Politecnico di Milano

School of Industrial and Information Engineering
Master of Science in Biomedical Engineering

Analysis of structural connectivity in brain disorders and validation of a novel tool for the investigation of brain connectivity

Supervisor: Prof. Giuseppe Baselli

Co-Advisors: M. Sc. Davide Coluzzi
Eng. Alice Pirastru
Eng/PhD. Laura Pelizzari

Master Sc. Thesis of:

Vanja Curcic 915165

Academic year 2020/2021

Acknowledgements

I wish to express my gratitude to Professor Giuseppe Baselli for giving me the opportunity to work on this topic, and on his availability and willingness to always provide constructive advice and ideas for the work.

I also want to express my deepest thankfulness to Davide Coluzzi who followed my work from the start to the end, always supporting me and inspiring me. It was an honest pleasure working with you.

Lastly, I thank Valeria Blasi, Alice Pirastru, Francesca Baglio, Laura Pelizzari and Monia Cabinio from MRI Laboratory of Fondazione Don Carlo Gnocchi for collaboration in this project and their advice and feedbacks.

To my two grandfathers who are not present to see me reaching this point.

Мојим родитељима који су увек чврсто стајали иза мене, пружајући ми безусловну подршку.
Мот брату, од кога свакодневно учим.

Abstract

Anatomical tracts build structural associations connecting brain areas into networks that constitute brain connectivity datasets and may be non-invasively explored via diffusion tensor imaging (DTI). In parallel, functional MRI (fMRI) correlations among brain areas permit to quantify the behavior of the brain network. Nonetheless, both the salient features of either information and their relationship is still an open research question. Complex network analysis, a novel multidisciplinary approach that stems from graph theory, provides an important framework and valuable insights in studying these datasets. As such, it aims to represent brain networks using a small number of topological indices, which are studied as potential neurological markers. It is common knowledge that connectivity abnormalities due to injuries and neurological disorders lead to a disruption of the regular brain connectivity pattern. Hereby, network metrics are employed to characterize network components, and compare networks between groups of subjects: e.g., patients and controls in clinical investigations.

This work is centered on the possible improvement in brain connectivity analysis by validating a novel software tool used for focused analysis on subgraphs of interest, we called SPIDER-Net tool (SNT), and also on the validation of the graph indexes in whole networks or in subnetworks extracted by SNT. In fact, this tool permits selecting specific subsets of brain parcels, seen as graph nodes, and compute network metrics with subsequent visualization of both intra-areal and inter-areal brain connections in a preselected region. The software was tested in a case study of stroke. A case presented with disruptions of connectivity in the frontal, temporal, and parietal regions of the right hemisphere, which was confirmed by employing the new software.

The present work addressed structural connectivity only, while a validation on functional connectivity data and structural to functional integration is left to future developments. Focusing on structural connectivity from DTI tractography, the present work addressed the analysis of uncertainties depending on the kind of chosen metrics, whether fractional anisotropy (FA) or the streamline number (SN), showing that a proper use of SNT, with its flexibility in comparing thresholding approaches and levels and with the immediate availability of graph indexes, is a valid mean to disentangle metrics dependent artifactual features from the real patho-physiological ones.

Furthermore, this work includes consideration about the utility of weights extracted from standard MRI methodologies, with particular attention to weak connections in the weighted graphs. The conducted analysis is based on a clinical study of early life adversity as a major risk factor for borderline intellectual functioning in children. A major finding was the relevance of

weighted graph analysis compared to the binarized one, which requires thresholding. Considering that DTI tractography is an inherently noisy technique, the usual pipeline in connectivity studies involves thresholding of low weight connections which are considered spurious. Conversely, the present work includes analysis on whether weak connections are valuable in connectivity studies. Significative difference was found between patients and controls, only if keeping the weak connections only. However, further tests are necessary for confirming these results.

Sommario

I tratti anatomici connettono aree cerebrali in reti che costituiscono dataset di connettività cerebrale e possono essere esploarate in modo non invasivo mediante le immagini del tensore di diffusione (DTi). In parallelo, le correlazioni fra aree cerebrali della MRI funzionale (fMRI) permettono di quantificare il comportamento della rete cerebrale. L'analisi delle reti complesse fornisce preziose informazioni nello studio di questi dataset. Inoltre è possibile rappresentare le reti tramite un piccolo numero di metriche topologiche e neurobiologicamente rilevanti che le descrivono e che sono, facilmente calcolabili. È risaputo lesioni dovute a disturbi neurologici portano a anomalie del normale modello di connettività cerebrale. Pertanto, le metriche delle reti vengono impiegate per caratterizzarne lecomponenti o confrontare gruppi di soggetti, e.g. pazienti e gruppo di controllo nelle indagini cliniche.

Questo lavoro è incentrato su un possibile miglioramento nell'analisi della connettività cerebrale mediante un nuovo software utilizzato per l'analisi su sottografi di interesse, SPIDER-Net tool (SNT), e pure una validazione degli indici dei grafi in reti complete o in sotto-reti estratte da SNT. Infatti, tale strumento consente di selezionare specifici sottoinsiemi delle aree cerebrali, viste come nodi del grafo, e di calcolare le metriche caratteristiche delle reti con successiva visualizzazione delle connessioni cerebrali sia intra-area che inter-area. Il software è stato testato in un caso di studio di ictus cerebrale. Il paziente presentava interruzioni della connettività nelle regioni frontale, temporale e parietale dell'emisfero destro, le quali sono state visualizzate e confermate tramite il software.

Il presente lavoro si è rivolto alla sola connettività strutturale, mentre una validazine relativa a quella funzionale e una integrazione strutturale-funzionale è lasciata a sviluppi futuri. Focalizzando la connettività funzionale, derivata dalla trattografia DTI, il presente lavoro ha considerato l'analisi dell'incertezza dipendente dalla metrica scelta, che può essere l'anisotropia frazionaria (FA) oppure il numero di fibre virtuali (streamlines, SN), mostrando che l'applicazione appropriata di SNT, con la sua flessibilità nella comparazione di metodi e livelli di sogliatura e con la disponibilità immediata degli indici di grafo, è un mezzo atto a separare le caratteristiche artefatti dipendenti dalla metrica da quelle pato-fisiologiche reali.

Inoltre, questo lavoro include considerazioni sulla utilità dei pesi estratti dalle metodologie MRI standard, con paricolare attenzione alla connessioni deboli nei grafici pesati. L'analisi condotta si basa su uno studio clinico sulle avversità precoci della vita come un importante fattore di rischio per il funzionamento intellettivo borderline nei bambini. Una scoperta importante è stata la rilevanza dell'analisi del grafico ponderato rispetto a quella binarizzata, dopo la sogliatura. Considerando che la trattografia DTI è una tecnica intrinsecamente rumorosa, la consueta

pipeline negli studi di connettività prevede il thresholding di connessioni a basso peso che sono considerate spurie. Al contrario, il presente lavoro include un'analisi sulla validità delle connessioni deboli negli studi sulla connettività. Una differenza significativa è stata trovata tra i pazienti e i controlli basandosi solo se le connessioni deboli erano mantenute. Tuttavia, sono necessari ulteriori test per confermare questi risultati.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Aims of the thesis work

Brain networks have been a long-standing intellectual current in theoretical neuroscience. Nonetheless, it was difficult, if not impossible, to directly examine and analyze the anatomical and functional networks of the brain until recently. The advanced techniques required for mapping extensive anatomical networks and recording functional brain activity across vast neuronal populations, or even the entire brain, took a long time to evolve. All of this has changed in the recent years. New tools for mapping and recording large-scale neuronal networks were introduced, yielding a wealth of data on neural system anatomical layout and functional dynamics.

Graph theory has the potential to give a relatively simple but strong quantitative framework for defining and comparing complete human brain structural and functional networks under a variety of experimental and clinical conditions. Hence, it is increasingly being used in analyzing neuroimaging data.

This thesis starts with describing the basic concepts of brain connectivity and graph theoretical methods. Next, focus is brought to the case study of stroke. With the goal of visualization and analysis of connectivity, a new approach is introduced.

This approach utilizes a novel software we called SPIDER-Net (Software Package for Insight and Enhanced Representation of a Net) for studying brain connectivity developed as a result of collaboration between Politecnico di Milano, Milano (IT) and CADiTeR, MRI Laboratory of Fondazione Don Carlo Gnocchi, Milano (IT). The main motivation for the overall project was to design a tool which will provide more convenient and advantageous mean for exploring brain connectivity, permitting to focus connection subsets, thus facilitating a systematic exploration of target networks, which was increasingly manually applied by the biomedical counterpart. Moreover, the requirement for this tool was to provide both structural and functional connectivity study, with computation of network topological properties and visualization of brain connectivity.

Accent is put on enhancing the visualization of brain networks. Taking into the consideration that brain consists of overwhelming number of intricate connections, cortical networks remain difficult to comprehend when visualized or assessed all together.

The massive quantity of data describing the connectedness of neural elements results in difficulties to meaningfully display or successfully perform visual inquiry. A major leap was made by the authors of the study [1] who introduced a new method for visualization of brain connectivity. In this work they introduced the term “connectogram” - circular representation of human cortical networks. Still, due to the high density of connectivity matrices stemming from the developed imaging modalities, connectograms may be “overcrowded” with connections making visual inquiry hard, or often impossible.

Our tool provides more efficient way of connectivity exploration by allowing the user to select hemispheric partitions of interest, and following, extracting a subgraph of selected region. In this way visualization can be focused, instead of being hindered by rest of the network connections.

Following this, study on weak connections is made. Common trend in connectivity study in past considered discarding weight information from connectivity matrices by binarizing them, which was supported by considering as spurious weak connections, compared to the noise of weights due to inherent limitations in both the MRI and the connection assessment technniques. However, in recent years weighted network analysis is gaining attention, advocating that broad range of weights in connectivity matrix is providing biological relevance, which might be obscured by arbitrary thresholding and binarization.

A major aim of this thesis was to keep the weak connections instead of thresholding them, and test their utility. This was motivated by many recent studies confirming that the weak connections possibly comprise valuable information. Analyses on binary networks performed by Basi et al. [Basi et al., 2020] were extended to weighted cases. Herby, network metrics calculated on the basis of weighted weak connections were used to compare a group of 32 children with borderline intellectual functioning disorder and 14 typical development children.

1.2 Brain connectivity

Human brain has a dense neuronal architecture forming one of the most complex network systems found in nature. It consists of billions of neurons linked by fibers and synapses organized over multiple scales of space and functionally interact over multiple scales of time. Units of nervous system can be represented by individual neurons, groups of neurons or brain regions segregated anatomically. The pattern of connectivity can be formed by structural connections (fiber pathways, synapses) or functional connections representing statistical or causal relationships measured as cross-correlations, coherence, or information flow. Brain connectivity

limits neural activity, and thus it is crucial in understanding how neural networks and neurons process information [2].

According to a key paradigm in modern neuroscience, anatomical and functional relations between brain regions are arranged in such a way that information processing is near optimal. However, obtaining brain connectivity patterns in vivo, with the elements and connections in different levels remains as outstanding challenge.

Development of imaging techniques made it possible to record neural activity dynamics across the whole brain, and to map entire nervous system anatomical connections. This resulted in the large amount of data on functional dynamic and anatomical organization. Simultaneously, statistical physics development resulted in emerging of interdisciplinary field of "complex network science" which provide conceptual and mathematical tools used for understanding structure and behavior of range of distinct networks [3]. This approach has its origin in the mathematical branch of graph theory, for the study of networks. Complex network science mainly focuses on real-life networks that are large and complex. Moreover, this network display properties that are neither random, nor regular [4]. These two factors together were the driving forces for the scientific research of connectomics. Connectome is a central concept in neuroscience. Importantly, the actual interest in connections, which aims at gaining insight in the whole brain coordination and, ultimately, in cognitive processes, relies on the centennial "structure-function" paradigm, i.e., the ubiquitous correspondence between functional and anatomical features. This concept, started by Brodmanns' [5] classification of cortical areas, permits to consider cortical parcells defined by atlases [6] as nodes of a network. Henceon, the new methods to assess functional and anatomical connectivity can be translated into a connectiivity matrix of connection weights between each parcel pair. However, this definition, which will be followed in this work, attains to the macroscale of brain description, which (with some differences from atlas to atlas) addresses order of 100 areas and connectivity nodes. A more general concept of connectome should consider the whole scales from brain areas down to single neurons (see Par. 1.2), which well sizes the enormous challenge to be faced by neurosciences in the decades or centuries to come [7].

1.1.1. Graph theoretical concepts for network analysis

Graph theory is used for modeling, estimation, and simulation of brain networks topology and dynamics. One of the benefits in using network analysis on neuroimaging data is the abstraction from local to global features and the availability of quantitative indexes despite the

complexity of neural networks. Brains show great variability (in size or shape of surface), hence, by masking these differences, network analysis helps in identifying general properties of neural organization. [Van Essen and Drury, 1997]. As previously introduced, a graph is a graphical representation of a connectivity matrix and is composed by a set of nodes (i.e., brain areas or parcels) connected by weighted edges representing the measured functional or structural connectivity.

1.3 Levels of brain networks

Brain connectivity can be described at different levels of scale:

1. *Microscale* brain connectivity is based on neuron theory [8] whereby neurons interconnected by synapses are represented as nodes of the graph and axonal projections and synapses are represented by edges.
2. *Mesoscale* networks are connecting neuronal populations organized into networks of columns and minicolumns [2].
3. *Macroscale* brain connectivity summarizes white matter connections between cortical areas. Diagrams of large-scale brain network organization consist of limited areas (nodes) that are spatially limited and interconnected by white matter tracts (edges).

This multiscale organization shows that there is no preferred scale for brain network analysis. Since the network is abstractly depicted as a set of nodes and edges, graph theory provides a unified vocabulary for understanding brain network topology, independent of scale or measuring technique. However, it is crucial to understand how nodes and edges are defined at each scale, because the definition is very reliant on the measurement methods used [7].

1.4 Classes of brain connectivity

There are three classes of brain connectivity:

1. *Structural connectivity* refers to the anatomical associations between neural elements or brain regions — the brain's actual (axonal and dendritic) wiring. Hence, white matter tracts between pairs of brain regions are forming structural network upon which neural activity unfolds. It is measured using different techniques (depending on the scale of measurement): electron microscopy, tract tracing and diffusion MRI.
2. *Functional connectivity* may occur between anatomically distant regions, and it correspond to magnitudes of temporal correlations. These correlations can be quantified by measures in time, frequency, or wavelet domain. Dynamical coupling between cortical regions brings up functional networks related to cognition, action, perception and spontaneous activity in resting state [4], [7], [9]. Workflow for construction of structural and functional networks is displayed in Figure 1.
3. *Effective connectivity* is defined at neuronal level and represents causal influences, direct or indirect, that neural elements exert over each other's activity [10]. Effective connectivity necessarily involves a model of how the measured signal is generated by neuronal dynamics. Effective connectivity may be inferred experimentally by perturbations or observation of the temporal ordering of neural events [11].

Functional and effective connectivity are time-dependent. Rapid changes occur in interactions between brain regions, depicting involvement of subsets of brain regions and pathways during different cognitive tasks and attention states. These three types of brain connectivity are interrelated by the fact that structural connectivity is a major constraint on the possible patterns of functional and effective connectivity that can arise in a network. On the other hand, functional relations are contributing to the shaping of the anatomical substrate [11]. The importance of directionality in structural, functional, and efficient connectivity cannot be overstated. Since every axon has a source and a target, structural connectivity should be directed. However, some techniques used for structural connectivity measurements, such as diffusion MRI, are limited in terms of directionality. Functional connectivity can be either directed or undirected, which is dependent to the method used for measurements. However, so far, fMRI is largely undirected since it relies on the correlation or the mutual information of the hemodynamic responses, which are in the orders of seconds, i.e. orders of magnitude slower than the underlying neural processes, known to be directed. For the above reasons, this work will be limited to undirected graphs.

Conversely, effective connectivity, out of our scope, is always directed, because it is based on a model of causal interactions among neural systems. These distinct types of connectivity are used to define edges that can be directed, weighted and dynamic, representing the heterogeneity in the type of connection that is made [7].

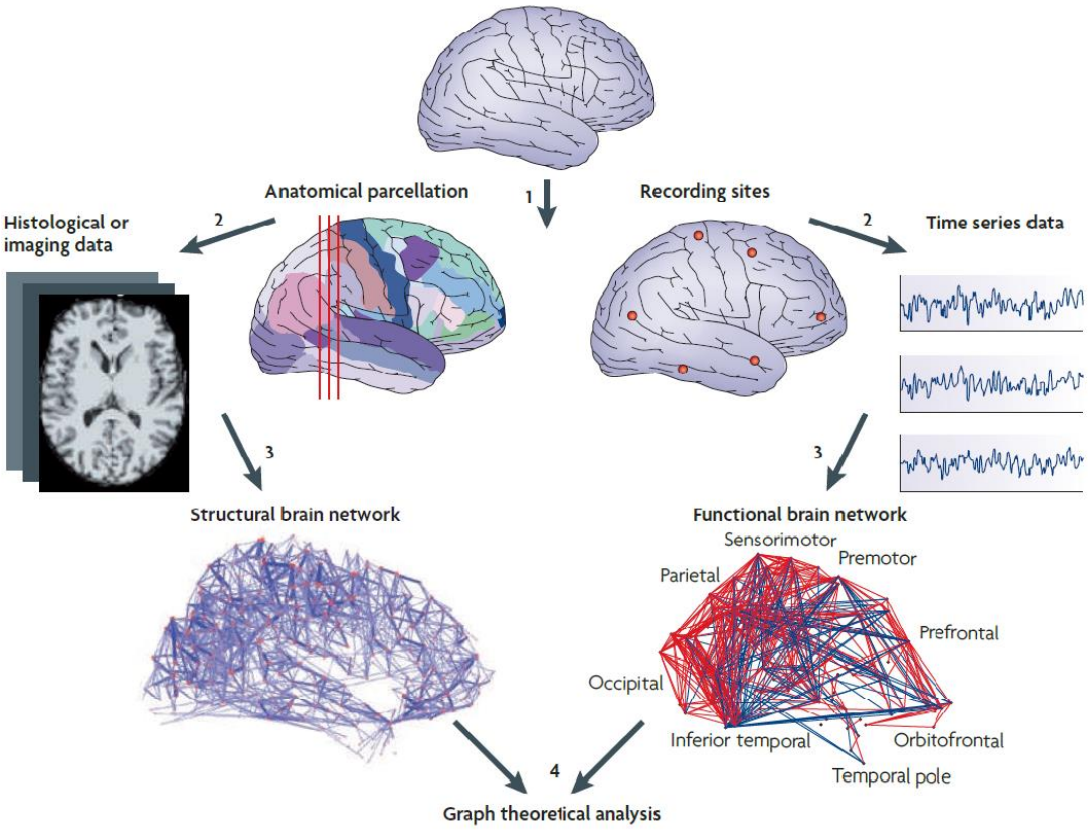


Figure 1. Workflow in construction of structural and functional networks.

1) define the network nodes; 2) estimate a continuous measure of association between nodes; 3) generate connectivity matrix by compiling all pairwise associations between the nodes; 4) calculate the network parameters of interest in current graphical model [3]

1.5 Defining nodes at macroscale

For networks constructed at microscale, neurons are represented by nodes, while axonal projections and synaptic junctions are represented by edges. However, there is no widely accepted means for defining nodes for connectomics analysis at macroscale. Indeed, the general structure-function paradigm may be conjugated in different degrees of segregation and suffers of many uncertainties, which explains the existence of various parcellation atlases.

A data-driven approach would be that each individual node corresponds to one measurement point. Thus, different nodes can be related to individual voxels in MRI dataset or different electrodes used in EEG, or sensors in MEG. Even though this approach does not require any additional processing, it cannot be guaranteed that measurement points coincide with borders of cell populations that are functionally specialized [7].

Other approaches include anatomical parcellation, where nodes are defined based on a priori anatomical information, however this parcellation results in low resolution and large variations in terms of node size. The problem of node size is solved in random parcellation, where brain is parceled into discrete nodes of similar size, but reliability of this method is in question. Another type of parcellation is functional, which relies on a priori functional information, but it is difficult to be applied on diffusion data [7], [12]. Parcellation schemes can be defined based on diverse criteria, but all of them should maintain key properties. Namely, parceled regions should not overlap, so that one location belongs to one region exclusively. Networks can be compared between each other only if they use the same parcellation scheme [13].

Since imaging methods have limited capacity in delineating homogenous cell populations, methods for defining nodes use heuristics. The heuristic nature of these approaches ultimately reflects the lack of gold standard for brain parcellation at the macroscale. However, gross topological properties such as small-worldness are maintained across different parcellations used, while that may not be the case for other topological parameters [7].

1.6 Brain connectivity in disease

Intra-brain connectivity is essential for the attainment and maintenance of life. Genes, proteins, neurons, cell assemblies, and gross brain regions are continually interacting to regulate the brain activities that support sensorimotor and cognitive processing. This delicate system, however, may be disrupted, particularly in the context of neurological and psychiatric diseases,

in which abnormal pathologic factors provoke massive changes in connectivity at all levels of the brain [14]. According to findings from structural and functional network research, network measures are heritable, they alter in the context of normal aging, and show abnormalities in clinical disorders. Thus, it can be concluded that these metrics reflect the brain organization and are hence investigated for their neurobiological importance [3].

Cognitive processes are dependent on interactions among distributed brain regions [15]. Modern imaging techniques have provided efficient way of investigating the functional and anatomical relationships between human brain regions, allowing for a better knowledge of the pathological foundations of neurological illnesses [16]. Different neurodegenerative illnesses are known to cause both abnormal (hyper-connections) and disturbed cerebral connections. It remains unclear how these connections form and coexist throughout illness progression. Damaged connections are frequently thought to be a direct result of neurodegeneration and demyelination, while hyper-connections are thought to represent compensatory mechanisms or spatiotemporal correlation [14].

Having regard to these facts, one application of complex network theory is to develop novel metrics able to quantify disparities between patient groups and suitable control groups [17]. Neuropsychiatric disorders such as schizophrenia and autism, for example, have been linked to a disruption and overall reduction in functional and structural connection [18]. Several studies have found that in individuals with schizophrenia or Alzheimer disease (AD), the characteristics of brain networks generated from fMRI, EEG, or structural MRI data, are disrupted. In an fMRI study of AD, it was noticed that clustering coefficient was significantly lower, at both local and global level, and furthermore able to discriminate AD patients from matching-age controls, providing high sensitivity and specificity. Hence, loss of small-world network properties potentially can provide clinically valuable diagnostic marker [19]. Similarly, an EEG study showed that path length in beta band functional networks is significantly higher in AD [20]. Another MEG investigation of resting-state functional networks corroborated a decline in small-world features in patients with Alzheimer's disease, and showed that this impact is caused by disease-related alterations in highly linked network hubs [21]. Konrad and Eickhoff [22] confirmed that white matter pathology and disrupted anatomical connectivity have both been linked to attention deficit hyperactivity disorder (ADHD). Furthermore, dysfunctional connectivity has been observed during rest and during cognitive tasks. Another study [23] suggests that mild cognitive impairment in Parkinson’s disease is characterized by increase in network modularity and small-world coefficients, as well as alterations in network hub regions. In addition, network disruptions accompany cognitive impairment in Parkinson's disease, represented by a weakening of long-range connections and an increase in local interconnectivity.

1.7 Magnetic resonance imaging

Brain activity can be recorded using different methods, which typically show a trade-off between spatial and temporal resolution, and the scalability and invasiveness of the technique used [7]. Attention will be brought to magnetic resonance imaging basic principles, with subsequent description of variants of this technique used in both structural and functional connectivity study.

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that provides detailed three-dimensional anatomical images. It is frequently used in the detection, diagnosis, and monitoring of diseases. MRI is based on the interaction between a magnetic field and a spin-containing particle. For several reasons, the ^{1}H nucleus, which consists of a single proton, is a natural choice for probing the body using MR techniques [24]. It is the most abundant hydrogen isotope, with a nuclear spin quantum number of $\frac{1}{2}$. Additionally, it has one of the largest responses to a magnetic field discovered in nature. Normal tissues will naturally generate a large MR signal because the body is made up of tissues that contain primary water and fat, both of which contain hydrogen [24]. Due to the nuclear spin, the ^{1}H nucleus rotates, producing a local magnetic field or magnetic moment oriented parallel to the rotation axis. Since the nuclear spin is constant in magnitude and orientation, the magnetic moment associated with will follow the same trend. This is analogous to bar magnet, with north and south pole aligning along rotation axis.

In general, MR measurements are performed on collections of spins that are similar. If arbitrary volume of tissue containing hydrogen atoms (protons) is considered, spin vectors of each proton composing the entire collection are oriented randomly in all directions. In this condition, there is no net magnetization in the tissue. When a strong external magnetic field ($\boldsymbol{B_0}$) is applied, the property of nuclei to have a spin causes them to align to it, producing net magnetization directed as $\boldsymbol{B_0}$. The spin vectors are slightly angled away from the magnetic field axis, but their precession axis is parallel to $\boldsymbol{B_0}$. Larmor frequency is the rotational velocity along the filed direction applied, proportional to the field strength and given by following equation:

$$w_0 = \frac{\gamma B_0}{2\pi}$$

where γ is physical constant called gyromagnetic ratio.

Within the static magnetic field, $\boldsymbol{B_0}$ nuclei with spin can be excited by applying a second radiofrequency (RF) magnetic field $\boldsymbol{B_1}$ that is perpendicular to $\boldsymbol{B_0}$. Short pulses of RF energy are normally used, each lasting a few microseconds. These pulses are called excitation pulses,

spanning narrow interval of frequencies [25]. When a proton is irradiated with energy of correct frequency w_0, it is excited from lower-energy state (spin-up) to a higher energy state (spin-down).

Without entering the details of MR signal generation and MRI reconstruction, it is important to recall that contrast in MR images is caused not only by differences in water density, but also by differences in basic nuclear magnetic processes known as *relaxation*. This is the process by which the protons release the energy absorbed from the RF pulse. Relaxation is characterized by different rates or *relaxation times*. T_1, T_2 and T_2^* are the three relaxation periods that are of main concern in MRI. These are the time constants for the magnetization to return to its equilibrium position aligned along the scanners' static magnetic field when it is disturbed (T_1 relaxation), as well as the time constants for signal loss once the magnetization has been sampled (T_2 and T_2^* relaxation) [26].]. Namely the blood oxygen level dependent (BOLD) signal component exploited in fMRI is due to T_2^* changes related to an increased oxygenation and washout of venous blood from activated gray matter areas. See a brief illustration of functional connectivity methods in Par. 1.7.2.

DTI imaging, conversely, is based on the diffusion weighted contrast, which is enhanced by bipolar gradient pulses. Thanks to the directionality of the gradient field a specific direction in space at a time can be sensed, thus inferring the anisotropic structure of the white matter and ultimately estimating the pathways of virtual neural fibers, called "streamlines". This technique is explained in detail in Par. 1.7.1.3, since the data considered in this thesis refer to DTI, tractography, and the consequent structural connectivity.

1.7.1 Methods for studying structural connectivity at the macroscale

Primary method for studying structural connectivity of the brain is diffusion MRI. First, attention is pointed out to the diffusion process itself, then to Diffusion Weighted MRI, Diffusion Tensor Imaging and Tractography.

1.7.1.1 Diffusion

Diffusion, also known as “Brownian motion”, refers to the constant random movement of molecules due to heat which leads to irreversible transport of matter. The molecular motion is related to thermal kinetic energy of the molecules, proportional to the temperature:

$$E_{kin} = \frac{3}{2k_bT}$$

where $k_b = 1.3810^{-23}\frac{K}{J}$ is the Boltzman constant. Higher E_{kin} is corresponding to faster molecular motion. This motion is completely stochastic. The relative amount of diffusion, or the mobility of the molecules is expressed by physical constant called the diffusion coefficient *D*, which is frequently given in units of $\frac{mm^2}{s}$. This constant describes mean displacement *d*, the stochastic motion within time τ (diffusion time). Thus, *d* is the mean diameter of the three-dimensional extension of molecular motion (or more exactly, the standard deviation of the position ($\sqrt{\langle x^2 \rangle}$) [27]:

$$d = \sqrt{\langle x^2 \rangle} = \sqrt{6D\tau}.$$

The molecular weight, intramolecular interactions (viscosity), and temperature are all factors that affect diffusion. In a homogeneous liquid such as water, the diffusion coefficient is the same in every direction, called *isotropic* [28]. On the other hand, cellular microstructure of tissue influences mobility of the molecules that diffuse across it, by creating number of strongly oriented barriers and compartments. In this case diffusion depends on direction and it is called *anisotropic*. This type of diffusion is of great interest because it carries much information about underlying anatomical architecture of living tissues. As a matter of fact, living tissues are composed of organized structures that guide water movement, which hence does not occur freely.

Diffusion Weighted Imagining ignores this complexity by reducing the diffusion coefficient to a single average value, the apparent diffusion coefficient. DTI, on the other hand, is a technique for measuring water molecule diffusion in distinct directions in each pixel of an MR image (and it involves DWI in at least six noncolinear directions) [28].

1.7.1.2 Diffusion Weighted Magnetic Resonance Imaging

Diffusion weighted imaging (DWI) uses magnetic resonance imaging to depict the diffusivity of water molecules in a defined voxel by applying motion-probing gradients. This method allows for a qualitative and quantitative evaluation of the diffusion process properties, as well as microstructural insights into the tissues state. Furthermore, DWI is unique method that gives different contrast mechanisms compared to conventional T_1 and T_2 weighted MR images. Signal intensity at diffusion weighted imaging is inversely proportional to the magnitude of water diffusion, and this is in turn influenced by cellular structure of the tissue where diffusion occurs [29]. Thus, relation can be made between the signal and histological structure of the tissue.

During the early 1990s diffusion MR imaging of the brain was starting to be used in clinical neuroradiology. Its utility was discovered immediately in evaluation of suspected ischemic stroke. Since then, major leaps forward in terms of technology of diffusion imaging had greatly improved quality of output images and enabled this technique to be used across various clinical applications. Main clinical application of diffusion weighted imaging remains in studying stroke, but it is also used in research of multiple sclerosis, dyslexia, schizophrenia and traumatic injuries. Changes in diffusion, which are an indicator of alterations in cellular homeostasis in acute ischemic stroke, can have a significant impact on treatment decisions and therapeutic outcomes for stroke victims if detected early. Various pathological processes affect tissue structural organization, causing destruction or regeneration of membranous components, as well as changes in cellularity. These changes induce variation in tissue permeability and osmolarity. Together, all these effects have impact on the extent of diffusion of water molecules which is captured by DWI. Anisotropic diffusion, on the other hand, is associated with fiber orientation explored with DTI, and can be used to better understand brain connectivity [30], [31].

A diffusion-weighted pulse sequence is made by adding a pair of diffusion-sensitizing gradients (motion-probing gradients) to a T_2-weighted spin echo sequence. These gradients are applied along the same directional axis before and after 180 degrees refocusing pulse. This method was proposed by Stejskal and Tanner [32] with the goal of quantifying the diffusion coefficient. Most diffusion weighting sequences still use this technique, either in its original or slightly modified form. The result of employing diffusion gradients is attenuation of transverse magnetization (i.e. the received MR signal intensity, depending of molecular motion magnitude). Attenuation effect is given by relation between precession frequency (Larmor) w_0, of a spin, and the external magnetic field $\boldsymbol{B}_0$: $w_0 = \gamma B_0$. A gradient pulse applied in *x* direction, with amplitude g_D causes spatially linear magnetic field $\Delta B(x) = g_D x$. Corresponding spatially dependent variation of Larmor frequency is given by:

$$\Delta w_0(x) = \gamma \Delta B(x) = \gamma g_D x.$$

After this gradient pulse with duration delta, there is an extra phase angle:

$$\Delta\varphi(x) = \Delta w_0(x)\delta = \gamma g_D x \delta$$

which depends linearly on position x. Thus, spins at different x positions have different additional phase angles after gradient pulse is applied – they are not precessing with the same phase. This process is called *dephasing*, while the opposite process of spins returning into identical precession phase is called *rephasing*. Dephasing and rephasing processing are used in diffusion weighting. By applying first diffusion gradient, spins are dephased and they have additional phase angle $\Delta\varphi(x)$. If spins do not move after applying the first gradient (those spins are called stationary spins), the additional phase angle is reverted by the second diffusion gradient. This means that the signal of transverse magnetization will remain the same, as if there was no diffusion sensitizing applied. On the other hand, if spins move between the two diffusion gradients, then each of them will be equipped with different phase angle after applying the first gradient pulse (dephasing). Hence, this effect cannot be compensated by applying second gradient pulse, since stochastically distributed phase angle remains. This random distribution of phase angles is observed as signal attenuation of magnetic resonance signal and represents a measure of diffusion. Faster molecule diffusion results in higher attenuation and weaker corresponding pixel signal intensity. This means that in regions where diffusion is limited, signal intensity will be greater than in regions where diffusion is rapid [29].

Attenuation is expressed as the ratio of measured signal $S(D, b)$ and the original signal S_0. It depends exponentially on diffusion coefficient and the properties of the diffusion gradients determined by b-value of pulse sequence, called *diffusion weighting*:

$$S(D, b) = S_0 e^{-bD},$$

where b-value can be calculated using properties of diffusion gradients: amplitude g_D, duration δ, and the time interval between two diffusion gradients Δ [27]:

$$b = (\gamma g_D \delta)^2 \left(\Delta - \frac{\delta}{3}\right)$$

By applying diffusion gradients in only one direction, only diffusion motion spanning parallel to this direction can be detected. This can be sufficient in isotropic liquids, however the biological tissues are much more complex and highly anisotropic. In this case diffusion is hindered by cell membranes in perpendicular direction to diffusion motion, which results in decreased apparent diffusion coefficient. In this case diffusion is no longer described with scalar but with *diffusion tensor*.

