coherence states were detected by both the K-means algorithm and the K-medoids algorithm across all partition models (Figures 4.4 and A.4). On the one hand, the algorithms detected the so-called Global Mode (FC state 1), which corresponded to a state of whole-brain BOLD signal phase coherence. Specifically, this FC state was found to represent a strong, large whole-brain connectivity network. Varying forms of this FC state were detected for all clustering solutions - suggesting that this pattern of connectivity is a hallmark of time-varying FC. On the other hand, the clustering algorithms consistently identified a set of non-global connectivity patterns which have been referred to as "ghost" attractor states [64, 76]. These states represented the activation of a subset of ROIs with an instantaneous BOLD phase in anti-phase compared to the bulk of the remaining BOLD phase orientations.

In this study, corrected significance thresholds were considered to reduce the likelihood of false positives when performing hypothesis tests. Specifically, the Bonferroni corrected significance threshold, $0.05/K$, accounted for the K independent hypothesis tests performed within each partition model with K FC states. Consistent with prior LEiDA analyses [21, 76–78], FC states detected across clustering solutions were found to represent approximately varying forms of the same FC configurations - evidencing the dependency among the tested null hypotheses across K. Consequently, a further conservative corrected significance threshold $(0.05/\sum_{K=2}^{20} K)$ was considered to encompass both dependent and independent hypotheses tests performed within and across clustering solutions. Given the observed dependency among hypotheses tests, the reported findings mainly concerned FC states for which the null hypothesis of no intergroup differences was rejected at a significance level of $0.05/\sum_{K=2}^{20} K$.

Overall, for all clustering solutions detected by the K-means algorithm, significant intergroup differences in the mean fractional occupancy were identified for a number of connectivity patterns with a high confidence level (Figure 4.1). Notably, a significant decrease in the mean fractional occupancy of the Global Mode in the schizophrenia group compared to the control group was consistently detected for all clustering solutions - suggesting that this state of global BOLD phase coherence recurred significantly

less in schizophrenia patients compared to healthy controls (Figure 4.2). This finding was in line with those of previous studies obtained using different approaches to investigate dFC [79,85,87,93,94]. Conversely, across partition models, the current study found that the mean fractional occupancy of a number of varying forms of the same underlying non-global connectivity patterns was significantly increased in schizophrenia patients compared to healthy controls. This indicated that a number of FC states characterised by the activation of a small subset of ROIs recurred significantly more often in schizophrenia patients than in healthy controls (Figure 4.2). In general, therefore, it seemed that the temporal evolution of brain activity in schizophrenia patients spent a significantly increased amount of time transitioning through the repertoire of "ghost" attractor states, as shown in [27, 79, 84, 86, 87, 93, 94]. These findings point to the direction that, even if the temporal state trajectories of schizophrenia patients and healthy controls span similar connectivity patterns, the recurrence of the FC states differs significantly between groups which could be speculated to explain the observed behavioural symptoms of schizophrenia, such as the collapse of psychological functions resulting in delusions and thought disorder.

In this study, the estimates of the dwell time of FC states were found to provide reduced sensitivity to identify discerning connectivity patterns between groups (Figure 4.3). In agreement with [78], a possible explanation for this might be the low temporal resolution of the neuroimaging data analysed in this research (TR = 2 s) - hindering the ability to capture the rapid fluctuations of FC, which have been revealed by MEG studies to occur at time scales of approximately 200 ms [118]. Despite this, the possibility that intergroup differences relate mostly to the fractional occupancy of connectivity patterns rather than their mean duration cannot be ruled out. Findings derived from estimates of this measure therefore need to be interpreted with caution. In agreement with previous studies [85, 93, 94], the mean dwell time of the Global Mode was found to be significantly decreased in the schizophrenia group compared to the control group in a number of partition models (Figure A.2). Furthermore, the mean dwell time of a reduced number of "ghost" attractor states was identified as significantly increased in schizophrenia patients in relation to healthy controls. In general, the estimates of the mean dwell time provided less significant and consistent differences between groups - raising questions as to the validity and efficacy of this measure [77].

Investigations of the overlap between the detected FC states and the seven reference RSNs defined in [49] suggested that the Global Mode did not represent the activation of any particular known functional network (Figure 4.4), as observed in previous studies [76, 78]. In contrast, the "ghost" attractor states were found to closely relate to the reference functional networks defined by [49] - suggesting that the phase shift in the BOLD signal of a given subset of brain areas, captured by LEiDA, revealed a separation of the connectivity patterns into distinct functionally meaningful systems. Overall, the results indicated that the significantly increased mean fractional occupancy in the schizophrenia group compared to the control group concerned "ghost" attractor states with functional relevance possibly related to that of the

Somatomotor, Dorsal Attention and Limbic reference RSNs (Table 4.1). These results supported previous research into altered dFC in brain functional subsystems [91]. Notably, despite the lack of a full understanding of the relationship between connectivity patterns observed in Electroencephalogram (EEG) and fMRI, these findings could be speculated to portray a temporal dynamics related to that observed with EEG microstates measured at a different time resolution. In fact, in line with the aforementioned findings, EEG studies have reported an increased recurrence of a microstate partially associated with the Limbic RSN in schizophrenia patients compared to healthy controls [119]. However, the same study showed that the recurrence of a microstate partially associated with the Attention RSN was reduced in schizophrenia patients [119]. This differed from the findings presented above.

One interesting finding was that the conceivably atypical leading eigenvectors belonging to the schizophrenia group did not considerably influence the detected FC states and the cluster vector assignments. In fact, using the pooled leading eigenvectors from all participants or only the leading eigenvectors from the control group as input to the K-means algorithm was found to produce similar FC states (and clusters) across the range of K. Furthermore, in both clustering approaches, significant intergroup differences in the mean fractional occupancy and in the mean dwell time were found for FC states which represented analogous connectivity patterns. These results further supported the idea that the temporal state trajectories of schizophrenia patients and healthy controls explored similar repertoires of connectivity patterns, but were characterised by properties such as the fractional occupancy, that led to the observed intergroup differences - underpinning the role of this measure as a valid tool to capture aberrant dFC in schizophrenia.

Here, the K-means clustering solution with 11 FC states was investigated in greater detail as it provided an adequate trade-off between the ability to identify significant intergroup differences and the validity of the partition of the data. Although improved clustering performance was obtained for clustering solutions with lower K, relying solely on internal clustering validation criteria was not considered suitable to detect schizophrenia-related abnormalities which could pertain to particular functional subsystems optimally defined for higher K [76]. In fact, partition models with a high number of FC states had successfully been employed in prior LEiDA analyses to investigate patients in remission from major depressive disorder [77] and the effects of psilocybin [78]. Furthermore, a large body of research had found evidence that brain FC may progress through a large number of connectivity patterns [33, 34, 49].

On the question of the validity of the optimal clustering solution, this study found that it provided satisfactory intra-cluster homogeneity (compactness) and inter-cluster heterogeneity (separability). Regarding its stability, the optimal clustering solution was considered appropriate for further investigation. Primarily, the performed 10-fold cross-validation approach allowed the entire collection of observations to be utilised as a validation sample. Additionally, indices of association, paired agreement and variation of information suggested that the agreement between partitions obtained from the validation samples

was adequate. In contrast, the validation process performed in this study was not exhaustive since other methods available in the literature were not explored. Furthermore, given the dimension and complexity of the dataset, the use of a simple Nearest Centroid classifier could have been a naive approach.

Analysis of estimates of the fractional occupancy of the optimal partition model, similarly to the remaining clustering solutions, suggested that schizophrenia patients showed decreased ability to access FC state 1 compared to healthy controls (Figure 4.7). This more frequently recurring pattern of global BOLD phase coherence [21, 76–78] has been linked to greater neural flexible switching via integration or segregation of different functional connections [21, 35]. Therefore, the Global Mode could be hypothesised to represent a cyclic equilibrium state to which the time-varying FC dynamics often evolves towards to facilitate the emergence of functionally meaningful RSNs [14, 64]. Following this rationale, the reduced recurrence of this state observed in patients with schizophrenia could potentially provoke reduced brain flexibility, i.e., the dynamical state trajectories of schizophrenia patients will, to a greater extent, stay restricted to a fixed set of "ghost" attractor states rather than flexibly transitioning through the full collection of FC states enabled by the recurrent transitions into the functionally integrative Global Mode. This hypothesis seemed to be consistent with previous research which found whole-brain integration of higher-order networks was impaired in schizophrenia [91]. Furthermore, this hypothesis was in line with previous findings indicating that schizophrenia patients exhibited diminished dynamic fluidity as they switched less frequently between connectivity patterns compared to healthy controls [84].

Another important finding was that the state trajectories of schizophrenia patients showed a significantly decreased mean probability of remaining in the Global Mode compared to healthy controls (Figure 4.10). Consequently, when in FC state 1, schizophrenia patients were significantly more likely not to remain in that state at the next time instant, in concordance with findings from previous studies [93]. Additionally, schizophrenia patients displayed a significantly increased mean probability of transitioning from the Global Mode to connectivity patterns 5 and 10 (Figure 4.10). Notably, the mean fractional occupancy of these FC states was found to be significantly increased in the schizophrenia group compared to the control group. It can therefore be suggested that the decreased fractional occupancy of FC state 1 resulted from the increased propensity of schizophrenia patients to rapidly transition to networks related to the Somatomotor (state 5) and Limbic (state 10) RSNs. This could have in turn led to the observed increased fractional occupancy (and limiting probability) of these states. These results provide further support for the hypothesis that schizophrenia patients present reduced brain integration of segregated functional connections.

Contrary to expectations, this study did not find an intergroup difference in the mean limiting probability of the Global Mode - suggesting that, in the long-run, the fraction of time points spent in a globally BOLD synchronised state was not different between groups (Equation (4.1)). This result raises the possibility that differences between schizophrenia patients and healthy controls could rely primarily in the

dynamical properties of the temporal state trajectories, i.e., the transitions between states, rather than in the fractional occupancy of FC states measured on restricted temporal windows. Nevertheless, it must be noted that a reduced subsample of participants was considered to assess intergroup differences in the group mean of this measure.

