Network mining based analysis of whole brain functional connectivity

André Amaral Chambel Leitão

Thesis to obtain the Master of Science Degree in

Electrical and Computer Engineering

Supervisors: Prof. Luís Miguel Teixeira D’Ávila Pinto da Silveira
Prof. Alexandre Paulo Lourenço Francisco

Examination Committee

Chairperson: Prof. Nuno Cavaco Gomes Horta
Supervisor: Prof. Luís Miguel Teixeira D’Ávila Pinto da Silveira
Member of the committee: Prof. Vasco Miguel Gomes Nunes Manquinho

October 2014
Acknowledgments

My first acknowledgement goes to my supervisors, Prof. Luís Miguel Silveira and Prof. Alexandre Lourenço Francisco, with whom I learned a lot during this past year, for their availability whenever I needed their help, for their support and also for their valuable advices. I would also like to thank to all the research team from the project HIFI-MRI: Whole Brain Functional Connectivity Analysis of Ultra-High Field fMRI (ref. PTDC/EEI-ELC/3246/2012) who have thought me a lot regarding the all fMRI basics.

I would also like to thank all my colleagues, in particular to Francisco Cary, David Nogueira and Zé Líbano Monteiro who I have been more close with, for their help, support and friendship throughout the past five years.

I am also very thankful to my parents for the effort that they have put on my education, for their patience with me and also for never stop encouraging me to be a better student, in particular when I was younger and had little concern about my future.

I also want to thank to all my family but in particular to my brothers and sisters, for always giving me a great time whenever I am with them at home.

Also a special thanks to all my friends with which I spend most of my free time for helping me throughout the past years to keep my mind off work and for all the amazing vacations we have spent together.

At last, but not least, I want to thank Benedita for all the support and encouragement she gave me, in particular when everything seemed more difficult she always managed to cheer me up.
Abstract

Doing a detailed mapping of the human brain has been a topic of interest in the scientific world throughout the last decades. However, since the brain is considered to be the most complex organ of the human body this has been proved to be a difficult task. An approach that has been widely used is based on using functional magnetic resonance images (fMRI) with high resolution which allows a deeper knowledge on the brain organization at a high granular level.

In this work we applied a method that is already widely used to map and understand the brain’s organization. This method consists in reconstructing brain functional connectivity networks from the fMRI data which are then analysed with standard graph theory algorithms. However, in this work the fMRI datasets are all acquired at high resolution, which increases the complexity of the problem both in spatial and temporal terms.

In order to address this problem we developed a tool that is able to reconstruct the brain connectivity network from the high resolution images and analyse it in terms of the network’s information flowing efficiency and also of the network’s organization in functional modules.

We were able to see that, although the networks are very complex, there is an apparent organization. Their structure allows the information to flow from one point to another in a very efficient way. We were also able to see that these networks have a modular structure.

Keywords

functional Magnetic Resonance Imaging, brain, network mining, high resolution, graph theory, functional connectivity
Resumo

Fazer um mapeamento detalhado do cérebro humano tem sido um tópico de interesse na comunidade científica ao longo das últimas décadas. No entanto, sendo o cérebro considerado o órgão mais complexo do corpo humano é esperável que não seja fácil. Uma abordagem bastante utilizada tem-se baseado no recurso à utilização de imagens de ressonâncias magnéticas funcionais (fMRI) de alta resolução que possibilitam um conhecimento mais profundo da organização do cérebro a um nível de maior granularidade.

Neste trabalho é aplicada uma técnica já utilizada anteriormente para mapear e perceber melhor como o cérebro está organizado. Esta técnica baseia-se em partir das fMRIs reconstruir redes de conectividade funcional do cérebro que são posteriormente analisadas recorrendo a conceitos vindos da teoria de grafos. No entanto, um aspecto distintivo deste trabalho é o facto de serem utilizadas fMRIs de alta resolução, o que implica que complexidade do problema cresce muito rapidamente, quer em termos de complexidade espacial como de complexidade temporal.

Para lidar com o problema foi desenvolvida uma ferramenta que é capaz de reconstruir a partir das imagens de alta resolução uma rede de conectividade cerebral e analisá-la tanto ao nível da eficácia de circulação de informação como ao nível da organização em módulos funcionais da rede.

Verificou-se que, embora a rede fosse muito complexa, existe uma aparente organização nestas redes. A sua estrutura permite que a informação circule de um extremo ao outro de forma muito eficaz e para além disso, verificou-se também a existência de uma organização da rede em módulos funcionais.

Palavras Chave

ressonância magnética funcional, cérebro, análise de redes, alta resolução, teoria de grafos, conectividade funcional
Contents

1 Introduction .................................................. 1
   1.1 Overview and Motivation ............................................... 2
   1.2 Thesis structure ..................................................... 3

2 Background .................................................. 5
   2.1 Functional Magnetic Resonance Imaging (fMRI) .................... 6
      2.1.1 Magnetic Resonance Imaging (MRI) fundamentals ............ 6
      2.1.2 fMRI and BOLD signal .......................................... 6
      2.1.3 Functional connectivity and BFC network .................. 8
   2.2 Graph theory ..................................................... 9
      2.2.1 Metrics ......................................................... 11
      2.2.2 Network topologies ............................................ 12

3 Related work ............................................... 15
   3.1 Data pre-processing .................................................. 16
   3.2 Building a network from the fMRI ................................. 17
   3.3 BFC network mining methods ...................................... 19
      3.3.1 Degree distribution ............................................ 19
      3.3.2 Modularity ..................................................... 20
      3.3.3 Clustering ..................................................... 22
      3.3.4 Small-world connectivity .................................... 23

4 Methodology ................................................. 25
   4.1 Building the correlation matrix ..................................... 26
   4.2 Building the graph .................................................. 28
   4.3 Analysing the BFC network ......................................... 29
      4.3.1 Basic Metrics .................................................. 29
      4.3.2 Community detection .......................................... 30
   4.4 Analysis of the resulting data ..................................... 31
# Contents

5 Results

5.1 Characterization of datasets ............................................. 34
5.2 BFC networks ............................................................... 34
  5.2.1 Number of Edges ...................................................... 35
  5.2.2 Connected components ............................................. 36
  5.2.3 Vertex degree distribution ...................................... 37
  5.2.4 Small worldness ..................................................... 38
    5.2.4.A Minimum average path .................................... 38
    5.2.4.B Clustering coefficient .................................... 39
  5.2.5 Community detection ............................................. 40
5.3 Validation ........................................................................ 44
5.4 Discussion

5.4.1 BFC networks basic properties .................................... 47
5.4.2 Networks' topology .................................................... 48
5.4.3 Community detection ................................................. 48

6 Conclusions ........................................................................ 51
6.1 Future work ...................................................................... 52

A Appendix

A.1 Abstract accepted in the Fourth Biennial Conference on Resting State / Brain Connectivity .................................................. 60
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>fMRI data explained: 1 - Brain slices that are made of a grid of voxels; 2 - Slices are acquired as shown, in the direction perpendicular to the image; 3 - Combination of all the slices that make the 3D brain. It is important to stress the fact that all the data on the Figure is acquired several times throughout different time points.</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>Example of a generic graph with 8 vertices and 14 edges</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>Graph of a ROI based brain functional network. The larger edges represent a higher correlation between the two vertices. Source: Institut National de Recherche en Informatique et en Automatique (2013)</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Correlation matrix before and after applying a certain threshold. Source: Rubinov and Sporns (2010)</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>Adjacency matrix. Source: Rubinov and Sporns (2010)</td>
<td>19</td>
</tr>
<tr>
<td>3.3</td>
<td>Brain network with the different modules. Source: Meunier et al. (2009)</td>
<td>21</td>
</tr>
<tr>
<td>3.4</td>
<td>Variation of the $\gamma$ metric with the correlation threshold $T$ and the average node degree $k$. Source: van den Heuvel et al. (2008)</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Areas of the brain covered by the datasets that will be used in this work. Source: Harvard Medical School (2013)</td>
<td>26</td>
</tr>
<tr>
<td>5.1</td>
<td>Number of created edges by each chosen correlation threshold</td>
<td>36</td>
</tr>
<tr>
<td>5.2</td>
<td>Degree distribution function for the BFC network of the subject 6 and with a correlation threshold of 0.4</td>
<td>38</td>
</tr>
<tr>
<td>5.3</td>
<td>Community detection for subject 6 at a correlation threshold of 0.4</td>
<td>41</td>
</tr>
<tr>
<td>5.4</td>
<td>Community detection for subject 6 at a correlation threshold of 0.45</td>
<td>41</td>
</tr>
<tr>
<td>5.5</td>
<td>Community detection for subject 6 at a correlation threshold of 0.5</td>
<td>42</td>
</tr>
<tr>
<td>5.6</td>
<td>Community detection for subject 1 at a correlation threshold of 0.45</td>
<td>42</td>
</tr>
<tr>
<td>5.7</td>
<td>Community detection for subject 2 at a correlation threshold of 0.45</td>
<td>43</td>
</tr>
<tr>
<td>5.8</td>
<td>Community detection for subject 3 at a correlation threshold of 0.45</td>
<td>43</td>
</tr>
<tr>
<td>5.9</td>
<td>Community detection for subject 4 at a correlation threshold of 0.45</td>
<td>43</td>
</tr>
<tr>
<td>5.10</td>
<td>Community detection for subject 5 at a correlation threshold of 0.45</td>
<td>44</td>
</tr>
</tbody>
</table>
List of Figures

5.11 Three modules found with LLP for subject 3 at a correlation threshold of 0.45 . . . . 46
5.12 Three IC found with ICA for subject 3 at a correlation threshold of 0.45 . . . . . . 46
List of Tables

5.1 Total number of vertices in the BFC network of each subject . . . . . . . . . . . . . 34
5.2 Number of created edges for each subject using different correlation thresholds . . 35
5.3 Number of created edges for each subject using different correlation thresholds . . 35
5.4 Percentage of the total vertices that are in the giant component of the network . . 37
5.5 Value of the exponent from the fitting function of the degree distribution . . . . . . 37
5.6 Estimation of the minimum average path for the networks of all subjects considering
   only three different correlation thresholds. ................................................... 39
5.7 Average characteristic path for the BFC networks, their respective random equiva-
   lents and value of the \( \lambda \) coefficient. .................................................. 39
5.8 Clustering coefficient from the networks of all subjects ................................. 39
5.9 Average clustering coefficient of the BFC networks, their respective random equiv-
   alent networks and value of the \( \gamma \) coefficient. ...................................... 40
5.10 Average small-worldness coefficient of all the BFC networks .......................... 40
5.11 Overlap between the modules found with LLP and the IC found with ICA. Only the
   overlaps over 60% are referenced .............................................................. 45
5.12 Normalized mutual information between the modules and the IC presented on Table

5.11 ............................................................... 47
List of Acronyms

BFC  Brain Functional Connectivity
BOLD  Blood-Oxygen Level-Dependent
CSF  Cerebrospinal fluid
fMRI  functional Magnetic Resonance Imaging
FMRIB  Oxford Centre for functional Magnetic Resonance Imaging of the Brain
FOV  Field of View
FSL  FMRIB Software Library
ICA  Independent Component Analysis
LLP  Layered Label Propagation
MELODIC  Multivariate Exploratory Linear Optimized Decomposition into Independent Components
MRI  Magnetic Resonance Imaging
NMI  Normalized Mutual Information
TE  Echo Time
TR  Time of Repetition
1. Introduction

In this chapter we give an introduction to the work pursued in this thesis explaining the main objectives and its scientific framework. We also describe the structure of this thesis.