1.7.1.3 Diffusion Tensor MR Imaging

The emergence of diffusion tensor imaging (DTI) and fiber tractography has created completely new non-invasive path of studying white matter connectivity of human brain. These techniques rapidly brought advancements in understanding of many neurological and psychiatric disorders and have been used clinically for mapping of the white matter tracts before intracranial mass resections. DTI imaging studies of nerves, spinal cord white matter and brain white matter are based on underlying phenomenon that water diffusion in nervous tissues is highly anisotropic [33]. Diffusion anisotropy measurements have been used to assess the structural integrity of white matter in the brain at various ages and in a variety of diseases [28]. Water tends to diffuse along axonal tracts in nervous tissues.

1) Diffusion tensor

A superior method to model diffusion in complex materials is to use the diffusion tensor which describes diffusion of water molecules using Gaussian model. It is proportional to the covariance matrix of three-dimensional Gaussian distribution which models molecular displacement [34]. Diffusion tensor is 3*3 symmetric, positive definite matrix:

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

The three diagonal elements (*Dxx, Dyy, Dzz*) represent diffusion coefficients measured along each of the three main laboratory axes (*x, y*, and *z*). The six off-diagonal terms show the correlation of random motions between each pair of major directions. If the diffusion is isotropic, all off-diagonal elements are zero. Optimal coordinate system for viewing diffusion tensor is based on diffusion ellipsoid. Axis of diffusion ellipsoid are defined by three orthogonal *eigenvectors* and three positive *eigenvalues*. In tensor model it is assumed that a water molecule undergoing diffusion for finite time is constrained to a volume of diffusion ellipsoid [35]. The major eigenvector is called *primary* eigenvector and it points the principal direction of diffusion within a voxel (direction of greatest diffusion). In anisotropic fibrous tissues the major eigenvector also defines the fiber tract axis of the tissue. Additionally, it is called *longitudinal diffusivity* because it specifies the rate of diffusion along the orientation of fibers. Three orthogonal eigenvectors together can be observed as a local fiber coordinate system [34]. Hence, second and third eigenvectors define diffusion along axis transverse to axonal fiber bundles. Three positive eigenvalues, λ_1, λ_2 and λ_3, give the magnitude of diffusion along three axes defined by eigenvectors. When λ_2 and λ_3, are averaged, so called *radial diffusivity* is calculated.

2) DTI – derived metrics

When diffusion tensor is obtained, three eigenvalues are used to calculate a number of diffusion metrics, which are then used to describe the microstructure in a specific voxel and produce grayscale and color maps. These measures are useful for describing the amount of tissue organization and for locating voxels likely to contain a single white matter tract (without crossing or fanning) [34].

a) Mean diffusivity and trace

The mean diffusivity represents average of diffusion coefficient along three principal axes. Sum of diffusion coefficients along principal directions (diagonal matrix elements) is called trace. When trace is divided by three, mean diffusivity, or so-called apparent diffusion coefficient, is obtained [27]. The trace and mean diffusivity are related to total amount of diffusion in voxel, which is related to amount of water in extracellular space. Lower values mean low diffusivity. The trace is useful in early detection of stroke because it leads to cytotoxic edema (cellular swelling) which restricts diffusion.

b) Fractional anisotropy

The diffusion anisotropy describes how much diffusion deviates from isotropic diffusion (0 corresponds to isotropic and 1 to anisotropic diffusion). The fractional anisotropy is most widely used anisotropy measure and can be thought of as the difference of the tensor ellipsoid's shape from the shape of a perfect sphere. It is calculated as a normalized variance of the eigenvalues:

$$FA = \sqrt{\frac{1}{2}} \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

c) Mode

Complementary to FA, measure called mode discriminates between linear and planar anisotropy. Mode relates to skewness of the eigenvalues.

$$\frac{(-\lambda_1 - \lambda_2 + 2\lambda_3)(2\lambda_1 - \lambda_2 - \lambda_3)(-\lambda_1 + 2\lambda_2 - \lambda_3)}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - \lambda_1\lambda_2 - \lambda_1\lambda_3 - \lambda_2\lambda_3)^{\frac{2}{3}}}$$

Other commonly used diffusion indices derived from decomposition of diffusion tensor are listed in Table 1.

Index	Acronym	Formula	Range	Unit
Mean diffusivity	MD	$\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}$	Not defined	$\frac{m^2}{s}$
Fractional anisotropy	FA	$\sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$	$0 \div 1$	Dimensionless
Relative anisotropy	RA	$\sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1^2 - \lambda_2^2)^2 + (\lambda_2^2 - \lambda_3^2)^2 + (\lambda_3^2 - \lambda_1^2)^2}}{\lambda_1 + \lambda_2 + \lambda_3}$	$0 \div 1$	Dimensionless
Volume ratio	VR	$\frac{\lambda_1\lambda_2\lambda_3}{\left(\frac{\lambda_1 + \lambda_2 + \lambda_3}{3}\right)^3}$	$0 \div 1$	Dimensionless
Axial diffusivity	AD	λ_1	Not defined	$\frac{m^2}{s}$
Radial diffusivity	RA	$\frac{\lambda_2 + \lambda_3}{2}$	Not defined	$\frac{m^2}{s}$

Table 1. Diffusion indices estimated from the decomposition of the diffusion tensor

d) Color maps

Another type of representation of eigenvector field is by mapping to colors. The principal eigenvector λ_1 is a unit vector consisting of *x*, *y* and *z* components, that satisfy criterion $x^2 + y^2 + z^2 = 1$, each scaled within 0 – 1 range. These x, y, and z components can be shown independently as grayscale maps, as illustrated in Figure 2A, with each vector component

allocated 256 (8 bit) grayscale steps. These vector component images can be multiplied by an anisotropy map such as FA to hide low-anisotropy regions that are presumed to have no dominating fibers, resulting in cleaner and more informative images. As illustrated in Figure 2B, these FA-weighted vector-component images can be shown as grayscale images. It is, however, difficult to understand fiber orientation information using three independent images. A 24-bit color presentation using RGB (8-bit each for red, green, and blue) channels has been proposed to better visualize fiber orientations in a single image. The x, y, and z component pictures are allocated to three RGB primary colors and concatenated to form a single color-coded map displayed in Figure 2C. The color scheme which is used for representing the orientation of the major eigenvectors is: blue is superior-inferior, red is left-right, green is anterior-posterior, while colors brightness is controlled by FA [34].

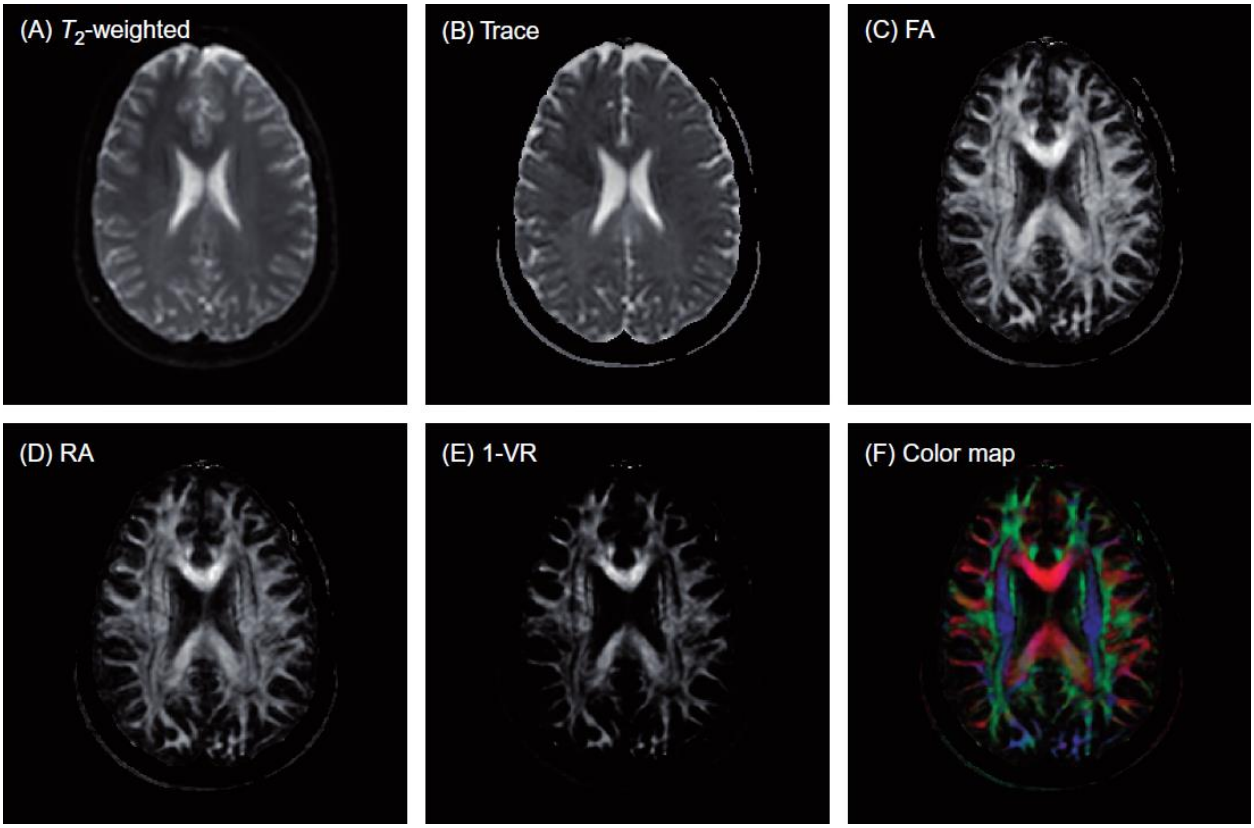


Figure 2. Various image contrasts obtained from DTI. (A) reference T2 -weighted image (least-diffusion-weighted image), (B) trace map, (C) fractional anisotropy map, (D) relative anisotropy map, (E) volume ratio map, and (F) color coded orientation map.

1.7.1.4 DTI fiber tracking

Magnetic resonance diffusion tractography is the main in-vivo technique used for identifying white matter pathways in the living human brain. These pathways have primary role for transferring information between regions of the brain, and therefore they are main point in understanding brain functioning in health and disease. Up to date, this technique has not only been used to infer valuable clues about brain development and functioning, but it has also been used in investigating white matter connectivity in neurological disorders, but also for surgical planning [36]. Diffusion tractography brought major technical progress, since it is the only technique capable of investigating white matter architecture in the central nervous system non-invasively. Previously, invasive tract tracing was the major method used to map structural connections in mesoscale connectomics. In a typical tract tracing experiment, a fluorescent dye or other tracer molecule is injected into a specific area of the brain. After the tracer has had enough time to cover the complete extent of the labeled neurons, the animal is euthanized, the brain is dissected, and the sites of tracer uptake are mapped. Hence, there are obvious ethical limits in performing tract tracing studies in humans. Although measurements made by DTI fiber tracking are prone to errors, the non-invasive nature and ease of measurements provide answers to scientific and clinical questions that cannot be answered by any other means [37].

DTI fiber tracking is used with purpose of determination intravoxel connectivity based on the anisotropic diffusion of water [38]. Diffusion of water molecules is obstructed in greater extent across axis shared by many axons aligned along it. This means that the diffusion is preferred in orientations that are matching with orientations of axonal fibers. In reconstructing fiber bundles, or drawing inferences about axonal connectivity, tractography algorithms try to find paths through the data field along which diffusion is least hindered [37]. In each voxel, the diffusion tensor provides two pieces of information: the magnitude of diffusion isotropy and the direction of maximal diffusion. This information is used by tractography algorithms for tracking entire white matter pathways by maintaining fiber continuity from one voxel to another.

However, diffusion tensor provides only simplification of diffusion of water molecules inside brain, and it can represent only one major fiber direction in voxel. Hence, DTI cannot be reliable in significant part of white matter voxels which contain fiber bundles that are oriented in various directions. Cases of complex architectures, where fiber tracts intersect or branch, affect the output of DTI fiber tractography in delineating pathways and may contribute to generation of spurious tracks. All decisions in streamline tractography are made locally, which leads to further accumulation of errors which induces false positive and false negative connections [34]. Another confound is limited spatial resolution. Voxel sizes in human diffusion MRI investigations are generally $2 * 2 * 2\ mm^3$, with state-of-the-art technology and software allowing for close to $1\ mm$ isotropic resolution [37].
However, there are approximately 10^5 axons per voxel at this resolution, and the chances that a single pathway is occupying a voxel without any contamination from other fiber paths, gray

matter or CSF is minimal [39]. Furthermore, low spatial resolution leads to partial volume artifacts, where MR signal from the voxel will be equal to weighted average of the signals of different tissues in that voxel. Main signal contamination is caused by CSF since it is characterized by isotropic diffusion (FA is around 0). Thus, CSF signal can dominate the tissue signal even when its fraction in voxel is low. This results in voxels containing both CSF and tissue have lower FA than voxels having only tissue. Consequently, these voxels are frequently overlooked during tractography because their FA value falls below the tracking threshold, making it difficult to recover fiber tracts close to CSF [40].

Thus, it should be noted that none of tractography methods are able to truly reconstruct nerve fibers or fiber bundles. There is no perfect method, and it is unlikely that it is possible, although it is a topic of active research. Algorithms are computing pathways or trajectories trough the data, such that a great portion of nerve fibers run reasonably in parallel to them. These pathways are used to derive direction and scope of fiber bundles, or as region of interest (ROI) which is used to infer quantitative metrics. Tractography algorithms are sufficiently effective in reconstructing major tracts, while reconstruction of lateral fibers of these pathways remains challenging topic [39]. Since DTI provides information at relatively low resolution, DTI fiber tracking is often combined with functional and/or higher resolution anatomic information to delineate specific pathways [41].

DTI fiber tracking algorithms can be divided into deterministic and probabilistic methods. Deterministic algorithms are constrained to choose one or more fiber orientations per voxel. On the other hand, probabilistic algorithms consider uncertainty in fiber orientations. Generally, they propagate a large number (typically > 1000) of streamlines iteratively, selecting the direction of propagation from a collection of fiber orientations in which there is a quantifiable trust.

1) Deterministic methods and limitations

Deterministic algorithms take one estimation of fiber orientation at each point of imaging space, hence they follow the primary eigenvector from voxel to voxel. They are initiated from user-defined voxel. When end of one voxel is reached, angular relationships between eigenvectors in voxel and in all its neighboring voxels is explored, and connection is formed by choosing the ones with smallest angles [42]. Most commonly these algorithms are guided by following the so-called *streamlines* trough the eigenvector field which provides valuable insights to connection between brain regions. However, the outputs of the algorithm will be either 0 or 1 depending on the existence of pathway between two data points. Thus, it can be said that deterministic algorithms study connections rather than connectivity.
The object of stopping criteria in deterministic tractography is to prevent streamline propagation trough the regions of high uncertainty and therefore to prevent false positives. If the angle

between directions of eigenvectors is less than *a priori* defined value, two voxels are connected, and the process is repeated to continue streamline trough white matter in brain image. On the other hand, if the angle is greater than the predefined threshold, the streamline is terminated in that point. Another constraint used to end the streamline propagation is fractional anisotropy value. This ending criterium ensures that the streamlines continue in regions where direction of maximal diffusion is well defined. Since the uncertainty in the principal diffusion direction decreases as the anisotropy in a voxel increases, streamlines will only continue if the anisotropy in each voxel exceeds an *a priori* specified threshold value [43]. Thus, results of tractography are dependent on several factors that are pre-set by the operator: angular and anisotropy thresholds, and the region chosen for seeding the tractography. Noise, patient movement, and imaging artefacts produce uncertainty in the orientation of three eigenvectors and together they negatively affect streamline fiber tracking.

Deterministic approach of streamline tracking using diffusion MRI data does not take into consideration either systematic or stochastic errors in predicting orientation of fiber. Systematic errors are due to complex fiber geometries with different orientations, such as crossing, twisting, bending, and kissing fibers cannot be adequately captured by using tensor model and the principal eigenvector orientation. Regarding the stochastic errors, it should not be overlooked that diffusion MRI is noisy technique. Susceptibility gradients, head motion, and eddy currents cause artifacts [36]. The noise in individually acquired images will further propagate trough the analysis and bring uncertainty in fiber orientation estimation, distortions in pathways and biased diffusion indices, which together lead to erroneous tractography [36]. Uncertainty in fiber orientation is not randomly distributed throughout the brain, but it will be correlated with specific area if the brain where tracking is done [42].

2) Probabilistic methods and limitations

Instead of terminating the path in regions and along trajectories where errors are probable, the goal of probabilistic tractography is to integrate the expected uncertainty into the tracking algorithm. Thus, differently from deterministic methods probabilistic algorithms are generating set of multiple pathways or streamlines that are passing through the seed-point. As a result, probabilistic tractography techniques can map fiber paths in areas where deterministic tractography methods will have to stop. The probability density function of fiber orientation can be estimated with an empiric function based on FA, a Bayesian model, or bootstrap statistics [38]. The accuracy of probabilistic methods is limited by the method used for constructing the probability density function and information contained in diffusion tensor. These algorithms are prone to systematic errors, both in acquisition of data and analysis, in the same measure as the deterministic ones. The most convenient way of interpreting resulting maps is that they are representing connectivity, and absolute numbers indicate connection strength. However, the majority of probabilistic tractography techniques use a voxel-wise streamline count as their connection metric. This metric is strongly correlated with successfulness of streamline

propagation, which is in turn determined by preformed acquisition, processing technique and fiber geometry. Connectivity index thus does not represent the actual strength of connection, but the confidence in a connection.
It worth noting than the estimation uncertainty of fiber orientation is influenced by SNR in the data and the total number of measurements used in estimation. The higher is SNR and number of measurements, the lower is the uncertainty. However, SNR linearly depends on voxel volume, i.e. reducing voxel volume from $2mm$ to 1.5 reduces the SNR by more than half. Thus, outputs of these methods directly depend on the parameters of the MRI experiment. Another contributor to accuracy decrease of probabilistic method is the length and the shape of the tract [44]. Streamline count is dependent on the pathway length, shape and size of the reference region and shape of the tract in question [39].

3) Streamlines and fractional anisotropy as a measure of connection strength

When streamlines are seeded and tracked between all pairs of nodes it is possible to derive several different measures of connectivity strength. The simplest measure is a count of number of streamlines that connect a pair of regions. The number of axonal fibers connecting two areas is assumed to be proportional to the number of connecting streamlines that can be reconstructed from the data in this calculation [7]. Thus, streamlines are not equivalent to axons; they are rather abstraction of the diffusion model and the tractography algorithm applied. Capacity to track a streamline can be influenced by numerous factors which are not essentially representative of axonal connectivity. This means that "raw" connection densities based on streamline are not biologically relevant [45]. However, streamline count is most used edge-weighing scheme and normalization by total brain volume, as well as streamline count with density and length correction, are examples of variants of this procedure.

Other most common measure of diffusion strength is fractional anisotropy (FA) measured along the connecting streamlines that is a more physiologically relevant measure of strength. Rather than an abstraction of trajectory count, this form of weight relies on tract integrity and myelination [46]. It quantifies degree of diffusion constriction along all directions within a voxel. Fractional anisotropy is high in areas where axon fibers are coherently oriented and densely packed since axons are acting as strong constrain to diffusion. Differently, fractional anisotropy is low in the areas where fibers are sparse, or they are organized in geometry that is at level of complexity that cannot be captured by diffusion model. Since FA depends on the direction in which it is measured, among the factors modulating the diffusion anisotropy there are axon density, distribution of axon diameter and fiber myelination. FA is not particularly specific measure if considering that regions with low anisotropy deep in white matter can often contain the largest axon number passing through them. This stems from the fact that orientation of fibers is incoherent and thus diffusion displacements are no longer preferentially hindered along a single resulting with anisotropy of the voxel-averaged displacements being low. These altogether can lead to paradoxical observations such us increased anisotropy that can be seen in types of pathologies where destruction of particular fiber orientation occurred [44].

Constructed weighted matrix can either be directly analyzed or it can be thresholded so that connections are maintained between nodes if they exceed minimum predefined number of streamlines or a minimum probability [13]. In functional connectivity, time series of brain activity in different voxels is derived and the correlation between the time series of different voxels is represented in form of correlation matrix where values are spanning the interval between -1 and 1. In dealing with negative weights, most common approaches include either filtering the negative weights which are considered as unsynchronized activations or using the absolute value.

Measure used for quantifying connectivity strength will also affect set of values present in connectivity matrix and thus impact interpretation of edge weights [7]. Both streamline count and fractional anisotropy are subject to a variety of factors that are not always related to biologically meaningful anatomical connectivity strength characteristics, such as axonal number, density, caliber, or myelination. Changes in the signal-to-noise of different voxels induced by head motion, physiological noise, or imaging artifacts, as well as complex fiber geometries within a voxel, all complicate the understanding of relation weights derived from diffusion MRI [7].

1.7.2 Methods for studying functional connectivity at the macroscale

Functional connectivity at the macroscale is typically measured using functional MRI, EEG or MEG.

1.7.2.1 Functional Magnetic Resonance Imaging

Functional magnetic resonance imaging (fMRI) is delicate and noninvasive tool that employs MRI to image patterns of activation in the working human brain caused by changes in neural metabolism. FMRI methods can be made sensitive to the changes in blood perfusion in a region, volume of blood, or blood oxygenation that follows the neuronal activity [26]. The BOLD (Blood Oxygenation Level Dependent) contrast is the most common fMRI contrast. It does not directly measure neuronal activity, but rather the metabolic demands (measured in terms of oxygen consumption) of active neurons. The origin of the BOLD signal changes relies on magnetic properties of hemoglobin that exists in two different magnetic states resulting in diverse local magnetic fields. Deoxygenated hemoglobin (*deoxyHB*) is slightly paramagnetic relative to brain tissue, while hemoglobin that carries oxygen (*oxyHb*) is isomagnetic. Thus, oxygenated arterial blood containing vessels cause no distortion to magnetic field in surrounding tissue, while deoxygenated-blood carrying vessels distort magnetic field in their close proximity. This leads to

destructive interference with signal within tissue voxel, resulting in shorter T_2^* relaxation time. The effect of T_2^* shortening is most noticeable near larger veins and is accentuated by using GRE sequences with TEs (echo times) close to T_2^*(at fields 3T or below).

During cerebral activation more oxygenated blood is supplied to activated region than it is required for its immediate metabolic needs. The relative concentration of deoxy hemoglobin in activated areas will decrease because of this "overshoot" of oxygenated blood. T_2^* shortening effects will be reduced, resulting in an increase in the BOLD signal in activated regions. The change in the MR signal caused by instantaneous neuronal activity is called hemodynamic response function. The BOLD effect is proportional to the amount of deoxy hemoglobin in the blood, which ranges from less than 2 percent in arterial blood to more than 40 percent in venous blood.

The regional BOLD response produced by a short peripheral stimulus is known as the Hemodynamic Response Function is the (HRF). The HRF typically exhibits a slight initial dip, followed by a tall peak, and then a variable post-stimulus undershoot (Figure 3). Initial increase in deoxy hemoglobin leads to a decrease in BOLD signal, displayed by initial dip. Next, an increase in blood flow dilutes the concentration of deoxy-Hb and shifts the balance in favor of oxy-Hb. This results in a peak in BOLD response 4 to 6 seconds after activation. After attaining its peak, the amplitude of the BOLD signal falls below the baseline level. This undershoot after stimulation is caused by a combination of decreased blood flow and increased blood volume.

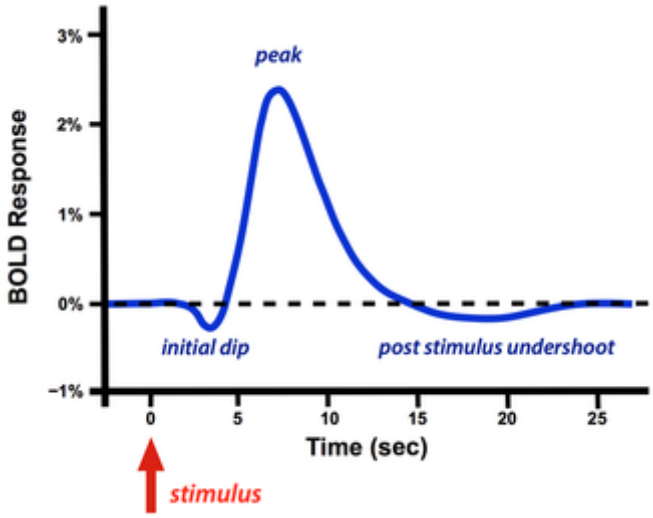


Figure 3. BOLD signal

Functional MRI offers a higher spatial resolution than EEG and MEG (~mm), although with diminished temporal resolution (~s). Coarse temporal resolution restricts measurable bandwidth to 0.001 to 0.5Hz.

Functional connectivity studies differ from task related studies since they are usually directed to detect spontaneous correlations of activities among brain areas. For this reason, functional connectivity is most frequently studied in “resting-state” condition [47], with the subject asked to stay still in the MRI scanner, without thinking and also not falling asleep, for a few minutes.

Importantly, resting state data can also offer data-driven methods for a functional parcellation of the set of coordinated regions in the brain, which may involve also dinstant regions, thus encompassing the structure-function paradigm. These region sets are called resting state networks [un paio di REF] and can be also addressed in terms of structural connectivity, thus fusing functional and structural features.

1.7.2.2 Other methods

Other methods for studying functional connectivity include Electroencephalography (EEG) and Magnetoenchephalography (MEG). EEG is one of the most well-known and commonly used techniques for studying the brain's electrical function. At a given position in the extracellular medium, electric current contributions from all active cellular processes within a volume of brain tissue superimpose and form a potential, measured by electrodes placed on the scalp or subdural, with respect to a reference potential. MEG employs an array of superconducting quantum interference devices (SQUIDs) to measure the weak magnetic fields that are generated by neuronal currents. Both EEG and MEG have high temporal resolution, enabling neural activation to be sampled on millisecond scales, which is comparable to neuronal signaling pace. Conversely, these methods suffer from poor spatial resolution predisposed by the fact that each sensor measures cumulative activity across large neurons populations. MEG and EEG have millisecond temporal resolution, but poor spatial resolution (~1cm) and show electro-magnetic activity at wide range of frequencies (1-100Hz).

2 Methods

The present chapter introduces commonly used network models. Next, connectivity matrix is defined, which is the starting point for brain network analysis. After the construction of the matrix, it is possible to evaluate a large number of metrics and topological characteristics of the network, which are potentially relevant indicators of health and disease (as already mentioned in the Chapter 1.5.) In this section, the most widespread metrics, with particular attention to the ones considered in the present work are described. Subsequently, widely used thresholding methods and new approaches to assess the role of weak connections are reported.

2.1 Network models

Although real-world connectivity networks are mainly compared by quantifying commonly used network measures, they can be classified according to the underlying structure that brings valuable insights to network spatial and topological organization. These simulated network models can be used in comparison and classification:

2.1.1 Regular networks

Regular networks, such as trees and lattices, are strictly homogeneous and very rarely representative of real-world networks, although important in the study of many human artifacts and also studied by graph theory methods. Accordingly, such networks are out of the scope of this thesis.

2.1.2 Random graphs

In past, random graphs were mainly used as a model for describing topology of various networks. They are generated by assigning connections between any pair of nodes in the network with uniform probability. If sufficiently large network is observed, connection density of this network is equivalent to this probability. Random graphs are distinguished by nodes of approximately same degree which results in Poisson degree distribution, also called exponential degree distribution. This configuration results to be particularly suitable both for computation and to derive theoretical results. Although they provide oversimplified approximations of connectivity structure in complex systems, still they are addressed as a benchmark to detect the organization hidden in complex networks.

2.1.3 Small world networks

The small-world phenomenon is an example of a surprisingly universal macroscopic behavior in complex systems, and it was firstly described for large social networks. Watts and Strogatz [48] investigated how path length and clustering behaved in a basic generative model (henceforth the WS model). The WS model iteratively rewires a binary lattice network of N nodes (Figure 4), each connected to the same number of nearest neighbors via edges of similar weight (unity). This rewiring process is performed by randomly deleting an existing edge between nodes *i* and *j,* and subsequently replacing it by a new edge between node *i* and any node $k \neq j$. They discovered that as the probability of random rewiring increased incrementally from zero, so that the initial lattice was progressively "organized", despite the random procedure.. Sparsely rewired networks showed both strong clustering (similar to a lattice) and short path length (like a random graph). These algorithmically created graphs were called small-world networks. According to findings from the paper [49] there is sufficient evidence to indicate that small-worldness is a practically universal feature of neurological systems. Moreover, brains appear to be only one of a wide "universality class" of small-world networks that includes many other non-neural or non-biological complex systems. Clearly, brain organization is far from being randomly organized and is built on important hubs, which are currently referred as the "rich club" (see Par. 2.2.7) . Nonetheless, the WS model has a core position in simulation to assess the main features of organized networks, avoiding any a-priori constraint.

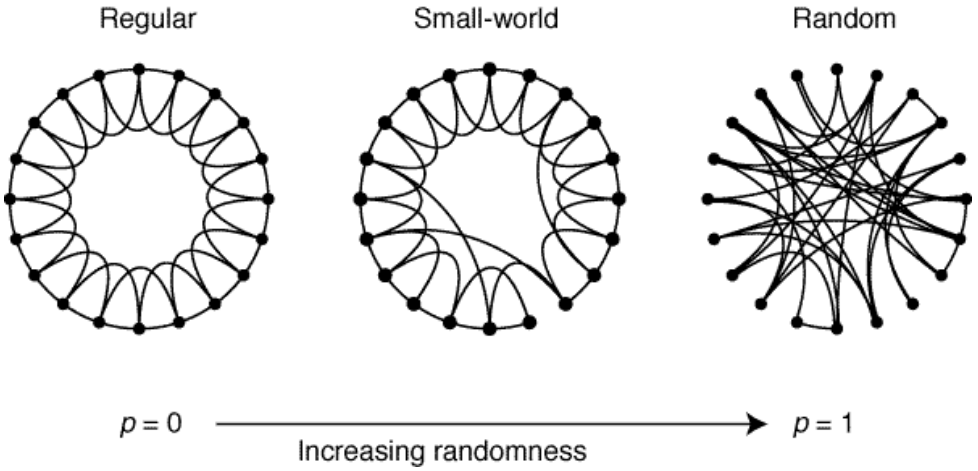


Figure 4. The random rewiring procedure of the Watts-Strogatz model

Later, this behavior was spotted in metabolic, information, genetic, signaling, computation and ecological networks. In neuroscience, small-world topology has been reported across multiple species and scales, from both structural and functional MRI studies of large-scale brain networks [49]. Brain networks' short path lengths may favor integrated information processing over the entire network, while high clustering may favor segregated processing across functionally specialized cliques of nodes. Hence, the small-world organization of the brain showed that apparently opposing tendencies of functional integration and segregation can in fact be joined in a single architecture. It should be noted that the presence of the small-world topology provides only a limited amount of information regarding network design. For example, two small-world networks may exhibit highly distinct patterns of connectedness [50]. A convenient way for summing small-wordness property is by the ratio of the clustering coefficient to the path length, but with prior normalization of these metrics by comparing them to the values obtained for random graph with equal number of nodes and edges (Par 2.2.6) [51]. However, this measure is not sufficient for certain determination whether a network is small-world or not because it can be influenced by a number of factors such as the density of the connectivity matrix [4].

2.1.4 Scale-free networks

At about the same time of small-world networks discovery, another type of complex networks was introduced: scale-free networks. This architecture is distinguished by a very broad and nonhomogeneous degree distribution, which results in the presence of nodes with far higher degrees than would be expected in a random, regular, or small-world network. These networks are found to exhibit degree distribution that follows power law, according to which the likelihood of discovering a node with a degree twice as great as an arbitrary number reduces by a constant factor. Many networks have power-law degree distributions, which indicate a "scale-free" organization [52]. The term "scale-free" refers to the fact that a power-law distribution has no characteristic scale—"zooming in" on any segment of the distribution has no effect on its shape and assigning a characteristic scale for the degree of network nodes is thus pointless [50]. Power law degree distribution can be generated by so called preferential attachment (Figure 5), which means that new nodes preferentially connect to nodes that already have high degree, so “the rich nodes get richer”. This process results in existence of at least few highly connected nodes, also called hubs, and many other less connected nodes. As the result, distribution that emerges from this model is identified by fat-tail.

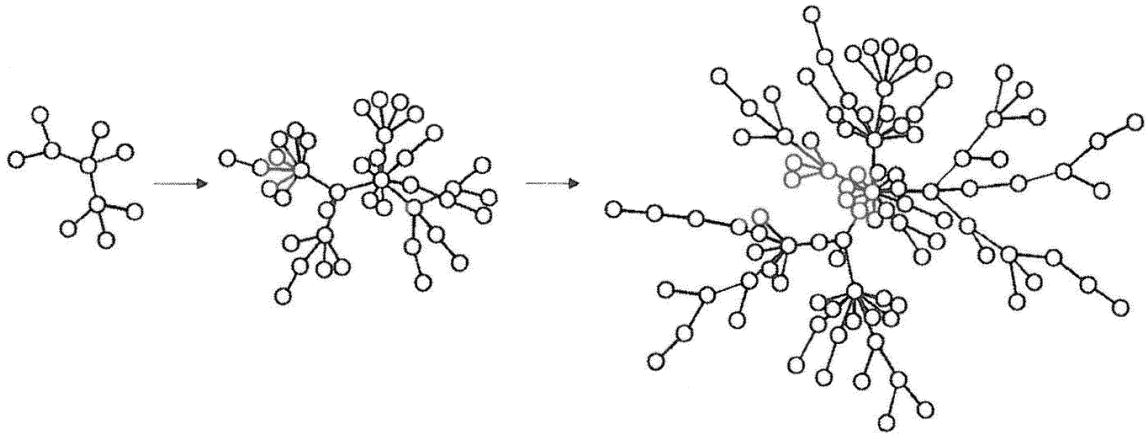


Figure 5. Illustration of early stages of network growth by preferential attachment. Nodes are added one by one, and a single new edge links the new node to an existing node chosen with a probability based on node degree.