With respect to the selected optimal clustering solution, the results of this study showed that the mean fractional occupancy of the "ghost" attractor states 5, 9 and 10 was significantly increased in schizophrenia patients compared to healthy controls (Figure 4.7). Consistent with the aforementioned findings, the mean limiting probability of connectivity patterns 5 and 10 was found to be significantly increased in schizophrenia patients compared to healthy controls (Equation (4.1)) - suggesting that the long-run fraction of time spent in states 5 and 10 was significantly increased in schizophrenia. One unanticipated finding was that the long-run fraction of time spent in FC state 9 did not differ between groups (Equation (4.1)). In detail, connectivity pattern 5 displayed BOLD phase coherence among the somatosensory cortex, parietal gyri and supplementary motor area (Figure 4.6(b)) - suggesting a possible association with primary sensory and motor areas relevant in top-down brain processing. State 9 consisted of an extensive network including ROIs from the latter state, the precuneus and the occipital gyri with potential implications to top-down brain processing, memory, self-consciousness and perception (Figure 4.6(b)). Lastly, FC state 10 included BOLD synchrony among areas from the temporal lobe, namely the hippocampus and amygdalas - displaying a pattern likely involved in memory, decision making and emotional response (Figure 4.6(b)). It is possible to hypothesise that the imbalanced recurrence of the mentioned connectivity patterns in schizophrenia is in line with the view of schizophrenia as a disorder affecting memory, attention and executive function [120]. It must be noted that a thorough analysis of the functional role of the FC patterns was not straightforward and was out of the scope of this project.

Overall, this study found that the mean probabilities of a number of state-to-state transitions were significantly altered in schizophrenia patients (Figure 4.10). One interesting finding was that, compared to healthy controls, schizophrenia patients presented a significantly increased mean probability of transitioning from a state sharing spatial similarities with the reference Dorsal Attention RSN (FC state 9) to a state sharing spatial similarities with the reference Limbic RSN (FC state 10). This finding was consistent with that of EEG studies which reported unexpectedly more transitions from microstate D (associated with the Attention RSN) to microstate C (associated with the Limbic RSN) [121]. Another important finding was the significant decrease in the mean probability of remaining in a state functionally related to the reference Default RSN observed in schizophrenia patients compared to healthy controls. This result seemed to be consistent with other fMRI studies reporting abnormal DMN dynamics [48,86,87]. Notably, this RSN has been linked to core processes of human cognition such as the integration of cognitive and emotional processing, monitoring the surroundings and mind-wandering [29, 32]. Therefore, this finding may be related to the cognitive symptoms observed in patients with schizophrenia [120]. Additional

significant intergroup differences were detected in a number of other mean state-to-state transition probabilities. It may be the case therefore that the observed abnormal dynamical state transitions provide potential biomarkers of this disorder as highlighted by studies using fMRI [93, 122] and MEG [85].

On the question of the influence of the selected clustering algorithm, this study found that employing either the K-medoids algorithm or the K-means algorithm produced similar findings. As with the K-means algorithm, the Global Mode detected using the K-medoids algorithm was found to recur significantly less in schizophrenia patients compared to healthy controls (Figure 4.11). Additionally, a significant increase in the group mean fractional occupancy of a number of "ghost" attractor states detected by the K-medoids algorithm was observed in schizophrenia patients (Figure A.4). Finally, the functional relevance of the mentioned "ghost" attractor networks (with significantly increased group mean fractional occupancy in schizophrenia patients) suggested a possible connection with the Somatomotor, Dorsal Attention and Limbic reference RSNs (Table 4.3) similarly with the findings from the K-means clustering analysis. These aforementioned findings suggested that the results from the K-medoids algorithm were qualitatively identical to those obtained using the K-means algorithm.

Surprisingly, despite the differences in the clustering solutions detected by each of the clustering algorithms, the FC patterns whose mean fractional occupancy and mean dwell time significantly differed between groups in both algorithms were found, through quantitative measures, to represent the coherence among BOLD signals from similar subsets of ROIs. This suggested that both clustering algorithms captured comparable intergroup differences and further reinforced the findings from the previous LEiDA analysis using the K-means algorithm (Section 4.1).

Importantly, although both clustering algorithms provided similar findings, the choice of an optimal clustering algorithm should rely not only on statistical and cluster validation analyses, but also on concepts and methods from dynamical systems theory [21, 64, 76]. In this study, the cluster centroids detected using K-means were considered to provide greater physiological interpretability than the cluster medoids detected using K-medoids. On the one hand, from the definition of the K-means algorithm, the detected FC states were not necessarily observations from the input dataset, but could rather be interpreted as averaged unobserved connectivity patterns to which the time-resolved state trajectories tended to transiently evolve towards and temporarily settle; hence, the designation of these FC patterns as "ghost" attractor phase locking states [64, 76]. However, the definition of the K-medoids algorithm implied the detected FC states were the most representative patterns of connectivity among all observed leading eigenvectors. Hence, under this framework, it could conceivably be hypothesised that the phase locking states of the time-varying state trajectories were not "ghost" states, but rather represented connectivity patterns which existed during the state trajectories of individuals. Research questions pertaining to the functional meaning of the detected FC patterns underline the need to employ tools from dynamical systems theory to provide further insights into the dynamical regime of brain activity [14, 64].

6

Conclusion

Contents

This final chapter reflects on the main findings and contributions of this thesis in Section 6.1. In Section 6.2 limitations of this study are discussed and recommendations for future studies are provided.

6.1 Concluding Remarks

In this thesis, resting-state dynamic functional connectivity comparisons were conducted between patients with schizophrenia and healthy controls by employing and extending the Leading Eigenvector Dynamics Analysis method. Through the characterisation of the temporal expression of different functional connectivity patterns, this analytical approach contributed with new insights into the dynamical foundations of brain activity in schizophrenia.

The investigation of intergroup differences has shown that patients with schizophrenia exhibited decreased fractional occupancy and mean duration of a strong, large-scale connectivity network representing BOLD signal phase coherence of the whole set of AAL90 brain regions. Conversely, schizophrenia patients showed increased fractional occupancy and mean duration of a number of functionally meaningful "ghost" phase locking states. One of the more significant findings to emerge from this study was that, from a dynamical perspective, the ability to remain in the Global Mode was reduced in schizophrenia patients, which was hypothesised to lead to decreased integration of segregated functional connections [21,35]. Additionally, this study found that a number of state-to-state transitions displayed abnormal dynamical behaviour in schizophrenia patients. In general, therefore, it seemed that, even in the absence of any explicit task, the dynamical behaviour of brain activity in schizophrenia was typified by more frequent transitions to network patterns that are commonly activated during specific tasks - supporting and extending the findings of previous research [84–86, 91, 93, 121].

Developing on from previous LEiDA analyses, this study proposed examining the limiting probability of the FC states which defined the state trajectories characterised by irreducible and aperiodic DTMCs. Notwithstanding the relatively reduced estimated effect sizes, this property offered valuable insights into the long-run fraction of time that the DTMCs spent in each of the FC states defined by the selected optimal state repertoire from a dynamical perspective. Furthermore, it highlighted the relevance of transition probabilities, which capture the underlying dynamical behaviour of brain activity, as a potential biomarker of schizophrenia [122]. In fact, this measure is based on the state-to-state transition probabilities. This study laid the groundwork for future research into this property of state trajectories. However, considerably more work will need to be conducted to determine its utility, which would greatly benefit from fMRI scanning sessions with high temporal resolution to capture more rapid state-to-state transitions.

In addition to investigating differences between schizophrenia patients and healthy controls, this study set out to examine the validity of the results outputted by the clustering procedure to a greater depth than in previous research employing LEiDA. Specifically, evaluating the performance of the K-

means clustering algorithm aided to guide the choice of an optimal clustering solution - guaranteeing the selected number of FC states facilitated the identification of intergroup differences, whilst reflecting a partition of the dataset with suitable compactness and separation properties. Importantly, this was the first LEiDA analysis to study the stability of the selected optimal clustering solution. The results of this investigation showed that the stability of its partition of the data was adequate. Although the stability analysis was not comprehensive, this finding reinforced that the estimated patterns of connectivity were accurate average representations of the dynamical fluctuations of brain activity. A natural progression of this investigation is to analyse clustering stability using other methods proposed in the literature.

Finally, this study aimed to assess the influence of using the K-medoids clustering algorithm to investigate dFC differences between schizophrenia patients and healthy controls. This study found that generally similar conclusions would have been drawn from conducting a K-medoids clustering analysis. This emphasised the need to interpret the results through the lens of dynamical systems theory to decide on the appropriate clustering analysis method [21, 64, 76].

6.2 Limitations & Future Work

The heterogeneity of characteristics of schizophrenia patients in research poses difficulties to the identification of reliable biomarkers of this disease [120]. Specifically, components of heterogeneity include the often unclear characterisation of patients and their clinical history [94, 120]. Since the data of this study was limited to a public repository [20], it was not possible to control the clinical heterogeneity of the schizophrenia sample. In combination with the analytical flexibility of fMRI research [4, 12], this could have influenced the findings of this study in an uncertain fashion.

A limitation of this study was that the detected FC states were strongly constrained by the selected parcellation atlas (AAL90). Despite having shown consistent results across studies employing LEiDA [21, 76–78], the AAL90 template is based on an anatomical parcellation and, therefore, may not provide an adequate framework to conduct an analysis of dFC. Accordingly, future studies could extend this analysis to other fMRI-derived anatomical or functional parcellations.

One source of weakness in this study which could have affected the measurements of the properties of the DTMCs was the low temporal resolution of the BOLD signals and the relatively short fMRI recording time window. This was most clearly observed from the inconsistencies found across state trajectories obtained from the optimal clustering solution where, often times, the occurrence of all FC states was not guaranteed. The referred limitation was overcome by assuming all recording sessions represented different observations of the same underlying dynamical system where, although all FC states existed, some could have not been observed. Future work should utilise lengthier fMRI scanning sessions with higher temporal resolution to improve the utility of properties such as the state-to-state

transition probabilities and the limiting probability of FC states as possible biomarkers of schizophrenia.

In schizophrenia research, it is unclear whether specific cognitive problems are schizophrenia specific or may be linked to concomitant factors such as age and gender [28, 120]. A limitation of this study is that the effect of age and gender was not taken into account while assessing intergroup differences. Consequently, the findings of this study were attributed solely to the effect of the group, independently of the effect of the mentioned variables. The effect of age and gender could be investigated by considering logistic regression models where group and age/gender correspond to the independent variables and a given property corresponds to the dependent variable. Future investigations should interpret the intergroup differences in light of age (e.g., by splitting the data into age quantiles and rerunning the permutation tests within each age group) and gender (e.g., by rerunning the permutation tests within each gender). Lastly, the interaction between age and gender should also be explored to determine its effect on the intergroup differences.