1.1 Overview and Motivation

Mapping the human brain has been a topic of interest for the last few decades. In spite of its incredible complexity it is now possible to map the brain using advanced techniques that often are only feasible thanks to the huge computational power that is available nowadays.

The human brain is estimated to have approximately $10^{11}$ neurons – basic brain cells that transmit information through electrochemical signals – which are connected by $10^{14}$ synapses – connections between the neurons (Swanson, 1995). With these numbers it is possible to understand how difficult it is to map the whole brain, in particular at a high resolution level.

In order to map the brain the first step that one is required to do is to use a brain imaging technique in order to view the activity of the brain, without being invasive. Then, one possible way to map the brain is to build the brain functional connectivity network from information provided by the imaging technique.

Having the whole brain mapped and the functional network built, it becomes possible to perform several types of analysis of the brain's network such as modularity, connectivity and clustering analysis. For instance, based on the outcome of these network mining techniques it is possible to compare the results taken from healthy subjects with others taken from subjects with a given pathology and compare them to extract useful information about how differences in the brain's functional networks are related with some specific brain diseases or behaviour.

It is believed that many psychiatric and neurological disorders are due to failures in the brain functional network and are also usually referred as disconnection syndromes (Catani et al., 2005). It has already been proven that brain disorders such as Alzheimer's (Stam et al., 2007) and schizophrenia (Liu et al., 2008) are directly related with the functional network topology.

This work intends to address the analysis and mapping problem, by starting with high quality functional Magnetic Resonance Imaging (fMRI), extracting from them the brain functional network based on them and applying network mining techniques to analyse them. It also intends to perform such analysis with high resolution resting state fMRIs obtained from experimental 7T machine scans provided through a collaboration with the Martinos Center at Massachusetts General Hospital. Having a high resolution image of the brain we hope will makes it possible to extract a more accurate and more detailed network. However, the increase in data size is also a problem as the amount of data can easily be hundreds of times bigger than usual fMRIs therefore one of the challenges of this work will also be to find efficient ways to build, represent and analyse this network.

The main objectives will be to process the 7T fMRI data obtained from six healthy subjects in...
resting state and reconstruct their brain functional connectivity network. Then we should identify and apply relevant network mining methods for the analysis of such networks and develop a tool that reconstructs the networks from the fMRI data.

This work has as main framework two scientific areas, algorithms and its optimization considering the data that is being handled, and also graph theory, that will explained further on. On graph theory there is a very high need for specific algorithms to analyse graphs in order to compute some particular metrics (this topic will be addressed in section 2.2). This makes these two scientific areas intrinsically related. Moreover most of the problems in graph theory are solved through algorithms which sometimes have very high computational complexity and require optimizations.

Algorithms are also related with graph theory as many of them were created specifically to solve such problems and are now very well known algorithms, such as the travelling salesman (that will be explained further on), the Dijkstra and the Bellman–Ford (for finding minimum length paths), and so on ([Leiserson et al.], 2001).

The work developed on this thesis was partially supported by national funds through FCT, Fundação para a Ciência e Tecnologia, under the project HiFiMRI - Whole Brain Functional Connectivity Analysis of UltraHigh Field MRI, PTDC/EEI-ELC/3246/2012.

1.2 Thesis structure

This thesis is organized in the following manner. In chapter 2 we provide some theoretical explanation on the background that is required in order to understand the following chapters.

In chapter 3 we present the most significant results that have been reported in related works. In chapter 4 we undertake a detailed explanation of the methods that will be used in order to achieve the results. In chapter 5 we present the results that were obtained using the methods explained on the previous chapter. Lastly, in chapter 6 we explain the main conclusions of this master thesis and provide some possible future work.
1. Introduction
Background

Contents

2.1 Functional Magnetic Resonance Imaging (fMRI) . . . . . . . . . . . . . . . . . 6
2.2 Graph theory . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
2. Background

In this chapter the basic concepts of fMRI are explained. We also provide some basic overview on graph theory concepts that are required in order to understand the next chapters and the remainder of the thesis.

2.1 Functional Magnetic Resonance Imaging (fMRI)

2.1.1 MRI fundamentals

Magnetic Resonance Imaging (MRI) is a brain imaging technique that was first used and reported by Carr (1953). This technique relies on detecting a signal in a radio frequency emitted by the hydrogen atoms which are excited using an oscillating magnetic field. The hydrogen atoms are present in every tissue of the body containing water molecules allowing the MRI to be applied to multiple areas of the body.

The contrast that will exist between different tissues in the MRI is determined by the time that is required for the excited atoms to return to an equilibrium state, which is also called the relaxation time. There are two different processes of relaxation, the T1 time relaxation, that is related with the relaxation of the magnetic field in the z direction and the T2 time relaxation, that is related with the relaxation of the transverse magnetization, i.e. the x-y plane. These two different processes are used to weight the MRIs in order to demonstrate different structures or to detect different pathologies because they will give a different color to those structures. For example, in T1 the white matter appears in a light grey color and the Cerebrospinal fluid (CSF) appears black whereas in T2 the CSF appears white and the white matter appears in a dark grey. This happens because different structures have different relaxation times (Buxton, 2009).

2.1.2 fMRI and BOLD signal

Functional magnetic resonance imaging (fMRI) is one of the most used brain imaging techniques. It relies on the magnetic properties of the hemoglobin measuring the Blood-Oxygen Level-Dependent (BOLD) signal. It differs from the usual Magnetic Resonance Imaging (MRI) which only measures the signal emitted by the hydrogen atoms that are excited by the magnetic field. The fMRI has as main advantages the fact that it is a non-invasive technique which has very good spatial and temporal resolution.

With fMRI the brain activity is measured based on the changes in the blood flow and on the fact that the blood flow in the brain is strongly correlated with neuronal activity (Huettel et al., 2004). This is explained by the hemodynamic response, which is the process by which the body adjusts its blood flow in order to deliver oxygen to meet the demand that is created by stressed cells (in this case the neurons), allowing them to work properly. This creates a difference in the magnetic susceptibility between the oxygenated hemoglobin and deoxygenated hemoglobin which will generate a signal variation that can be detected by the MRI machine.
2.1 Functional Magnetic Resonance Imaging (fMRI)

When the fMRI is applied to the brain, the BOLD signal will therefore be more intense in the areas of the brain that are active at a given time. Thus, the fMRI will provide a spatial map of the 3D brain where each volume division (voxel) will have associated to it a different BOLD signal. This allow us to know how active was that specific volume unit of the brain through the time course of the test.

Using a stronger magnetic field makes it possible to get higher quality spatial resolution. That property is consequentially reflected on the size and number of voxels, i.e., the smaller the voxel the higher the resolution and also the larger the total amount of voxels. Each voxel is a 3D part of the brain, whose dimensions depend on the thickness of each fMRI slice and the grid used on each slice. Bigger voxels contain more neurons and more blood flow, hence having a stronger signal than the small ones.

The fMRI data is basically a 4D dataset that consists of a sequence of 3D brain images, obtained over a sequence of time points. At each time point, the 3D brain image consists of a set of 2D brain slices which can be combined into a single volume, as shown in Figure 2.1.

The fMRI scan acquires a single slice at a time (some times more than one but never the whole brain), which is divided into a grid of voxels and on each slice the BOLD signal for each of the voxels is measured. Combining the information of the BOLD signal intensity for a given voxel at each timepoint gives us the BOLD signal fluctuation for that given voxel.

There are two main types of fMRI. The first one is the task-based or event-related fMRI, which is obtained while the subject is doing some task. It measures the response of the brain to that specific task, once again relying on the BOLD signal variation. This type of fMRI allows to observe the differences that exist in the neural activity associated with the task to which the subject is doing. Moreover, it also allows to observe what areas of the brain are activated by the task that is being performed.

The other type of fMRI is the resting state fMRI. For many years this type of fMRI was not a research topic because all the spontaneous low-frequency fluctuations (usually < 0.1Hz) were considered to be noise of the data. Only in 1995 [Biswal et al., 1995] the resting state fMRI signal was for the first time considered and studied. It is currently used to study the activity of the brain while the subject is resting, i.e., not performing any explicit task. This activity study gives a better understanding about the functional organization of the brain and also allows to examine how this organization is altered by neurological diseases.

This work will be focused only on resting state fMRIs obtained with a 7T magnetic field, which is considered to be a high resolution fMRI as compared with the usual clinical fMRIs machines that only operate between 1.5T and 3T. Such a high magnetic field is usually expected to yield isotropic voxels with a size varying between 1 mm and 2 mm, as opposed to the 50mm³ obtained with standard 3T machines [De Martino et al., 2011].
2. Background

Figure 2.1: fMRI data explained: 1 - Brain slices that are made of a grid of voxels; 2 - Slices are acquired as shown, in the direction perpendicular to the image; 3 - Combination of all the slices that make the 3D brain. It is important to stress the fact that all the data on the Figure is acquired several times throughout different time points.

With the fMRI images it is possible, using some image processing and mathematical techniques, that will be discussed further on, to build the Brain Functional Connectivity (BFC) network.

2.1.3 Functional connectivity and BFC network

Functional connectivity can be defined as the temporal correlation between spatially remote neurophysiological events (Friston 1994). This is now the most commonly accepted definition (Horwitz 2003). In other words the brain functional connectivity network will give us an insight on how the different brain regions are functionally related, i.e., share functional properties. This can be applied to both task-based and resting state fMRIs. However, what is computed when evaluating the functional connectivity in not always the same because the fMRI data allows to perform several different correlation analysis. The evaluated functional connectivity may differ depending on whether the complete time series is used or just part of it and also on whether one uses the data from a single subject or the data obtained by averaging across subjects. All these different approaches may yield different functional connectivity networks even though the same datasets are being used. This network will have as its basic elements each of the voxels and using the information about voxels functional connectivity to know if they are connected or not.
Functional connectivity should not be misunderstood with the concept of anatomical connectivity: the later only refers to the physical connections that actually exist on the brain and not if they share functional properties. Anatomical connectivity tends to be determined only by its ontogeny and phylogeny, i.e., the brain's development at the embryo formation and its natural evolution through generations whereas functional connectivity, depending on experience or environment of the brain, can be changeable and adaptive [Park et al., 2008]. If two brain regions are anatomically connected that does not mean that they have to be also functionally connected and vice versa [Biswal et al., 2010].

2.2 Graph theory

To perform all the network mining analysis that are required for the previously described functional network, a branch of mathematics and computer science called graph theory will be used. A generic graph \( G \) consists of a set of nodes, or vertices, \( (V) \) connected to each other through a set of edges \( (E) \), i.e., \( G = (V,E) \). These connections can either exist or not based on the pairwise relation between the vertices.

Graphically a graph is commonly represented by drawing a point for each vertex and a line connecting two vertices for each edge. An example of a graph with 8 vertices and 14 edges is presented in Figure 2.2.

![Figure 2.2: Example of a generic graph with 8 vertices and 14 edges](image)

This concept of graph was first used and introduced by Euler in its Seven Bridges of Königsberg problem. This problem consisted on doing a walk through the city of Königsberg and cross all its seven bridges exactly once. Euler proved this to be impossible and to do so he used a graph abstraction for the problem where the vertices represented the land and the edges the bridges. [Bigg et al., 1976]

Often a graph that models the brain functional network can have as vertices the brain’s regions of interest (ROI), that are usually known from a brain atlas, which is a three-dimensional map of the human brain. If a more detailed analysis is desired the vertices can be the voxels that come
2. Background

directly from the fMRI. The connections between either the ROIs or the voxels will be the edges of the graph, thus all the graphs that model the BFC network will fall under the category of simple graphs, i.e., graphs with no loops - edges that start and end in the same vertex - nor multiple edges - more than on edge with the same set of start and end vertices. A graph representing a BFC ROI-based network will be similar to the one shown in Figure 2.3.