Small- world and scale- free properties are compatible, but they are not equivalent to each other, meaning that network can be small-world while not being scale-free. Structural networks can display scale-free structure, although with low maximum degrees, whereas functional brain networks always exhibit scale-free distribution as well as small-world characteristics [11]. As said in the Introduction, a full comprehension or the brain aims at the ambitious goal of integrating the whole connectivity scales, which can be hypothesized to show scale free (alias, fractal) features, as in many natural structures. However, our starting point given by neuroimages limits this study to the macroscale of brain areas and scale free networks are thus out of our scope.

2.2 Connectivity matrix and measures of connectivity

The connectivity matrix (or adjacency matrix) is the simplest mathematical representation of network connectivity [53]. This matrix is used to create a graph-based representation of network, in a way that each row and column in this matrix represents a node. The matrix element at the intersection of the *i-th* row and the *j-th* column stores information about the link between nodes *i* and *j,* thus edge between two nodes. Edges can be differentiated based on weight and directionality. In undirected edges information flows in both directions – every connection is bi-directional (i.e., the connectivity matrix is symmetric), whereas in directed it is restricted in going from one node to the other, but not in the opposite direction.

In this matrix form, connectivity between all nodes of the network is compactly described [7]. If number of the nodes in the network is equal to N, connectivity matrix is N^2 sized.

Diagonal elements of the matrix are sometimes viewed as indicating the connectedness of each node with itself, but they may also be used to express a node intrinsic feature. However, many graph theoretical metrics used to analyze brain networks neglect the matrix diagonal, by forcing its values to zero Conversely, the off-diagonal elements correspond to obtained pairwise connectivity, with values of these elements being strongly dependent on the method used for connectivity estimation. If the upper and lower triangles (above and beneath main diagonal) of connectivity matrix are identical, matrix is symmetric, and it depicts undirected graph of a network. Networks unidirectionality implies that we cannot draw any assumptions about probable information flow directions or the causal impact that one neuronal element may have on the activity of another. This is possible in graphs that are directed, with directionality in the matrix representation being encoded by asymmetric elements in connectivity matrix [7].

Since each axonal projection initiates from a cell body and terminates at one or more synapses, brain anatomical networks are naturally directed. However, mapping the directionality requires application of invasive methods addressing functional or structural features of single fibers. Noninvasive methods, such as diffusion fMRI or DTI, do not provide directional information, as previously explained.

Variations in connectedness between neural elements is described trough connectivity weights. The range of weights observed and the way they should be interpreted is determined by the method used for connectivity estimation. Furthermore, sensitivity, noise, artifacts and resolution limitation also affect weights values. Unweighted or binary networks provide simplification of weighted case, where edges can take either values of zero or one, depending on the presence of the link between two nodes [53]. Most of the past studies, and correspondingly networks metrics developed, are made for binary analysis. Although it has been the most widespread and reliable approach for the first-generation tract-tracing studies, binary graph models are considered far

from optimal strategy too. Nonetheless, evaluating the binary topology of a connection matrix may frequently provide useful insights into network organization [7]. Common workflow for simplification of network data is displayed in Figure 6.

Tract tracing data, based on the number of detected streamlines, show values of weights span several orders of magnitude [54], highlighting a great variability that is completely discarded in binary analysis. The weighted edges carry relevant information about strength of connection between any pair of nodes in the network. In structural networks, weights can correspond to size, density, or coherence of anatomical tracts, probability that a node can be reached from another node (in probabilistic tracking), or the number of fibers between the brain regions (in streamline count of deterministic tracking).). Besides, indexes based on the quality of the tract, such as fractional anisotropy (FA), are compressed in small ranges, with a lower cutoff determined by the tractography reconstruction rules (e.g., a streamline is stopped, when FA falls below 0,2). Conversely, weights can represent some statistical measure of association in functional and effective networks, such as magnitude of correlation or causal interactions [3], [7], [13]. Depending on the level of brain network which is studied, nodes and edges are defined. When connectivity matrix is obtained, graph theoretical measures can be computed.

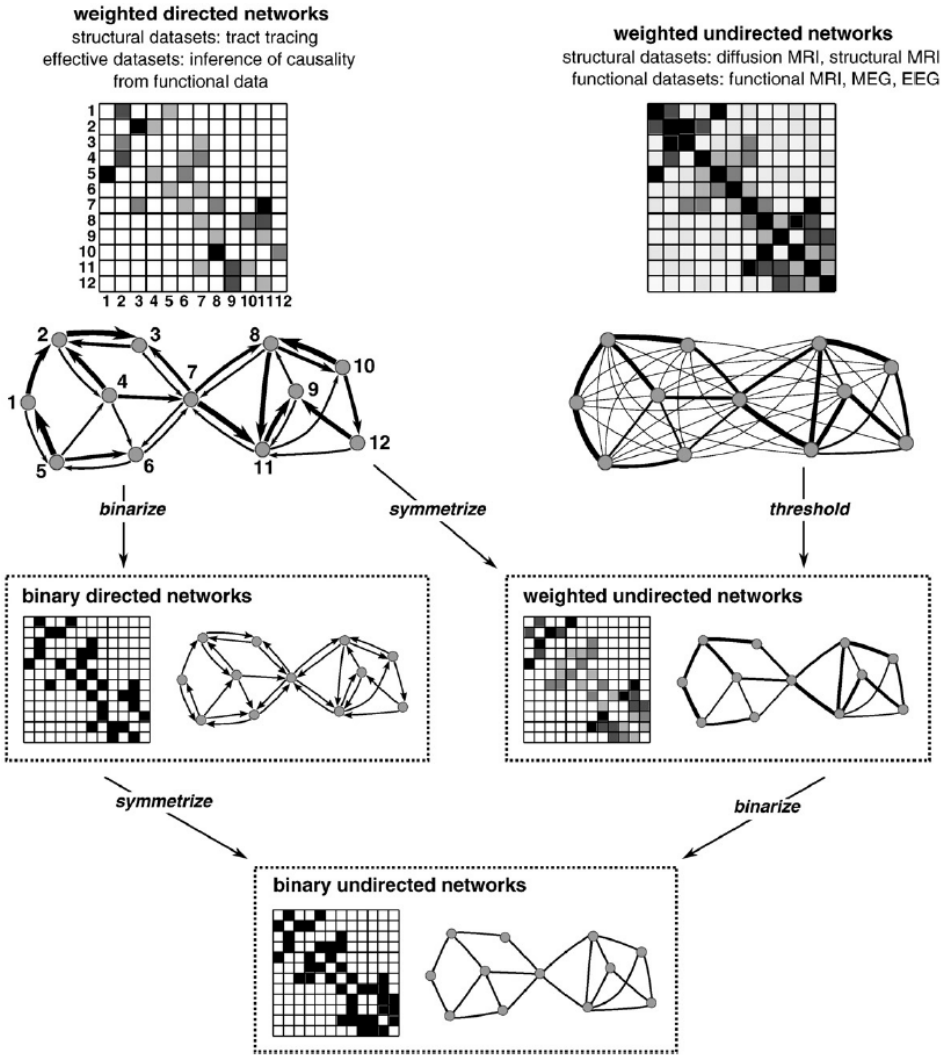


Figure 6. Networks are typically represented using connectivity matrices, with rows and columns representing nodes and matrix entries representing edges. To make analysis easier, networks are frequently converted to a sparse binary undirected form by thresholding, binarizing and symmetrizing.

2.2.1 Degree and Strength

Networks may be classified at several levels. Local measurements focus on individual component attributes, whereas global measures at macroscale look at the network as a whole. Although global metrics can give valuable summaries of network features, they are not sufficient in inferring if the effects are diffused across the brain or limited to a subset of nodes and/or edges [7].

Components of a network, starting at the local scale, are nodes and edges. The most fundamental network measure, inextricably linked to most other measures, is node degree. This measure represents number of connections that link one node to the rest of the network. The node degree k_i of a node *i* in binary graph is defined as:

$$k_i = \sum_{i \neq j} A_{ij}$$

Definition for node degree in undirected binary network is equal to summing all the nonzero elements through the rows or the columns of the binary adjacency matrix (*A*) since it is symmetric for undirected network. The mean degree in an undirected network is the average of all network degrees, and it is most typically employed as a measure of density or the overall "wiring cost" of the network [4].

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i$$

In undirected networks, every connection between pair of nodes is bi-directional. On the other hand, in directed networks, distinction is made between incoming (afferent) and outgoing connections (efferent), resulting in two diverse types of node degree: *in-degree* and *out-degree.* Together, all network nodes' degrees combine to generate a degree distribution, which is a simple depiction of network connectedness and a key indicator of network development and resilience[4], [53]. The degree distribution in most real-world networks is heterogeneous. Numerous nodes have a limited number of links, while a smaller number of nodes are included in many interconnections, which are referred to as *network hubs* [7]. Network hubs connect groups of regions, known as network communities, that are densely intrinsically connected and highly coupled within the community, but have sparse connections between members of other communities [50]. Therefore, network hubs are important in facilitating integration throughout the network.

Regarding weighted networks, measure that is analogous to degree is the node connection strength, s_i:

$$s_i = \sum_{j=1}^{N} A'_{ij} = \sum_{j=1}^{N} A_{ij} w(e_{ij})$$

In binary networks, edge can assume a value of either zero or one. On the other hand, edges in weighted networks can assume a range of different values (expressed by *w* in the formula) which captures variations in the strength of connectivity between pairs of neural elements. Likewise degree distribution, the node connection strength distribution provides a simple representation of connectivity in a weighted network [7], [53].

2.2.2 Clustering coefficient

Local measurements can also relate to the general neighborhood of a node. All nodes that directly project to or receive projections from a node are referred to as node neighbors. To analyze local clustering, the connection between neighbors is utilized [13]. The local clustering coefficient is calculated as the ratio between the number of edges between nodes' nearest neighbors and the total number of possible connections between neighbors [13]. Thus, local clustering coefficient can be thought as likelihood of finding a connection between any two neighbors of node *i*. This is equivalent to computing the proportion of closed triangles that are attached to node *i*, relative to the total number of closed triangles,

$$C_l(i) = \frac{2t_i}{k_i(k_i - 1)},$$

where k_i is the degree of the node *i* and t_i is number of closed triangles that are connected to *i.* Numerator of the fraction is doubled because an undirected network is considered.

On the other hand, one of the most commonly used global measures is based on average connectivity of nodes' and it is called global clustering coefficient. To quantify the clustering of whole network, local clustering coefficients are averaged:

$$C_g = \frac{1}{N} \sum_{i \in N} C_l(i) = \frac{1}{N} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)}$$

The clustering coefficient is a measure of functional segregation [17]. This is a property that associates the clustering coefficient with specialized processing (e.g., sensory input analysis, such as visual and auditory) where nearby nodes work together to achieve complex tasks [55]. Therefore, high clustering of the neighboring nodes ultimately allows efficient communication and complex task processing [53]. The values of both the local and global clustering coefficients are in the range of zero to one. A value of zero implies no clustering, whereas a value of one implies full connectedness between node neighbors, or a completely linked network in the case of the global clustering coefficient [7].

To be a component of a triangle, the degree of each node must be more than two, implying that each node must have at least two neighbors. Clustering coefficient is 0 if node degree is less than two [56]. Nodes with no connections (isolated nodes) or with only one link (leaf nodes) are unlikely to be found in structural connectivity, but such nodes might occur for functional and effective connectivity [13]. Differently from random networks, complex networks have high clustering which is associated with high local efficiency of information transferring [17]. It has been proven that the clustering coefficient provide a clinically useful diagnostic marker in various pathologies such as schizophrenia, Alzheimers' disease, Parkinsons' disease [19], [22].

There are different purposes for extending the concept of unweighted clustering coefficient described above to weighted clustering coefficient, reported in the literature [57]–[59].

In this work, focus was on definition of weighted clustering coefficient given in paper [58]. In this definition generalization from unweighted to weighted clustering coefficient is developed on the basis of the idea of subgraph intensity. Intensity $I(g)$ of the subgraph *g* with vertices v_g and links l_g is derived as the geometric mean of its weights,

$$I\,(g) = \left(\prod_{(i,j)\in l_g} w_{ij} \right)^{\frac{1}{|l_g|}}$$

Triangles are one of the most fundamental nontrivial motifs, and as previously stated, they play a crucial role in determining the clustering coefficient as one of the basic parameters of network characterization. As triangles are one type of subgraphs, definition of intensity can be used to give weighted clustering coefficient by replacing the number of triangles t_i in the in local clustering coefficient equation with the sum of triangle intensities as:

$$C_{i,O} = \frac{2}{k_i(k_i-1)} \sum_{j,k} \left(w_{ij} w_{jk} w_{ki}\right)^{\frac{1}{3}}.$$

Edge weights are normalized by maximal weight in the network, $w_{ij} = w_{ij} / \max\left(w_{ij}\right)$, and contribution of each triangle depends on all its edge weights. Triangles in which edge weight is

maximal will contribute unity to the sum, while the ones with negligible weight will have negligible contribution.

2.2.3 Network density

Another global measure, edge density (connectance or connectivity) is the proportion of nonzero elements in the connectivity matrix compared to the total number of connections that can be formed in the network. In undirected network, the overall number of potential connections is given by *N(N-1)/2*, with *N* being number of nodes in the network and *N(N-1)* the number of off-diagonal elements in the adjacency matrix [7]. The *N* diagonal elements are obviously not included. Hence, edge density in undirected network can be calculated as:

$$D = \frac{e}{N^2 - \frac{N}{2}}$$

where *e* is number of edges. Density can vary in interval between zero and one, where former indicates no connections present, while latter means that all potential edges are present, or the network is fully connected.

2.2.4 Path length

The edge density is a preliminary indicator of network connectivity. However, depending on the network topology, the time needed to travel from one node to another, may vary significantly. This property is depicted by the measure called characteristic path length.
The capacity of the brain to quickly incorporate specialized information from dispersed brain areas is known as functional integration. The concept of path is commonly used to assess how well brain areas interact and how easy it is to transport information through the network [60]. Hence, path length is an inverse measure of network integration.

Furthermore, it is a key metric, representing number of edges that must be traversed in order to go from node *i* to any other node *j* of the network and hence offers a potential pathway for information flow. Paths in functional networks are statistical association sequences that do not necessarily correspond to information flow on anatomical connections [4].
When path lengths over all network nodes is averaged, resulting number represents characteristic path length [49], which is given by:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij}$$

where d_{ij} is the distance between any given pair of nodes *i* and *j*.

The length of a binary path is equal to the number of links in the path, while the length of a weighted path is equal to the sum of individual link lengths. Link lengths are typically calculated as reciprocal of weights, since large weights typically represent strong connections and close proximity in brain networks [4].

2.2.5 Local and Global Efficiency

Local efficiency depicts the integration between nodes' neighbors. This measure can also be observed as generalized clustering coefficient, taking into account only direct links between neighbors of a node. Local efficiency of the *i*-th node, $E_{loc}(i)$, is defined in the paper [61] such that,

$$E_{loc}(i) = \frac{1}{N_{Gi}(N_{Gi}-1)} \sum_{j,h \in G_i} \frac{1}{l_{jh}}$$

where Gi represents the subgraph containing all nodes that are the *i*-th nodes' immediate neighbors. In other words, the local efficiency of node *i* is defined in relation to the subgraph that contains all of *i*'s neighbors after removing *i* and its incident edges.

Alternative measure of network integration is global efficiency. The reciprocal of the harmonic mean of the networks' path lengths (or the mean of all pairwise efficiencies), is representing the network's global efficiency [17],

$$E_{glob} = \frac{1}{L'} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{l_{ij}}.$$

Differently from path length, the global efficiency may be meaningfully used for characterizing disconnected graphs. Since paths between disconnected nodes are considered as infinite lengths, this value corresponds to efficiency equal to zero.

2.2.6 Small-worldness metric

Summarized properties of small-world networks are described by the small-worldness coefficient. This metric compares the average clustering coefficient and path length values of the network of interest relative to clustering coefficient and path length values in random network characterized by same number of nodes and edges.

Small-worldness metric is obtained by taking ratio of two ratios:

1. the ratio of average clustering coefficient of the brain network c_g to clustering coefficient of random network c_n;
2. the ratio of average path length of brain network l_g, to the average path length of a random network, l_n:

$$\lambda = \frac{l_g}{l_n}$$

and

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

Then, the metric of small-worldness, S_w is given by ratio:

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

where γ and λ are measures of clustering coefficient and path length normalized by these measures obtained from random network of same size.
Generally, binary networks small worldness shows γ > 1 and λ ~ 1 and they have a high degree of connections among the neighbors of any node, compared to random network, but small average path length is preserved.

2.2.7 Modularity, Hubs, Rich-Clubs and Core-Periphery

As previously mentioned, segregation and integration are important information processing mechanisms in the brain that promote functional specialization and effective global communication. Clustering coefficients and modularity are two related measures that assess topological segregation properties in brain networks (Figure 7A). On the other hand, integration is often assessed by a network's characteristic path length (Figure 7B).
Moreover, moving our focus between the overall graph level and the node level, structures, communities, and families are addressed, which are subnetworks with strong intra-connections. In this perspective, common network metrics are, besides previously mentioned modularity, rich-clubs and core-periphery.

1) Modularity

Many real-world networks' nodes aggregate into tightly connected subgroups known as modules or communities. Nodes inside these modules are more tightly connected to one another than to the rest of the network. The high interconnection of nodes inside a module shows that they perform similar tasks. Weak connectivity between nodes in distinct modules indicates that these functions are separated. Small-world features are common in modular networks. Strong within-module connectivity leads in a high clustering coefficient, whereas a small number of intermodular links is enough to keep the network's characteristic path length low [7]. Index of modularity of the network is defined as the difference between the degree of intra-module connectivity observed in network and that which is expected by chance.

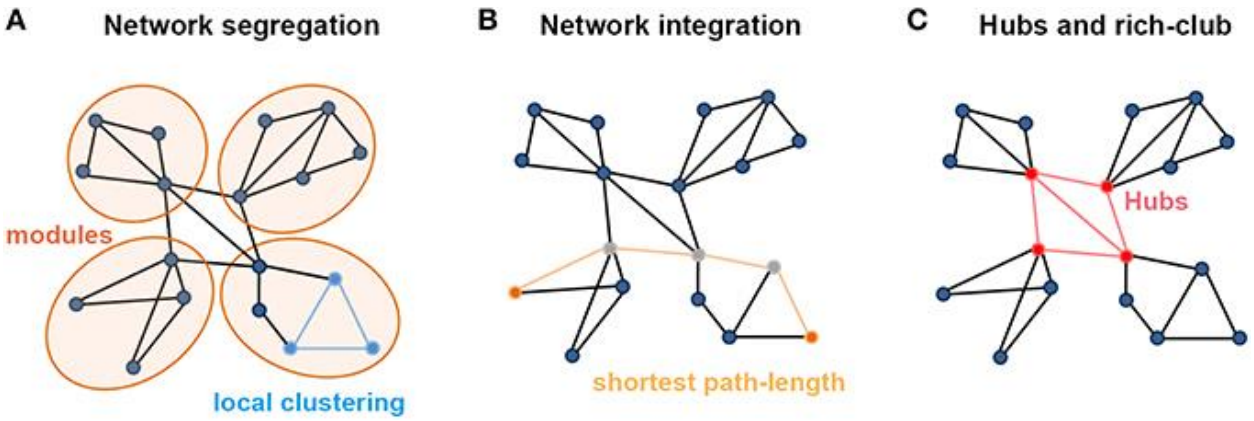


Figure 7. A) Metrics describing the segregation of a network. Local clustering shows nodes' tendency to form triangles; there are four modules in the graph within which connectedness is much higher than between them. B) Metrics describing the integration of the network. The shortest path length is the minimum number of steps needed to travel between two nodes (dots in orange). C) Small set of high degree nodes in central position of the network suggest existence of hub nodes. High-level connectivity (denoted in red) between hubs may suggest the existence of central rich club in the overall network structure [62].

2) Hubs and Rich-Clubs

Using various graph measurements, nodal locations in brain networks that are positioned to make significant contributions to global network communication can be classified as network hubs (Figure 7C).

Degree centrality, which measures the number of connections associated with a given node, is the most basic graph measure used to locate hubs. Another metric is *betweenness centrality*, which is defined as how many of the shortest pathways between all other node pairs in the network flow via a given node, reflecting its capacity to change information.

Importantly, in the brain architecture, these high-degree or high-centrality hubs are densely coupled, forming very well-connected subnetworks that constitute a rich-club structure. These hubs and rich-clubs are discovered to play essential roles in global information transformation despite significantly higher wiring, operating expenses, and vulnerability [62].

3) Core-periphery score

A core-periphery structure is defined by network being divided in two subsets of nodes. Core nodes are located in the networks' topological center, and they are highly connected to one another. On the other hand, peripheral nodes are at least moderately connected to core nodes, but only sparsely connected to one another (Figure 8) [63]. The ideal structure is composed by core that is completely connected, and the periphery that is fully disconnected.

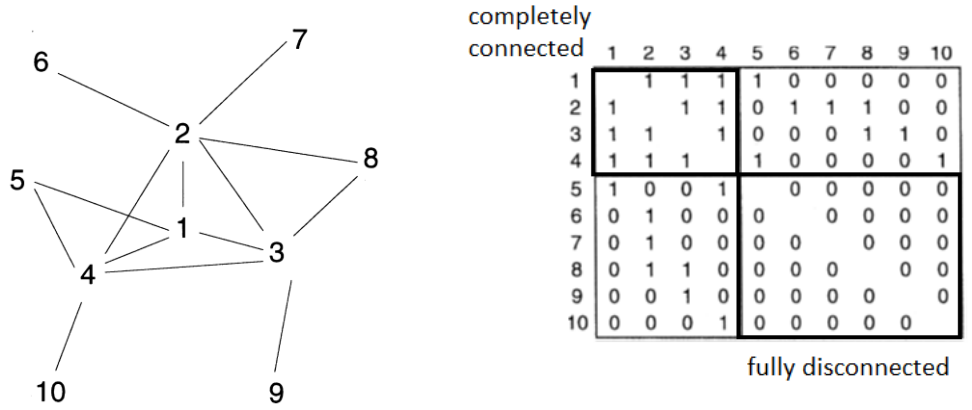


Figure 8. A network with a core-periphery structure (left) and corresponding adjacency matrix (right) [63]

All network metrics described in present chapter are listed in the Table 2.

Index	Mathematical expression	Definition
Average degree	$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i$	*Average of all degrees (number of connections that link one node to the rest of the network) gives a measure of "wiring cost".*
Strength	$s_i = \sum_{j=1}^{N} A'_{ij} = \sum_{j=1}^{N} A_{ij} w(e_{ij})$	*Analogue to degree in weighted networks (where edges can assume a range of different values w).*
Global clustering coefficient	$C_g = \frac{1}{N}\sum_{i \in N} C_l(i) = \frac{1}{N}\sum_{i \in N} \frac{2t_i}{k_i(k_i-1)}$	*Quantifies average connectivity of nodes' neighborhood.*
Weighted clustering coefficient	$C_{i,O} = \frac{2}{k_i(k_i-1)}\sum_{j,k} \left(w_{ij} w_{jk} w_{ki}\right)^{\frac{1}{3}}.$	*Generalization of clustering coefficient in weighted case.*
Density	$D = \frac{e}{N^2 - \frac{N}{2}}$	*Proportion of nonzero elements in connectivity matrix compared to the total number of connections that can be formed in the network.*
Characteristic path length	$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij}$	*Average of all path lengths (number of edges that must be transverse to reach node any node j starting from the node i) of the network.*

Local efficiency	$E_{loc}(i) = \frac{1}{N_{Gi}(N_{Gi}-1)} \sum_{j,h \in G_i} \frac{1}{l_{jh}}$	*Quantifies integration between nodes' neighbors (connected by direct links).*
Global efficiency	$E_{glob} = \frac{1}{L'} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{l_{ij}}.$	*Quantifies the capability of concurrently exchanging information via shortest paths.*
Modularity	$Q = \sum_{u \in M} \left[e_{uu} - \sum_{v \in M} e_{uv} \right]^2$	*The degree to which the network may be subdivided into such clearly delineated and nonoverlapping groups.*
Degree centrality	$C_D(i) = k_i = \sum_{i \neq j} A_{ij}$	*Simplest measure of centrality equivalent to the node degree.*
Betweenness centrality	$b_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N, h \neq j, \\ h \neq i, j \neq i}} \frac{\rho_{hj(i)}}{\rho_{hj}}$	*It measures the proportion of the shortest paths between all node pairs in the network that pass through a given index node.*
Core-periphery score	$\rho = \sum_{i,j} a_{ij}\delta_{ij}$ $\delta_{ij} = \begin{Bmatrix} 1 \; if \; c_i = core \; or \; c_j = core; \\ 0 \; otherwise \end{Bmatrix}$	*A measure of how well the real structure approximates the ideal (core fully connected and periphery fully disconnected).*

Table 2. Network metrics with mathematical expressions and definitions.

2.3 Thresholding

Since it is commonly considered that the traditional fiber-tracking methods cannot reliably determine the strength of connections in the brain [44], many studies have chosen to ignore connection weights entirely and instead perform graph-theoretical analysis on binary connectomes. Structural connectivity workflow consists of computing the weighted connectome, with each edge depicting the number of streamlines connecting any two regions or nodes, and then binarizing it by employing thresholding by zero (edges with at least one streamline are set to one, others to zero). Clearly, the aim is also to capture the topology of connections, leaving out their quantification. However, if binary connectomes are computed in this manner, weak connections will be indistinguishable from strong connections [64]. According to study [17], using binary graph model is unlikely to be appropriate approach for network analysis of tract tracing data. Considered that connectivity weighs are spanning over six orders of magnitude in mammalian cortical networks, effectively removing the seeming biological diversity in fiber connection strengths can be a serious limitation.

Advancements in probabilistic tractography enabled the computation of dense weighted connectomes which provide more reliability in connection strengths. This progress leaded to preserving weights information and directly employing it in graph theoretical analysis instead of

binarizing it [49], [64]. However, measurement noise and the probabilistic nature of the tracking process, on the other hand, result in an unknown fraction of false white matter connections [65]. As a consequence, network thresholding techniques are commonly employed to eliminate purportedly spurious connections, but it is still unclear how different thresholding tactics affect basic network features. However, different studies suggest various thresholding strategies which will be briefly described in the next section.

2.3.1 Absolute thresholding

One of the strategies is absolute thresholding, where one threshold is applied uniformly across the network with the goal of retaining only connections with weights higher than the predefined one [12]. Setting an absolute threshold, however, can result in varying numbers of network edges across datasets, as well as varied level of network density between control and patient cases, which is relevant for disease analyses.
Many graph metrics have been shown to be affected by network density [66], which could lead to statistical differences in network metrics between patient and control populations. However, these effects should be assigned to underlying different number of links in network rather than disease-related variation in network topology. Appropriately, it has been proposed that this method is less suitable for case-control investigations [47].

2.3.2 Density-based thresholding

To tackle this issue of different number of connections between groups, studies have suggested an alternative approach of using a density-based threshold [67]–[69]. This strategy aims to keep the number of connections constant across all individuals and eliminate the impact of network density on graph metrics computation and comparison across groups.

In this thresholding approach, the strongest percentage of connections in each individual network are chosen and retained, while in binary case surviving connections set to 1 and all other connections are set to 0.

Therefore, this method sets the density, or so called “network cost”, to be fixed across patients and healthy controls. Potential between-group differences in graph metrics (e.g. clustering, path length) are assumed to result from differences in the topological organization of edges and not due to differences in number of edges. Density thresholding has been shown to successfully distinguish density from topological effects [70] and provide more stable network metrics, making it a popular method for network construction and analysis in disease connectome research.

However, in order to reach the preset density in patient group, the inclusion of lower and thus potentially less reliable connections is indisputable. Inclusion of more random edges means creating shortcuts and decreasing overall shortest path length. Considering that patient population has overall lower connectivity, this effect will lead to increased global efficiency, decreased clustering, and topology closer to random network configuration. In contrast, if patients have higher levels of connectivity than healthy controls, this can result in decreased global efficiency and higher local clustering. Thus, proportional thresholding should be cautiously used when there are between-group differences.

This strategy is mostly used in functional connectivity analysis with the goal of producing more sparse matrices. It is less commonly used in the reconstruction and analysis of structural graphs (since the matrix is usually already sparse), the inclusion of false-positive edges may, in theory, influence anatomical network reconstruction and thus introduce influence between-group differences in graph metrics. Summarizing, it is known that topological qualities vary with the number of edges in a network, implying that density-based thresholding should be used. On the other hand, individual or group differences in connection density may themselves transmit useful information, implying that any effect these variations have on network topology is real and should not be erased by density thresholding [7].

2.3.3 Proportional thresholding

Another technique commonly used is proportional thresholding which provides retaining the connections which are present in predefined proportion of subjects [71]. The connections whose prevalence (i.e., detection count) meets or exceeds this group threshold are then subjected to further study [12]. Low group criteria allow the inclusion of existing connections that are difficult to reconstruct, preventing false negatives, whereas high group thresholds require a link to be present in many participants, enhancing the accuracy of reported connections and therefore avoiding false positives. This decision to define a specific group threshold is usually made empirically and, indeed can lead to bias and impair the comparability of group-based results between investigations.
In the study [71] the threshold value has a significant impact on regularly used network metrics, emphasizing the necessity of methodical knowledge in this parameter selection.

2.3.4 Consistency thresholding

Another approach given in [72], [73] called consistency thresholding keeps the connections with stable weights across individuals, assuming that connections with the most inter-subject variability are false. This method is used for constructing group-representative networks which is a single network that represents the connection data of many different subjects.

2.4 Weak connections

With the use of cutting-edge fiber tracking technologies, dense weighted structural connectomes are becoming more abundant, however there is no agreement on how to treat the many weak connections that these connectomes include [64]. On the other hand, several studies suggest that thresholding is not necessary in weighted connectomes [53]. Some other studies have argued for employing raw (unthresholded) matrices with no connections removed, claiming that the existence of weak connections has no effect on topological network features. The authors of paper [64] showed that the elimination of weak-to-moderate strength connections in dense weighted connectomes has no effect on value of various graph metrics computed. By opting not to eliminate weak connections, the computational pipeline is simplified and the requirement to choose an arbitrary threshold is removed. Such arbitrary or heuristic-based decisions make it difficult to compare studies and, worse, might lead to statistical results that are unstable [64].

Another point is, although the lowest weighted network connections (e.g., those involving fewest streamlines) are often false-positives, low weights do not necessarily correspond to implausible connections [65]. Considering the intrinsic uncertainty in the extraction of the weights, weak connections have not been given valuable attention in brain network studies and have frequently been eliminated by the process of thresholding.

Despite the fact that long-distance connections are generally weak, they are very consistent among brains. Also, as an example, imagery and multisensory integration are thought to be enhanced by long-distance connections in the occipital brain [54].

In many recent studies attention is brought to potential utility of weak connections because of the notion that they might be relevant for cognition and behavior. According to study of resting state fMRI data by authors of paper [74], individual variations in the strength of weak functional connections from the lateral prefrontal cortex to areas inside and beyond the frontoparietal network have been found to predict individual variations in fluid intelligence. Another study [75] showed that individual differences in relatively weak, long-distance functional connections at

rest were substantially linked with full scale, verbal, and performance IQ. Furthermore, another study by Basset et al. [76] showed that the significance of weak connections is extending to psychiatric illness, as the highly organized topology of weak functional connections in resting-state fMRI was able to accurately and specifically differentiate people with schizophrenia from healthy controls, while this conclusion could not be derived from analysis of strong connections. Moreover, Individual variations in these weak connections were shown to be strongly associated to variations in cognitive scores and symptomatology. Weighted small-world organization, which acknowledges the relevance of weak connections, plays a vital role in system functions that are particularly important to neural systems, such as coherence, computing, control, and resilience [77].

Aside from the issue of erroneous connections, it is still unclear whether network weighting best reflects the underlying biological connectedness. Diffusion MRI structural networks have yielded a range of weights that reflect various conceptions of link strength. Interregional streamline counts and measures of FA are the most common weighting methods [12], which have also been employed for constructing matrices used in the present thesis work.

3 Protocol

This chapter describes tools used in the present work, together with the protocols employed in data acquisition and analysis. Furthermore, motivates the construction and provides the description of a novel software tool we called SPIDER-Net for studying brain connectivity (currently under development finalization) is provided. SPIDER-Net tool (SNT) was validated on a case study of stroke. This software was employed for visualization of brain connectivity trough construction of “connectograms”. Following software description, an introduction is made to the Borderline Intellectual Functioning study, with the focus on the weak connections in weighted networks derived from Magnetic Resonance Diffusion Tractography.

3.1 Analysis of the tools used and developed

3.1.1 MATLAB

All calculations in the present work were carried out in MATLAB.

Calculation of networks metric is based on Brain Connectivity Toolbox (http://www.brain-connectivity-toolbox.net), a freely available Matlab toolbox containing a collection of complex network measures both in binary and weighted forms. Brain Connectivity Toolbox contains measures that variously detect aspects of functional integration and segregation, quantify the importance of individual brain regions, characterize patterns of local anatomical circuitry, and test resilience of networks to insult.

Furthermore, visualization of connectograms in developed software is based on freely available Matlab toolbox circularGraph (https://github.com/paul-kassebaum-mathworks/circularGraph).

