ISTITUTO ITALIANO
DI TECNOLOGIA

UNIVERSITÀ DEGLI STUDI
DI GENOVA

DEPARTMENT OF PATTERN ANALYSIS AND COMPUTER VISION (PAVIS)

DEPARTMENT OF ELECTRICAL, ELECTRONIC AND TELECOMMUNICATIONS
ENGINEERING AND NAVAL ARCHITECTURE (DITEN)

PhD in Science and Technology for Electronic and Telecommunication Engineering

Curriculum: Computational Vision, Automatic Recognition and Learning -
Cycle XXXIII (2017-2020)

Investigating Brain Functional Networks in a Riemannian Framework

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Muhammad Abubakar Yamin

Supervisor: Prof. Vittorio Murino
Co-Supervisor: Dr. Diego Sona

Coordinator of the PhD Course: Prof. Mario Marchese

MARCH 2021

This thesis is wholeheartedly dedicated to my beloved parents, who has always been a source of inspiration and gave me strength when I thought of giving up, who continuously provide their moral, spiritual and emotional support.

Acknowledgements

Many people deserve an acknowledgment due to their help and support, professionally and personally throughout my Ph.D. degree.

First of all, I would like to thank Prof. Dr. Vittorio Murino, director of PAVIS (IIT, Genova) for giving me a chance to join this group of brilliant people and his continuous support throughout this journey. I am deeply thankful and would like to express my gratitude to my supervisor Dr. Diego Sona for his guidance and help in all moments of my Ph.D., particularly in writing, reviewing papers & this thesis, and his dedication to do so anytime. Under the shadow of his command, I was able to accomplish this milestone. I am greatly indebted to them for accepting me as their student and allowing me to grow under their guidance. I am also thankful to Dr. Alessio Del Bue for his support in this tenure.

Moreover, very special gratitude to Dr. Michael Dayan for his professional guidance in medical image analysis, owing to his rich experience and excellent expertise. I would also like to thanks all my collaborators Vaibhav Diwadkar, Maria A Rocca, Jacopo Tessadori, Paola Valsasina, Latezia, and Paolo Brambila, who helped me figuring out the basic knowledge and background related to neuroimaging.

I want to say a big thanks to my wife (Ramsha) for showing your utmost care and support throughout this duration and also to my Son and daughter (Usaym & Nooreh) for making me feel relax and comfortable with their presence. A special thanks to my father (Yamin) who always supported me throughout my journey of achieving this goal. Lots of love for all of you. I also want to thanks my sister(Qurat ul Ain) and brother (umer) for his care and support.

I would like to thanks all members and friends of PAVIS & VGM for their cooperation and suggestions, I had a wonderful time with you people. Many thanks to all who I met in IIT for your support and care, especially Shahid, Muhammad Dahy, Usman, Shafiq, Zohaib, Nabeel, and Waqar for your valuable suggestions and helpful discussions. I wish them big good luck in their career.

Abubakar Yamin
November 2020, Genova

Abstract

The brain is a complex system of several interconnected components which can be categorized at different Spatio-temporal levels, evaluate the physical connections and the corresponding functionalities. To study brain connectivity at the macroscale, Magnetic Resonance Imaging (MRI) technique in all the different modalities has been exemplified to be an important tool. In particular, functional MRI (fMRI) enables to record the brain activity either at rest or in different conditions of cognitive task and assist in mapping the functional connectivity of the brain.

The information of brain functional connectivity extracted from fMRI images can be defined using a graph representation, i.e. a mathematical object consisting of nodes, the brain regions, and edges, the link between regions. With this representation, novel insights have emerged about understanding brain connectivity and providing evidence that the brain networks are not randomly linked. Indeed, the brain network represents a small-world structure, with several different properties of segregation and integration that are accountable for specific functions and mental conditions. Moreover, network analysis enables to recognize and analyze patterns of brain functional connectivity characterizing a group of subjects.

In recent decades, many developments have been made to understand the functioning of the human brain and many issues, related to the biological and the methodological perspective, are still need to be addressed. For example, sub-modular brain organization is still under debate, since it is necessary to understand how the brain is functionally organized. At the same time a comprehensive organization of functional connectivity is mostly unknown and also the dynamical reorganization of functional connectivity is appearing as a new frontier for analyzing brain dynamics. Moreover, the recognition of functional connectivity patterns in patients affected by mental disorders is still a challenging task, making plausible the development of new tools to solve them.

Indeed, in this dissertation, we proposed novel methodological approaches to answer some of these biological and neuroscientific questions. We have investigated methods for analyzing and detecting heritability in twin's task-induced functional connectivity profiles. in this approach we are proposing a geodesic metric-based method for the estimation of similarity between functional connectivity, taking into account the manifold related properties of symmetric and positive definite matrices.

Moreover, we also proposed a computational framework for classification and discrimination of brain connectivity graphs between healthy and pathological subjects affected by mental disorder, using geodesic metric-based clustering of brain graphs on manifold space. Within the same framework, we also propose an approach based on the dictionary learning method to encode the high dimensional connectivity data into a vectorial representation which is useful for classification and determining regions of brain graphs responsible for this segregation. We also propose an effective way to analyze the dynamical functional connectivity, building a similarity representation of fMRI dynamic functional connectivity states, exploiting modular properties of graph laplacians, geodesic clustering, and manifold learning.

Keywords: Neuroimaging, brain imaging, functional connectivity, dynamics functional connectivity, connectome, graph theory, functional MRI, task-induced, resting state, schizophrenia, autism, multiple-sclerosis, laplacian, clustering, manifold, machine learning

Contents

List of Figures

List of Tables

Common Abbreviations

MRI Magnetic Resonance Images

fMRI Functional Magnetic Resonance Images

BOLD Blood Oxygen Level Dependen

T1w T1-weighted MRI

T2w T2-weighted MRI

WM White Matter

CSF Cerebrospinal Fluid

GM Gray Matter

ML Machine Learning

MEG Magnetoencephalography

EEG Electroen-cephalograph

ICA Independent Component Analysis

AAL Automated Anatomical Labelling

FC Functional Connectivity/Connectome

DFC Dynamic Functional Connectivity

SPD Symmetric Positive Definite

GL Graph Laplacian

RM Riemannian Manifold

Log-E Log Euclidean

PSD Positive Semi-Definite

MZ Mono-zygotic

DZ Di-zygotic

HCP Human Connectome Project

H_0 Heritability

MV Motor Vision

NMV Non-Motor Vision

SEM Standard Error & Mean

NoI Network of Interest

SVM Support Vector Machine

HC Healthy Controls

ASD Autism Spectrum Disorder

SCHZ Schizophrenia

DB Davies Bouldin

MS Multiple Sclerosis

RRMS) Relapsing Remitting -MS

SPMS Secondary Progressive -MS

PPMS primary Progressive -MS

PMS Progressive -MS

EDSS Expanded Disability Status Scale

RoI Region of Interest

DS Dominant Set

CV Cross Validation

LR Logistic Regression

RS fMRI Resting State fMRI

LOOCV Leave One Out -CV

SSE Sum of Squared Error

Chapter 1

Introduction

The human brain is among the most advanced and intricate network in nature which transfer signals between different regions in order to respond to an external/internal event. The anatomical structure of the brain enumerates a vast number of interlinked neurons (nervous cells), which interpret brain composition at various levels. Responsibilities of this complex system include mental processes, begin from basic sensorimotor processing to cognitive and executive functions. The conception and advance exploration of such a composite structure has drawn the attention of different fields of research, from biology, psychology and neuroscience to mathematics, computer science and engineering. Of course, medical disciplines played a vital role in this exploration and revealed most of the information regarding how this complex system of brain works and how it manages the information flow. For the purpose of analysis, this dense interchanging of information needs to be modelled and in the field of neuroscience, the idea of describing the brain as a complex network has deep roots [1].

The advancements made in the last century described the brain as a system where each region is delegated to perform a specific function. Thanks to the developments in neuro-anatomy, white matter connections became gradually more refined in the brain map, illustrating the links between different anatomical regions of the brain. Later with the advent of neurophysiological methods, it became possible to explore the cortical architecture of the brain with a precise functional meaning. Finally, now in the era of modern neuroscience, neuroimaging methods such as Magnetic Resonance Imaging (MRI) have revolutionized the way to understand and study the functional and anatomical connectivity, by providing multidimensional (2D, 3D & 4D) images of the brain with different modalities. Morphometry measures have been traditionally applied on 3D images, which allows the inspection of local differences in brain anatomy. Differences in brain volumes have been commonly highlighted through the drawing of Regions of Interest (ROIs). However, these methods can only provide a computation of rather large areas and small differences in volume may be missed.

Later in [2], Voxel-Based Morphometry (VBM) has been suggested as an add-on of ROI based approach, which focus on providing a whole-brain analysis over multiple subjects. VBM has been widely acquired to define changes in the brain volumes, which is linked to

the structural analysis of the brain. Comparably, functional analysis has been also carried out using ROI based approach, trying to map functional connections from single regions to the rest of the brain [3]. Moreover, simple data-driven or linear model-based methods have been developed to explore the activated functional brain networks [4, 5, 6]. Even though these methods have constituted and still represent an important strategy to study the brain functionality, in the last decade, thanks to the introduction of novel developments in medical imaging and to the growing interest in the network science, many major studies have been proposed focusing at understanding the brain under different characteristics, taking into account also the principle of network science.

The study of connectivity between different regions of the brain is known as "connectomics", which is a relatively recent field of research that enable neuroscientists to explore the interplay between different regions of the brain by modelling it as a network or "connectome" [7]. The term connectome recognizes a wide range of distinct connectivities, which are mostly depending on the brain scale: the microscale expresses single neurons connectivity; the mesoscale express the neuronal population connectivity; the macroscale describing the connectivity at brain regions level. Concerning the connectome, there are different characteristics that should be taken into consideration [8]. This term is mostly used to define structural connectivity which describes the *physical* connections between neural elements, providing an anatomical interpretation of the brain. Besides the structural connectome, the functional connectome is referred to as *functional* connectivity between brain regions, which is mainly considered as an undirected statistical estimate between neural elements that change over time.

Indeed, the connectome is a way to describe brain connectivity and to compress the quantity of brain data in an easy framework to emphasize the main goals. This way of describing is independent of the scale, and the overall examination of the brain at different levels should aim at structuring a full map of neural connections. Finally, the connectome is an illustration of brain architecture but it is also a mathematical object that fit well with complex network theory, focusing at exploration neuroscience process. As a matter of fact, graph theory-based approaches have been successfully applied to analyze the connectome's, bringing awareness in studies of neuroscience and on biological system [9].

Nowadays, independently from scales, the construction of connectome depends on the *imaging* method used to obtain brain/neuronal images. In spite of the advancements made in histological segmentation and microscopic imaging, microscale connectome focus at mapping the connectivity between every single neuron. This view is quite impractical considering the whole brain, due to a vast number of neurons to be mapped (approx. 10^{11} neurons) [10]. However, considering the possibility to trace every single neuronal connection, it has been

indicated that cognitive functions depend on the activity of a huge group of neurons forming a network [11]. With the advent of multi-electrode array technology (MEA), it permits a further step to distinguish the connections between patterns describing a mesoscale to study neuronal connections activities. At mesocale, the analysis includes a local population of neurons. The analysis of a small cluster of neurons, drawn out from a brain region manages a specific brain process, provides proof of a specialized sub-networks organization [12, 13].

Stepping towards a higher level, the macroscale connectome makes it simpler to understand cognition of brain functions. Neurons are clustered into anatomical and particular brain regions focusing at simplifying the formation of the brain network. However, the lacking of universal brain parcellation creates a significant barrier for such understanding. However, an anatomical parcellation of brain regions and the recognition of the inter-region connections, integrate with the increasing interest in neuroinformatics and advancements in medical imaging, made the macroscale a key tool to study the brain functions.

In fact, progress in MRI and especially in diffusion Magnetic Resonance Imaging (dMRI) and functional Magnetic Resonance Imaging (fMRI) allow connecting basic biophysical method with particular anatomical features of the brain, specifically the macroscopic architecture of the brain. These recent advances in neuroimaging, have made it feasible to examine human brain connectivity systematically and across the whole brain in large numbers of individual subjects. At the macroscale, the functional and structural aspects of the brain have drawn huge interest and researchers have initiated to exploration on the relationship between these two connectivities. The structural (or anatomical) connectome defines the white matter fibre links between different regions of the brain (cortical and subcortical). Whereas, the functional connectome is typically formulated as a representation of the statistical pairwise measure between regional brain dynamics.

The development of advanced methods to understand the brain connectivity and indeed evaluation of integration or segregation (i.e. group detection) are important, both at a single-subject level and also for group-wise studies. In recent years, there has been a rise in the number of fMRI studies which includes patients affected by mental disorders and, aims to classify between groups and identification of regions responsible for this segregation. Such analysis requires advanced tools and framework to locate anomalous connections and a deep understanding of the brain architecture underlying functional process in healthy and pathological states. Moreover, the trending dynamical analysis of functional connectivity requires advance tools to better understand the complex dynamic nature of brain connectivity and use this information to identify/classify between groups of subjects.

Motivations

In recent years with the development of advanced methods of analysing brain connectivity characterization, the curiosity of the neuroscience community to explore more and more is increasing. The concept that the brain could be narrated as a network is quite entrenched for many years, however only in the last few years, there has been a rapid increase in the efforts to analyze the brain connectivity. The structural or functional connectivity information can be considered as a graph, a mathematical object which defines a set of complex networks. In literature, the term “network” has several definitions. In the context of graph theory and complex networks, "network" explicitly implies a bundle of nodes and pair-wise edges, by which the nodes are linked. This perception is referred to as Graph theoretical exploration of brain networks. "Network", in neuroimaging, may identify a group of voxels or Regions of Interest (ROIs), that at resting state or in particular cognitive tasks, act identically [14].

The Connectome naturally acquired from this belief allows to use graph theory tools to analyze the structure and the function of the brain. The expanding interest in brain connectivity is also attributed to the recent advances made in Magnetic Resonance Imaging (MRI), with particular regards to functional MRI (fMRI), which is used to compute the pair-wise relations between brain dynamics of different regions using the formation of a functional connectome. This estimation of brain networks do not have an arbitrary organization but they show segregation and integration properties, which can be evaluated through topological measures, like the modular structure [15]. Exploring the relation between these two properties is still challenging and its interpretation is playing a crucial role in understanding brain connectivity under different circumstances/pathologies. The structural connectivity is almost static while functional connectivity has been revealed to be dynamic and highly task-dependent [4]. This aspect makes functional connectivity highly variable between subjects both during resting state and task analysis [14].

Although the graph presentation of connectivity focus at simplifying the complication of fMRI data, the encoding of such data for analysis purpose in a more robust way is still an active field of research. Indeed, the functional connectivity information of the brain requires an in-depth analysis and with advanced tools to manage the complexity of this multi-dimensional data. Moreover, a graph theory-based analysis applied to brain functional connectivity graphs may enhance the knowledge about the working of brain network and the corresponding structuring in healthy and pathological conditions for neuro-markers recognition’s. Identification of assessable neuro-markers may help the prompt diagnosis of mental disorders, making feasible early access to clinical care. With connectomics-based methods it is possible to explore different brain alterations, identifying each of them in term of connectivity alterations.

However, the discussion about the relation between connectivity alterations linked with specific symptoms and cognitive impairments is still ongoing. Indeed many questions are still unsolved, which are primarily related to the understating of the brain processes, like how the changes in functional connectivity reflect the characterization of brain networking for different brain disorders or understanding of brain modulation/integration during a cognitive task. Such information can be helpful in understanding hidden aspects of brain functioning under multiple scientific conditions e.g. effect of genetic heritability, integration of regions during task etc. In this scenario, this thesis focus on establishing new methodological approaches to address those questions.

Identification of brain sub-modular structuring might divulge the existence of interconnected brain regions which might be included in different cognitive functions. A robust group-analysis with functional data aiming at identifying distinct brain regions in response to a specifically designed task to see the effect of genetic heritability in twin's using neuroimaging based functional connectivity is still missing and it has been tackled in this dissertation. Moreover, the dynamical reorganization over time of brain modular structure has been dealt with focusing at studying the changes in functional connectivity and verifying the genetic effect in dynamic functional connectivity. To this aim, the embedding of functional connectome by using geometrical properties of connectivity matrices and graph Laplacian on the Riemannian Manifold has been explored to describe the dynamical and static reconfiguration of fMRI functional connectivity.

Assuming that alterations in functional connectivity are directly related to the brain diseases [16], classification and identification of functional neuro-markers in healthy and pathological subjects are becoming a very plausible exploration as it needs robust tools to quantify such alterations. Geodesic methods on manifold properties of connectivity have been implemented taking into account the geometrical and properties of graph Laplacians and symmetric positive definite matrices. However, neuro-markers identification in functional connectivity alterations are pretty difficult and challenging to analyze. To achieve this goal, an automatic tool has been proposed which allows the identification of disruption in connectivity between between brain regions in pathological condition avoiding the manual intervention that is prone to errors. In this proposed method, a graph-based clustering method which deploys Game Theory has been adopted at the multi-subject level to cluster the data and then finding reference networks which help to encode high dimensional data into vectorial representation. Such representation is meaningful and easy to adapt to perform classification between groups and identifying the neuro-markers.

Outline & Contribution

This dissertation is structured into five main chapters, mainly focused on brain functional connectivity analysis. Basically, we have specified some current important issues (i.e. neuro-markers detection, brain graphs classification) and propose novel methods based on graph analysis combining the geometrical properties of functional connectomes, geodesic methods on manifold approaches and clustering-based data encoding.

Chapter 2 contains a general introduction to functional-MRI techniques that permits the study of brain connectivity. It follows a detail section on how to build a functional brain connectome, introducing some methodological outlooks to deal with brain connectivity issues. Then geometrical properties of functional connectivity matrices and implementation of geodesic methods in analysis are discussed in greater detail.

Chapter 3 is primarily focused on the problem of analyzing the effect of genetic heritability on twin's using task-induced fMRI. First, a brief introduction and background related to twins zygosity and genetics have been discussed. This is followed by the experimental part which employs geodesic methods to analyze twins' connectomes on a manifold and proves the effect of genetics on functional connectivity. In the end, we compute the heritability index using Falconer's formula using the functional connectomes [17, 18].

In *Chapter 4*, we have extended the geodesic metric-based methodology explains in the previous chapter and is mainly focused on the classification of brain connectome. This chapter explains insights on connectivity-based discrimination within groups, using data encoding and machine learning methods. Then we introduce our contribution, a more general framework based on geodesic clustering of functional connectome using manifold approaches with the k-means algorithm with their application to perform classification between groups. The proposed method has been tested on real data for functional brain connectivity classification between healthy and pathological subjects [19].

Chapter 5 addresses another crucial problem in brain imaging, the identification of neuro-markers responsible for segregation and integration of brain organization in multiple pathologies. In this chapter, we further extend the methodology presented in previous chapters and our contribution includes the introduction of a graph-based clustering algorithm, based on Game Theory, which able to group the subjects in multiple clusters in a more efficient way. Then centroid of such cluster serve as reference connectome and helps to encode the data for neuro-marker identification and classification purpose. The proposed method has been tested on real data of healthy and patients affected with multiple sclerosis [20, 21].

Chapter 6 extends the algorithm described in previous chapters to develop multiple ap-

proaches which allow exploring the dynamic of functional connectivity. Following recent evidence about the dynamical nature of functional connectivity, we propose an approach to compute the similarity between multiple windows of dynamic functional connectivity and then use this similarity measure as a feature to discriminate between two groups. In the second method, we are using a geodesic metric based k-means to cluster dynamic functional connectivity matrices and then encoding data using geodesic methods for classification purpose. Both approaches were tested on task-induced twin's data to classify two groups [22, 23].

Finally, at the end in *Chapter 7*, conclusions and suggestions for future research are summarized and discussed.

Publications

The papers published or accepted during the PhD period.

- **Yamin A.**, Dayan, M., Squarcina, L., Brambilla, P., Murino, V., Diwadkar, V., Sona, D. (2019, April). "Comparison of brain connectomes using geodesic distance on manifold: A twins study". In Proceedings of 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 1797-1800). IEEE.

  The work presented in this paper is based on material presented in Chapter 3

- **Yamin A**., Dayan, M., Squarcina, L., Brambilla, P., Murino, V., Diwadkar, V., Sona, D. (2019, May). "Investigating the impact of genetic background on brain dynamic functional connectivity through machine learning: a twins study". In Proceedings of 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) (pp. 1-4). IEEE.

  The work presented in this paper is based on material presented in Chapter 6

- **Yamin A.**, Dayan, M., Squarcina, L., Brambilla, P., Murino, V., Diwadkar, V., Sona, D. (2019, September). Analysis of Dynamic Brain Connectivity Through Geodesic Clustering. In Proceedings of International Conference on Image Analysis and Processing (pp. 640-648). Springer, Cham.

  The work presented in this paper is based on material presented in Chapter 6

- **Yamin A.**, Tessadori, J., Akbar, M. U., Dayan, M., Murino, V., Sona, D. (2020, July). "Geodesic Clustering of Positive Definite Matrices For Classification of Mental Disorder Using Brain Functional Connectivity". In Proceedings of 2020 International Joint Conference on Neural Networks (IJCNN 2020) (Virtual) (pp. 1-5). IEEE.

  The work presented in this paper is based on material presented in Chapter 4

- **Yamin A.**, Valsasina P., Dayan, M., Vascon S., Tessadori J., Filippi M., Rocca M., Murino V., Sona D.. “Encoding Brain Networks through Geodesic Clustering of Functional Connectivity for Multiple Sclerosis Classification”. In Proceedings of Of the 2020 25th International Conference on Pattern Recognition (ICPR 2020) (Virtual), Milan Italy.

  The work presented in this paper is based on material presented in Chapter 5

- **Yamin A.**, Dayan, M., Valsasina P., Rocca M., Filippi M., Murino, V, Sona, D.

"Multiple-Sclerosis phenotype classification and Neuro-marker Identification using Manifold Approach" **Submitted** in Neuroimage (Nov 2020) .

The work presented in this paper is based on material presented in Chapter 5

- **Yamin A.**, Dayan, M., L. Squarcina, P. Brambilla, Murino, V., Diwadkar V., Sona, D. "The heritability of brain network profiles is contextual: Converging evidence from two analytical frameworks" **Submitted** in (Nov 2020) in Proceedings of the National Academy of Sciences.

  The work presented in this paper is based on material presented in Chapter 3

- Tessadori, J., **Yamin A.**, Valsasina P., Rocca M., Filippi M., Sona, D. "Dynamic functional connectivity for the classification of multiple sclerosis phenotype: a hidden Markov model approach." In Proceedings of 2021 IEEE International Symposium on Biomedical Imaging (ISBI 2021) (Virtual) April 13-16, 2021.

Chapter 2

Background

2.1 Functional MRI

Functional MRI is a non-invasive technique which provides a means and advanced opportunity to understand connectivity organization and functioning of the brain. fMRI based on the principle that, when the brain becomes active in response to particular stimuli (task fMRI) or during resting-state analysis (no task), the amount of blood flow to the involved brain regions increases. As a result, due to the increase in the metabolism of those regions and more oxygenated blood flows in the regions. The fMRI measures the increase in oxygenated blood demand of a group of neurons. The changes in the proportion of oxygenated and de-oxygenated blood are measured via the hemodynamic response in each voxel. Due to the paramagnetic properties of blood, MRI can capture this act, which changes with the activation of such regions. This effect is captured by MR as the Blood Oxygen Level Dependent (BOLD) signal, which is comparable to the haemodynamic response of the grouped neurons activity.

In recent two decades, fMRI has been developed and considered as one of the most prominent methods being used for functional brain imaging [24, 25] and it is growingly being used to explore the functional integrity of brain networks. Besides fMRI, there are other techniques which yield functional features as Magnetoencephalography (MEG) or Electroencephalography (EEG). However, fMRI is more admired rather than MEG and EEG since it is a non-invasive technique and it has a good trade-off between spatial and temporal resolution. Typically an fMRI image is a 4D image where the fourth dimension illustrates the time, each volume expresses brain functions at a definite time and after primary preprocessing steps, the time-series data of each voxel are extracted. One of the main goals of fMRI studies is the extraction of activation patterns during stimuli or in resting state.

General Linear Model (GLM) is one of the foundations of statistical analysis of task-based fMRI data. The configuration of a GLM can be written in a matrix shape as $Y = X\beta + \epsilon$, where Y represent a matrix of time-courses in each voxel, X is a design matrix reflecting the stimuli extracted at each time point and ϵ shows the error with constant or non-constant variance [26]. In many cases, the evaluation of β can be acquired with ordinary least squares.

Generally, the model is quite flexible and extendable to group analysis. However, GLM needs some initial assumptions as of the independence of the voxels and time points which are nearly impractical. For such reasons, alternatives and multi-variate approaches have been proposed.

Indeed, advanced algorithms such as Independent Component Analysis (ICA) or Principal Component Analysis (PCA) have been adopted to study the fMRI. ICA is one of the most considered technique to explore fMRI and analyze the functional networks [27], which normally examine for linear combinations of the original data and assuming it to be non-normal and maximally independent. It does not presume normality of the data, as in PCA, and it does not apply any dimensionality reduction. Regarding the application of ICA on fMRI data, studies related to temporal and spatial versions have materialized in literature. As an example of temporal and spatial ICA check [28], while [29] for spatial ICA. Despite several versions of ICA, a major step forward has been the development of group-ICA analysis presented by [30] as a model for group inferences, where the individual subject results are combined into a group-map.

Alternatives approach to analyze the fMRI, are based on Canonical Correlation Analysis, Maximum Correlation Analysis and several modifications of ICA have been applied in functional brain studies (see [31] for a review). All of these methodologies can be used with different fMRI protocols. However, resting-state or cognitive tasks analysis are the most widely used. To analyze the information submerged in fMRI images, it requires advanced tools for a precise examination. Many times functional alterations underlie brain disorders such as Autism, Schizophrenia, Multiple-Sclerosis or Alzheimer. Indeed, the development of methods focusing at a complete mapping of functional connectivity might refine the knowledge of such diseases allowing a translation into clinical studies.

2.2 Functional Connectome

Connectome has its basics in the concept of network. Nowadays network science has come over different disciplines, from social science to neuroscience. In general, each connected system with links is considered as a network. In neuroimaging, we refer the brain network to describe a set of brain regions linked with each other, or as a group of brain regions activated at the same time, [32]. In general, the term network is mostly appropriate for structural connectivity since it represents physical connections, while the functional connectivity is linked to statistical measures. However, we can use the word "graph" that is more general and it describes the mathematical objects, over which, analysis of brain connectivity can be done.

Whenever a graph can model the association between objects, the choice of nodes and links becomes the preeminent step. Nodes are objects with meaning, which explicate a part of the entire system (i.e. anatomical or functional) while keeping an independent relationship with each other. In brain graphs, with regards to the macro-scale connectome, each node represents a brain region with a specified meaning. The major question concerns how these brain regions should be defined. This problem of defining the brain in different regions based on its functionality is known as *brain parcellation* and various attempts have been made to steady the choice of the brain atlas. Typically, brain atlas depends on the modality; the first basic method to define the nodes of the brain graph is to consider every single voxel of the MRI image (structural, diffusion or fMRI) as a vertex. However, it prompts computational problems due to the vast number of voxels in MRI images and it disobeys the presumption that each node should be independent of each other.

Grouping voxels according to common behaviour has been presented to segment the brain in notable regions; one of the most usual brain atlas used in functional studies is the Automated Anatomical Labelling (AAL) [33], which mainly focus on anatomical features and does not consider functional information. There are other brain atlas constructed using functional features, probabilistic approaches or other anatomical features and also taking into account the information on brain anatomy [34]. Predominantly, even though there are several ways to define the nodes, the main advantage of the connectome approach is that once defined the most suitable brain atlas, functional and structural connectivity analysis can be simultaneously performed through the forming of a *connectivty matrix*. The nature of the links typically defines how we want to narrate our system. Binary links only indicate the presence or absence of connections, while weighted links also denote the strength of the connection. On the other side, links can indicate the directionality which makes a graph directed. However, undirected weighted graphs $G = (V, E)$ is usually constructed to estimate the brain connectivity, where a node/vertex (V) in the graph delineates a brain region (i.e. ROI) and an edge/link (E) between two nodes is indicative of brain regions being functionally connected (connection weights) [35].

Functional connectivity graphs have different construction techniques. Considering the brain parcellation, which recognizes the set of brain regions, functional links between nodes are defined following different rudiments. After the acquisition, preprocessing of fMRI image and extraction of corresponding time-series signal, which is extracted for each voxel and averaged to describe a single time-series for each node in the graph (check [36] for a review on fMRI pre-process), one of the ways to build the functional connectivity is through the computation of covariance between each pair of time-series, which reflects signal transmission

from one region to another via indirect links [37] and is defined as:

$$cov_{x,y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{N-1} \tag{2.1}$$

where $\bar{x}$ and $\bar{y}$ is the mean of time-series of regions x and y respectively. Covariance matrices enjoy the property of being symmetric & positive semi-definite, however, in a real scenario, they are almost always symmetric & positive definite (SPD) (described in detail in section 2.3). Another and one of the most common way to compute functional connectivity matrix is by using Pearson Correlation between each pair of time-series, defined as:

$$\rho_{x,y} = \frac{cov_{x,y}}{\sigma_x \sigma_y} \tag{2.2}$$

where σ_x and σ_y is the standard deviation of time-series of regions x and y respectively. As it can be seen in Eq. 2.2 that correlation matrix ρ, which is also symmetric and positive semi-definite, is a normalized version of covariance and represent a subset or submanifold of SPD. So, due to this reason, correlation can be considered as a constrained type of SPD matrix that builds a submanifold, assuming generic geometric properties of normal SPD manifold. This does not reduce the significance of SPD manifold and to study functional connectivity since many depictions of functional connectivity are equipped with SPD properties and the suggested methods in this dissertation are presented for a broader class of SPD matrices.

Hence, we construct the functional *connectivity matrix* which extract a statistical measure of each pair of time-courses over time. However, to date, no accordance has been reached about which functional connectivity measures is able to define the similarity between two functional time-series. There are some alternative measures such as, partial correlation, which reflects the direct level of interactivity between two regions removing the impact of all other regions.

In recent years, **Graph Laplacian** has been widely used to analyze the brain connectivity due to its geometrical properties based on symmetric & positive semi-definite nature of graphs. Different graph laplacians exist and they can be computed from the corresponding connectivity matrix. Assuming W as the functional connectivity matrix, which describes the brain network as an undirected weighted and symmetric graph with positive weights $(w_{ij} > 0)$, we divide graphs Laplacian into two classes: unnormalized and normalized graphs. The Unnormalized graph Laplacian is defined as follow:

$$L = D - W \tag{2.3}$$

where D is the diagonal degree matrix $(D = \text{diag}(\sum_i w_{ij}))$.

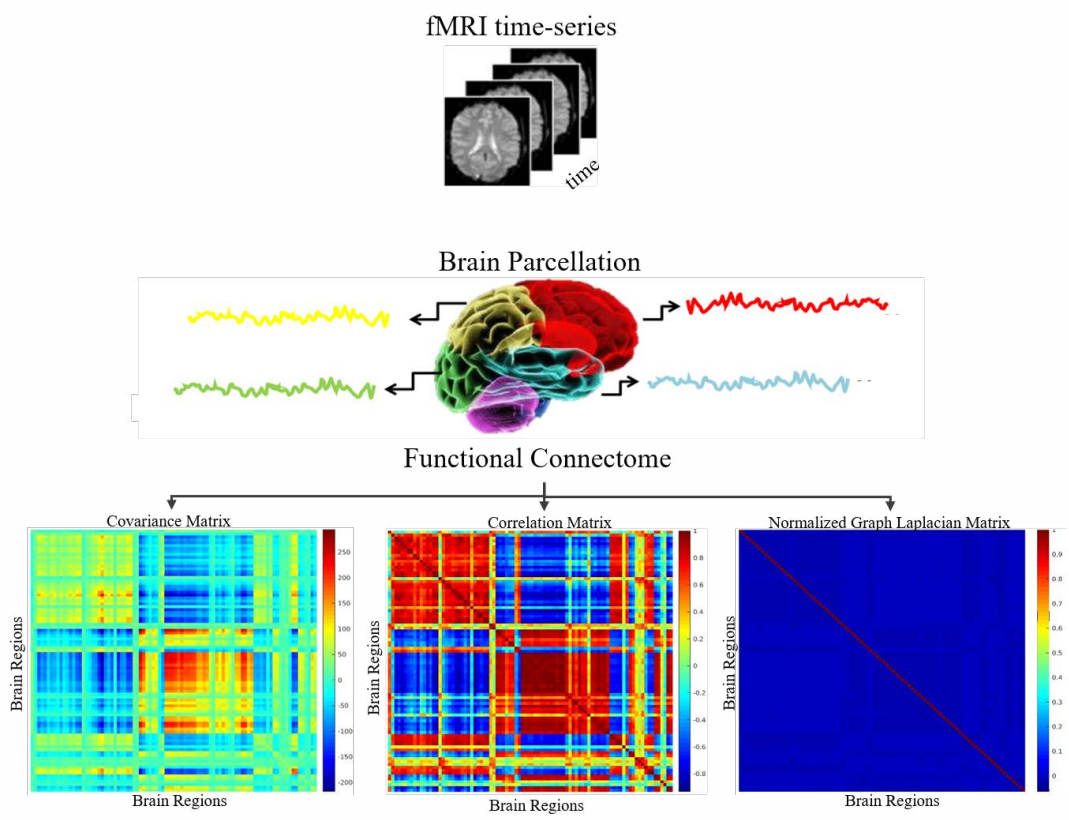

Figure 2.1: The pipeline to build functional brain connectome. **a)** A generic example of fMRI images of the human brain. **b)** A generic brain parcellation of the human brain. **c)** fMRI functional connectivity matrix computed using covariance, correlation and normalized graph Laplacian

Beside, there are two Normalized Graph Laplacian, defined as follow:

$$L_{sym} = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}} \tag{2.4}$$

$$L_{rw} = D^{-1}W \tag{2.5}$$

where L_{sym} is the symmetrized version of the Laplacian and L_{rw} is the Random Walk Laplacian. Figure 2.1 shows the pipeline to obtain the fMRI functional connectome using multiple methods starting from the raw images.

However, fMRI functional connectome construction is undergoing change since recently it has been demonstrated that functional brain connectivity has a dynamical behaviour [38], and that along with an fMRI scan, different patterns of FC can occur [4, 39]. So, recently the term "Chronnectome" has been lodged to identify the new frontier of studies related with brain functionality, which should take into account the changes occur in the network over time [40]. Basically, instead of providing only one connectivity matrix expressing the correlation between each pair of the entire fMRI time-courses, scientists have started to

ponder local correlations along with the temporal FC, sampling the whole time-series at a different time point, allowing to define FC over time with multiple connectivity matrices. The simplest and most widely used approach to build an fMRI functional chronnectome is through sliding windows technique [38, 4], which allows capturing local FC correlations. In the next two sections, we introduce one of the important property of functional connectivity connectome and related approaches to study and explore brain connectivity.

2.3 Manifold Approaches

A crucial issue in brain connectivity analysis is related to the geometrical properties of original connectivity matrices. Especially in the case of methods for classification and discrimination task, a measure to define the similarity index is needed. Assuming as functional weights the Pearson correlation or covariance between region time-series for functional connectivity and/or the number of fibres connecting each pair of brain regions for structural connectivity. In such space, the euclidean distance is the natural way to measure the similarity between two objects. However, referring to brain graphs, this choice is suboptimal since it does not well describe the complex geometry of the data input.

An alternative embedding of the input data is given by the properties of manifold. Of particular interest are the *Riemannian Manifold* with the corresponding geometrical properties. In the following sections, we will revise the main properties of these approaches.

2.3.1 Riemannian Manifold

A manifold is a topological space which is locally similar to an Euclidean space. A Riemannian manifold $\mathcal{M}$ is a differentiable manifold which is outfitted with a smooth inner product on each tangent space. The Geodesic distance on the manifold defines the distance between two generic points as the length of the shortest curve connecting each other. Of our interest is the Riemannian manifold defined by all symmetric & positive definite matrices (SPD) of the same size. A typical example of SPD matrix is the covariance or correlation matrix, which represents a possible way to describe the relations between fMRI region time-series for functional connectome (see section 2.2). Hence, we introduce Riemannian manifold and its properties for a generic set of SPD matrices.