![Graph of a ROI based brain functional network. The larger edges represent a higher correlation between the two vertices. Source: Institut National de Recherche en Informatique et en Automatique (2013).](image)

Sometimes the definition of graph does not accurately describe the presented problem, for instance in a traffic flow problem the roads may have only one possible way. Hence, it is necessary to define a different type of graphs in which each edge has an assigned orientation, the directed graphs or digraphs.

With graph theory there are several metrics that can be computed for a given graph. In order to understand all of them there are some baseline concepts that need to be defined first. One simple concept is that of degree of a vertex, that is the number of edges that are connected to it. All the vertices that are connected through an edge to a given vertex are said to be in the neighbourhood of that vertex. If the graph is directed then there are two different degrees in each vertex, the in-degree - number of edges in which the vertex is endpoint - and the out-degree - number of edges in which the vertex is the starting point.

Another important concept is that of a path, that is the sequence of vertices and edges that are crossed to get from a vertex of the graph to another. The length of a path can be measured by the number of edges that are crossed and this yields the concept of distance between two vertices, that is the shortest of all paths that connects them. There is also the concept of the "longest shortest path" which is known as the diameter of the network.
2.2 Graph theory

2.2.1 Metrics

With all these concepts in hand it is now possible to define the following different metrics:

**Degree distribution:** This may be the simplest metric that can be used in a graph; it usually comes in the form of a function $P(k)$ that gives the fraction of vertices of the network with degree $k$. With the degree distribution it is possible to know which vertices have more edges connected to it. These types of vertices are called hubs for their influence in the network is large. For a directed graph there are two different distributions, de in-degree and the out-degree distribution.

**Connectivity:** A graph is said to be connected if there is at least one path from any vertex to any other vertex. If this does not happen the graph is said to be disconnected. Moreover a graph is said to be $k$-connected if after removing $k-1$ vertices of it there is still a path from any vertex to any other; this implies that there must exist $k$ vertex disjoint paths between all the vertices. Two paths are said to be vertex disjoint if other than perhaps the initial and final vertices they do not share any other vertices.

**Clustering coefficient:** It is used to measure how much the vertices tend to cluster together. This measure can either be global, concerning the whole graph or local concerning only a single vertex. Without going into too much detail on the calculation of the coefficient, at a global level, we have

$$C = \frac{3N_\Delta}{N_3} \quad (2.1)$$

where $N_\Delta$ is the total number of triangles in the network – sets of three vertices with edges between each pair of vertices – and $N_3$ is the total number of connected triplets – sets of three vertices where each vertex can be reached from each other (directly or indirectly), i.e., two vertices must be adjacent to another vertex (the central vertex). The factor three assures that $0 \leq C \leq 1$, as each triangle can be seen as three different connected triplets. That will give an overall indication of the clustering that exists in the network. At a local level, it will be calculated by

$$C_i = \frac{N_\Delta(i)}{N_3(i)} \quad (2.2)$$

where $N_\Delta(i)$ is the number of triangles involving vertex $i$ and $N_3(i)$ is the number of connected triplets having $i$ as its central vertex. This gives an insight of how likely it is for in any given node that their neighbours are also neighbours of each other. Having the clustering coefficient for all the vertices it is also possible to do an average and get another coefficient for the whole network.
2. Background

\[ \tilde{C} = \frac{1}{N} \sum C_i \]  \hspace{1cm} (2.3)

where \( N \) is the total number of vertices in the network and \( C_i \) the clustering coefficient at node \( i \). The difference between the two definitions is that the first one gives the same weight to each triangle of the network while the second gives the same weight to each vertex. The two definitions will therefore yield different values as vertices with higher degree can be involved in a larger number of triangles than vertices with smaller degree (Costa et al., 2007).

**Modularity**: Designed to measure how much a graph can be divided into different components, or modules. In general a graph with high modularity has a high edge density within the modules but a sparse one between the vertices that are on different modules. The most used modularity measure is the one presented by Newman and Girvan (2004) which defines it for a network that has already been partitioned into communities as:

\[ Q = \sum_{u \in M} \left[ e_{uu} - \left( \sum_{v \in M} e_{uv} \right)^2 \right] \]  \hspace{1cm} (2.4)

where the network is fully divided in a set of non-overlapping modules \( M \), and \( e_{uv} \) is the proportion of all links that connect vertices in module \( u \) with vertices in module \( v \).

2.2.2 Network topologies

In addition to the previously described measures, graphs can also be classified according to their structural properties. In the context of this work the most commonly used classifications of a network will be:

**Small-world network**: A graph can be classified as having a small-world topology if the large majority of the vertices are not neighbours of each other but there is a path connecting almost all of them and usually with a small number of hops, i.e., passing through a small number of vertices. As defined by Watts and Strogatz (1998) the distance \( L \) between two vertices of such a network grows at least proportionally with the logarithm of the number of vertices \( N \), i.e.:

\[ L \propto \log N \]  \hspace{1cm} (2.5)

**Scale-free network**: This type of network is characterized by having a very disproportional degree distribution function \( P(k) \). There is a small number of vertices that have a very high
degree in comparison with all the other remaining vertices. Mathematically speaking these networks are said to have a degree distribution that asymptotically follows a power law, i.e.:

$$P(k) \sim k^{-\gamma}$$  \hspace{1cm} (2.6)

Common situations are such that $2 < \gamma < 3$ (Grigorov, 2005). Such a network will have some vertices that are considered to be hubs and all the others will have a low degree. Such organization makes the network robust to failures that may occur, as the percentage of vertices with low degree is much higher the probability of the failure to occur also in one of those vertices is also consequently larger.

With all these different analysis techniques that can be applied to a network represented as a graph it is in our interest to have the brain functional connectivity network also represented as one. A more detailed analysis of graph mining techniques can be seen in the works of Costa et al. (2007) and Fortunato (2010).
2. Background
3 Related work

Contents

3.1 Data pre-processing ................................................. 16
3.2 Building a network from the fMRI ............................... 17
3.3 BFC network mining methods ................................. 19
3. Related work

In this chapter we will present the main results and achievements of related works in this subject, in particular concerning the pre-processing of the fMRI data, the building of the BFC network and also the subsequent analysis with network mining methods.

3.1 Data pre-processing

The pre-processing of the data is basically a standard procedure that consists in a number of steps in order to remove errors in the datasets that are mainly due to limitations in the data acquisition process, namely noise in the data and head movements. The main steps of the procedure are related to the scan time correction, the non-brain removal, the high-pass temporal filtering and the motion correction (De Martino et al., 2011; Ferrarini et al., 2009).

Briefly describing each of these pre-processing steps, the first one – the slice timing correction – is necessary because the magnetic resonance machine acquires the different brain slices at different points in time, and so each of the slices from the same volume is related to a different time point. If this is taken into account, the analysis is much more complex as each of the slices will also be acquired several times. In order to rule out this problem each 3D brain structure will be considered to be taken at the same time point and so there will be a 4D dataset where each 3D structure has different values through time.

The second step is the non-brain removal. It is usually performed using the Brain Extraction Tool (BET) that will remove all the non-brain data from the fMRI. This method uses a deformable model that will evolve until fitting on the brain’s surface and can be applied with or without any other pre-processing steps (Smith, 2002).

Next, temporal filtering is also a very important step as it will eliminate all the components of the signal that are considered to be noise, i.e., only the frequencies of the signal that come from the voxels are taken into account for the fMRI. This is done using a band-pass filter which is a filter that removes all the frequencies of the signal that are outside the specified frequency range.

The last important pre-processing technique is the head motion correction. This is required because while the fMRI is acquired the subject may unwittingly move his head and when that happens the voxels are no longer in the same place that they were in the last time point in which the previous image was acquired. The adjustment can be done using the whole 3D body from the data or a single slice both taken from the first time point and do rotations and shifts to it. Then all the possible transformations are compared using a cost function like correlation or simply doing a sum of the squared differences between voxels’ signal intensity. The chosen transformation is the one that has the lowest value for the cost function.

After the datasets are pre-processed it is required for each voxel’s BOLD signal to be transformed in its relative BOLD signal variation. This is required because we are not interested in connecting voxels with similar signal intensities but instead with similar signal variations.
that in a resting state have a similar signal variation are the ones that we consider to be functionally connected, whereas the intensity of the signal is not of our interest as far as functional connectivity is concerned.

3.2 Building a network from the fMRI

Once the image is pre-processed it is now possible to start analysing the fMRI in order to extract the brain functional connectivity (BFC) network.

In the BFC network each ROI or each voxel will be a vertex and their pairwise functional connectivity will be an edge, so one needs to establish a way to determine whether two vertices of the graph are functionally connected and so if there should be an edge connecting them. The most common way to do this is by measuring the correlation between two vertices of the graph \cite{Smith_2012}. Having the correlation between all pairs of vertices a threshold is set and only pairs with a correlation above that level are accepted as functionally connected. It is also possible to connect all the vertices in the graph and use the correlation coefficient to weight each of the edges \cite{Margulies_2010}. With non-weighted connections it is only possible to know whether a connection exists or not whereas in a weighted connection is also possible to know information about the strength of the connection, that in this case means how much the pair of voxels shares functional properties. The connections can also have directionality however this is only used when we are dealing with anatomical networks, which are networks whose connections take into consideration only the physical aspects of the neuronal structure.

As mentioned before there are two main types of BFC networks, the ROI-based and the voxel-based, hence there are also two main ways of obtaining the correlation in order to build the network.

The first technique, ROI-based, used by Ferrarini et al. \cite{Ferrarini_2009} and based on Salvador et al. \cite{Salvador_2005} consists in computing the correlation matrix between any generic pair of voxels \((i, j)\). In order to compute the correlation matrix \(R\), the sample covariance matrix \(S\) is required to be calculated first. Having the data matrix \(Y\), with dimensions \(M \times T\) where \(T\) represents the number of samples taken in during the time series, the element \(s_{i,j}\) of the matrix \(S\) is given by:

\[
s_{i,j} = \frac{1}{T} \sum_{t=1}^{T} (y_{i}(t) - \bar{y}_{i})(y_{j}(t) - \bar{y}_{j})
\]

where \(\bar{y}_{i}\) denotes, for a given voxel \(i\), the average over time of the observations and \(y_{i}(t)\) denotes the element \(i\) from the matrix \(Y\) at a given time \(t\).

Once the covariance matrix \(S\) is estimated it is now necessary to calculate the inverted matrix \(S^{-1}\) and then obtain the final matrix \(R\) where each generic \(r_{i,j}\) is given by:

\[
r_{i,j} = \frac{-\hat{s}_{i,j}}{\sqrt{s_{i,i}s_{j,j}}}
\]
where each element $s_{i,j}$ denotes the element $(i,j)$ of the inverted matrix $S^{-1}$.

It is easy to understand why this technique is only applicable to the ROI-based networks, as the covariance matrix is required to be inverted. In general, this is only possible when there are more time samples than brain regions, or at least if they are of the same order of magnitude (Salvador et al., 2005). As expected in the voxel-based network the number of regions is several orders of magnitude greater than the number of time points and therefore another technique should be used.