3.1.2 Motif for circular representation of brain connectivity and Circos Software

The portrayal of human connectomics through the use of a circular visualization method is well suited to the investigation of central nervous system architecture. This sort of representation, known as a 'connectogram,' is capable of identifying neuro-connectivity

relationships in an intuitive and elegant manner. This methodology gives an anatomically informed organization of nodal cortical and subcortical architecture, geometric aspects of these architectures, and the degree and attributes of regional connection.

The architecture of cortical networks has primarily been visualized using graph theory representations and symbols. Although they provide precise representations, they are however not always appropriate in the context of human connectomics. Canonical methods employed for node-edge relationship representation does not always provide optimal insight into functional and anatomical brain connections. Such representations have been adapted to the study of cortical structure in a multitude of ways. Still, due to the overwhelming number of intricate interactions whose overarching meanings can easily be missed during a visual inquiry, cortical networks remain difficult to comprehend. In addition to that, connectivity analysis typically includes a certain amount of dimensionality reduction and simplification [1]. The idea of representing neuroanatomical structure and connectivity information graphically in a circular diagram stems from Irimia et al. [1]. In this paper freely available software called Circos, is used for visualizing brain connectivity.

Circos is a Perl-based[1] cross-platform program that uses a circular layout to enable the visualization of connections between pairs of coordinates using various graphical components such as connections and heat maps. While originally it was constructed for the purpose of rendering genomic information, Circos can be efficiently adapted to the analysis of data sets comprising complicated interactions between large numbers of elements [78]. Links are generated between any two parcellations connected by a white matter tract. Connectivity representation is based on links that depict fibers connecting two regions. Additionally, links are color-coded by the average value of fractional anisotropy (FA). This value is normalized to the interval between 0 and 1 and divided to three equal-sized bins of different colors according to the value of FA.

Furthermore, fiber count was encoded as link transparency. Within the bin, the connection with the highest fiber count has been rendered fully bright, while the connection with the lowest fiber count has been colored as transparent as possible, though avoiding making it invisible. Five cortical metrics are represented within the circular structure, each represented by a different heat map: the total gray matter (GM) volume, total area of surface associated with the interface between gray and white mass, mean cortical thickness, mean curvature, and connectivity per unit volume. An example of Circos connectogram is given in Figure 9 together with cortical metrics.

---

[1] Perl is a general-purpose programming language originally developed for text manipulation and now used for a wide range of tasks including system administration, web development, network programming, GUI development, and more.

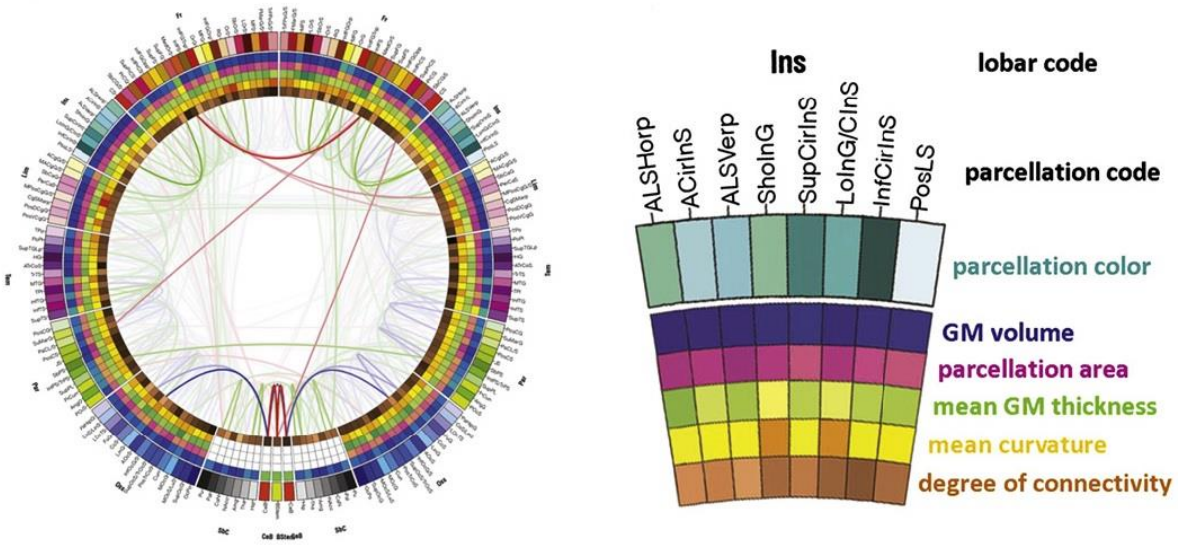


Figure 9. Sample connectogram created for a single subject and legend of the representation of the corresponding cortical metrics

3.1.3 SPIDER-NET Tool (SNT)

SNT was developed as a collaborative project between Politecnico di Milano, Milano (IT) and CADiTeR, MRI Laboratory of Fondazione Don Carlo Gnocchi, Milano (IT). The idea for software development started with the motive of improving visualization of the brain connectivity and extend current functionalities provided by Circos.

The visualization and quantification of brain connectivity networks represents a major issue, since the human brain comprises a large number of nodes and a dense network of links. In particular, it involves order of 100 nodes relevant to the parcellation of cortical areas and sub-cortical structures, and hence order of 100^2 potential edges. This may result in an inability to visualize and analyze connections within a specific part of the brain. Moreover, connections within the region of interest can be hidden by the huge number of network links.

This software provides superior visualization properties by allowing the user to select regions of interest in brain and subsequently extract the corresponding subgraph.

At the initial point of software development, the construction of connectograms was based on the Circos tool (Figure 10). In the latest version, instead of relying on Circos, visualization was modified by using the open Matlab function circularGraph (Figure 11). In this way, computation time is reduced and problems with permissions and constraints are bypassed. Software is made more flexible and user-friendly.

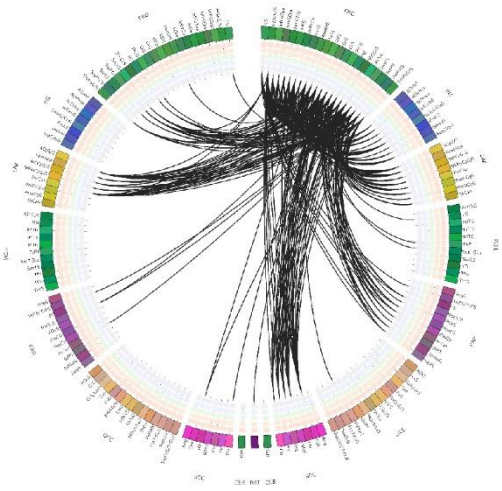


Figure 10. Connectogram visualization using Circos

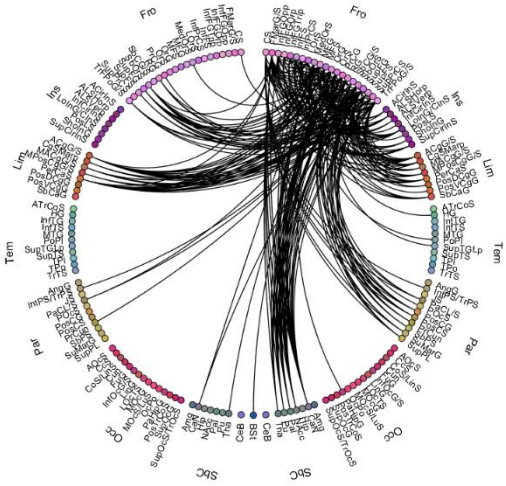


Figure 11. Connectogram visualization using Circular graph

The software enables both exploration of structural and functional connectivity, with possibility, at the same time, to compute large number of binary and weighted network indices and visualize connectivity by plotting connectograms. Starting interface is displayed in the Figure 12. User is supposed to make a selection between structural and functional connectivity and browse atlas according to parcellation used to construct connectivity matrix.

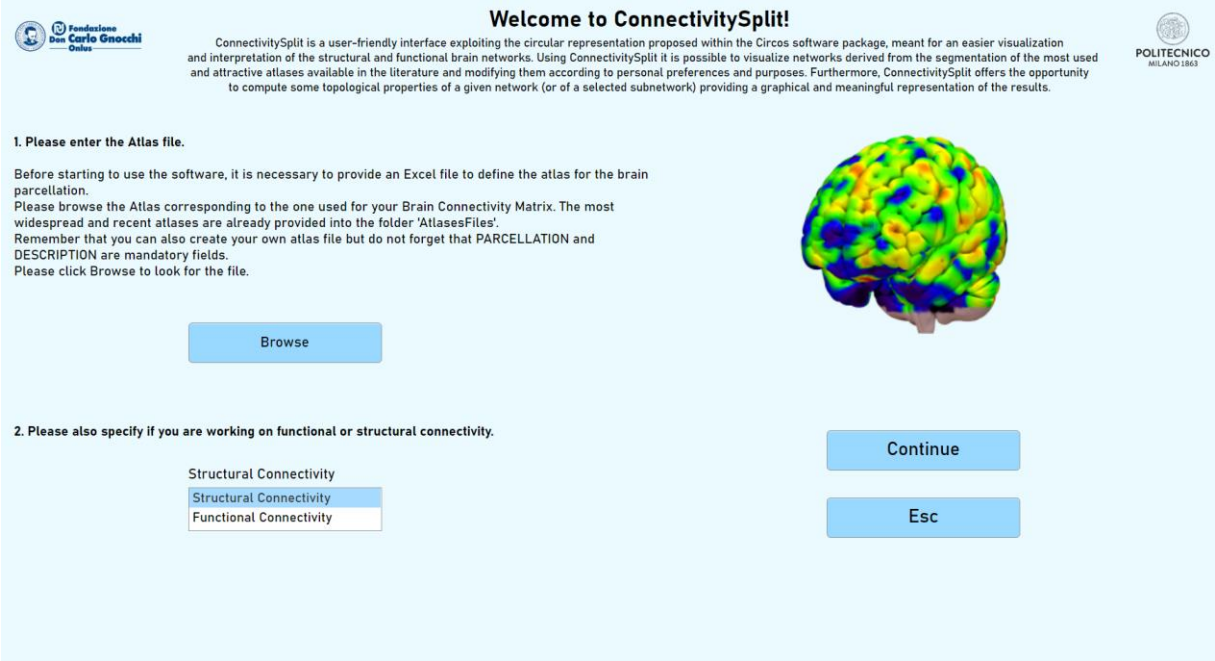


Figure 12. The starting software interface

Next, main software interface is opened (Figure 13). As an input, the user passes a connectivity matrix with corresponding node labels. Furthermore, the tool is adjusted for several anatomical atlases, and thus supports different brain parcellations.

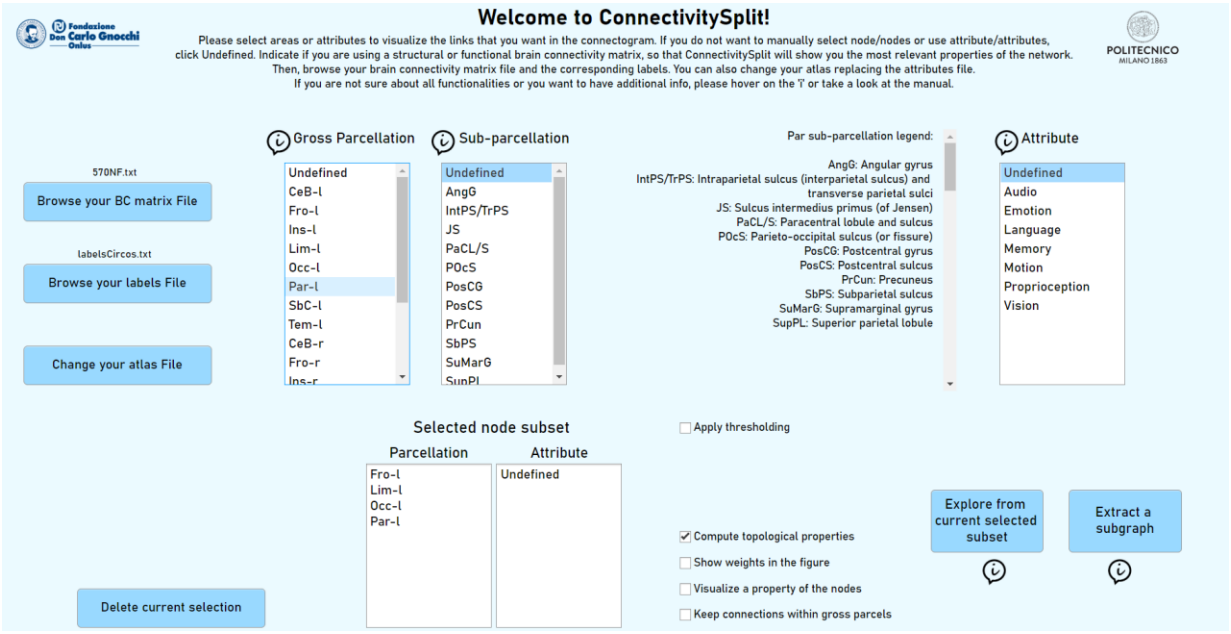


Figure 13. Main software interface

The first possibility offered by the tool is to allow the selection of gross parcels of interest (hemispheric partition), or finer parcels within the specific lobe. In this way, a subgraph is extracted from the connectivity matrix. This enables a more precise and focused connectivity study adjusted to the users' area of interest.

An example of selection is displayed on Figure 14. Here by selecting Parietal lobe in left hemisphere, user can examine list of all finer sub-parcels of the Parietal lobe on the right side, together a dynamic legend for the abbreviations.

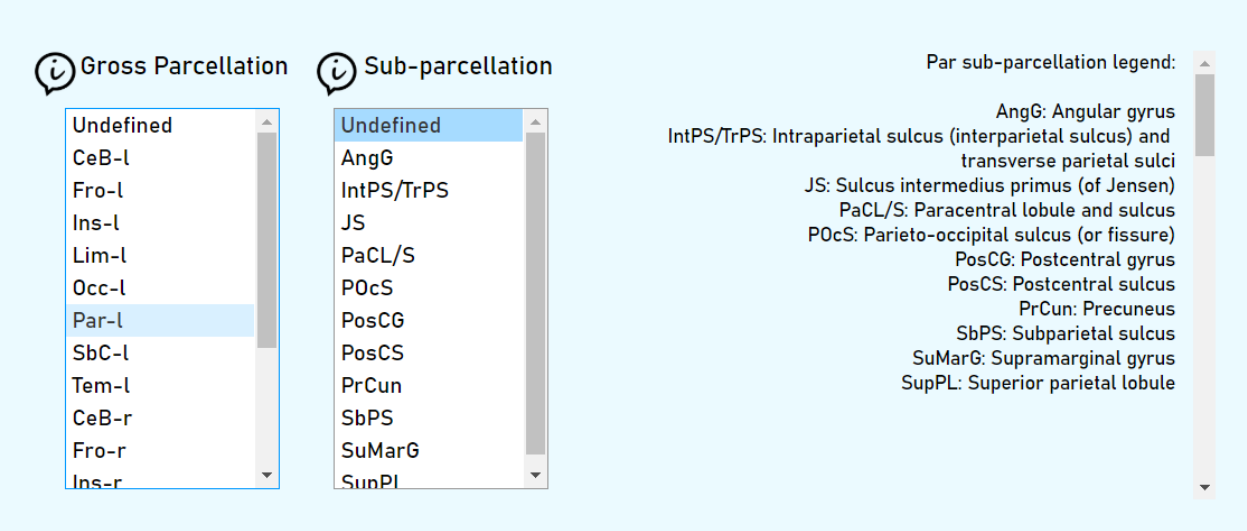


Figure 14. Selection of gross parcellations and sub parcellation

Two of the main functionalities of this software are:

1) *Calculation of topological properties*

Calculation of networks metric in the software is based on Brain Connectivity Toolbox (BCT). Metrics that are computed:

1. Density;
2. Average node degree;
3. Average node strength;
4. Global unweighted clustering coefficient;
5. Global weighted clustering coefficient;
6. Unweighted characteristic path length;

7. Weighted characteristic path length;
8. Unweighted global efficiency;
9. Weighted global efficiency;
10. Modularity;
11. Coreness statistics;
12. Rich-club coefficient.

As an output, together with values of network measures, this option enables the visualization of connectivity matrix of the region of interest (*selected connectivity submatrix*) and distribution of local coefficients (*binary and weighted local efficiency, clustering and degree*) values across the gross-parcels of the brain. An example of output is displayed in Figure 15.

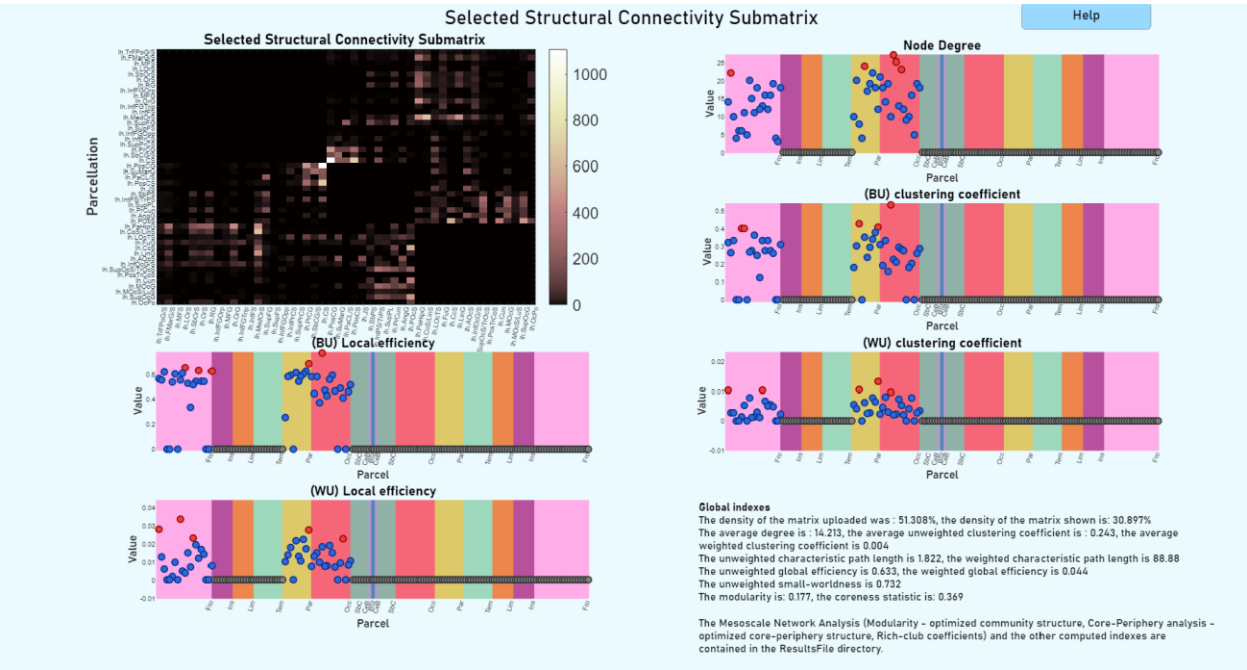


Figure 15. Output of topological analysis (selection of Frontal, Insular and Parietal lobe in the left hemisphere)

2) Visualization of connectivity

There are two modes of visualizing the connectivity between brain regions. Selection is made in right bottom corner in Figure 14.

a) Explorative mode (explore from currently selected subsets) allows the selection of one priority parcel or sub-parcel of the brain. Next, edges from this source to the selected target parcels or sub-parcels are drawn. If the target is not selected, all links between exploration source and all other parcels will be drawn. Example of this mode of visualization is given on the Figure 16.

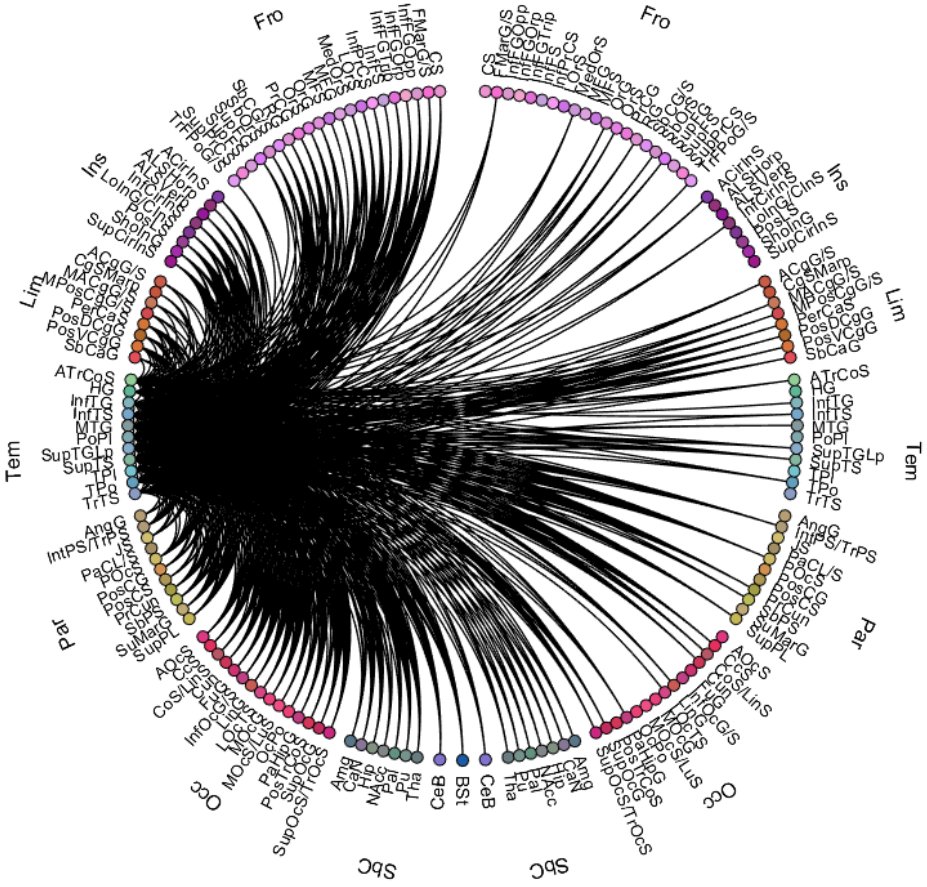


Figure 16. Explorative mode: Connections from Temporal left lobe to all other nodes in the network

b) Another mode instead provides *extracting a subgraph* of all common links within the selected parcels or sub-parcels. Example is given in Figure 17.

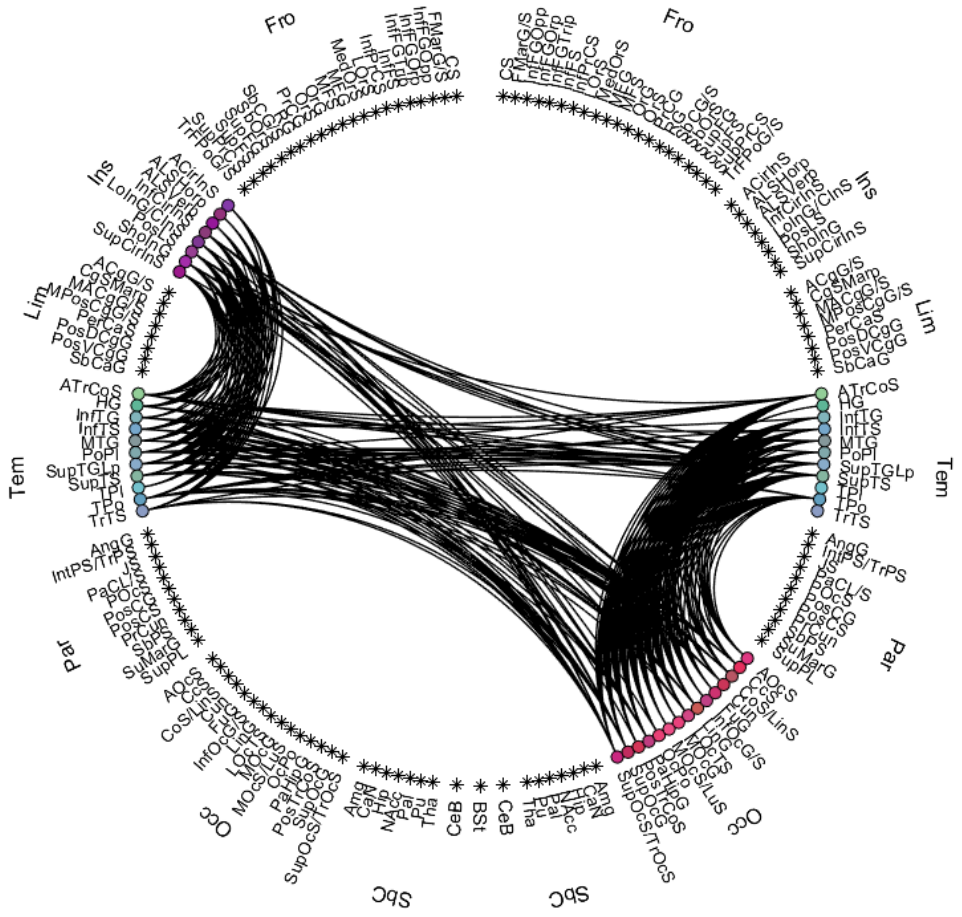


Figure 17. Extract subset: Common links between Temporal and Insular lobe in left hemisphere and Occipital and Temporal lobes in right hemisphere.

3) Additional software functionalities

a) Additionally, the user can select whether to obtain the intra-parcel edges or to remove them (checkbox *keep connections within gross parcels* in the main software interface). Examples of this functionalities are displayed in Figure 18.

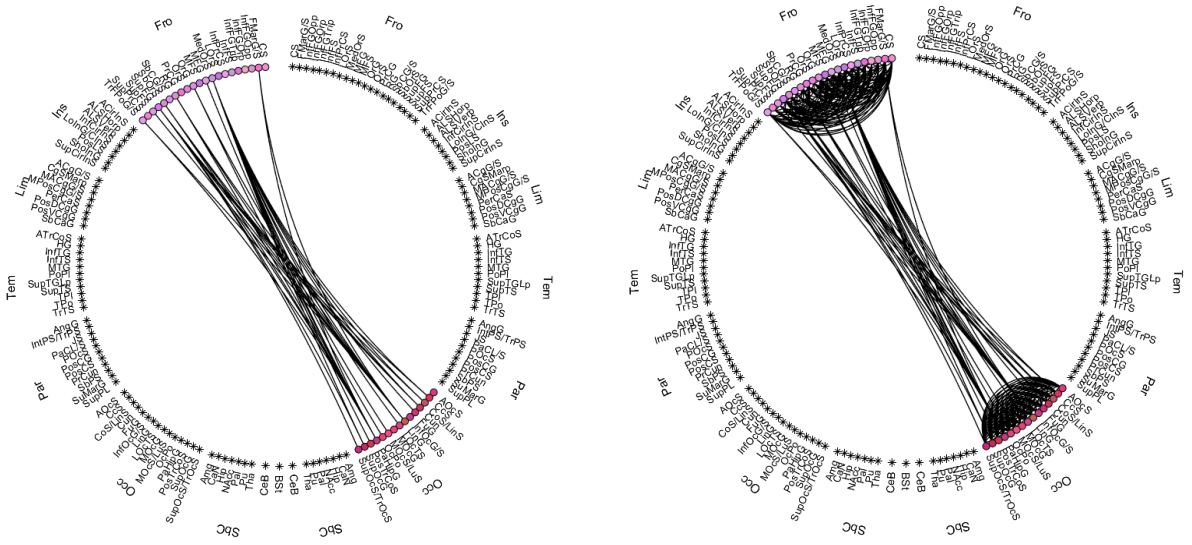


Figure 18. Connectogram without intra-lobe connections (left); connectogram with intra-lobe connections (right) for selection of Frontal left and Occipital right lobe

b) Software provides the possibility to perform density thresholding (checkbox *thresholding* in the main software interface) which can be done both in the binary and the weighted case. Initially, the goal density is selected. Then, the software extracts the threshold that in the best way approximates the initial density to the one that is selected by user. In this process, starting threshold is equal to zero, and as it is increasing weak connections are the first to be removed.

At the start of thresholding process, original matrix is thresholded and then selected subsection is extracted. Hence, topological indexes are calculated on the basis of thresholded matrix, and connectogram provides visualization of thresholded connectivity submatrix (Figure 19).

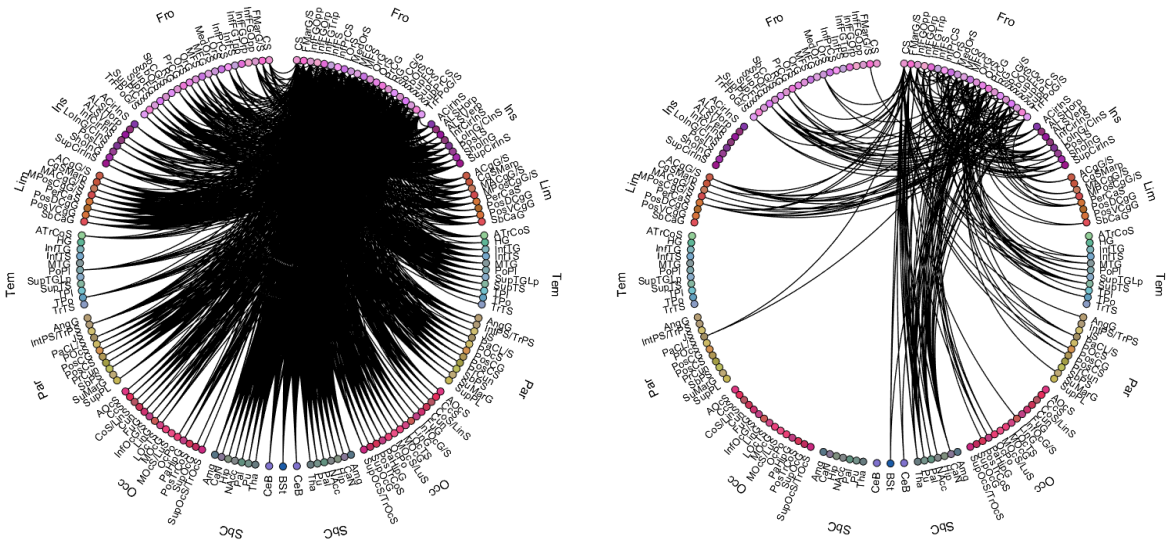


Figure 19. Connectogram with original density (left) and 10% density (right) for selection from Frontal right lobe to all other nodes in the network.

c) Another visualization option is to display weights (checkbox *show the weights in the figure* in the main software interface) in connectogram. Here, links are color coded according to the normalized values of their weights (Figure 20). Weight values normalized interval (from 0 to 1) is divided into three equal-sized bins denoted by different colors. Weakest connections are colored in yellow, middle in blue, while the strongest connections are in red.

If negative connections are present (in the case of study of functional connectivity) they are colored in green.

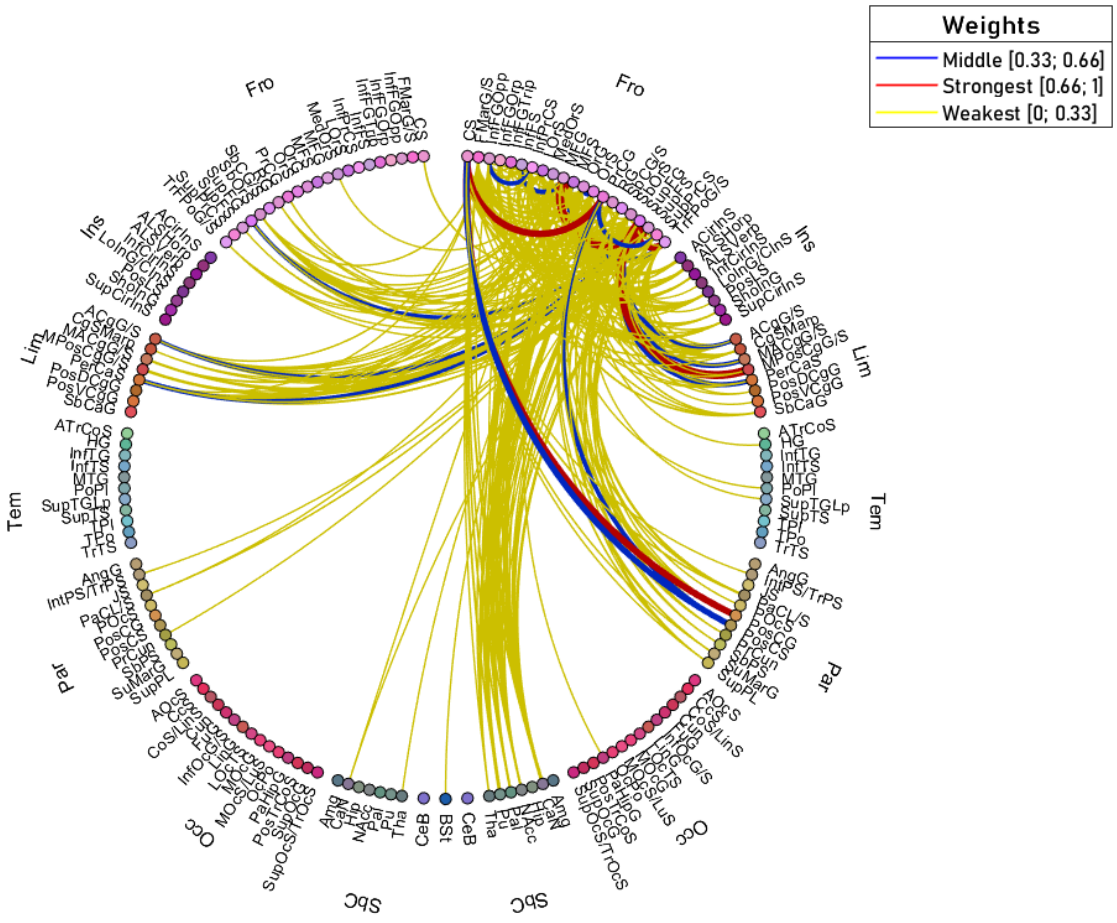


Figure 20. Connectogram with displayed weight values for selection from Frontal right lobe to all other nodes in the network.

d) Furthermore, the software provides possibility to display a specific property of a single node (checkbox *visualization of a node property* in the main software interface). This visualization is based on the heatmap tool used for representing cortical metrics within the circular structure in Circos Software.