The results presented in this project were obtained by considering the full frequency spectrum of the ROI-averaged BOLD time courses (see Section 3.2.2). Contrary to the requirements when working with PC-based approaches [70], it was unknown whether the BOLD signals analysed herein were narrowband. Further research is necessary to determine the influence of not temporally filtering the data on the ability to differentiate schizophrenia patients from healthy controls. This could be achieved by following the procedure of [76] and may contribute with additional valuable insights into this controversial preprocessing procedure [39, 40].

One of the more significant findings to emerge from this study was that the dynamical behaviour of a state representing global BOLD phase coherence was found to be altered in schizophrenia. Interestingly, this FC state has been suggested to possibly relate to the "global signal" [21, 77] (see Section 3.2.3). Here, the "global signal" was not regressed from the neuroimaging data. Therefore, the study should be repeated using GSR to understand whether schizophrenia patients would still be distinguishable from healthy controls based on the properties of a whole-brain BOLD phase coherence FC pattern.

In this project, motion-related artefacts were not removed from the neuroimaging data (see Section 3.2.3). Therefore, an additional uncontrolled factor was the possibility that motion parameters could have biased the estimates of dFC [100]. Further research needs to be conducted to fully understand the influence of these artefacts on the results. Furthermore, this investigation could usefully evaluate the robustness of LEiDA to motion-related confounds as this remains to be elucidated. On the one hand, this could be achieved by applying motion regression methods to the data and repeating the statistical analysis. On the other hand, following the procedure proposed to study the effect of age and gender, FD [96] could be used as an independent variable of a logistic regression model to assess the overall impact of head motion. In fact, an initial analysis showed that the hypothesis of zero-effect of head motion on the estimates of the fractional occupancy and on the estimates of the dwell time was rejected.

Following previous LEiDA studies, intergroup comparisons were performed by utilising permutation tests. In contrast to previous research [21, 77, 78], this study used permutation tests based on two different statistics, depending on the homogeneity between the group variances. A limitation of the statistical tests used to perform intergroup comparisons was that the standard deviation of the difference of the group means was not estimated independently of the calculated means difference. Therefore, this study should be repeated using a modified version of the permutation-based statistical procedure. The alternative version could explore the possibility of using bootstrap samples within each permutation sample to estimate the standard deviation of the difference of the means. A preliminary analysis showed that this modified statistical procedure produced relatively similar results to the ones presented in this study - leading to similar conclusions regarding dynamic brain activity in schizophrenia patients. A thorough and comparative analysis of both these procedures merits further investigations. Alternatively, statistical tests based on random sampling with replacement, i.e., bootstrap tests, could be considered in future investigations to assess intergroup differences. Another alternative to study intergroup differences could be to build logistic regression models with the group of participants as the dependent variable and the estimates of a given measure as the independent variable. The latter approach would enable assessing the effect of other concomitant variables in a more flexible fashion.

This research has thrown up the use of the limiting probability of FC states as a dynamic property to differentiate diseased patients from healthy subjects. In fact, the measurements of this property were derived from the estimation of the stationary distribution of the state trajectories defined as irreducible and aperiodic DTMCs. As stated in Section 3.3.4.D, this study did not examine intergroup differences in the stationary distributions. A natural progression of this work is to analyse whether this property may provide further insights into differences between the limiting dynamic behaviour of brain activity in diseased and healthy populations. This could be achieved by employing the two-sample goodness of fit χ^2 test.

With respect to the clustering algorithm, inspection of the clustering solutions detected by K-medoids suggested that the partition model with 7 clusters provided an adequate trade-off between a higher number of FC states and clusters with a relatively high compactness and separation (results not presented). The exploration of the K-medoids algorithm as a suitable alternative to the K-means algorithm, in light of the methods and concepts from dynamical systems theory, merits further examination.

Finally, this study attempted to provide unbiased and statistically rigorous evidence for differences between schizophrenia patients and healthy controls. Nevertheless, the potential to serve as biomarkers of schizophrenia and the clinical implications of the results derived herein remain to be analysed. Future research should gather a diverse panel of experts to explore how these findings could be applied to diagnose the disorder to determine its prognosis and to predict and monitor a patient's response to interventions.

Bibliography

[1] P. M. Matthews, G. D. Honey, and E. T. Bullmore, "Applications of fMRI in translational medicine and clinical practice," *Nature Reviews Neuroscience*, vol. 7, no. 9, pp. 732–744, 2006.

[2] A. Griffa, P. S. Baumann, J.-P. Thiran, and P. Hagmann, "Structural connectomics in brain diseases," *NeuroImage*, vol. 80, pp. 515–526, 2013, mapping the Connectome. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811913004035

[3] D. M. Lydon-Staley and D. S. Bassett, *Network Neuroscience: A Framework for Developing Biomarkers in Psychiatry*. Cham: Springer International Publishing, 2018, pp. 79–109. [Online]. Available: https://doi.org/10.1007/7854_2018_41

[4] E. Canario, D. Chen, and B. Biswal, "A review of resting-state fMRI and its use to examine psychiatric disorders," *Psychoradiology*, vol. 1, no. 1, pp. 42–53, 05 2021. [Online]. Available: https://doi.org/10.1093/psyrad/kkab003

[5] A. Fornito, A. Zalesky, C. Pantelis, and E. T. Bullmore, "Schizophrenia, neuroimaging and connectomics," *NeuroImage*, vol. 62, no. 4, pp. 2296–2314, 2012, connectivity. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811912002133

[6] M. P. Van Den Heuvel and A. Fornito, "Brain networks in schizophrenia," *Neuropsychology review*, vol. 24, no. 1, pp. 32–48, 2014.

[7] E. Mennigen, B. Rashid, and V. D. Calhoun, "Chapter 7 - connectivity and dysconnectivity: A brief history of functional connectivity research in schizophrenia and future directions," in *Connectomics*, ser. The Elsevier and MICCAI Society Book Series, B. C. Munsell, G. Wu, L. Bonilha, and P. J. Laurienti, Eds. Academic Press, 2019, pp. 123–154. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128138380000078

[8] D. J. Fisher and D. F. Salisbury, "The neurophysiology of schizophrenia: Current update and future directions," *International Journal of Psychophysiology*, vol. 145, pp. 1–4, 2019, the Neurophysiology of Schizophrenia: A Critical Update. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167876019304593

[9] 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, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, 2013, mapping the Connectome. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105381191300579X

[10] M. G. Preti, T. A. Bolton, and D. Van De Ville, "The dynamic functional connectome: State-of-the-art and perspectives," *NeuroImage*, vol. 160, pp. 41–54, 2017, functional Architecture of the Brain. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811916307881

[11] V. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The chronnectome: Time-varying connectivity networks as the next frontier in fmri data discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0896627314009131

[12] J. M. Soares, R. Magalhães, P. S. Moreira, A. Sousa, E. Ganz, A. Sampaio, V. Alves, P. Marques, and N. Sousa, "A hitchhiker's guide to functional magnetic resonance imaging," *Frontiers in Neuroscience*, vol. 10, p. 515, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2016.00515

[13] D. J. Lurie, D. Kessler, D. S. Bassett, R. F. Betzel, M. Breakspear, S. Kheilholz, A. Kucyi, R. Liégeois, M. A. Lindquist, A. R. McIntosh, R. A. Poldrack, J. M. Shine, W. H. Thompson, N. Z. Bielczyk, L. Douw, D. Kraft, R. L. Miller, M. Muthuraman, L. Pasquini, A. Razi, D. Vidaurre, H. Xie, and V. D. Calhoun, "Questions and controversies in the study of time-varying functional connectivity in resting fMRI," *Network Neuroscience*, vol. 4, no. 1, pp. 30–69, 02 2020. [Online]. Available: https://doi.org/10.1162/netn_a_00116

[14] J. Cabral, M. L. Kringelbach, and G. Deco, "Exploring the network dynamics underlying brain activity during rest," *Progress in Neurobiology*, vol. 114, pp. 102–131, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0301008213001457

[15] 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*, vol. 23, no. 5, pp. 351–366, 2010.

[16] W. R. Shirer, H. Jiang, C. M. Price, B. Ng, and M. D. Greicius, "Optimization of rs-fmri pre-processing for enhanced signal-noise separation, test-retest reliability, and group discrimination," *NeuroImage*, vol. 117, pp. 67–79, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105381191500395X

[17] K. Murphy, R. M. Birn, and P. A. Bandettini, “Resting-state fmri confounds and cleanup,” *NeuroImage*, vol. 80, pp. 349–359, 2013, mapping the Connectome. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811913003170

[18] T. Matsui, T. Murakami, and K. Ohki, “Neuronal Origin of the Temporal Dynamics of Spontaneous BOLD Activity Correlation,” *Cerebral Cortex*, vol. 29, no. 4, pp. 1496–1508, 03 2018. [Online]. Available: https://doi.org/10.1093/cercor/bhy045

[19] S. L. James, D. Abate, K. H. Abate, S. M. Abay, C. Abbafati, N. Abbasi, H. Abbastabar, F. Abd-Allah, J. Abdela, A. Abdelalim *et al.*, “Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017,” *The Lancet*, vol. 392, no. 10159, pp. 1789–1858, 2018.

[20] P. Bellec, “COBRE preprocessed with NIAK 0.17 - lightweight release,” November 2016. [Online]. Available: https://figshare.com/articles/dataset/COBRE_preprocessed_with_NIAK_0_17_-_lightweight_release/4197885

[21] J. Cabral, D. Vidaurre, P. Marques, R. Magalhães, P. S. Moreira, J. M. Soares, G. Deco, N. Sousa, and M. L. Kringelbach, “Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest,” *Scientific reports*, vol. 7, no. 1, pp. 1–13, 2017.

[22] M. Cloutier, M. S. Aigbogun, A. Guerin, R. Nitulescu, A. V. Ramanakumar, S. A. Kamat, M. DeLucia, R. Duffy, S. N. Legacy, and C. Henderson, “The economic burden of schizophrenia in the united states in 2013,” *The Journal of clinical psychiatry*, vol. 77, no. 6, pp. 0–0, 2016.

[23] C. Hjorthøj, A. E. Stürup, J. J. McGrath, and M. Nordentoft, “Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis,” *The Lancet Psychiatry*, vol. 4, no. 4, pp. 295–301, 2017.