The voxel-based BFC network is going to need a simpler correlation and the most common definition is the Pearson product-moment correlation coefficient (Lee Rodgers and Nicewander, 1988) in which the difference from the previous method lies in the fact that for the calculation of each element of the correlation matrix, the covariance matrix is used instead of its inverse yielding the following formula:

$$r_{i,j} = \frac{s_{i,j}}{\sqrt{s_{i,i} s_{j,j}}}$$

(3.3)

Once the correlation matrix is built, the most common thing to do next is to apply a threshold to it (Rubinov and Sporns, 2010). This means that a given pair of voxels, or a given pair of ROIs, is only considered to be functionally connected if the correlation between them is above the applied threshold. As one can see in Figure 3.1 after the threshold is applied it becomes much easier to extract information from the correlation matrix. In this figure the matrix corresponds to a ROI-based network; if it was a voxel-based network it would have a size that could be between hundreds of thousands up to millions and therefore the decrease in the amount of data would be even more significant when a threshold is applied.

Figure 3.1: Correlation matrix before and after applying a certain threshold. Source: Rubinov and Sporns (2010)

It is now possible to understand what has been previously said about the same fMRI being
The thresholded correlation matrix the network can be built with weighted connections or not. If we chose to weight the connections, then the weight of a connection between any two vertices (ROIs or voxels) is the correlation between those two vertices. On the other hand, if the network is not intended to have weights in the connections, then the thresholded matrix can be seen as a binary matrix where zeros mean no connection and ones mean there is one connection of some appreciable strength. In graph theory this is called the adjacency matrix and it is shown in Figure 3.2.

### 3.3 BFC network mining methods

The most commonly used metrics concerning the BFC network are related with measuring the functional segregation of the brain, i.e., the brain’s ability for the denser and interconnected groups to do some specific and specialized processing. This can be identified by the presence of clusters or modules in the network and they also can be applied to anatomical networks. Furthermore, the functional integration of the brain it is also a very important aspect, as it shows how much is the network able to combine information from distributed brain regions. This is usually measured with the average shortest path of the network, also known as the characteristic path.

All the previously referenced metrics have already been applied in several previous works and in this section we provide some insight on the most significant results regarding the analysis of the BFC network done in previous works.

#### 3.3.1 Degree distribution

The degree distribution of the BFC network is interestingly related with the type of model that is used to represent the network. As mentioned before, the two network construction models that
3. Related work

are used to perform this type of analysis are either ROI-based or voxel-based.

Voxel-based BFC networks have been reported in several works as having a degree distribution function that follows a power law distribution and therefore fall in the category of scale-free networks [van den Heuvel et al. 2008, Eguiluz et al. 2005]. On the other hand in ROI-based networks this is not true, the degree distribution function has been observed to follow a truncated power law and therefore it is not truly scale-free, implying that the probability of one node to be highly connected is higher than in an equivalent random network but less that would be expected in a scale-free network [Achard et al. 2006, Bullmore and Sporns 2009].

3.3.2 Modularity

The first analysis that reported modularity in resting-state functional connectivity was conducted by Ferrarini et al. (2009) and it provided very interesting results. It is important to stress the fact that the analysis was made in a ROI-based functional network, extracting the ROIs from the Tzourio-Mazoyer atlas (Tzourio-Mazoyer et al. 2002), and using the methodology from Salvador et al. (2005) for building the BFC network leading to a network with 90 vertices and 256 edges (average number of links per node 5.87).

The findings in the study conducted by Ferrarini et al. (2009) suggest that the modularity in the brain is mostly a small-scale phenomenon, as the three major modules that they found only represented 43% of the total number of areas in the used atlas. They also found that the highly connected vertices of the network tend to connect to other highly connected vertices, which suggests the presence of self-organized modularity in the BFC network. Modularity was found between the frontal, sub-cortical, parietal and temporal regions of the brain.

Another important modularity analysis that is usually made, as far as state of the art works are concerned, is the hierarchical modularity analysis, also called nested modularity. In this type of analysis each of the modules that is obtained at the highest level is further decomposed into sub-modules, which can also be further decomposed into sub-submodels, and so on. Such type of analysis was made by Meunier et al. (2009). In their study they used an algorithm called the "Louvian method" to perform the hierarchical study of the modularity.

The algorithm starts by assigning to each vertex a module, doing a partition of N modules, with N being the total number of vertices. Then it sequentially selects all of the vertices, such that when a vertex is selected it may leave the current module and join another one or stay in the same one, depending on whether this change increases the modularity, which is calculated with the method presented in Newman and Girvan (2004). The second step of the algorithm consists in building a new network whose vertices are the modules created in the previous step. Both steps are repeated iteratively until a maximum modularity is reached.

Meunier et al. (2009) found five main modules corresponding to different parts of the brain and each of these modules was further decomposed into sub-modules. The main modules that are
possible to see on Figure 3.3 were the central module of the brain – represented in red – with 239 vertices and 11 sub-modules, the parieto-frontal – represented in dark blue – with 138 vertices and 10 sub-modules, the medial occipital – represented in green – with 132 vertices and 1 sub-module, the lateral occipital – represented in light blue – with 101 vertices and 1 sub-module and the fronto-temporal module – represented in yellow – with 89 vertices and 24 sub-modules. These major modules, as expected, represent functionally and/or anatomically related regions of the brain, and this pattern was also evident at some of the sub-modules.

The role of the vertices in the inter-modular connectivity was also studied by Meunier et al. (2009) measuring it and classifying the vertices according to its inter-modular connectivity. The non-hub vertices were classified in ultra-peripheral vertices, peripheral vertices, connector vertices and killness vertices, as the inter-modularity increases and the hubs were classified into provincial hubs, connector hubs and killness hubs. This allows more precision about the topological role of each node in the network. The inter modular connectivity is measured by the participation coefficient that is given by

\[ P_i = 1 - \sum_{n=1}^{N} \left( \frac{k_{ni}}{k_i} \right)^2 \]  

(3.4)

where \( k_i \) is the degree of the node \( i \) and the \( k_{ni} \) is the number of edges connecting the \( i \)th node to other vertices in the same module. \( P \) will be close to 1 if the node is linked to all other modules and 0 if it is connected only to vertices in its own module.

This measure also allows us to do a resilience test to the network as the first vertices to be disconnected are the killness hubs for their role in the network’s connectivity is of high importance.
3. Related work

3.3.3 Clustering

As explained above the clustering coefficient describes the connectivity of the direct neighbours of a given node and gives information on the formation of sub-graphs within the full network. In order to compare the clustering coefficient of a BFC network relative to an equivalent random network, there is a very commonly used metric \( \gamma \) defined by:

\[
\gamma = \frac{C_{BFC}}{C_{random}}
\]  

(3.5)

where \( C_{BFC} \) is the clustering coefficient from the BFC network and \( C_{random} \) is the clustering coefficient from the random equivalent network.

As mentioned before, all the BFC networks are proven to have a small-world topology and therefore \( \gamma \) must be always greater than 1.

Several studies have been made using this metric. One of them that is strongly related with this work was done by van den Heuvel, Stam, Boersma, and Hulshoff Pol (2008) using voxel-based BFC networks from resting-state fMRIs, which will be the scope of our work. They tested how the \( \gamma \) factor varies with the different thresholds that are applied on the correlation matrix and with the average node degree of the network. As can be seen on Figure 3.4, as the correlation threshold increases the \( \gamma \) also increases, which means that the higher the threshold the higher will be the difference between the clustering coefficient in the BFC network compared with an equivalent random network. Moreover it is also possible to extract the inverse conclusion for the average node degree (\( k \)). As \( k \) decreases \( \gamma \) will increase, meaning that in BFC networks with a high average node degree the clustering coefficient is similar with its equivalent random network. On the other hand for lower average degree networks the same does not happen, the difference between the clustering coefficient on the BFC network and the equivalent random network is very high and therefore \( \gamma \) will also be higher. This can be explained with the fact that random networks with low average degree do not tend to form clusters.

![Figure 3.4: Variation of the \( \gamma \) metric with the correlation threshold \( T \) and the average node degree \( k \). Source: van den Heuvel et al. (2008)](image-url)
3.3 BFC network mining methods

3.3.4 Small-world connectivity

Regarding connectivity it has already been proven that both ROI-based and voxel-based networks have a small-world structure (Hayasaka and Laurienti, 2010) and therefore the network combines the presence of functional modules with a high number of inter modular edges. This type of connectivity is associated with the BFC networks because of the brain’s ability to flow the information from one point to another in a very fast and efficient way, which is by definition the most basic property of small-world networks.

For a network to have such structure its clustering coefficient has to be higher than the one from a random equivalent network, i.e., a network with the same number of vertices and the same average number of links per node, and also needs to have a characteristic path similar (or a little higher in sparser networks) to the one found also in the random equivalent network.

In order to compare the characteristic path of the network with its random equivalent there is a metric $\lambda$ defined by:

$$\lambda = \frac{L_{BFC}}{L_{random}}$$

(3.6)

where $L_{BFC}$ is the characteristic path from the BFC network and $L_{random}$ is the characteristic path from the random equivalent network.

To combine the two metrics already used to measure the small-worldness of a network there is a metric $\sigma$ that is given by:

$$\sigma = \frac{\gamma}{\lambda}$$

(3.7)

Relying on the fact that small-world networks should have $\gamma$ greater than 1 and $\lambda$ close to 1, hence for a network to have a small-world topology, $\sigma$ should be greater than 1 (Humphries et al., 2006).

An example of this can be found in several works such as Ferrarini et al. (2009), Sporns and Zwi (2004) and van den Heuvel et al. (2008). In these works the results regarding the structure of the network always yielded a higher cluster coefficient than in the equivalent randomly generated network, a very low average minimum path length, when compared with the total number of vertices, and similar in both random and BFC networks, i.e., $C_{rnd} < C_{BFC_n}$ and $L_{BFC_n} \propto \log N$. 

23
3. Related work
4
Methodology

Contents

4.1 Building the correlation matrix .................................................. 26
4.2 Building the graph ................................................................. 28
4.3 Analysing the BFC network ..................................................... 29
4.4 Analysis of the resulting data ................................................... 31
4. Methodology

In this section we will provide an overview of the techniques that are used in this work. The work to be described compared with all the other ones previously mentioned shows improvements in two main aspects. First of all the fMRI data will be acquired with 7T machine scans, provided by the Martinos Center at Massachusetts General Hospital, and therefore it will be high resolution fMRI, which will yield BFC networks at an unusually high level of granularity. Moreover the scans will cover all of the brain plus the cerebellum and the brainstem. This will be a differentiating factor as, to the best of our knowledge, BFC network construction and analysis of such a complete and granular brain representation has never been reported. The included areas of the brain that are available in the datasets to be used are shown in Figure 4.1.

![Figure 4.1: Areas of the brain covered by the datasets that will be used in this work. Source: Harvard Medical School (2013)](image)

4.1 Building the correlation matrix

As previously explained, in order to build the whole voxel-based brain functional network we need to compute a correlation matrix that correlates every pair of voxels. Usually, each fMRI machine yields a different number of horizontal brain slices, that in our case will be 240, and in each slice we have the BOLD signal for every brain voxel at different time points (the voxels that are outside of the brain do not have BOLD signal associated with them and consequentially are not considered). Two voxels that have a high correlation should have a similar fluctuation in their BOLD signal. Therefore, assuming that all the pre-processing of the data (such as non-brain removal, head motion correction, temporal filtering, etc.) is already done, the first processing step
4.1 Building the correlation matrix

is to transform the fMRI data in a way that instead of every voxel having its BOLD signal intensity they now would have the relative variation of the signal through time. This is done by applying the following transformation to the signal of every voxel:

\[ \delta_j = \frac{I_j - \bar{I}_j}{\bar{I}_j} \]  

(4.1)

where \( I_j \) denotes the BOLD signal intensity on voxel \( j \) and \( \bar{I}_j \) denotes, for voxel \( j \), the average signal through all time points. The new data that we want to correlate will now have for each voxel its normalized BOLD signal variation with respect to its average.