By selecting this option, it is possible to code by color a property of the nodes uploaded by the user. Hence, the colors are chosen according to the value of the local node property provided, that typically can be volume or thickness (or any other node attribute uploaded by the file). Moreover, a color-bar corresponding to normalized node property value is

shown in the resulting figure. Example is displayed in Figure 21. Conversely, if this flag is not selected the colors are randomly generated in the manner that nodes with the same gross parcel have similar colors.

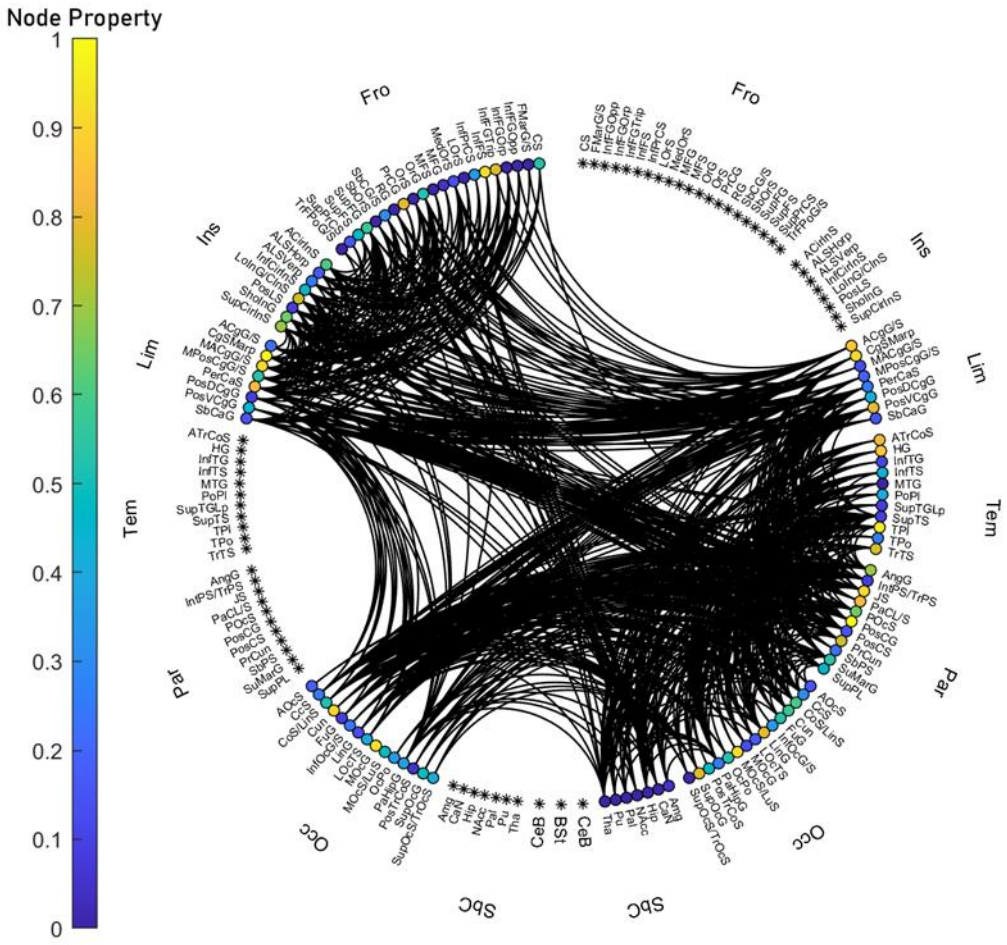


Figure 21. Visualization of a node property (one of the mentioned node indexes) by a color scale (edges are conversely in black)

All previously described functionalities of SNT render visualization of brain connectivity completely adjustable to the users' needs.

3.2 Simulations and data

3.2.1 Test matrices

First tests were performed on two example connectivity matrices of real subjects, although sent us in blind, for testing only. They are denoted as "M_1" and "M_2". Great discrepancy in their density was immediately noticed, which was assumed to be a consequence of the previous methodological pre-processing applied. The number of links in the former is 5424 while in the latter it is 13849, which correspond to density values of 0.2 and 0.5. Both matrices are 165*165 in size and are symmetric and undirected.

Brain Connectivity toolbox was employed for calculation of most commonly used network measures:

1. Density;
2. Average node degree;
3. Average node strength;
4. Global clustering coefficient;
5. Weighted global clustering coefficient;
6. Characteristic path length;
7. Weighted characteristic path length;
8. Weighted global efficiency;
9. Small-worldness.

For the purpose of computing binary network indices, both matrices were binarized, so all the elements of the matrix that are > 0 were set to 1, while all the others were set to 0. On the other hand, if weighted network indices are computed, weight information was retained.

Values of measures computed locally, for each node of the network, compose diverse distributions which are visualized in Chapter 4.

Next, density thresholding was performed on matrix of higher density (M_2). This process was performed for both binary and weighted case. Matrix was thresholded by using multiple values of threshold: 3, 4 and 10 which were set to reach densities reduction to 0.4, 0.35 and 0.3.

3.2.2 Stroke case study

1) Study population

The data set consists of connectivity matrices of two subjects provided by CADiTeR, MRI Laboratory of Fondazione Don Carlo Gnocchi, Milano (IT). Matrices correspond to one stroke patient and one healthy control. Case and control are considered as matching subjects (matching criteria is age):

1. Case: male, 44 years old; stroke patient with a lesion in the right hemisphere, displayed in Figure 22;
2. Control: male, 43 years old.

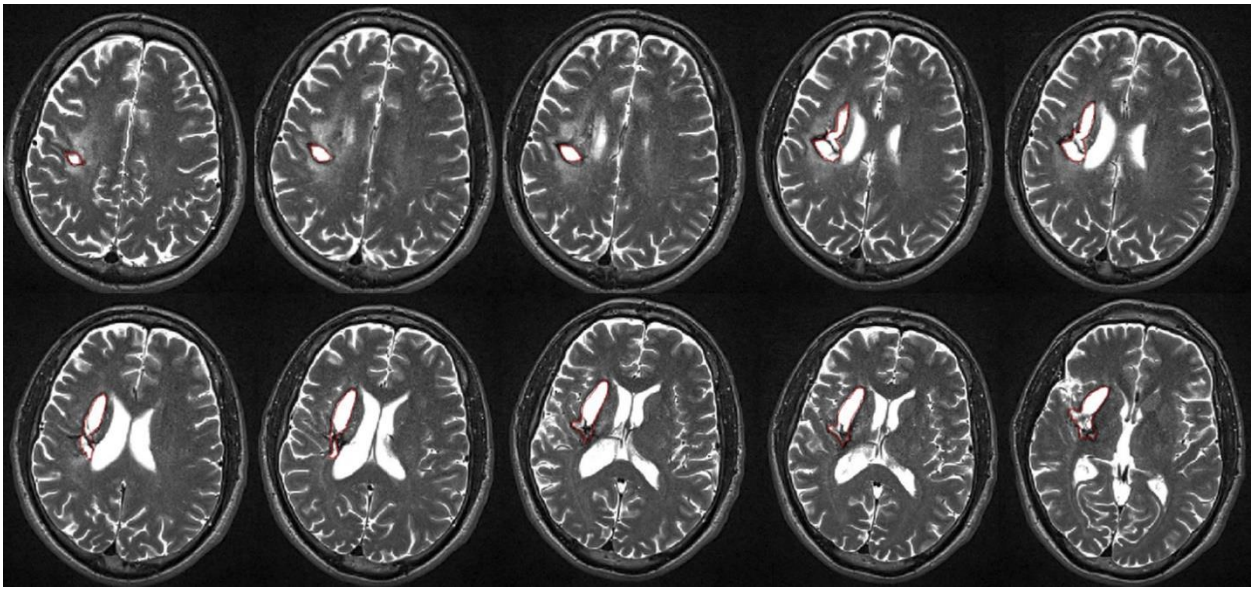


Figure 22. Stroke patient; lesion in the right hemisphere

Moreover, each matrix has two variants based on the different methods used for the evaluation of connectivity strength from tractography data. Therefore, one variant is constructed by using the number of streamlines (or the number of fibers – denoted by NF) as connectivity measure, while the second variant is constructed by evaluating the values of fractional anisotropy (denoted by FA).

2) Data acquisition

The acquisition protocol and methods reported hereunder were used both for stroke patient and matched healthy controls subject.
The MRI protocol was acquired on a 1.5 T Siemens Magnetom Avanto scanner equipped with a 12-channels head coil. The acquisition protocol included:

- a high-resolution 3D T1-weighted Magnetization Prepared Rapid Gradient-Echo (MPRAGE) image, (repetition time (TR)/echo time (TE)=1900/3.37 ms, Filed of View (FoV) = 192x256 mm2, voxel size = 1 mm isotropic, 176 axial slices);
- a diffusion-weighted echo planar images (EPI) image along 64 directions (b- value 1500s/mm2, TR/TE 7800/109 ms, matrix size102x102x46, resolution 2.5 x 2.5 x 2.5 mm3) and 3 b0 images (2 with AP, and 1 with PA encoding direction);
- two conventional anatomical sequences (axial PD/T2) to segment brain lesions (for patients)

The lesions were segmented by an experienced neuroradiologist on the PD/T2 volumes. The MPRAGE volumes were parcellated and automatically labeled using FreeSurfer (version 6 – Laboratory for Computational Neuroimaging at the Athinoula and A. Martinos Center for Biomedical Imaging, HMS and MGH) into 75 cortical parcels for each hemisphere (150 in total) according to the Destrieux atlas [6]. Seven subcortical regions per hemisphere (thalamus, caudate, putamen, pallidum, nucleus accumbens, amygdala and hippocampus) and the brain stem were also segmented using the FreeSurfer automatic labeling process [79] for a total of 165 parcels.

DTI volumes were preprocessed using the FMRIB's Software Library tools. First, the susceptibility-induced geometric distortions were corrected using the acquisition along opposite phase encoding directions with the topup tool; then the DTI images were corrected for eddy current distortion and head motion. The diffusion tensor was estimated for each voxel using the FSL DTIFIT toolbox.

Cortical and subcortical parcels were registered to the DTI space using the FSL flirt toolbox and the WM tracts, connecting each pair of registered parcels (nodes) were reconstructed with TrackVis software. The connectivity matrices were derived by computing the edges as the number of the reconstructed fibers (NF).

3) Experiments and analysis

All network measures in the present work were computed primarily in MATLAB using the Brain Connectivity Toolbox (BCT), and subsequently results were compared with SPT outputs.

Most commonly used network metric are computed both binary and weighted matrices (Par. 4.1.1. and 4.1.2.) For performing binary network analysis, both control and case matrices were binarized by zero. This means that all the matrix entries that have values higher than 0 are set to 1, otherwise they are set to 0. Before obtaining weighted network measures, the diversity between matrices constructed by using fractional anisotropy and number of streamlines as a connection strength measure is emphasized.

Next, accent is put on thresholding process. Disparity in density between two matching subject was immediately clear. Taking this into consideration, matrix of control was thresholded starting from its' initial density and finally reaching the density of the patient.

Thresholding was performed for both the binary and the weighted case (Par. 4.1.3. and 4.1.4.). The difference in outcomes of thresholding between two variants of weighted matrices was analyzed. Further analysis of thresholding effect was made by comparing network metrics on local level. Firstly, comparison is made between original control matrix and stoke case. Next, control matrix is thresholded so that the density corresponds to the density of patients' matrix (Par. 4.1.5.). Furthermore, effect of thresholding on global network measures' values was evaluated (Par. 4.1.6.).

Following this, sensitivity analysis was performed on global indices (Par. 4.1.7) by assessing how global measures were affected by the deletion of a single link between any given pair of nodes. This was implemented by removing only one edge between two nodes in each loop iteration, and subsequently computing the global measure. This process was repeated for each possible pair of nodes in the network. Next, the values of the newly calculated global network metrics were subtracted from the original value of the corresponding metrics, hence the difference in value due to deleting a link was obtained. These values were sorted in descending order, and the highest values extracted to a display table, together with the relevant node labels.

Next, for both control and patient matrices corresponding to left and right hemisphere were extracted. Goal of this test was obtaining metrics differences between the hemispheres of a single subject but also the differences between the control and the patient (Par. 4.1.8.).

In the present work both previously described modes of visualization (*explorative mode* and *extraction of a subgraph*) of SNT were tested on the stroke case. Connectograms depicting connectivity between different brain regions was obtained for stroke patient and control and

subsequently compared. Hence, both the patient's healthy hemisphere and the healthy subject were used as control for the lesioned hemisphere.

Explorative mode was used to visualize connections starting from both left and right Frontal, Temporal and Parietal lobe and reaching all other hemispheric parcellations (Figures 49, 50 and 52).

Moreover, in further tests the focus of this analysis is put on connections located in the right hemisphere lobes, since it is known that the patients' stroke lesion is located there. For this purpose, subgraphs were extracted for visualizing connections between right Temporal and Frontal lobes (Figure 51) and Parietal and Frontal lobes (Figure 53).

Finally, all common connections between Frontal, Parietal and Temporal lobes (in the left and the right hemisphere) were visualized, both with and without including intra-lobe connections (Figures 54 and 55).

As a final point of this analysis, thresholding was employed in the same manner as previously in this work, equalizing density between patient and control (Figure 56).

3.2.3 Borderline Intellectual Functioning Study

1) Motivation for the analysis and study population

Idea for this analysis was based on the study made by the authors of the paper [Blasi et al., 2020] The goal of this study was to investigate brain network connectivity in children who had been exposed to adverse social settings, and display border intellectual functioning (BIF), as well as its relationship to cognitive functioning. In the original study, the connectivity matrices were thresholded (in a way that at least one-third of subjects share the same connections) and subsequently binarized.

Instead of thresholding and binarizing the matrices, the aim of this work was to extend the analysis to the weighted approach. Furthermore, focus was put on the potential importance of keeping the weak connections instead of thresholding them out.

The main motivation to this approach came from multiple studies that, as stated in Chapter 2.3, advise about the importance of weak connection in weighted connectomes [3], [54], [74], [75], [80]. Moreover, recent findings reveal that moderately weak and long-distance connections have

effect in the variance in IQ levels between the subjects. Taking this into consideration, the data set of children with border intellectual functioning and healthy controls was considered as appropriate for this analysis.

The study population included children with BIF associated with significant ELA were recruited from the Child and Adolescent Neuropsychiatry Unit of IRCCS Don Carlo Gnocchi Foundation and the ASST S. Paolo and S. Carlo Hospital. The data set is composed of 32 children with borderline intellectual functioning (BIF) and 14 age-matched typical development (TD) children. Inclusion criteria were:

1. age range comprised between 6 to 11 years old;
2. attendance of a primary mainstream school;
3. a Full-Scale Intelligence Quotient (FSIQ) scores ranging from 70 to 85.

All children underwent a neuropsychological evaluation including the Wechsler Intelligence Scale for Children-III (WISC-III); the Child Behavioral Checklist (CBCL 6-18); the Socio-Economic status - SES; the Environmental Stress Checklist - ESCL, an ad hoc developed checklist to explore the environmental stress the children were exposed to. Among the 32 children with BIF, 25 had an Adjustment Disorder, 4 had a Generalized Anxiety Disorder and 1 had a Disruptive, Impulse-Control, and Conduct Disorder. Moreover, in 14 children a Specific Learning Disorder was associated, in 14 there was a history of Language Development Disorder. Demographic data relative to the 32 BIF and 14 TD children are summarized in Table 3. No significant differences were found for age, sex, and (SES), while the Child Behavior Checklist CBCL ($p=0.005$), the ESCL, and as expected the Full-Scale Intelligence Quotient FSIQ ($p<0.0001$) were significantly different between the two groups.

	TD (n=14)	BIF (n=32)	p-value
Age in years, median (IQR)	9.2 (8.5-9.6)	8.6 (8.2-9.9)	0.543
Male number (%)	7 (50%)	17 (53%)	0.9
FSIQ, median (IQR)	119 (111.5 -121.5)	80 (75 - 84)	**<0.0001**
CBCL, median (IQR)	49.5 (41-53)	59 (52 - 66.5)	**0.005**
SES, median (IQR)	25 (25-48.8)	23.5 (16.3 - 31.4)	0.119
ESCL, median (IQR)	0.0 (0.0 -1.0)	2.0 (1.0 - 4.0)	**<0.0001**

Table 3. FSIQ, Full-Scale Intelligence Quotient; CBCL,Child Behavior Checklist, SES, Socio-Economic Status; ESCL, Environmental Stress Check List. The correction for multiple comparisons was implemented with the Bonferroni Correction, setting the significance threshold at $p \leq 0.0125$. Significant p-values are highlighted in bold [81]

2) Data acquisition

MRI was performed on a 1.5 T Siemens Magnetom Avanto scanner equipped with a 12-channels head coil. The acquisition included:

- a 3D T1-weighted Magnetization Prepared Rapid Gradient-Echo (MPRAGE) image, (repetition time (TR)/echo time (TE)=1900/3.37 ms, Filed of View (FoV) = 192x256 mm2, voxel size = 1 mm isotropic, 176 axial slices);
- DTI included a diffusion-weighted EPI image along 30 directions with b-value=1,000 s/mm2 and the reference image without diffusion weighting (TR/TE = 6,700/100 ms, FoV = 200x200 mm2, voxel size 1.6x1.6x2.5 mm3, 40 axial slices, two runs);
- two conventional anatomical sequences (axial PD/T2 and coronal FLAIR) were performed to exclude gross brain abnormalities.

The 3D-T1 images were segmented and parcellated using FreeSurfer [79] version 5.31 into 148 cortical areas (74 for each hemisphere) according to the Destrieux atlas [6]. Furthermore, the FreeSurfer automatic labeling process was used to extract seven subcortical regions per hemisphere (thalamus, caudate, putamen, pallidum, and nucleus accumbens, amygdala and hippocampus) and the brain stem for a total of 163 parcels.

Using FMRIB's Software Library tools3, the DTI images were corrected for eddy current distortion. The motion evaluation was performed by checking the relative movement estimated by eddy toolbox and excluding subjects exciding a threshold fixed to 0.5. Then, using the FSL DTIFIT toolbox4 the tensor was estimated for each voxel. The cortical/subcortical parcels were registered to the DTI space using the FSL flirt tool. Finally, the WM tracts connecting each pair of registered cortical and subcortical parcels (nodes) were reconstructed with TrackVis software5.

The connectivity matrices were derived by computing the edges as the number of the reconstructed fibers normalized by the sum of the nodes volumes (NFn) in order to consider the effect of anatomical variability.

3) Experiments and analysis

Firstly, a two-sample t-test was performed. The t-test is an inferential statistic method used to determine if there is a significant difference in the means of two groups, under the hypothesis of Gaussian dispersion. For this purpose, Matlab function *ttest2* is used. This function returns the test decision for previously set null hypothesis stating that the data in two subject groups have equal mean. Moreover, assumptions made for this test are that both groups' data are drawn from independent random samples that form normal distribution and have equal but unknown variances. Null hypothesis is rejected if 5% significance level is reached.

In present work, t-test was employed with the goal of evaluating the potential importance of weak connections. This was done by performing and subsequently comparing the two t-tests results:

a) Comparing control and patient original matrices;
b) Comparing weak connections only for control and patient.

The level of significance was set to 0.05. According to Santarnecchi et al. [75] connections with weights lower than 0.2 were considered weak.

Following this, bootstrapping is utilized. This approach necessitates the creation of an ensemble of simulated data (surrogates) that is both similar to, and consistent with the hypothesis of interest. It allows to approximate the mean and variance of an estimator, construct confidence intervals and calculate p-values of tests when the distribution of the statistic of interest is unknown.

More precisely, bootstrapping method was applied for constructing surrogates of control subjects. In the construction of one surrogate matrix, each entry of surrogate matrix is equal to

the value at the corresponding entry position from one of the 14 randomly selected control connectivity matrices corresponding to controls (Figure 23). This procedure was repeated 1000 times resulting in 1000 surrogate matrices. After every surrogate matrix construction, network metrics were calculated and saved in the vector. In this way, 1000 control matrices are simulated, providing more reliable analysis.

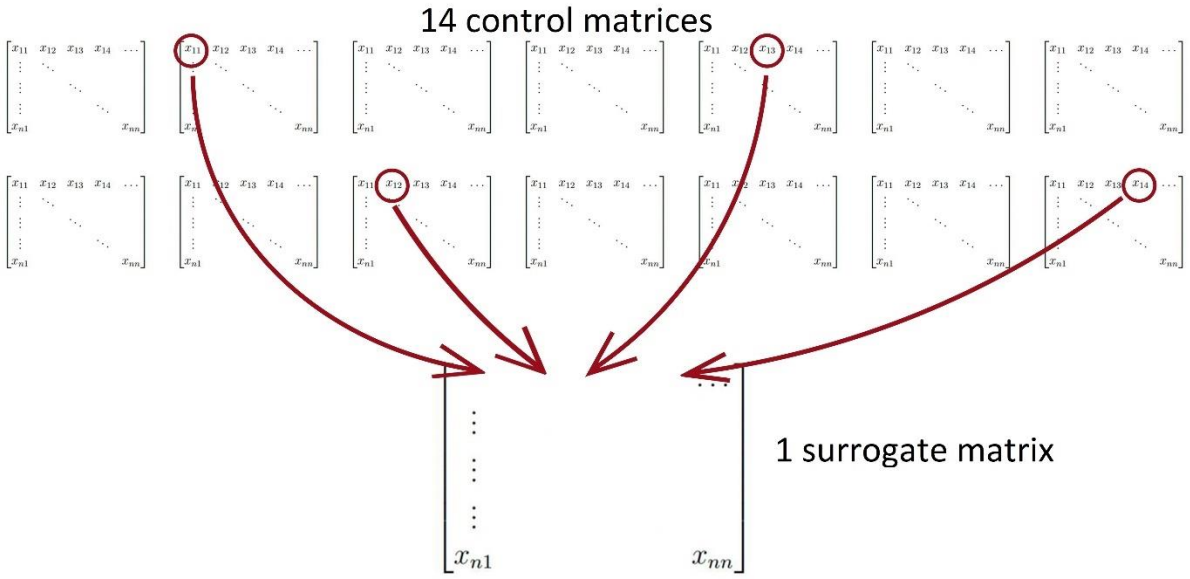


Figure 23. Construction of one surrogate matrix

The purpose of this method is obtaining a confidence interval for the normality range of diverse network measures computed for the data corresponding to the control subjects. In this way significant differences to Borderline Intellectual Functioning condition can be observed.

4 Results and discussion

The present chapter provides results of the tests performed.

Firstly, result of analysis is displayed for two sample test subjects introduced in Par. 3.4.1. These analyses were performed to test the functions implemented in the Brain Connectivity Toolbox (BCT) and run simulations for SPIDER-Net tool (SNT) development.

Following this, the stroke case study results are given. These include calculated binary and weighted network measures. Furthermore, the effect of density thresholding is widely analyzed evaluating the disparities between thresholding outcomes in case of connectivity matrices based on fractional anisotropy (FA) or on the number of streamlines (NF). Next, a sensitivity analysis was performed to evaluate the robustness of the network to the possible errors that can be introduced by tractography methodology. As a final point of this analysis, SPIDER-NET software was tested for visualizing connectivity between diverse brain regions and compared between the stroke patient and healthy control.

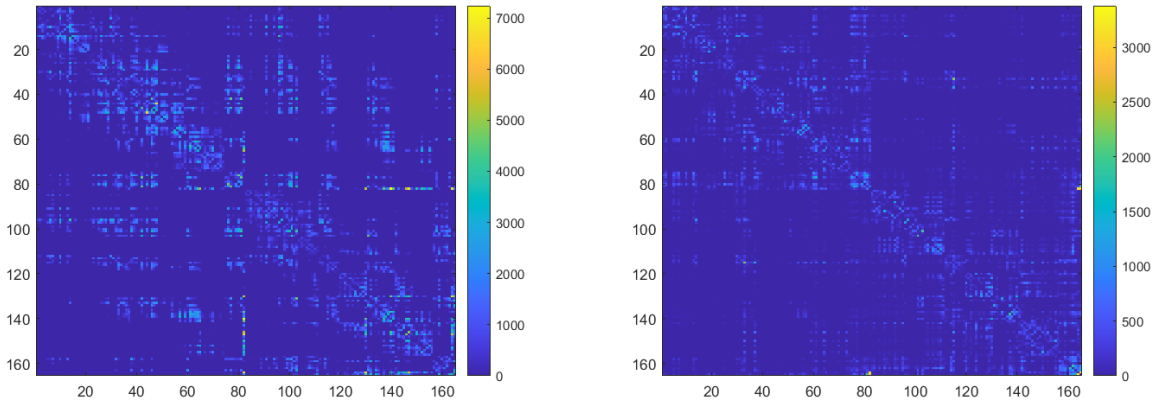
As a final point of this chapter, results of Borderline Intellectual functioning study are given. These results include the outcomes of two sample t-test performed, also considering the case of only weak connections. Furthermore, preliminary results in the application of bootstrapping method to confirm and extend the analysis are given.

4.1 Illustrative results

Values of network measures are calculated for the two example connectivity matrices, denoted by M_1 and M_2. As mentioned in Par. 3.2.1. these two matrices are largely different in density since that of M_2 matrix is twofold that of M_1. The trend of measure changes under thresholding process was analyzed for M_2 matrix. Since background information on these matrices is not known and hence cannot be connected to any physical meaning, focus is put on different visual representation of calculated measures distributions to be included in the SNT and analysis of the local measures. All results are obtained by using the Brain Connectivity Toolbox (BCT) functions. A detailed analysis is given below, relevant to both local and global indexes, while the latter ones are next summarized in Table 4.

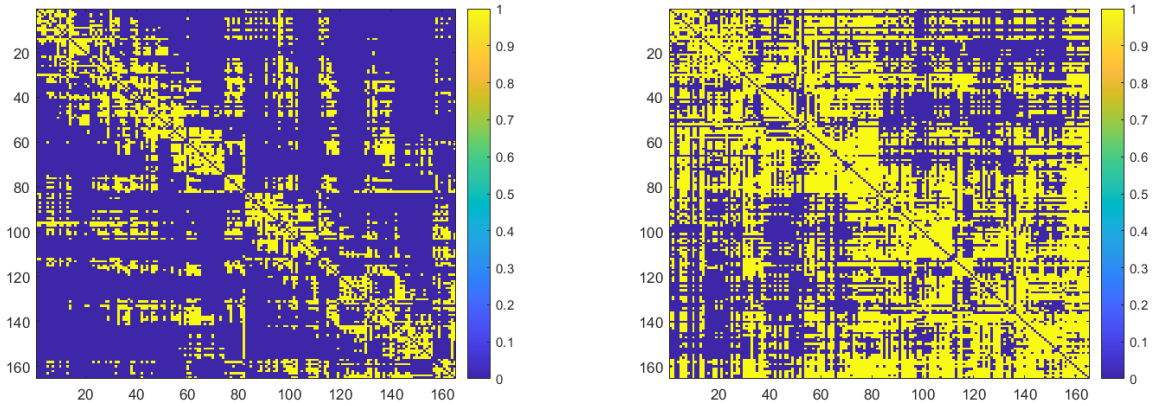
4.1.1 Binary and weighted matrix analysis

Matrices before the binarization process are displayed in Figure 24, a) corresponds to the M_1 matrix, while b) corresponds to M_2. Here, it can be seen that weights in both matrices have great variability. Moreover, M_1 weights have much higher value (reaching up to 7000) compared to M_2 (weights reaching up to 3000). Figure 25 shows results for the binarization process, which enhances the density difference.



a) M_1 matrix *b) M_2 matrix*

Figure 24. Matrix representation before binarization

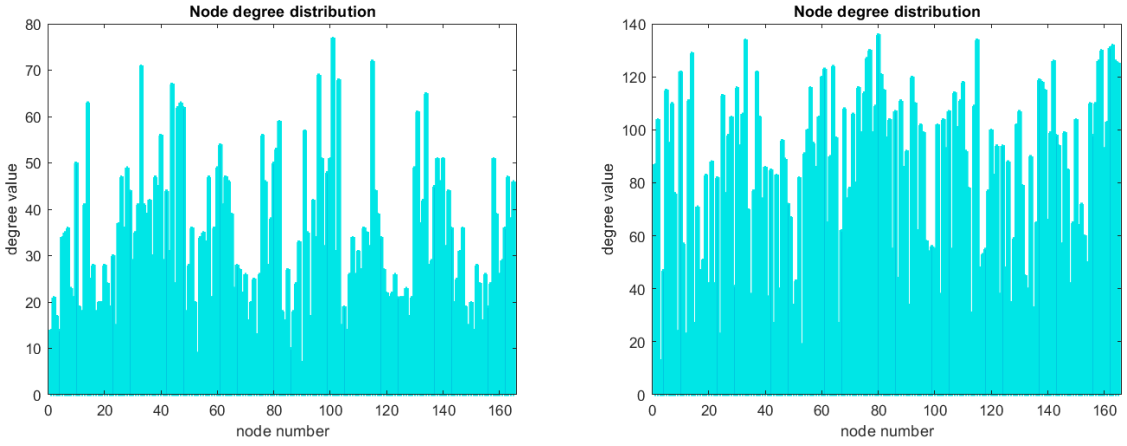


a) M_1 matrix *b) M_2 matrix*

Figure 25. Matrix representation after binarization

4.1.1.1 Node degree and node strength

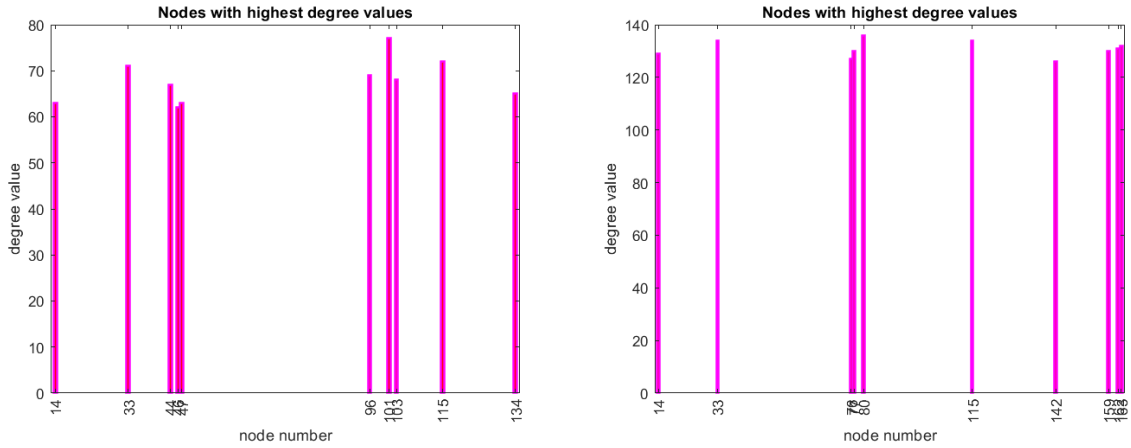
Degree of all nodes calculated locally is displayed in Figure 26. The average degree value in case of M_1 matrix is equal to 33.87, minimal degree value is 7, while the maximum is 77. On the other hand, average degree for M_2 matrix is 84.2, minimum is 14 and maximum 136.



a) M_1 matrix *b) M_2 matrix*

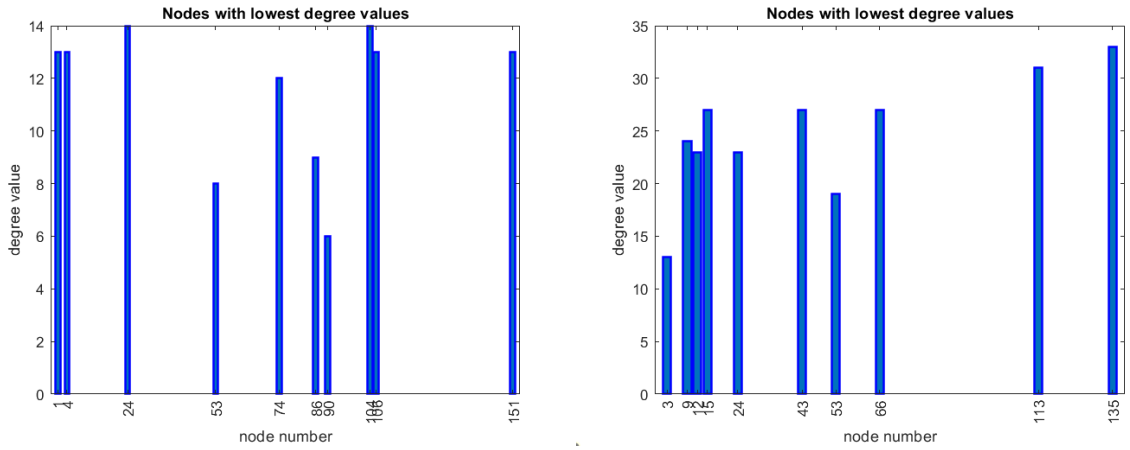
Figure 26. Node degree distribution

All degrees of the network form the degree distribution plot, which provides a simple representation of network connectivity. It can be clearly seen that the node degree is generally higher for M_2 matrix. This visualization provides valuable insights regarding the overall connectedness in the network. The most connected nodes (nodes with the highest degree value) for M_2 matrix are SupFG (Superior frontal gyrus), PerCaS (Pericallosal sulcus), Pu (Putamen), Hip (Hipothalamus) in the left hemisphere, and PerCaS (Pericallosal sulcus), PaHipG (Parahippocampal gyrus), Hip (Hipothealamus) and Tha (Thalamus).



a) M_1 matrix *b) M_2 matrix*

Figure 27. Highest degree nodes

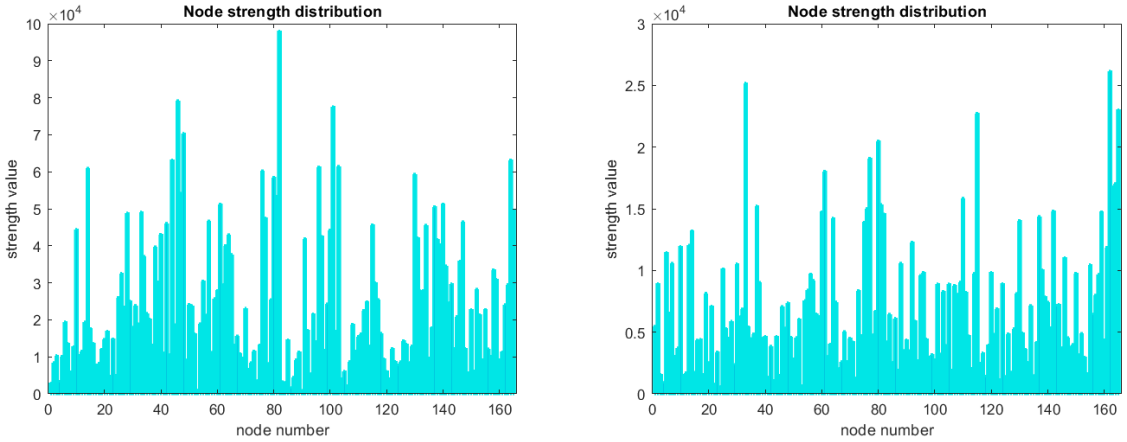


a) M_1 matrix *b) M_2 matrix*

Figure 28. Lowest degree nodes

Figures 28 displays the 10 highest degree values in the *x*-axis positions relevant to the original order of all parcels. Using this visualization for extracting the highest degree nodes, as in Figure 28, hubs of the network can be easily detected and retrieved. Conversely, extracting lowest degree nodes may be useful when studying the effect of neurological disorder or injury, where nodes with mostly compromised degree value can be detected.