[24] A. Karow, L. Wittmann, D. Schöttle, I. Schäfer, and M. Lambert, “The assessment of quality of life in clinical practice in patients with schizophrenia,” *Dialogues in clinical neuroscience*, vol. 16, no. 2, p. 185, 2014.

[25] K. Friston and C. Frith, “Schizophrenia: a disconnection syndrome?” *Clinical neuroscience (New York, N.Y.)*, vol. 3, no. 2, p. 89—97, 1995. [Online]. Available: http://europepmc.org/abstract/MED/7583624

[26] J. Cabral, E. Hugues, M. L. Kringelbach, and G. Deco, "Modeling the outcome of structural disconnection on resting-state functional connectivity," *NeuroImage*, vol. 62, no. 3, pp. 1342–1353, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811912005848

[27] M.-E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore, "Functional connectivity and brain networks in schizophrenia," *Journal of Neuroscience*, vol. 30, no. 28, pp. 9477–9487, 2010. [Online]. Available: https://www.jneurosci.org/content/30/28/9477

[28] R. A. McCutcheon, T. Reis Marques, and O. D. Howes, "Schizophrenia—An Overview," *JAMA Psychiatry*, vol. 77, no. 2, pp. 201–210, 02 2020. [Online]. Available: https://doi.org/10.1001/jamapsychiatry.2019.3360

[29] M. P. van den Heuvel and H. E. Hulshoff Pol, "Exploring the brain network: A review on resting-state fmri functional connectivity," *European Neuropsychopharmacology*, vol. 20, no. 8, pp. 519–534, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924977X10000684

[30] 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*, vol. 34, no. 4, pp. 537–541, 1995. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910340409

[31] R. A. Poldrack, J. A. Mumford, and T. E. Nichols, *Handbook of Functional MRI Data Analysis*. Cambridge University Press, 2011.

[32] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, no. 1, pp. 253–258, 2003. [Online]. Available: https://www.pnas.org/content/100/1/253

[33] 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*, vol. 103, no. 37, pp. 13 848–13 853, 2006. [Online]. Available: https://www.pnas.org/content/103/37/13848

[34] 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*, vol. 106, no. 31, pp. 13 040–13 045, 2009. [Online]. Available: https://www.pnas.org/content/106/31/13040

[35] J. S. Nomi, S. G. Vij, D. R. Dajani, R. Steimke, E. Damaraju, S. Rachakonda, V. D. Calhoun, and L. Q. Uddin, "Chronnectomic patterns and neural flexibility underlie executive function," *NeuroImage*, vol. 147, pp. 861–871, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811916305791

[36] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proceedings of the National Academy of Sciences*, vol. 87, no. 24, pp. 9868–9872, 1990. [Online]. Available: https://www.pnas.org/content/87/24/9868

[37] N. K. Logothetis and B. A. Wandell, "Interpreting the bold signal," *Annual Review of Physiology*, vol. 66, no. 1, pp. 735–769, 2004, pMID: 14977420. [Online]. Available: https://doi.org/10.1146/annurev.physiol.66.082602.092845

[38] M. G. Bright, C. R. Tench, and K. Murphy, "Potential pitfalls when denoising resting state fmri data using nuisance regression," *NeuroImage*, vol. 154, pp. 159–168, 2017, cleaning up the fMRI time series: Mitigating noise with advanced acquisition and correction strategies. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811916307480

[39] J. E. Chen and G. H. Glover, "Bold fractional contribution to resting-state functional connectivity above 0.1hz," *NeuroImage*, vol. 107, pp. 207–218, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914010040

[40] S. R. Gohel and B. B. Biswal, "Functional integration between brain regions at rest occurs in multiple-frequency bands," *Brain Connectivity*, vol. 5, no. 1, pp. 23–34, 2015, pMID: 24702246. [Online]. Available: https://doi.org/10.1089/brain.2013.0210

[41] R. Ciric, D. H. Wolf, J. D. Power, D. R. Roalf, G. L. Baum, K. Ruparel, R. T. Shinohara, M. A. Elliott, S. B. Eickhoff, C. Davatzikos, R. C. Gur, R. E. Gur, D. S. Bassett, and T. D. Satterthwaite, "Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity," *NeuroImage*, vol. 154, pp. 174–187, 2017, cleaning up the fMRI time series: Mitigating noise with advanced acquisition and correction strategies. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917302288

[42] D. M. Lydon-Staley, R. Ciric, T. D. Satterthwaite, and D. S. Bassett, "Evaluation of confound regression strategies for the mitigation of micromovement artifact in studies of dynamic resting-state functional connectivity and multilayer network modularity," *Network Neuroscience*, vol. 3, no. 2, pp. 427–454, 2019.

[43] L. Parkes, B. Fulcher, M. Yücel, and A. Fornito, "An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional mri," *NeuroImage*, vol.

171, pp. 415–436, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917310972

[44] S. Arslan, S. I. Ktena, A. Makropoulos, E. C. Robinson, D. Rueckert, and S. Parisot, "Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex," *NeuroImage*, vol. 170, pp. 5–30, 2018, segmenting the Brain. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917303026

[45] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811901909784

[46] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, February 2009. [Online]. Available: http://dx.doi.org/10.1038/nrn2575

[47] A. G. Garrity, G. D. Pearlson, K. McKiernan, D. Lloyd, K. A. Kiehl, and V. D. Calhoun, "Aberrant "default mode" functional connectivity in schizophrenia," *American Journal of Psychiatry*, vol. 164, no. 3, pp. 450–457, 2007, pMID: 17329470. [Online]. Available: https://ajp.psychiatryonline.org/doi/abs/10.1176/ajp.2007.164.3.450

[48] M. S. E. Sendi, E. Zendehrouh, C. A. Ellis, Z. Liang, Z. Fu, D. H. Mathalon, J. M. Ford, A. Preda, T. G. M. van Erp, R. L. Miller, G. D. Pearlson, J. A. Turner, and V. D. Calhoun, "Aberrant dynamic functional connectivity of default mode network in schizophrenia and links to symptom severity," *Frontiers in Neural Circuits*, vol. 15, p. 17, 2021. [Online]. Available: https://www.frontiersin.org/article/10.3389/fncir.2021.649417

[49] B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of Neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011, pMID: 21653723. [Online]. Available: https://doi.org/10.1152/jn.00338.2011

[50] Y. Du, Z. Fu, and V. D. Calhoun, "Classification and prediction of brain disorders using functional connectivity: Promising but challenging," *Frontiers in Neuroscience*, vol. 12, p. 525, 2018. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2018.00525

[51] C. Chang and G. H. Glover, “Time–frequency dynamics of resting-state brain connectivity measured with fmri,” *NeuroImage*, vol. 50, no. 1, pp. 81–98, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811909012981

[52] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010, computational Models of the Brain. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105381190901074X

[53] H. Lv, Z. Wang, E. Tong, L. Williams, G. Zaharchuk, M. Zeineh, A. Goldstein-Piekarski, T. Ball, C. Liao, and M. Wintermark, “Resting-state functional mri: Everything that nonexperts have always wanted to know,” *American Journal of Neuroradiology*, vol. 39, no. 8, pp. 1390–1399, 2018. [Online]. Available: http://www.ajnr.org/content/39/8/1390

[54] Y. Zang, T. Jiang, Y. Lu, Y. He, and L. Tian, “Regional homogeneity approach to fmri data analysis,” *NeuroImage*, vol. 22, no. 1, pp. 394–400, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811904000035

[55] Z. Yu-Feng, H. Yong, Z. Chao-Zhe, C. Qing-Jiu, S. Man-Qiu, L. Meng, T. Li-Xia, J. Tian-Zi, and W. Yu-Feng, “Altered baseline brain activity in children with adhd revealed by resting-state functional mri,” *Brain and Development*, vol. 29, no. 2, pp. 83–91, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0387760406001549

[56] Q.-H. Zou, C.-Z. Zhu, Y. Yang, X.-N. Zuo, X.-Y. Long, Q.-J. Cao, Y.-F. Wang, and Y.-F. Zang, “An improved approach to detection of amplitude of low-frequency fluctuation (alff) for resting-state fmri: Fractional alff,” *Journal of Neuroscience Methods*, vol. 172, no. 1, pp. 137–141, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027008002458

[57] V. Calhoun, T. Adali, G. Pearlson, and J. Pekar, “A method for making group inferences from functional mri data using independent component analysis,” *Human Brain Mapping*, vol. 14, no. 3, pp. 140–151, 2001. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.1048

[58] V. Calhoun, J. Sui, K. Kiehl, J. Turner, E. Allen, and G. Pearlson, “Exploring the psychosis functional connectome: Aberrant intrinsic networks in schizophrenia and bipolar disorder,” *Frontiers in Psychiatry*, vol. 2, p. 75, 2012. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyt.2011.00075

[59] Y. Du, S. L. Fryer, D. Lin, J. Sui, Q. Yu, J. Chen, B. Stuart, R. L. Loewy, V. D. Calhoun, and D. H. Mathalon, “Identifying functional network changing patterns in individuals at clinical high-risk for psychosis and patients with early illness schizophrenia: A group ica study,” *NeuroImage: Clinical*, vol. 17, pp. 335–346, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158217302590

[60] B. B. Biswal, M. Mennes, X.-N. Zuo, S. Gohel, C. Kelly, S. M. Smith, C. F. Beckmann, J. S. Adelstein, R. L. Buckner, S. Colcombe, A.-M. Dogonowski, M. Ernst, D. Fair, M. Hampson, M. J. Hoptman, J. S. Hyde, V. J. Kiviniemi, R. Kötter, S.-J. Li, C.-P. Lin, M. J. Lowe, C. Mackay, D. J. Madden, K. H. Madsen, D. S. Margulies, H. S. Mayberg, K. McMahon, C. S. Monk, S. H. Mostofsky, B. J. Nagel, J. J. Pekar, S. J. Peltier, S. E. Petersen, V. Riedl, S. A. R. B. Rombouts, B. Rypma, B. L. Schlaggar, S. Schmidt, R. D. Seidler, G. J. Siegle, C. Sorg, G.-J. Teng, J. Veijola, A. Villringer, M. Walter, L. Wang, X.-C. Weng, S. Whitfield-Gabrieli, P. Williamson, C. Windischberger, Y.-F. Zang, H.-Y. Zhang, F. X. Castellanos, and M. P. Milham, "Toward discovery science of human brain function," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, pp. 4734–4739, 2010. [Online]. Available: https://www.pnas.org/content/107/10/4734

[61] J. Cabral, M. Kringelbach, and G. Deco, "Functional graph alterations in schizophrenia: a result from a global anatomic decoupling?" *Pharmacopsychiatry*, vol. 45 Suppl 1, p. S57—64, May 2012. [Online]. Available: https://doi.org/10.1055/s-0032-1309001

[62] B. Rashid, M. R. Arbabshirani, E. Damaraju, M. S. Cetin, R. Miller, G. D. Pearlson, and V. D. Calhoun, "Classification of schizophrenia and bipolar patients using static and dynamic resting-state fmri brain connectivity," *NeuroImage*, vol. 134, pp. 645–657, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105381191630091X

[63] G. Deco, V. K. Jirsa, and A. R. McIntosh, "Emerging concepts for the dynamical organization of resting-state activity in the brain," *Nature Reviews Neuroscience*, vol. 12, no. 1, pp. 43–56, 2011.