Let Sym_n^{++} denote the set of SPD matrices of size $n \times n$, that is the set of all symmetric $n \times n$ matrices W such that the quadratic form $\mathbf{x}^{\mathbf{T}} W \mathbf{x} > 0 \ , \forall \ \mathbf{x} \neq 0 \in \mathbb{R}^n$. A crucial property of the set Sym_n^{++} is that it is not a vector space but forms a Riemannian manifold [41]. As a consequence of the manifold structure of Sym_n^{++}, computational methods that simply rely on the Euclidean distances between SPD matrices are generally suboptimal [42]. Figure 2.2

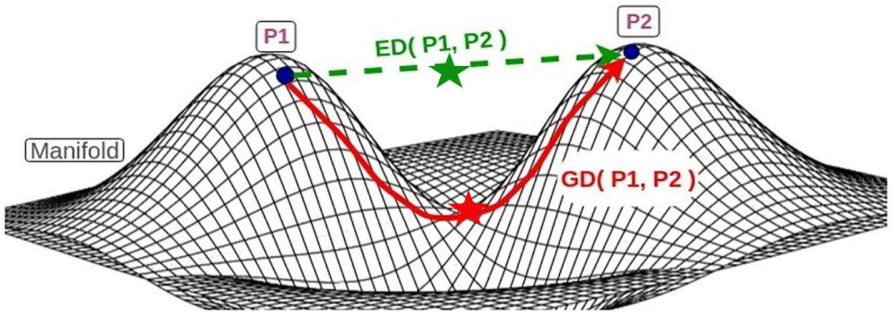

Figure 2.2: The conceptual difference between Euclidean distance (green line), Euclidean mean (green start) and geodesic distance (red line), geodesic mean (red start) between two points P1 and P2 on a manifold.

visualizes the difference between geodesic and Euclidean distance on a manifold.

Over the last few years, geometric properties of Riemannian manifold of SPD matrices have been exploited in brain connectomics but have not yet been fully recognized. The main use of these properties was to capture and discriminate complex or non-linear patterns which can occur in healthy or mental disorders. In [43], a probabilistic model based on the manifold of covariance matrices has been used to separate post-stroke patients from a group of healthy control. Also, taking into account the manifold structure, the concept of transport on the manifold of covariance matrices of fMRI time-series has been successfully applied in longitudinal studies to classify changes in functional connectivity after a particular task [44].

In [45], Gaussian kernels based on Log-Euclidean and Stein divergence metrics on Riemannian manifolds have been used to classify subjects between healthy and those with pathological disorders. In [6], Grassmannian geometry on graph Laplacians have been used to highlight sub-networks that, in turn, allow FC to be used for classification purposes. Similarly, Riemannian geometry approaches allowed the classification of sub-connectivity patterns [46], the diagnosis of mild cognitive impairment [47] and the identification of auditory stimuli from the evoked functional states [48].

Moreover, in [49, 50, 51, 52, 53, 54] kernel-based methods and in [49, 54, 55] manifold based methods related to supervised dictionary learning were used by assuming that all the samples were on the Riemannian manifold of SPD matrices. In kernel-based dictionary learning methods, [53] employed multiple kernels to get better results, unlike [50, 51, 52] where only single kernels were used. Riemannian manifold based methods were defined in [54, 55] and Grassman manifold based methods were deployed in [49].

Also in biomedical applications, some examples have demonstrated the advantages of those

techniques. For example, in Brain-Computer Interface (BCI), Riemannian manifold kernel classification has been adopted in EEG-based studies [56]. In [57] a combination of locally linear embedding and Log-Euclidean Riemannian metric [58] was proposed to embed functional brain network into a lower-dimensional space. Although recently, methods relying on Riemannian geometry are receiving particular interest in functional connectivity, originally, Riemannian framework has first been applied on diffusion tensor imaging, aiming at designing the tensor structure of white matter fibres [41, 58].

Riemannian metrics and kernel methods are quite adopted for image classification since the combinations of such frameworks overcomes standard approaches in Euclidean space [59]. In computer vision tasks, as image classification, Riemannian manifold properties and kernel methods are extensively used, thanks to the fact that their combination outperforms kernel methods developed in Euclidean space [59]. The main reason for such improvement has to ascribe to the metric and to the smoothness of such manifold defined by PSD matrices. Riemannian metrics have the great advantages to describe much better the distances between a point on the manifold and they easily allow to define a kernel matrix.

Graph Laplacian Recently, graph laplacian has been widely adopted to analyze the connectivity of the brain. In [45, 6] Kernel-based method using graph laplacian has been successfully deployed to perform classification between subjects and identifying discriminative brain connection. Details regarding the computation of graph laplacian have already been discussed in section 2.2. Focusing on the unnormalized and on symmetric normalized Laplacian, we revise the main properties. Laplacian matrix is a *semi-definte positive matrix* (PSD) and it can be decomposed in eigenvalues and eigenvectors as follow:

$$L = U'\Lambda U \tag{2.6}$$

where $U = [u_1, ..., u_n]$ is the eigenspace and Λ is the diagonal matrix with corresponding eigenvalues. Since Laplacians are *semi-definte positive matrices*, λ_n are always positives $0 = \lambda_1 \leq ..... \leq \lambda_n$ and can be easily regularized to a positive definite (PD) matrix by the modification $\hat{L}_t = L_t + \gamma I$, where $\gamma > 0$ is a regularization parameter and I is the identity matrix.

2.4 Geodesic Metric on Manifold

To fully exploit the advantage of the manifold structure, it is essential to consider the notion of geodesic distance, which measures the length of the shortest curve on the manifold connecting two points (two matrices). Among the different Riemannian metrics that have been considered on Sym_n^{++}, the one that has been most studied and analysed is the classical

affine-invariant metric, which is a full geodesic distance on the manifold between two SPD matrices W_1 and W_2 is given by:

$$d_{\text{affine}}(W_1, W_2) = \left\| \log \left(W_1^{-\frac{1}{2}} W_2 W_1^{-\frac{1}{2}} \right) \right\|_F \tag{2.7}$$

with $\log(\cdot)$ denoting the principal matrix logarithm and $\| \cdot \|_F$ denoting the Frobeniuous matrix norm. While being invariant concerning all affine transformations, a drawback of the affine-invariant metric is that it is computationally intensive, especially for large scale datasets. This latter problem might not be so relevant in neuroimaging, where usually the amount of data available is not so large. However, a significant drawback of an affine metric is its impossibility to generate any kernel matrix, which makes it useless as a similarity function for kernel methods.

To overcome this problem, other metrics have been developed as approximation of full geodesic distance on Sym_n^{++} . The most common one is the *Log-Euclidean* distance [58], which is simple and fast to compute. Equation 2.8 and 2.9 describe, respectively, the log-E distance formula between two SPD matrices W_1 and W_2 and the closed form formula to compute the mean [60] (illustrated in Figure 2.2) of two or more SPD matrices with this metric.

$$d_{logE}(W_1, W_2) = ||\log(W_1) - \log(W_2)\,||_F. \tag{2.8}$$

$$W_L = \exp\left\{ \arg\inf_W \sum_{i=1}^{K} \| \log(W_i) - \log(W) \|^2 \right\} = \exp\left\{ \frac{1}{n} \sum_{i=1}^{K} \log(W_i) \right\}, \tag{2.9}$$

The matrix can be described as a spectral decomposition $W = U\Lambda U^{\mathrm{T}}$ where U is an orthogonal matrix and, because W is positive definite, Λ is a diagonal matrix with strictly positive entries. Hence, the logarithm of a matrix is given by $logW = Ulog\Lambda U^{\mathrm{T}}$, where the logarithm of the diagonal Λ is computed as the logarithm of all elements in the diagonal. Similarly, the exponential of a matrix can be computed as $exp(W) = Uexp(\Lambda)U^{\mathrm{T}}$.

Log-Euclidean is a full geodesic distance on the Riemannian manifold. This metric is computationally faster compared to the Affine distance (Eq. 2.7) and it defines a positive kernel which can be easily combined with a kernel-based classification algorithm.

An alternative metric, which is much faster compared to Eq. 2.7 and 2.8 is the *Stein Diver-*

gence between two SPD matrices. Stein Divergence [61] is defined as follows:

$$d_{stein}(W_1, W_2) = \left\{ \log\det\left(\frac{W_1 + W_2}{2}\right) - \frac{\log\det(W_1 W_2)}{2} \right\}^{1/2} \tag{2.10}$$

Stein divergence is not a Riemannian metric and it is based on the convex structure of Sym_n^{++}. As well as the Log-Euclidean distance, it defines a positive kernel suitable for kernel-based methods.

As in [62], instead, a component of the Frechet distance is describes known as Wasserstein distance. The Wasserstein distance is defined between two Gaussian distributions, which is built on two components. One component considers the means, and the other component is between the covariance's defined as:

$$d_t^2(W_{(x)}, W_{(y)}) = tr[W_{(x)} + W_{(y)} - 2(W_{(x)} W_{(y)})^{1/2} \tag{2.11}$$

Where $W_{(x)}$ and $W_{(y)}$ are the covariance's of two distributions x and y respectively, and tr is the trace operator. It has been proven that the distance expressed by Eq. (2.11) is a metric on covariance, hence, on positive semi-definite matrices. The advantage of the Wasserstein metric is that it allows the computation of the geodesic distance between two positive semi-definite matrices without introducing the regularization and it is easy and fast to compute as compared to other geodesic distances.

Chapter 3

A Riemannian Framework for the Comparison of Brain Connectomes

3.1 Introduction

The study of Twins provides a foundational perception of the respective contributions of genes and the environment on immeasurable aspects of the unfolding lifespan [63]. Consequences of Heritability are as diverse as to incorporate (but not limited to) traits of personality [64], brain structure [65], intersection between regions of brain [66], and the functional operations of the basic motor systems of brain [67]. Indeed, the genetic effect on human behaviour is so considerable, extensive and replicable [68], that genetic impact on behaviour is casually preserved in "first law of behavioural genetics [69]. As observed in some studies that mono (MZ)- and di-zygotic (DZ) twin samples are important with notable power for concluding additional and non-additional contributions of genetic/environment effects [70, 71, 72]. Generally, at the time of birth, MZ twins share 100% of their genetic data (against 50% in DZ twins), and a simplifying presumption is that shared genetic material will deploy effects throughout the lifespan.

Particularly, if on average, MZ pairs are more similar on some "biological trait" as compared to DZ pairs, and given that the in utero and post-birth environmental histories are identical to similar (between members of the pair), then mono-zygosity may be deduced to "cause" the effect of that increased trait similarity. Customary methods of understanding brain activities based on brain networks. Due to a variety of reasons related to genetics and formulation of brain organization, a complex question is that either the functional brain networks can be considered as "biological traits" or not? Typically, functional brain networks are inherently dynamic [73].

Some recent studies in twins have addressed the heritability in brain networks using resting-state fMRI data [74], proposing moderate to high heritability of intrinsic resting-state functional connectivity[75, 76]. The heritability of task-induced brain network profiles may themselves be highly contextual. To be more precise, a determined visuo-motor task (coupled to a more general behavioural aspect such as visuo-motor function), will induce network

connectivity in task-related sub-networks of regions [77]. In the case of twin studies, this proposes that the resemblance in network profiles between MZ twin-pairs (as compared to DZ pairs) may be greater for task-related sub-networks (in contrast with task-orthogonal sub-networks). This can be suggested because the effect of genetics may be contextualized by the properties of task, and arbitrate specific connectivity patterns. Use of such a provoking and novel framework has not been suggested before. If task-relevant brain functioning is dynamic itself, as is expected in an open system like the human brain[78], then consequences of genetics on brain networks will also be sensitive to selective network effects. Experimentally we have assessed this complex question within the conventional twinning design structure (pairs of MZ and pairs of DZ twins).

A primary methodological novelty in this work, was the use of a concise analytical framework [45] that allowed the mapping of covariance-based functional connectomes (which are symmetric and positive semi-definite) induced by the task to Riemannian manifold space (see chapter 2, subsection 2.3.1) [79]. This mapping to a smooth geometrical manifold gives a simple outline of network activity within a manifold space. As discussed earlier in section 2.4, the use of euclidean metric on such matrices is suboptimal and hence use of geodesic distances between mapped connectome within the manifold space is suggested which provide a metric of similarity between network profiles. A few studies of analysis of functional connectomes on Riemannian manifolds of SPD matrices includes, evaluation of average and variability of group-level covariance matrices and statistical testing [43, 80], detection of change point [81], covariance matrix estimation on manifold for individual [82], functional connectivity regression for estimating structural connectivity [83] and dimensionality reduction of connectivity data for machine learning application [55, 84, 57, 85].

By mapping the connectome profiles of task-relevant and task-orthogonal sub-networks induced by the visuo-motor integration task, it enables us to evaluate similar metrics of profiles in MZ and DZ twin pairs. The resultant findings were further confirmed using a supportive (and established) method for evaluating heritability (H_0) index, based on Falconer's formula[86]. H_0, the heritability of any trait is built on the difference between correlations of pairs of twins (where it is expected that the correlations in MZ twins on an inheritable trait are higher than in DZ twins). In this work, H_0 was computed for each component in the functional connectome of task-relevant and task-orthogonal networks Figure 3.1.

For comparison purpose, the heritability was explored in the same sub-networks using task-evoked fMRI data in twins from the Human Connectome Project (HCP) data set [87]. We specifically considered two task domains: a) visually guided motor mapping (henceforth HCPMotor) wherein similar to the primary analyses, visual stimuli were used to evoke effector responses (hand, foot or tongue) and b) a standard working memory task (with verbal

and visual stimuli; henceforth HCPWM). Across twin samples of our dataset, tasks, and analytic methods, the accumulation of results (Figure 3.5 - 3.10), presents clear validation that the heritability of brain network profiles is contextual.

3.2 Material and Methods

3.2.1 Data Acquisition

Thirteen twin pairs (7 MZ, 6 DZ) have been recruited from the population-based Italian Twin Registry. MRI data were acquired on a 3-Tesla MR imaging unit Siemens Allegra system (Siemens, Erlangen, Germany) with a standard head coil. T2*-weighted images were acquired using a gradient-echo EPI-BOLD pulse sequence (TR: 2000 ms; TE: 30 ms; flip angle 75°; FOV: 92x192; 31 axial slices; thickness: 3 mm; in-plane: 3 mm2; matrix: 64x64). High-resolution MPRAGE T1-weighted structural images were acquired in the same session (TR: 2300 ms; TE: 3.93 ms; flip angle 12°; FOV: 256x256; 160 axial slices; slice thickness: 1 mm; matrix 256x256).

During fMRI, subjects were positioned with adjustable padded restraints employed for head stabilization. Details of basic visuo-motor paradigm and stimuli are given in [88]. Responses were made with either the Right or Left hand and the use of hand was blocked across scans; Thus, two separate scans (182 volumes each) were consecutively acquired and blocking response hand across scans was designed to preempt ancillary effects of excitatory and inhibitory signalling response competition [88], as these processes were tangential to the goals of the study.

3.2.2 Data Pre-Processing

Pre-processing of MRI data for investigating network profiles across the entire cerebrum was performed using a combination of MATLAB and shell scripts. These scripts collectively integrated the functionalities of AFNI [89], FSL [90], and FreeSurfer [91]. T1 structural images were processed using the FreeSurfer recon-all command-line tool to obtain the grey matter (GM), white matter (WM) and Cerebrospinal fluid (CSF) masks. This method identified ROIs from the Destrieux atlas [12, 13, 92], and then GM masks were created by integrating the resulting ROIs accordingly. GM for all subjects was parcellated using the AAL atlas defined in MNI space [33]. Each subject's T1 image of each subject was transformed to the T1 MNI152 template with FSL FNIRT Non-linear transformation [93]. The transformed T1 GM mask was directly applied to the AAL atlas to match the subject's GM anatomy.

For the purpose of computing the functional connectivity (FC) matrices, we considered the

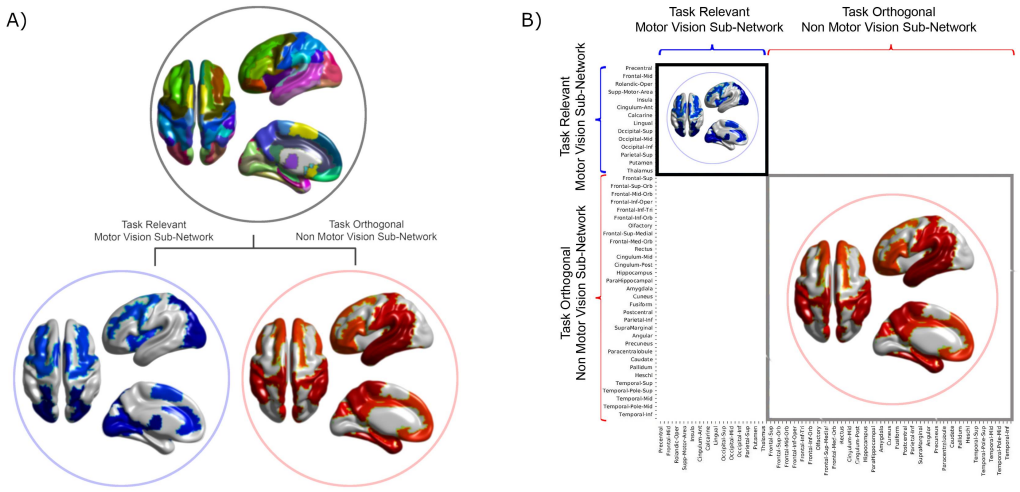

Figure 3.1: A) Parcellation of 90 cerebral ROI of AAL atlas (top) into task-relevant MV (blue) and task-orthogonal NMV (red) sub-networks. B) Division of FC matrices into task-relevant (blue bracket, black inset) and task-orthogonal (red brackets, grey inset) sub-networks.

set of 90-ROI in the cerebrum only as defined in the AAL atlas. The visuo-motor attributes of the task prompt a priori clustering of the ROI set into two classes of sub-networks. First, the task-relevant sub-network was influenced by two criteria: a) an ALE-based meta-analyses of motor paradigms [94] acquired from visual-guided finger tapping studies and b) previous work documenting effects associated with probe detection tasks using a similar paradigm [88]. This task-relevant motor vision (MV) sub-network was composed of n=28 bilateral ROIs and was complemented by a task-orthogonal non-motor vision (NMV) sub-network comprising of n=62 bilateral ROIs that were orthogonal to task-induced activity. The resultant classification of sub-networks (as illustrated in Figure 3.1) was maintained for the remainder of the study and established the bases for the main analyses of network profiles.

From each of these 90 ROIs, the mean time series across voxels in unsmoothed fMRI images were extracted using FSLMEANTS [95], individually for each task acquisition (Left-hand and Right-hand tasks). These time series were then used to construct two symmetric 90 x 90 FC matrices for each subject (C_i^{Right} and C_i^{Left}) by using the covariance coefficient (Eq. 2.1) between each possible pair of ROIs. Hence, These matrices represent the functional connectomes showing co- and counter-activations of ROIs, including self-connections in the diagonal which indicates the overall variance of ROIs. The estimated spatial variance-covariance matrices were configured in a manner that cleaved apart rows (and columns) into Task-Relevant and Task-Orthogonal sub-networks. For this analysis, we did not exclude any negative values from the connectome and used the whole connectome with all positive and negative covariance values.

3.2.3 Manifold Mapping of Functional Connectome

The advantage of using covariance-based functional connectomes is that they form a set of symmetric positive semi-definite matrices [79] and can be easily made symmetric positive definite (SPD) with a small regularization (see section 2.3 for derails). In practice, most of the time covariance-based matrices are already SPD, and regularization is not needed. The set of SPD matrices of size enjoys the important property of being represented by a Riemannian manifold [41] which is a smoothed manifold that in our case described the set of connectomes with nodes. An advantage of the Riemannian manifold representation is that the distance between the mapped points on its curved structure can be captured and measured using the notion of geodesic distance, that is the length of the shortest curve connecting two points on the manifold and the use of the Euclidean distance metric is sub-optimal for capturing true distance in this space [42].

For this work, we have deployed the Log-Euclidean distance (Log-E) to quantify the distance between profiles. The estimation of Log-E distance is characterized by computational simplicity as it can be estimated using matrix logarithms (as described in Eq. 2.8 of section 2.4). Moreover, formula describing the geodesic mean of matrices w.r.t Log-E distance is described in Eq. 2.9 in section 2.4.

3.2.4 Normalizing Log-E Distance Estimates Based on Sub-Network Size

Suppose two sub-networks of N brain regions, with their covariance-based functional connectivity matrices mapped to a manifold space. The distance d^N_{logE} in the manifold space between any two corresponding (e.g. Task-Relevant$_{Twin1}$ and Task-Relevant$_{Twin2}$) functional connectivity matrices $\hat{C}^N_1$ and $\hat{C}^N_2$ is based on the Log-E distance distance defined as

$$d^N_{logE} = ||\log\hat{C}^N_1 - \log\hat{C}^N_2||_F. \tag{3.1}$$

The computed d^N_{logE} is an increasing function of N, indicating that the estimated distance of this metric can be confounded by sub-network size (i.e., N). Due to the differences in the size of sub-network (N=28, Task-Relevant; N=62, Task-Orthogonal), it was essential to quantify the effects of network size N on computed distance d^N_{logE} to make sure that the computed distance was not affected by differences in network size N.

To evaluate how d^N_{logE} varied with changing network size N and to avoid any confounding effects of zygosity, we computed d^N_{logE} between pairs of random unrelated subjects, after selecting random connectivity matrices of size N, where N varied from 1 to 90 (the total number of cerebral regions in the AAL atlas). More precisely, we set up an iterative experiment in which we selected each pair of unrelated subjects (from a total of 312 unrelated

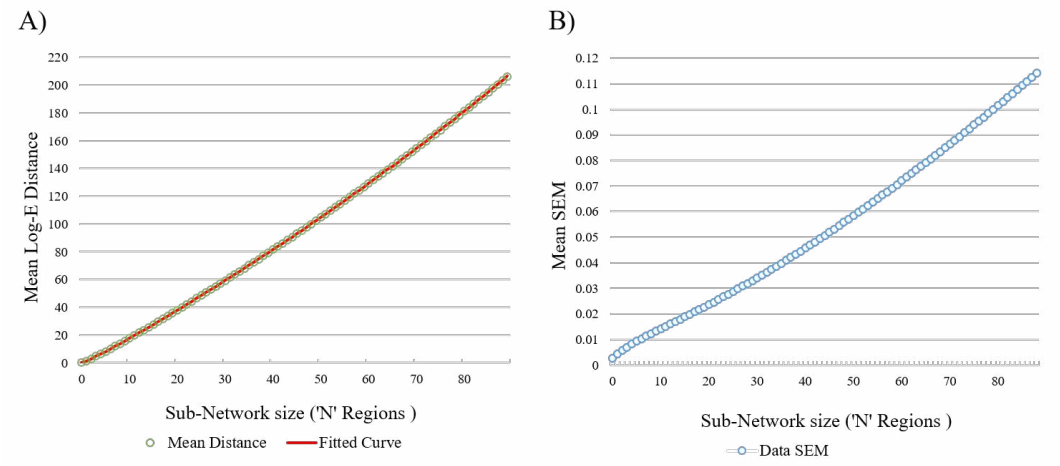

Figure 3.2: A) The mean Log-E distance from the experimental procedures (for 10 unrelated subjects) is plotted as a function of sub-network size (green open circles). The red curve depicts the power law function fit to the data (see text) to use parameters for subsequent normalization. B) The SEM from the data for 10 unrelated subjects (Figure 3.4) is plotted as a function of sub-network size. As seen, the SEM (as the mean Log-E distance) increased as a function of sub-network size.

pairs of subjects). Then, for each pair, we randomly selected 10 groups of N regions for which we computed d^N_{logE}, while iteratively increasing N from 1 to 90. For each value of N, we therefore had 100 d^N_{logE} measurements. The resulting data were fit to estimate the relationship between d^N_{logE} and N.

In Figure 3.2, A) depicts the mean Log-E distance as a function of N, and b) depicts the standard error of the mean, derived from the experimental simulations. As observed both the estimated mean Log-E distance and the SEM were confounded by increases in the sub-network size (N) of the connectivity matrices.

We fitted a power law function defined below to the data shown in Figure 3.2.A,

$$y = a \times x^b \tag{3.2}$$

where the parameter a represent the proportionality constant and the parameter b is the power (or exponent). Both were considered functions of N. The fitting procedure was based on the least squares method. The parameters a and b were selected to minimize the squared difference between the measurements and the fitted data. The least square fit resulted in coefficients $a = 1.053(1.036, 1.069)$ & $b = 1.171(1.168, 1.175)$. The fit parameters from Eq. 3.2 were used to obtain the normalization factor for each number of region N, such that the distance between the connectivity matrices of size N was divided by the normalization factor $1.053 \times N^{1.171}$.

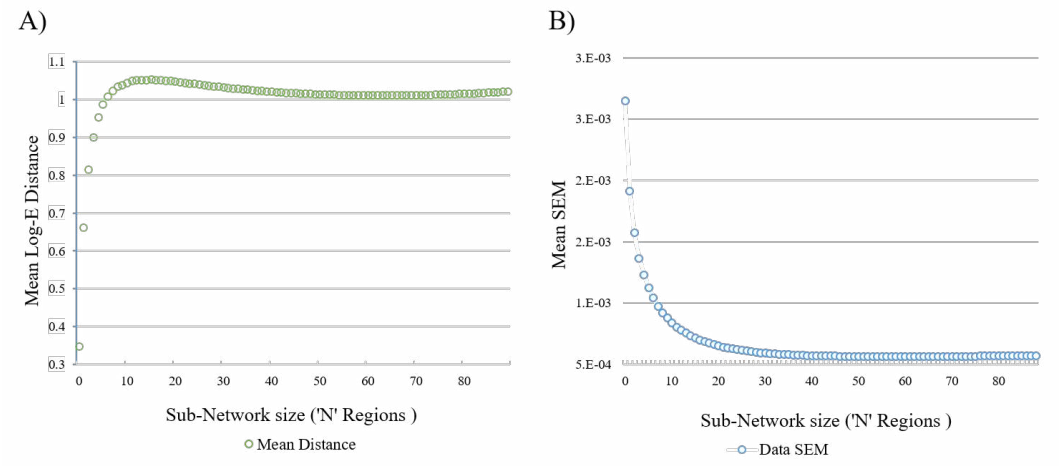

Figure 3.3: A) The Log-E distance (Figure 3.2) expressed after applying the normalization (for network size) factor. As seen, the Log-E distance estimate is impervious to network size for $N \geq 15$. B) The SEM of normalized data. As seen, the SEM values are very small and no changing for $N \geq 15$ as a function of sub-network size.

Figure 3.3.A depicts the data after applying the normalization factor to the data in Figure 3.2.A. As observed following normalization, Log-E distance was invariant of sub-network size for $N \geq 15$, indicating that normalization of the true obtained data for each of the task-relevant and task-orthogonal sub-networks corrected for confounds related to network size and hence, Log-E distance were normalized to make them comparable on task-relevant and task-orthogonal networks. Figure 3.3.B depicts that SEM of normalized data and it can be seen that SEM values area very small and unchangeable for $N \geq 15$.

3.3 Experiments and Statistical Analysis

In this work, our dependent variable of interest was the Log-E distance between functional networks. We used the Log-E distance to estimate the similarity in network profiles between twin pairs (MZ_x - MZ_y; DZ_x - DZ_y), because the metric represents the shortest path in the manifold between pairs. Log-E distance between each twin pair was independently calculated for the a priori selected Task Relevant (MV) and Task Orthogonal (NMV) networks, and when responding with either the Right or Left hand. Thus, each pair contributed four Log-E distance values forwarded for subsequent analyses.

A three-way mixed Analyses of Variance was employed to evaluate the effects of the three independent variables of interest on the calculated Log-E distance. The three modelled factors were 1) Zygosity (Mono vs. Di) modelled as a between subject's factor; 2) Network sub-type (Task Relevant vs. Task Orthogonal) modelled as a within subject's factor and 3) Response hand (Right vs. Left) modelled as a within subject's factor. To achieve our aims,

we were focused on the main effects of Zygosity and Network sub-type, and second-order interactions (Zygosity x Network). The task-relevant (MV) and task-orthogonal (NMV) networks of interest included 28 and 62 ROIs respectively, both of which larger than the minimum threshold required for the Log-E distance to be comparable across networks (see subsection 3.2.4 and Figure 3.3). The factor of Response hand was not of specific interest but was modelled to ensure no contribution from hypothesis-neutral sources. All significant effects were clarified in a post-hoc analysis using non-parametric paired t-tests (Wilcoxon).

3.3.1 Alternative Approach using Graph Laplacian:

To validate the result & findings, we suggested using an additional approach, where we have used correlation matrices instead of covariance matrices for analysis. As defined in section 2.2 correlation matrices define a submanifold of SPD matrices as it is a normalised version of the covariance matrix. So, for all given subjects we computed person correlation-based functional connectomes and applied r-z transform to normally distribute the data and only retaining the positive correlation values as commonly practised in functional connectivity analysis [96].

For these symmetric undirected weighted connectivity graphs, it is possible to define Graph Laplacian (GL) which enjoys some useful properties of being symmetric and positive semi-definite. Specifically we used Normalized Symmetric Laplacian (see section 2.2 and Eq. 2.4). Once matrices were obtained we implemented to these graph Laplacian based connectivity matrices the same method defined in subsection 3.2.3 but using Wasserstein distance (see section 2.4 and Eq. 2.11). We also applied normalization of the distance values process (see subsection 3.2.4 and Figure 3.3) on the data obtained with Wasserstein distance, to make them normalize and comparable. For this comparison we applied post-hoc analyses using non-parametric paired t-tests (Wilcoxon) on the data achieved with this alternate analysis method and results are illustrated in Table 3.2.

3.4 Results

3.4.1 Characterization of Network Profiles (Covariance matrices)

Figure 3.4 illustrate covariance matrices for all the subjects (n=26) in the initial study. For each participant, the task-relevant and task-orthogonal network are delineated (see Figure 3.1). Within each matrix, the covariance between time series across all pairs of regions within the task-relevant sub-network (MV, 28 regions) and the task-orthogonal sub-network (NMV, 62 regions) are portrayed. Visual inspection suggests higher similarities in covariance a) between corresponding regions of the task-relevant (MV) relative to task-orthogonal

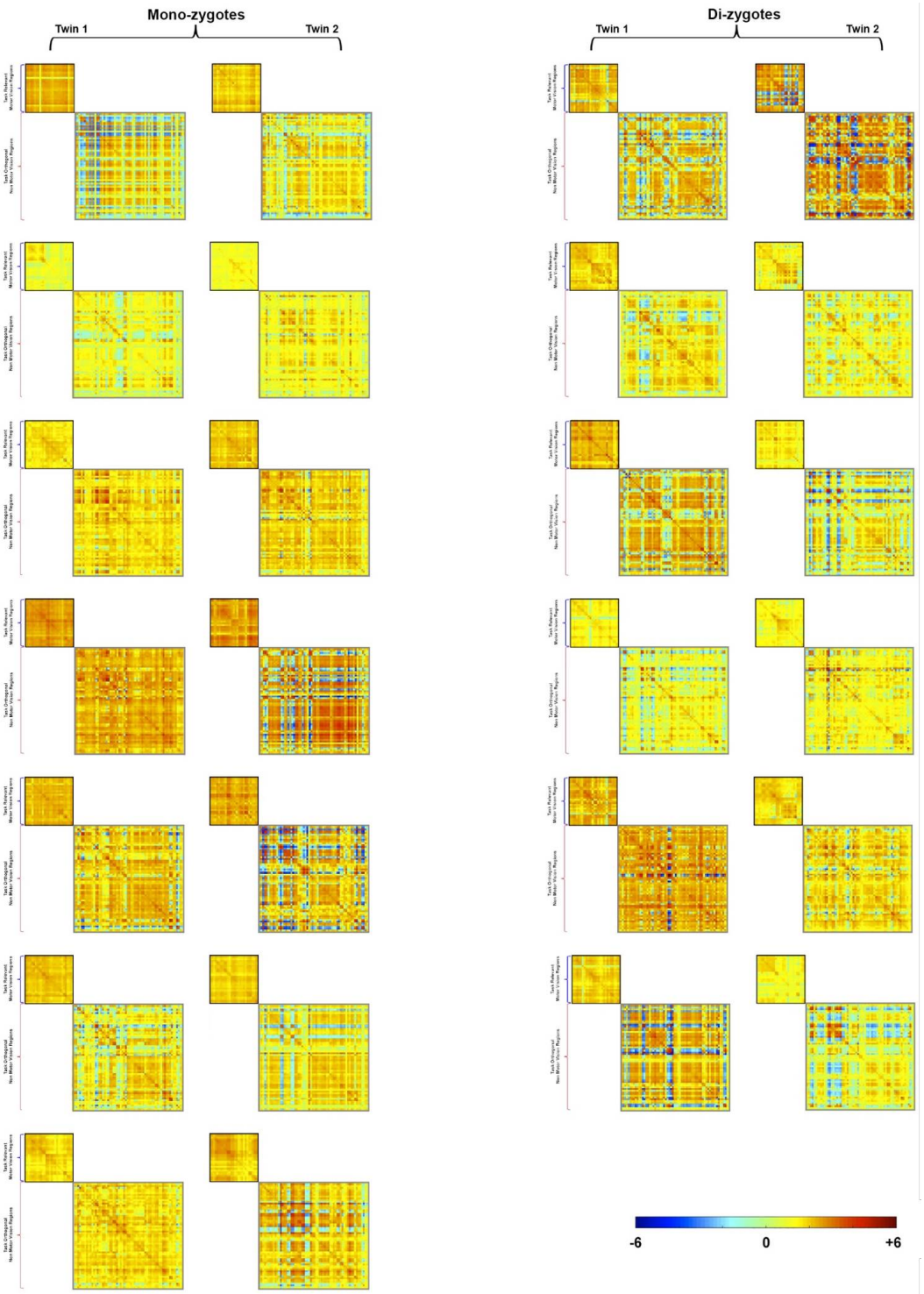

Figure 3.4: The Covariance matrices of all 26 subjects, for each of the monozygotic (left) and dizygotic (right) twin pairs are shown. The color scale represents the covariance between each of the sub-network pairs. Visual inspection suggests greater similarity in covariance between mono-zygotic twins, and particularly within task-relevant (MV) sub-networks.

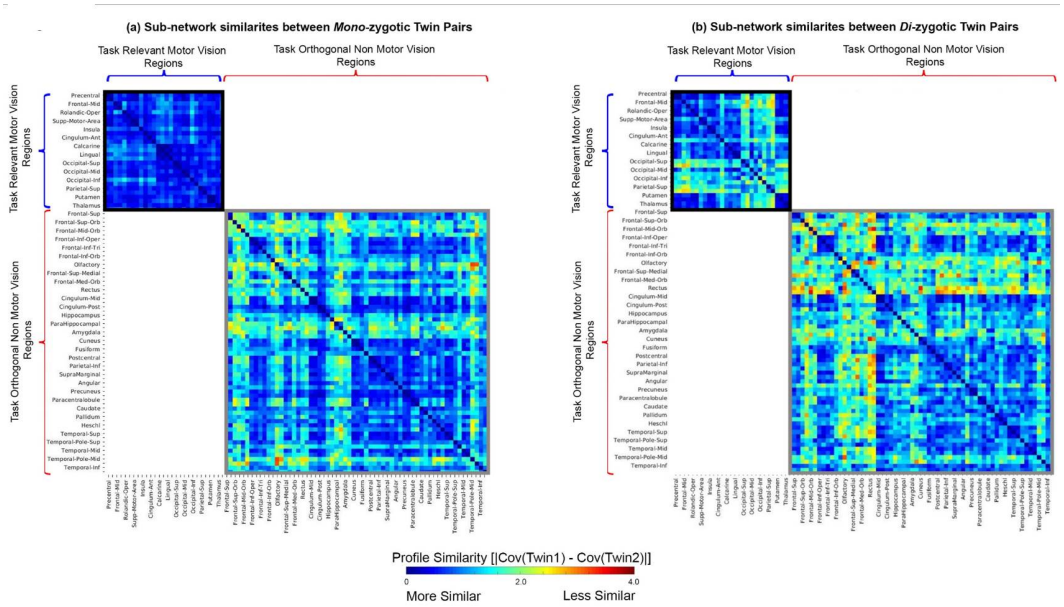

Figure 3.5: The symmetric heat maps summarize similarities in the cross-covariance matrices (See Figure 3.4) across all pairs of mono-zygotic twins (a) and di-zygotic twins (b). The heat maps were obtained after first computing for each twin pair the absolute difference in covariance matrices, and then, computing the difference in the geodesic mean across twin pairs. This was done separately for the monozygotic (a) and dizygotic group (b). The greater similarity between twin covariance matrices is represented by a lower value on this metric (coded by cooler colours). Mono-zygotic pairs are characterized by greater similarity across the network space overall, with highly salient effects in the task-relevant sub-network (black inset). These data were submitted for subsequent manifold mapping to optimally reduce the comparison between network profiles to a single scalar distance metric, hence facilitating the assessment of global differences between groups.

(NMV) sub-matrices, and b) between mono-zygotic twins, relative to di-zygotic twins.

These visual intuitions are formalized in Figure 3.5. We calculated the similarities of the covariance matrix elements between twins of each pair, and then averaged the results (using the geodesic mean, Eq. 2.9) for each group (i.e. the mono- and di-zygotic groups). For each twin pair, the similarity between their covariance matrices (from Figure 3.4) was computed as a simple absolute difference. The (geodesic) mean similarity was then computed for each group (i.e. the mono- and di-zygotic groups) as shown on Figure 3.5. Greater similarity (i.e., the smaller absolute difference in covariances between pairs of twins) is represented by a lower value on this metric (cooler colours). As seen, mono-zygotic twins show greater similarity across space in general, with highly salient effects in the task-relevant (MV) subspace.