Similarly, the local strength of all nodes is displayed in Figure 29.



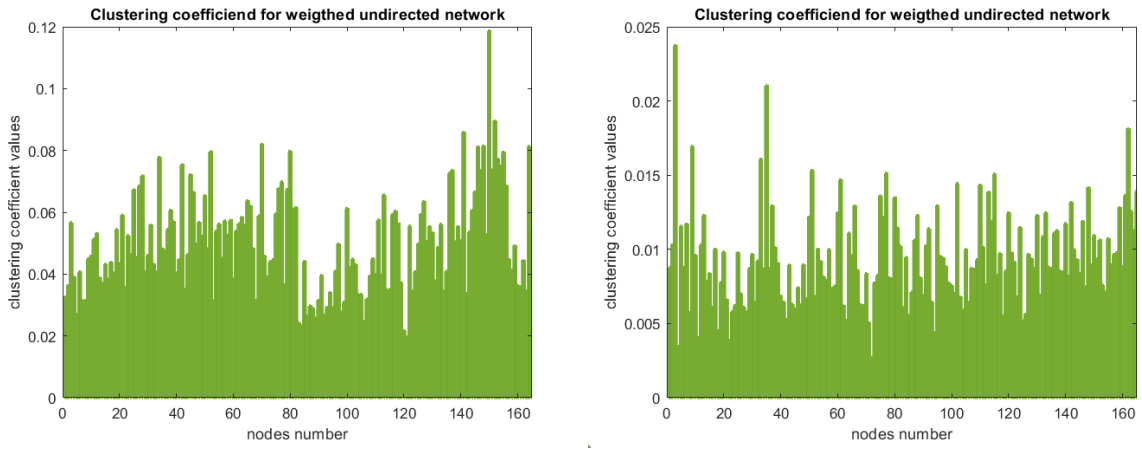
a) M_1 matrix *b) M_2 matrix*

Figure 29. Node strength distribution

Like degree distribution, the node connection strength distribution provides a simple representation of connectivity in a weighted network. Observing the obtained range of strengths, it can be seen that matrix M_1, even though less dense, is composed by generally higher weights values in comparison to Matrix M_2. More precisely, the average node strength value for M_1 matrix is 24148, whereas for M_2 matrix it is 7009,6.

4.1.1.2 Clustering coefficient

The binary clustering coefficient computed over all nodes in the network is displayed in Figure 30. Additionally, the global clustering coefficient for M_1 is 0.6603, whereas for M_2 it is 0.7371. The same analysis was performed for the weighted clustering coefficient which highlighted stronger values in the left hemisphere for M_1. This difference is confirmed by Figure 30, where it is obvious that highest strength nodes are located exactly in this part of the brain.



a) M_1 matrix *b) M_2 matrix*

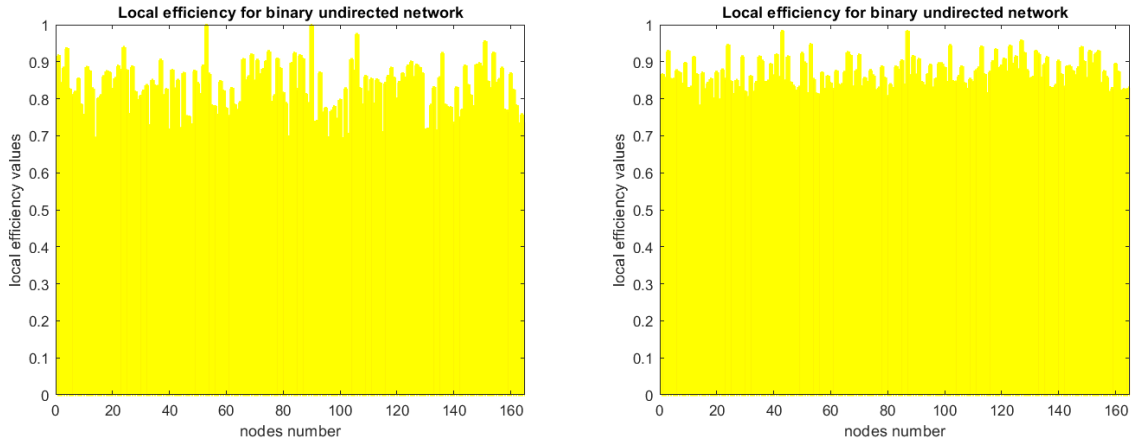
Figure 30. Clustering coefficient for weighted undirected network (Onnela et al. implementation)

4.1.1.3 Characteristic path length

The characteristic path length for M_1 was 2.086, while for M_2 it was 1.4856. Taking into the consideration that the density of M_1 matrix is only 0.2, the calculated path length was unexpectedly low, meaning that to reach a node from any other node of the network only ~2 steps on average are needed. From this, it can be concluded that the low number of edges, a 20% of potential ones, is highly efficient for the whole network integration. On the other hand, the value of characteristic path length for M_2 was expected to be lower compared to M_1, which is more dense. The weighted characteristic path length obtained for the M_1 was 0.0022 and for M_2 was 0.0073. These results can possibly be a consequence of the weight distribution in the adjacency matrix (as mentioned, weights in M_1 are reaching twice as high values compared to the M_2 matrix).

4.1.1.4 Global and local efficiency

The global efficiency was inversely related to characteristic path length. The global efficiency calculated for binary M_1 matrix was equal to 0.5532 and 0.7554 for M_2. For the weighted networks, the global efficiency for M_1 was 0.0858 and 0.0514 for the M_2. The local efficiency values distribution across all nodes of the binary network is displayed in Figure 31.



a) M_1 matrix *b) M_2 matrix*

Figure 31. Local efficiency for binary undirected network

4.1.1.5 Small-worldness

The first step in computing small-worldness was creating 100 random networks with the same number of nodes and edges as the networks of interest. Subsequently, values of the global clustering coefficient and characteristic path length for random networks are obtained and averaged (generated for both M_1 and M_2 matrix). The average global clustering coefficient of generated random networks were 0.2067 (generated for M_1) and 0.5142 (generated for M_2). The characteristic path length values were 1.7895 and 1.4775.

The value of small-worldness metrics is obtained for the M_1 matrix was 0.9125, whereas for M_2 matrix it was 1.4292. Hence, it can be concluded that M_1 does not have small-world properties, since the value of small-worldness metric is less than 1. Since the small-world property is considered nearly-universal for real-world data, it can be assumed that the differences between the two matrices, especially in the density of M_1, could be attributed to possible differences in the acquisition protocol of DTI methodology (e.g. different predefined threshold for deterministic/probabilistic approaches) that likely removed important information. This result should explain the poor M_1 efficiency, despite its higher density.

4.1.2 Results summary

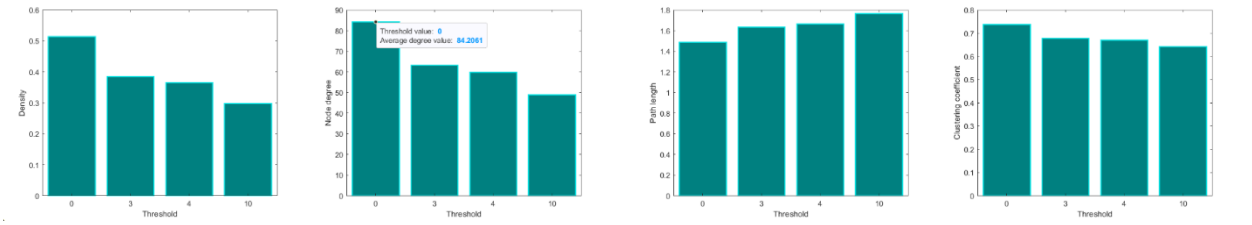
Results of all computed global indexes for both M_1 and M_2 matrix are displayed in Table 4.

	M_1	M_2
Density	0.2	0.5
Average node degree	33.87	22.8
Average node strength	24148	7009.6
Global clustering coefficient	0.6603	0.7371
Weighted global clustering coefficient	0.0502	0.0094
Characteristic path length	2.086	1.4856
Weighted characteristic path length	0.0022	0.0073
Global efficiency	0.5532	0.7554
Weighted global efficiency	0.0858	0.0514
Small-worldness	0.9125	1.4292

Table 4. Network measures computed for incomplete and complete matrix

4.1.3 Binary thresholding

Density thresholding was performed on the M_2 binary matrix. As stated in the Chapter 2.3.1, this approach is mainly used in canceling out the differences in number of connections between the subjects, and it was used in further analysis made the present work. Starting from the original density, M_2 was thresholded by using multiple values of threshold: 3, 4 and 10 which were set to reach densities reduction to ~ 0.4, 0.35 and 0.3. These values of densities were chosen arbitrarily, to observe the trend of network measures under thresholding. Results are displayed in Figure 32.



a) Density *b) Average node degree* *c) Characteristic path length* *d) Global clustering coefficient*

Figure 32. Binary thresholding: Density, node degree, characteristic path length and global clustering coefficient for binary matrix

Figure 32 a) shows how the increase in the threshold leads to the decrease in density. In Figure 32 b) it can be seen that the node degree is decreasing with increasing threshold, since increase in the threshold means decrease in the edge numbers. Furthermore, Figure 32 c) shows that the characteristic path length is increasing as the threshold increases. This could be expected since the distribution shifts to longer paths lengths when weak edges serving as "shortcuts" are removed. Distribution of the binary clustering coefficient varies from 0.7371 to 0.6418 (Figure 32 d)) which is a consequence of weak edges removal and the loss of strength in some nodes, eventually resulting in a clustering coefficient decrease.

Figures 33 and 34 provide the visualization of node degrees across all nodes of the network before and after the thresholding process.

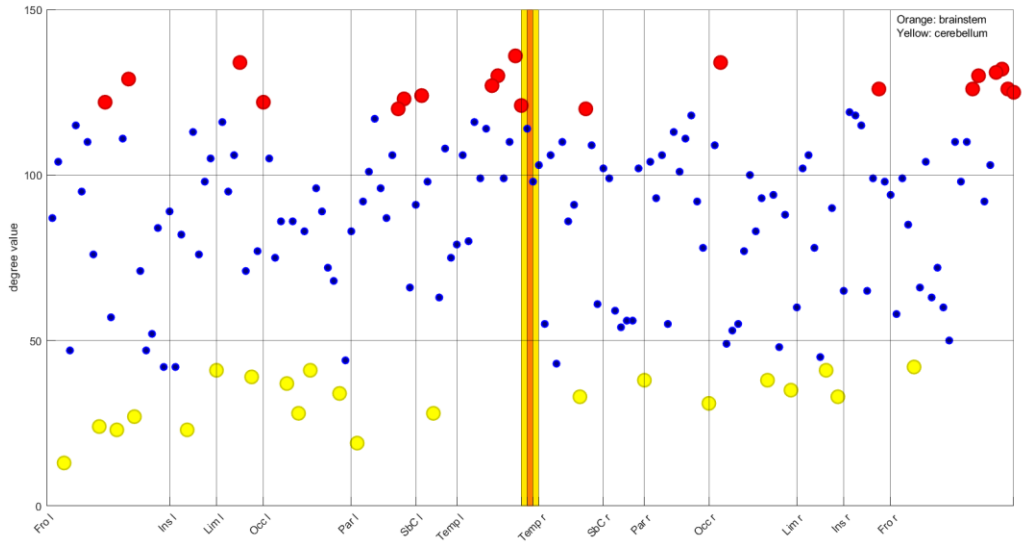


Figure 33. Node degree distribution for original M2 matrix. Red color denotes highest degree nodes, blue middle values and yellow the lowest values

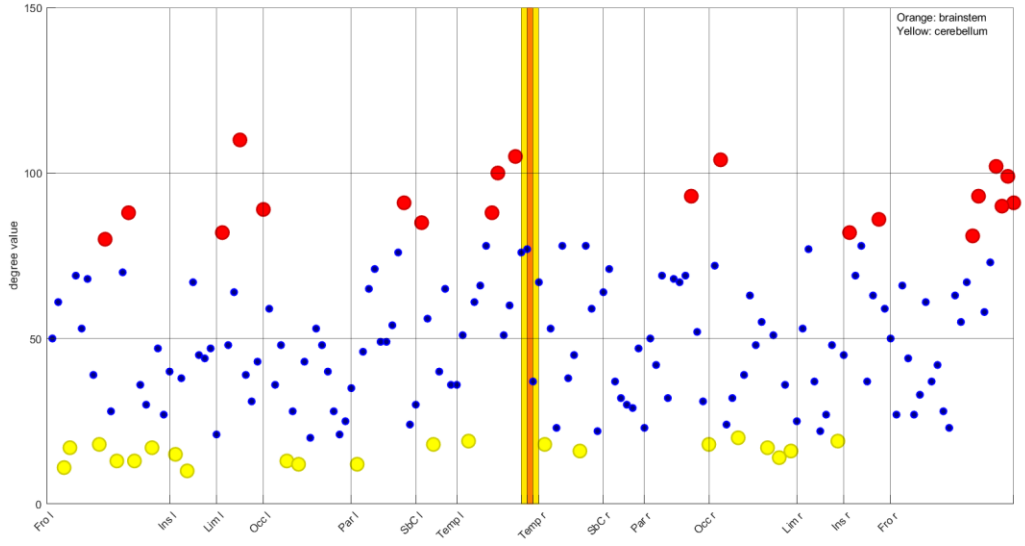


Figure 34. Node degree distribution for thresholded M2 matrix. Red color denotes highest degree nodes, blue middle values and yellow the lowest values

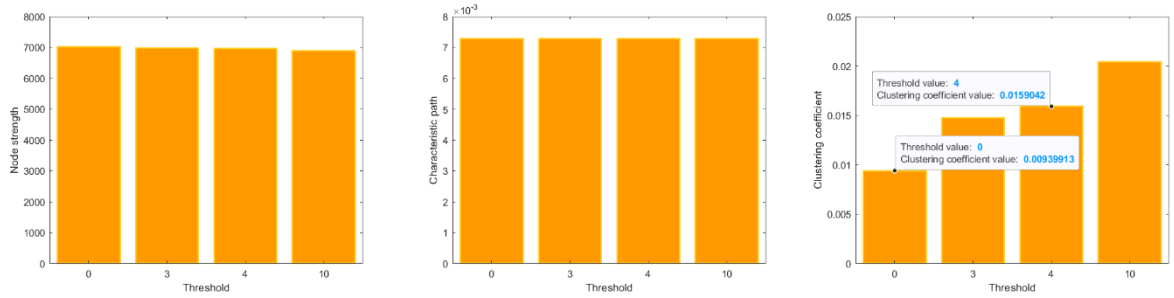
Nodes are divided in 3 different color-coded groups according to their degree value. Red color denotes highest degree nodes, blue middle values and yellow the lowest values. Additionally, on the *x*-axis of the graph hemispheric parcellations are listed so degree can be visualized across the lobes of left and right hemisphere, respectively.

Figure 33 represents the original degree values in M_1, while Figure 34 was obtained by a threshold of 10 (i.e., the highest value of threshold used in previously described binary thresholding).

Comparing these two figures it is evident that there is high difference in distribution of node degrees between without a threshold and the maximal considered threshold. The maximal node degree is decreasing by thresholding, as previously stated.

4.1.4 Weighted thresholding

The same thresholding procedure previously described was performed on weighted M_2 matrix and results are displayed in Figure 35.



a) Average node strength *b) Characteristic path length* *c) Global clustering coefficient*

Figure 35. Weighted thresholding: Average node strength, characteristic path length and global clustering coefficient for weighted matrix

In case of weighted matrix, node strength (corresponding to the binary measure of node degree) is relatively constant for all nodes as the threshold is increasing. (Figure 35.a). The characteristic path length appears to be constant under various thresholds (Figure 35.b). This is because thresholding removes the weakest weight connections which are several orders of magnitude lower than the highest weights values. Weighted clustering coefficient varies only slightly with varying thresholds which can also be explained by specific weight distribution. Behavior of both binary and weighted indices under thresholding will be analyzed in further detail in thestroke case study (Par. 4.2.2 and 4.2.3).

4.2 Stroke case study

4.2.1 Binary network

Results of computation of global network measures are displayed in the Table 5.

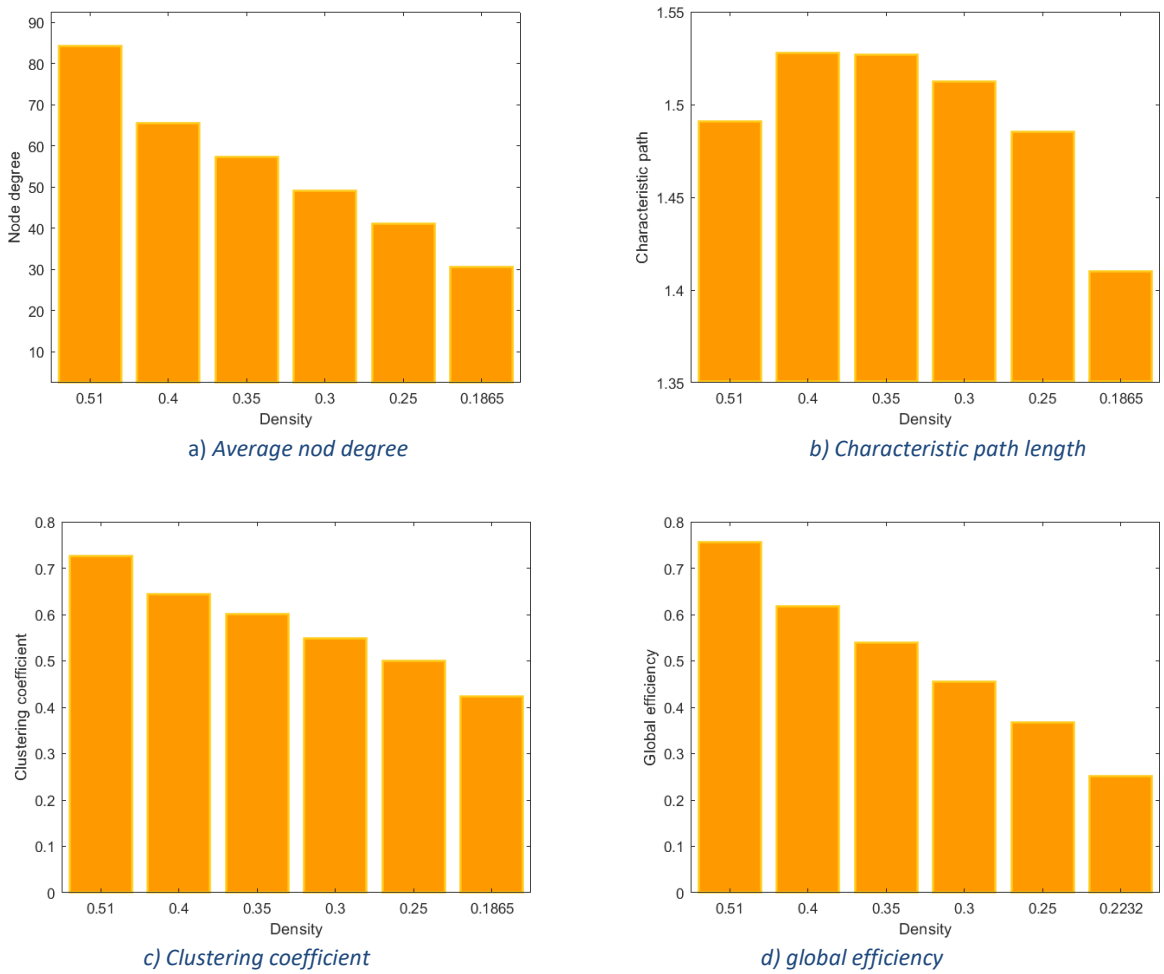
	Average node degree	Density	Characteristic path length	Global clustering coefficient	Global efficiency
Patient	36.6060	0.1865	1.9154	0.6043	0.5866
Control	84.1455	0.5153	1.4817	0.7270	0.7559

Table 5. Global measures for binary network

Regarding Table 5, the density is obviously compromised in case of stroke. However, the strongest difference is seen in average node degree. Moreover, the characteristic path length is significantly higher, indicating the lower network integration and thus lower ability for efficient information transfer trough the network. Accordingly, global efficiency (inverse of path length is lower in case of stroke.

4.2.2 Binary matrix thresholding

As it can be seen in Table 5, the matrix density of the control case is more than two times higher compared to the density of the stroke patient. As stated in Par. 2.3.1., various graph measures are dependent on the network density. Therefore, differences in computed measures may reflect this density discrepancy instead of reflecting the disease-related outcomes. Taking this into consideration, matrix of control was thresholded starting from initial density and finally reaching the density of the patient. This method is known as density thresholding (see Par. 2.3.2). Results of binary thresholding are displayed in the Figure 36.



a) *Average nod degree*

b) *Characteristic path length*

c) *Clustering coefficient*

d) *global efficiency*

Figure 36. Binary Thresholding: Node degree, characteristic path length, global clustering coefficient and global efficiency for binary matrix

The average binary network degree is decreasing as the threshold is increasing (Figure 36.a), which is expected since increase in threshold results with reduction of the edges number in the graph. This is in keeping to the findings by Colon-Perez [53]. Regarding the path length it is expected that removing links by thresholding will result in longer path lengths, since in binary network path between each pair of connected nodes is equal to 1. However, results displayed in Figure 36.b), show a different behavior. Initially the trend is upward, but after reaching the density of 0.25 it starts decreasing. Additional tests were performed where threshold was set to values even lower than the lowest value displayed on the graph, which resulted in abrupt increase of characteristic path length. High values of threshold imposed disconnection of many nodes, which, in turn, abruptly raised the average path length.

4.1. 2. Weighted network

Before observing the values of global coefficients for weighted networks, attention should be brought to the effects on indices due to different connection metrics, namely FA and NF. It is worth recalling that in weighted analyses the length of an edge is conventionally assumed as the inverse of the weight. Figure 37 displays box plots of normalized data sets. In Figure 37 a) weights distribution is displayed. Regarding the FA, median value is 0.44. The weight distribution is such that the 98th percentile is equal to 0.8. On the other hand, median value for NF matrix is $3.34 * 10^{-4}$ and 98th percentile is equal to 0.15. Figure 37 b) shows lengths distribution. Blue color scale corresponds to the FA matrix, while red to NF one. Here, 75th percentile for FA matrix is equal to 1.57 and median value is 1.20. In case of the NF matrix 75th percentile is 186.94 while the median value is 4.56. Global measures for weighted network are displayed in Table 6. Prior to calculations, all matrices were normalized to the interval between 0 and 1.

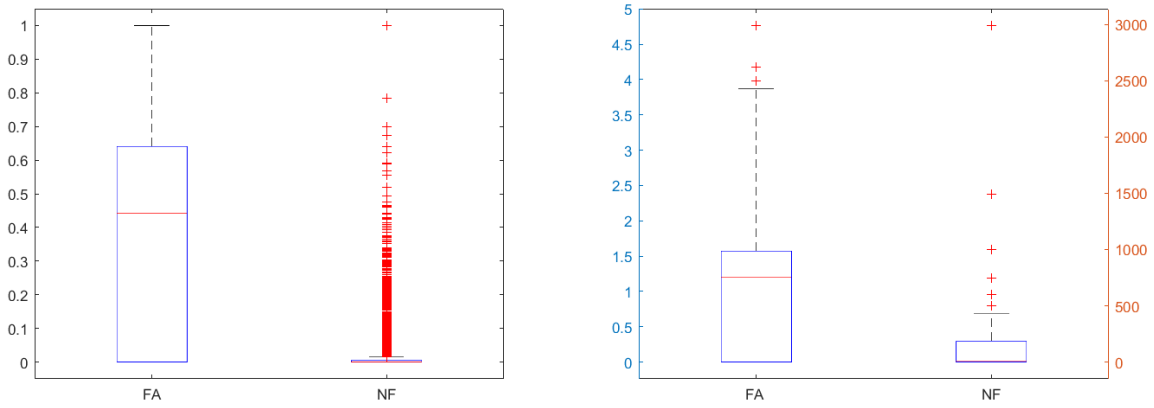


Figure 37. Boxplot of weights (a) and lengths (b) distributions for FA and NF matrix

	Average node strength	Density	Characteristic path length	Global clustering coefficient	Global efficiency
Patient (FA)	20.1079	0.1865	3.0194	0.3927	0.3784
Patient (NF)	1.1384	0.1865	36.7087	0.0152	0.0384
Control (FA)	53.9081	0.5131	2.1258	0.4618	0.5098
Control (NF)	2.3918	0.5131	20.7833	0.0107	0.0602

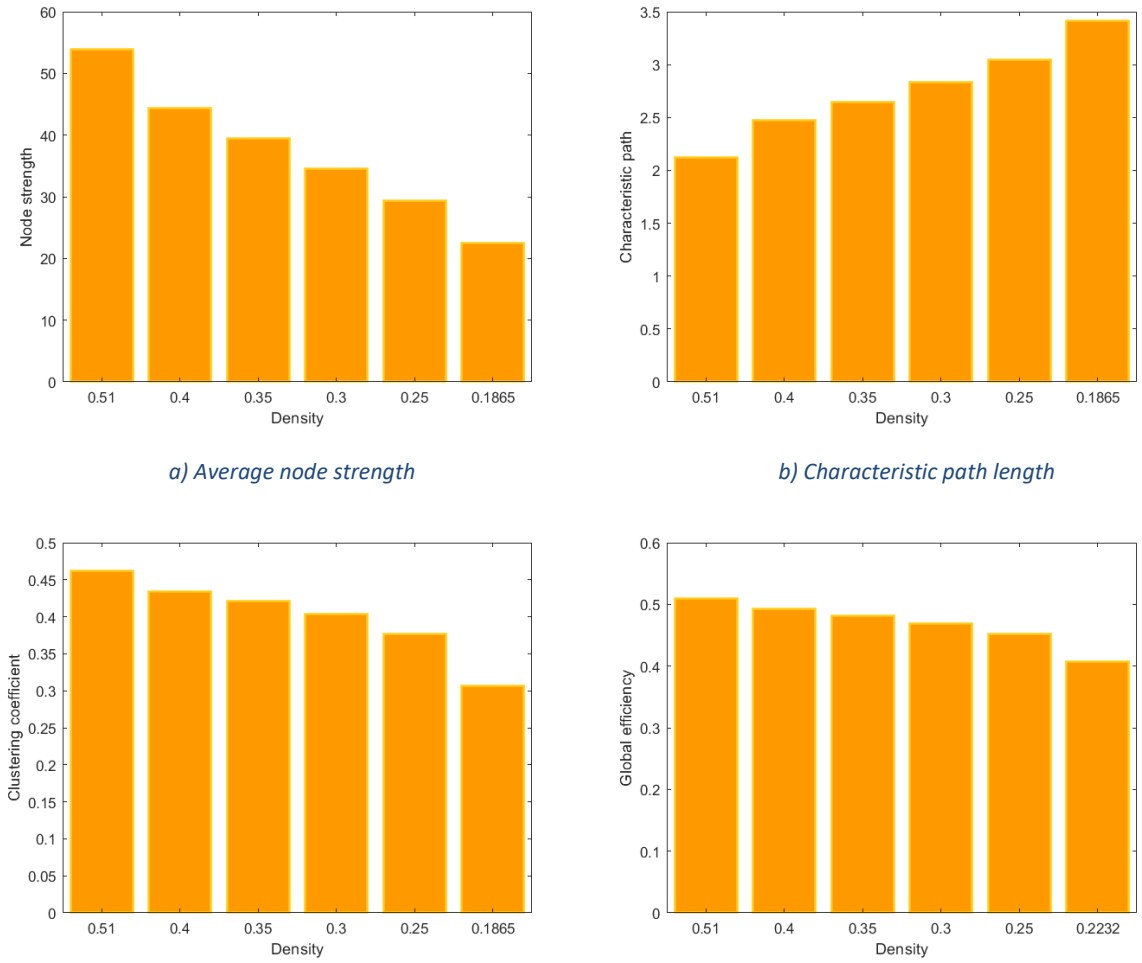
Table 6. Global measures for weighted networks

As to Table 6, it can be seen that the index values highly differ (on order of magnitude) depending on whether FA (i.e., fractional anisotropy) or SN (i.e., tract quantity) was the applied metrics. This behavior of global network measures reflects the large difference or the two weight distributions, as described in the previous section.

Despite the highly different absolute values with metrics, stroke vs. control differences observed in binary indexes were maintained by the weighted approach. The average node strength (analogous to node degree, in binary networks) is about two-fold lower in case of stroke. Similarly, to the binary case, the stroke patient displays a higher characteristic path length and lower clustering coefficient, depicting compromised integration and segregation in the network.

4.2.3 Weighted matrix thresholding

Different outcomes are produced by thresholding FA and NF matrices. Results are displayed in Figures 38 and 39.



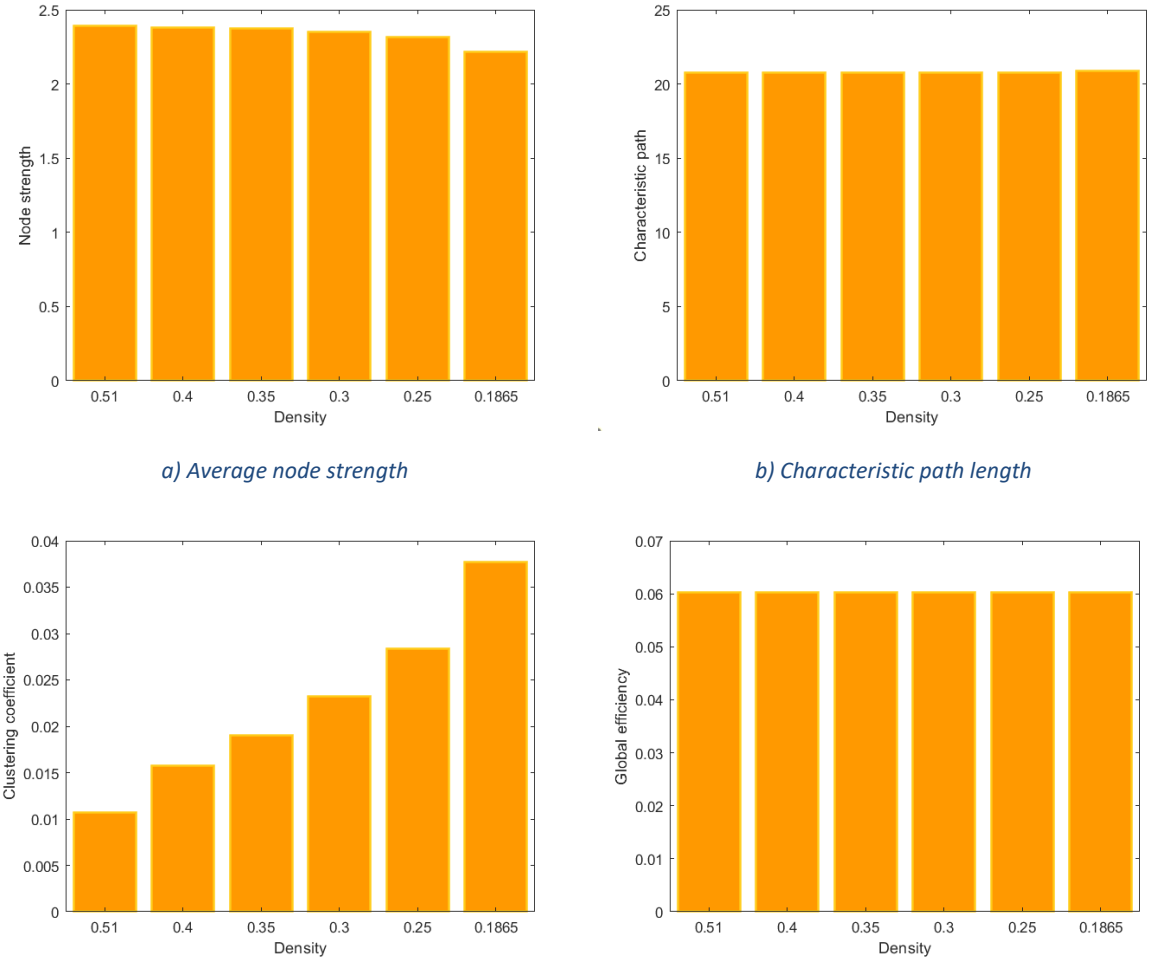
a) Average node strength

b) Characteristic path length

c) Global clustering coefficient

d) Global efficiency

Figure 38. Thresholding FA matrix: a) average node strength, b) characteristic path length, c) global clustering coefficient and d) global efficiency for weighted control matrix across different thresholds



a) Average node strength

b) Characteristic path length

c) Global clustering coefficient

d) Global efficiency

Figure 39. Thresholding NF matrix: a) average node strength, b) characteristic path length, c) global clustering coefficient and d) global efficiency

In the next section the outcomes shown in the Figures 38 and 39 are discussed for each index computed:

1. Node strength

The average node strength for NF is only slightly decreasing as the threshold is increasing. This trend can also be attributed to the limited impact of the thresholded weak connections, as previously mentioned.