[64] G. Deco and V. K. Jirsa, "Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors," *Journal of Neuroscience*, vol. 32, no. 10, pp. 3366–3375, 2012. [Online]. Available: https://www.jneurosci.org/content/32/10/3366

[65] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking Whole-Brain Connectivity Dynamics in the Resting State," *Cerebral Cortex*, vol. 24, no. 3, pp. 663–676, 11 2012. [Online]. Available: https://doi.org/10.1093/cercor/bhs352

[66] E. C. Hansen, D. Battaglia, A. Spiegler, G. Deco, and V. K. Jirsa, "Functional connectivity dynamics: Modeling the switching behavior of the resting state," *NeuroImage*, vol. 105, pp. 525–535, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914009033

[67] J. Cabral, M. L. Kringelbach, and G. Deco, "Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms,"

NeuroImage, vol. 160, pp. 84–96, 2017, functional Architecture of the Brain. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917302537

[68] N. Leonardi and D. Van De Ville, "On spurious and real fluctuations of dynamic functional connectivity during rest," *NeuroImage*, vol. 104, pp. 430–436, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914007496

[69] A. Ponce-Alvarez, G. Deco, P. Hagmann, G. L. Romani, D. Mantini, and M. Corbetta, "Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity," *PLOS Computational Biology*, vol. 11, no. 2, pp. 1–23, 02 2015. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1004100

[70] E. Glerean, J. Salmi, J. M. Lahnakoski, I. P. Jääskeläinen, and M. Sams, "Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity," *Brain Connectivity*, vol. 2, no. 2, pp. 91–101, 2012, pMID: 22559794. [Online]. Available: https://doi.org/10.1089/brain.2011.0068

[71] G. Deco and M. L. Kringelbach, "Metastability and coherence: Extending the communication through coherence hypothesis using a whole-brain computational perspective," *Trends in Neurosciences*, vol. 39, no. 3, pp. 125–135, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166223616000023

[72] G. Deco, J. Cabral, M. W. Woolrich, A. B. Stevner, T. J. van Hartevelt, and M. L. Kringelbach, "Single or multiple frequency generators in on-going brain activity: A mechanistic whole-brain model of empirical meg data," *NeuroImage*, vol. 152, pp. 538–550, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S105381191730232X

[73] A. M. Bastos and J.-M. Schoffelen, "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls," *Frontiers in Systems Neuroscience*, vol. 9, p. 175, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnsys.2015.00175

[74] M. Feldman, *Hilbert transform applications in mechanical vibration*. John Wiley & Sons, Ltd, 2011. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119991656.ch2

[75] C. C. Aggarwal and C. K. Reddy, *Data Clustering: Algorithms and Applications*, 1st ed. Chapman & Hall/CRC, 2013.

[76] J. Vohryzek, G. Deco, B. Cessac, M. L. Kringelbach, and J. Cabral, "Ghost attractors in spontaneous brain activity: Recurrent excursions into functionally-relevant bold phase-locking states," *Frontiers in Systems Neuroscience*, vol. 14, p. 20, 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnsys.2020.00020

[77] C. A. Figueroa, J. Cabral, R. J. T. Mocking, K. M. Rapuano, T. J. van Hartevelt, G. Deco, P. Expert, A. H. Schene, M. L. Kringelbach, and H. G. Ruhé, "Altered ability to access a clinically relevant control network in patients remitted from major depressive disorder," *Human Brain Mapping*, vol. 40, no. 9, pp. 2771–2786, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.24559

[78] L.-D. Lord, P. Expert, S. Atasoy, L. Roseman, K. Rapuano, R. Lambiotte, D. J. Nutt, G. Deco, R. L. Carhart-Harris, M. L. Kringelbach, and J. Cabral, "Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin," *NeuroImage*, vol. 199, pp. 127–142, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811919304525

[79] E. Damaraju, E. Allen, A. Belger, J. Ford, S. McEwen, D. Mathalon, B. Mueller, G. Pearlson, S. Potkin, A. Preda, J. Turner, J. Vaidya, T. van Erp, and V. Calhoun, "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia," *NeuroImage: Clinical*, vol. 5, pp. 298–308, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213158214000953

[80] T. E. Nichols and A. P. Holmes, "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Human Brain Mapping*, vol. 15, no. 1, pp. 1–25, 2002. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.1058

[81] F. Pesarin and L. Salmaso, *Permutation tests for complex data: theory, applications and software*. John Wiley & Sons, 2010.

[82] M. Marozzi, "Some remarks about the number of permutations one should consider to perform a permutation test," *Statistica*, vol. 64, no. 1, p. 193–201, Jan. 2004. [Online]. Available: https://rivista-statistica.unibo.it/article/view/32

[83] 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*, vol. 8, p. 897, 2014. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnhum.2014.00897

[84] R. L. Miller, M. Yaesoubi, J. A. Turner, D. Mathalon, A. Preda, G. Pearlson, T. Adali, and V. D. Calhoun, "Higher dimensional meta-state analysis reveals reduced resting fmri connectivity dynamism in schizophrenia patients," *PLOS ONE*, vol. 11, no. 3, pp. 1–24, 03 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0149849

[85] L. Sanfratello, J. M. Houck, and V. D. Calhoun, "Dynamic functional network connectivity in schizophrenia with magnetoencephalography and functional magnetic resonance imaging: Do

different timescales tell a different story?" *Brain Connectivity*, vol. 9, no. 3, pp. 251–262, 2019, pMID: 30632385. [Online]. Available: https://doi.org/10.1089/brain.2018.0608

[86] S. Ma, V. D. Calhoun, R. Phlypo, and T. Adalı, "Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis," *NeuroImage*, vol. 90, pp. 196–206, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914000093

[87] Y. Du, G. D. Pearlson, Q. Yu, H. He, D. Lin, J. Sui, L. Wu, and V. D. Calhoun, "Interaction among subsystems within default mode network diminished in schizophrenia patients: A dynamic connectivity approach," *Schizophrenia Research*, vol. 170, no. 1, pp. 55–65, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920996415300645

[88] M. S. Sendi, G. D. Pearlson, D. H. Mathalon, J. M. Ford, A. Preda, T. G. van Erp, and V. D. Calhoun, "Multiple overlapping dynamic patterns of the visual sensory network in schizophrenia," *Schizophrenia Research*, vol. 228, pp. 103–111, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0920996420306423

[89] J. Su, H. Shen, L.-L. Zeng, J. Qin, Z. Liu, and D. Hu, "Heredity characteristics of schizophrenia shown by dynamic functional connectivity analysis of resting-state functional mri scans of unaffected siblings," *Neuroreport*, vol. 27, no. 11, pp. 843–848, 2016.

[90] Y. Du, S. L. Fryer, Z. Fu, D. Lin, J. Sui, J. Chen, E. Damaraju, E. Mennigen, B. Stuart, R. L. Loewy, D. H. Mathalon, and V. D. Calhoun, "Dynamic functional connectivity impairments in early schizophrenia and clinical high-risk for psychosis," *NeuroImage*, vol. 180, pp. 632–645, 2018, brain Connectivity Dynamics. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811917308492

[91] D. Dong, M. Duan, Y. Wang, X. Zhang, X. Jia, Y. Li, F. Xin, D. Yao, and C. Luo, "Reconfiguration of Dynamic Functional Connectivity in Sensory and Perceptual System in Schizophrenia," *Cerebral Cortex*, vol. 29, no. 8, pp. 3577–3589, 10 2018. [Online]. Available: https://doi.org/10.1093/cercor/bhy232

[92] Y. Du, G. D. Pearlson, D. Lin, J. Sui, J. Chen, M. Salman, C. A. Tamminga, E. I. Ivleva, J. A. Sweeney, M. S. Keshavan, B. A. Clementz, J. Bustillo, and V. D. Calhoun, "Identifying dynamic functional connectivity biomarkers using gig-ica: Application to schizophrenia, schizoaffective disorder, and psychotic bipolar disorder," *Human Brain Mapping*, vol. 38, no. 5, pp. 2683–2708, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23553

[93] L. Rabany, S. Brocke, V. D. Calhoun, B. Pittman, S. Corbera, B. E. Wexler, M. D. Bell, K. Pelphrey, G. D. Pearlson, and M. Assaf, "Dynamic functional connectivity in schizophrenia and

autism spectrum disorder: Convergence, divergence and classification," *NeuroImage: Clinical*, vol. 24, p. 101966, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221315821930316X

[94] K. Lottman, N. Kraguljac, D. White, C. Morgan, V. Calhoun, A. Butt, and A. Lahti, "Risperidone effects on brain dynamic connectivity– a prospective resting state fmri study in schizophrenia," *Frontiers in Psychiatry*, vol. 8, p. 14, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyt.2017.00014

[95] MATLAB, *MATLAB version 9.7.0.1190202 (R2019b))*. Natick, Massachusetts: The MathWorks Inc., 2019.