To evaluate the similarity of whole covariance matrices rather than matrix element by element, the same data were utilized for subsequent manifold mapping (see subsection 3.2.3). In brief, manifold mapping allows summarization of network profiles (based on covariance matrices) into a single point on a geometrical manifold using geodesic distance in-

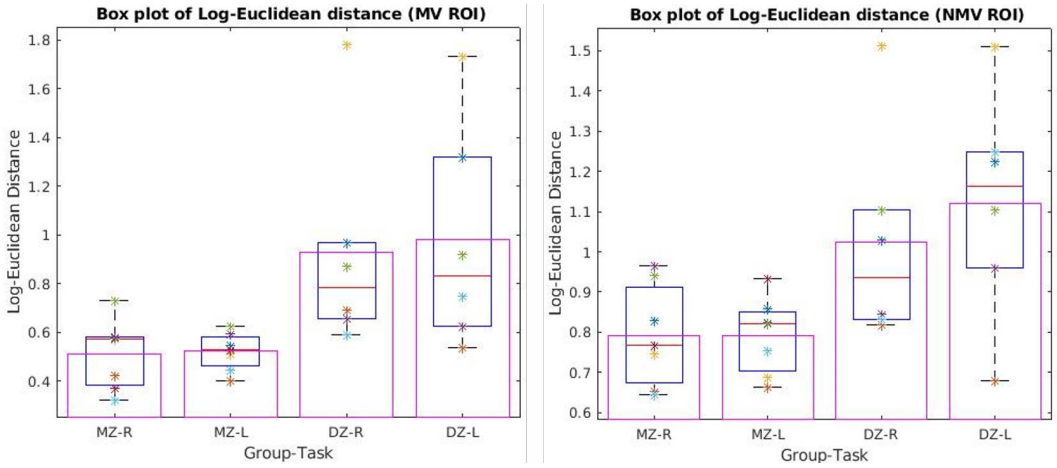

Figure 3.6: Boxplot's represent the distribution of Log-E distance of each twin pair of each group (Monozygotic (MZ) vs Dizygotic (DZ)) and for given task (R=Right and L=Left) and for task-relevant (MV) (left plot) and task-orthogonal (NMV) (right plot)

stead of Euclidean distance computations. The distance between points on the manifold (Log-Euclidean distance in our case) provides an efficient scalar description of the similarity between network profiles for a given pair of twins subjects (e.g., $MZ_{1(\text{Task Relevant})} \longleftrightarrow MZ_{2(\text{Task Relevant})}$). Moreover, geodesic calculations allow computing the mean of a set of covariance matrices more optimally than a simple Euclidean mean. The geodesic mean (Eq. 2.9) was therefore used to compute the mean similarity for each group as shown in Figure 3.5.

3.4.2 Manifold Mapping Analysis of Twin Profiles

To perform the Analyses of Variance (see Methods), the unit of analyses was the Log-E distance between each twin pair. Thus, in initial analyses for each Network of Interest (**NoI**: Task Relevant vs. Task Orthogonal), Log-E distance between pairs was estimated separately for each **Response Hand** (Right or Left). In these initial analyses, each pair contributed four Log-E distance's to the overall analyses of variance (52 Log-E distance's estimated across pairs). This approach permitted examination of the effects of three distinct factors and their potential interactions.

These Log-E distance's data were submitted to a mixed repeated measures analyses of variance and illustrated in Figure 3.6 in the form of boxplot for task relevant (left) and task-orthogonal (right) for each group (Mono- vs. Di-zygotic) and given task (Right or Left). In the 2nd level model, Network of Interest and Response Hand were modelled as repeated measures factors, and **Zygosity** (Mono- vs. Di-zygotic) as a between pairs factor. In these analyses, we observed a highly significant effect of **NoI**, $F_{1,11} = 25.31, p < .0001, MSe = .03$, with a very large effect size (Partial η^2=.697). A second significant effect was the ef-

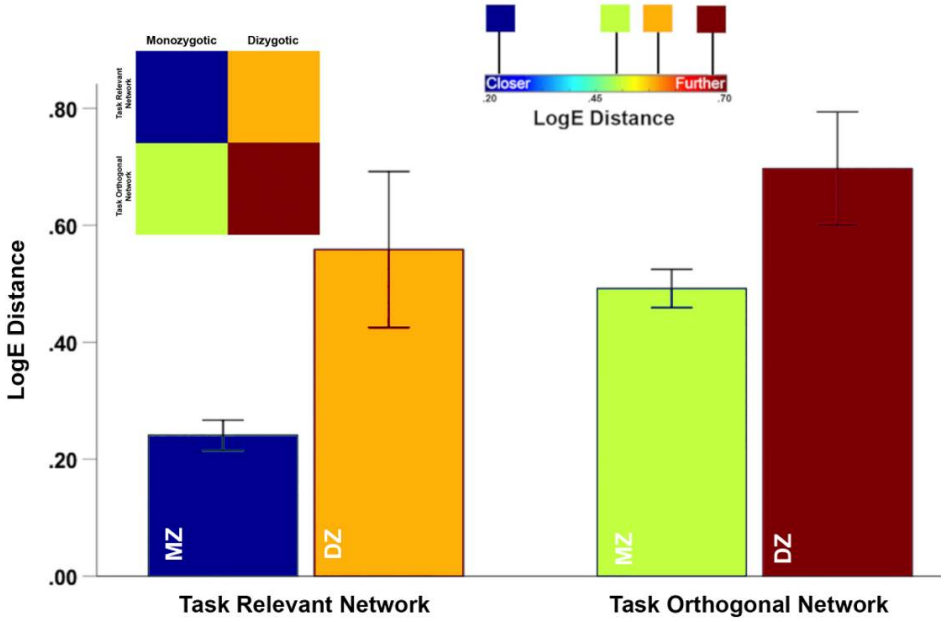

Figure 3.7: The heatmap represents the mean LogE Distance between twin pairs (columns) estimated in each of the Task-Relevant and Task-Orthogonal networks (rows). The data are re-expressed for each group and network in the adjoining bar graph (Error bars are $\pm$ SEM). The bars are colour-coded for consistency with the corresponding cells in the heatmap. As seen, smaller LogE distances between pairs were observed in task-relevant networks, relative to task-orthogonal networks. This indicates that network profiles in task-relevant networks mapped to locations that were closer in manifold space, than network profiles in task-orthogonal networks. In addition, the effect of Zygosity was significant (see text). As can be seen in the bar graphs, network profiles in monozygotic (MZ) twin pairs mapped to locations closer in manifold space, than network profiles in dizygotic (DZ) twin pairs. Finally, the Zygosity x Nol interaction was marginally significant. As can be inferred from the graphs, this interaction resulted from the similarity between monozygotic twins compared to dizygotic twins being higher in Task Relevant, compared to Task Orthogonal networks.

fect of **Zygosity**, $F_{1,11} = 7.09, p < .02, MSe = .29$, with a moderate effect size (Partial η^2=.39). Notably, none of the effects associated with **Response Hand** (main effect or any of the 2- or 3-way interactions involving this variable) were significant (.79 $\leq$ F $\leq$ 1.12) and were characterized by small effect sizes (.02 $\leq$ Partial $\eta^2 \leq$ .09).

Following this clarification of effects, Log-E distance was re-computed by averaging (using the geodesic mean) across the covariance matrices associated with the **Response Hand**. Figure 3.7, represents the results of the analyses of the Log-E distance for each of the task-relevant and task-orthogonal sub-networks. The factorial nature of the analyses (Group x Sub-Network) is represented in the heat map (top left) where the colours code the mean Log-E distance (see accompanying colour bar) between pairs of mono- or di-zygotic twins. The data are re-expressed for each group and network in the adjoining bar graph (Error bars are $\pm$ SEM). The colour coding is maintained for consistency.

In these summative analyses, each twin pair contributed two Log-E distance (one for each

Repeated Measure Two-Way ANOVA		
Source	Log-E	Euclidean
Zygocity	0.029100	0.0943
NoI	0.000026	0.0056
Zygocity x NoI	0.045000	0.7107

Table 3.1: Result of Repeated Measure Two-Way ANOVA for comparing Zygocity (MZ vs DZ), sub-networks (NoI) (MV & NMV) and two way affect.

Non-Parametric Paired t-tests P Values				
	Covariance		Graph Laplacian	
Pair	Log-E	Euclidean	Wasserstein	Euclidean
DZ-MV - MZ-MV	0.00081	0.1375	0.00031	0.0526
DZ-NMV - MZ-NMV	0.1014	0.1806	0.11792	0.6307
DZ - MZ	0.0126	0.0526	0.00014	0.0795

Table 3.2: Result of non-parametric paired t-tests (Wilcoxon) for comparing sub-networks (NoI) (MV & NMV) of two groups (MZ & DZ), for Covariance matrices with Log-E and Euclidean distances and for Graph Laplacian with Wasserstein and Euclidean distances.

NoI) to the overall mixed repeated measures analyses of variance with two factors, **Zygosity**, and **NoI** (again modeled as the repeated measures factor, Table 3.1). The main effect of **Zygosity** was significant, $F_{1,11} = 5.85, p < .03, MSe = .076$, with a moderate effect size (Partial η^2=.35). The main effect of **NoI** was also significant,$F_{1,11} = 47.5, p < .0001, MSe = .005$, with a very large effect size (Partial η^2=.89). Finally, and notably, the interaction between **Zygosity** and NoI was significant, $F_{1,11} = 3.97, p < .05$, one-tailed, $MSe = .005$, with a moderate effect size (Partial η^2=.28).

For the sake of comparison, we have also run the statistical tests on the data acquired using the Euclidean distance. Table 3.1 show the results of mixed repeated measure analysis of variance obtained with Log-E distance and Euclidean distance and Table 3.2 shows the results for non-parametric paired t-tests (Wilcoxon). It can be seen that, a highly significant effect between two groups of (MZ vs DZ) in task-relevant (visuomotor, MV) network has been observed when analyzing covariance-based functional connectome using Log-E distance and also with the Wasserstein distance on graph Laplacian. Whereas for task-orthogonal (non-visuomotor, NMV) there is no significant effect observed with any metric. A significant effect can also be seen when comparing overall groups (MZ vs DZ) with Log-E and Wasserstein distance. Whereas with Euclidean distance no significant effect has detected. These findings strongly support our hypothesis that genetics mediation has an affect on brain FC, which can be analyzed only when using correct representation of data (manifold mapping) along with an appropriate metric to compute similarity on the manifold.

3.4.3 Heritability Analysis (H_0)

Traditional estimates of heritability derived from mono- and di-zygotic twin designs rely on Falconer's formula [86] and is defined as

$$H_0 = 2(r_{MZ} - r_{DZ}) \tag{3.3}$$

H_0 is a measure designed to estimate the respective contributions of genetics and the environment to alteration in any trait, where heritability of any trait is based on the difference between trait correlations in mono-zygotic twins (r_{MZ}) compare to dizygotic-twins (r_{DZ}) (where it is expected that for an inheritable trait the correlations in mono-zygotic twins is higher than in dizygotic twins). From the mass univariate data (406 matrix elements from the 28 cerebral regions in the task-relevant motor vision network; 1953 elements from the 62 cerebral regions in the task-orthogonal non-motor vision network), the heritability of all these matrix elements were computed. To achieve this, first, the correlations of covariances were computed (from the data presented in Figure 3.4) from the set of mono- and di-zygotic twin pairs.

The resultant maps (Figure 3.8.a, first and second columns) encode the network (i.e., trait) similarities within each group of twins (MZ and DZ, see colour bar at the bottom). From these correlations (uncorrected for significance across the mass univariate space), we computed H_0 for each matrix element in each sub-network. As seen in the resultant H_0 maps (Figure 3.8.a, third column), higher heritability values are observed across the task-relevant motor vision, than the task-orthogonal sub-network (see the colour bar at right). These effects are formalized in subsequent statistical analyses (below and Figure 3.11).

Given the mass univariate nature of these analyses, additional H_0 analyses were restricted only to matrix elements that survived Bonferroni correction. To achieve this, non-significant correlations (null hypothesis of $r = 0$) from each of the MZ and DZ (Figure 3.8.a) were filtered out ($p < .05$, Bonferroni). The Bonferroni corrected probability maps for each of the MZ and DZ groups (Figure 3.8.b, top row) represent only those cells with significant correlations. The intersection of these p-maps (MZ $\cap$ DZ, Figure 3.8.b, bottom row) was used as a statistical filter. Figure 3.8.c, represents the probabilistically filtered matrix elements from Figure. The resultant H_0 map (Figure 3.8.c, third column) exhibited similar features than the uncorrected analyses, with higher heritability observed in the task-relevant, compared to the task-orthogonal sub-network. The estimated mean heritability within motor vision networks was expectedly greater (0.97) than in non-motor vision networks (0.52).

3.4.3.1 Extending the framework to twin data in the Human Connectome Project (HCP):

The exploration of the heritability of network profiles was extended to an independent sample of fMRI data collected in 21 pairs of mono and di-zygotic twins in the HCP dataset [87, 97]. The aim was to expand this framework to an independent data set that would permit assessment of H_0 in tasks with degrading levels of similarity to our original model. Accordingly, two distinct task domains in the set were chosen satisfying the criteria for level of similarity (and dissimilarity) with the originally employed paradigm: a) A visually guided motor mapping task(henceforth HCPMotor)[74]. As with the initially employed task, stimulus presentation during the HCPMotor was used to evoke effector responses (hand, foot or tongue) and is expected to have some overlap with the relative specificity of the assigned sub-networks; b) a standard working memory task (with verbal and visual stimuli; henceforth HCPWM) with extensive sub-network engagement not pinched to our a priori motor-vision network [98]. The pipeline for estimating H_0 (initially with uncorrected data, followed by Bonferroni correction) was identical to that presented in Figure 3.8.

Analyses for HCPMotor are presented in Figure 3.9 and for HCPWM in Figure 3.10. Visual inspection of the H_0 maps in each figure are suggestive of a graded effect of heritability in task-relevant sub-networks (H_0:HCPMotor > H_0:HCPWM). Within the HCPMotor task (Figure 3.9), visual inspection suggests that H_0 within task-relevant sub-networks was greater than task-orthogonal sub-networks.

We pursue to formally investigate the H_0 data across Figure 3.8, 3.9 & 3.10 within a single unified parametric analysis. The goal of this analysis was to assess the effects of the task (Visuo-motor, HCPMotor and HCPWM) and each of the sub-networks (Task-Relevant vs. Task-Orthogonal) on H_0. To achieve this, each matrix element was treated as the primary unit of analyses, and the estimated H_0 data were submitted to a two-factor mixed analysis of variance. The task was modelled as a within-units factor (each matrix element had an H_0 estimate from each of the three tasks) and sub-network was modelled as a between-units factor (each matrix element was uniquely assigned to one or the other sub-network).

The results of this analysis of variance was unequivocal and are illustrated in Figure 3.11. A significant main effect of task was evident, $F_{2,4714} = 470.67, p < .0001, MSe = .606$, with a medium effect size (Partial η^2=.167)(Figure 3.11.a). Pairwise comparisons (based on LSD) indicated a graded effect of Task on H_0 (Visuo-motor > HCPMotor > HCPWM). The main effect of Sub-Network was also significant, $F_{1,2357} = 114.34, p < .0001, MSe = .586$, with a small effect size (Partial η^2=.046) (Figure 3.11.b). Finally, we observed a significant interaction, $F_{2,4714} = 32.71, p < .0001, MSe = .606$, with a small effect size (Partial η^2=.014) (Figure 3.11.c).

3.4.4 Discussion

In this work, we proposed a computational framework to analyze the influence of genetics on twins in term of similarity between brain functional network profile. To this aim, we exploit the property of functional connectome to be symmetric and positive definite which allows analyzing these connectomes on Riemannian manifold using the notion of geodesic distance along the curves of manifold instead of a vector space representation with the euclidean distance which is suboptimal to use on such data. For a better comparison of analyzing genetic influence, we implemented this framework using two different approaches of connectome representation and with two different geodesic metric.

In the first approach (section 3.3), the similarity between covariance-based functional connectome of twins pairs is being computed using Log-E distance. One advantage of this method is that we used all values of connectome (including negative and positive covariances) for this analysis. Hence a combination of these SPD matrices forms a wider Riemannian manifold which allows an appropriate analysis of brain connectivity. Statistical tests (Table 3.2) suggested that, in the task-relevant regions, functional connectivity of monozygotic twins is more similar as compared to dizygotic twins (the difference is statistically significant) when analyzed with Log-E distance as compared to the Euclidean distance. Whereas for the task-orthogonal regions there was no significant difference between two groups of twins.

In the second approach (subsection 3.3.1), graph Laplacian of correlation-based functional connectome (positive values only) were used along with Wasserstein distance to compute the similarity between brain connectivity profiles. The advantage of Wasserstein distance is that it can also be used for symmetric and positive semi-definite matrices. Statistical tests results (Table 3.2) suggested that monozygotic twins are more similar in task-relevant regions as compared to dizygotic twins (the difference is statistically significant), only when analyzed with Wasserstein distance as compared to Euclidean distance.

To evaluate the plausibility of this framework, we relied on a combination of a) multiple analytic methods that could efficiently summarize profiles of large brain networks, applied these to b) task-based fMRI studies where task-loading on the brain networks varied in a priori known ways and where c) the data were collected in independent sets of mono- and di-zygotic twin samples.

Our salient results were: 1) Using manifold mapping (Figure 3.4 & 3.5) to compute similarities in network profiles between mono- and di-zygotic twins, we demonstrated that profiles in mono-zygotic (compared to di-zygotic) twins were significantly more similar in visuo-spatial task-relevant, than in task-orthogonal sub-networks (Figure 3.7)[88]. We demon-

strated, that these computations of manifold mapping were independent of differences in sub-network size. 2) These effects were corroborated computing mean H_0 across regions within task-relevant and task-orthogonal networks (Figure 3.8) before and after statistical correction of correlation coefficients. 3) In an independent data set of mono- and di-zygotic twins, we confirmed that H_0 was higher in MZ twins for task-relevant sub-networks during tasks with strong visual-spatial compared to working memory demands (Figure 3.9 & 3.10). 4) Finally, we collated the analyses of H_0 (Figure 3.11) and using an omnibus statistical analysis, confirmed contextual effects on the heritability of brain network profiles.

These results obtained with both approaches, clearly show the influence of genetics and we can conclude that the brain network of monozygotic twins is genetically and functionally more similar as compared to dizygotic twins. The results of this work also clearly demonstrate that use of Euclidean distance is not the best choice, as it is not properly managing the complex structure of graphs, indeed the similarity effect is more visible when using the geodesic distance. This allowed us to discover scientifically relevant questions related to genetics, and its impact on brain network function e.g. analysis of heritability index. In subsection 3.4.3 we presented our approach to explore the heritability in twins data by using Falconer's formula. Heritability values obtained in monozygotic twins are significantly higher as compared to dizygotic twins specifically in task-relevant regions.

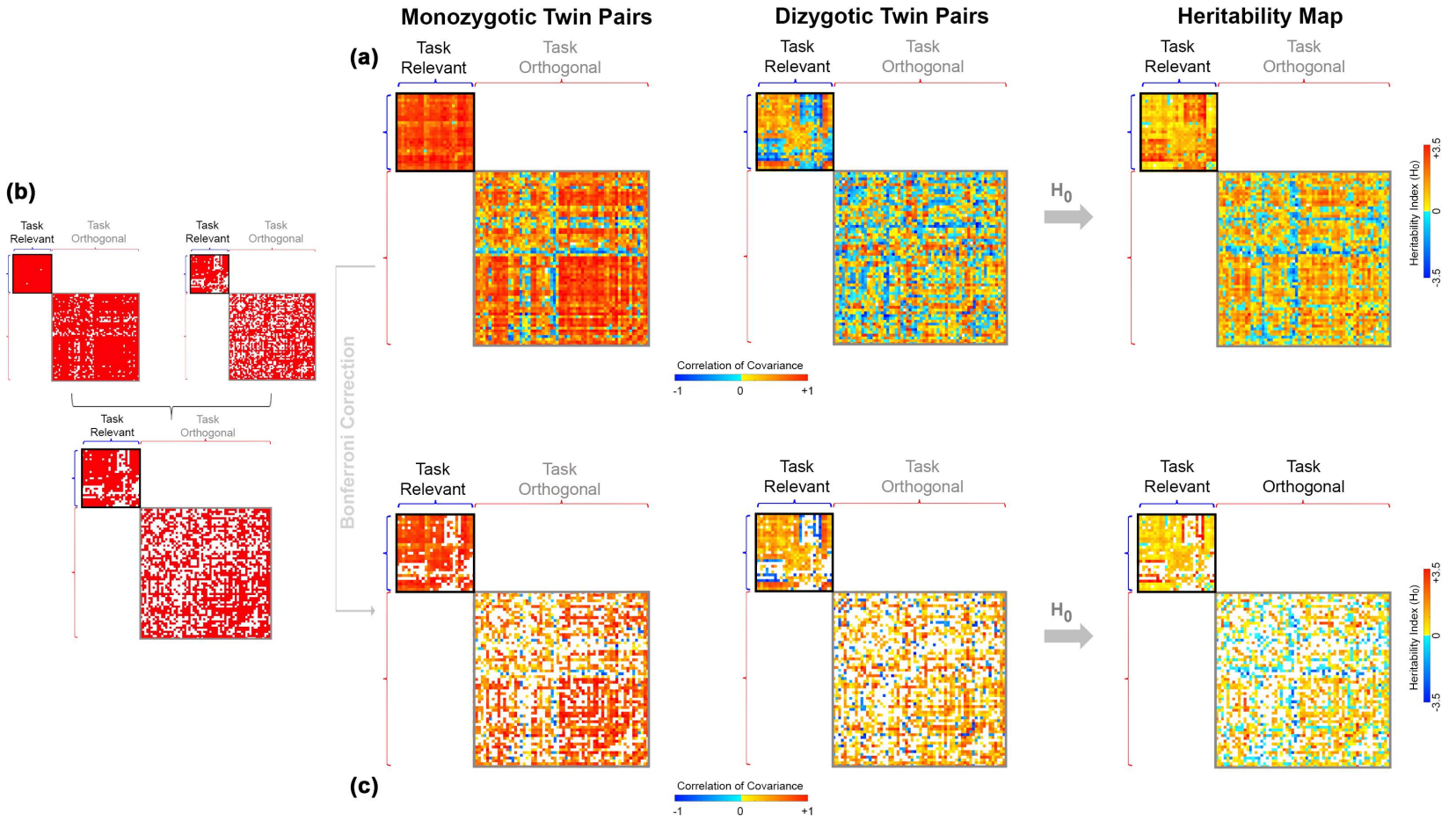

Figure 3.8: Heritability estimates (based on Falconer's formula) were derived in task-relevant motor vision and task-orthogonal non-motor vision sub-networks. (a) Initial computations of correlations of covariance were computed for each group of twin pairs (MZ and DZ). Monozygotic twin pairs are characterized by an increase in correlations, notably within task-relevant sub-network pairs (black inset). By comparison, the correlations observed in Dizygotic pairs is lower and more variable. Note that correlation coefficients (r) when treated as "veridical" measures of similarities in network profiles are uncorrected for significance, and were forwarded for computation of heritability, H_0 (based on Falconer's formula). Mean heritability within motor vision networks was expectedly greater than in non-motor vision networks (1.11 vs. 0.62). (b) To filter out non-significant correlations, and to ensure that H_0 was only computed for correlations significant in both Mono- and Dizygotic twin pairs, the coefficients in (a) were corrected for multiple comparisons (Bonferroni). The resultant probability maps depict only significant correlations in each class of twin pair, feeding into a final intersection (MZ $\cap$ DZ) map. (c) Adjacency matrices from (a) were masked by the intersection matrix in (b). Note the loss of large spurious negative correlations in Dizygotic twins. The resultant H_0 map is a subset of that depicted in (a) and maintains evidence of greater average heritability in motor vision relative to non-motor vision networks (0.97 vs. 0.53) even when H_0 is restricted to only significant correlation coefficients

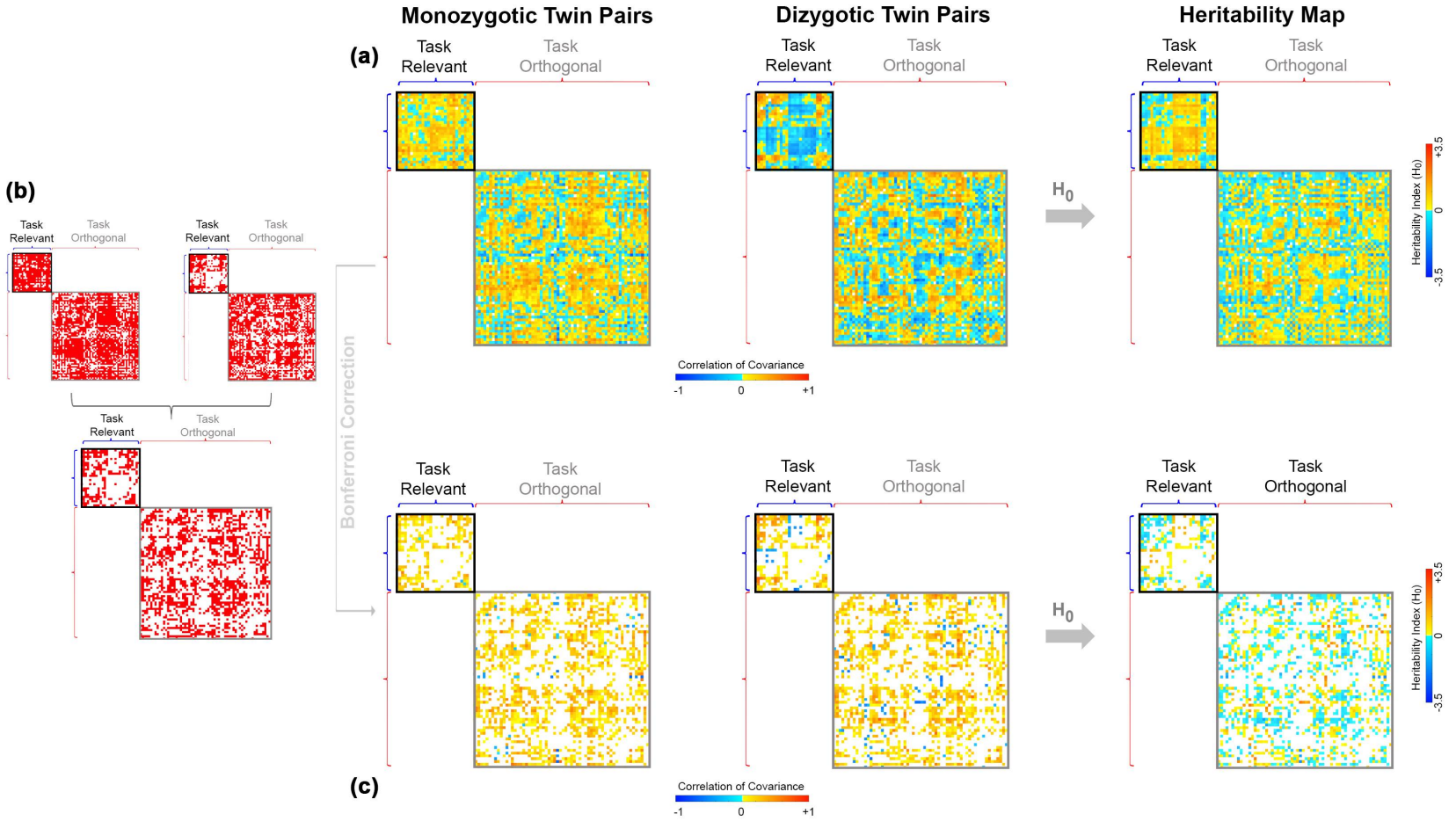

Figure 3.9: Heritability estimates, H_0 (based on Falconer's formula) derived from the application of the motor task to mono- and di-zygotic twins in the HCP data (HCPMotor, see Results) are presented in an identical format to Figure 3.8 and derived using an identical pipeline. There remains some evidence for the contextual heritability of network profiles, even though the sub-networks were not derived from this task. Thus, H_0 within the task-relevant sub-network remained higher than for the task-orthogonal sub-network for both (a) uncorrected (0.15 vs. 0.06) and (c) Bonferroni corrected (0.05 vs. 0.04) analyses.

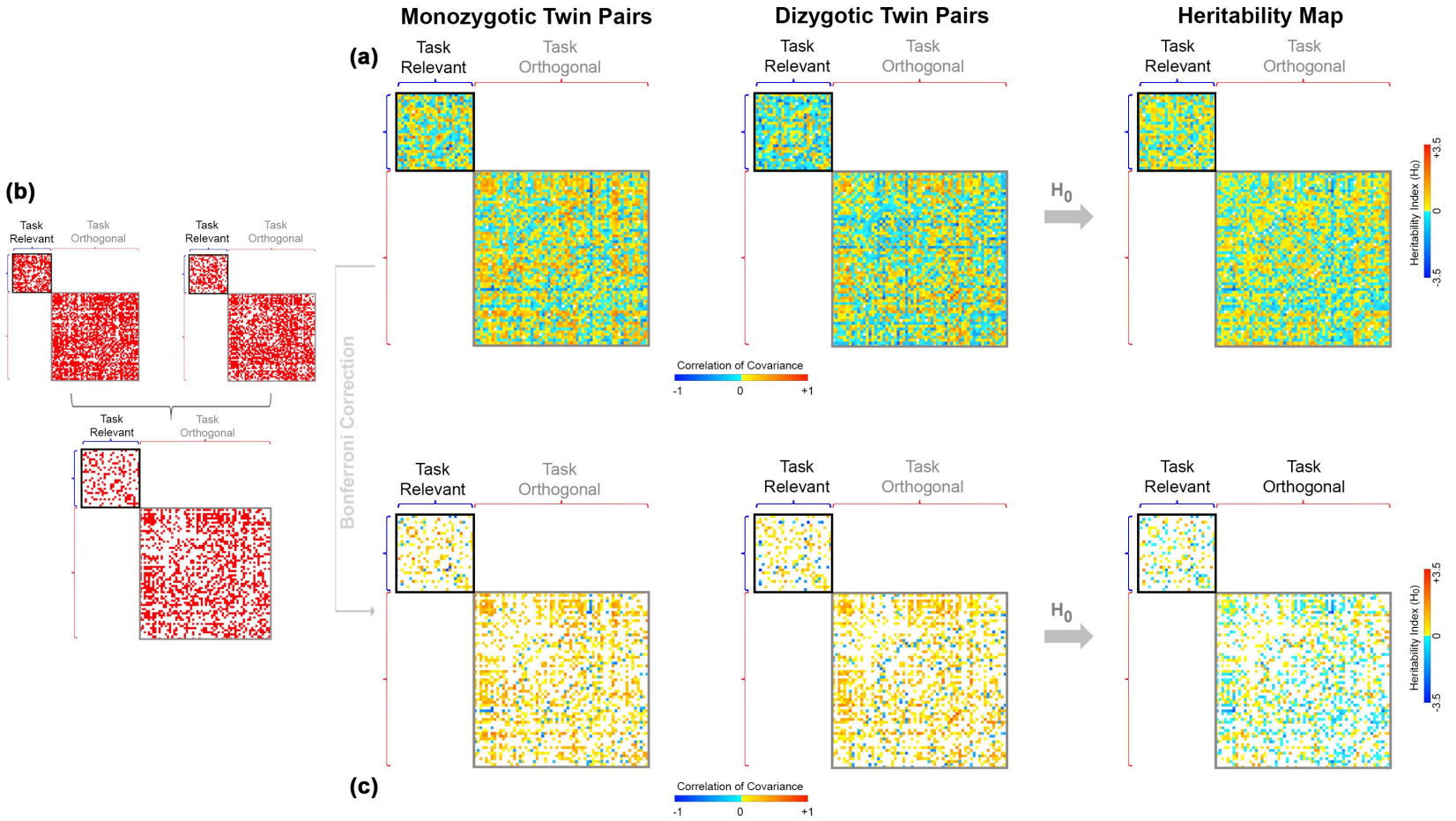

Figure 3.10: Heritability estimates, H_0 (based on Falconer's formula) derived from the application of the working memory task to mono- and di-zygotic twins in the HCP data (HCPWM)are presented in an identical format to Figure 3.8 & 3.9, and derived using an identical pipeline. A feature of these results, is further evidence for the contextualized heritability of brain network profiles. As seen (and in comparison, to Figure 3.8 & 3.9), here we did not observe differences in H_0 across the sub-networks either for the uncorrected (0.1 vs 0.07) or the corrected (0.1 vs. 0.05) data. H_0 values across each of Figure 3.8 - 3.10 compared in a subsequent omnibus analysis (Figure 3.11) accentuate evidence for contextual inheritance.

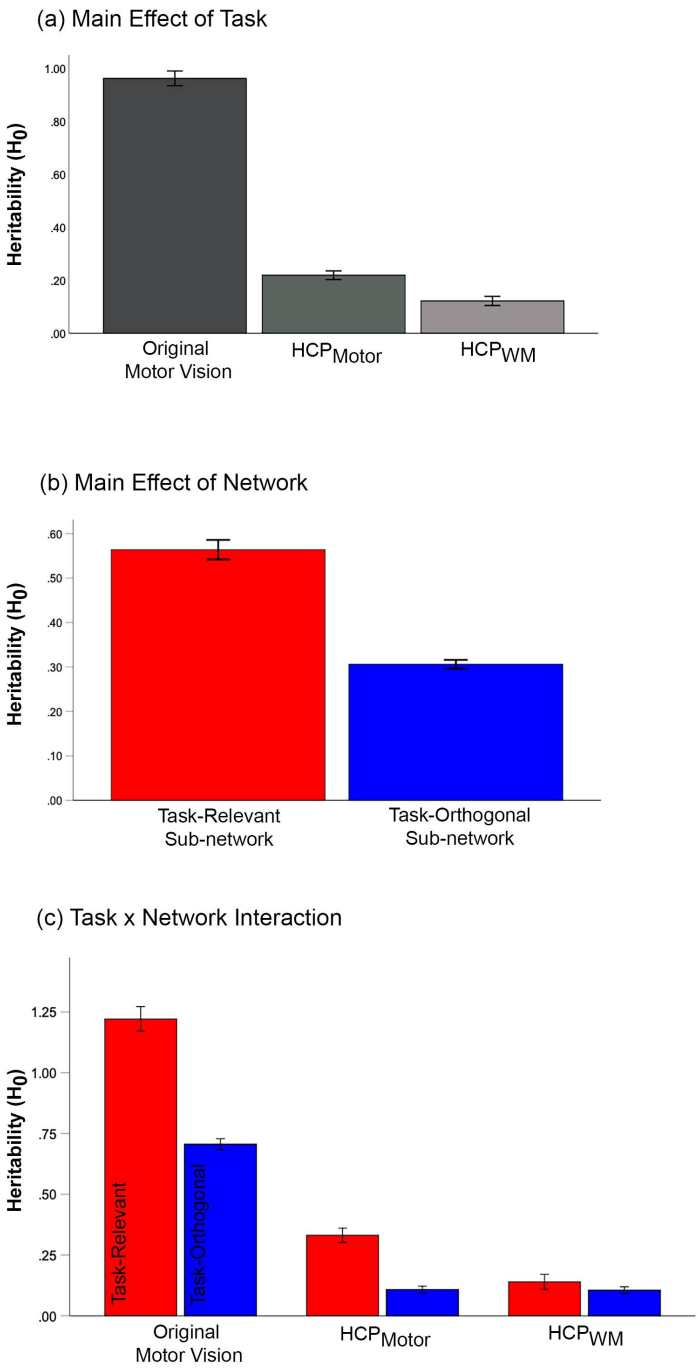

Figure 3.11: We sought to formally investigate the H_0 data across Figure 3.8 - 3.10 within a single unified omnibus repeated measures parametric analysis. Here we assessed the effects of the task (Visuo-motor, HCPMotor and HCPWM) and each of the sub-networks (Task-Relevant vs. Task-Orthogonal) on H_0. Because each cell in the adjacency matrices was treated as the primary unit of analyses, data from the uncorrected estimates of H_0 were used. The task was modelled as a within-units factor (each cell had an H_0 estimate from each of the three tasks) and sub-network was modelled as a between-units factor (each cell was uniquely assigned to one or the other sub-network). The bar graphs represent the significant effects of (a) Task, (b) Network and c) the interaction of Task x Network. As is evident, the original task from which network classification was derived expectedly exerted the strongest effects on H_0 (a). Regardless of the task, the effects were strongly driven by the network (b). Finally, the relative effects of network on H_0 were strongest for the original motor-vision task, in less evidence for HCPMotor, and non-existent for HCPWM (c).