It is known that the weights are spanning over 6 orders of magnitude in matrices derived by reconstructing number of streamlines (i.e., NF) in mammalian cortex [54]. Considering that NF matrix is composed by mainly low connection weights, as previously mentioned, their effect in computation of average strength is non negligibly lower comparing to multiple orders of magnitude weights of strongest connections. Hence, node strength remains almost invariant to thresholding.

However, this is not the case regarding the FA matrix, where for the same reason of specific weights distribution mentioned earlier, node strength is showing downward trend with increasing threshold.

2. Characteristic path length

Characteristic path length value for FA and NF matrices also shows diverse trends when computed across the thresholds. As mentioned in Par 2.2.4, in the computation of path length related indexes, connection lengths are commonly associated to the inverse of connection weights [17]. Hence, a path of stronger connections (higher weights) is assumed to be shorter. As a result, the shortest path between two nodes is characterized either by a low number of leaps, or by high weights, or both. As a result, even if two regions were connected by a direct weak edge, a path composed of more strong connections will be shorter [64].

This fact provides an explanation why path length for NF matrix is invariant to thresholding – since weak connections are effectively bypassed during the metric computation removing them has no effect and thus path length remains constant. Proven results are in coincidence with findings by [64]. Conversely, path length for FA matrix shows slight increase as threshold is increased. This can be attributed to the fact that higher values of threshold imply deletion of stronger weights, having in mind the weighs distribution in this case.

3. Clustering coefficient

Comparing the thresholding results for FA and NF matrices, it can be seen the trend of clustering coefficient is opposite. In case of FA, the global clustering coefficient is decreasing as the threshold is increasing, while for NF the opposite holds.

The clustering coefficient for weighted network was calculated by the formula given in the work by Onnela et al. [58] (Par. 2.2.2). Here, the edge weights are normalized by the maximal weight, $w_{ij} = w_{ij} / \max(w_{ij})$, and the contribution of each triangle depends on all its edge weights [82].

Concerning Figure 37, it is confirmed that the NF matrix is structured so that the majority of entries corresponds to extremely low weights, while the number of large weights (98^{th} percentile is 0.15) is very low. Hence, removing the low weights by thresholding will show a negligible contribution, since triangles composed by very low weights have insignificant contribution to the clustering coefficient. On the other hand, the thresholding process affects the node degree, k_i, which is decreasing as links are removed. Observing that node degree appears in the denominator in global clustering coefficient formula, this will lead to an overall effect of increase in clustering coefficient.

FA matrices' structure is, unlike NF matrices, characterized by higher abundance of large weight connections (98^{th} percentile is 0.8). Since in this case, according to the formula by Onnela et al. [58], thresholding means removing triangles with weights whose contribution to the clustering coefficient value cannot be considered as negligible, overall effect will be instead decreasing trend of clustering coefficient.

It is worth remarking that the tractography procedure always implies at least a low threshold, with the goal of ensuring accurate and efficient outcomes. E.g., a threshold (typically around 0.2) is set on FA, bellow which propagation of the pathway is stopped [39]. And this is ahead of considering more selective thresholding for the connectivity analysis. Appropriate choice of FA threshold for tractography analysis of neurodegenerative diseases is known to be essential [83]. Study by Kunimatsu et al. [84] suggests that optimal trackability threshold for corticospinal tract (CST) tractography is 0.2 Another study [83] stated that for uncinate fasciculus tracing optimal FA threshold value is in range between 0.15 and 0.20. It can be assumed that putting threshold on FA value will contribute largely to the distribution of weights in connectivity matrix, and these properties will propagate trough further analysis. The problem is even worst considering the SN metrics, where the effects of FA thresholding are superimposed to further constraints given to the minimal length of counted fibers. In addition, quantitative tract measures are highly influenced by geometrical factors (extension of connected areas, their distance, interference of other crossing tracts, etc.), which explains the large range of weights, which makes further thresholding even more critical [42].

4. Global efficiency

Since global efficiency is calculated as the reciprocal value of the harmonic mean of the path lengths [17], it is expected that for both FA and NF matrices it will display opposite behavior to characteristic path length, which can be evidenced in Figures 38 and 39.

4.2.4 Local differences analysis

In this section, network metrics are compared on local level. Values of metrics are displayed across the lobes of left and right hemisphere, respectively. Firstly, comparison was made between the original control matrix and the stroke case. Next, the control matrix was thresholded to drop its density to that of the patient. All local indices were analyzed in this way and the results are displayed in Figures 40-47. Furthermore, this procedure was performed for both FA and NF matrices for the purpose of visualization of thresholding effect differences on the local level in case of diverse weight distributions.

Strength

a) FA matrix

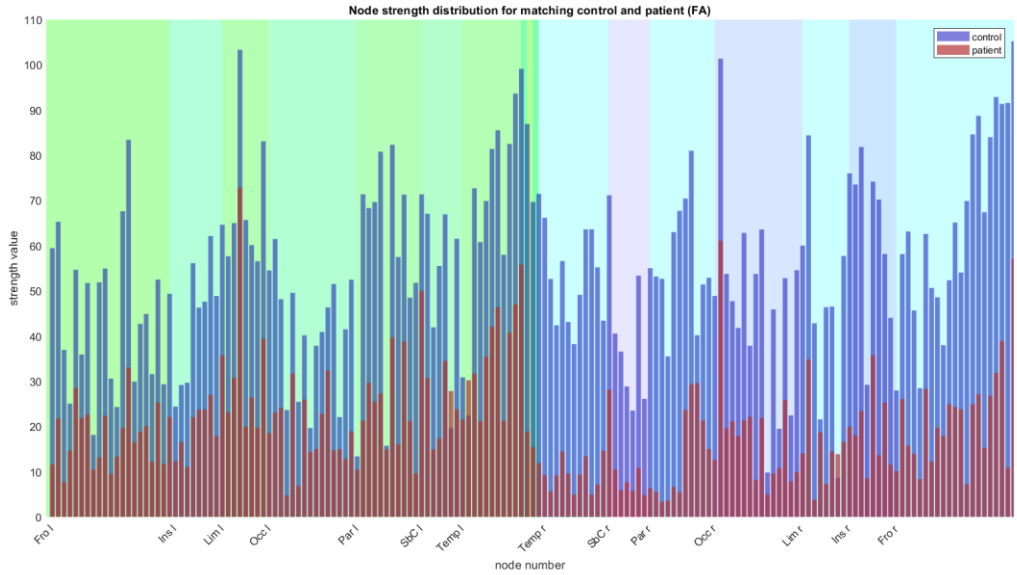


Figure 40. Node strength distribution: control (blue) and patient (red) (FA)

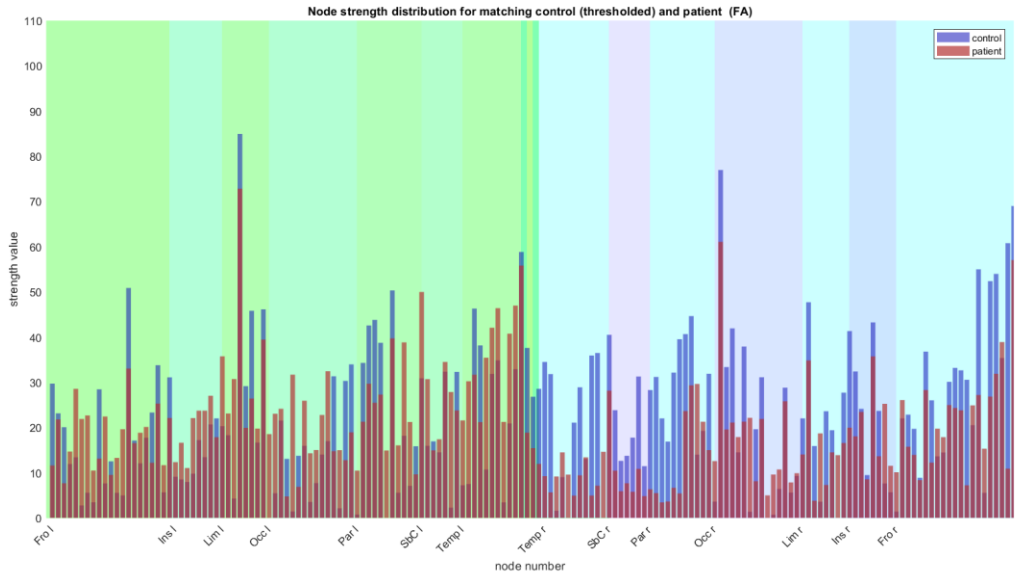


Figure 41. Node strength distribution: thresholded control (blue) and patient (red) (FA)

b) NF matrix

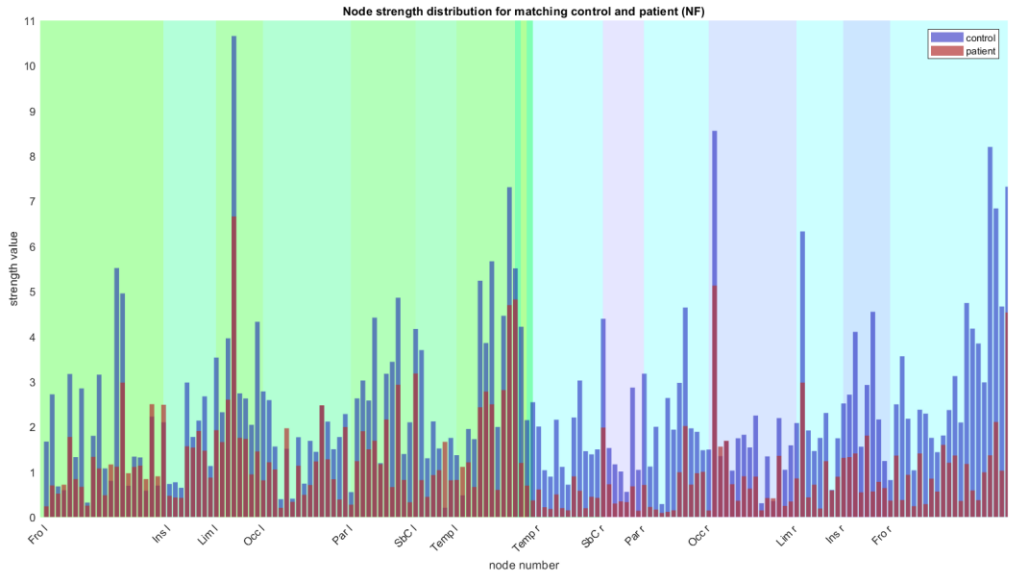


Figure 42. Node strength distribution: control (blue) and patient (red) (NF)

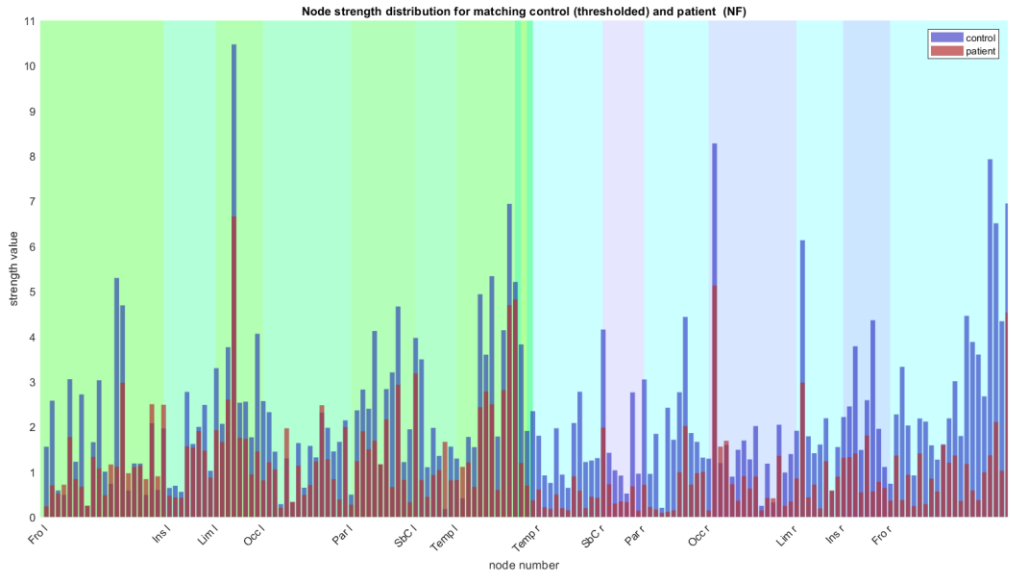


Figure 43. Node strength distribution: thresholded control (blue) and patient (red) (NF)

1. Clustering coefficient
   a) FA matrix

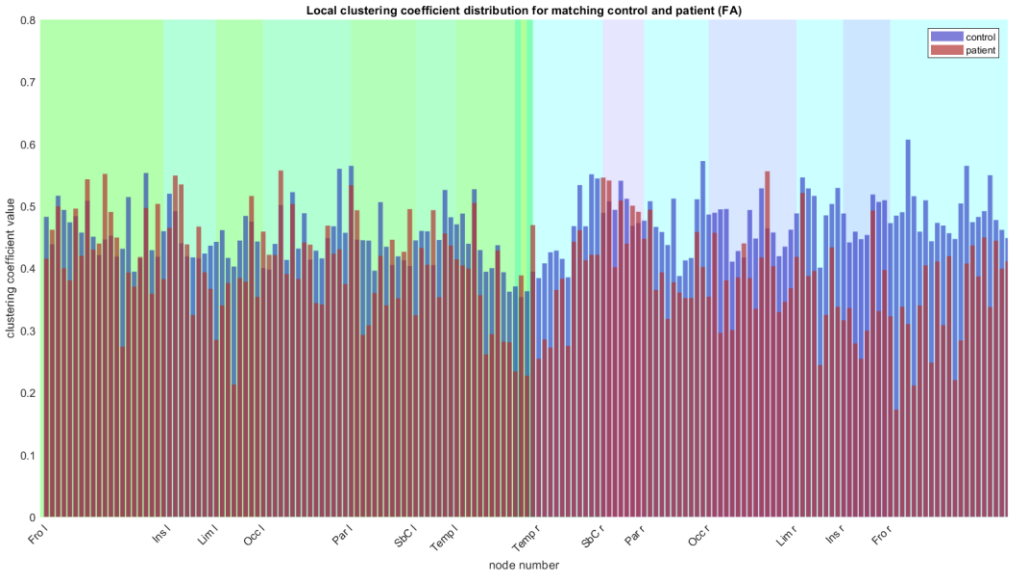


Figure 44. Clustering coefficient distribution: control (blue) and patient (red) (FA)

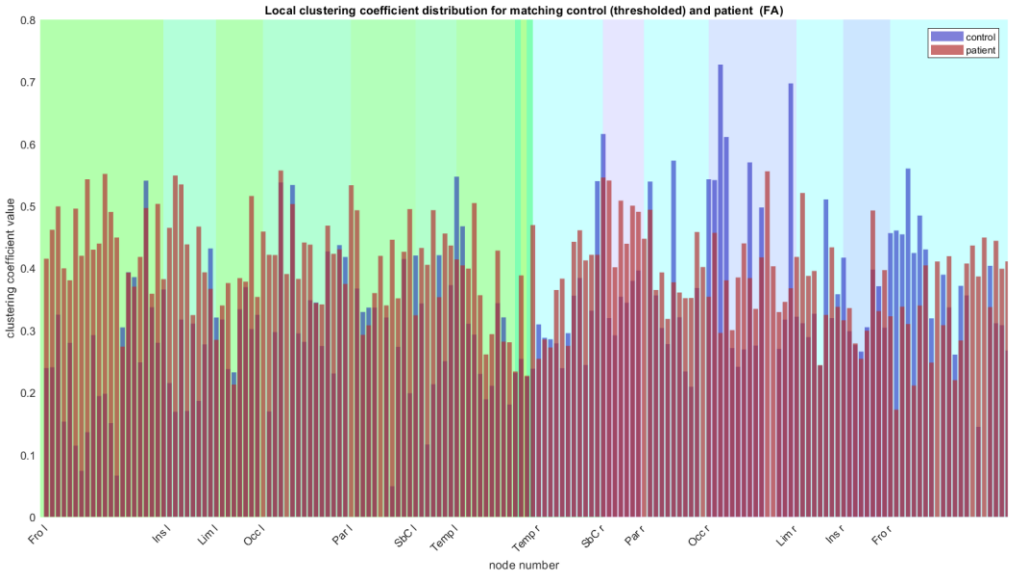


Figure 45. Clustering coefficient distribution: thresholded control (blue) and patient (red) (FA)

b) NF matrix

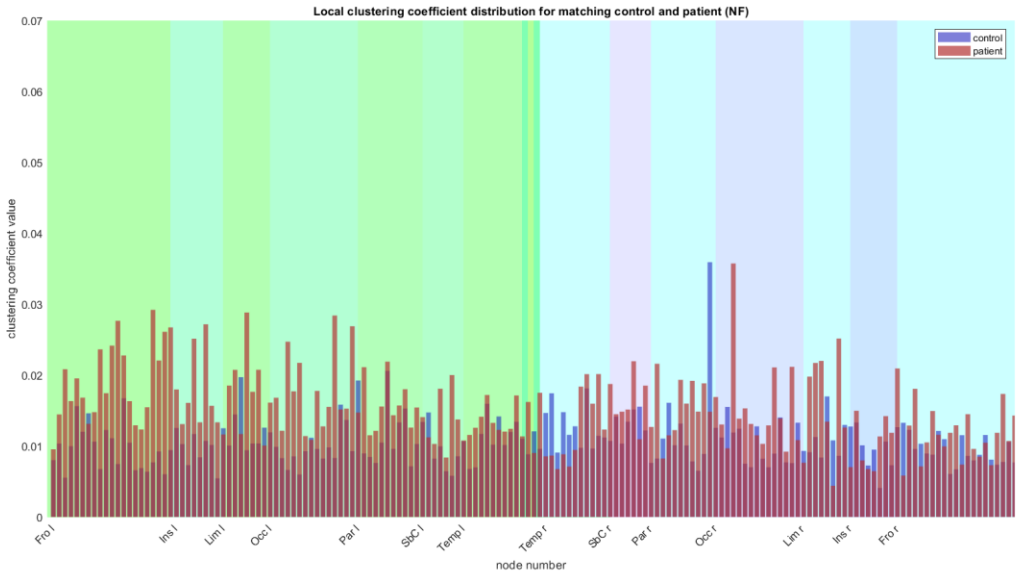


Figure 46. Clustering coefficient distribution: control (blue) and patient (red) (NF)

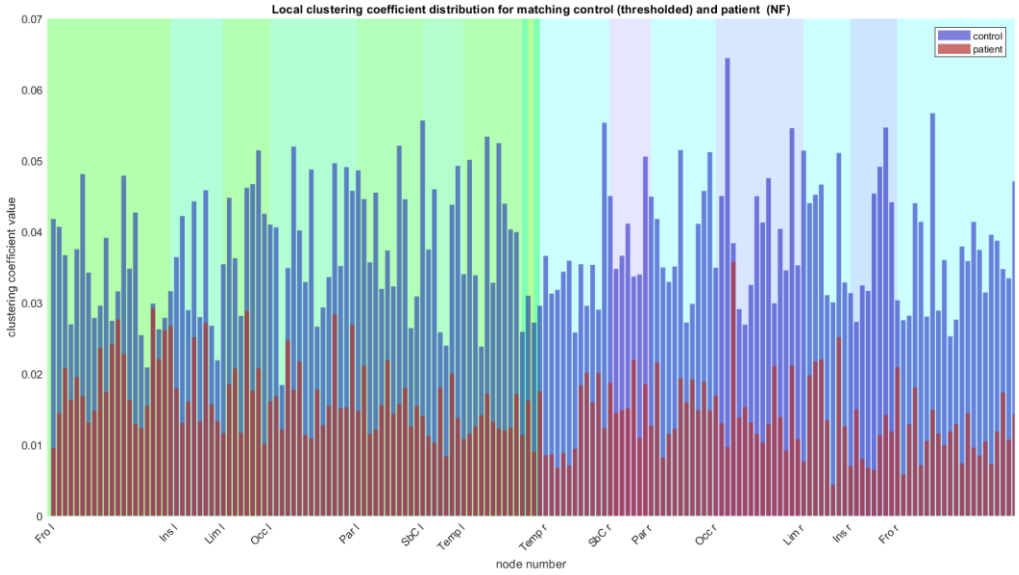


Figure 47. Clustering coefficient distribution: thresholded control (blue) and patient (red) (NF)

In the next section the outcomes shown in the previous figures are discussed for each local index computed:

1. Node strength

In Figure 40 highest difference in strength between the patient and the control is found in in the right hemisphere (lesioned) and spans across the temporal region, sub-cortical structures, and the parietal region. This analysis confirmed the brain regions where connectivity was compromised. Figure 41 shows that this difference persists even upon thresholding the control matrix to the density of the patient. On the other hand, in case of NF matrix node strengths show no visible differences before (Figure 42) and after thresholding (Figure 43) which is in accordance with the previously described weight distribution (Par. 4.1.2) for the NF matrices. Thresholding is removing low weight connections which are much more abundant compared to only few large weight connections, hence thresholding has no overall effect to node strength values.

1. Clustering coefficient

Figures 44 and 45 show that thresholding of the FA control matrix leads to the decrease of clustering coefficient. Most visible differences are across the frontal region in the right hemisphere, and they are maintained even after thresholding the control matrix. This is, again, in line with the fact that one of the regions affected by the lesion is the right frontal region. In Figures 46 and 47, it can be observed that the clustering coefficient trend for NF matrix is opposite to the one of FA matrix (Par. 4.2.3). Here, thresholding leads to an increase in clustering coefficient, which again can be attributed to the specificity of weight distribution in the NF matrices as previously described (Par. 4.1.2).

2. Local efficiency

Regarding the local efficiency analysis, results showed some changes after thresholding the FA matrix, in the control subject, but overall trend remained the same. The highest discrepancy was maintained across the right frontal lobe. On the contrary, in the case of the NF matrix, results showed an abrupt change in control local efficiency after thresholding. Local efficiency for the control was increased more than two-fold after thresholding.

4.2.5 Sensitivity analysis

Sensitivity analysis was performed on control matrix according to the procedure described in Par 3.2.1 (Experiments and analysis).

Prior to performing the sensitivity analysis, we inspected which nodes have the highest local degree value in the binary network, and analogously highest average node strength in the weighted network. For the binary network, nodes with the highest degrees were:

1. BSt - Brain stem
2. lh.Tha - Thalamus
3. lh.Pu - Putamen
4. lh.PerCaS - Pericallosal sulcus (S of corpus callosum)
5. rh.PerCaS - Pericallosal sulcus (S of corpus callosum)
6. rh.Tha - Thalamus
7. lh.CeB - Cerebellum
8. rh.CeB - Cerebellum
9. rh.Pu – Putamen
10. rh.CaN - Caudate nucleus

In weighted network, nodes with the highest node strengths were:

1. lh.PerCaS
2. rh.PerCaS
3. rh.Pu
4. BSt
5. lh.Pu
6. rh.Tha
7. rh.SupTS
8. lh.Hip
9. lh.MedOrS
10. lh.Tha

Subsequently, sensitivity analysis was performed on weighted characteristic path length, weighted global clustering coefficient, average node strength and small-worldness metrics. Results are displayed in Tables 7, 8, 9 and 10.

Node 1	Node 2	Value	lobes
lh. Pu	lh.InfCirInS	0.1351	Sbc/Ins
rh. Pu	BSt	0.1128	Sbc/Bst
lh.InfCirInS	lh.HG	0.1044	Ins/Tem
lh.Tha	BSt	0.1012	SbC/BSt
lh.InfCirInS	BSt	0.0970	Lim/BSt

Table 7. Sensitivity of weighted characteristic path length. Pu - Putamen; InfCirInS - Inferior segment of the circular sulcus of the insula; BSt - Brain Stem; HG - Heschl's gyrus (anterior transverse temporal gyrus); Tha - Thalamus; PerCaS - Pericallosal sulcu sulcus (S of corpus callosum)

Node 1	Node 2	Value	Lobes
rh. JS	rh.AOcS	$3.0839 * 10^{-5}$	Par/Occ
rh. JS	rh.InfTG	$2.7752 * 10^{-5}$	Par/Tem
rh. JS	rh.SupPL	$2.2379 * 10^{-5}$	Par/Par
rh. JS	rh.SbCG/S	$2.1168 * 10^{-5}$	Par/Fro
rh. JS	rh.CS	$2.0470 * 10^{-5}$	Par/Fro
rh.PosCS	rh. JS	$-1.8771 * 10^{-5}$	Par/Par
rh.MTG	rh. JS	$-1.9781 * 10^{-5}$	Tem/Par
lh.PerCaS	rh.PerCaS	$-2.1916 * 10^{-5}$	Lim/Lim
rh.AngG	rh. JS	$-2.2566 * 10^{-5}$	Par/Par
rh.SuMarG	rh. JS	$-2.6213 * 10^{-5}$	Par/Par

Table 8. Sensitivity of global weighted clustering coefficient. JS - Sulcus intermedius primus (of Jensen); AOcS - Anterior occipital sulcus and preoccipital notch (temporo-occipital incisure); CS - Central sulcus (Rolando's fissure); PosCS -.0 Postcentral sulcus; MTG – Middle temporal gyrus; PerCaS - Pericallosal sulcus (S of corpus callosum); AngG - Angular gyrus; SuMarG - Supramarginal gyrus

Node 1	Node 2	Values	Lobes
rh.PerCaS	lh.PerCaS	-0.0061	Lim/Lim
BSt	rh.CeB	-0.0048	BSt/CeB
lh.CeB	rh.CeB	-0.0042	CeB/CeB
BSt	lh.Tha	-0.0041	BSt/SbC
rh.CS	rh.PrCG	-0.0039	Fro/Fro

Table 9. Sensitivity of the average node strength. PerCaS - Pericallosal sulcus (S of corpus callosum); BSt - Brain Stem; CeB - Cerebellum; Tha - Thalamus; CS - Central sulcus (Rolando's fissure); PrCG - Precentral gyrus

Node 1	Node 2	Values	Lobes
lh.InfFGOrp	lh.MOcS/LuS	0.0089	Tem/Occ
rh.HG	rh.SupOcG	0.0089	Par/Occ
rh.JS	rh.AOcS	0.0089	Par/Occ
lh.PosTrCoS	lh.MOcS/LuS	0.0088	Occ/Occ
lh.MOcS/LuS	lh.InfFGOrp	0.0088	Occ/Fro

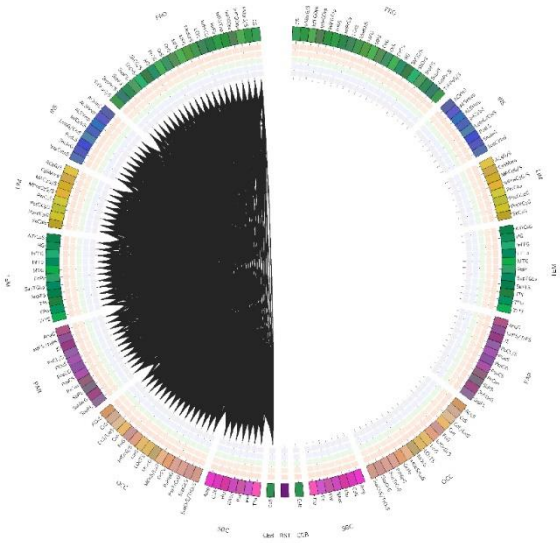
Table 10. Sensitivity small worldness measure. InfFGOrp - Orbital part of the inferior frontal gyrus; MOcS/LuS - Middle occipital sulcus and lunatus sulcus; HG - Heschl's gyrus (anterior transverse temporal gyrus); SupOcG - Superior occipital gyrus; JS – Sulcus intermedius primus (of Jensen); AOcS - Anterior occipital sulcus and preoccipital notch (temporo-occipital incisure); PosTrCoS - Posterior transverse collateral sulcus; MOcS/LuS - Middle occipital sulcus and lunatus sulcus; InfFGOrp - Orbital part of the inferior frontal gyrus

This analysis indicates which nodes are strategically important in the network. These nodes, whose deleting leads to valuable difference obtained in global measures, are the network hubs – most important nodes generally with the highest number of connections.

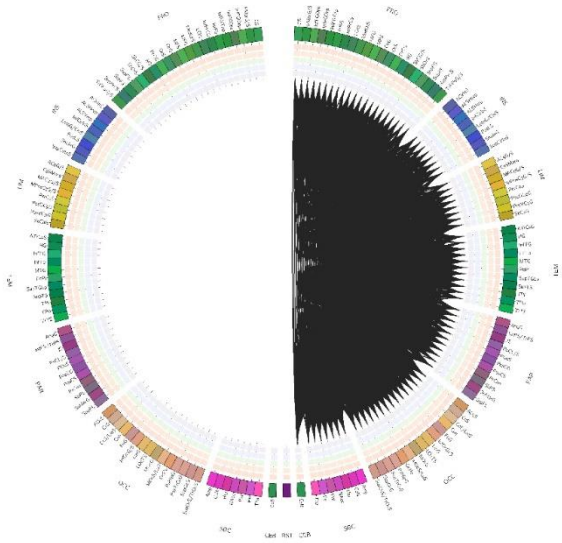
According to the work by van den Heuvel and Sporns [85] network analyses have consistently identified the precuneus, anterior and posterior cingulate cortex, insular cortex, superior frontal cortex, temporal cortex, and lateral parietal cortex as densely anatomically connected regions with a central position in the overall network. These results are confirmed by the sensitivity analysis presented in this work. Moreover, brain hub regions are more densely associated than indicated by their degree alone, resulting in the establishment of a highly interconnected ‘core' or rich club (Par. 2.2.7). This form of organization may have substantial functional significance by increasing the robustness of inter-hub connection and encouraging efficient communication and functional integration across the brain. Further neurophysiological investigations on the brain regions, starting from anatomical knowledge, are in progress with cooperation of neurologists and physiologists at Don Carlo Gnocchi Foundation.

4.2.6 Visualization of connectivity

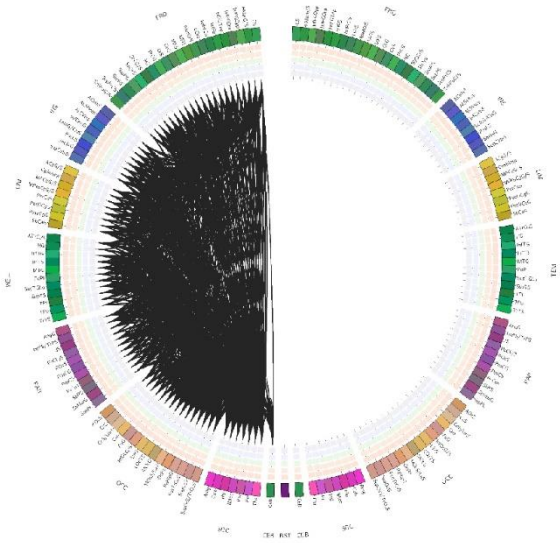
SPIDER-NET software was used to select and subsequently visualize connectivity in left and right hemispheres for both the control and the patient. Resulting connectograms, together with number of links in specific hemisphere, for both control and patient, are displayed in Figure 48.



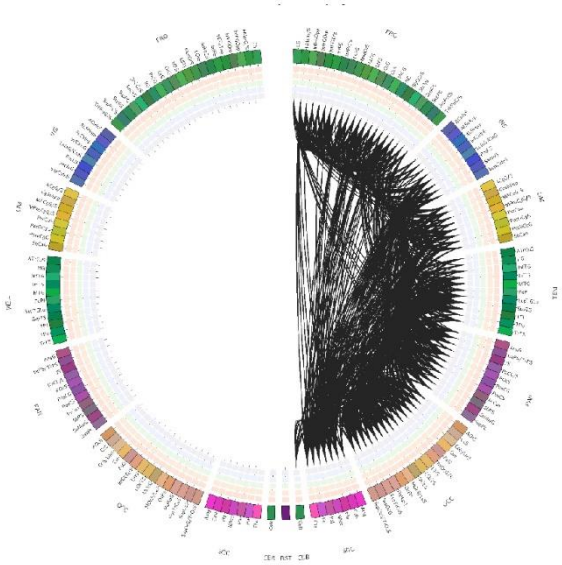
a) Control left hemisphere 2017 links



b) Control right hemisphere 2187 links



c) Patient left hemisphere 1193 links



d) Patient right hemisphere 805 links

Figure 48. Connectograms of the controls' and the patients' left and right hemispheres, intra-lobe connections included

Figure 48 displays visible differences in connectivity between control and patient for both hemispheres. Additionally, connections within the single lobes are included in this visualization. The difference between the control and the patient is more prominent for the right side, which can be expected due to the right hemisphere lesion. For the control, in the left hemisphere 2107 links were drawn, while in the right there were 2187. In contrast, for the patient 1193 links were drawn in the left hemisphere, while in the right there were 805 links.