[96] 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*, vol. 59, no. 3, pp. 2142–2154, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811911011815

[97] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: https://www.R-project.org/

[98] A. Geissler, R. Lanzenberger, M. Barth, A. R. Tahamtan, D. Milakara, A. Gartus, and R. Beisteiner, "Influence of fmri smoothing procedures on replicability of fine scale motor localization," *NeuroImage*, vol. 24, no. 2, pp. 323–331, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811904005051

[99] A. M. Triana, E. Glerean, J. Saramäki, and O. Korhonen, "Effects of spatial smoothing on group-level differences in functional brain networks," *Network Neuroscience*, vol. 4, no. 3, pp. 556–574, 07 2020. [Online]. Available: https://doi.org/10.1162/netn_a_00132

[100] T. O. Laumann, A. Z. Snyder, A. Mitra, E. M. Gordon, C. Gratton, B. Adeyemo, A. W. Gilmore, S. M. Nelson, J. J. Berg, D. J. Greene, J. E. McCarthy, E. Tagliazucchi, H. Laufs, B. L. Schlaggar, N. U. F. Dosenbach, and S. E. Petersen, "On the Stability of BOLD fMRI Correlations," *Cerebral Cortex*, vol. 27, no. 10, pp. 4719–4732, 09 2016. [Online]. Available: https://doi.org/10.1093/cercor/bhw265

[101] K. Murphy and M. D. Fox, "Towards a consensus regarding global signal regression for resting state functional connectivity mri," *NeuroImage*, vol. 154, pp. 169–173, 2017, cleaning up the fMRI time series: Mitigating noise with advanced acquisition and correction strategies. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811916306711

[102] A. Nalci, B. D. Rao, and T. T. Liu, "Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fmri," *NeuroImage*, vol. 184, pp. 1005–1031, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811918308097

[103] M. G. Bright and K. Murphy, "Is fmri "noise" really noise? resting state nuisance regressors remove variance with network structure," *NeuroImage*, vol. 114, pp. 158–169, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811915002669

[104] 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*, vol. 44, no. 3, pp. 893–905, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811908010264

[105] M. N. Hallquist, K. Hwang, and B. Luna, "The nuisance of nuisance regression: Spectral misspecification in a common approach to resting-state fmri preprocessing reintroduces noise and obscures functional connectivity," *NeuroImage*, vol. 82, pp. 208–225, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811913006265

[106] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E*, vol. 74, p. 036104, Sep 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.74.036104

[107] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07. USA: Society for Industrial and Applied Mathematics, 2007, p. 1027–1035.

[108] V. G. Kulkarni, *Introduction to Modeling and Analysis of Stochastic Systems*, 2nd ed., ser. Springer Texts in Statistics. Springer-Verlag New York, 2011.

[109] A. M. Winkler, G. R. Ridgway, M. A. Webster, S. M. Smith, and T. E. Nichols, "Permutation inference for the general linear model," *NeuroImage*, vol. 92, pp. 381–397, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053811914000913

[110] R. Durrett, *Probability: Theory and Examples*, 5th ed., ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[111] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0377042787901257

[112] J. C. Dunn†, "Well-separated clusters and optimal fuzzy partitions," *Journal of Cybernetics*, vol. 4, no. 1, pp. 95–104, 1974. [Online]. Available: https://doi.org/10.1080/01969727408546059

[113] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, “Understanding of internal clustering validation measures,” in *2010 IEEE International Conference on Data Mining*, 2010, pp. 911–916.

[114] L. Hubert and P. Arabie, “Comparing partitions,” *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985.

[115] M. Meilă, “Comparing clusterings—an information based distance,” *Journal of Multivariate Analysis*, vol. 98, no. 5, pp. 873–895, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0047259X06002016

[116] R. M. McIntyre and R. K. Blashfield, “A nearest-centroid technique for evaluating the minimum-variance clustering procedure,” *Multivariate Behavioral Research*, vol. 15, no. 2, pp. 225–238, 1980. [Online]. Available: https://doi.org/10.1207/s15327906mbr1502_7

[117] M. C. Martins and M. G. M. S. Cardoso, “Evaluation of clusters of credit cards holders,” *Revista de Ciências da Computação*, vol. 3, no. 3, pp. 1–11, 2008. [Online]. Available: http://inqueritos.lead.uab.pt/OJS/index.php/RCC/

[118] A. P. Baker, M. J. Brookes, I. A. Rezek, S. M. Smith, T. Behrens, P. J. Probert Smith, and M. Woolrich, “Fast transient networks in spontaneous human brain activity,” *eLife*, vol. 3, p. e01867, mar 2014. [Online]. Available: https://doi.org/10.7554/eLife.01867

[119] J. N. Ramos da Cruz, O. Favrod, M. Roinishvili, E. Chkonia, A. Brand, C. Mohr, P. Figueiredo, and M. Herzog, “Eeg microstates are a candidate endophenotype for schizophrenia,” *Nature Communications*, vol. 11, no. 3089, pp. 1–11, 2020. [Online]. Available: http://infoscience.epfl.ch/record/278140

[120] M. Fioravanti, V. Bianchi, and M. E. Cinti, “Cognitive deficits in schizophrenia: An updated meta-analysis of the scientific evidence,” *BMC Psychiatry*, vol. 12, 2012.

[121] K. Rieger, L. Diaz Hernandez, A. Baenninger, and T. Koenig, “15 years of microstate research in schizophrenia – where are we? a meta-analysis,” *Frontiers in Psychiatry*, vol. 7, p. 22, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyt.2016.00022

[122] M. S. E. Sendi, E. Zendehrouh, Z. Fu, B. Mahmoudi, R. L. Miller, and V. D. Calhoun, “A machine learning model for exploring aberrant functional network connectivity transition in schizophrenia,” in *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, 2020, pp. 112–115.

[123] R. A. Johnson, D. W. Wichern *et al.*, *Applied multivariate statistical analysis*. Pearson London, UK:, 2014, vol. 6.

Supplementary Material

A.1 Supplementary Material of Section 2.6

Table A.1: State of the art of resting-state dynamic functional connectivity in schizophrenia.

Study	Main Objective	Sample	Methods	Main Findings
[86]	Examine dynamic connectivity changes among pairs of spatial components estimated from rs-fMRI data.	10 SZ 10 HC	Markov modelling analysis: (1) Sliding-window analysis to segment data; (2) IVA to extract spatial components; (3) K-means clustering ($K = 3$).	Overall, SZ patients presented higher FC values than HCs. Significant spatial changes in network connectivity in the SZ group mainly occurred between the frontoparietal, cerebellar and temporal lobe regions. SZ patients presented more FC dynamics.
[79]	Assess whole brain FC differences between HCs and SZ patients and evaluate if any observed differences primarily occurred in certain connectivity states during the scan duration.	151 SZ 163 HC	(1) Group spatial ICA to decompose imaging data; (2) dFC via SWA; (3) K-means clustering ($K = 3$).	On average, SZ patients spent much less time than HCs in states characterised by strong, large-connectivity and abnormal connectivity patterns were more pronounced during these connectivity states. SZ patients showed hypoconnectivity within sensory regions, hyperconnectivity of the thalamus with these regions.
[83]	Provide a whole-brain characterisation of regional differences in FC variability and distinction of discrete FC states among HCs, SZ and BP patients.	60 SZ 38 BP 61 HC	(1) Group spatial ICA to decompose imaging data; (2) dFC via SWA; (3) K-means clustering ($K = 5$).	Patients made fewer transitions to some states compared to HCs, with most such differences confined to a single state. SZ patients showed more differences from HCs than did BP patients, including both hyper- and hypoconnectivity in one common connectivity state.
[87]	Evaluate the differences in dynamic connectivity within DMN between HCs and SZ patients using rs-fMRI.	82 SZ 82 HC	dFC within DMN. (1) Spatial ICA to decompose imaging data; (2) dFC via SWA; (3) K-means clustering ($K = 2$).	Overall, SZ patients displayed within-network disruptions of the DMN. Connectivity measures suggested HCs showed stronger DMN inter-subsystem interaction than patients. Compared to HCs, patients spent more time in the states with nodes sparsely connected. SZ patients showed lower values than HCs in the averaged node strength, clustering coefficient, global efficiency, and local efficiency.

Continued on next page

Table A.1 – continued from previous page

Study	Main Objective	Sample	Methods	Main Findings
[84]	Present evidence that multiple explicitly dynamical properties of time-varying whole-brain network connectivity are strongly associated with schizophrenia.	151 SZ 163 HC	(1) Group ICA to decompose imaging data; (2) dFC via SWA; (3) Parametrised vector with distance between each windowed FC matrix to each cluster centroid.	Time-varying whole-brain network connectivity patterns are found to be less dynamically active in SZ patients. Compared to HCs, SZ patients exhibited diminished dynamic fluidity: (1) SZ patients occupied fewer meta-states and (2) SZ patients switched less frequently between meta-states.
[62]	Classification of SZ, BP and healthy subjects based on their static and dynamic FC features.	60 SZ 38 BP 61 HC	(1) Group spatial ICA to decompose imaging data; (2) sFC via covariance matrix; (3) dFC via SWA; (4) Linear Support Vector Machine.	Results showed that the dFC features significantly outperformed the sFC features in terms of predictive accuracy. Combining sFC and dFC features did not significantly improve the classification performance over the dFC features alone.
[89]	Explore the possibility that there are similar transient connectivity patterns between SZ patients and their unaffected siblings (SB).	25 SZ 25 SB 25 HC	(1) dFC via SWA; (2) PCA to reduce data dimensionality; (3) K-means clustering ($K = 8$).	Five FC states, with reduced connectivity strength in both sibling and patient groups, were recognised as trait-related connections (cingulo-opercular, occipital, and default mode networks) and reflected shared connectivity alterations between SZ patients and their unaffected siblings. SZ patients and their unaffected siblings shared common transient functional dysconnectivity.
[94]	Prospective study to evaluate dFC in unmedicated patients with SZ (N) and to test if, and how, antipsychotic medications change brain network dynamics after 1 (1W) and 6 (6W) weeks of treatment in an effort to disentangle medication effects form intrinsic illness characteristics.	SZ: 33 N 29 1W 24 6W 35 HC	(1) Group spatial ICA to decompose imaging data; (2) dFC via SWA; (3) K-means clustering ($K = 3$).	Significantly increased connectivity was present only between the thalamus and somatomotor network in one state in unmedicated patients compared to HCs, but no evidence of decreased connectivity in any states was found. In comparison to controls, unmedicated patients had shorter mean dwell times and fraction of time spent in the sparsely connected state, and longer dwell times and fraction of time spent in the intermediately connected state. Antipsychotic medication appeared to normalise mean dwell times, but not fraction of time.