The Pearson correlation is then applied to the transformed data, since we are dealing with voxel-based BFC networks and it is the most widely used correlation technique for such networks. In order to build the BFC network all the brain voxels are taken into account and a connection is made only between those with a pairwise correlation above a certain threshold.

The amount of data and our ability to deal with it when we compute a matrix that correlates every pair of voxels is a challenging problem. For instance, if we are dealing with the previously mentioned 240 horizontal slices, each of which will have a grid of \( 179 \times 123 \) voxels (this will represent a \( 1.1 \text{mm}^3 \) voxel, corresponding to high resolution) the correlation matrix size will be \( 5284080 \times 5284080 \). Doing the math, assuming that each entry of the matrix will be a float (32 bits - 4 bytes), this matrix would occupy almost one hundred terabytes in memory and therefore it is totally impossible to compute and store the whole matrix.

Although the matrix is impossible to store that does not mean that we cannot extract a BFC network from it; however in order to do so there are several data processing steps that we are going to need to do. First of all the most obvious step to do is on each slice to only consider the voxels that actually belong to the brain, i.e., any voxels that do not have any BOLD signal or that its signal is only noise may be removed from the group of voxels that need to be considered as being part of the BFC network.

Even tough this first step may reduce the amount of data, it is not yet good enough for our needs because even if it reduces by ten times the amount of data, most computers do not have dozens of terabytes to store such data. However there is another simple, yet significant, simplification that can be done. It consists of, instead of making the computation of the whole correlation matrix at once, doing it by chunks. Each of those chunks is then processed to extract the pairs of voxels which have correlation above the chosen threshold. Those pairs are stored and all other data can be discarded. This is the most significant processing step as far as decreasing the amount of data is concerned and will be applied as follows:

Initially the data matrix is divided in chunks of equal size

\[ Y = [ Y_1 \ Y_2 \ldots \ Y_n ] \]  

(4.2)

where each matrix \( Y_i \) has dimensions \( m \times t \), with \( m \) being the number of voxels in that given matrix.
4. Methodology

and \( t \) the number of different time points at which the BOLD signal of those voxels is measured.
Then each of the chunks is correlated pairwise with all the other chunks yielding the correlation
matrix that is formed by the sub-matrices of the chunks pairwise correlations.

\[
R = \begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1n} \\
R_{21} & \ddots & \cdots & R_{2n} \\
\vdots & \ddots & \ddots & \vdots \\
R_{n1} & R_{n2} & \cdots & R_{nn}
\end{bmatrix}
\] (4.3)

All of these sub-matrices will have the same size \( m \times m \) where \( m \) is equal to the number of
voxels that are present in each of the data chunks \( Y_i \). It is important to stress the fact that as all
the chunks from the initial matrix have the same number of voxels, it obliges all these sub-matrices
to also have the same size. Additionally, the correlation matrix \( R \) is symmetric, hence only the
sub-matrices \( R_{i,j} \) with \( i \geq j \) are required to be computed. This will decrease almost in half the
amount of required computation and storage.

4.2 Building the graph

Having the previously mentioned list of pairs of voxels, for which the correlation is above
the specified threshold, the BFC network can be built. This network will of course depend on the
chosen threshold: the higher it is the lower the number of connections that are kept.

This network, as any generic graph, can be stored in two different data structures. An ad-
jacency matrix with a size of \( n \times n \), \( n \) being the number of vertices, where each \((i,j)\) entry of
the matrix is 1 if there is a connection between voxels \( i \) and \( j \) and if it does not the value is 0.
An alternative to this data structure is the adjacency list, that has a list of all the \( n \) vertices and
for each node it stores all the vertices that are connected to it. The structure to be used usually
depends on the networks’ edge density that, on a unweighted graph, can be calculated by

\[
D = \frac{2E}{V(V-1)}
\] (4.4)

where \( E \) denotes the total number of edges on the graph and \( V \) denotes the total number of
vertices or vertices. The maximum number of edges in a graph is \( \frac{V(V-1)}{2} \) and therefore if the
number of edges in the graph is equal to that then \( D = 1 \). A graph in which \( E = O(V^2) \) it is
classified as dense and a graph in which \( E = O(V) \) is classified as sparse (Preiss, 2008). For a
dense graph the most adequate data structure is an adjacency matrix and for a sparse one should
be used an adjacency list should be used.

In this work the BFC network will be a sparse graph and therefore the data structure to be used
will be the adjacency list. Besides that, as already mentioned, it is unfeasible to store a matrix
with size \( n \times n \) for this case as \( n \) has dimensions of \( 10^6 \).
4.3 Analysing the BFC network

4.3.1 Basic Metrics

The first analysis that will be performed in the BFC network concerns the choice of the correlation thresholds and its influence on the network structure. For each chosen threshold the BFC network is analysed in order to calculate the number of disjoint components, i.e., network areas that are not connected to other areas. The disjoint component can be a single node or a group of vertices. If there is only one (disjoint) component it means that any node can be reached from any other. It is also relevant to know how much of the total numbers of vertices in the network are in the biggest connected component, this will allow us to know how much information is included, or lost, in the network when going from one threshold to another. This information will allow us to know which correlation thresholds yield networks with the least possible loss of information.

Before choosing the final thresholds that will yield the networks to be analysed with all the previously mentioned algorithms, we check another structural property of the BFC networks yielded from all the used thresholds. This will double check if we are choosing the right threshold. In order to perform this we will compute the node degree distribution function of each network, which should follow a power law distribution with a specific exponent and allow the network to be classified as scale-free, as previously suggested given that it is a voxel-based BFC network [Hayasaka and Laurienti, 2010].

Combining the results obtained from the two previous analysis it is possible to chose a set of a few adequate thresholds and then build, for all the subjects, the final networks to work on. For all the subjects it is important that we chose more than only one threshold because as the networks are different from subject to subject the threshold that yields the best results may also vary, hence we will chose for each subject the three most adequate thresholds.

We are now interested in doing a more detailed analysis of the structure of these networks that we found. To do so we will investigate if the small-world structure, that has been already verified in a lot of previously published works on voxel-based BFC networks [Hayasaka and Laurienti, 2010], is also present in these networks that we are dealing with (the main difference from the usual voxel-based networks is the complexity in terms of number of vertices and edges).

For a network to be classified as having a small-world structure it must satisfy the properties described in section 3.3.4, which basically means that it needs to have a small-world coefficient ($c$) greater than one. To compute this coefficient there are two main metrics that have to be computed: the average shortest path and the clustering coefficient. After computing these metrics for all the subjects’ networks it is also required to compute for all of them its random equivalent network - a network with the same number of vertices whose edges are randomly created in order to ensure the same degree distribution function.

As these networks are very complex, the usual algorithms for computing minimum paths are
4. Methodology

not feasible, therefore some more sophisticated and optimized ones are needed. The algorithms that we used in our work are detailed in [Boldi and Vigna (2013)] and [Boldi et al. (2011b)], and allowed us to have a very good estimation of the average shortest path in a small amount of time. The cluster coefficient is computed using the definition presented in Section 2.2.1

4.3.2 Community detection

After all the high level and more simple analysis is made then a more detailed one can be performed. We will perform a modularity analysis that is intended to find separate modules on the network. The most common modularity measure is the one explained in Section 2.2.1 and it can be computed iteratively, as previously referenced, in order to find the optimal partition of the network. The best partition is the one that concentrates more edge density within its modules. Once the best partition in communities is found for the whole network it is possible to compute the same measure for each of those communities, also for the partitions found in them, and so on. Computing the modularity will also allow us to find motifs in the network, i.e., identical communities that statistically repeat themselves through the different datasets sharing the same vertices and the same interactions between them. This will enable us to find functional patterns in the network.

However the modularity measure explained in section 2.2.1 is very inefficient for networks with the size of ours and therefore a more efficient community detection algorithm is required, more specifically an algorithm that only scales linearly with the number of edges. To do so we used a parallel algorithm developed by [Boldi et al. (2011a)] called Layered Label Propagation (LLP) that is based on the usual label propagation algorithms.

The main idea behind the usual label propagation algorithms is the following: the algorithm execute in a loop, and at the beginning of each iteration every vertex of the network has a label assigned to it that represents the cluster where that vertex belongs (when the algorithm starts each vertex has a different label). At each iteration every vertex updates its label and changes it to the label that occurs more frequently in its neighbourhood, if there is a tie then a random choice is performed and if the vertex already has the most frequent label in its neighbourhood then it remains unaltered. This algorithm will stop when in a given iteration no more vertices change labels and then all the vertices with the same labels are grouped together as one cluster.

This algorithm also allows us to perform a hierarchical modularity analysis as we can run the algorithm on a given community to divide it into other sub modules. The hierarchical modularity will also be a very interesting metric as far as this work is concerned due to the fact that we are dealing with more brain regions than most of the usual works in this area. This will probably lead to finding a higher number of high level communities and also to explore the inner community modularity in these modules of the brain, an area that has not yet been studied.

In order to use some of the previously referenced algorithms we are required to compress the network in order to allow it to be stored on memory; such compression techniques were already
used by [Boldi and Vigna (2004)] to study the world wide web network as a graph and we also used them for this work.

4.4 Analysis of the resulting data

All these analysis will be made on datasets from different subjects in order to allow us to extract more accurate conclusions based on the most relevant statistical patterns that will be observed throughout the datasets. The conclusions regarding the different functional modules that will be drawn from these datasets will then be compared with different state of the art analysis of the same datasets.

The most commonly used approach to analyse resting state fMRI is the Independent Component Analysis (ICA), and will therefore be the one that we are interested in comparing with. ICA is a statistical method to separate a signal into its different components. It relies on the assumption that the components of the signal follow non-Gaussian distributions and it tries to maximize the independence between the independent components. After application of this method, the voxels are separated in ICs where all the voxels in the the same IC were found to have a similar BOLD signal fluctuation.

The results from the ICA will yield a set of different modules of the brain with which we can compare our results in order to check for their validity. If some of the modules found by our community detection algorithm have a significant overlap with the modules found by ICA it means that we may have found a functional module of the brain.

All the other structural aspects of the networks will be compared to other analysis performed on BFC networks in related works.
4. Methodology
## Contents

5.1 Characterization of datasets ........................................... 34  
5.2 BFC networks .......................................................... 34  
5.3 Validation .............................................................. 44  
5.4 Discussion .............................................................. 47
5. Results

In this chapter the most significant results that were obtained during this thesis work are presented.

5.1 Characterization of datasets

The resting-state fMRI datasets were collected from a group of six healthy volunteers on a 7T Siemens whole-body scanner equipped with a custom-built 32-channel receive coil, using a 3 slice accelerated SMS EPI sequence (Echo Time (TE)/Time of Repetition (TR)=32ms/2500ms, 210 volumes), with 1.1mm isotropic voxels covering the whole-brain in 123 sagittal slices (FOV=264x198mm^2). The average duration of the fMRI acquisition for all the subjects is 574.5s.

Standard pre-processing was performed using FMRIB Software Library (FSL), including: slice scan time correction, motion correction, non-brain removal, spatial smoothing and high-pass temporal filtering.

5.2 BFC networks

As the size of the brain varies from one person to another, each of the subjects has a different number of voxels with BOLD signal and consequently the BFC networks all have a different total number of vertices.