Chapter 4

Geodesic Clustering of SPD Matrices for Data Representation

4.1 Introduction

An important task in functional connectome studies aims at classifying pathological states from healthy controls (HC). Traditionally, univariate statistical hypothesis testing on connectivity graphs focus on analyzing changes in the functional connectivity. Nowadays FC plays an important role to characterize brain connectivity in many psychiatric and neurodegenerative disorders. With the growing importance of machine learning approaches, the task of classification in neuroimaging took the name of "neuroimaging data encoding" and it has been proved in various studies that it is possible to exploit more precise relationships compared to standard statistical approaches.

From a methodological perspective, an adequate classification algorithm should be able to discriminate as much as possible between samples belonging to different classes. In terms of classification, there are two main lines to obtain good performances: the development of new algorithms which use standard features or the extraction of novel features to embed in a richer feature space to be used as input of a standard classifier, for example, *Support Vector Machine* (SVM), *k-nearest neighbour* (k-NN), *Linear Discriminant Analysis* (LDA) or *Random Forrest* (RF).

Many classification studies on brain disorders affecting connectivity such as schizophrenia, autism or Alzheimer have been carried out and most of the time a unique description of the features has been preferred rather than the evolution of a new classifier. Support Vector Machine [99] is considerably used in neuroimaging for classification of mental disorder, often combined with dimensionality reduction or feature selection algorithms. For example, in [100, 101] have performed classification of functional graphs using SVM classifier and the vectorization of functional connectivity matrices. However, since a given brain graph with N nodes has $N(N-1)/2$ connections, authors also come up with feature selection strategy to decrease the number of connections, focusing at extracting only the most appropriate. [102] and [103] have also adopted SVM classifier along with an univariate analysis as feature

selection.

To reduce the number of features, sparsity techniques have been also used. An example is *L1-norm* on the connections which force sparsity defining a Gaussian graphical model [104]. In [105], authors proposed a combination of SVM classifier and the sparse inverse covariance matrix to classify the mental disorder. Another novel technique is the identification and classification of "sparse" networks using matrix decomposition/factorization focusing at describing the connectivity matrix as an amalgamation of sub-patterns [106]. The resultant framework, when combined with SVM, gives promising results which are comparable to traditional approaches but with much fewer features.

Dimensionality reduction algorithms have also been introduced to attenuate the high dimensionality of the connectomes data. In [107], Local Linear Embedding (LLE) has been used to express fMRI data in three dimensions over which, *k-NN* classifier has been adopted. In [108], an alternative method for brain connection discrimination was presented, which relies on classification trees (an RF version) and learns a discriminative weight w_i for each connection and then it fit a hyperplane using the most discriminative feature. The algorithm learns the most discriminative feature through a statistical step of feature selection, defining complex and non-linear boundaries.

Different methods have been proposed to classify groups of subjects using the geometrical properties of symmetric positive definite (SPD) matrices. The set of all SPD matrices of the same size forms a Riemannian manifold, so several approaches have been developed to leverage this manifold structure during the analysis. In [43], a probabilistic model for covariance matrices was used to distinguish post-stroke patients from HC. In [85] manifold transportation of covariance matrices was applied in longitudinal studies to determine changes in FC after a task. In [45] a kernel-based classification approach has been deployed which analyzed the FC matrices using Log-Euclidean Gaussian kernel and Stein Gaussian kernels. In [6], an approach based on Grassmannian geometry and low-rank graph Laplacian has been used for a classification task exploiting a set of sub-networks that was then used to identify connectivity biomarkers. Correctly taking into account the properties of positive semi-definite matrices allowed to classify sub-connectivity patterns [46], functional states generated from auditory stimuli [47] or mild cognitive impairment [48].

In this work, we employ a geodesic clustering algorithm which uses geodesic metrics on a Riemannian manifold to cluster FC matrices. The computed centroids are then used to generate a representation allowing to discriminate between classes. More specifically, using a two-fold approach, functional connectivity matrices of brain activity during rest are clustered in a “predefined” number of clusters and in a second step, the geodesic distances of the

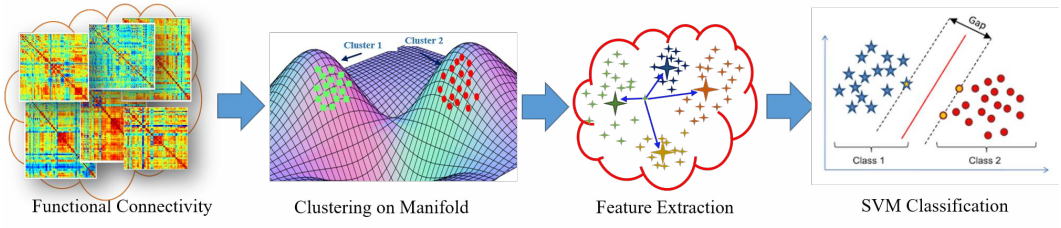

Figure 4.1: The pipeline of proposed method start with clustering of FC matrices on manifold, followed by feature vector extraction in term of distance of each FC from each centroid and then training and testing of SVM based classifier

connectivity matrices from the cluster centroids are used as features to train a linear-SVM. The proposed method has been tested on two different problems: HC vs. subjects affected by Autism Spectrum Disorder (ASD), and HC vs. subjects affected by Schizophrenia (SCHZ). To show the benefit of using the Riemannian properties, the same experiments have been done using the Euclidean metrics, comparing the results in terms of both clustering and classification performance. The Figure 4.1 show the pipeline of the proposed methodology.

4.2 Methods

4.2.1 Manifold Representation of Connectivity Matrices

Let $\mathcal{X}_\rho = \{\rho_1, \ldots, \rho_n\}$ be the set of correlation matrices describing the brain functional connectivity of all N subjects. The correlation matrices are symmetric and positive semi-definite in nature and can be easily regularized into SPD matrices by adding a small constant to the main diagonal ($\rho_i = \rho_i + \lambda I$, with λ very small, e.g., $\lambda = 10^{-5}$). As described in subsection 2.3.1, the set of all SPD matrices of the same size form a Riemannian manifold, which allows the analysis of such matrices on a manifold space. To take the full advantage it is recommended to use the notion of geodesic distances which allows a description of this data better than using Euclidean metrics [45, 6].

There are several possible alternative geodesic distances on the Riemannian manifold of SPD matrices which are defined in section 2.4; we decided to adopt the Log-Euclidean (Log-E) distance, which is simple and fast to compute. Equation 2.8 and 2.9 describe, respectively, the log-E distance formula between two SPD matrices and the closed-form formula to compute the mean [60] of two or more SPD matrices with this metric.

4.2.2 Geodesic Clustering Analysis

In this proposed approach, the aim is to divide FC matrices into homogeneous groups of subjects presenting similarities in their connectivity. The underlying assumption is that there are

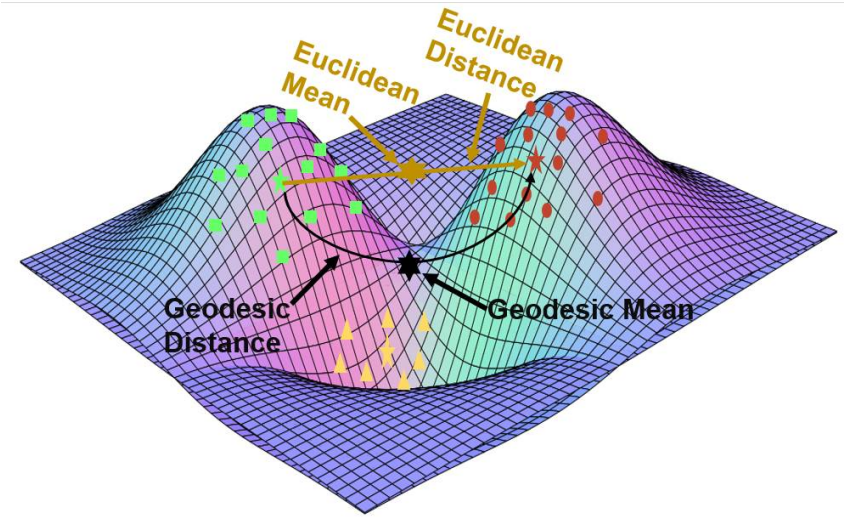

Figure 4.2: The difference between Euclidean distance and Euclidean mean of two points (brown straight line and star) on manifold and the corresponding geodesic distance and geodesic mean (black curve and star along the manifold)

some alterations in brain connections of the patients [109] that can be grasped by the organization of the cluster. To this aim, we have opted to use the ideas from the $k-means$ clustering algorithm with geodesic metric [110] to cluster the FC matrices into k disjoint clusters $\mathcal{C} = \{c_1, \ldots, c_k\}$ (different groups). The $k-means$ clustering algorithm is a popular and widely used method for clustering [111]. In dynamic FC analysis (dFC), $k-means$ has been deployed as baseline approach to estimate the set of highly replicable patterns of whole brain dFC which is also known as dFC brain states [40, 112, 113, 114, 115]. Some literature that attempt to use geometric structures, such as in [116] they consider a soft geodesic kernel $k-means$ algorithm that adapts to geodesic distance to cluster when the data have a geometric structure. In [117], proposed clustering and dimensionality reduction on Riemannian manifolds. In [118] introduce a class of geodesic distances and extend the $k-means$ clustering algorithm using this metric. In[119] use $k-means$ clustering algorithm in order to cluster the summary data of different stocks by their Realized Trading Volatility (RTV) model

In order to describe a $k-means$ clustering algorithm, one need to define a similarity measure for any two data points and centroids of observations in clusters. Since the FC matrices are SPD and define the a Riemannian structure, we can choose a similarity measure and centroids based on this structure. So, the $k-means$ clustering algorithm was implemented using the Log-E distance [45] as defined in Eq. 2.8, with the centroids computed as the geodesic mean, which can be computed in the closed-form by Eq. 2.9. Figure 4.2 shows the conceptual difference between geodesic distance and geodesic mean as compared to Eu-

clidean distance and Euclidean mean on a manifold.

In order to choose the optimal number of clusters (K) we used the *Davies-Bouldin* (DB) index as criterion [120, 121]. This index evaluates the consistency using the distance of all points within a cluster to the corresponding centroids and the separation between clusters using the distance between centroids. The lower is the index, the better are the clustering results. In this work, the DB index is computed for every considered number of clusters (i. e. K=[2,3,4,5,6]) and the minimum value suggests the natural partition of data.

Given a set of correlation matrices $\mathcal{X}_\rho = \{\rho_1, \ldots, \rho_n\}$ and a set of clusters $\mathcal{C} = \{c_1, \ldots, c_k\}$ partitioning $\mathcal{X}_\rho$ in K groups, cluster representatives are defined as

$$\overline{c_k} = \frac{1}{|c_k|} \sum_{\rho_i \in c_k} \rho_i \tag{4.1}$$

and the distance between matrices $d(\rho_i, \rho_j)$ used in our analysis is the Log-E distance. The equation for the DB index is given as follow

$$S(c_k) = \frac{1}{|c_k|} \sum_{\rho_i \in c_k} d(\rho_i, \overline{c_k}) \tag{4.2}$$

and

$$DB(\mathcal{C}) = \frac{1}{K} \sum_{c_i \in \mathcal{C}} \max_{c_j \in \mathcal{C} \setminus c_i} \left\{ \frac{S(c_i) + S(c_j)}{d(\overline{c_i}, \overline{c_j})} \right\} \tag{4.3}$$

Since the goal is to achieve minimum within-cluster dispersion and maximum between-cluster separation, the number of clusters $\mathcal{C}$ that minimizes DB index is taken as the optimal value of $\mathcal{C}$ [122].

4.3 Experiments

The working hypothesis is that we can cluster the FC matrices preserving the alteration of brain connectivity characterizing the groups. This would allow therefore compression of graphs into a smaller vectorized representation retaining the group differences while filtering the intrinsic variability of subjects in the same group. Indeed, using the cluster representatives as a dictionary, we built a vector representation for each subject, computing the features as the distances of the subject FC matrix from all cluster centroids.

4.3.1 Feature Extraction and Classification

In our experiments, we performed geodesic clustering multiple times with a variable number of clusters ranging from $K = 2$ to $K = 6$ to find the best K. Once convergence was achieved we computed the Log-E distance between the samples in the training set and all K centroids (e.g. for $K = 2$ each sample was described by 2 distance values and for $K = 4$ each sample was represented by 4 features). These distance values were used as feature vectors to train a classifier. In the test phase, each sample in the test set was described by the distances of the corresponding FC matrix from all cluster centroids computed during training. For classification purpose we opted to use linear-SVM [99] which classify the data by generating a separation model [123]. Selection of linear-SVM is based on the several facts, e.g. it works well and in-line with some other classifiers in term of accuracy and speed when the data is based on the numerical attributes [124], also SVM is based on geometrical properties of the data which makes it more suitable for the data acquired through clustering on manifold [125, 126]. According to the literature [45, 6, 101, 5], for two class classification problem, SVM also works well when there are mild number of samples and features are more sparse, and under such conditions selection of a more complex or simpler classifier can leads towards either over-fitting or under-fitting. Furthermore, unlike the traditional statistical parametric mapping approaches, SVM is a multivariate approach that combines information from multiple features for the purpose of classification [127]. So, for being widely used in the field of neuroscience with considerable performance [128, 129, 130, 131, 132], we opted systematically to use linear-SVM classifier in this work.

To avoid double-dipping we made all the experiments using 5-fold cross-validation, randomly selecting the samples and preserving the proportion between the classes in each fold. For statistical reasons, we repeated this cross-validation process 100 times with a randomized selection of folds. The Figure 4.3 illustrate the process of extracting training and testing distance vector from clustered data and feeding them to train the SVM classifier.

In the end, we evaluated the results in terms of average accuracy and confusion matrix averaging over all 100 iterations. In our experiments, all distances were computed using the Log-E distance (Eq. 2.8) and the corresponding geodesic mean (Eq. 2.9). Also, to show the advantage of using the geodesic distance on the manifold containing the data, we conducted the same experiments using Euclidean metrics, allowing them to evaluate the differences in performance. To check the significance level of the performance of our classifier we performed a permutation test on labels. For this purpose, we generated a null distribution by randomly shuffling the labels 1000 times and in each iteration, we performed L-SVM classification using 5-fold cross-validation and computed the mean cross fold accuracy.

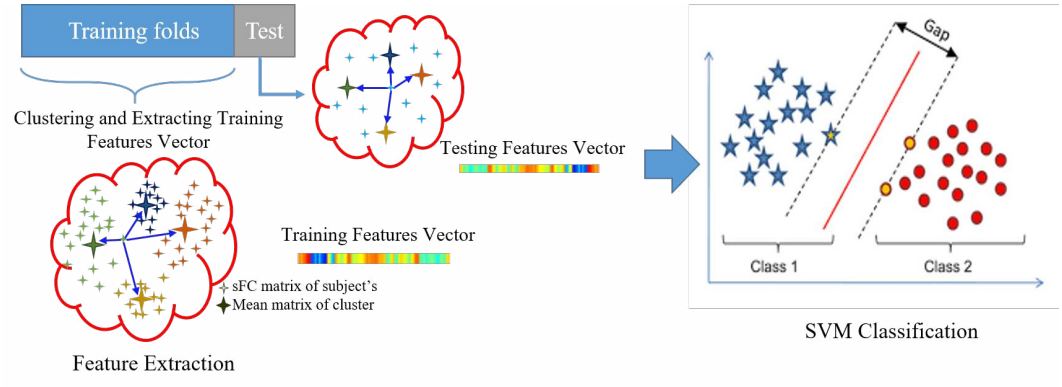

Figure 4.3: An illustration of division of FC data into training and testing using 5-folds CV and using training data to perform clustering of FC matrices into different clusters using k-means, followed by computing the distance of each training sample from each cluster centroid to encode the data into vectorial representation. whereas testing fold data is only being used to compute the test vector in term of distance from each centroid. Finally, this train and test data were used to train a linear SVM for classification purpose.

4.3.2 Dataset

To test the proposed method, we have used two publicly available functional connectivity datasets. The first dataset is from the ASD connectome database released by UCLA [133]. This dataset is composed of the rs-fMRI of 37 HC and 42 ASD patients. Further details of the acquisition and pre-processing are described in [96]. FC matrices were obtained from the Power atlas, which defines 264 regions of interest (ROIs) in the brain [134]. These 264x264 FC matrices are estimated for each subject by computing the pairwise Pearson correlation between average time-series of brain ROIs. Furthermore, we analyzed the FC dataset released by the Network-Based Statistic (NBS) toolbox. It is composed of 15 HC and 12 SCHZ subjects [135]. In this dataset, FC matrices were built using a subset of regions from the AAL atlas (90 ROIs without cerebellum) using the same pairwise Pearson's correlation approach. In our method, we are considering the whole connectivity matrix including negative values.

4.4 Results

The Figure 4.4 depicts box plots showing the classification accuracy over 100 iterations with the proposed method on the ASD dataset (Figure 4.4.A) and on the SCHZ dataset (Figure 4.4.B). Blue and orange bars represent results obtained with geodesic and Euclidean metrics respectively. The grey line shows the average over 100 iterations of the DB index for the geodesic clustering. For the ASD dataset (Figure 4.4.A) it can be seen that the highest mean accuracy (67.12%) is achieved with Log-E distance for $K = 4$ clusters, whereas with Euclidean distance the maximum obtained mean accuracy is 61.59% with $K = 6$ clusters.

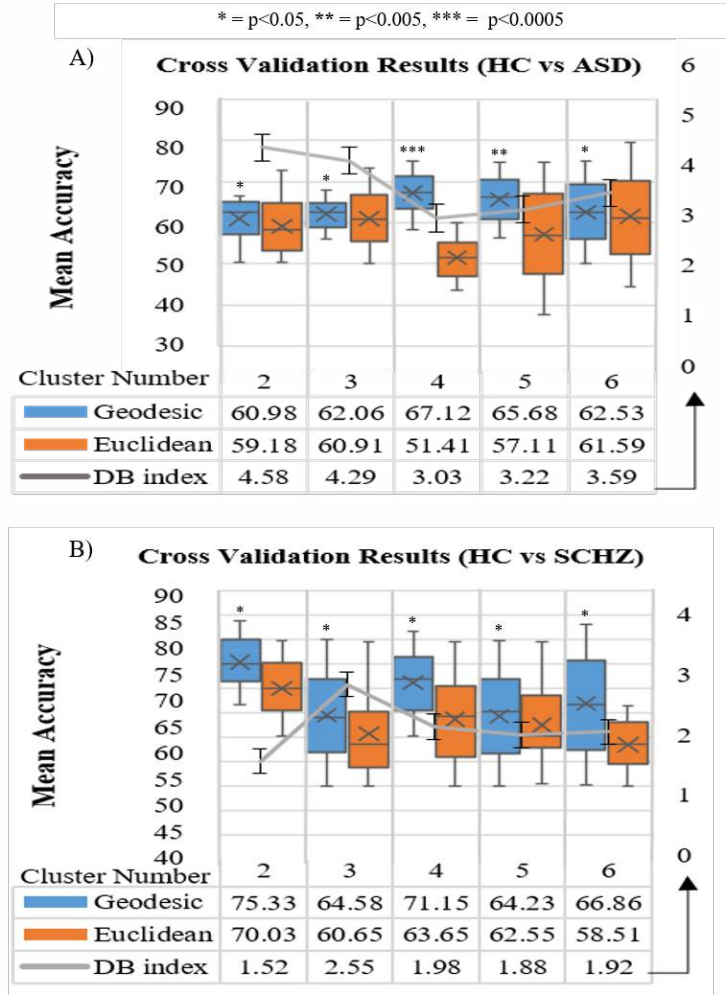

Figure 4.4: Boxplot representing the mean classification accuracy for A) HC vs ASD and B) HC vs SCHZ dataset with geodesic (blue box) and Euclidean (orange box) metrics based k-means clustering. Line plot shows the mean DB index value for each cluster of geodesic k-means clustering. selection. Stars on the bar shows the significance level obtained through permutation test.

Figure 4.4.A also shows that the DB index (line plot) has a minimum value in $K = 4$, suggesting that this is the optimal number of clusters. This is reinforced by the fact that this is the same number of clusters with peak accuracy for geodesic clustering. Results from the proposed methodology also outperform the results presented in [45, 6] using the same dataset, where achieved accuracy was 60.76% & 63.29%.

For the Schizophrenia dataset, (Figure 4.4.B) shows that, with the geodesic metric, maximum mean accuracy (75.33%) is achieved with $K = 2$. Similarly, in (Figure 4.4.B), the DB index (silver line) for this dataset also has the minimum value for $K = 2$. On the other hand, for Euclidean metrics, the maximum accuracy is achieved (70.03%), on this dataset, for $K = 4$. Results from the proposed methodology also outperform the result presented

Mean Confusion Matrix of HC vs ASD (for K=4)			
		Predicted Class	
		HC	Pathological Subjects
Actual Class	HC	22	15
	Pathological Subjects	12	30

Mean Confusion Matrix of HC vs SCHZ (for K=2)			
		Predicted Class	
		HC	Pathological Subjects
Actual Class	HC	10	5
	Pathological Subjects	2	10

Table 4.1: Confusion matrix of average classification results for the proposed approach based on geodesic clustering for HC vs. ASD and HC vs. SCHZ datasets

in [45] using the same dataset, where achieved accuracy was 74.07%. The embedded tables in both figures summarize these results. Table 4.1 shows the average confusion matrix for geodesic clustering results for both datasets between HC and pathological subjects.

In order to assess the statistical significance of our obtained results, we implemented a permutation test. The results of the permutation test are represented in form of p-values in Figure 4.4. The p-value is computed as the ratio between the number of accuracy values greater than the tested accuracy and the total number of permutations (1000 in our case). These results strongly support the principle according to which the use of geodesic metric on SPD matrices, which form a Riemannian manifold, gives better results in term of accuracy, whereas the use of Euclidean metric on SPD matrices is suboptimal.

4.5 Discussion

In this chapter, we have presented a Riemannian manifold based computational framework, which allows the classification of HC and patients using static FC matrices obtained from rs-fMRI. To achieve this goal, we performed $k-means$ clustering by taking advantage of the properties of SPD matrices: in this context, using geodesic metrics proved to be superior to the Euclidean approach.

In particular, classification features have been constructed with a subject-wise graph similarity representation by using a geodesic metric (Log-E) based on $k-means$ clustering. In order to evaluate our proposed algorithm, we made a similar experiment but using $k-means$ clustering with Euclidean metric instead of the geodesic metric. The results of this study noticeably reveal that the use of Euclidean metrics on the manifold of SPD matrices is suboptimal, as it is causing a data representation leading to decreased accuracy. Indeed, the classification performance improved when using the geodesic metric which computes the

shortest possible distance along the curvature of the manifold, thus offering an optimal data representation. Hence to compare and analyze FC SPD matrices it is suggested to consider a geodesic metric exploiting the properties of the Riemannian space on which these matrices lie. This study also reveals that a specific encoding of the FC matrices, describing them according to their distances from cluster centroids, allows good performance in distinguishing between HC and patients.

K-means algorithm is one of the traditional and widely being used algorithm to perform the clustering. Besides the good classification results, k-means clustering algorithm has some disadvantages e.g. it needs prior information about the number of clusters 'K' which make it rigid and computationally expensive when it comes to choosing multiple numbers of clusters. K-means clustering is sensitive to initialization (Initial data points to start clustering) and outliers present in the data. These reasons can make the k-means algorithm to give different shapes of clusters if run multiple times with different initialization as it tends to terminate at local optimum. So this approach with these disadvantages is not suitable to distinguish aberrant brain connectivity patterns in pathological subjects. Hence we need a better clustering algorithm which overcomes all these issues and performs the clustering in a more appropriate way.

Chapter 5

Encoding Brain Connectome for Classification and Neuromarkers Identification

5.1 Introduction

Multiple sclerosis (MS) is a chronic, inflammatory and neurodegenerative disease of the central nervous system (CNS) which mainly affects young adults and is characterized by non-uniform clinical manifestations and a variable progression [136]. In most patients, reversible episodes of neurological deficits characterize the first phase of the disease, called relapsing-remitting (RRMS). Over time, the development of permanent neurological deficits and the progression of clinical disability become prominent; this phase is known as secondary progressive (SPMS). A minority of patients experience a progressive disease course from the clinical onset: this phenotype is referred to as primary progressive (PPMS). Recent studies suggested that PPMS and SPMS share similar characteristics in terms of disease activity and disease progression [137]; as such, they can be grouped together when investigating functional brain reorganization. Studies applying connectomic techniques highlighted that functional reorganization occurs in MS, and varies according to disease phenotype [138].

The application of functional MRI (fMRI) techniques to the study of patients with MS holds great promise to provide an accurate characterization of brain injury across different disease stages and to improve the understanding of brain response to the progressive accumulation of disease-related damage. In the last two decades, the advent of resting-state (RS) fMRI allowed to avoid the behavioural confounds of task-based studies, and to study functional connectivity (FC) abnormalities also in severely disabled patients. RS fMRI studies of MS patients often showed trajectories of FC changes mirroring those detected by task-based studies. Among the more advanced approaches introduced to the analysis of neuroimaging data, network-based analysis has recently received great attention. In the field of Connectomics main focus is the exploration of brain FC [139], in order to characterize the alterations in the functional organization of the brain. A common approach adopted in these studies is to discriminate healthy subjects from those suffering from a neurological condition e.g. MS. To achieve the aim of discriminating between two groups of subjects, for instance, healthy controls (HC) and MS patients, different methods have been proposed re-

cently which take into account the geometrical properties of SPD matrices and discussed in detail in section 2.3

In this work, as the first step towards classification, participating subjects were clustered into a limited number of groups with similar FC covariance matrices, in order to reduce the impact of individual variability on the characterization of MS phenotype. Dominant set clustering has been adopted as the technique of choice, in which intrinsic data properties are used to define the "natural" number of clusters and it does not require the definition of the number of clusters, which is unknown a priori in this case. All that is required for clustering is, instead, the similarity matrix between subjects, easily defined once the geodesic distance (Log-Euclidean, section 2.4) on a Riemannian manifold is computed between subjects. For comparison, we also used the geodesic version of k-means (chapter 4), which requires a fixed number of clusters decided a priori and is also sensitive to the initialization.

Once clusters have been assigned, the centroids of each cluster are subsequently used as a dictionary (reference point) to build a low-dimensional representation: the FC of each subject, at this stage, is described by the distance from each centroid (typically between 5 and 7 in number), as opposed to the covariance matrix itself, which is defined by several thousand independent values. The resulting low-dimensional representation is then fed to a linear SVM which provides the actual classification. This sequence of operations is carried out both on the entire covariance matrix at once and on specific sub-networks (defined by neuroscientists/collaborators), in order to evaluate the impact of specific brain-regions on the classification accuracy (and hence their relevance in MS progression). Finally, for comparison of classification performance, we repeated the same procedure using euclidean distance in both clustering algorithms. Similarly, the weights of the trained SVMs have been analyzed in order to extract the connections most likely to be affected and provide the definition of possible neuro-markers to monitor in order to evaluate MS evolution.

5.1.1 Dataset Acquisition and Pre-Processing

Subjects used for the current analysis are part of a prospective cohort followed at the Neuroimaging Research Unit (Hospital San Raffaele, Milan, Italy). To be included, subjects had to satisfy the following criteria: 1) right-handedness; 2) have no other major systemic, psychiatric or neurological disorders; 3) no history of drug/alcohol abuse; 4) for patients, to be relapse- and steroid-free for at least 3 months before MRI acquisition and have a stable disease-modifying treatment during the past 6 months. Within 48 hours from MRI acquisition, MS patients underwent a complete neurological evaluation, with a rating of clinical disability using the Expanded Disability Status Scale (EDSS) score [140].

MRI scans used in this analysis were collected from all study subjects using a 3.0 T Philips

Intera scanner (Philips Medical Systems, Eindhoven, The Netherlands) and included the following MRI sequences: a) T2*-weighted single-shot echo-planar imaging (EPI) sequence for RS fMRI (repetition time [TR]=3000 ms/echo time [TE]=35 ms, field-of-view=240 mm2, matrix=128x128, slice thickness=4 mm, 200 sets of 40 contiguous axial slices); and b) 3D T1-weighted turbo field echo (TR/TE=7/3.2 ms, inversion time [TI]=900 ms, field-of-view=256x240 mm2, matrix=256x240, slice thickness=1 mm, 192 sagittal slices).

For RS fMRI, standard preprocessing, including motion correction and registration to MNI space, was performed using fMRIPrep [141] and selecting the non-aggressive ICA-based Automatic Removal Of Motion Artifacts (AROMA) denoising output [142]. 3D T1-weighted scan was processed using FSL FAST [143] for tissue segmentation and grey matter (GM) tissue mask was matched with a subset of regions derived from the AAL atlas ($n = 86$ regions considering cerebrum only, and excluding bilateral putamen and insula for misregistration issues) and applied to the processed fMRI scan of each subject to extract the mean time-series signal of all ROIs. The 86×86 RS FC matrices Σ_i were computed for each subject using the covariance between time-series, describing in this way the brain connectivity in term of both signal co-activation between ROIs and its amplitude.

5.2 Methods

5.2.1 The Riemannian Manifold of SPD Matrices

Let $\mathcal{S} = \{\Sigma_1, \ldots, \Sigma_K\}$ be the set of covariance matrices describing the brain connectivity of all subjects. As described in detail in section 2.3 of chapter 2, covariance matrices enjoy the property of being symmetric positive semi-definite, however, in a real scenario, they are almost always symmetric positive definite (SPD). To fully exploit the advantage of the manifold structure, it is essential to consider the notion of geodesic distance section 2.4, we opted to use the Log-Euclidean (Log-E) distance Eq. 2.8 which is simple and fast to compute and the related mean is defined in a closed-form Eq. 2.9

5.2.2 Encoding and Classification

Recent studies (discussed in section 2.3) proved that there is some alteration in FC of the brain due to the pathological condition, which in principle, can be useful for classification between HC and MS patients. These alterations are concealed behind the intrinsic high variability between subjects which makes them problematic to recognize. So, the underlying hypothesis of this work is to reduce this high variability by collecting all subjects into homogeneous groups and compute a unique representative connectome for each group. Hence, the problem can be considered as a clustering task on the FC matrices and the resulting

cluster centroids of each cluster become reference networks, representing groups of people with similar conditions. The ensemble of these reference networks represents, therefore, a dictionary that can be used to compress the high dimensional connectivity patterns into a lower-dimension vectorial descriptor retaining the difference between groups while filtering the intrinsic variability of subjects in the same group. More specifically, the FC matrix of each subject can be represented by the set of distances from all cluster centroids (As shown in previous chapters).

With such representation, any classifier can be employed for classification. In particular, in all our experiments, we have used a linear SVM and logistic regression with LASSO regularization for comparison and analysis of features. This approach is similar to [39], where PCA was adopted to determine the representative brain networks. This work is based on the speculation that being a complex and geometrical nature of data structure, FC matrices cannot be fully characterized using Euclidean metric and can be better expressed by a geodesic distance on the manifold of FC matrices. Moreover, we have deployed the DS clustering algorithm, which has two main advantages: it is easily adaptable to any metric, and it automatically determines the number of clusters. Another interesting property of DS clustering is that it does not only consider the points in isolation with respect to the centroids (as in k-means), but it also exploits the relations with all other nodes in the graph. This property generates more robust-to-noise clusters which are completely explicit to the initialization (contrary to k-Means). Furthermore, in the past, DS clustering algorithm has been successfully applied in other partially related contexts [144, 5, 145].

5.2.3 Geodesic Dominant Set Clustering

The Dominant Set (DS) clustering [146] is a graph-based method that generalizes the problem of finding maximal cliques in edge-weighted graphs. For this reason, the DS can be used for partitioning (clustering) a graph into coherent, well-separated and compact disjoint sets. To perform data clustering [147][1], the dataset at hand is modeled as an undirected edge-weighted graph $G = (V, E, w)$ with no self loops, in which the nodes V are the items of the dataset (SPD matrices in this case), the edges $E \subseteq V \times V$ are the pairwise relations between nodes and the weight function $\omega : E \rightarrow \mathbb{R}\geq 0$ calculates pairwise similarities:

$$\begin{aligned} \omega(i,j) &= \begin{cases} 1 - \frac{d_L(\Sigma_i,\Sigma_j)}{\gamma} & \text{if } (i,j) \in E \\ 0 & \text{otherwise.} \end{cases} \\ \gamma &= max(d_L(\Sigma_i,\Sigma_j)) \quad \forall i,j \in \{1,\dots n\} \end{aligned} \tag{5.1}$$

where d_L is defined in Eq. 2.8 and γ is the maximum pairwise geodesic distance in the entire

[1] we used the implementation available here: https://github.com/xwasco/DominantSetLibrary

graph. The $n \times n$ adjacency matrix $A = (\omega_{ij})$ summarizes G.

To obtain a well separated and compact cluster, the DS method optimizes the so-called *cohesiveness*, a quadratic function considering the similarity matrix of the graph and the cluster assignment of all nodes. Such measure quantifies the goodness of a cluster based on the elements within it and all the other items in the graph. Optimizing the cohesiveness means that, once an optimal solution (local/global) is found, no other nodes can enter the cluster without lowering such measure. This explicitly satisfies two assumptions for good clusters: have a high *intra-cluster* homogeneity while having a low *inter-cluster homogeneity*. This is one of the main motivation for choosing this method over other alternatives.

The optimization problem is formalized as:

$$\begin{aligned} \text{maximize} \quad & \mathbf{x}^T A \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \in \triangle^n \end{aligned} \tag{5.2}$$

where $\mathbf{x}^T A \mathbf{x}$ is the cohesiveness, A is the similarity matrix of the graph and $\mathbf{x}$ is the so-called *characteristic vector*. The vector $\mathbf{x}$ lies in the n-dimensional simplex $\triangle^n$, defined as $\sum_i \mathbf{x}_i = 1, \forall i \; x_i \geq 0$. The entries of the vector $\mathbf{x}$ represent the likelihood of each node of being assigned to the cluster.

The optimization of Eq. 5.2 is performed through a dynamical system which mimics a Darwinian selection process. Such dynamical system is an important result from evolutionary game-theory [148] and is known as *replicator dynamic* (RD) expressed by Eq. 5.3).

$$x_i(t+1) = x_i(t) \frac{(A\mathbf{x}(t))_i}{\mathbf{x}(t)^T A \mathbf{x}(t)} \tag{5.3}$$

The RD operates a selection process over the components of the vector $\mathbf{x}$. At convergence of Eq. 5.3 ($|| \, (\mathbf{x}(t) - \mathbf{x}(t+1)) \, ||_2 \leq \epsilon$), certain components will emerge ($x_i > 0$) while others get extinct ($x_i = 0$). The support $\sigma(\mathbf{x}) = \{i \in V | x_i > 0\}$ is a dominant set. The dynamical system starts at the barycenter of the simplex, giving all the nodes the same probability of being part of a cluster, thus letting groups emerge spontaneously from the data. The components of the characteristic vector are iteratively updated using Eq. 5.3. At convergence, a dominant set is extracted and its subsets of nodes are removed from the graph (*peeling-off* strategy) making the algorithm faster at each extraction. The process starts again on the remaining nodes extracting the next cluster. This procedure is performed until all nodes are clustered, hence the number of clusters is automatically found. DS algorithm extracts the clusters in a sequential manner and determines the number of clusters automatically.

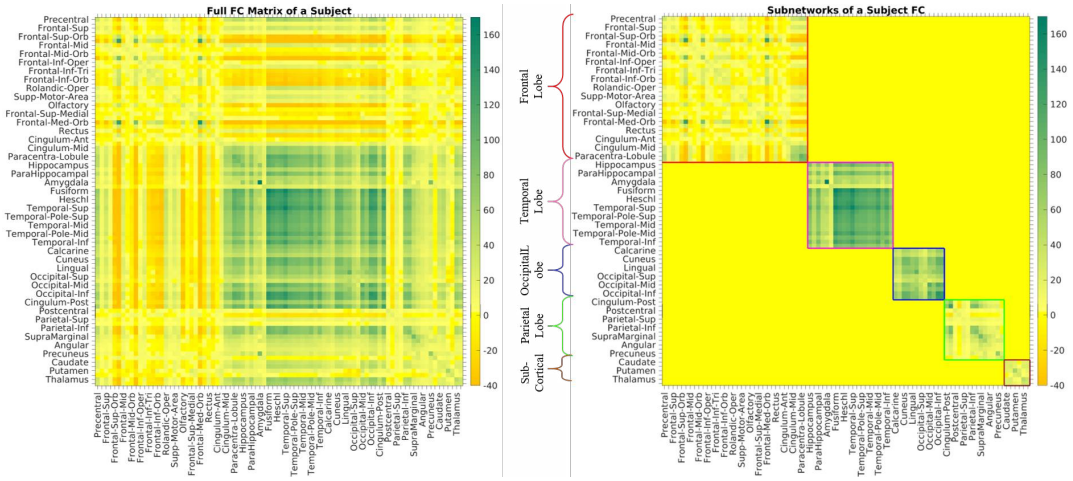

Figure 5.1: Illustration of full FC matrix (86 x 86) of a subject (left) and division of regions into different subnetwokrs (right)

Since its ability to preserve the internal coherency, DS grouped data into a cluster with high pairwise similarity which prevents the outliers to be grouped into any clusters.