Continued on next page

Table A.1 – continued from previous page

Study	Main Objective	Sample	Methods	Main Findings
[92]	Explore dFC across a spectrum of symptomatically-related disorders including BP with psychosis, Schizoaffective Disorder (SAD) and SZ using the GIG-ICA method.	113 SZ 132 SAD 140 BP 238 HC	(1) dFC via SWA of whole-brain BOLD time courses; (2) GIG-ICA to estimate group-level and subject-specific connectivity states.	Progressive abnormalities from HCs to BP patients to SAD patients to SZ patients with respect to hypoconnectivities and hyperconnectivities were found. Diagnosis-related states with varied activation were present in mental disorders. Overall, results suggested SAD and SZ showed common impairments in frontal connectivities, compared to HCs and BP.
[90]	Investigate group differences among HCs, clinical high-risk for psychosis (CHR) and early illness SZ groups by applying the group information guided GIG-ICA method.	58 SZ 53 CHR 70 HC	(1) dFC via SWA of whole-brain BOLD time courses; (2) GIG-ICA to estimate group-level and subject-specific connectivity states.	Multiple functional connections in the dominant state were altered in CHR and SZ groups. The dominant connectivity state showed greater connectivity changes in the SZ group compared to the CHR group. The results suggested disease-related connectivity states occurred in CHR and SZ groups.
[91]	Identify voxel-level, region-level and network-level FC variability alterations in order to map the temporal dynamics of resting-state FC in SZ.	102 SZ 124 HC	dFC via SWA at every level.	Patients showed consistent increased FC variability in the sensory and perceptual system, including the visual, sensorimotor, attention networks, and thalamus. Findings highlighted the rudimentary role of elevated instability of information communication in sensory and perceptual system and attenuated whole-brain integration of high-order networks in SZ.
[93]	Compare multiple aspects of whole-brain resting state temporal dynamics in young adults with SZ and ASD, and HCs, providing a detailed account of convergence and divergence.	33 SZ 34 ASD 34 HC	(1) Group spatial ICA to decompose imaging data; (2) dFC via SWA; (3) K-means clustering ($K = 4$).	SZ and ASD spent an increased portion of the time in a state of weak intra-network connectivity, and a decreased portion of the time in the state of strong connectivity. When SZ patients transitioned into the state of strongest connectivity, they switched states very rapidly. SZ group presented reduced number of transitions between states. These results suggested a severe and pervasive pattern of altered temporal dynamics in SZ.

Continued on next page

Table A.1 – continued from previous page

Study	Main Objective	Sample	Methods	Main Findings
[85]	Identify overlaps and differences between the dFC of HCs and patients with SZ using different imaging modalities (fMRI and MEG).	46 SZ 45 HC	(1) PCA to reduce data dimensionality; (2) Group spatial ICA on retained principal components; (2) dFC via SWA; (3) K-means clustering ($K = 4$).	On average, dFC analysis predicted on fMRI data revealed HCs and SZ patients remained in different overall brain states for significantly different periods of time, with SZ patients spending less time in a state typified by strong, large-scale connectivity. fMRI and MEG revealed between-group FC differences in distinct ways, highlighting the utility of using each of the modalities individually, or potentially a combination of modalities.

A.2 Supplementary Material of Section 3.3.2

This section aimed to confirm the robustness of the approach used by LEiDA when analysing rs-fMRI data from [20]. The derivations presented below followed the methodology of [78].

Let B denote the phase coherence matrix at time t ($B(t)$ is equivalent to $dFC(t)$ used in the main text, and was simply used here to improve the readability of the equations). The phase coherence between brain areas n and p at time t is written as:

$$B_{n,p}(t) = cos(\theta_n(t) - \theta_p(t)) \tag{A.1}$$

The phase coherence matrix B is real-valued and symmetric. Then, from the spectral theorem it follows that there exists an eigenbasis $\{v_i\}_{i=1}^N$ with associated eigenvalues $\{\lambda_i\}_{i=1}^N$, such that the spectral decomposition of B at time t is given by Equation (A.2):

$$B(t) = Q\Lambda Q^\intercal = \sum_{i=1}^{N} \lambda_i v_i v_i^\intercal = \begin{bmatrix} | & & | \\ v_1 & \dots & v_N \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \begin{bmatrix} — & v_1^\intercal & — \\ & \vdots & \\ — & v_N^\intercal & — \end{bmatrix} \tag{A.2}$$

where N is the number of brain areas, Λ is the diagonal matrix of eigenvalues and the columns of Q are the eigenvectors of B. Here it was assumed that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Using PCA it is possible to quantify the variance explained by each eigenvector of B, and thus the proportion of the signal explained by the leading eigenvector.

The j^{th} column of B represents the pairwise similarities between the phase of ROI j and all the other

ROIs. For a given column j, the centered phases were obtained by subtracting the mean of that column from all its entries. The matrix of the centered similarities was denoted with a subscript c:

$$B_c = B - \mathbf{1}B_m^\intercal \tag{A.3}$$

where B_m is the vector of the mean of the columns of B. The sample variance-covariance matrix at time t of the centered similarities between each pair of phases is approximated by $S_c = B_c^\intercal B_c$. This matrix is symmetric since $S_c^\intercal = (B_c^\intercal B_c)^\intercal = B_c^\intercal (B_c^\intercal)^\intercal = B_c^\intercal B_c = S_c$. Consequently, for each time t, the spectral theorem holds and the spectral decomposition of S_c is written in Equation (A.4):

$$S_c(t) = Q_c \Lambda_c Q_c^\intercal = \sum_{i=1}^{N} \lambda_{c,i} v_{c,i} v_{c,i}^\intercal \tag{A.4}$$

where Λ_c is the diagonal matrix of eigenvalues and the columns of Q_c are the eigenvectors of S_c so that $Q_c Q_c^\intercal = Q_c^\intercal Q_c = \boldsymbol{I}$. From the properties of the trace of matrices it follows for each time t:

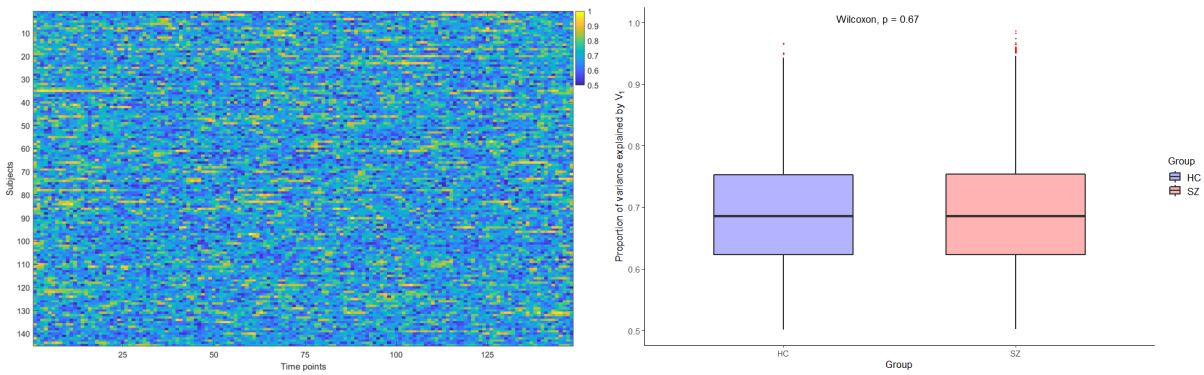
$$\mathrm{tr}(S_c) = \mathrm{tr}(Q_c \Lambda_c Q_c^\intercal) = \mathrm{tr}(\Lambda_c Q_c^\intercal Q_c) = \mathrm{tr}(\Lambda_c) = \lambda_{c,1} + \lambda_{c,2} + \ldots + \lambda_{c,N} \tag{A.5}$$

From standard PCA, the proportion of total variance in the coherence of the phases due to (explained by) the i^{th} eigenvector of the sample variance-covariance matrix is expressed as the ratio [123]:

$$\begin{pmatrix} \text{Proportion of the total sample} \\ \text{variance due to the i}^{\text{th}} \text{ eigenvector} \end{pmatrix} = \frac{\lambda_{c,i}}{\lambda_{c,1} + \lambda_{c,2} + \ldots + \lambda_{c,N}} \tag{A.6}$$

As the eigenvalues are ordered in decreasing order, the eigenvector $\lambda_{c,1}$ will explain the most variance, and if $\lambda_{c,1} > \sum_{i>1}^{N} \lambda_{c,i}$ it explains most of the variance in the similarities between phases. Following this reasoning, it was observed that the leading eigenvectors V_1 computed at each time point $t \in \{1, \ldots, T\}$ and for all subjects consistently explained more than 50% of the variance in the similarities between BOLD phases in each group, as evidenced in Figure A.1(a). To assess whether the mean proportion of variance explained by the leading eigenvectors across all time points differed between healthy controls and schizophrenia patients, a two-sided Wilcoxon Rank-Sum test with Bonferroni correction was performed, since the QQ-plots suggested non-normality and the Levene's test heteroscedasticity of the data. This test did not detect intergroup differences in the mean proportion of variance explained by V_1 across all time points. The boxplot in Figure A.1(b) corroborates the previous result.

Therefore, as in previous work [21, 76–78], considering only the leading eigenvector of each phase coherence matrix is a robust and representative approach of the dominant FC pattern at each time point.



(a) Proportion of variance explained by the leading eigenvector at all time points for all participants.

(b) Boxplot of the proportion of variance explained by the leading eigenvector at each time point for each group. Comparison of means performed using the two-sided Wilcoxon Rank-Sum test. Red points represent outliers, according to the Tukey criterion.

Figure A.1: Proportion of phase coherence explained by the leading eigenvectors.