The total number of vertices in each BFC network is presented in Table 5.1:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Horizontal slices</th>
<th>Number of vertices</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>144</td>
<td>1 365 082</td>
</tr>
<tr>
<td>2</td>
<td>120</td>
<td>1 080 702</td>
</tr>
<tr>
<td>3</td>
<td>133</td>
<td>1 282 836</td>
</tr>
<tr>
<td>4</td>
<td>138</td>
<td>1 305 160</td>
</tr>
<tr>
<td>5</td>
<td>145</td>
<td>1 365 120</td>
</tr>
<tr>
<td>6</td>
<td>135</td>
<td>1 262 244</td>
</tr>
<tr>
<td>Average</td>
<td>136</td>
<td>1 276 857</td>
</tr>
</tbody>
</table>

Table 5.1: Total number of vertices in the BFC network of each subject

In order to build the BFC network for each subject it is required to compute a correlation matrix with dimensions $N \times N$, where $N$ is the total number of voxels. Here it is possible to understand what has already been explained in the previous chapters, that the correlation matrix for a voxel-based BFC network is almost impossible to compute all at once. In this particular case, taking the average number of vertices, our matrix would occupy almost 12TB. Thus, the previously explained optimizations were applied to our data.

The array containing all the voxels' data from the BOLD signal intensity in the different time points was divided into 64 sub-arrays of voxels with equal size, with a zero padding in the last sub-array if the total number is not a multiple of 64. Then the full correlation between all the sub-arrays...
5.2 BFC networks

is computed.

In order to minimize the size of the stored data while performing the computation, we previously define a threshold and as the correlation is computed only the information about pairs of voxels with correlation greater or equal to that threshold is stored, all the other information is discarded.

5.2.1 Number of Edges

For each subject, when we chose a different correlation threshold it will consequently yield a different BFC network and this difference can be easily observed when computing the number of created edges. The lower the threshold the more pairs of voxels are considered as functionally connected thus resulting in a higher number of edges for the lower thresholds.

In Table 5.2 it is possible to observe the number of created edges with different correlation thresholds for all the six subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>643 726 854</td>
<td>533 490 068</td>
<td>314 593 674</td>
<td>854 248 894</td>
<td>513 294 664</td>
<td>1 154 584 928</td>
</tr>
<tr>
<td>0.45</td>
<td>293 179 725</td>
<td>232 804 710</td>
<td>139 860 962</td>
<td>340 341 230</td>
<td>242 322 242</td>
<td>565 804 017</td>
</tr>
<tr>
<td>0.50</td>
<td>133 673 620</td>
<td>102 734 100</td>
<td>61 358 154</td>
<td>136 584 841</td>
<td>112 480 949</td>
<td>274 622 469</td>
</tr>
<tr>
<td>0.55</td>
<td>60 108 492</td>
<td>44 917 200</td>
<td>26 252 603</td>
<td>55 106 150</td>
<td>50 457 139</td>
<td>129 264 717</td>
</tr>
<tr>
<td>0.60</td>
<td>26 310 936</td>
<td>19 051 725</td>
<td>10 927 042</td>
<td>22 321 640</td>
<td>21 575 159</td>
<td>57 915 679</td>
</tr>
<tr>
<td>0.65</td>
<td>11 086 248</td>
<td>7 665 391</td>
<td>4 473 639</td>
<td>9 063 264</td>
<td>8 705 395</td>
<td>24 329 169</td>
</tr>
<tr>
<td>0.70</td>
<td>4 435 876</td>
<td>2 863 345</td>
<td>1 825 110</td>
<td>3 651 258</td>
<td>3 307 007</td>
<td>9 497 373</td>
</tr>
<tr>
<td>0.75</td>
<td>1 659 048</td>
<td>978 146</td>
<td>739 892</td>
<td>1 430 911</td>
<td>1 176 343</td>
<td>3 453 477</td>
</tr>
<tr>
<td>0.80</td>
<td>565 141</td>
<td>302 048</td>
<td>282 529</td>
<td>525 568</td>
<td>385 525</td>
<td>1 179 969</td>
</tr>
</tbody>
</table>

Table 5.2: Number of created edges for each subject using different correlation thresholds

As one can observe on Table 5.3 even for a low correlation threshold the edge density (computed as explained in section 4.2) is very low, which makes the network sparse. This was an expected result since in previous state of the art works all the BFC networks were found to be sparse.

<table>
<thead>
<tr>
<th>T</th>
<th>Edge average</th>
<th>Edge density</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>668 989 847</td>
<td>8.2067e-04</td>
</tr>
<tr>
<td>0.45</td>
<td>302 385 481</td>
<td>3.7095e-04</td>
</tr>
<tr>
<td>0.50</td>
<td>136 909 022</td>
<td>1.6795e-04</td>
</tr>
<tr>
<td>0.55</td>
<td>61 017 717</td>
<td>7.4852e-05</td>
</tr>
<tr>
<td>0.60</td>
<td>26 350 364</td>
<td>3.2325e-05</td>
</tr>
<tr>
<td>0.65</td>
<td>10 887 184</td>
<td>1.3356e-05</td>
</tr>
<tr>
<td>0.70</td>
<td>4 263 328</td>
<td>5.2300e-06</td>
</tr>
<tr>
<td>0.75</td>
<td>1 572 970</td>
<td>1.9296e-06</td>
</tr>
<tr>
<td>0.80</td>
<td>540 130</td>
<td>6.6260e-07</td>
</tr>
</tbody>
</table>

Table 5.3: Number of created edges for each subject using different correlation thresholds
5. Results

Figure 5.1 also illustrates how much the threshold influences the total number of edges, in particular it allow us to yield the conclusion that there is a relation between the increment in the threshold and the relative number of added edges that remains constant throughout the range of chosen thresholds.

![Graph showing the number of edges vs. correlation threshold](image)

**Figure 5.1: Number of created edges by each chosen correlation threshold**

5.2.2 Connected components

In order to chose an appropriate threshold it is required to check how much information about the network is lost when passing from a low threshold to a higher one. To check for this information the size of the largest connected component of the network was computed and compared with the total number of vertices in the network. If for a given correlation threshold the biggest component has a low percentage of total vertices of the network it means that it is not suitable to use that threshold.

The results regarding the number of vertices of the biggest component with relation to the total number of vertices in the network are presented in Table 5.4.

From the results presented above it is easy to conclude that if the threshold is too high then the network loses its connectivity and the amount of information lost is too high. We could infer that, on average, for a correlation threshold between 0.4 and 0.5 little information is lost, whereas above that we will starting to have is a significant lost of information. This effect is even more relevant for the networks that have a lower number of edges. In order to check which thresholds
5.2 BFC networks

Table 5.4: Percentage of the total vertices that are in the giant component of the network

<table>
<thead>
<tr>
<th>Subject</th>
<th>T 1</th>
<th>T 2</th>
<th>T 3</th>
<th>T 4</th>
<th>T 5</th>
<th>T 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>0.45</td>
<td>98%</td>
<td>100%</td>
<td>99%</td>
<td>100%</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>0.50</td>
<td>77%</td>
<td>100%</td>
<td>91%</td>
<td>100%</td>
<td>88%</td>
<td>99%</td>
</tr>
<tr>
<td>0.55</td>
<td>53%</td>
<td>96%</td>
<td>58%</td>
<td>99%</td>
<td>58%</td>
<td>92%</td>
</tr>
<tr>
<td>0.60</td>
<td>37%</td>
<td>77%</td>
<td>44%</td>
<td>94%</td>
<td>35%</td>
<td>71%</td>
</tr>
<tr>
<td>0.65</td>
<td>26%</td>
<td>48%</td>
<td>32%</td>
<td>74%</td>
<td>23%</td>
<td>55%</td>
</tr>
<tr>
<td>0.70</td>
<td>16%</td>
<td>28%</td>
<td>22%</td>
<td>46%</td>
<td>14%</td>
<td>42%</td>
</tr>
<tr>
<td>0.75</td>
<td>8%</td>
<td>14%</td>
<td>13%</td>
<td>24%</td>
<td>7%</td>
<td>30%</td>
</tr>
<tr>
<td>0.80</td>
<td>3%</td>
<td>4%</td>
<td>6%</td>
<td>8%</td>
<td>3%</td>
<td>16%</td>
</tr>
</tbody>
</table>

Table 5.4: Percentage of the total vertices that are in the giant component of the network

meet the structural properties found in related works we also computed the degree distribution before choosing the final networks to work on.

5.2.3 Vertex degree distribution

Regarding the vertex degree distribution for the BFC networks it is also dependent on the chosen correlation threshold. As it affects the total number of edges in the network it will also have an effect on the degree of the vertices, thus for each network there will be a different degree distribution.

To check if our BFC networks exhibit properties similar to the ones already studied in other state of the art works their degree distribution should follow a power law, with an exponent between 2 and 3 according to Equation 2.5. In order to know which correlation threshold yields a BFC network closer to the ones found in other related works, for each subject and for each threshold the degree distribution function was computed and fitted with a power law. The results of the power law exponent that fits each degree distribution function are shown on Table 5.5.

Table 5.5: Value of the exponent from the fitting function of the degree distribution

<table>
<thead>
<tr>
<th>Subject</th>
<th>T 1</th>
<th>T 2</th>
<th>T 3</th>
<th>T 4</th>
<th>T 5</th>
<th>T 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>2.09</td>
<td>2.04</td>
<td>2.15</td>
<td>2.16</td>
<td>2.04</td>
<td>2.28</td>
</tr>
<tr>
<td>0.45</td>
<td>2.00</td>
<td>1.92</td>
<td>2.04</td>
<td>2.06</td>
<td>1.97</td>
<td>2.18</td>
</tr>
<tr>
<td>0.50</td>
<td>1.92</td>
<td>1.83</td>
<td>1.95</td>
<td>1.95</td>
<td>1.89</td>
<td>2.09</td>
</tr>
<tr>
<td>0.55</td>
<td>1.85</td>
<td>1.66</td>
<td>1.86</td>
<td>1.84</td>
<td>1.77</td>
<td>1.96</td>
</tr>
<tr>
<td>0.60</td>
<td>1.80</td>
<td>1.64</td>
<td>1.79</td>
<td>1.72</td>
<td>1.77</td>
<td>1.88</td>
</tr>
<tr>
<td>0.65</td>
<td>1.77</td>
<td>1.63</td>
<td>1.74</td>
<td>1.66</td>
<td>1.75</td>
<td>1.70</td>
</tr>
<tr>
<td>0.70</td>
<td>1.73</td>
<td>1.60</td>
<td>1.66</td>
<td>1.59</td>
<td>1.73</td>
<td>1.65</td>
</tr>
<tr>
<td>0.75</td>
<td>1.81</td>
<td>1.56</td>
<td>1.61</td>
<td>1.54</td>
<td>1.72</td>
<td>1.63</td>
</tr>
<tr>
<td>0.80</td>
<td>1.78</td>
<td>1.52</td>
<td>1.55</td>
<td>1.51</td>
<td>1.67</td>
<td>1.59</td>
</tr>
</tbody>
</table>

Table 5.5: Value of the exponent from the fitting function of the degree distribution

It is possible to see from Table 5.5 that the networks whose degree distribution is closer to the ones reported in other state of the art works are the ones corresponding to lower thresholds. This
5. Results

is expected as the edge density for the networks with higher correlation thresholds is very low. An example of the resulting degree distribution function can be seen in Figure 5.2.

![Degree distribution function for the BFC network of the subject 6 and with a correlation threshold of 0.4](image)

Figure 5.2: Degree distribution function for the BFC network of the subject 6 and with a correlation threshold of 0.4

5.2.4 Small worldness

Based on the previously obtained results, from this point on, the only networks that will be considered for further analysis are the ones obtained with a correlation threshold between 0.4 and 0.5.

For the small-worldness of those networks to be proved we are required to compute the minimum average path - that should be proportional to the logarithm of the total number of vertices of the network - and the clustering coefficient - that should be much higher than its random equivalent network.