To compare the proposed method using geodesic DS clustering algorithm, we also tested geodesic k-means clustering, which is already discussed in detail in chapter 4. In this approach we used Log-Euclidean distance Eq. 2.8 as distance metric and Log-E based mean formula Eq. 2.9 to compute the group centroid. We tested the k-means clustering with a number of clusters ranging from 2 to 10 whereas with DS clustering number of clusters were automatically determined by the algorithm.

5.3 Experiments

The working hypothesis is that the FC matrices can be clustered into groups while preserving the alterations in brain connectivity which characterize these groups. So, in the proposed pipeline, clustering of FC matrices is the first step and we used DS clustering with geodesic distance because the number of clusters is not known a priori. In our experiments we used the FC data in two ways: 1) using the full FC matrix for whole-brain connectivity analysis (Figure 5.1, left) and 2) using 5 sub-networks for the analysis of within sub-networks connectivity (Figure 5.1, right). Moreover, we addressed the task in two steps with the classification task followed by the identification of discriminative neuro-markers.

5.3.1 Classification Experiments

For classification purpose, we have deployed a 5-fold cross-validation (CV) setup with the constraint of preserving the proportion between classes. In each iteration of the CV setup, we

used DS clustering to cluster the training data and computing the centroid of each cluster. Then we computed the geodesic distance of each training and testing sample from these centroids and represented these matrices as a set of distance vectors (features). We used these distance vectors to train and test classifiers. For classification purpose we opted to use linear-SVM [99] which classify the data by generating a separation model [123], which has relative weights associated to each input feature that represent their importance for distinguishing between groups, thus contributing to the characterization of the data [127]. This way, instead of analyzing predefined regions, there is an objective evaluation of the whole brain, which may highlight relevant features (reference network/cluster) which can leads toward the brain regions that might not have been otherwise considered. Selection of linear-SVM is for the same motivations discussed in subsection 4.3.1

The number of features depends on the number of clusters extracted automatically by the DS algorithm (usually in the range of 5-7). We repeated this task only once for full FC matrices analysis and one by one for each sub-network and concatenating the features obtained from each sub-network together in the end. In these experiments all distances were computed using Log-Euclidean distance Eq. 2.8 and corresponding geodesic mean Eq. 2.9. In order to evaluate the impact of the fold selection procedure, this process was repeated 100 times and results reported are the averages over these iterations. To check the significance level of classification results we also performed a permutation test on labels: a null distribution was generated by randomly changing the labels 1000 times (generated 1000 random labelling sequences) and for each distribution, we performed the same linear SVM classification using 5-folds CV approach.

5.3.2 Neuro-marker Identification Experiments

A neuro-marker can be interpreted as a region of the brain whose connections are especially relevant for a classification task. To this aim, as a first step, we considered the whole connectome to compute the group representatives (cluster centroids) using DS clustering. Distances of all subjects from these representative networks were then used in a 5-fold CV setup to train linear SVM and LASSO regularized logistic regression (LR). Finally, feature weights were observed in each repetition. Indeed, for statistical reasons, CV was repeated 100 times. We repeated this task only once for full FC matrices analysis and one by one for each sub-network and concatenating the features together in the end. Two types of the classifier were used so that a better understanding and comparison about feature weights can be developed.

To select the set of relevant features, we applied a permutation test on feature values (by randomly shuffling features value for each subject) and selecting only those features having

weight value higher than 95 percentile of permutation test value Selected features actually describe the network which has most discriminant connections, i.e., contributing most in classification For each selected feature we determined the threshold best separating the two group of the subjects. We then used the subjects correctly separated by this univariate criterion to compute for each group the mean FC matrix by Eq. 2.9, which actually represent the reference connectomes for the two groups corresponding to the selected feature. By computing the difference of these reference connectome we can actually identify the neuro-markers which shows the difference of FC between two groups that contributed to the classification.

5.4 Results

We applied our proposed algorithm to classify between HC and MS patients with the different disease phenotype, i.e., HC versus RRMS, HC versus progressive MS (PMS) and RRMS vs PMS. DS clustering resulted in a different number of clusters in considered experiments ranging between 5-7 for both, full FC matrices and sub-network matrices.

5.4.1 Classification Results

Figure 5.2 shows the boxplot of classification accuracy over 100 iteration of CV. Blue and orange boxplot represent the results for geodesic and euclidean k-means respectively. Green and brown boxplot shows the result for geodesic and euclidean DS clustering. Moreover table Table 5.1, 5.2 and 5.3 show the average confusion matrix, precision, recall and F1 score for HC vs RRMS, HC vs PMS and RRMS vs PMS respectively. In Figure 5.2.A it can be seen that, for HC vs RRMS achieved accuracy with geodesic DS is **72.51%** and precision, recall and F1 score of **72.85%**, **66.50%** and **69.40%** respectively (Table 5.1). With geodesic k-mean, Euclidean k-mean & Euclidean DS achieved accuracy is less. For HC vs PMS average accuracy achieved with geodesic DS is **85.19%** with precision, recall and F1 score of **87.11%**, **88.38%** and **87.21%** respectively (Table 5.2), whereas with geodesic k-means and Euclidean k-means and DS achieved accuracy is far lower. For RRMS vs PMS achieved accuracy with geodesic DS is **76.04%** and precision, recall and F1 score of **77.85%**, **77.83%** and **74.65%** respectively (Table 5.3), with geodesic k-means and Euclidean k-means and DS it is, again, lower.

Figure 5.3 shows the comparison of mean accuracy achieved with two types of classifiers, SVM (yellow) and logistic regression (green), and for two types of analysis A) full FC matrices analysis and B) analysis with sub-network connectivity only. It can be seen that the mean accuracy achieved with SVM is slightly higher as compared to logistic regression, as the latter is a probabilistic method, which produces highly interpretable statistical models, but also make it vulnerable to overfitting. On the other hand, SVM is based on geometrical

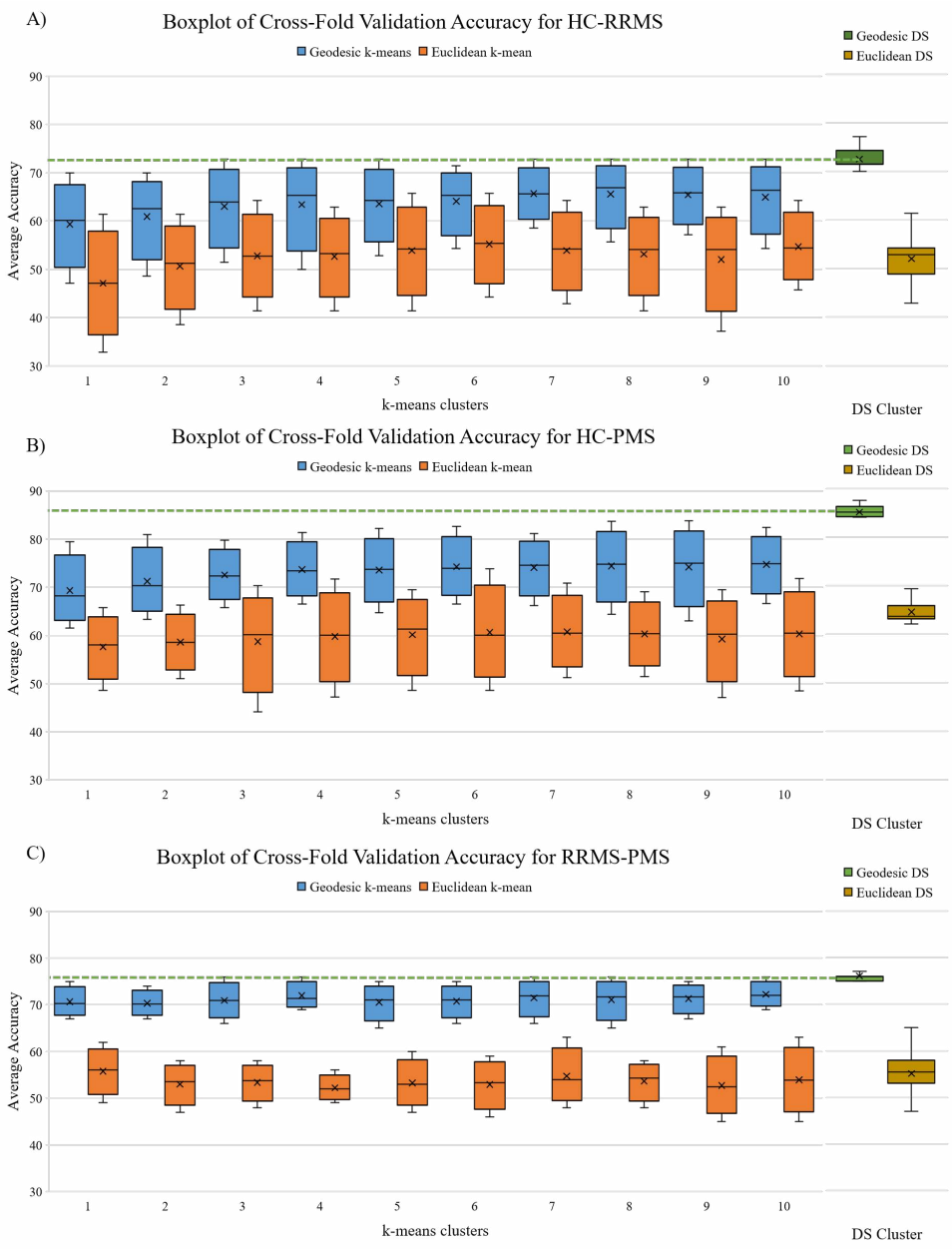

Figure 5.2: Boxplots represent the average cross-fold validation accuracy (with SVM, full FC matrices) with geodesic DS (green), Euclidean DS (brown), geodesic $k-means$ (blue) and Euclidean $k-means$ (orange) with $K = 2:10$) for A) HC-RRMS, B) HC-PMS and C) RRMS-PMS. 'x' show the mean of accuracy

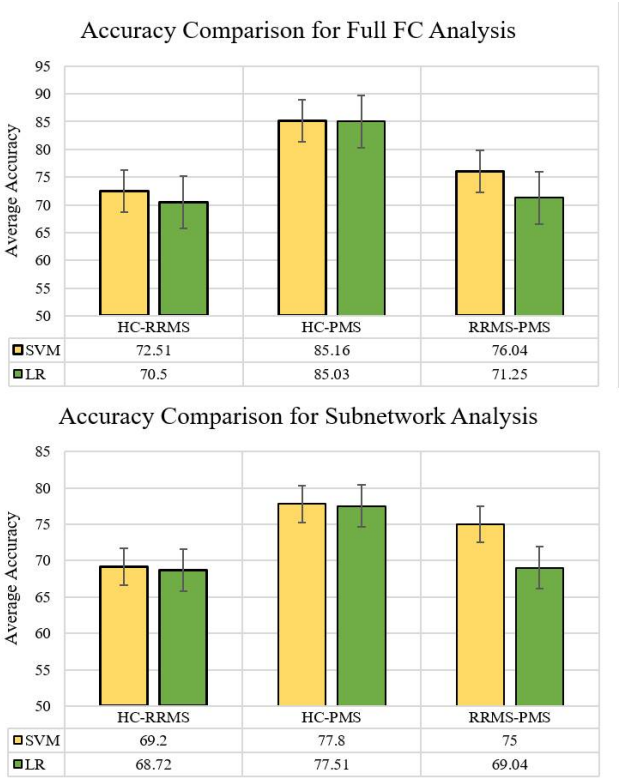

Figure 5.3: Mean accuracy achieved with SVM (yellow) and logistic regression(green) for full FC analysis (top) and subnetwork analysis (bottom).

properties of the data which makes it more suitable for the data acquired through clustering on manifold [125]. The difference in accuracy achieved by the two types of analysis (full FC matrices vs. sub-networks only) is due to the fact that in full FC matrices we are using values of all nodes as input (within sub-network and also between network connections values), hence providing more information to the classifier, whereas, in the sub-networks analysis, we are only providing the node values of within sub-networks connection as input and hence providing less information to the classifier.

To check the statistical significance of our results, we have performed a permutation test on labels. Figure 5.4 shows the distribution of performance of the proposed method (green) in comparison to the permutation test (red). Permutation results are also summarized in form of p_values, which show the ratio of a total number of accuracy values greater than the accuracy value with the true label divided by the total number of permutation (1000 in our case). For HC vs RRMS, HC vs PMS & RRMS vs PMS p_val is zero, because there is no accuracy value from permutation test that is higher than the mean accuracy with our experiment. These results strongly support the hypothesis that alteration in brain FC can be utilized to discriminate between HC and MS phenotype especially when using an appropriate

Table 5.1: Average Confusion matrix of classification results for HC vs. RRMS

Confusion matrix for HC-RRMS				
		Predicted Class		
		HC (-ve)	RRMS (+ve)	
Actual Class	HC (-ve)	25.68	11.32	**Recall 66.50**
	RRMS (+ve)	11.92	37.08	
		Precision 72.85		**F1 Score 69.40**

Table 5.2: Average Confusion matrix of classification results for HC vs. PMS

Confusion matrix for HC-PMS				
		Predicted Class		
		HC (-ve)	PMS (+ve)	
Actual Class	HC (-ve)	29.84	7.16	**Recall 84.18**
	PMS (+ve)	5.93	45.07	
		Precision 89.11		**F1 Score 86.71**

Table 5.3: Average Confusion matrix of classification results for RRMS vs. PMS

Confusion matrix for RRMS-PMS				
		Predicted Class		
		RRMS (-ve)	PMS (+ve)	
Actual Class	RRMS (-ve)	35.01	13.99	**Recall 77.83**
	PMS (+ve)	9.97	41.03	
		Precision 77.85		**F1 Score 74.65**

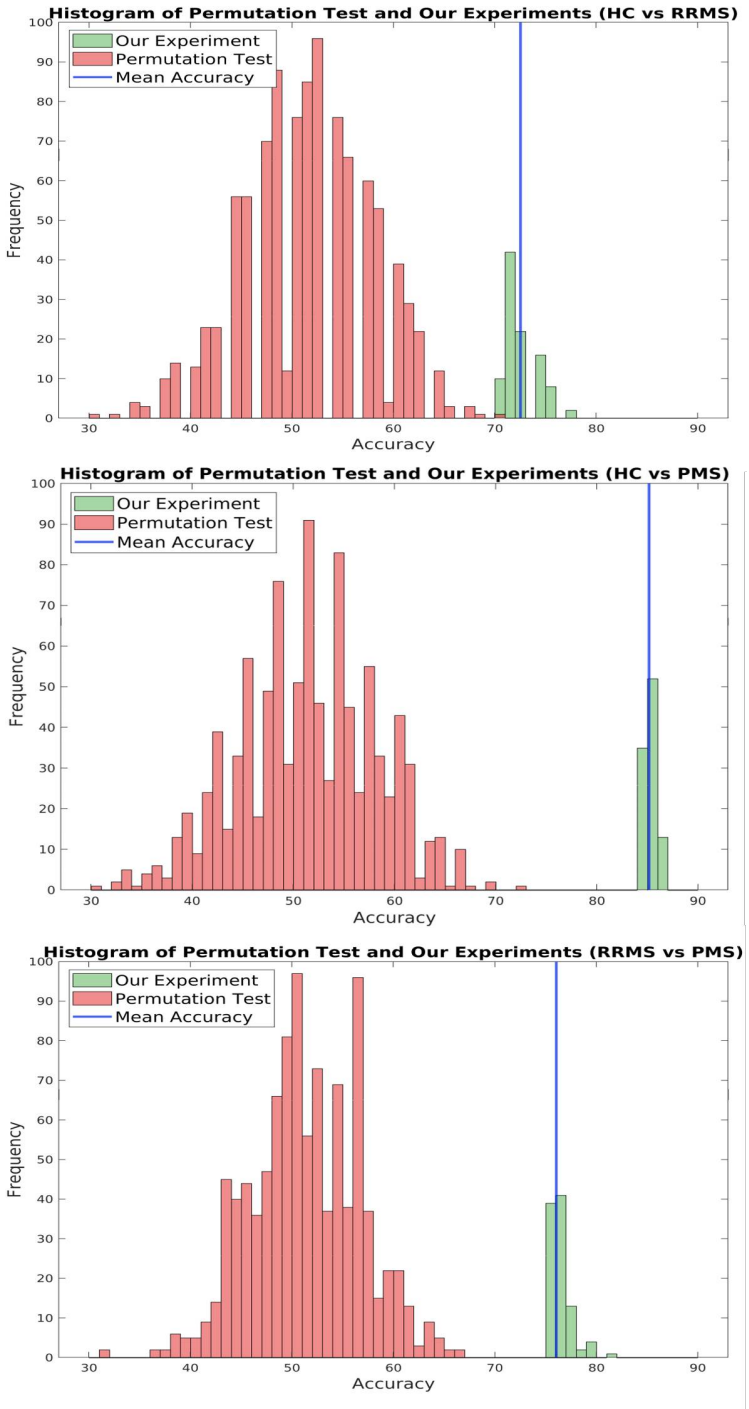

Figure 5.4: Histogram of accuracy (with SVM) distribution achieved with permutation test(red) obtained shuffling 1000 times the labels as compared to geodesic DS (green). Blue line represents the mean accuracy achieved with geodesic DS, which is always higher than the overall permutation distribution, demonstrating our algorithm accuracy was significantly higher than the accuracy obtained by chance only.

encoding technique which follows the manifold nature of SPD matrices

5.4.2 Neuro-marker Identification

Figure 5.5 shows the boxplot of classifier weights for 100 iterations for full FC analysis. Yellow bar represent the weights of the features for SVM classifier and green bar represent the weights of the features for LR classifiers for a combination of experiments A) HC-RRMS, B) HC-PMS and C) RRMS-PMS. Gray and orange lines show the threshold obtained with permutation test for SVM and LR classifier respectively. Figure 5.7, 5.9 and 5.11 show the classifier weights for subnetwork analysis of HC v RRMS, HC vs PMS and RRMS vs PMS experiments respectively. It can be seen that weights values obtained with SVM and LR look quite stable and identical. In LR we are using an L1 regularization which shrinks the weights of irrelevant features to zero and hence only provide high values to the weights which has significance importance in classification.

So for feature selection, we have considered the feature weights obtained with LR analysis and we have selected the features which have the maximum distance from the threshold. So in full FC matrices analysis, for HC vs RRMS, HC vs PMS and RRMS vs PMS we have selected the cluster number 1, 2 and 4 respectively. Figure 5.6 shows the selected features and the distribution of subjects in term of distance from centroid for each combination of phenotype classification. For HC vs RRMS and HC vs PMS, we are selecting only HC subjects below the threshold and RRMS & PMS subjects only above the threshold to form two separate groups. For RRMS vs PMS, we are selecting RRMS subjects only below threshold and PMS subjects only above the threshold for distinguishing groups formation.

For subnetwork matrices analysis we have selected feature number 5, 8, 19 and 32 for HC vs RRMS, feature number 1, 8, 22 and 29 for HC vs PMS and feature number 12, 16 and 22 for RRMS vs PMS. Red line over distribution shows the threshold to divide the subjects into two groups. Figure 5.8, 5.10 and 5.12 shows the distribution of subjects around the centroid of selected clusters for HC vs RRMS, HC vs PMS and RRMS vs PMS respectively. For HC vs RRMS (Figure 5.8), for cluster number 5,8 & 32 we are selecting HC subjects before threshold and RRMS subjects after the threshold. whereas in cluster 19 we are selecting RRMS before threshold and HC after the threshold. For HC vs PMS(Figure 5.10), for cluster 1 and 8 we are selecting HC subjects only before threshold and PMS subjects only after the threshold. whereas for cluster 22 & 29 we are selecting PMS subjects only before threshold and HC subjects after the threshold. For RRMS vs PMS (Figure 5.12), for cluster 12 & 22 we are selecting PMS subjects only before thresholding and RRMS subjects only after the threshold. For cluster 16 we are selecting RRMS subjects only before the threshold and PMS subjects after the threshold.

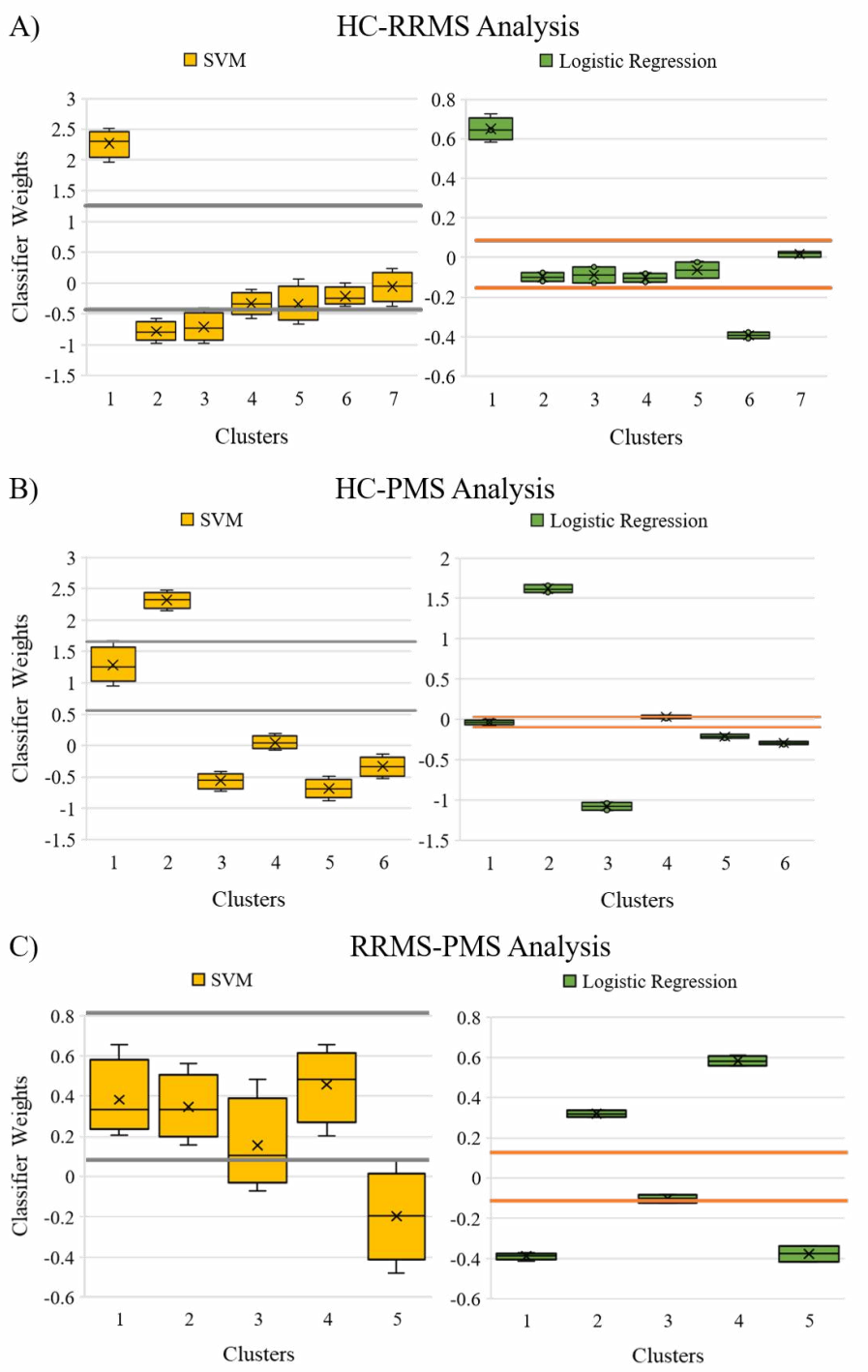

Figure 5.5: Results of sensitivity analysis (Full FC matrix) for feature weights of SVM classifier (in yellow, left side) and logistic regression classifier (in green, right side) for each combination of experiments A) HC vs RRMS, B) HC vs PMS and C) RRMS vs PMS. Gray lines represent the threshold obtained with permutation of features values for SVM classifier and orange line show the threshold for logistic regression classifier

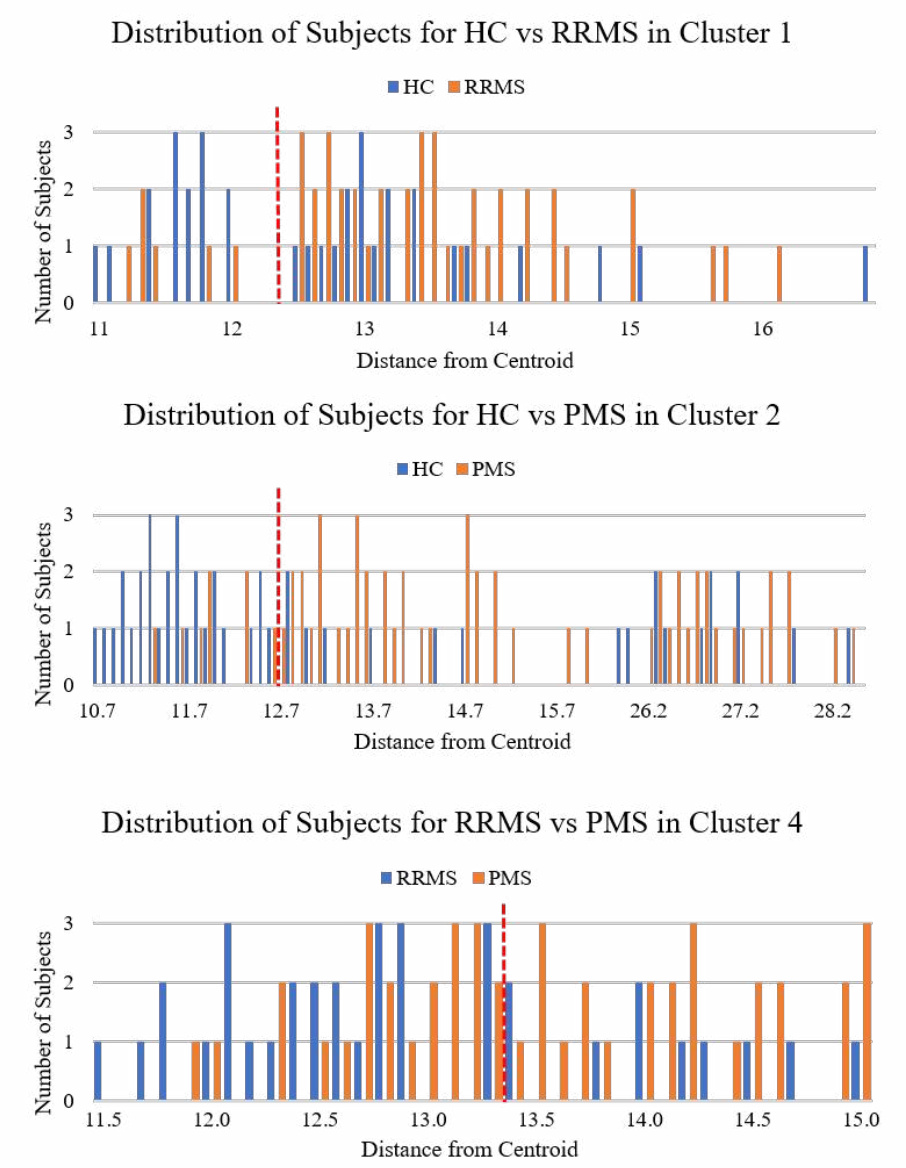

Figure 5.6: Distribution of all subjects around the selected centroid for selected cluster 1 of HC vs RRMS (top), cluster 2 of HC vs PMS (middle) and cluster 4 of RRMS vs PMS (bottom). Blue and orange color shows the subjects of two groups distributed around the centroid of selected cluster in term of their geodesic distance (log-E). Red doted line show the calculated threshold to separate two groups.

As described in subsection 5.3.2, for each cluster identified by sensitivity analysis, we are computing the difference between reference connectomes of two groups to highlight the neuro-markers mainly responsible for discrimination between two groups. Figure 5.13, 5.14 and 5.15 shows the results obtained for the analysis of HC vs RRMS, HC vs PMS and RRMS vs PMS respectively. Part A of each figure shows the reference connectome of each group along with the identification of subnetworks. Part B shows the significant connection changes obtained with the subtraction of two reference connectomes given in part A. Regions with blue colour represent a decrease of connectivity in the connectome of the second group whereas red colour represents an increase in the connectivity. Brain network images illustrate the mapping of connectivity changes within subnetworks (left) and between subnetworks (right). Part C shows the subnetworks with significant changes selected from sub-network analysis. Connecting arrows between part B and C show the resemblance of connectivity changes detected between full FC analysis and sub-networks analysis.

As shown in Figure 5.13 and 5.14, RRMS and PMS patients were mainly characterized by increased RS FC in the basal ganglia subnetwork (especially between the bilateral thalami) vs HC, as well as by decreased RS FC within the temporal, frontal and parietal subnetworks. In these networks, RS FC decrease mainly involved the fusiform gyrus for the temporal network and the medial frontal cortex for the frontal network, respectively. Conversely, an increased RS FC between the bilateral paracentral lobule and other regions of the frontal subnetwork was detected. With regards to the parietal subnetwork, decreased RS FC among the posterior cingulate cortex, angular gyrus and precuneus were detected. Finally, both RRMS and PMS patients showed decreased RS FC within the occipital subnetwork vs HC.

When looking at the comparison between PMS patients and HC, as highlighted in Figure 5.14, the increased RS FC in the basal ganglia subnetwork was more extensive than that in RRMS vs HC and involved also the bilateral caudate nuclei. Moreover, in the parietal subnetwork, evidence of increased RS FC between the superior parietal lobule and other parietal regions was detected.

The direct comparison of PMS vs RRMS patients showed, as highlighted in Figure 5.15, a stronger decrease of RS FC within occipital and temporal sub-networks in PMS vs PPMS patients, and a higher RS FC between the posterior cingulate cortex and precuneus. Interestingly, as shown in Figure 5.13, connectivity among sub-networks was markedly increased in RRMS patients vs HC, while between-network RS FC increase was not so evident in PMS patients vs HC (Figure 5.14). This was reflected by a decreased RS FC between parietal, occipital, temporal and frontal networks in PMS vs RRMS (Figure 5.15). The neuro-scientific results presented in this section is analyzed, verified and approved for presentation by our collaborators in Neurophysiology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy.

5.5 Discussion

In subsection 5.2.3 we have presented our contribution of a computational framework to distinguish between HC and MS subjects based on the FC matrices computed from rs-fMRI. In chapter 4 we presented a k-means clustering algorithm based framework, along with the limitations that make it a sub-optimal choice in this context. We then proposed a computational framework of graph clustering which deploys DS clustering algorithm to cluster FC graphs on Riemannian manifold and considering the properties of SPD matrices we used geodesic metric, which, in this context, proved to be superior to the Euclidean metric. We also used k-means for comparison purpose. Beside this clustering approach, a novel idea of data encoding was also suggested.

More in detail, classification features have been obtained as the geodesic distances between each subject's FC matrix and the centroids of the clusters defined either through k-means or dominant set clustering. In both cases, log-Euclidean distances on the Riemannian manifold space of SPD matrices of subjects were adopted. Both clustering and classifier training occurred on training folds, in order to prevent double-dipping. In testing, we computed the distances of each test sample to the centroids defined on the training set and then used these distance values to test the performance of the trained classifier. In order to minimize the impact of fold selection, the results presented in this work are relative to 100 repetitions of the training algorithm.

In order to evaluate our proposed algorithm, we performed the same comparison relying on Euclidean rather than geodesic distances for clustering. The result of this study supports the hypothesis that using Euclidean metric on manifold space is sub-optimal and generates lower system performance, whereas geodesic metric offers an optimal data representation allowing a better system performance. This study also proved the viability of the encoding scheme of FC matrices by using DS clustering and then defining a vector space representation according to their distance from cluster centroids. Results defined in Figure 5.2 also supports the fact that encoding of FC matrices using DS clustering algorithm always gives better results as compared to the k-means clustering algorithm.

From a clinical point of view, the results of this study showed that MS patients, considered as a whole, were well distinguished from HC in terms of FC. This reinforces the notion that FC reorganization does occur in MS, probably to counteract the accumulation of structural damage. When looking at disease phenotypes, PMS patients were better classified from HC than RRMS patients. This is also not surprising since PMS is characterized by longer disease duration and a worse clinical disability. Overall, classification of MS according to their connectivity profiles seems to be a rewarding strategy to characterize MS clinical heterogeneity.

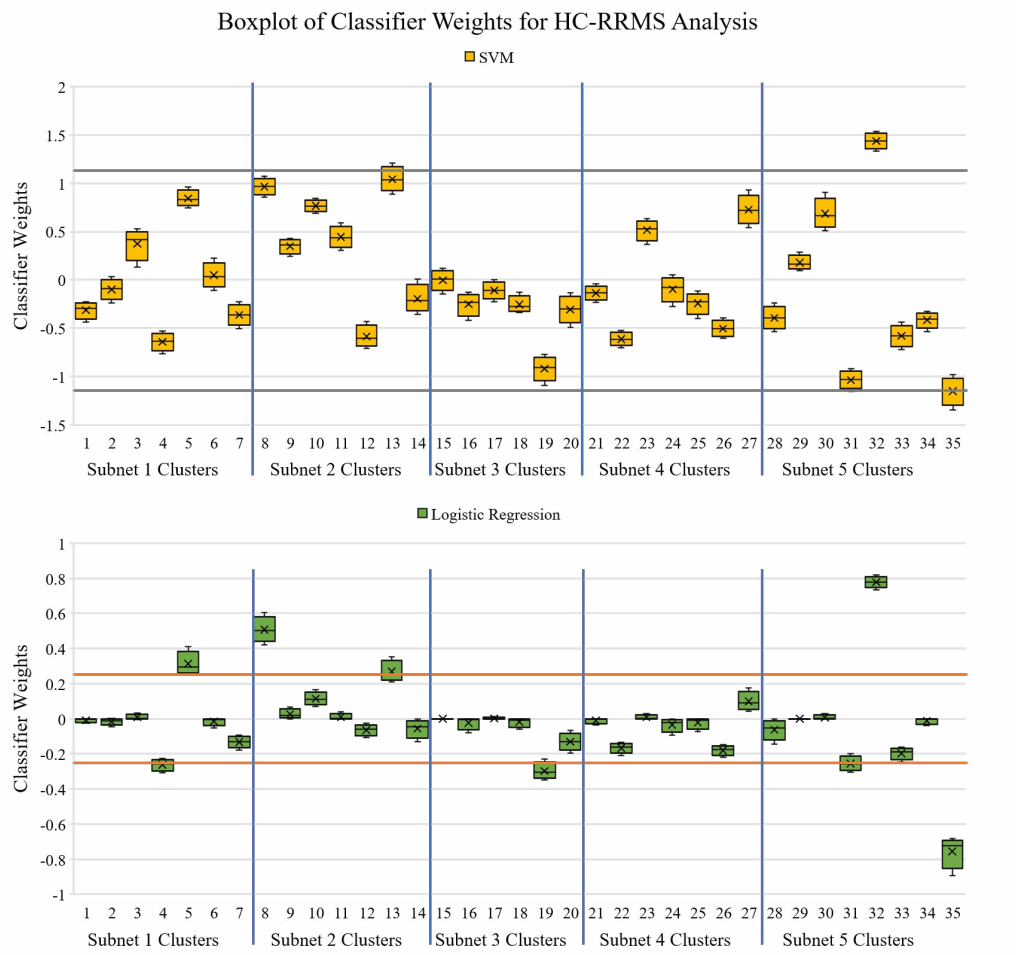

Figure 5.7: Results of sensitivity analysis (Subnetwork analysis) for feature weights of SVM classifier (in yellow, top) and logistic regression classifier (in green, bottom) for HC vs RRMS. Gray lines represent the threshold obtained with permutation of features values for SVM classifier and orange line show the threshold for logistic regression classifier

This approach might be rewarding in monitoring disease evolution and optimizing patients' management.