A.3 Supplementary Material of Section 4.1.2

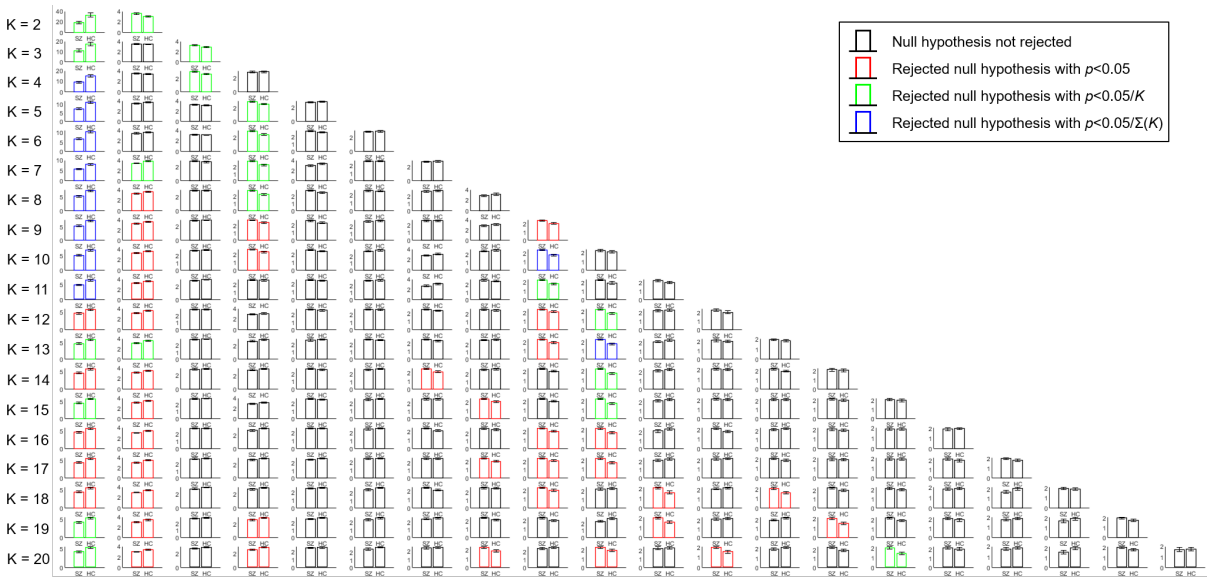
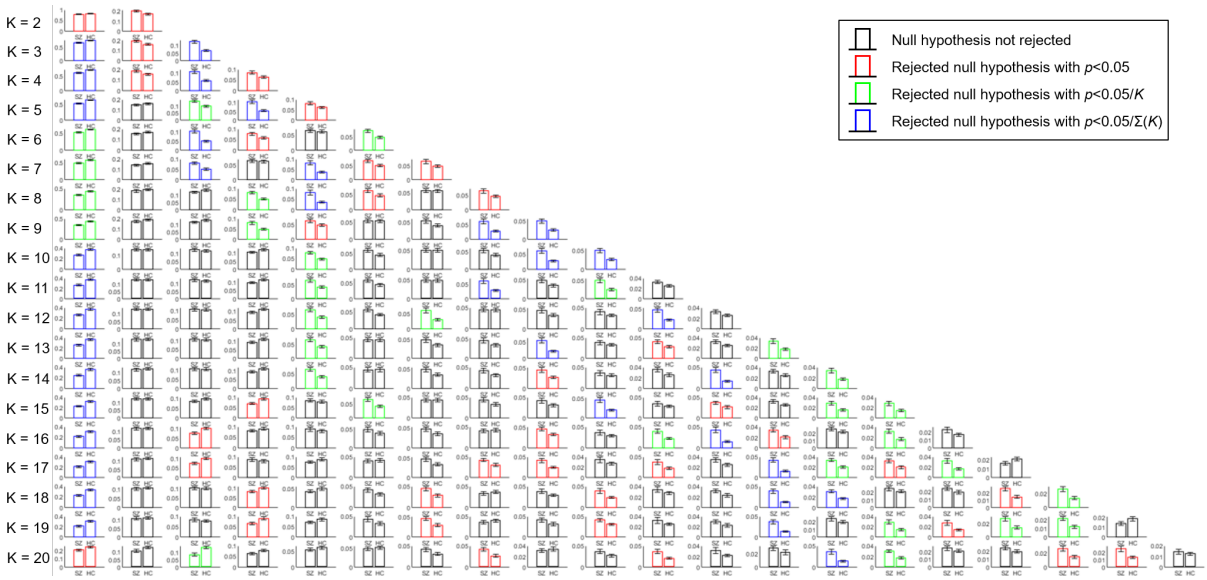
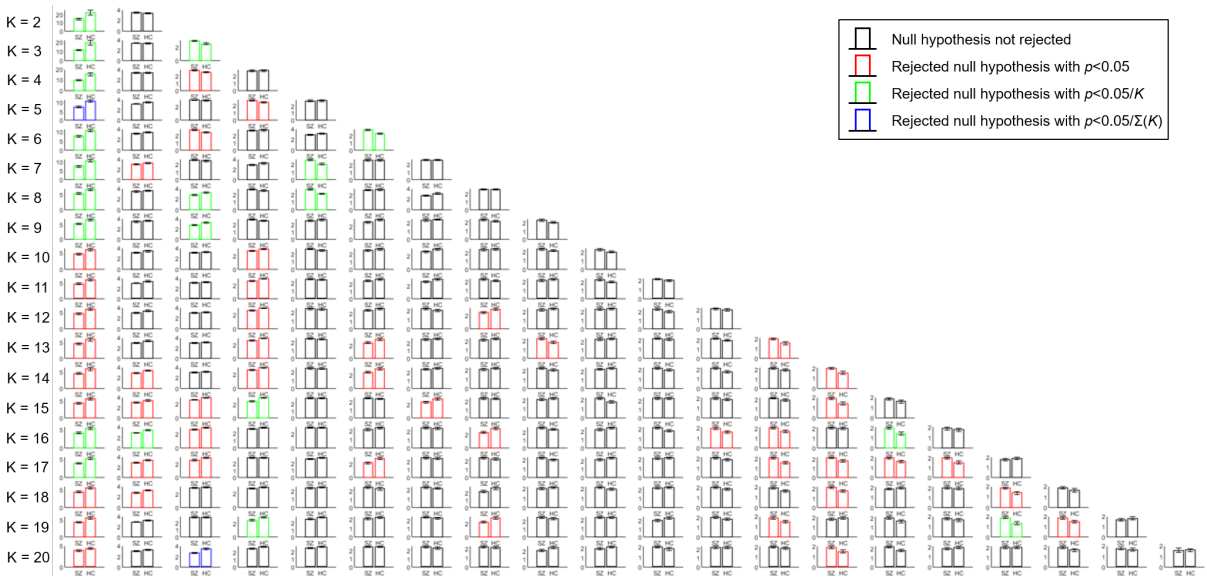


Figure A.2: Barplot of the estimated mean dwell time with associated standard error of each FC state detected by the K-means algorithm for each group. For each FC state, the colour of the bars indicates whether the null hypothesis of no intergroup differences in the mean dwell time was rejected (two-tailed tests). Black bars indicate the null hypothesis was not rejected at a 5% significance level. Red, green and blue bars indicate the null hypothesis was rejected at a 0.05, α_2 and α_3 significance level, respectively. The standard error of each bar was calculated as the standard deviation of the sample data divided by the square root of the sample size.

A.4 Supplementary Material of Section 4.2.1



(a) Probability of occurrence.



(b) Dwell Time.

Figure A.3: Barplot of the estimated mean fractional occupancy and of the estimated mean dwell time with associated standard error of each FC state detected by the K-medoids algorithm for each group. For each FC state, the colour of the bars indicates whether the null hypothesis of no intergroup differences in the mean of each measure was rejected (two-tailed tests). Black bars indicate the null hypothesis was not rejected at a 5% significance level. Red, green and blue bars indicate the null hypothesis was rejected at a 0.05, α_2 and α_3 significance level, respectively. The standard error of each bar was calculated as the standard deviation of the sample data divided by the square root of the sample size.

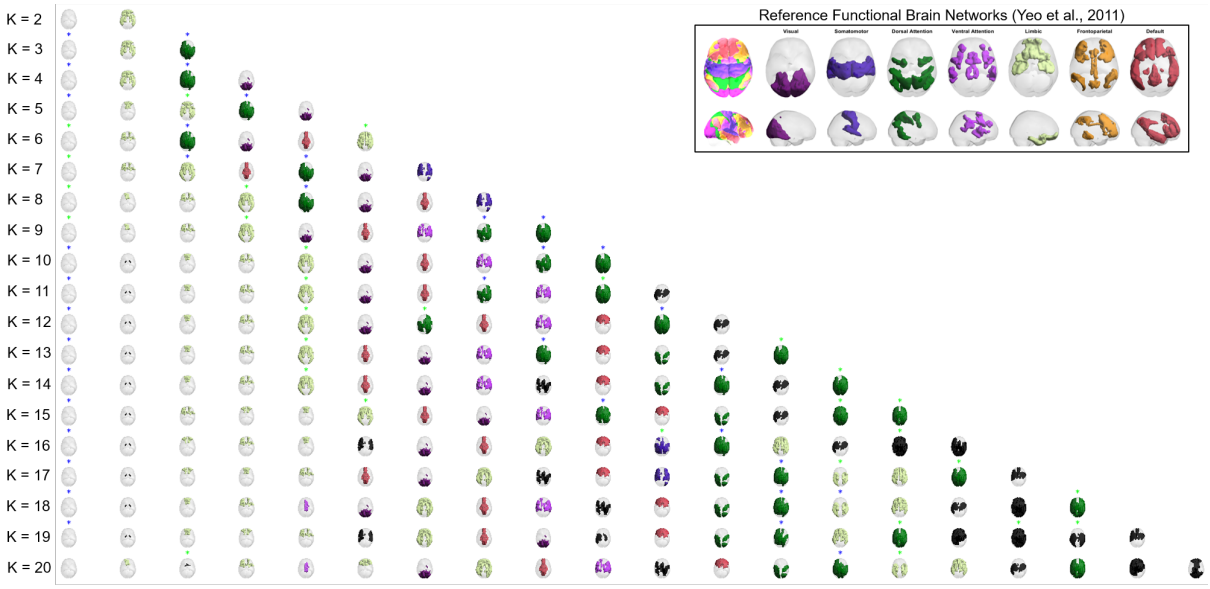


Figure A.4: Cortical space representation of the cluster medoids obtained for each clustering solution produced by the K-medoids algorithm. The rendered ROIs correspond to positive elements in the vectors of the cluster medoids. Medoids (FC states) are coloured according to the reference functional network defined by [49] whose p-value obtained from computing the Pearson's correlation coefficient was lowest (with $p < 0.05/K$). Medoids not significantly overlapping with any of the reference RSNs are coloured in black. For each FC state, title asterisks indicate whether significant intergroup differences in the mean fractional occupancy were detected. Green and blue asterisks indicate $p < \alpha_2$ and $p < \alpha_3$ (two-tailed tests), respectively.