5.2.4.1 Minimum average path

For each of the six subjects the minimum average path was estimated with the results presented on Table 5.6. It is easy to observe that for all the networks, even those with a smaller number of edges, the minimum average path is very small when compared with the total number of vertices in the network.

This is an important property of the networks that have a small-world topology, it is possible to go from any vertex to any other with a small number of steps.
5.2 BFC networks

To prove the small-world topology we need to compute the minimum average path in all the respective random equivalent networks. With this information we are now able to compute the \( \lambda \) coefficient. All the results regarding these computations are presented on Table 5.7.

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>4.47</td>
<td>3.63</td>
<td>4.95</td>
<td>3.24</td>
<td>4.81</td>
<td>3.58</td>
</tr>
<tr>
<td>0.45</td>
<td>6.64</td>
<td>4.68</td>
<td>6.96</td>
<td>4.04</td>
<td>6.90</td>
<td>4.66</td>
</tr>
<tr>
<td>0.5</td>
<td>7.53</td>
<td>6.25</td>
<td>9.94</td>
<td>5.22</td>
<td>9.75</td>
<td>6.23</td>
</tr>
</tbody>
</table>

Table 5.6: Estimation of the minimum average path for the networks of all subjects considering only three different correlation thresholds.

<table>
<thead>
<tr>
<th>T</th>
<th>BFC</th>
<th>rand</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>4.113</td>
<td>3.251</td>
<td>1.265</td>
</tr>
<tr>
<td>0.45</td>
<td>5.650</td>
<td>3.509</td>
<td>1.610</td>
</tr>
<tr>
<td>0.5</td>
<td>7.487</td>
<td>4.698</td>
<td>1.498</td>
</tr>
</tbody>
</table>

Table 5.7: Average characteristic path for the BFC networks, their respective random equivalents and value of the \( \lambda \) coefficient.

As one can see from Table 5.7, the minimum average path of all the BFC networks is almost as low as the one from their random equivalents, which is exactly what usually happens in small-world networks (Watts and Strogatz 1998).

5.2.4.4 Clustering coefficient

Regarding the clustering coefficient, in order to prove the small-world topology it is also required to compute the coefficient for every network and also for its respective random equivalent network. For these networks to have a small-world topology they should always have a higher cluster coefficient than its random equivalent. With both of them computed it is possible to compute the \( \gamma \) coefficient.

The results regarding all these computations are presented in Table 5.8 and 5.9.

<table>
<thead>
<tr>
<th>T</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.181</td>
<td>0.240</td>
<td>0.178</td>
<td>0.243</td>
<td>0.186</td>
<td>0.251</td>
</tr>
<tr>
<td>0.45</td>
<td>0.159</td>
<td>0.223</td>
<td>0.159</td>
<td>0.243</td>
<td>0.158</td>
<td>0.239</td>
</tr>
<tr>
<td>0.5</td>
<td>0.146</td>
<td>0.190</td>
<td>0.145</td>
<td>0.220</td>
<td>0.136</td>
<td>0.201</td>
</tr>
</tbody>
</table>

Table 5.8: Clustering coefficient from the networks of all subjects.

From the previous results it is possible to estimate the \( \sigma \) coefficient, presented in Equation 3.7 with the results shown in Table 5.10.
5. Results

<table>
<thead>
<tr>
<th>T</th>
<th>BFC</th>
<th>rand</th>
<th>γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.213</td>
<td>0.053</td>
<td>4.102</td>
</tr>
<tr>
<td>0.45</td>
<td>0.197</td>
<td>0.042</td>
<td>4.690</td>
</tr>
<tr>
<td>0.5</td>
<td>0.173</td>
<td>0.039</td>
<td>4.436</td>
</tr>
</tbody>
</table>

Table 5.9: Average clustering coefficient of the BFC networks, their respective random equivalent networks and value of the γ coefficient.

<table>
<thead>
<tr>
<th>T</th>
<th>γ</th>
<th>λ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>4.102</td>
<td>1.265</td>
<td>3.243</td>
</tr>
<tr>
<td>0.45</td>
<td>4.690</td>
<td>1.610</td>
<td>2.913</td>
</tr>
<tr>
<td>0.5</td>
<td>4.436</td>
<td>1.498</td>
<td>2.961</td>
</tr>
</tbody>
</table>

Table 5.10: Average small-worldness coefficient of all the BFC networks

With these final results of the σ coefficient we are now able to know that all the studied BFC networks have a small-world topology, as for all of them the σ coefficient is higher than 1, which, as shown in the work of van den Heuvel et al. (2008), is enough to prove it.

5.2.5 Community detection

For each of the three different BFC networks of all the subjects a community detection algorithm was applied with the purpose of finding functional modules of the brain. In every network it was found a number of different clusters, however as far as this analysis is concerned only the 6 major clusters of each network were considered, i.e., the clusters with higher number of vertices.

In each subject the results of the community detection algorithm were intrinsically related with the correlation threshold that was chosen to generate the network. In Figures 5.3, 5.4 and 5.5 the results of running the algorithm are shown for a single subject with each of the networks being generated with a different correlation threshold.

The colours representing each community are related with the total number of vertices in it, having in increasing order: light blue, red, blue, green, yellow and pink.
5.2 BFC networks

Figure 5.3: Community detection for subject 6 at a correlation threshold of 0.4

Figure 5.4: Community detection for subject 6 at a correlation threshold of 0.45
5. Results

It is possible to notice several differences between the three figures, the most obvious is that the lower the correlation threshold the more volume of the brain is covered by the communities that are represented. As the threshold increases the size of the modules decreases and therefore from one figure to another some of the modules are split in smaller ones.

Besides the variation of results associated with the choice of the correlation threshold, there is also a difference in the results according to each subject. Not only because the shape of the brain is different but also the size and location of the communities varies from subject to subject. This is an expected result because, as it has been shown in previous sections, although the BFC networks of all subjects share the same topological properties there are some clear differences regarding the structure (e.g.: number of edges and number of vertices).

In the following figures it is illustrated the differences between subjects, as far as community detection results are concerned.
5.2 BFC networks

Figure 5.7: Community detection for subject 2 at a correlation threshold of 0.45

Figure 5.8: Community detection for subject 3 at a correlation threshold of 0.45

Figure 5.9: Community detection for subject 4 at a correlation threshold of 0.45
5. Results

From Figures 5.6, 5.7, 5.8, 5.9, 5.10 and 5.4 that represent the six major modules found in all six subjects it is possible to see very clearly that the brain’s shape varies throughout all subjects. This will affect directly the location and shape of the found modules and it is also an evidence on why some subjects’ BFC networks have more vertices and edges than others. On the other hand, it is possible to see that some modules are in similar locations, such as the brainstem, the frontal and the occipital areas of the brain.

However, in order to validate these results the found modules must be compared with the resulting data provided by other analysis to the same datasets. We used the independent component analysis (ICA) and computed it for all subject networks in order to find 20 different ICs. The analysis was conducted using the Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) made available by the Oxford Centre for functional Magnetic Resonance Imaging of the Brain (FMRIB) software library (FSL) version 5.0.6 with MELODIC version 3.14 (Beckmann and Smith, 2004).

For the LLP analysis only the six major modules were represented because on average the other modules were very small when compared with the average size of the IC. Even tough we only chose the six largest some of them are up to five times smaller than the usual IC. The results regarding the comparison between the IC and the modules found with the LLP are presented in the next section.

5.3 Validation

After conducting the ICA we have for each of the networks 20 different IC. Having these IC computed we have compared them with the modules already found with the LLP algorithm. To measure if the results of both analysis are similar, we used two different metrics. First we measured the overlap between the modules found with LLP and the ICA, then we computed their
Normalized Mutual Information (NMI). The NMI is a metric that tells us for an arbitrary random variable A and another random variable B how much information one variable tells about the other, in our case it will tell us how much of the information about one given module found with LLP is contained on another found using the ICA. The value of the NMI is 1 if all the information about one random variable is contained on the other and 0 if the variable A has absolutely no information about variable B.

The first analysis made was intended to find the match between the modules of the LLP and the IC which had the most overlap between them. The modules are numbered from m1 to m6 and the IC are numbered from IC1 to IC20. The results from that matching and its respective result are presented in Table 5.11.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Subject1</th>
<th>Subject2</th>
<th>Subject3</th>
<th>Subject4</th>
<th>Subject5</th>
<th>Subject6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>m4 &amp; IC13</td>
<td>no overlap</td>
<td>m6 &amp; IC15</td>
<td>no overlap</td>
<td>no overlap</td>
<td>m2 &amp; IC12</td>
</tr>
<tr>
<td></td>
<td>68.38%</td>
<td>over 60%</td>
<td>74.30%</td>
<td>over 60%</td>
<td>over 60%</td>
<td>62.92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>m5 &amp; IC20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64.50%</td>
</tr>
<tr>
<td>0.45</td>
<td>m4 &amp; IC13</td>
<td>m3 &amp; IC9</td>
<td>m3 &amp; IC7</td>
<td>m2 &amp; IC4</td>
<td>m2 &amp; IC14</td>
<td>m3 &amp; IC12</td>
</tr>
<tr>
<td></td>
<td>76.32%</td>
<td>63.66%</td>
<td>69.13%</td>
<td>85.56%</td>
<td>60.34%</td>
<td>73.41%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>m3 &amp; IC5</td>
<td>m4 &amp; IC8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>66.04%</td>
<td>91.59%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m5 &amp; IC15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>79.68%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>m6 &amp; IC13</td>
<td>m3 &amp; IC5</td>
<td>m5 &amp; IC8</td>
<td>m2 &amp; IC4</td>
<td>m4 &amp; IC12</td>
<td>m5 &amp; IC16</td>
</tr>
<tr>
<td></td>
<td>90.57%</td>
<td>93.23%</td>
<td>91.46%</td>
<td>93.01%</td>
<td>80.88%</td>
<td>66.95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>85.00%</td>
<td>m5 &amp; IC13</td>
<td>m5 &amp; IC16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>69.61%</td>
<td>61.82%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.11: Overlap between the modules found with LLP and the IC found with ICA. Only the overlaps over 60% are referenced.

As one can see from Table 5.11 there is some significant overlap between some modules found by LLP and IC found by ICA, with some of them up to 90%. This is a very relevant result as it proves that our analysis made with the LLP algorithm has very likely found relevant modules of the brain because it is supported by the results of ICA. It is important to stress the fact that what Table 5.11 represents is the percentage of vertices from a given module that are contained on a given IC.

It is also possible to see that there is a significant overlap of the modules with the ICs in almost every subjects’ networks at all the three chosen threshold levels; however it is noticeable that some thresholds have better results than others.

On Figure 5.11 the three modules from the BFC network of subject 3, that are referenced on Table 5.11 as having over 60% of their vertices contained in some IC, are represented. Figure 5.12 shows the ICs where those modules are contained. As can be seen, both images are very similar, with the modules that have a higher overlap with the IC being the ones that seem almost
the same. Module m3 and the IC7 are represented in red; module m4 and the IC8 are represented in green and module m5 and the IC15 are represented in yellow.