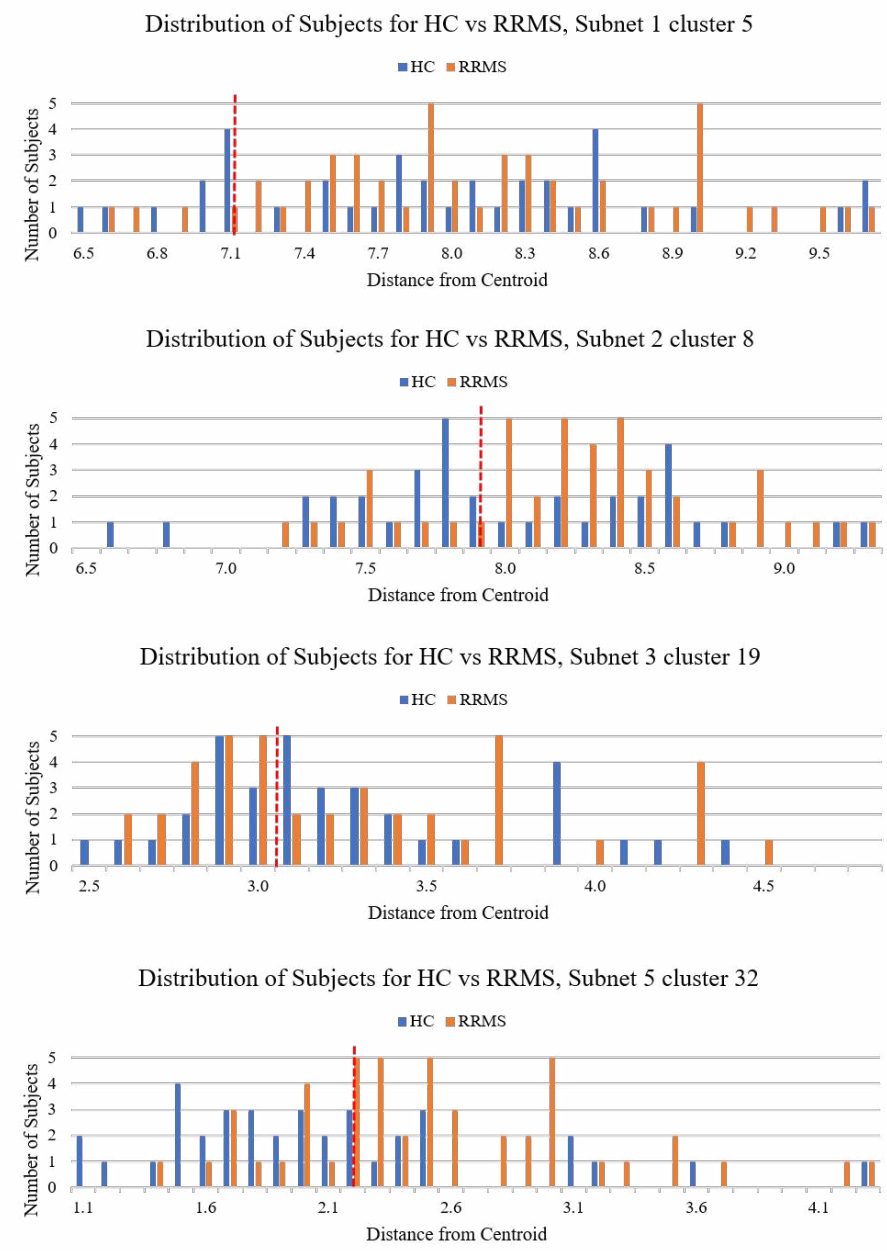

Figure 5.8: Distribution of all subjects (HC vs RRMS) around the selected centroid of subnetworks 1 cluster 5, subnetwork 2 cluster 8, subnetwork 3 cluster 19 and subnetwork 5 cluster 32. Blue and orange color shows the subjects of two groups (HC and RRMS) distributed around the centroid of selected cluster in term of their geodesic distance (log-E). Red doted line show the calculated threshold to separate two groups.

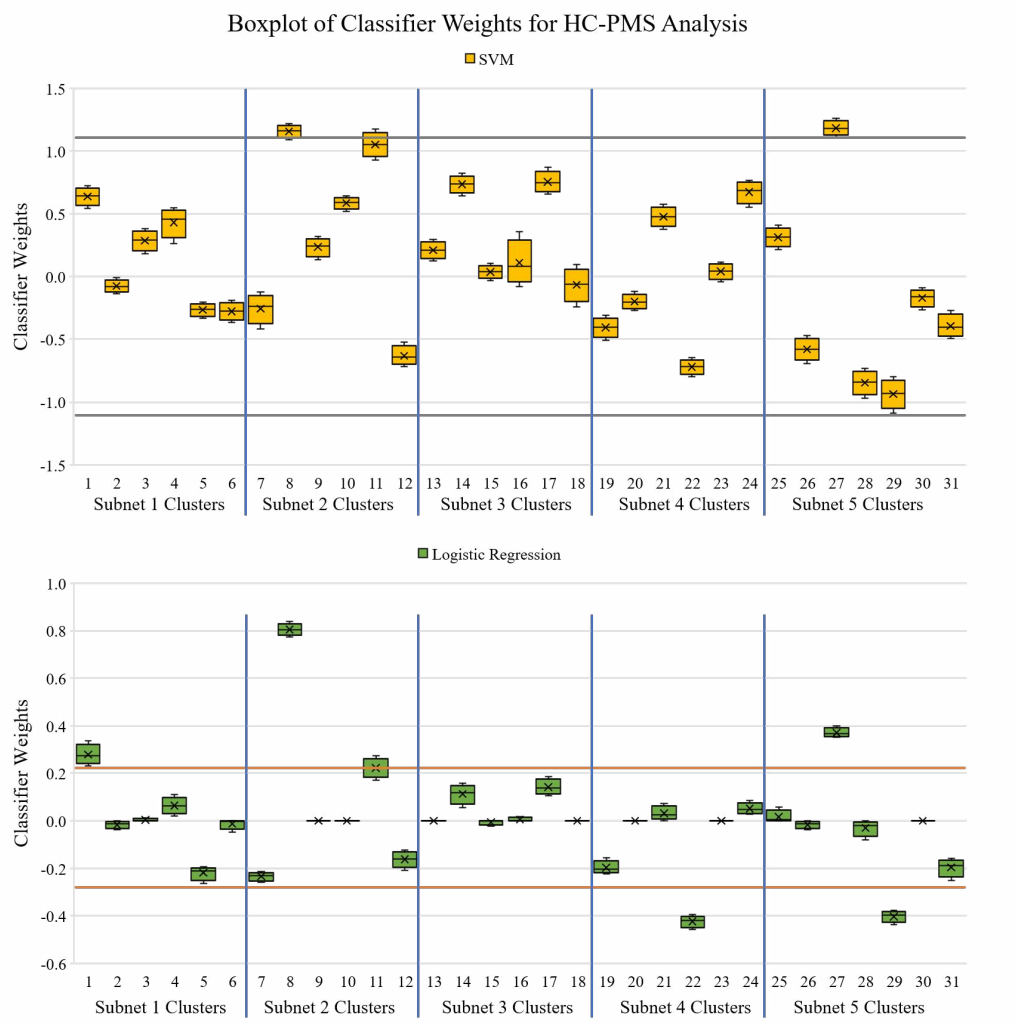

Figure 5.9: Results of sensitivity analysis (Subnetwork analysis) for feature weights of SVM classifier (in yellow, top) and logistic regression classifier (in green, bottom) for HC vs PMS. Gray lines represent the threshold obtained with permutation of features values for SVM classifier and orange line show the threshold for logistic regression classifier.

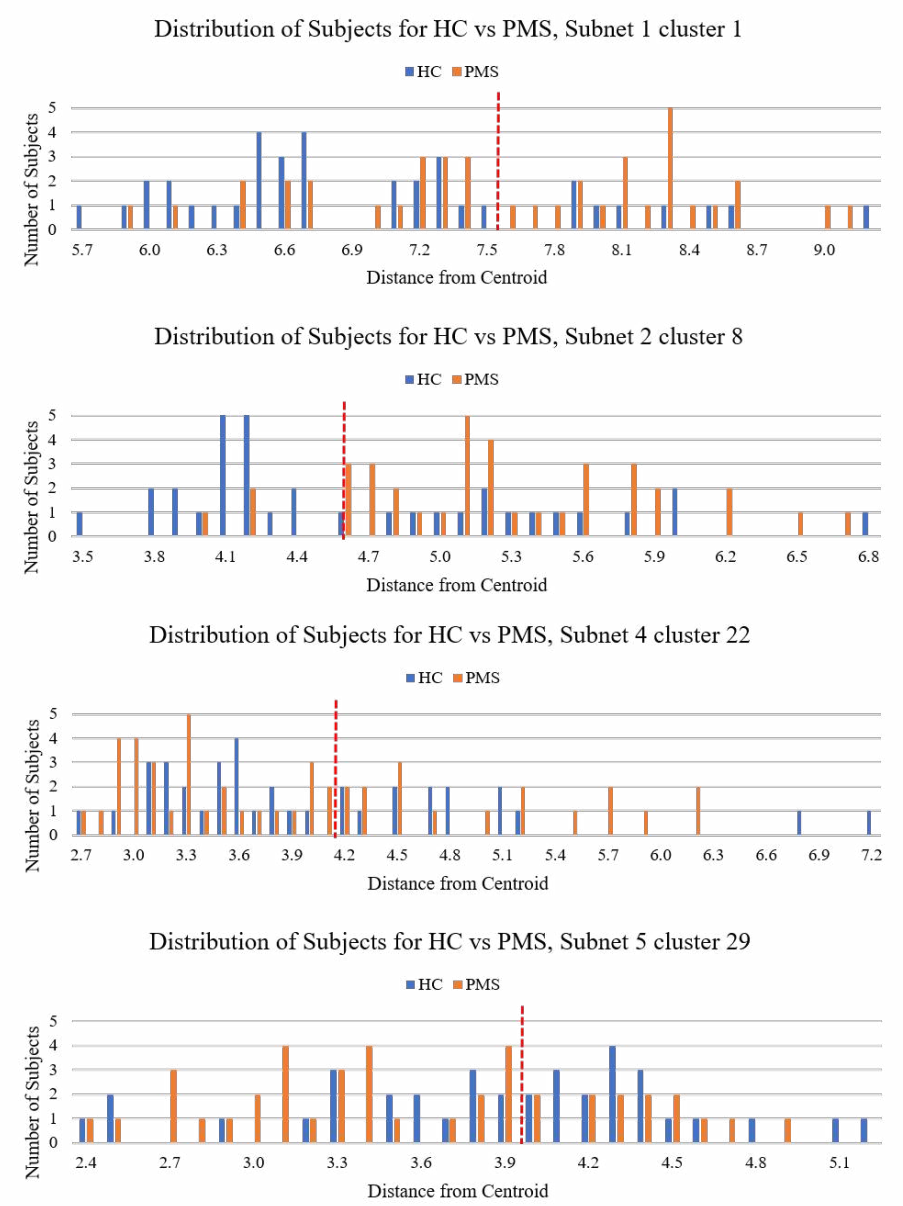

Figure 5.10: Distribution of all subjects (HC vs PMS) around the selected centroid of subnetworks 1 cluster 1, subnetwork 2 cluster 8, subnetwork 4 cluster 22 and subnetwork 5 cluster 29. Blue and orange color shows the subjects of two groups (HC and PMS) distributed around the centroid of selected cluster in term of their geodesic distance (log-E). Red doted line show the calculated threshold to separate two groups.

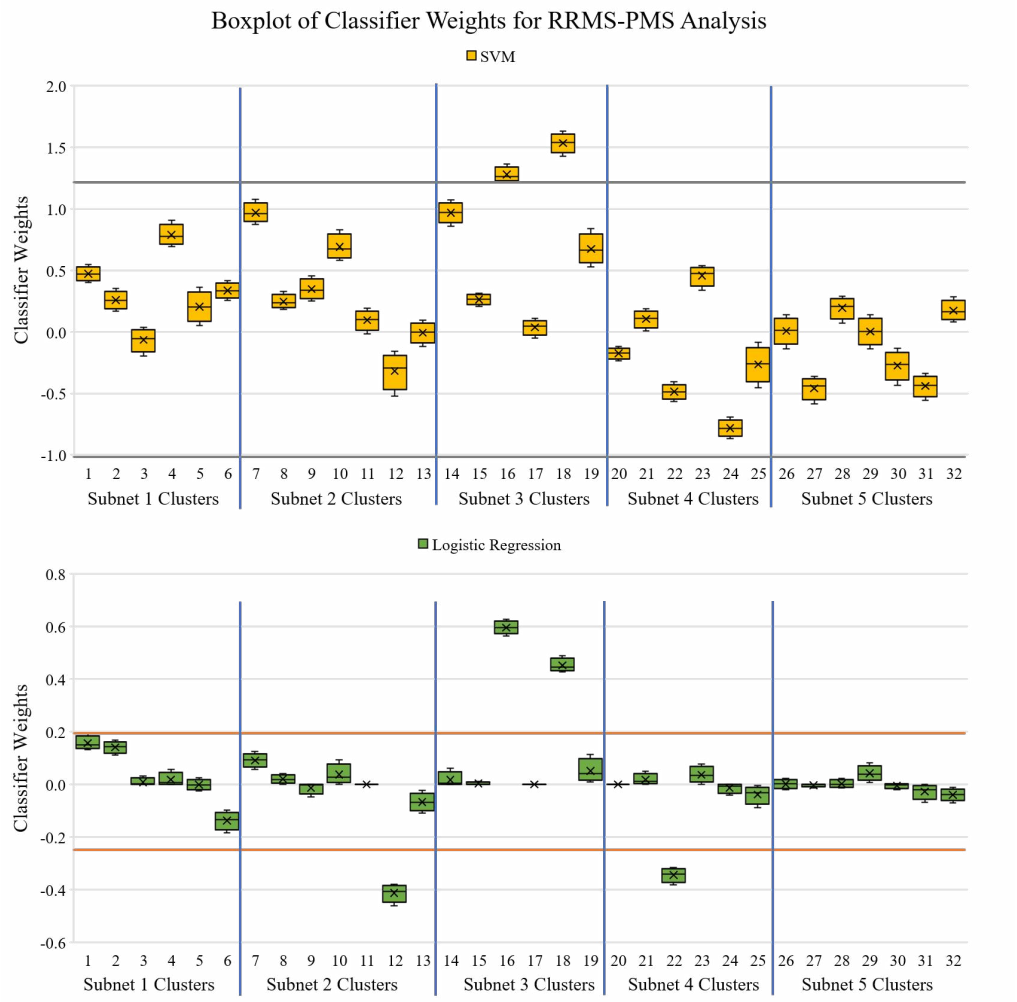

Figure 5.11: Results of sensitivity analysis (Subnetwork analysis) for feature weights of SVM classifier (in yellow, top) and logistic regression classifier (in green, bottom) for RRMS vs PMS. Gray lines represent the threshold obtained with permutation of features values for SVM classifier and orange line show the threshold for logistic regression classifier

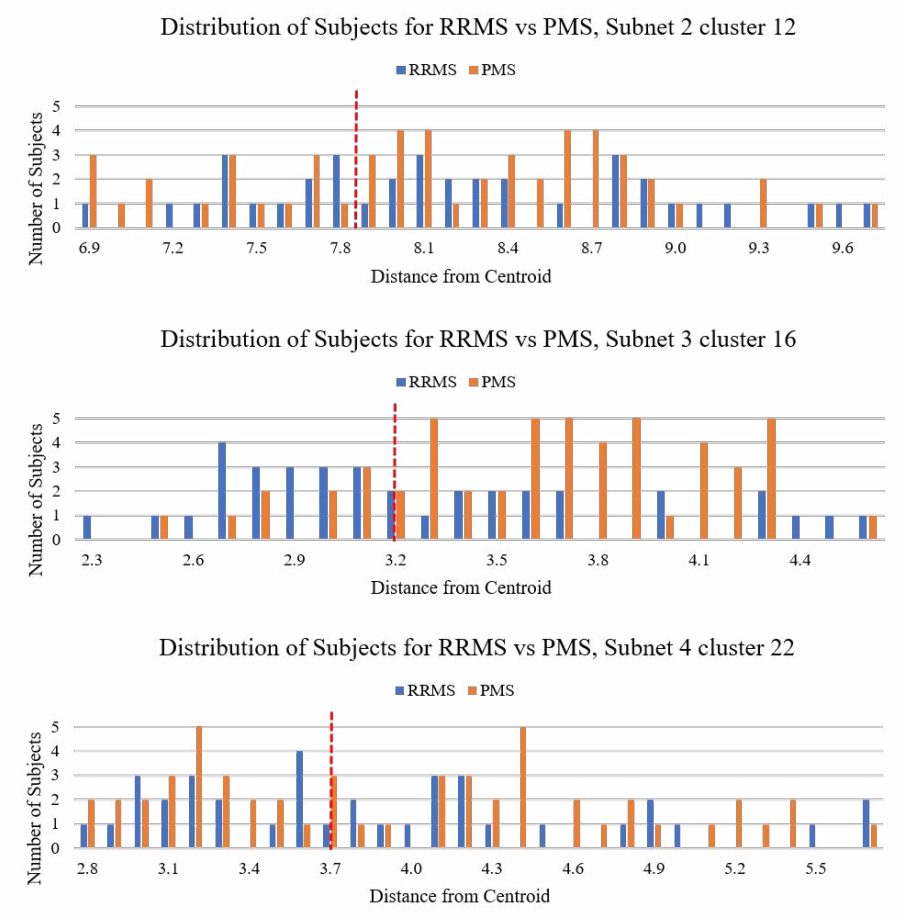

Figure 5.12: Distribution of all subjects (RRMS vs PMS) around the selected centroid of subnetworks 2 cluster 12, subnetwork 3 cluster 16, subnetwork 4 cluster 22 and subnetwork 5 cluster 29. Blue and orange color shows the subjects of two groups (RRMS and PMS) distributed around the centroid of selected cluster in term of their geodesic distance (log-E). Red doted line show the calculated threshold to separate two groups.

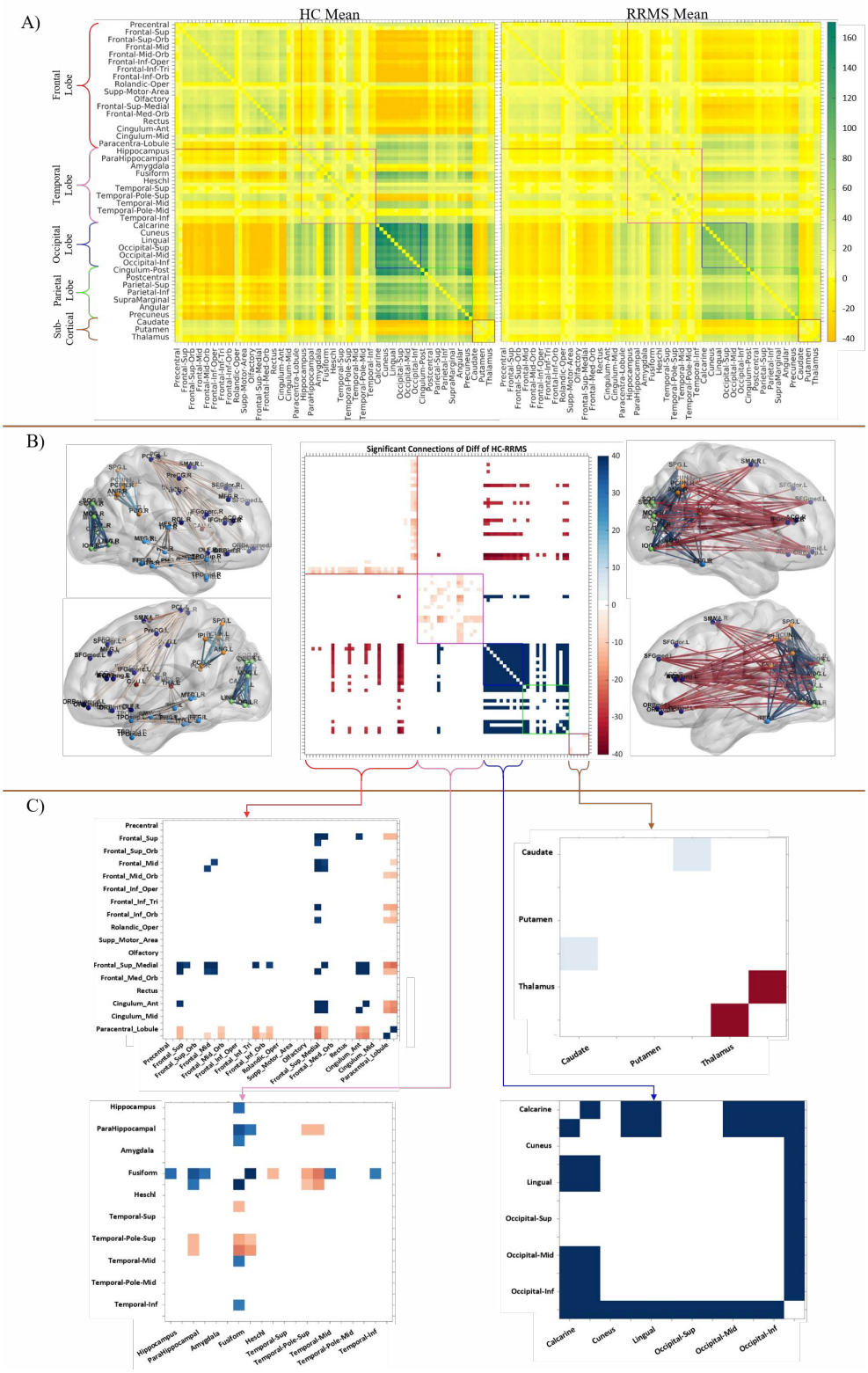

Figure 5.13: A) Reference connectome of HC and RRMS group along with the division of full FC matrix into subnetworks. B) Showing the difference between reference connectomes with only significantly changing connections, where blue colour represents the decrease in the connectivity of RRMS and red colour show the increase in connectivity of RRMS group. Mapping of significant connections on brain image is illustrated for within subnetwork connections (left side) and between subnetworks (right side). C) visualizing the significant connection changes for selected subnetworks and comparing the significant changes with full FC analysis

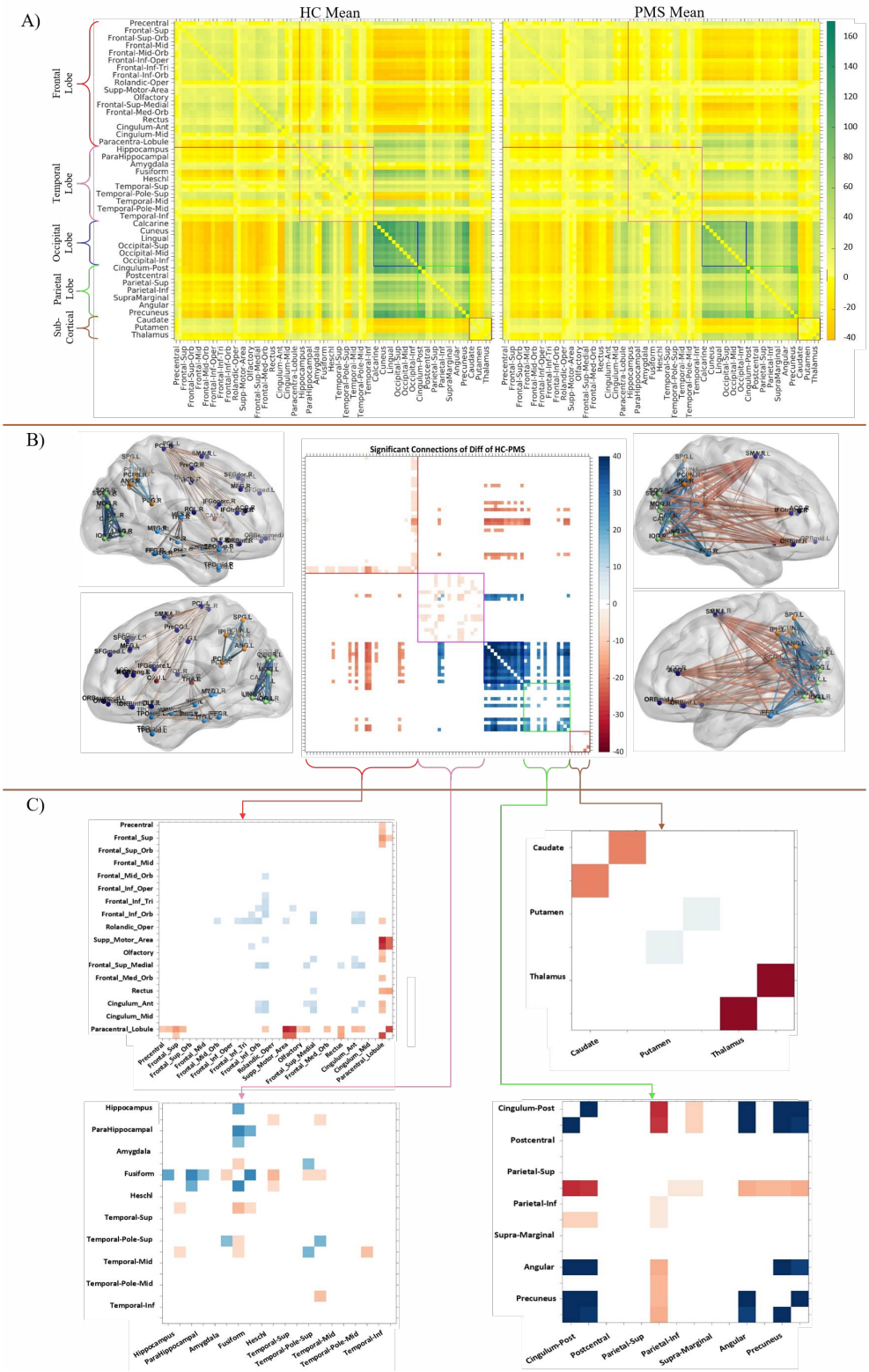

Figure 5.14: A) Reference connectome of HC and PMS group along with the division of full FC matrix into subnetworks. B) Showing the difference between reference connectomes with only significantly changing connections, where blue colour represents the decrease in the connectivity of PMS and red colour show the increase in connectivity of PMS group. Mapping of significant connections on brain image is illustrated for within subnetwork connections (left side) and between subnetworks (right side). C) visualizing the significant connection changes for selected subnetworks and comparing the significant changes with full FC analysis

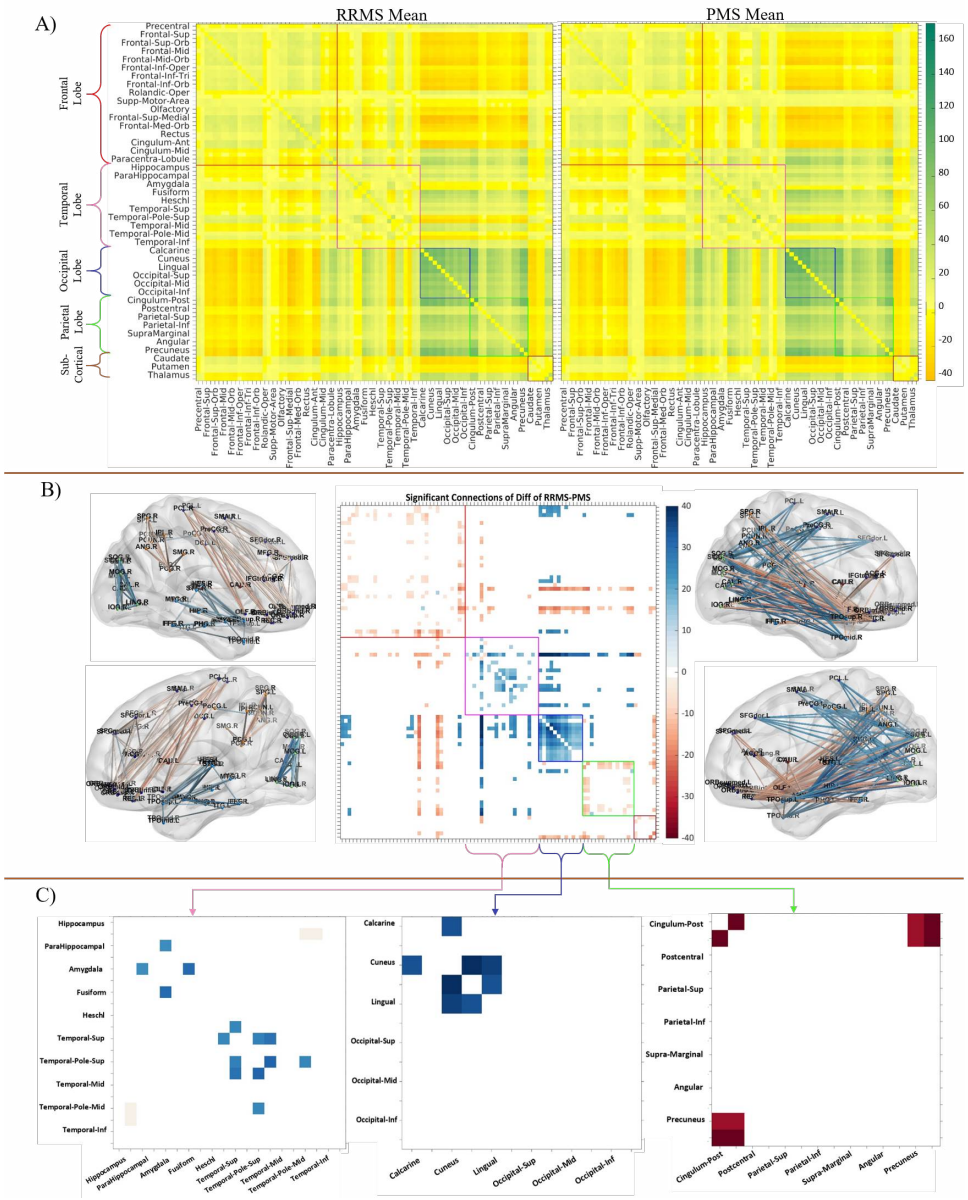

Figure 5.15: A) Reference connectome of RRMS and PMS group along with the division of full FC matrix into subnetworks. B) Showing the difference between reference connectomes with only significantly changing connections, where blue colour represents the decrease in the connectivity of RRMS and red colour show the increase in connectivity of RRMS group. Mapping of significant connections on brain image is illustrated for within subnetwork connections (left side) and between subnetworks (right side). C) visualizing the significant connection changes for selected subnetworks and comparing the significant changes with full FC analysis

Chapter 6

Exploring Dynamic Brain Connectivity using Riemannian Framework

fMRI functional connectivity (FC) is being widely used to investigate the functionality of the brain and there is emerging evidence of the dynamic nature of FC. Recent studies showed that the human brain is intrinsically organized into the large-scale functional network by measuring the Spontaneous spatiotemporal fluctuations in the brain connectivity [149]. Usually, fMRI data has been examined by using a mean of time-average statistics. Pair-wise Pearson correlation or covariance between regions of the brain is widely used to describe FC, which enable to represent the modular structure of the brain. In past studies, a common assumption was that the FC doesn't change its pattern over the data acquisition time period although brain roams between different states of brain connections, known as brain states and hence FC is non-stationary in nature. In several recent studies, growing importance has been given to explore the dynamical features of fMRI data to study non-stationary changes in connectivity and to discover the relevant organization of brain function during the fMRI data acquisition period which is typically from 5-15 minutes [150, 151, 152].

Initial studies demonstrated the reiteration of specific brain regions over time [38] that was not only changing during the task-related activity but also while resting. With the non-stationary assumption of FC, analyzing whole-brain connectivity becomes more complex and hence required a different framework as compared to those commonly used. The most widely used method to compute dynamic functional connectivity (DFC) consist of sliding a temporal window across time points across the fMRI sequence and compute a correlation matrix within each resultant window [151, 153]. This gives a three-dimensional stack of FC matrices, which then analyze by using several available methods to summarize statistics of brain DFC. In a state-based approach, DFC matrices are concatenated across the subjects and then being clustered by using some algorithm e.g. k-means to identify canonical centroids which represent different "brain-states"[114, 154]. By using the states as the fundamental unit, this approach allows analyzing higher-order summary statistics such as dwell time in computed states [155, 156] or computing meta-state vector which shows the contribution of various states to each DFC matrix[157].

In [39] proposed another sliding window-based data-driven approach to extract coherent

FC patterns by applying principal component analysis (PCA) on concatenated FC matrices over time and subjects. In [158] dictionary learning approach has been used to extract meaningful dynamic-FC (DFC) patterns in healthy subjects or in specific patients group. Moreover, [159] have proposed a total activation based method to extract innovation-driven co-activation (iCAPs) maps from resting-state fMRI which are both spatially and temporally overlapping. Cognition needs complex and dynamic interplay between distributed regions of the cortical and subcortical area because the brain at rest us usually defined in terms of significantly less number of networks compared to the number of functions it can perform [160, 114]. So to explore the whole brain FC in a more systematical way, in this study, on the contrary, we emphasize to analyze task-based DFC, which help to investigate the integration mechanism between task-related regions of the brain and is useful to exploit reorganization of whole-brain network[34, 108].

Specifically, in this chapter we have analyzed DFC on task-based fMRI dataset acquired on twins (same as in chapter 1), to investigate the relationship between genetic heritability and dynamic functional brain networks. Indeed, in chapter 1 we have shown that MZ twins pair are more similar in term of FC connectivity because MZ twins share the 100% of their genetics, as compared to the DZ twins which share 50% of genetics information. The main focus here is to evaluate whether the alterations of FC based on genetic heritability captured by dynamics can be meaningful to discriminate one group of twins from the other. For such reason, in next sections we will introduce our two contributions of design of frameworks 1) to classify between MZ and DZ twin pairs and 2) to classify a given pair as twin or unrelated pair. These frameworks are based on the utilization of Riemannian manifold properties of symmetric & positive definite (SPD) matrices to measure the similarity between DFC overtime at untangling patterns of FC.

6.1 Classification of Twins Pairs through Graph Laplacian on Manifold

Based on the results presented in chapter 1, it is possible to differentiate the overall functional connectivity of brain networks between monozygotic (MZ) and dizygotic (DZ) twins and then exploit this property to perform classification between twin pairs. In this section, we addressed the discrimination task of twin pairs with the help of machine learning algorithms applied to dynamic functional networks. Use of Machine learning in this task is not novel. For example, in [113] linear support vector machines (SVMs) have been used on DFC analysis based on k-means clustering, to discriminate traumatic brain injury. Also, in [161] enhanced FC variability has been used as a feature to classify autism spectrum disorder subjects from healthy controls using machine learning algorithms.

To this aim we used graph Laplacians representations of DFC matrices, to move the problem to the Riemannian space of semi-positive definite matrices (see section 2.2 and 2.3) This allowed us to compare the connectivity matrices using a geodesic distance. In our approach, we have used the dynamics of a component of the Frechet distance [62] known as Wasserstein distance. Each pair was represented as a sequence of differences (features) between the dynamic connectomes and linear SVMs have been used to discriminate the two groups. To get improved results and to reduce the dimensionality of our feature set we have also used the Laplacian Score to perform an unsupervised feature selection. Our result suggests that the use of geodesic distance on graph Laplacian based DFC matrices is a valuable method to evaluate the variations in the brain regions.

6.1.1 Dynamic Functional Connectivity Estimation

After applying the preprocessing steps including realignment, time slice correction, motion correction, normalization and smoothing, we extracted the time-series of regions defined from the Automated Anatomical Labeling atlas [33] (90 ROIs in the Cerebrum). Based on prior studies (chapter 1), the 90 ROIs were classified into two groups [88]. The first subgroup included task-relevant visuomotor (MV) ROIs (28 regions), whereas the second subgroup included complementary task-orthogonal non-visuomotor (NMV) ROIs (62 regions). The working hypothesis is that being the first group task-relevant, the designed protocol should allow to better highlight the networks similarities between genetically identical twins.

We computed the DFC matrices using Pearson correlation $\rho = corr(x[t, t+\triangle t], y[t, t+\triangle t])$ with a sliding window of $\triangle t$=30 TR (60 sec) and a step size of 8 TR (16 sec) [162]. This resulted in 41 DFC matrices W_t describing the modulation of connectivity along the entire recorded sequence. In total, we had $(N \times N) \times T$ matrices for each subject, where T is the total number of window (41 in our case) and N is the number of regions (N=90 when considering ALL ROI's, N=28 for MV ROI's and N=62 for NMV ROI's). We then applied the Fisher r-z transform to normally distribute the data, keeping only positive values of correlations as commonly done in FC analysis [96]. Figure 6.1. shows how dynamical functional connectivity is built compared to the common approach which is usually known as static functional connectivity. As described in chapter 1, there was no effect of task between the two groups. So, for each subject, we averaged the DFC matrices across tasks. Taking the average led to higher SNR and hence more stable results.

6.1.2 Graph Laplacian and Riemannian Manifold

Several methods allow us to take advantage of graph modularity from symmetric undirected weighted matrices W_t [163, 164]. In our case, we transformed DFC matrices into graph Laplacian to enjoy some properties. Specifically, for each temporal window, we used the

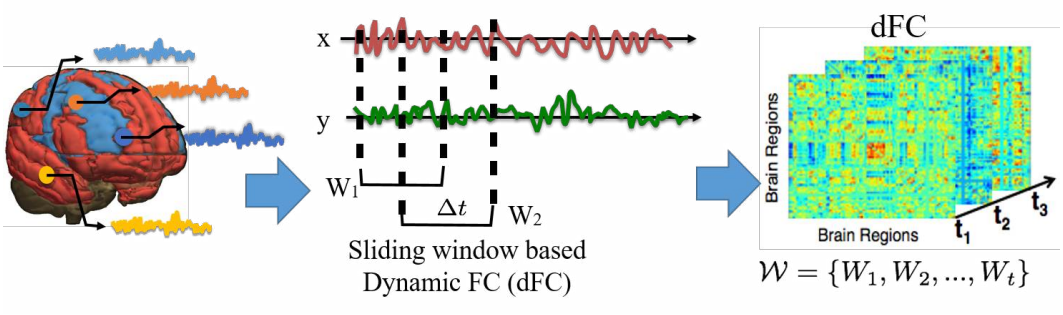

Figure 6.1: An illustration of brain parcellation, extraction of time series signal and computation of Dynamic Functional Connectivity using sliding window based approach

Normalized Symmetric Graph Laplacian [165] defined as:

$$L_t = D_t^{-1/2}(D_t - W_t)D_t^{-1/2} \tag{6.1}$$

where D_t=$diag(\Sigma_j w_{ij})$ is the degree matrix of W_t. The main advantage of graph Laplacian matrices is that they are symmetric and positive semi-definite(PSD) and it can be easily regularized to a positive definite (PD) matrix by the modification $\hat{L}_t = L_t + \gamma I$, where $\gamma > 0$ [1] is a regularization parameter and I is the identity matrix. The set of PD matrices describe a Riemannian manifold, which is a descriptor richer than a vector space [41]. The advantage is that graphs can be compared with a geodesic distance. Indeed, a metric based on Euclidean distance is suboptimal when applied to positive semi-definite matrices [42], because this metric is not responsive to the geometry in the data. For this reason we used a component of Frechet distance, which is a metric in the space of positive semi-definite matrices [62] known as Wasserstein distance (for details check section 2.4):

$$d_t^2(W_{t(x)}, W_{t(y)}) = tr[W_{t(x)} + W_{t(y)} - 2(W_{t(x)}W_{t(y)})^{1/2}] \tag{6.2}$$

Where $W_{t(x)}$ and $W_t(y)$ are two DFC matrices of subject x and y respectively, and tr is the trace operator. By using Equation 2.11, we determined the distance between each paired DFC matrices for each twin couple. So for each twin pair, we have a distance-vector of length T=41 which actually represent the evolution in time of connectivity similarity. For the sake of comparison, we have also performed the same experiments with Euclidean distance. The temporal dynamics of the distribution of Wasserstein distances (mean and standard deviation) of the two groups is shown in Figure 6.2 (above), which shows that connectivity of MZ pairs is closer in each time-epoch as compared to DZ pair. Smaller distance represents more similar connectomes. In particular, this figure shows that MZ twins are more similar to each other than DZ twins independently of the networks considered. Figure 6.2

[1] Imposing $\gamma = 10^{-9}$ in our settings [22]

(below), shows the distribution of Euclidean distance of two groups. It can be seen that with Euclidean distance-based difference between two groups is not considered good.

Moreover, with Wasserstein distance, independently of the twin class, when considering the task-related network (middle) the distances are generally smaller than distances between networks including task-orthogonal ROIs (bottom) and all ROIs (top). Finally, not only the distances are generally smaller when considering the MV ROIs, but also the variance is much smaller, especially for MZ twins.

6.1.3 Feature Selection and Classification

In this analysis, our working hypothesis was that connectivity between MZ twin pairs would be more similar than between DZ pairs, and particularly for the MV sub-network. By encoding similarity as the distance between functional networks, we, therefore, expected to predict twin groups using DFC network distances as features. The classification was performed on all possible combination of ROIs (ALL, MV & NMV) using the linear support vector machine (SVM) classifier with leave-one-out cross-validation (LOOCV).

Laplacian Score for Feature Selection: To address the curse of dimensionality in our dataset characterized by a number of features (41) greater than the number of observations (13), we adopted Laplacian Score (LS) based unsupervised feature selection [166]. Our aim was to choose the DFC time-epochs best differentiating the two groups. In this feature, selection approaches each feature was ranked by computing relevance based on variance and separability e.g. by computing the Laplacian score of each feature in the nearest neighbour graph by their formula and then ranking the features based on highest score e.g. feature with highest Laplacian score will be ranked 1st and vice versa. The LS worked on the basic idea of assessing the features according to their locality preserving power. We choose the top 10 feature using this method and tested the classifier with different numbers of selected feature (3 to 10) to check the stability of the system.

Due to the small number of samples, we did not use any supervised feature selection approach as not enough data was available. In order to check the significance level of the performance of our classifier, we also implemented a permutation test on labels (twin class). For this, we generated a null distribution by randomly changing the labels 8000 times and in each iteration, we performed SVM classification using LOOCV and computed the cross fold accuracy. This test was useful to validate the selection of features and to quantify the significance of classification accuracy. Figure 6.3 Illustrate the pipeline used for this analysis.

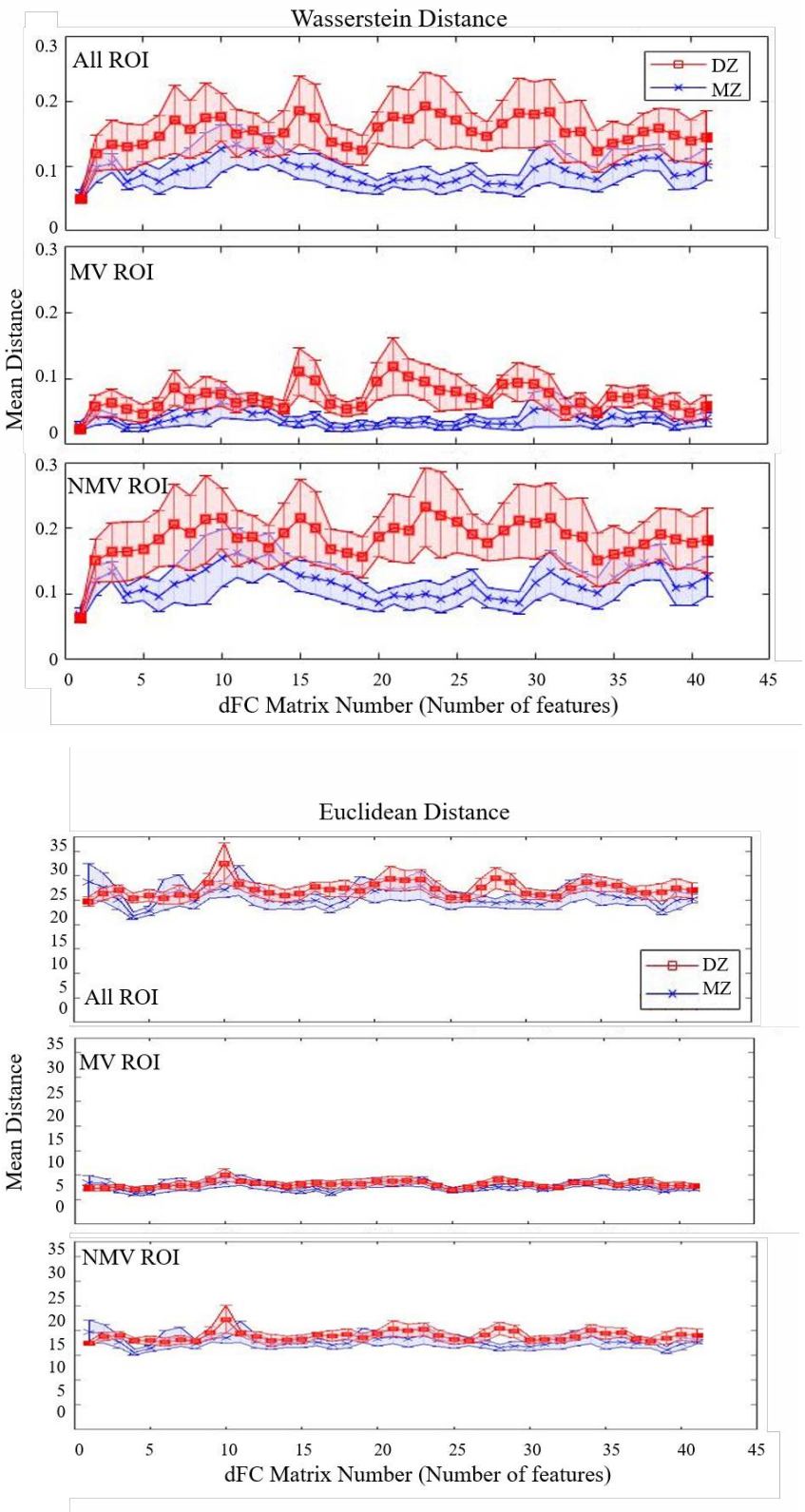

Figure 6.2: (Mean graph distance & standard error of each DFC matrix computed through the sliding window over fMRI time-series with Wasserstein distance(top) and Euclidean distance(bottom). Marker in blue color shows the mean distance and standard error for MZ group and red square shows mean distance and standard error for DZ group. The plots are respectively for the network with (top) all nodes (middle) visuomotor nodes (bottom) non-visuomotor nodes.

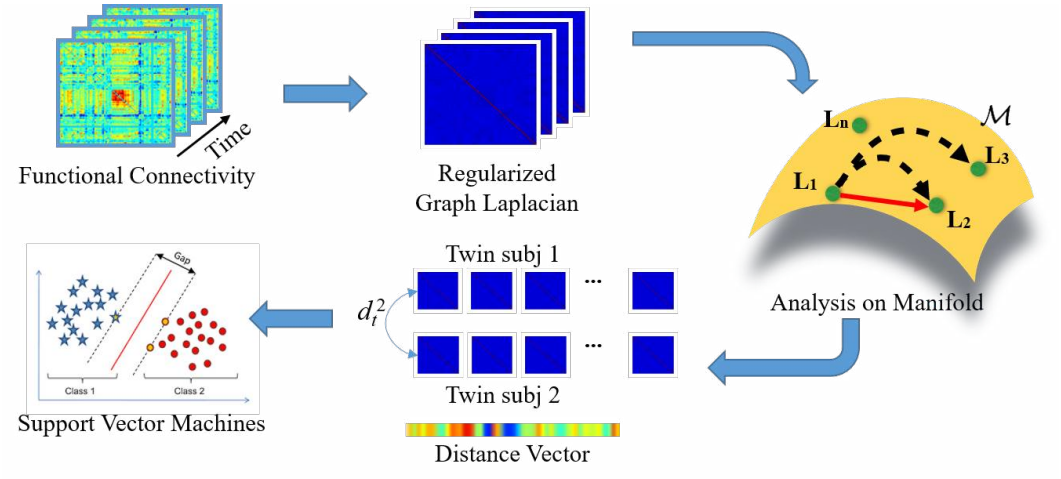

Figure 6.3: Proposed pipeline start with conversion of dFC into dynamic Laplacian connectivity, then computation of distance between each dFC of each pair on Riemannian manifold (distance vector building) and then classification based on distance vector.

6.1.4 Results and Discussion

Figure 6.4 shows the SVM classification accuracy for different combinations of networks and feature selection with Wasserstein distance (above) and Euclidean distance (below). The plot in the Figure 6.4.a (above) describes the cross fold SVM accuracy when using all 41 distances as a feature for the three networks (All, MV & NMV). As seen the DFC of task-relevant ROI (MV) are giving better accuracy (77%) as compared to task-orthogonal (NMV) & All ROIs (62%) with Wasserstein distance. For Euclidean distance, it can be seen that accuracy with all type of network is below chance level (50%) and also with feature selection accuracy is very less. Due to the high dimensionality issue of the data (more features than observations) these results are not statistically significant. So feature selection was performed for all three networks, selecting from 3 to 10 features by using unsupervised Laplacian Score. The classification was then performed using the same schema (linear SVM with LOOCV).

The cross-fold classification accuracy of SVM classifier with a different number of features for all three networks with Wasserstein distance and Euclidean distance can be seen in Figure 6.4.b (above) It can be observed that with a small number of features (3 to 5) when comparing the network with task-relevant nodes only there is a surprising rise of performance (accuracy of 92%) whereas this doesn't happen when using task-orthogonal nodes. This shows that task-relevant ROIs are playing a vital role in discriminating the two groups. FOr Euclidean distance accuracy with the selected feature is still not compare with the Wasserstein distance approach. It can also be seen that as the number of features increases, the cross-validation SVM accuracy declines. This is due to the integration of some redundant

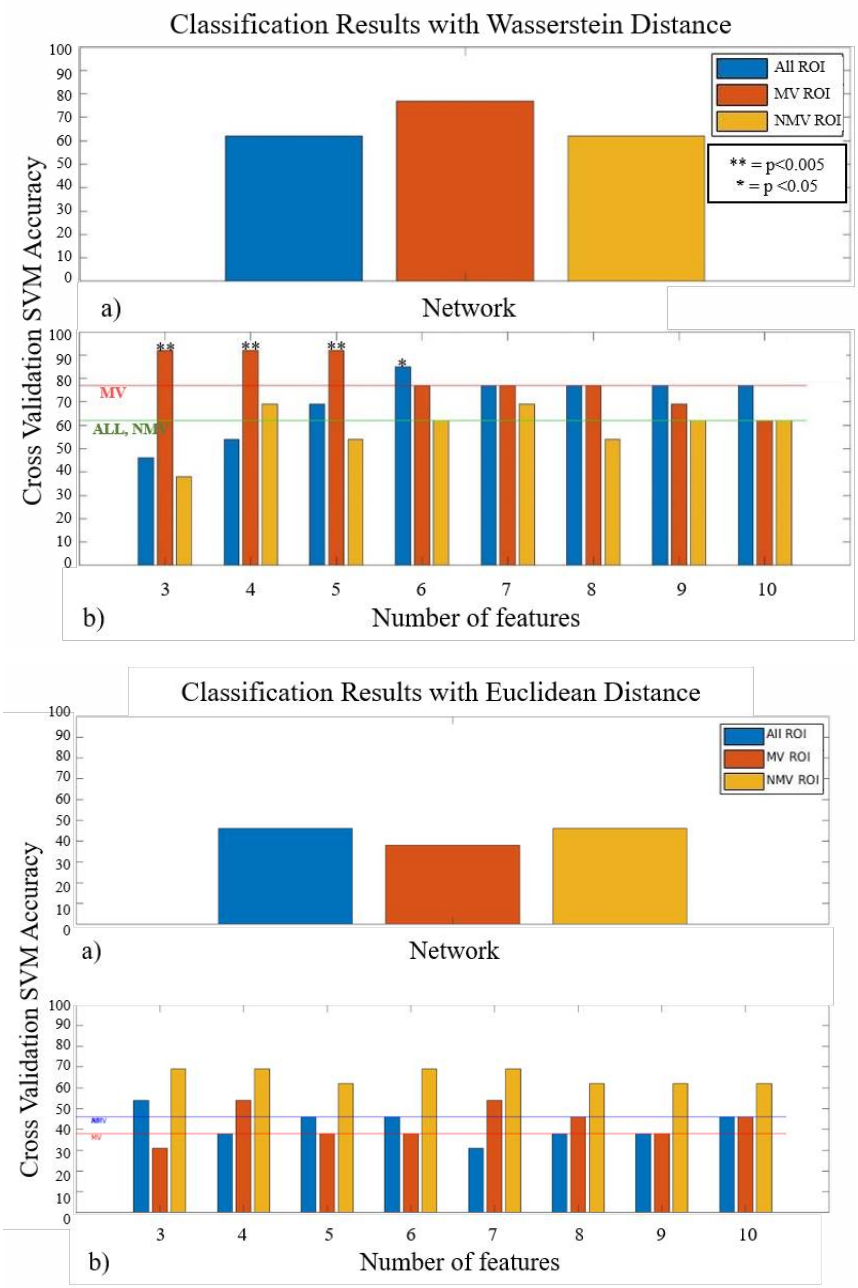

Figure 6.4: (a) Classification accuracy obtained with Wasserstein distance (above) and Euclidean distance(below) for all features for networks composed of different sets of ROIs: (b) Classification performance with feature selection with Wasserstein distance and Euclidean distance. Stars on bar represent the significance level obtained through permutation test.

features.

In order to assess the significance of our results obtained with feature selection, we implemented a permutation test on labels. The results of the permutation test along with the

accuracy with each feature selection are represented in figure 2 with ** on the bars having significance level p_value <0.005 and * with p_value <0.05. P values represent the ratio of the total number of accuracy values greater than the accuracy value with true label divided by the total number of permutations (8000 in our case). These results support the hypothesis that genetic similarity has an impact on brain network similarity which can be examined using the DFC especially for the purpose of classification between two groups.

To conclude, the results of our study demonstrate that genetic influences on brain network profiles can be detected using dynamic functional connectivity. In MZ twin pairs, the task-relevant visuomotor networks are more similar than they are in DZ twin pairs and this genetic difference is very useful to be used as a feature to classify a pair of twins either MZ or DZ. On the other hand, when considering only the task- orthogonal non-visuomotor network or whole network, they don't play a significant role in terms of classification. These findings imply that Zygocity modulates the connectivity of task-relevant networks which can be used to distinguish one group from others.

In summary, in this work, we have presented a computational framework to compare the dynamic functional connectivity of twins. The proposed approach includes a novel method to computes the difference between two graphs by finding the graph Laplacian of DFC matrices (considering only positive correlations) and then applying Wasserstein distance to find the similarity between DFC matrices. This similarity information was adopted to compute the features in the classification task, where the aim was to discriminate between MZ and DZ twins. To address the issue of dimensionality an unsupervised feature selection based on Laplacian score was used to reduce the number of features. This approach allowed us to obtain better accuracy with linear SVM classifier.

6.2 Classification of Dynamic Brain Connectivity through Geodesic Clustering

In this section, we have investigated Twin's task-based fMRI, which encourages the identification of integration mechanism between specific task-related brain regions and is useful to identify task-related networks in brain connectivity [34, 108]. The main purpose of this study is to evaluate if there are DFC patterns shared among twins, allowing discriminating twin pairs from unrelated pairs and to investigate the effect of genetic heritability on the dynamics of functional brain networks. In [Chapter 2 reference], we have shown that differentiation between two groups was measurable when using the graph Laplacian representation of non-dynamic FC matrices, which transforms the representation of data into the smoothed space of positive semi-definite matrices [167]. In section 6.1 we have further extended the use of graph Laplacian and Wasserstein metric to classify between MZ & DZ twin pairs using DFC matrices. In this section, we want to investigate the DFC analysis exploiting the concept of brain states. To perform this investigation, we exploited the similarity of DFC patterns associated with the brain states of the two groups (twins and non-twins). To this aim, we clustered the DFC matrices into reference states and then we used a compact representation to perform classification.

Methods described in [113, 161], exploiting clustering to generate a set of reference states and are based on similarities computed in a vectorial space. Using metrics on the vectorial space – like the Euclidean distance frequently used in k-means – is sub-optimal section 6.1. We know, however, that FC matrices can be managed to form a manifold of positive definite matrices, and a more appropriate choice of similarity is to use a geodesic metric defined on the smooth manifold. Therefore, in our approach, we used a geodesic metric both to cluster the matrices with k-means and to extract the features to be used by the classifier. For this analysis, we used the same twin's task-based data defined in [chapter 2 reference] which includes 13 Twin pairs. We made some experiments both using the Euclidean metric in a vectorial space and a geodesic metric (Log Euclidean distance) on the Riemannian space of symmetric positive definite matrices.

6.2.1 Dynamic Functional Connectivity Estimation

Given N the number of regions in the atlas (in our case AAL is made by N=90) we estimated the $N \times N$ covariance matrices $\Sigma_i(w)$, for all subject $i = 1 \dots M$, (M is the total number of subjects) and for all sliding windows $w = 1 \dots W$ over the fMRI time-series (W is the total number of windows). In our experiments, we used a sliding window of size $\triangle t$=30 TR (60 sec) and a step size of 4 TR (8 sec) [167]. This resulted in W=83 DFC matrices Σ_i describing the modulation of connectivity along the entire recorded sequence. Due to the

relatively small windows size, the estimation of the covariance matrices might be unstable and heavily affected by the limited amount of information. To overcome this issue a more robust estimate of the covariance with small data can be obtained from the estimate of a sparse version of the inverse of the covariance matrix $\Sigma_i^- 1(w)$ [168, 169, 37]. This sparse precision matrix can be obtained regularizing the estimated parameters with the graphical LASSO as described in [23]. This method has proven to be very effective when there is a limited number of observations at each node [114, 170], such as in our case where we have small intervals of fMRI scan.

The covariance matrices are always guaranteed to be symmetric & positive semi-definite (SPD), however, in real applications they are frequently also symmetric & positive definite (PD). If some matrices are not PD we can apply a small regularization ($\Sigma_i = \Sigma_i + \lambda I, \lambda = 10^{-9}$ in our settings) making them PD. In this way, they form a Riemannian manifold of PD matrices [170] which enable us to analyze the DFC matrices on the manifold instead of using the vector space [45]. To take full advantage of the manifold structure of PD matrices, it is essential to consider a geodesic distance, which measures the shortest path between two points (two matrices in our case) along the smooth and curved manifold [45]. There are some possible alternative geodesic distances on the Riemannian manifold of PD matrices [45, 6], we decided to adopt the Log-Euclidean distance (see Eq. 2.8), which is simple, and fast to compute.

As described in our previous work chapter 2, there is no effect of task (left and right hand) between the two groups. So, for each subject, we averaged the DFC matrices across tasks by using geodesic mean Eq. 2.9 [60] so that the geometric nature of matrices is maintained.

6.2.2 The Dynamic States and Geodesic Clustering Analysis

In order to define a set of states describing intrinsic brain network patterns, we have used geodesic k-means clustering on SPD matrices [110] to associate a state to each cluster. To initialize the cluster centroids, we first selected a set of exemplar matrices [114] from the data (8 matrices per subject in our case) maximizing the distance from the rest of the exemplars of the same subject. The geodesic k-means was then applied on the set of exemplars to obtain the initial centroids, which were then refined running again the geodesic k-means on all DFC matrices of all subjects. In order to run the geodesic k-means, we used the Log-Euclidean distance [45] as defined in Eq. 2.9 for which the mean of multiple covariance matrices can be computed in a closed form:

In order to choose the optimal number K of clusters we used two criteria. The first criterion

was based on the minimization of the Sum of Squared Error (SSE):

$$SSE = \sum_{i=1}^{k} \sum_{\Sigma \in C_i} d^2(m_i, \Sigma), \tag{6.3}$$

where Σ is a dFC matrix associated to cluster C_i and m_i is the corresponding centroid.

The second criterion was based on the necessity of having in any cluster some matrices for all subjects, due to the encoding framework explained below. We, therefore, computed the SSE by using Eq. 6.3 ranging over a number of clusters $(K = 2 \dots 10)$. In our case, the best solution fulfilling with the two criteria resulted to be with $K = 2$. This result also supports the hypothesis that in a task-based fMRI there appear at least two macrostates, one is a task-related state and the second one is a no-task state.

6.2.3 Feature Extraction and Classification

The working hypothesis is that dynamic connectivity between twin pairs would be more similar than between unrelated pairs, and based on this we can classify pairs either as twins or as unrelated. Therefore, we need to encode the subject's similarity taking into account the brain state. To this aim, subjects' representatives were computed for all clusters. More specifically, the subset of all DFC matrices of a subject associated with a cluster was averaged with a geodesic mean Eq. 2.9 creating a subject representative for that cluster. In this way, a subject has a representative for each cluster.

At this point, we could compute the features characterizing the similarity between the subjects. For each pair of subjects, we measured the inter-subject geodesic distance between the two subject representatives of each cluster. In addition, we computed the geodesic distance of each subject representative from the cluster centroids. Therefore, for each pair of subjects (twins or unrelated), there are 3 distances per cluster (features) and in our experiments, the data representation included a total of 6 features because we have K=2 clusters. Figure 6.5 illustrate the phenomena of extracting the features in term of distance from each centroid. In our data set, we have 13 twin pairs, corresponding to 26 subjects that can be recombined to form 312 unrelated pairs. In short, we have a dataset composed of 13 samples from the twin's class and 312 samples from the unrelated class. Due to the high unbalanced dataset, we opted to use the weighted SVM [171] with three cross-fold validation. To this aim, we divided our data into 3 chunks randomly selecting the samples while maintaining the proportion between the classes (each fold was composed by 104 samples from the unrelated pairs and 4 from twin's pairs). For statistical purpose, we repeated 100 times this cross-validation procedure with the randomized selection of folds. We evaluated the results in terms of average accuracy, precision, recall, F1 score and confusion matrix.

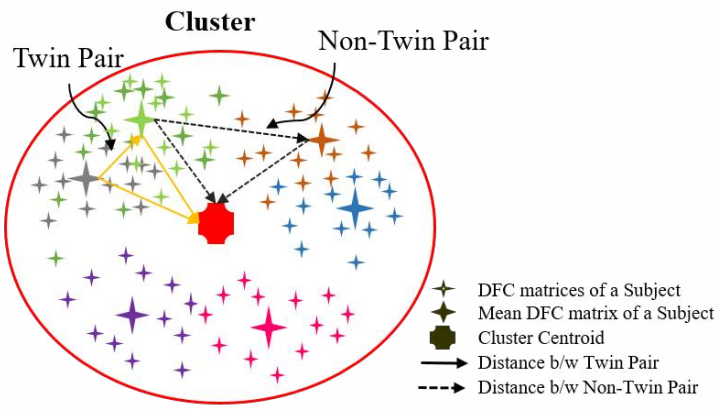

Figure 6.5: An illustration of extraction of feature in term of the distance between each subject representative, and between-subject representative and cluster centroid.

In all our experiments, all distances between graphs and means of graphs were computed using the Log-Euclidean distance Equation 2.8 and the corresponding geodesic mean 2.9 respectively. However, for the sake of comparison, we performed identical experiments using the Euclidean distance and the corresponding Euclidean mean.

6.2.4 Results and Discussion

Figure 6.6 shows the results of classification with the weighted SVM when using the Log-Euclidean distance (blue bars) and the Euclidean distance (red bars). It can be observed that using the geodesic metric to describe the data considerably boosts the performance during classification, i.e., during the exploitation of the encodings. In particular, the accuracy with "geodesic encoding", 87.21%, is much higher than the "Euclidean encoding" accuracy, 66.35%. Similar differences can be observed for the precision (88.35% versus 67.42%) and F1 score (92.92% versus 79.14%). Higher and similar recall for both metrics could be due to the higher unbalance in the classes. The embed table in Figure 1 summarizes these results.

The mean confusion matrix for both distance metrics is given in Table 6.1. It can be observed that when using geodesic distance during the data encoding the rate of correctly classified pairs is much better than using Euclidean distance. These results strongly support the fact that the use of Euclidean metric on symmetric positive definite matrices is suboptimal. Hence, a better way to compare and process the undirected weighted graphs described by SPD is to use a geodesic distance on the Riemannian space. Figure 6.7 show the centroid of both cluster and it can be observed that activation of connection between different regions of the brain is more prominent in cluster 2 as compared to cluster 1. These connectivity matrices also support our assumption of two brain states for this task-based fMRI, one with no-task (cluster 1) and second is task state (cluster 2).

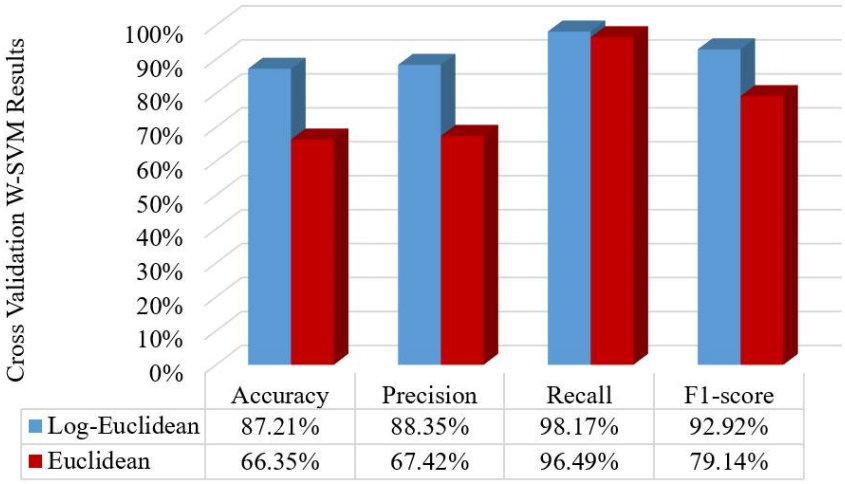

Figure 6.6: Comparison of average performance of classification with weighted linear SVM classifier with Log-Euclidean distance (blue bars) and with Euclidean distance (Red bars)

Log-Euclidean Distance				**Euclidean Distance**		
		Predicted Class			**Predicted Class**	
		Non-Twin Pair	Twin Pair		Non-Twin Pair	Twin Pair
Actual Class	Non-Twin Pair	275	37	Non-Twin Pair	210	102
	Twin Pair	4	9	Twin Pair	8	5

Table 6.1: Average confusion matrix showing the performance of classification when using the Log-Euclidean distance or the Euclidean distance

To conclude, in this work, we have presented a novel computational framework, which allows distinguishing between twins and unrelated pairs of subjects using their dynamic functional brain connectivity. To this aim, we designed a specific encoding of graphs into subjects' similarities, exploiting the concept of geodesic metric on the Riemannian manifold of SPD matrices. In particular, for the encoding of data, we derived a subject-wise graph similarity representation exploiting a geodesic k-means clustering. Indeed, the algorithm uses the Log-Euclidean metric on the space of functional brain graphs. Once the clusters were generated, the Log-Euclidean metric was also used to calculate the similarity of two subjects in terms of the distance between subjects and distance from cluster centroid. These distances were used as features for the data representation. Due to the highly unbalanced dataset to solve the classification task we used the weighted SVM.

In order to evaluate whether, beyond having a good estimation of covariance matrices, it is important to use metrics working on the space of data, we made an identical experiment using the Euclidean distance in place of the geodesic distance. The results of our study clearly demonstrate that use of Euclidean distance is not the best choice, as it is not properly managing the complex structure of graphs, indeed the classification performance is boosted when using the geodesic distance. This study also reveals that a careful encoding of the

dynamic functional connectivity allows a clear distinction of twin pairs from non-twin pairs.

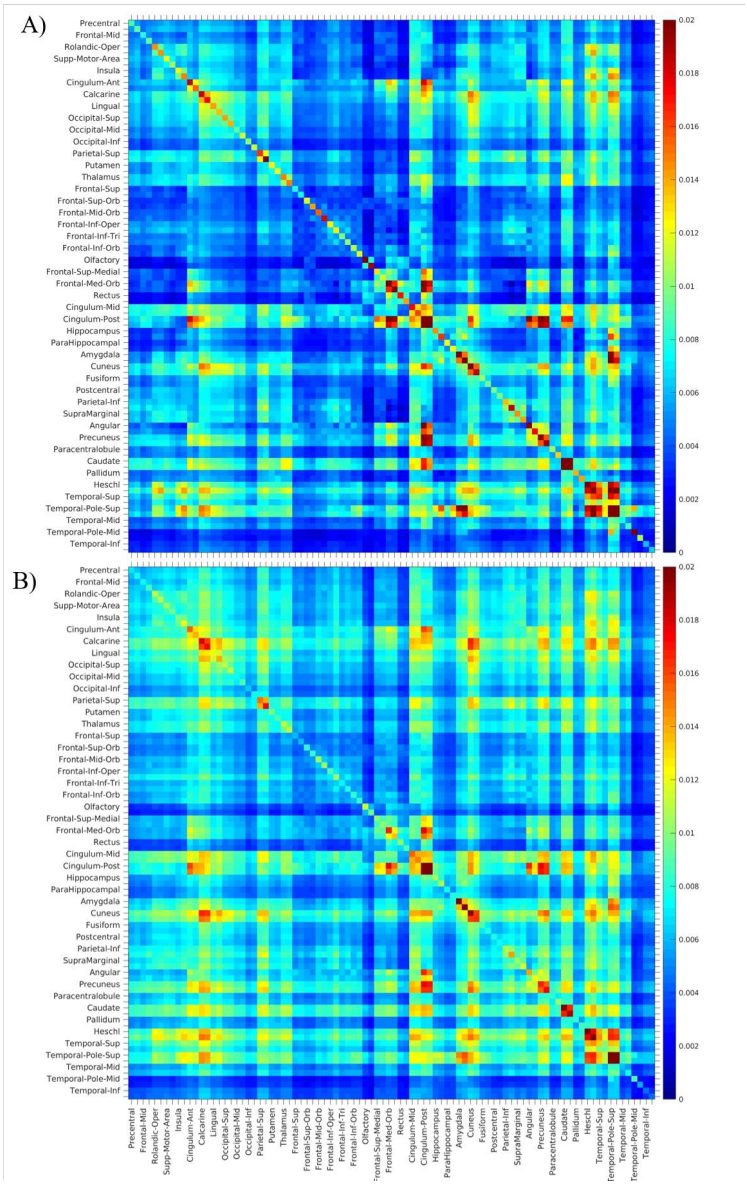

Figure 6.7: Mean Connectivity matrix of A) cluster 1 and B) cluster 2 computed by using Equation 2.9

Chapter 7

Conclusions and Future Work

7.1 Summary and Conclusions

This thesis is motivated by the increasing interest in brain network analysis. The emerging advances in magnetic resonance imaging allow obtaining multiple images associated with different characteristics of the brain, like the networks functional or structural organization. All of these aspects can be investigated with graph-based approaches, which however have some constraints or are not able to respond to all neurological questions. Indeed, the key goal of this work has been the implementation of new methods to cover open questions in brain connectivity. Regardless of the different applications covered in this thesis, the common insight between them is the graph-based inspection carried out to study the brain functional connectivity.

A first contribution regards the investigation of functional connectome (FC) of twin's task-induced data in terms of analysing the effect of genetic heritability on the functional organization of the brain. Taking into account the symmetric and positive definite properties of covariance-based functional connectivity matrices which allow considering these matrices on Riemannian manifold, we performed a similarity analysis between pairs using geodesic methods on Riemannian manifold. Geodesic approaches follow the geometry of manifold and hence gives a more appropriate estimation of similarity between connectome as compared to Euclidean distance which is suboptimal to use. Moreover, we have also introduced a framework to compute heritability of twin's based on falconers formula using functional connectome. The results of this study demonstrate how our analytic innovations reveal genetic influences on brain network profiles. In monozygotic twin pairs, the task-relevant networks are more similar than they are in dizygotic twin pairs which clearly depicts the influence of genetic sharing. On the other hand, there is no significant difference between these two groups when considering the task-orthogonal network. These findings imply that zygosity modulates the connectivity of task-relevant networks, emphasizing a value of task-based fMRI. The result of the heritability index computed using the proposed method for the task-orthogonal network is comparatively high as compared to task-orthogonal network. Which support the effect of genetic mediation in twins especially in the task-induced network.

As a second aim of this work, we have developed a novel classification framework to discriminate brain functional connectome between healthy and pathological subjects. Classification of brain graphs between healthy and pathological subjects is an important task to solve. In order to obtain a classifier able to capture the discriminative patterns of brain connectivity, we have worked out on a computational framework, which allows the classification of HC and patients using static functional connectome. In this work, we proposed geodesic $k-means$ clustering method which clusters the covariance-based functional connectome by taking advantage of the properties of SPD matrices: in this context, using geodesic metrics proved to be superior to the Euclidean approach. In particular, cluster centroid considered as reference connectome and classification features has been constructed with a subject-wise graph similarity representation by using a geodesic metric between each sample and each reference connectome. This scheme surpasses standard approaches for graph classification and it is also effective to find irregular connectivity patterns which can be potentially notable in patients affected by some syndromes like Autism or Schizophrenia.

Furthermore, as an extension to our previous work, we suggested another computational framework to distinguish between HC and different phenotype of patients affected by multiple sclerosis using resting-state functional connectome. The descriptors to be used by the classifier were determined with dominant set clustering approaches, again considering the properties of SPD matrices we used geodesic metric. Data encoding was performed in a similar way by computing the similarity between each sample and reference networks (cluster centroids). For classification purpose, geodesic metric offers an optimal data representation allowing a better system performance. Results also support the fact that encoding of functional connectivity matrices using DS clustering algorithm always gives better results as compared to the geodesic k-means clustering algorithm. To highlight the variations between two groups of HC and MS, a difference of reference connectome of both groups in each prominent clusters (selected based on sensitivity analysis) was computed. These variations show the prominent changes in functional connectivity between two groups and helped to identify neuro-markers. From a clinical point of view, the results of this study showed that MS patients, considered as a whole, were well distinguished from HC in terms of functional connectivity and this approach might be rewarding in monitoring disease evolution and optimizing patients management

Finally, we developed a framework to analyze the dynamic functional connectivity (DFC). Furthermore, sub-network interpretation in dynamical systems, as fMRI functional connectivity, allows revealing different patterns of connectivity can arise over time, especially during the task-induced fMRI. To deal with this problem, our proposed method is to inspect such dynamics by computing the similarity between connectome in each time frame window of dynamic functional connectivity, which also allows measuring the temporal variability of the

brain organization. In particular, for the encoding of DFC data, we derived a subject-wise graph similarity representation between twin's pair (to classify between MZ & DZ twin pair) and we also consider exploiting a geodesic k-means clustering (to classify between Twin's and un-related pair). The results of our study clearly demonstrate that, that a careful encoding of the dynamic functional connectivity allows a clear distinction of combinations of pairs, especially when using the geodesic distance on SPD representation of data.