Figure 5.11: Three modules found with LLP for subject 3 at a correlation threshold of 0.45

Figure 5.12: Three IC found with ICA for subject 3 at a correlation threshold of 0.45

Now that the overlap has been computed it is in our interest to compute the NMI, as it will give us a more accurate measure of how much information does the modules and the IC actually share. This can be difficult to measure with the overlap alone because for instance the size of the modules and of the IC can be different and while a module may be 90% contained in the IC if the module is smaller than the IC then it may happen that the IC is only 50% contained in the module. To account for such differences the NMI has been computed for all the modules and IC that are present in Table 5.11 with the results presented in Table 5.12.
5.4 Discussion

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Subject 1</th>
<th>Subject 2</th>
<th>Subject 3</th>
<th>Subject 4</th>
<th>Subject 5</th>
<th>Subject 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>m4 &amp; IC13</td>
<td>not</td>
<td>m6 &amp; IC15</td>
<td>not</td>
<td>not</td>
<td>m2 &amp; IC12</td>
</tr>
<tr>
<td></td>
<td>0.345</td>
<td>considered</td>
<td>0.476</td>
<td>considered</td>
<td>considered</td>
<td>0.390</td>
</tr>
<tr>
<td>0.45</td>
<td>m4 &amp; IC13</td>
<td>m4 &amp; IC9</td>
<td>m3 &amp; IC7</td>
<td>m2 &amp; IC4</td>
<td>m4 &amp; IC14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.341</td>
<td>0.311</td>
<td>0.393</td>
<td>0.410</td>
<td>0.258</td>
<td></td>
</tr>
<tr>
<td></td>
<td>m3 &amp; IC5</td>
<td>m4 &amp; IC8</td>
<td>m5 &amp; IC15</td>
<td>0.414</td>
<td>m3 &amp; IC12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.389</td>
<td>0.531</td>
<td>0.498</td>
<td></td>
<td>0.414</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>m6 &amp; IC13</td>
<td>m3 &amp; IC5</td>
<td>m5 &amp; IC8</td>
<td>m2 &amp; IC4</td>
<td>m4 &amp; IC12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.323</td>
<td>0.487</td>
<td>0.427</td>
<td>0.375</td>
<td>0.442</td>
<td></td>
</tr>
<tr>
<td></td>
<td>m5 &amp; IC15</td>
<td>m6 &amp; IC15</td>
<td>m5 &amp; IC16</td>
<td>0.313</td>
<td>m5 &amp; IC16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.465</td>
<td>0.465</td>
<td>0.313</td>
<td>0.188</td>
<td>m5 &amp; IC16</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.12: Normalized mutual information between the modules and the IC presented on Table 5.11

From Table 5.12 it is possible to see that almost all the modules and ICs that were chosen have an NMI between 0.3 and 0.5. This may seem an unexpected result because of the high percentage of vertices from the modules that are contained in the ICs. However it is important to stress the fact that the size of the modules sometimes is quite different from the size of the ICs and that means that although the majority of the vertices from the module overlaps the IC there is still a number of vertices from the IC that is outside of that given module.

However, as far as the clustering evaluation is concerned, the expected value of the NMI for us to be able to say that two clusters are similar is usually above 0.5, which means that in fact our method yielded some interesting results, because some of them are very close to that. However, there are still improvements to be done.

5.4 Discussion

After all the analysis of the networks are performed we will now evaluate the achieved results taking into account other state of the art results and the breakthroughs that have been achieved in this work.

5.4.1 BFC networks basic properties

The networks that have been drawn from the initial 7T fMRI data have a much higher granular level than all the others studied in related works. Regardless of that fact several studies have been performed on lower resolution voxel-based BFC networks. Therefore we have some possible comparisons that can be done in order to evaluate our results.

Regarding the number of vertices and edges of the networks we know that they were supposed to be much higher than all the networks in other related works, nevertheless the number of edges
5. Results

should not be too high because all the voxel-based networks are sparse. This fact is verified in all our networks, even for the ones generated with a low correlation threshold, which is a good result.

Regarding the choice of the best correlation threshold we have to check for the loss of information that happens when we increase the threshold but also the properties that the network should have. In our case, on average, there was a significant loss of information for a threshold higher than 0.5, which is a fairly good result if we consider that most of the works with voxel-based networks also work only with these thresholds. For instance on the work of van den Heuvel et al. (2008), where a detailed characterization of voxel-based networks is performed (using networks with around 9500 vertices), the thresholds used are between 0.7 and 0. In our case the thresholds could not be too low because the networks would become dense and even more difficult to handle.

Now to check which networks had not only a low loss of information but also a similar structure to what has been reported in related works, we computed the degree distribution function and the respective exponent of the function. The best results obtained were also for the thresholds above 0.5, thus we concluded that the most accurate networks to work on, as far as this work is concerned, are those with a correlation threshold between 0.4 and 0.5.

5.4.2 Networks’ topology

Having for each subject only three networks we then analysed them in order to check if their topology also matched other state of the art works. This analysis was specifically intended to find if these networks have a small-world structure, which is a very important aspect of BFC networks, as previously explained.

For us to know if the networks had a small-world topology we computed the average shortest path and their clustering coefficient, which from graph theory should follow the rules specified in Section 3.3.4 and have a small-world coefficient greater than one. After computing all of these metrics the results were quite good because all of the initially chosen networks for us to work on had a small-world topology. This means that even at this high resolution it is still possible to see some clear organization in the brain’s network in order to flow information from one place to another in a very short number of hops.

5.4.3 Community detection

Detecting communities in the BFC network will correspond to finding clusters of voxels that have a high edge density between them. This means that they have similar BOLD signal fluctuations at rest and therefore we can say that they may share functional properties.

For us to detect such communities we used a cluster detection algorithm and found several different clusters for each subject’s network. In order to evaluate the significance of the clusters that were found we also conducted an ICA to the network and found 20 ICs. These IC will also
consist in groups of voxels with similar BOLD signal fluctuations but that were found using a different approach that based on the properties of the signal tries to isolate voxels with similar signals into ICs.

From all the clusters that we found only the six larger ones were considered for the analysis because those were the ones which had a size with a better match with the size of the average IC.

Having all the clusters and the IC computed we measured, for each BFC network, which pairs of clusters and ICs had the most overlap, where the measured overlap is the proportion of vertices from the cluster that is inside the IC. This gave us for each network a set of pairs of clusters and ICs. Then the NMI was computed in order to give a more accurate statistic measure of similarity between those pairs.

The results of the overlap were very interesting, with some of the clusters almost fully contained in some given IC of the same network. However when the NMI was computed only few of the pairs had a NMI close to 0.5, which means that despite the overlap in most of the cases there was not a very clear statistical similarity and therefore the cluster did not have enough information about the IC and vice versa. This might be related with the fact the sizes of the ICs are often different from the clusters and therefore, in spite of the fact that the cluster is fully contained on the IC, there are still a number of vertices that are in the IC alone.
5. Results
6 Conclusions

Contents

6.1 Future work ........................................... 52
6. Conclusions

Resting state fMRI connectivity is increasingly becoming an interesting and widely studied topic of research and nowadays a lot of works are conducted in this field. This fact is allowing a deeper and more accurate knowledge of the human brain, which for the scientific community is still the most complex organ of the human body.

Increasingly, new research efforts concentrate on dealing with high resolution images of the brain because they allow more granular and precise results. Thus, this master’s thesis also uses high resolution fMRIs with the intent of producing a detailed mapping of the brain and to reconstruct for each image the brain functional connectivity network.

The results of our work were very interesting, as far as reconstructing the BFC network from high resolution fMRIs is concerned, because to the best of our knowledge it has not been built any tool that allows a reconstruction of such high resolution networks. Furthermore, our results also showed that the structural properties of the networks are similar to the ones found in low resolution networks, which is also an important aspect because even at high resolution we found that there is an evident ability of the brain’s network to flow information in a very efficient way.

Regarding the detection of functional modules our approach revealed to be efficient but with results that still have to be improved, namely regarding the similarity with the ICs found for the same data.

6.1 Future work

The work developed during this project still has some parts that are inefficient and time consuming, namely in what concerns the reconstruction of the BFC networks that even though it has been very optimized it takes too much computational time. Therefore, one possible way is to use parallel computation for instance while computing all the sub-matrix chunks of the correlation matrix used to generate the network since all of them are independent the computation can be carried out in parallel. However, there is always a trade-off between the number of cores used and the number of chunks in the correlation matrix, because if the chunks are too large then they will consume a lot of memory and therefore it is only possible to parallelize with few cores. If the chunks are small they occupy less memory and we are able to use more cores but there are a lot more chunks to compute.

Regarding the BFC network analysis there are also more advanced metrics that can be computed and more detailed modularity analysis that can be made. For instance, it can be performed, for each cluster that we found, another modularity analysis and check for clusters within the clusters.

It is also important to stress the fact that all the performed analysis are based on all the previous pre-processing of the data, hence the better and more advanced are the methods used to remove noise from the data the more interesting and accurate the results will be. However, all
these techniques are also complex specially in such datasets that cover areas of the brain, such as the brainstem, that are very exposed to noise and they also need improvements.
6. Conclusions


Appendix

Contents

A.1 Abstract accepted in the Fourth Biennial Conference on Resting State Brain Connectivity ........................................ 60
A. Appendix

A.1 Abstract accepted in the Fourth Biennial Conference on Resting State / Brain Connectivity

Title: Large functional connectivity network analysis of whole-brain high-resolution resting-state fMRI

Authors: A. C. Leitão¹, A. P. Francisco¹, R. Abreu², P. Figueiredo², L. L. Wald³, L. M. Silveira¹

Author affiliations: ¹INESC-ID/ Instituto Superior Técnico – Universidade de Lisboa, Portugal; ²ISR/ Instituto Superior Técnico – Universidade de Lisboa, Portugal; ³Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA USA.

Background: It has recently become possible to collect fMRI data from the whole-brain for the analysis of resting-state BOLD fluctuations at ultra-high-field (e.g. at 7T). The resulting high-resolution datasets cover a wider volume of the brain, including the cerebellum and the brainstem. In this work, we apply graph theory algorithms to investigate the brain functional connectivity (BFC) networks obtained from high dimensionality data based on full voxel-based correlation analysis. Using such techniques we can measure basic properties of the network's structure and find functional modules of the brain. The high resolution enabled a much higher granularity than previous works.

Methods: Resting-state fMRI data were collected from a group of healthy volunteers on a 7T Siemens whole-body scanner equipped with custom-built 32-channel receive coils, using a 3x slice accelerated SMS EPI sequence (TE/TR=32ms/2500ms, 210 volumes), with 1.1mm isotropic voxels covering the whole-brain in 123 sagittal slices (with a Field of View (FOV) of 264x198mm2). Standard pre-processing was performed using FSL, including: slice scan time correction, motion correction, non-brain removal, spatial smoothing (FWHM=1.5mm Gaussian kernel) and high-pass temporal filtering (100ms cut-off). BFC networks were obtained with vertices representing voxels and edges functional connections, computed as correlations between voxels' BOLD signals. Basic properties of the network's structure were extracted including degree distribution, connected components, average distance and clustering coefficient. A preliminary modularity analysis of the network was conducted in order to find brain functional modules.

Results: BFC graphs with over one million vertices were obtained. The unusually high dimensionality of the data required specialized algorithms to be implemented to enable construction and subsequent analysis of the networks. We found that these networks have: 1) a small-world topology (average minimum path is low compared with the number of vertices and the clustering coefficient is high compared with the coefficient of the equivalent random network); and 2) a scale-free organization (taking into account computed centrality measures distributions). The BFC modules obtained were consistent with those reported in previous works, although showing finer-grained modules.

Conclusions: We obtained large voxel-based BFC networks from whole-brain 1.1mm isotropic...
resting-state fMRI data, and found small-world and scale-free properties as previously reported for smaller networks. We did find a much higher module granularity, which should be further investigated. A systematic analysis of the network modularity will be performed using appropriate methods that we have found to be efficient in such big networks, and compared with conventional analysis.

Acknowledgements: This work was partly supported by national funds through FCT – Fundação para a Ciência e Tecnologia, under projects PTDC/EEI-ELC/3246/2012 and PEst-OE/EEI/LA0021/2013.