7.2 Future Directions

One of the primary challenges in connectomics is the interpretation of alteration occurs in functional connectivity under different circumstances. Our work on a geodesic method based functional connectivity analysis is an attempt to describe the modular structure of the brain in term of functionality. Besides functional, structural connectivity also plays a vital role in term of comparing the alterations occurs in brain networking. Neurological studies, independently from the scale, indicate that the two connectivities (structural and functional) are not directly associated, suggesting an intense restructuring of the information flow among the brain regions. Indeed, till today very few efforts have been made to implement mathematical models which could be able to merge the two aspects of the brain connectivity. Despite the capacity of each proposed model, the key constraints are the prior assumptions which have to be made to define a relation between the connectomes. This reality makes it essential to continue the analysis of brain connectivity focusing at obtaining multiple-views of the brain, in order to test predictive models with the real connectivity estimated from MRI images.

In future, we can extend the current framework of clustering based encoding to perform identification of neuro-marker and analysing alteration in structural connectivity between HC and patient affected with multiple phenotypes of multiple-sclerosis disease. Moreover, we can merge the encoding results obtained with functional and structural connectivities and try to implement a better classification method. Connectivity data is usually very high dimensional which is normally difficult to process and analyze. So as a future perspective we can implement a dimensionality reduction method applied on SPD matrices of functional or structural connectome directly on manifold (without projecting to a tangent space, which is normally being practised). With this approach, we will be able to reduce the amount of data and hence can perform the clustering and encoding in a more appropriate way.

Regarding functional connectivity, recently, it has been revealed that spontaneous fluctuations can arise across time, showing the strong dynamical nature of brain functional connectivity even at resting. This view of functional connectivity opened a new era of research and could be helpful in answering many unsolved questions, such as how to model these fluctuations in order to investigate the dynamical restructuring of the brain. Our proposed method of

comparing the dynamic functional connectivity matrices using geodesic metric-based methods goes in that direction and introductory results show that the functional connectivity is influenced by the brain states during task-induced fMRI. We intended to extend our analysis of comparing HC group with multiple-sclerosis patients using the dynamic functional connectivity method. In this approach, we want to implement DS clustering to perform clustering on dynamic matrices and to compute the cluster which represents unique brain pattern (brain state). Then we can compare obtained brain states of two groups to get the neuromarkers occurring at different brain states and hence can easily find the regions responsible for discrimination.

In conclusion this dissertation proposed novel methods to carry out different connectivity analysis. Each of the suggested solutions takes into account the graph-based properties of the functional connectome, and of course, they lay strengths and limitations. However, the proposed methods can be extended both from a methodological perspective and from the neurological side, making them suitable for different problems.

Bibliography

[1] Marco Catani, Michel Thiebaut de Schotten, David Slater, and Flavio Dell'Acqua. Connectomic approaches before the connectome. *NeuroImage*, 80:2–13, 2013.

[2] John Ashburner and Karl J Friston. Voxel-based morphometry—the methods. *NeuroImage*, 11(6):805–821, 2000.

[3] Joset A Etzel, Valeria Gazzola, and Christian Keysers. An introduction to anatomical roi-based fmri classification analysis. *Brain research*, 1282:114–125, 2009.

[4] Elena A Allen, Eswar Damaraju, Sergey M Plis, Erik B Erhardt, Tom Eichele, and Vince D Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, page bhs352, 2012.

[5] Luca Dodero, Sebastiano Vascon, Vittorio Murino, Angelo Bifone, Alessandro Gozzi, and Diego Sona. Automated multi-subject fiber clustering of mouse brain using dominant sets. *Frontiers in neuroinformatics*, 8:87, 2015.

[6] Luca Dodero, Fabio Sambataro, Vittorio Murino, and Diego Sona. Kernel-based analysis of functional brain connectivity on grassmann manifold. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 604–611. Springer, 2015.

[7] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS Comput Biol*, 1(4):e42, 2005.

[8] Olaf Sporns. The human connectome: a complex network. *Annals of the New York Academy of Sciences*, 1224(1):109–125, 2011.

[9] Edward T Bullmore and Danielle S Bassett. Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*, 7:113–140, 2011.

[10] W Lange. Cell number and cell density in the cerebellar cortex of man and some other mammals. *Cell and tissue research*, 157(1):115–124, 1975.

[11] Vernon B Mountcastle. *Perceptual neuroscience: the cerebral cortex*. Harvard University Press, 1998.

[12] Anders M Dale, Bruce Fischl, and Martin I Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999.

[13] Bruce Fischl, Martin I Sereno, and Anders M Dale. Cortical surface-based analysis: Ii: inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, 1999.

[14] Junfeng Sun, Shanbao Tong, and Guo-Yuan Yang. Reorganization of brain networks in aging and age-related diseases. *Aging and disease*, 3(2):181, 2012.

[15] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52(3):1059–1069, 2010.

[16] Cheryl L Grady, Maura L Furey, Pietro Pietrini, Barry Horwitz, and Stanley I Rapoport. Altered brain functional connectivity and impaired short-term memory in alzheimer's disease. *Brain*, 124(4):739–756, 2001.

[17] Muhammad Abubakar Yamin, Michael Dayan, Letizia Squarcina, Paolo Brambilla, Vittorio Murino, V Diwadkar, and Diego Sona. The heritability of brain network profiles is contextual: Converging evidence from two analytical frameworks. *Submitted in (Nov 2020) in Proceedings of the National Academy of Sciences*, 2020.

[18] Muhammad Abubakar Yamin, Michael Dayan, Letizia Squarcina, Paolo Brambilla, Vittorio Murino, V Diwadkar, and Diego Sona. Comparison of brain connectomes using geodesic distance on manifold: A twins study. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1797–1800. IEEE, 2019.

[19] Muhammad Abubakar Yamin, Jacopo Tessadori, Muhammad Usman Akbar, Michael Dayan, Vittorio Murino, and Diego Sona. Geodesic clustering of positive definite matrices for classification of mental disorder using brain functional connectivity. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–5. IEEE, 2020.

[20] Muhammad Abubakar Yamin, Paola Valsasina, Michael Dayan, Sebastiano Vascon, Jacopo Tessadori, Massimo Filippi, Mara A Rocca, Vittorio Murino, and Diego Sona. Encoding brain networks through geodesic clustering of func-tional connectivity for multiple sclerosis classification. In *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR 2020)*. IEEE, 2020.

[21] Muhammad Abubakar Yamin, Paola Valsasina, Michael Dayan, Sebastiano Vascon, Jacopo Tessadori, Massimo Filippi, Mara A Rocca, Vittorio Murino, and Diego Sona. Multiple-sclerosis phenotype classification and neuro-marker identification using manifold approach. *Submitted in NeuroImage (Nov 2020)*, 2020.

[22] Muhammad Abubakar Yamin, Michael Dayan, Letizia Squarcina, Paolo Brambilla, Vitorio Murino, Vaibhav Diwadkar, and Diego Sona. Investigating the impact of ge-

netic background on brain dynamic functional connectivity through machine learning: a twins study. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 1–4. IEEE, 2019.

[23] Muhammad Abubakar Yamin, Michael Dayan, Letizia Squarcina, Paolo Brambilla, Vittorio Murino, V Diwadkar, and Diego Sona. Analysis of dynamic brain connectivity through geodesic clustering. In *International Conference on Image Analysis and Processing*, pages 640–648. Springer, 2019.

[24] Peter A Bandettini. Twenty years of functional mri: the science and the stories. *NeuroImage*, 62(2):575–588, 2012.

[25] Alistair M Howseman and Richard W Bowtel. Functional magnetic resonance imaging: imaging techniques and contrast mechanisms. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 354(1387):1179–1194, 1999.

[26] Karl J Friston, Andrew P Holmes, Keith J Worsley, JP Poline, Chris D Frith, Richard SJ Frackowiak, et al. Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 2(4):189–210, 1994.

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

[28] VD Calhoun, T Adali, GD Pearlson, and JJ Pekar. Spatial and temporal independent component analysis of functional mri data containing a pair of task-related waveforms. *Human brain mapping*, 13(1):43–53, 2001.

[29] Bharat B Biswal and John L Ulmer. Blind source separation of multiple signal sources of fmri data sets using independent component analysis. *Journal of computer assisted tomography*, 23(2):265–271, 1999.

[30] VD Calhoun, T Adali, GD Pearlson, and JJ Pekar. A method for making group inferences from functional mri data using independent component analysis. *Human brain mapping*, 14(3):140–151, 2002.

[31] Nicole Lazar. *The statistical analysis of functional MRI data*. Springer Science & Business Media, 2008.

[32] Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1):253–258, 2003.

[33] Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated

anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage*, 15(1):273–289, 2002.

[34] Gaël Varoquaux and R Cameron Craddock. Learning and comparing functional connectomes across subjects. *NeuroImage*, 80:405–415, 2013.

[35] Lorena Deuker, Edward T Bullmore, Marie Smith, Soren Christensen, Pradeep J Nathan, Brigitte Rockstroh, and Danielle S Bassett. Reproducibility of graph metrics of human brain functional networks. *NeuroImage*, 47(4):1460–1468, 2009.

[36] Stephen C Strother. Evaluating fmri preprocessing pipelines. *Engineering in Medicine and Biology Magazine, IEEE*, 25(2):27–41, 2006.

[37] Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich. Network modelling methods for fmri. *NeuroImage*, 54(2):875–891, 2011.

[38] Catie Chang and Gary H Glover. Time–frequency dynamics of resting-state brain connectivity measured with fmri. *NeuroImage*, 50(1):81–98, 2010.

[39] Nora Leonardi, Jonas Richiardi, Markus Gschwind, Samanta Simioni, Jean-Marie Annoni, Myriam Schluep, Patrik Vuilleumier, and Dimitri Van De Ville. Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *NeuroImage*, 83:937–950, 2013.

[40] Vince D Calhoun, Robyn Miller, Godfrey Pearlson, and Tulay Adalı. The chronnectome: time-varying connectivity networks as the next frontier in fmri data discovery. *Neuron*, 84(2):262–274, 2014.

[41] Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A riemannian framework for tensor computing. *International Journal of computer vision*, 66(1):41–66, 2006.

[42] Oncel Tuzel, Fatih Porikli, and Peter Meer. Pedestrian detection via classification on riemannian manifolds. *IEEE transactions on pattern analysis and machine intelligence*, 30:1713–27, 11 2008.

[43] Gaël Varoquaux, Flore Baronnet, Andreas Kleinschmidt, Pierre Fillard, and Bertrand Thirion. Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 200–208. Springer, 2010.

[44] Bernard Ng, Gael Varoquaux, Jean Poline, Michael Greicius, and Bertrand Thirion. Transport on riemannian manifold for connectivity-based brain decoding. 2015.

[45] Luca Dodero, Ha Quang Minh, Marco San Biagio, Vittorio Murino, and Diego Sona.

Kernel-based classification for brain connectivity graphs on the riemannian manifold of positive definite matrices. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 42–45. IEEE, 2015.

[46] Harini Eavani, Theodore D Satterthwaite, Roman Filipovych, Raquel E Gur, Ruben C Gur, and Christos Davatzikos. Identifying sparse connectivity patterns in the brain using resting-state fmri. *NeuroImage*, 105:286–299, 2015.

[47] Sandro Vega-Pons, Paolo Avesani, Michael Andric, and Uri Hasson. Classification of inter-subject fmri data based on graph kernels. In *2014 International Workshop on Pattern Recognition in Neuroimaging*, pages 1–4. IEEE, 2014.

[48] Biao Jie, Daoqiang Zhang, Chong-Yaw Wee, and Dinggang Shen. Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification. *Human brain mapping*, 35(7):2876–2897, 2014.

[49] Mehrtash Harandi, Richard Hartley, Chunhua Shen, Brian Lovell, and Conrad Sanderson. Extrinsic methods for coding and dictionary learning on grassmann manifolds. *International Journal of Computer Vision*, 114(2-3):113–136, 2015.

[50] Shenghua Gao, Ivor Wai-Hung Tsang, and Liang-Tien Chia. Sparse representation with kernels. *IEEE Transactions on Image Processing*, 22(2):423–434, 2012.

[51] Weiyang Liu, Zhiding Yu, Meng Yang, Lijia Lu, and Yuexian Zou. Joint kernel dictionary and classifier learning for sparse coding via locality preserving k-svd. In *2015 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2015.

[52] Huaping Liu, Jie Qin, Hong Cheng, and Fuchun Sun. Robust kernel dictionary learning using a whole sequence convergent algorithm. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[53] Ashish Shrivastava, Vishal M Patel, and Rama Chellappa. Multiple kernel learning for sparse representation-based classification. *IEEE Transactions on Image Processing*, 23(7):3013–3024, 2014.

[54] Mehrtash Harandi and Mathieu Salzmann. Riemannian coding and dictionary learning: Kernels to the rescue. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3926–3935, 2015.

[55] Yuchen Xie, Baba C Vemuri, and Jeffrey Ho. Dictionary learning on riemannian manifolds. In *MICCAI workshop on STMI*, 2012.

[56] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Bci

signal classification using a riemannian-based kernel. In *20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2012)*, pages 97–102. Michel Verleysen, 2012.

[57] Anqi Qiu, Annie Lee, Mingzhen Tan, and Moo K Chung. Manifold learning on brain functional networks in aging. *Medical image analysis*, 20(1):52–60, 2015.

[58] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Fast and simple calculus on tensors in the log-euclidean framework. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 115–122. Springer, 2005.

[59] Sadeep Jayasumana, Richard Hartley, Mathieu Salzmann, Hongdong Li, and Mehrtash Harandi. Kernel methods on the riemannian manifold of symmetric positive definite matrices. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 73–80, 2013.

[60] Ian L Dryden, Alexey Koloydenko, Diwei Zhou, et al. Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *The Annals of Applied Statistics*, 3(3):1102–1123, 2009.

[61] Suvrit Sra. A new metric on the manifold of kernel matrices with application to matrix geometric means. In *Advances in Neural Information Processing Systems*, pages 144–152, 2012.

[62] DC Dowson and BV Landau. The fréchet distance between multivariate normal distributions. *Journal of multivariate analysis*, 12(3):450–455, 1982.

[63] Bruno Sauce and Louis D Matzel. The paradox of intelligence: Heritability and malleability coexist in hidden gene-environment interplay. *Psychological bulletin*, 144(1):26, 2018.

[64] Thomas J Bouchard et al. Genes, environment, and personality. *Science-AAAS-Weekly Paper Edition*, 264(5166):1700–1701, 1994.

[65] Paul M Thompson, Tyrone D Cannon, Katherine L Narr, Theo Van Erp, Veli-Pekka Poutanen, Matti Huttunen, Jouko Lönnqvist, Carl-Gustaf Standertskjöld-Nordenstam, Jaakko Kaprio, Mohammad Khaledy, et al. Genetic influences on brain structure. *Nature neuroscience*, 4(12):1253–1258, 2001.

[66] Sofie L Valk, Felix Hoffstaedter, Julia A Camilleri, Peter Kochunov, BT Thomas Yeo, and Simon B Eickhoff. Personality and local brain structure: their shared genetic basis and reproducibility. *bioRxiv*, page 645945, 2020.

[67] Patricia Cowell and Jennifer Gurd. Handedness and the corpus callosum: A review and further analyses of discordant twins. *Neuroscience*, 388:57–68, 2018.

[68] Robert Plomin, Michael J Owen, and Peter McGuffin. The genetic basis of complex human behaviors. *Science*, 264(5166):1733–1739, 1994.

[69] Eric Turkheimer. Three laws of behavior genetics and what they mean. *Current directions in psychological science*, 9(5):160–164, 2000.

[70] Daniëlle Posthuma and Dorret I Boomsma. A note on the statistical power in extended twin designs. *Behavior genetics*, 30(2):147–158, 2000.

[71] Jiska S Peper, Rachel M Brouwer, Dorret I Boomsma, René S Kahn, and Hilleke E Hulshoff Pol. Genetic influences on human brain structure: a review of brain imaging studies in twins. *Human brain mapping*, 28(6):464–473, 2007.

[72] Dorret Boomsma, Andreas Busjahn, and Leena Peltonen. Classical twin studies and beyond. *Nature reviews genetics*, 3(11):872–882, 2002.

[73] Hae-Jeong Park and Karl Friston. Structural and functional brain networks: from connections to cognition. *Science*, 342(6158), 2013.

[74] BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, Mert R Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L Roffman, Jordan W Smoller, Lilla Zöllei, Jonathan R Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 2011.

[75] Benjamin Sinclair, Narelle K Hansell, Gabriëlla AM Blokland, Nicholas G Martin, Paul M Thompson, Michael Breakspear, Greig I de Zubicaray, Margaret J Wright, and Katie L McMahon. Heritability of the network architecture of intrinsic brain functional connectivity. *NeuroImage*, 121:243–252, 2015.

[76] Zhi Yang, Xi-Nian Zuo, Katie L McMahon, R Cameron Craddock, Clare Kelly, Greig I de Zubicaray, Ian Hickie, Peter A Bandettini, F Xavier Castellanos, Michael P Milham, et al. Genetic and environmental contributions to functional connectivity architecture of the human brain. *Cerebral cortex*, 26(5):2341–2352, 2016.

[77] Karl J Friston. Models of brain function in neuroimaging. *Annual Review of Psychology*, 56:57–87, 2005.

[78] Ludwig von Bertalanffy. General system theory: Foundations, development, applications. 1969.

[79] Bart Vandereycken, P-A Absil, and Stefan Vandewalle. A riemannian geometry with

complete geodesics for the set of positive semidefinite matrices of fixed rank. *IMA Journal of Numerical Analysis*, 33(2):481–514, 2013.

[80] Cedric E Ginestet, Jun Li, Prakash Balachandran, Steven Rosenberg, Eric D Kolaczyk, et al. Hypothesis testing for network data in functional neuroimaging. *The Annals of Applied Statistics*, 11(2):725–750, 2017.

[81] Mengyu Dai, Zhengwu Zhang, and Anuj Srivastava. Testing stationarity of brain functional connectivity using change-point detection in fmri data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 19–27, 2016.

[82] Mehdi Rahim, Bertrand Thirion, and Gaël Varoquaux. Population shrinkage of covariance (posce) for better individual brain functional-connectivity estimation. *Medical image analysis*, 54:138–148, 2019.

[83] Fani Deligianni, Gael Varoquaux, Bertrand Thirion, Emma Robinson, David J Sharp, A David Edwards, and Daniel Rueckert. A probabilistic framework to infer brain functional connectivity from anatomical connections. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 296–307. Springer, 2011.

[84] Alireza Davoudi, Saeed Shiry Ghidary, and Khadijeh Sadatnejad. Dimensionality reduction based on distance preservation to local mean for symmetric positive definite matrices and its application in brain–computer interfaces. *Journal of neural engineering*, 14(3):036019, 2017.

[85] Bernard Ng, Martin Dressler, Gaël Varoquaux, Jean Baptiste Poline, Michael Greicius, and Bertrand Thirion. Transport on riemannian manifold for functional connectivity-based classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 405–412. Springer, 2014.

[86] Douglas Scott Falconer. *Introduction to quantitative genetics*. Pearson Education India, 1996.

[87] David C Van Essen, Kamil Ugurbil, Edward Auerbach, Deanna Barch, TEJ Behrens, Richard Bucholz, Acer Chang, Liyong Chen, Maurizio Corbetta, Sandra W Curtiss, et al. The human connectome project: a data acquisition perspective. *NeuroImage*, 62(4):2222–2231, 2012.

[88] Vaibhav A Diwadkar, Marcella Bellani, Asadur Chowdury, Silvia Savazzi, Cinzia Perlini, Veronica Marinelli, Giada Zoccatelli, Franco Alessandrini, Elisa Ciceri, Gianluca Rambaldelli, et al. Activations in gray and white matter are modulated by uni-manual

responses during within and inter-hemispheric transfer: effects of response hand and right-handedness. *Brain imaging and behavior*, 12(4):942–961, 2018.

[89] Robert W Cox. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3):162–173, 1996.

[90] Mark Jenkinson, Christian F Beckmann, Timothy EJ Behrens, Mark W Woolrich, and Stephen M Smith. Fsl. *NeuroImage*, 62(2):782–790, 2012.

[91] Bruce Fischl. Freesurfer. *NeuroImage*, 62(2):774–781, 2012.

[92] Florent Ségonne, Anders M Dale, Evelina Busa, Maureen Glessner, David Salat, Horst K Hahn, and Bruce Fischl. A hybrid approach to the skull stripping problem in mri. *NeuroImage*, 22(3):1060–1075, 2004.

[93] Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy EJ Behrens, Heidi Johansen-Berg, Peter R Bannister, Marilena De Luca, Ivana Drobnjak, David E Flitney, et al. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23:S208–S219, 2004.

[94] Suzanne T Witt, Angela R Laird, and M Elizabeth Meyerand. Functional neuroimaging correlates of finger-tapping task variations: an ale meta-analysis. *NeuroImage*, 42(1):343–356, 2008.

[95] Mark W Woolrich, Saad Jbabdi, Brian Patenaude, Michael Chappell, Salima Makni, Timothy Behrens, Christian Beckmann, Mark Jenkinson, and Stephen M Smith. Bayesian analysis of neuroimaging data in fsl. *NeuroImage*, 45(1):S173–S186, 2009.

[96] J.D. Rudie, J.A. Brown, D. Beck-Pancer, L.M. Hernandez, E.L. Dennis, P.M. Thompson, S.Y. Bookheimer, and M. Dapretto. Altered functional and structural brain network organization in autism. *NeuroImage: Clinical*, 2:79 – 94, 2013.

[97] Peter Kochunov, Brian Donohue, Braxton D Mitchell, Habib Ganjgahi, Bhim Adhikari, Meghann Ryan, Sarah E Medland, Neda Jahanshad, Paul M Thompson, John Blangero, et al. Genomic kinship construction to enhance genetic analyses in the human connectome project data. *Human brain mapping*, 40(5):1677–1688, 2019.

[98] Deanna M Barch, Gregory C Burgess, Michael P Harms, Steven E Petersen, Bradley L Schlaggar, Maurizio Corbetta, Matthew F Glasser, Sandra Curtiss, Sachin Dixit, Cindy Feldt, et al. Function in the human connectome: task-fmri and individual differences in behavior. *NeuroImage*, 80:169–189, 2013.

[99] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[100] R Cameron Craddock, Paul E Holtzheimer, Xiaoping P Hu, and Helen S Mayberg. Disease state prediction from resting state functional connectivity. *Magnetic resonance in Medicine*, 62(6):1619–1628, 2009.

[101] Mariana Zurita, Cristian Montalba, Tomás Labbé, Juan Pablo Cruz, Josué Dalboni da Rocha, Cristián Tejos, Ethel Ciampi, Claudia Cárcamo, Ranganatha Sitaram, and Sergio Uribe. Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion mri data. *NeuroImage: Clinical*, 20:724 – 730, 2018.

[102] Ling-Li Zeng, Hui Shen, Li Liu, Lubin Wang, Baojuan Li, Peng Fang, Zongtan Zhou, Yaming Li, and Dewen Hu. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain*, 135(5):1498–1507, 2012.

[103] Longlong Cao, Shuixia Guo, Zhimin Xue, Yong Hu, Haihong Liu, Tumbwene E Mwansisya, Weidan Pu, Bo Yang, Chang Liu, Jianfeng Feng, et al. Aberrant functional connectivity for diagnosis of major depressive disorder: a discriminant analysis. *Psychiatry and clinical neurosciences*, 68(2):110–119, 2014.

[104] Bernard Ng, Gaël Varoquaux, Jean Baptiste Poline, and Bertrand Thirion. A novel sparse group gaussian graphical model for functional connectivity estimation. In *International Conference on Information Processing in Medical Imaging*, pages 256–267. Springer, 2013.

[105] Maria J Rosa, Liana Portugal, Tim Hahn, Andreas J Fallgatter, Marta I Garrido, John Shawe-Taylor, and Janaina Mourao-Miranda. Sparse network-based models for patient classification using fmri. *NeuroImage*, 105:493–506, 2015.

[106] Harini Eavani, Theodore D Satterthwaite, Raquel E Gur, Ruben C Gur, and Christos Davatzikos. Discriminative sparse connectivity patterns for classification of fmri data. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 193–200. Springer, 2014.

[107] Hui Shen, Lubin Wang, Yadong Liu, and Dewen Hu. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fmri. *NeuroImage*, 49(4):3110–3121, 2010.

[108] Jonas Richiardi, Hamdi Eryilmaz, Sophie Schwartz, Patrik Vuilleumier, and Dimitri Van De Ville. Decoding brain states from fmri connectivity graphs. *NeuroImage*, 56(2):616–626, 2011.

[109] Yuan Zhou, Peter Zeidman, Shihao Wu, Adeel Razi, Cheng Chen, Liuqing Yang, Jilin

Zou, Gaohua Wang, Huiling Wang, and Karl J Friston. Altered intrinsic and extrinsic connectivity in schizophrenia. *NeuroImage: Clinical*, 17:704–716, 2018.

[110] Haesung Lee, Hyun-Jung Ahn, Kwang-Rae Kim, Peter T Kim, and Ja-Yong Koo. Geodesic clustering for covariance matrices. *Communications for Statistical Applications and Methods*, 22(4):321–331, 2015.

[111] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108, 1979.

[112] Eswar Damaraju, Elena A Allen, Aysenil Belger, Judith M Ford, S McEwen, DH Mathalon, BA Mueller, GD Pearlson, SG Potkin, A Preda, et al. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage: Clinical*, 5:298–308, 2014.

[113] Victor M Vergara, Andrew R Mayer, Kent A Kiehl, and Vince D Calhoun. Dynamic functional network connectivity discriminates mild traumatic brain injury through machine learning. *NeuroImage: Clinical*, 19:30–37, 2018.

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

[115] Anees Abrol, Eswar Damaraju, Robyn L Miller, Julia M Stephen, Eric D Claus, Andrew R Mayer, and Vince D Calhoun. Replicability of time-varying connectivity patterns in large resting state fmri samples. *Neuroimage*, 163:160–176, 2017.

[116] Jaehwan Kim, Kwang-Hyun Shim, and Seungjin Choi. Soft geodesic kernel k-means. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, volume 2, pages II–429. IEEE, 2007.

[117] Alvina Goh and René Vidal. Clustering and dimensionality reduction on riemannian manifolds. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7. IEEE, 2008.

[118] Nima Asgharbeygi and Arian Maleki. Geodesic k-means clustering. In *2008 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008.

[119] Xiao-Wei Ai, Tianming Hu, Xi Li, and Hui Xiong. Clustering high-frequency stock data for trading volatility analysis. In *2010 Ninth International Conference on Machine Learning and Applications*, pages 333–338. IEEE, 2010.

[120] David L Davies and Donald W Bouldin. A cluster separation measure. *IEEE transactions on pattern analysis and machine intelligence*, (2):224–227, 1979.

[121] Victor M Vergara, Anees Abrol, Flor A Espinoza, and Vince D Calhoun. Selection of efficient clustering index to estimate the number of dynamic brain states from functional network connectivity. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 632–635. IEEE, 2019.

[122] Nikhil R Pal and J Biswas. Cluster validation using graph theoretic concepts. *Pattern Recognition*, 30(6):847–857, 1997.

[123] Vladimir Vapnik. *The nature of statistical learning theory*. Springer science & business media, 2013.

[124] Fabrice Colas and Pavel Brazdil. Comparison of svm and some older classification algorithms in text classification tasks. In *IFIP International Conference on Artificial Intelligence in Theory and Practice*, pages 169–178. Springer, 2006.

[125] Jair Cervantes, Farid Garcia-Lamont, Lisbeth Rodriguez-Mazahua, and Asdrubal Lopez. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, 408:189–215, 2020.

[126] Michael E Mavroforakis and Sergios Theodoridis. A geometric approach to support vector machine (svm) classification. *IEEE transactions on neural networks*, 17(3):671–682, 2006.

[127] Igor Griva, Stephen G Nash, and Ariela Sofer. *Linear and nonlinear optimization*, volume 108. Siam, 2009.

[128] Claudio Stamile, Gabriel Kocevar, Salem Hannoun, Françoise Durand-Dubief, and Dominique Sappey-Marinier. A graph based classification method for multiple sclerosis clinical forms using support vector machine. In *Medical learning meets medical imaging*, pages 57–64. Springer, 2015.

[129] Jidan Zhong, David Qixiang Chen, Julia C Nantes, Scott A Holmes, Mojgan Hodaie, and Lisa Koski. Combined structural and functional patterns discriminating upper limb motor disability in multiple sclerosis using multivariate approaches. *Brain Imaging and Behavior*, 11(3):754–768, 2017.

[130] Muthuraman Muthuraman, Vinzenz Fleischer, Pierre Kolber, Felix Luessi, Frauke Zipp, and Sergiu Groppa. Structural brain network characteristics can differentiate cis from early rrms. *Frontiers in neuroscience*, 10:14, 2016.

[131] Kerstin Bendfeldt, Stefan Klöppel, Thomas E Nichols, Renata Smieskova, Pascal Kuster, Stefan Traud, Nicole Mueller-Lenke, Yvonne Naegelin, Ludwig Kappos, Ernst-Wilhelm Radue, et al. Multivariate pattern classification of gray matter pathology in multiple sclerosis. *NeuroImage*, 60(1):400–408, 2012.

[132] Gabriel Kocevar, Claudio Stamile, Salem Hannoun, François Cotton, Sandra Vukusic, Françoise Durand-Dubief, and Dominique Sappey-Marinier. Graph theory-based brain connectivity for automatic classification of multiple sclerosis clinical courses. *Frontiers in neuroscience*, 10:478, 2016.

[133] Jesse A Brown, Jeffrey D Rudie, Anita Bandrowski, John D Van Horn, and Susan Y Bookheimer. The ucla multimodal connectivity database: a web-based platform for brain connectivity matrix sharing and analysis. *Frontiers in neuroinformatics*, 6:28, 2012.

[134] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.

[135] Andrew Zalesky, Alex Fornito, and Edward T Bullmore. Network-based statistic: identifying differences in brain networks. *NeuroImage*, 53(4):1197–1207, 2010.

[136] Massimo Filippi, Bar-Or Amit, Piehl Fredrik, Preziosa Paolo, Solari Alessandra, Vukusic Sandra, and Rocca Maria A. Multiple sclerosis. *Nature Reviews Disease Primers*, 2018.

[137] Fred D Lublin, Stephen C Reingold, et al. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology*, 46(4):907–911, 1996.

[138] Maria A Rocca, Paola Valsasina, Alessandro Meani, Andrea Falini, Giancarlo Comi, and Massimo Filippi. Impaired functional integration in multiple sclerosis: a graph theory study. *Brain Structure and Function*, 221(1):115–131, 2016.

[139] KJ Friston, CD Frith, PF Liddle, and RSJ Frackowiak. Functional connectivity: the principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism*, 13(1):5–14, 1993.

[140] John F Kurtzke. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (edss). *Neurology*, 33(11):1444–1444, 1983.

[141] Oscar Esteban, Christopher J Markiewicz, Ross W Blair, Craig A Moodie, A Ilkay Isik, Asier Erramuzpe, James D Kent, Mathias Goncalves, Elizabeth DuPre,

Madeleine Snyder, et al. fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, 16(1):111–116, 2019.

[142] Raimon HR Pruim, Maarten Mennes, Daan van Rooij, Alberto Llera, Jan K Buitelaar, and Christian F Beckmann. Ica-aroma: A robust ica-based strategy for removing motion artifacts from fmri data. *NeuroImage*, 112:267–277, 2015.

[143] Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57, 2001.

[144] Jian Hou, Huijun Gao, and Xuelong Li. Dsets-dbscan: A parameter-free clustering algorithm. *IEEE Transactions on Image Processing*, 25(7):3182–3193, 2016.

[145] Sebastiano Vascon, Marco Cristani, Marcello Pelillo, and Vittorio Murino. Using dominant sets for k-nn prototype selection. pages 131–140, 2013.

[146] Massimiliano Pavan and Marcello Pelillo. Dominant sets and pairwise clustering. *IEEE transactions on pattern analysis and machine intelligence*, 29(1):167–172, 2006.

[147] Sebastiano Vascon, Samuel Rota Bulò, Vittorio Murino, and Marcello Pelillo. Dslib: An open source library for the dominant set clustering method. *arXiv preprint arXiv:2010.07906*, 2020.

[148] Jörgen W Weibull. Evolutionary game theory. 1997.

[149] Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.

[150] Daniel A Handwerker, Vinai Roopchansingh, Javier Gonzalez-Castillo, and Peter A Bandettini. Periodic changes in fmri connectivity. *NeuroImage*, 63(3):1712–1719, 2012.

[151] R Matthew Hutchison, Thilo Womelsdorf, Joseph S Gati, Stefan Everling, and Ravi S Menon. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human brain mapping*, 34(9):2154–2177, 2013.

[152] Waqas Majeed, Matthew Magnuson, Wendy Hasenkamp, Hillary Schwarb, Eric H Schumacher, Lawrence Barsalou, and Shella D Keilholz. Spatiotemporal dynamics of low frequency bold fluctuations in rats and humans. *NeuroImage*, 54(2):1140–1150, 2011.

[153] Maria Giulia Preti, Thomas AW Bolton, and Dimitri Van De Ville. The dynamic

functional connectome: State-of-the-art and perspectives. *NeuroImage*, 160:41–54, 2017.

[154] Joana Cabral, Diego Vidaurre, Paulo Marques, Ricardo Magalhães, Pedro Silva Moreira, José Miguel Soares, Gustavo Deco, Nuno Sousa, and Morten L Kringelbach. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. *Scientific reports*, 7(1):1–13, 2017.

[155] David T Jones, Prashanthi Vemuri, Matthew C Murphy, Jeffrey L Gunter, Matthew L Senjem, Mary M Machulda, Scott A Przybelski, Brian E Gregg, Kejal Kantarci, David S Knopman, Bradley F. Boeve, Ronald C. Petersen, and Clifford R. Jack Jr. Non-stationarity in the "resting brain's" modular architecture. *PloS one*, 7(6):e39731, 2012.

[156] Junchao Li, Delong Zhang, Aiying Liang, Bishan Liang, Zengjian Wang, Yuxuan Cai, Mengxia Gao, Zhenni Gao, Song Chang, Bingqing Jiao, et al. High transition frequencies of dynamic functional connectivity states in the creative brain. *Scientific reports*, 7:46072, 2017.

[157] Barnaly Rashid, Mohammad R Arbabshirani, Eswar Damaraju, Mustafa S Cetin, Robyn Miller, Godfrey D Pearlson, and Vince D Calhoun. Classification of schizophrenia and bipolar patients using static and dynamic resting-state fmri brain connectivity. *NeuroImage*, 134:645–657, 2016.

[158] Nora Leonardi, William R Shirer, Michael D Greicius, and Dimitri Van De Ville. Disentangling dynamic networks: Separated and joint expressions of functional connectivity patterns in time. *Human brain mapping*, 35(12):5984–5995, 2014.

[159] Fikret Işik Karahanoğlu and Dimitri Van De Ville. Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks. *Nature communications*, 6(1):1–10, 2015.

[160] Viviana Betti, Stefania Della Penna, Francesco de Pasquale, Dante Mantini, Laura Marzetti, Gian Luca Romani, and Maurizio Corbetta. Natural scenes viewing alters the dynamics of functional connectivity in the human brain. *Neuron*, 79(4):782–797, 2013.

[161] Ravi Tejwani, Adam Liska, Hongyuan You, Jenna Reinen, and Payel Das. Autism classification using brain functional connectivity dynamics and machine learning. *arXiv preprint arXiv:1712.08041*, 2017.

[162] On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage*, 104:430 – 436, 2015.

[163] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.

[164] Danielle S. Bassett, Nicholas F. Wymbs, Mason A. Porter, Peter J. Mucha, Jean M. Carlson, and Scott T. Grafton. Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, 108(18):7641–7646, 2011.

[165] Andrew Y Ng, Michael I Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856, 2002.

[166] Xiaofei He, Deng Cai, and Partha Niyogi. Laplacian score for feature selection. In *Advances in neural information processing systems*, pages 507–514, 2006.

[167] Kaiming Li, Lei Guo, Jingxin Nie, Gang Li, and Tianming Liu. Review of methods for functional brain connectivity detection using fmri. *Computerized medical imaging and graphics*, 33(2):131–139, 2009.

[168] Guillaume Marrelec, Alexandre Krainik, Hugues Duffau, Mélanie Pélégrini-Issac, Stéphane Lehéricy, Julien Doyon, and Habib Benali. Partial correlation for functional brain interactivity investigation in functional mri. *NeuroImage*, 32(1):228–237, 2006.

[169] Gaël Varoquaux, Alexandre Gramfort, Jean-Baptiste Poline, and Bertrand Thirion. Brain covariance selection: better individual functional connectivity models using population prior. In *Advances in neural information processing systems*, pages 2334–2342, 2010.

[170] Barnaly Rashid, Eswar Damaraju, Godfrey D Pearlson, and Vince D Calhoun. Dynamic connectivity states estimated from resting fmri identify differences among schizophrenia, bipolar disorder, and healthy control subjects. *Frontiers in human neuroscience*, 8:897, 2014.

[171] Xulei Yang, Qing Song, and Aize Cao. Weighted support vector machine for data classification. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, volume 2, pages 859–864. IEEE, 2005.