Hence, previously described discrepancies in density between the subjects can be visualized in this representation. Additionally, comparing connectograms b) and d) it is visible that mainly connections from/to the frontal lobe are affected.

4.2.7 Comparison between left and right hemisphere

For both control and patient matrices corresponding to left and right hemisphere were extracted. Goal of this test was obtaining metrics differences between the hemispheres of a single subject but also the differences between the control and the patient.

Matrix division into two parts according to the hemispheric partition was made by using SPIDER-NET tool (SNT). This selection is graphically displayed in Figure 48. Brain Stem is excluded. Results are displayed in Table 11.

	Control (FA)	Patient (FA)	Control (NF)	Patient (NF)
Average node strength (L)	31.6714	19.2193	1.8998	1.2744
Average node strength (R)	33.6441	11.8211	1.8030	0.6918
Density (L)	0.6344	0.3592	0.6344	0.3592
Density (R)	0.6585	0.2423	0.6585	0.2423
Characteristic path length (L)	2.0489	2.3730	19.9705	23.3525
Characteristic path length (R)	1.9883	3.1656	18.5068	42.0641
Global clustering coefficient (L)	0.4690	0.4411	0.0152	0.0203
Global clustering coefficient (R)	0.4994	0.3755	0.0147	0.0158
Global efficiency (L)	0.5219	0.4615	0.0657	0.0553
Global efficiency (R)	0.5389	0.3642	0.0665	0.0336

Table 11. Values of metrics for the left and the right hemisphere

Regarding the values of measures obtained for the matrices or the control subject, as expected, there were minor differences in metrics between the left and the right hemisphere. In contrast to this, the values of measures for the stroke case significantly differed between the left and the right hemisphere.

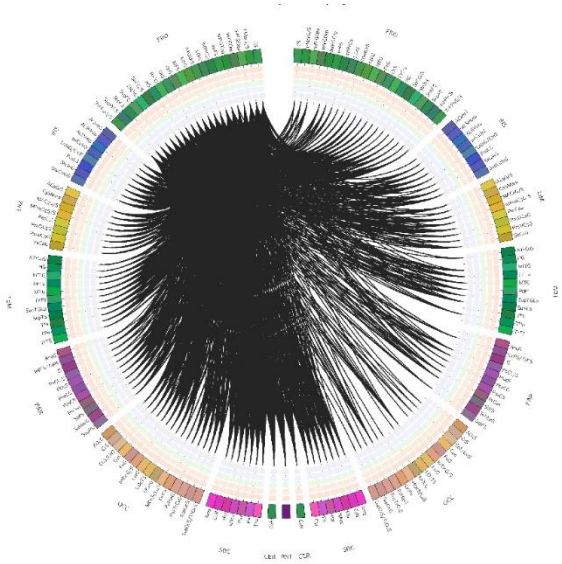
Additionally, by comparing values for each hemisphere between the two subjects, further insight was obtained. For example, differences in characteristic path lengths in the left hemisphere between the control and the patient is much lower compared to the difference seen in the right hemisphere. The same pattern can also be observed for all other measures computed.

These results advocate that the patient's connectivity is strongly compromised, especially in the right hemisphere, in keeping with this localization of the, spanning multiple lobes. It is also clearly visible how the right damage is also projected to the left hemisphere via long range connections, although the differences in connectivity metrics are less evident than in the lesioned side. Hence, intra-lobe connections are the ones mostly affected, which results with network measures for affected hemisphere showing remarkable discrepancies compared to the corresponding measures for the healthy control.

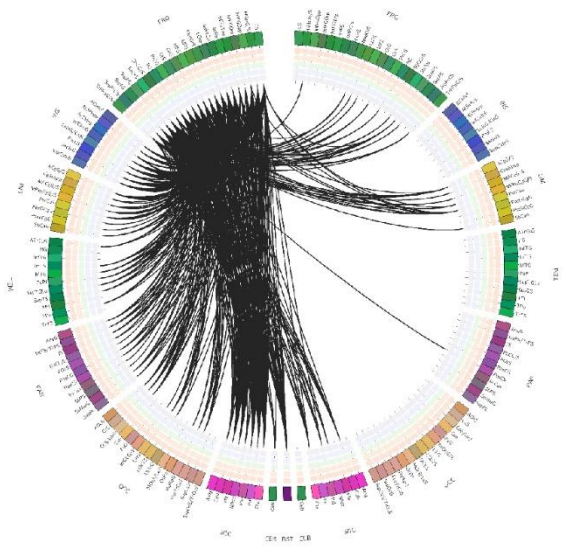
4.2.8 The effect of lesion to the brain connectivity

The stroke patient observed in this study is known to have a lesion in the right hemisphere, localized among the frontal, temporal, and parietal lobe. The goal of this analysis is to visualize the differences in connectivity between the two subjects (control and patient) across the brain regions and try to localize the lesion by relying on visual representation given by connectograms.

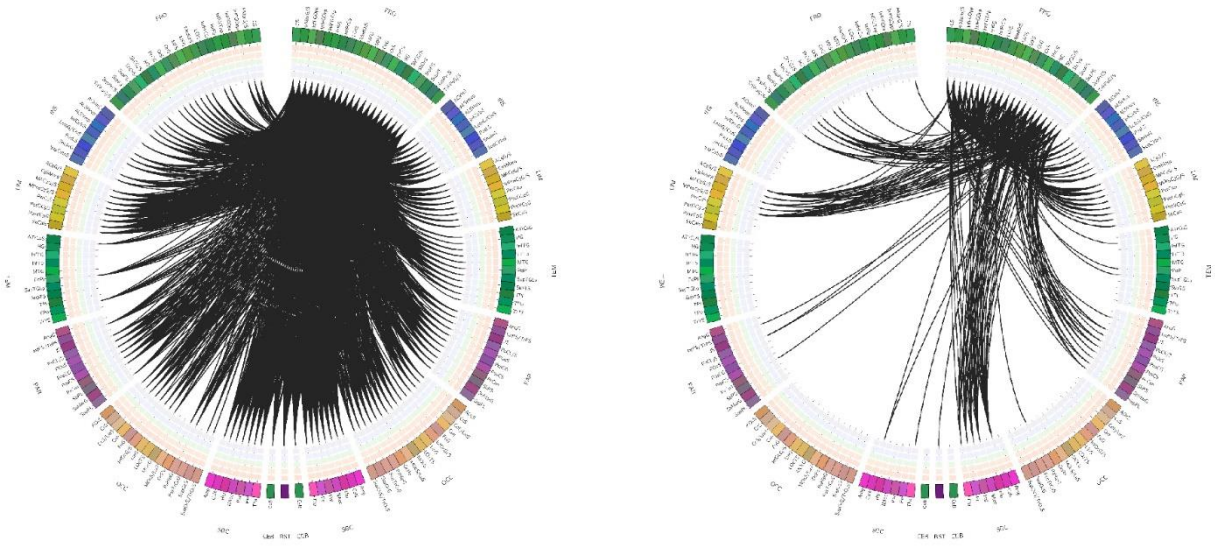
1. Connections from/to frontal lobe



a) Control Fro L -all 1352



b) Patient Fro L - all 487



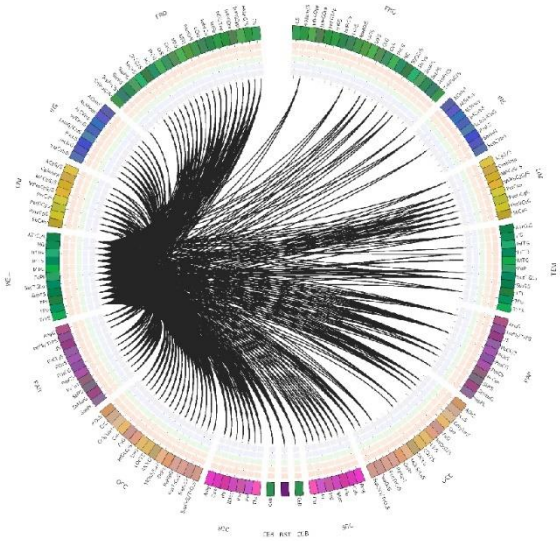
c) Control Fro R -all 1051 *d) Patient Fro R - all 279*

Figure 49. Connectograms displaying the connections from/to the frontal lobe

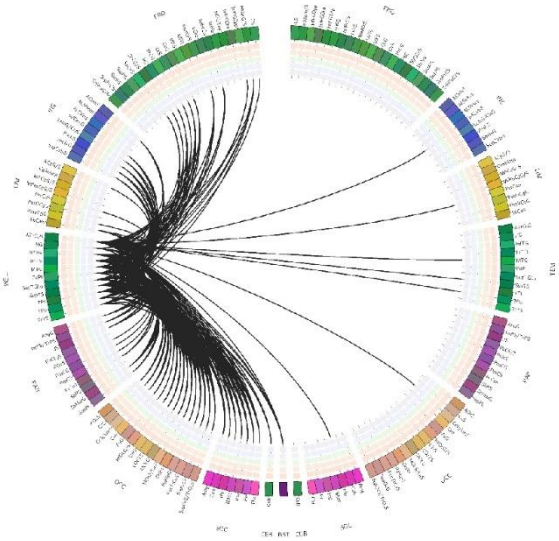
In Figure 49, connections from or to the frontal lobe are displayed (taking into consideration that the matrix is undirected). Selection of submatrix is made through *explorative mode* of SNT software, by selecting Frontal lobe as starting point (exploration source) and leaving target *undefined.* This means that all other gross parcels will be observed as target of connections using the Frontal lobe as priority region.

By comparing these figures it can be concluded that the overall connectivity in the stroke case is lower compared to the connectivity of the control. This can also be expected considering that the connection density of control is more than twice larger than the density of the patient. Furthermore, even though the left hemisphere of the patient is not affected directly by the lesion, still the discrepancy in connectedness between the patient and the control persists. Thus it can be concluded that the lesion is not disrupting only the connectivity in the hemisphere where it is located, but also affects the intra-hemispheric connections. Additionally, the difference in connectivity is present also considering the number of links within the frontal lobe. For control number of links is equal to 129, while for the patient it is 82 suggesting interruption of intra-lobe connectivity.

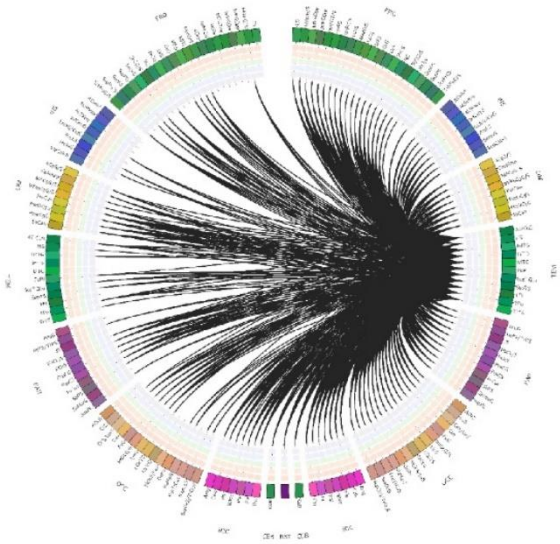
2. Connections from/to temporal lobe



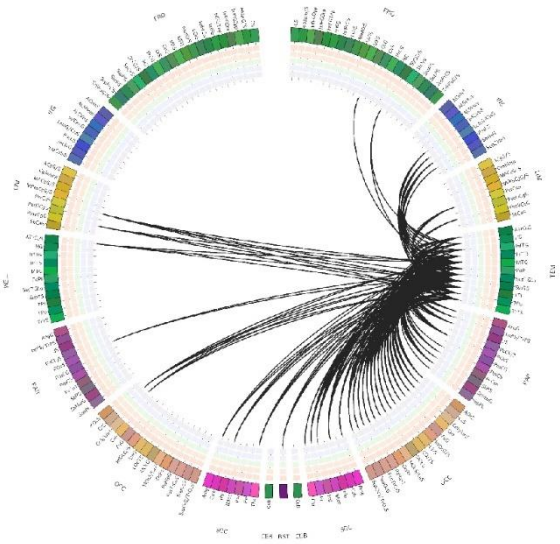
a) Control Temp L - all 701



b) Patient Temp L - all 303



c) Control Temp R - all 780



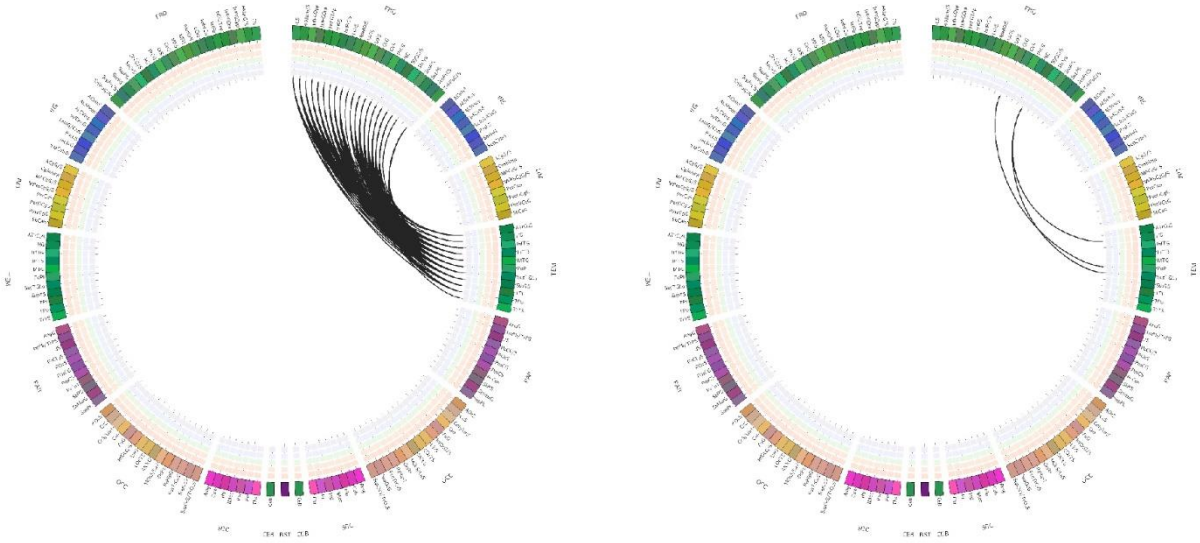
d) Patient Temp R - all 246

Figure 50. Connectograms displaying the connections from/to the temporal lobe

The analysis of the connections from the frontal lobe was also extended to a comparison between the two subjects. An extraction of the relevant submatrix was performed in the same way as for the Frontal lobe. Results are displayed in Figure 50.

Similarly, the connectivity was compromised in both hemispheres regardless of the lesion location. Furthermore, comparing the differences in Figure 50 a) and c) and Figure 50 b) and d) it can be appreciated how a large majority of the inter-lobe connections from the temporal lobe was disrupted. Conversely, comparing these results with the decrease in the intra-lobe connections (control: 43 connections, patient: 34 connections), it can be concluded that connections within the temporal lobe were less affected.

3. Connections between temporal and frontal lobe



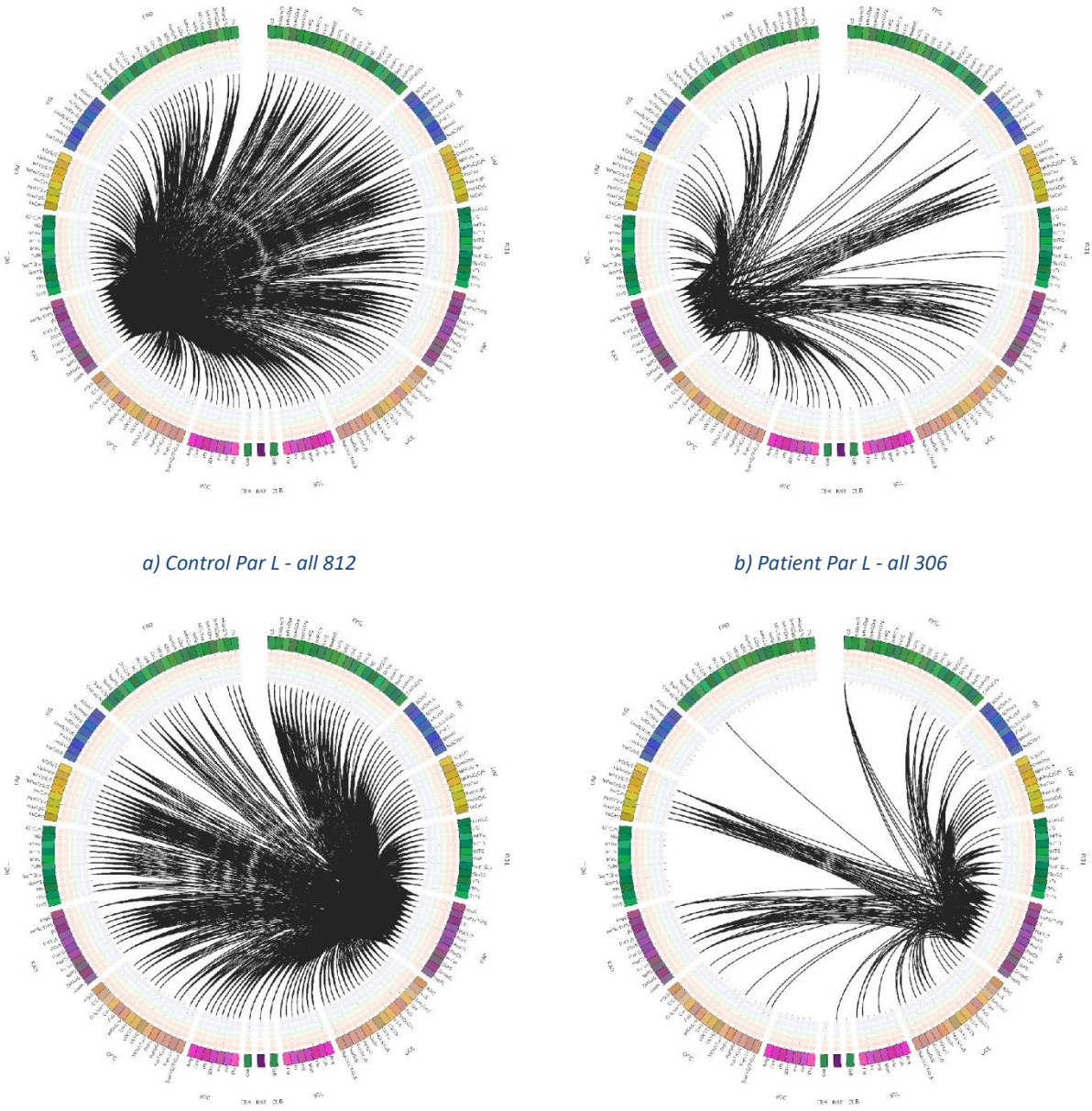
a) Control Temp R – Fro R 113 *b) Patient Temp R – Fro R 3*

Figure 51. Connectograms displaying the connections the temporal right and the frontal right lobe

Figure 51 provides a clearer look at the connections between the temporal and the frontal lobe in the right hemisphere. In this way, the already described pattern from Figure 50 (generally lower number of inter-lobe connections) is more visible for this specific selection of lobes. Here, the ease of visualization provided by a selection of interhemispheric parcels in this novel software can be valued.

Here, submatrix is selected trough *extraction mode* of SNT, whereby selecting Temporal and Frontal lobe in the right hemisphere, common edges between mentioned lobes are drawn.

4. Connections from/to parietal lobe



a) Control Par L - all 812

b) Patient Par L - all 306

c) Control Par R - all 820

d) Patient Par R - all 251

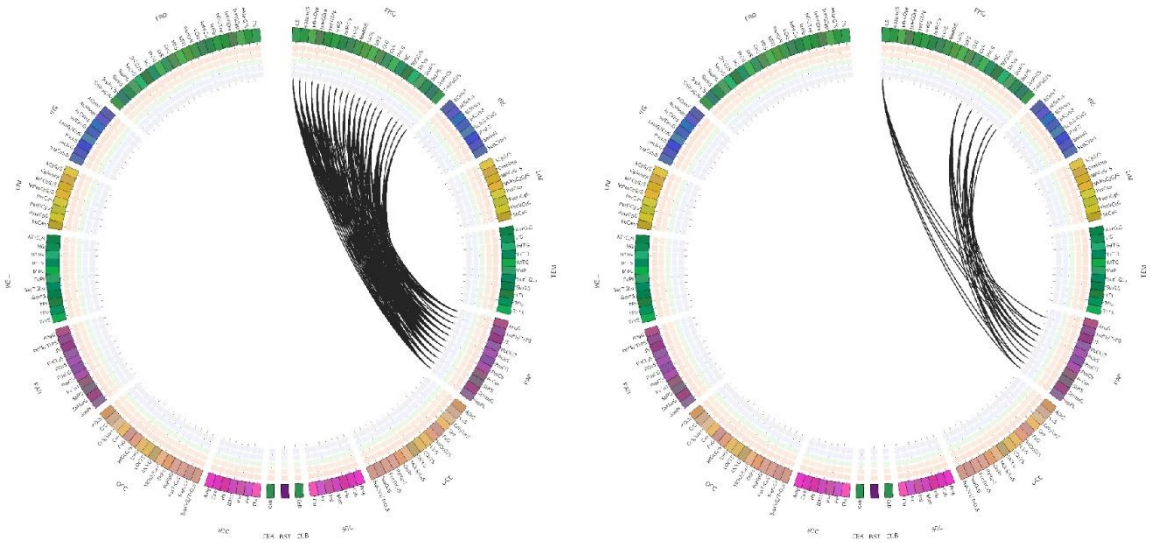
Figure 52. Connectograms displaying the connections from/to the parietal lobe

After the *exploration* of connections from Frontal and Temporal lobe, procedure was repeated using Parietal lobe as exploration source, since it is also known to be affected by the lesion.

By comparing number of connections from right Parietal lobe, Figure 52 c) and d), it can be concluded that the discrepancies in connection numbers were expectedly higher than in the left side of the brain. Furthermore, comparing connectograms c) and d), it can be seen that mostly compromised connections were the ones between right Parietal and left Frontal lobe, where only one link were present in connectogram corresponding to the patient, d). It can be also quickly visualized that, the link between right Parietal lobe and both left and right hemisphere Cerebellum was disturbed.

Intra-hemispheric connections were 42 and 36 in the control and the stroke subject. Hence, it can be said that intra-hemispheric connections within Parietal lobe had negligible damage.

5. Connections between Parietal and Frontal lobe



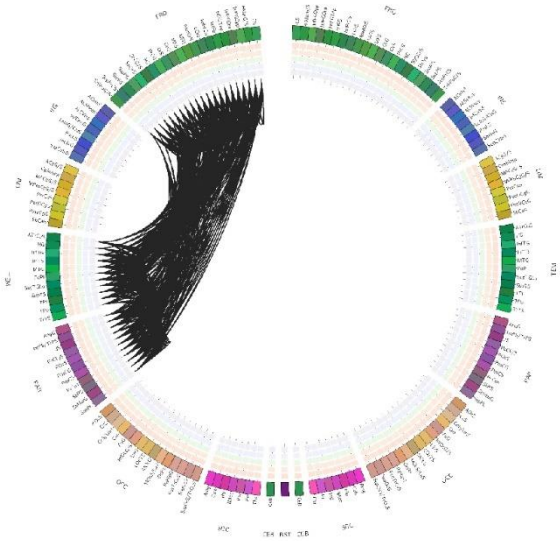
a) Control Par R – Fro R 119 *b) Patient Par R – Fro R 55*

Figure 53. Connectograms displaying the connections from/to the parietal lobe

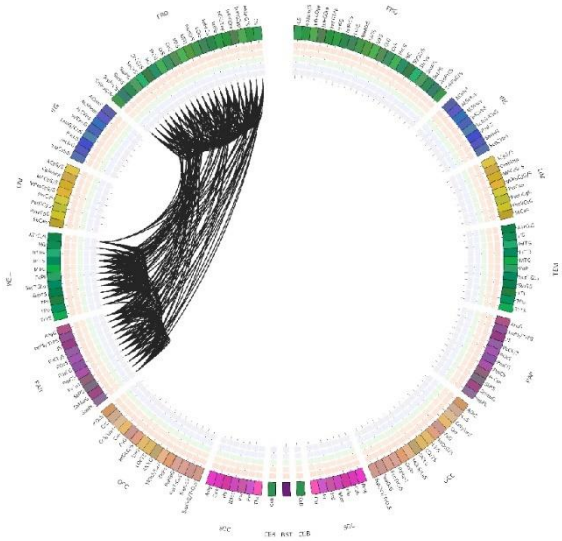
Next, connections between Parietal and Frontal lobe in the right hemisphere were explored. Resulting connectograms are displayed in Figure 53. Selection of submatrix was made through *extraction mode* in the software, by selecting Parietal and Frontal lobe in the right hemisphere.

Number of connections in the control subject connectogram was 119 and was two-fold lower, 55, in the stroke one. Zooming into connectogram b), and by comparing with the control connectogram a), links connecting lesioned areas are effectively enhanced.

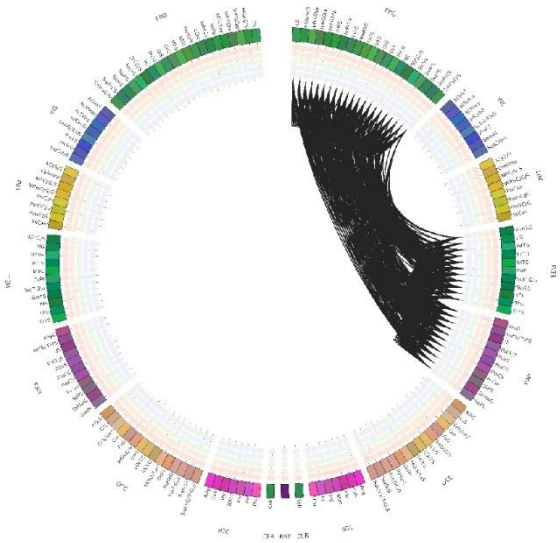
6. Connections between Temporal, Parietal and Frontal lobe (including connections within the lobe)



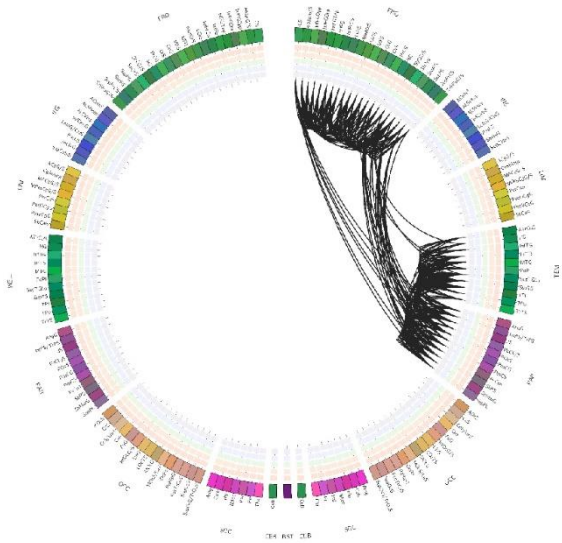
a) Control Tem-Par-Fro (L) 493 links



b) Patient Tem-Par-Fro (L) 261 links



c) Control Tem-Par-Fro (R) 563 links



d) Patient Tem-Par-Fro (R) 62 links

Figure 54. Connectograms of common links between Temporal, Parietal and Frontal lobe (including intra-lobe connections) for control and patient

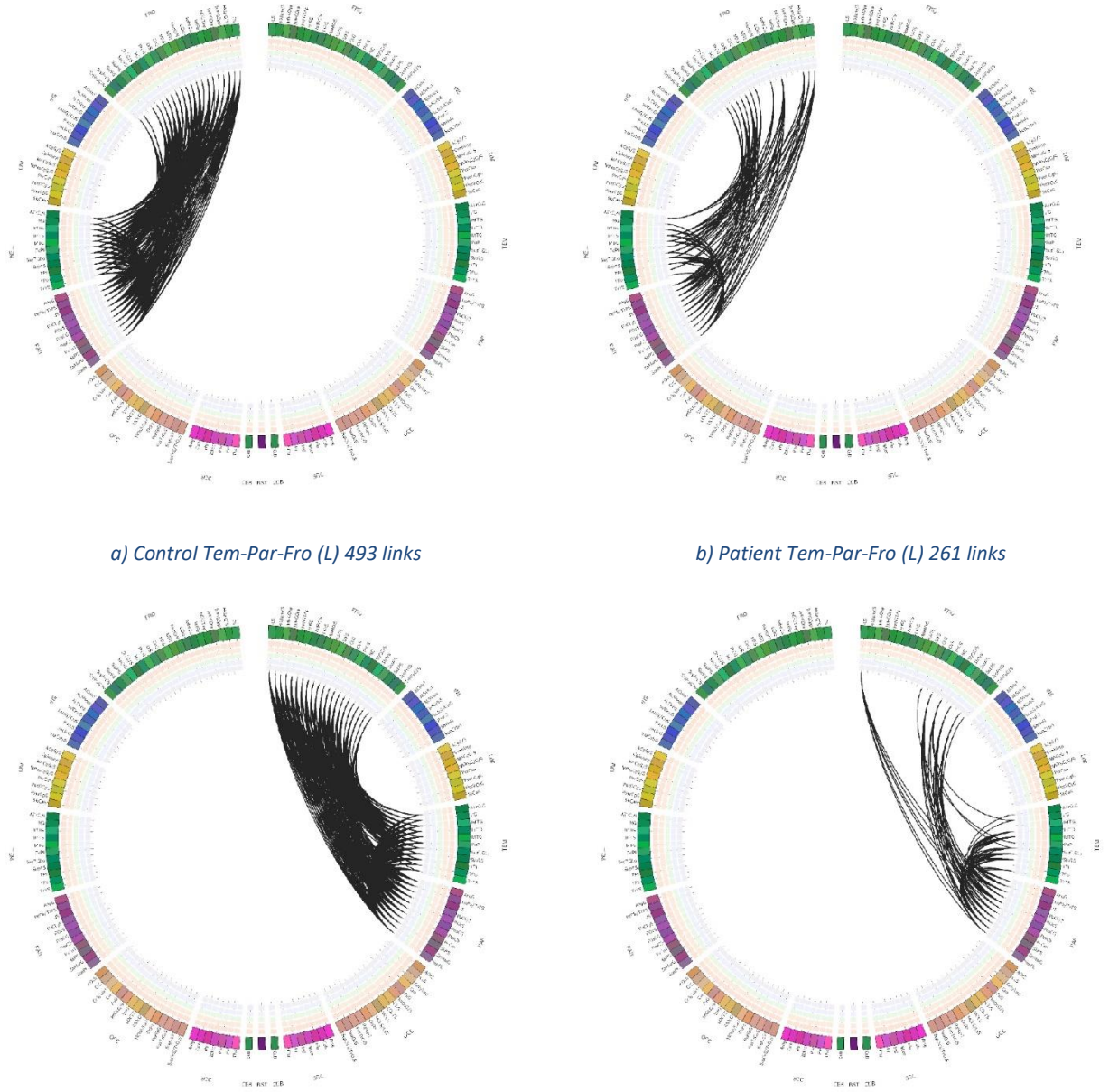
In the following analysis, common connections between Temporal, Parietal and Frontal lobe were extracted for both the left and the right hemisphere. Additionally, the connections within each of the lobes were included in the analysis. Resulting connectograms are displayed in Figure 54.

This selection was made by selecting the addressed lobes in either hemisphere separately, next using *extraction mode* of the software. Furthermore, by selecting the checkbox *keep connections within gross parcels*, provides intra-lobe connections to be included in the analysis.

In this way, a sub-matrix was extracted for the regions of interest and the relevant subgraph was displayed thus evidencing the outcomes of stroke on connectivity in the lesioned areas and the symmetrical ones.

In the Figure 54 a) number of connections in the controls' left hemisphere is 493 while for the patient number of connections is 261 (Figure 54 b)). Next, in Figure 54 c) connections among three mentioned lobes in the controls' right hemisphere are displayed. Number of connections is 563. In the contrast to this, Figure 54d) shows abrupt decrease in patients' connectivity between among these three lobes, number of connections is only 216. Furthermore, comparing c) and d) it can be confirmed, as previously said, that connections between diverse lobes are more obstructed than the ones within the lobe.

7. Connections between Temporal, Parietal and Frontal lobe (excluding connections within the lobe)



a) Control Tem-Par-Fro (L) 493 links

b) Patient Tem-Par-Fro (L) 261 links

c) Control Tem-Par-Fro (R) 563 links

d) Patient Tem-Par-Fro (R) 62 links

Figure 55. Connectograms of common links between temporal, parietal and frontal lobe (intra-lobe connections are excluded) for control and patient

The Temporal, Parietal, and Frontal lobe analysis was repeated excluding within lobe links.

Connectograms are displayed in the Figure 55. This representation provides clearer visualization of links between the lobes since they are not overlapped with more dense intra-lobe connections. Thus, sharp differences in connectivity can be more efficiently explored.

Regarding the control (Figure 55a), number of connections in the left hemisphere was 274, while the same lobes in case of the patient share 101 common connections. For the right hemisphere, Figure 55c) displays connections for the control, whose number is 322. On the contrary, patients' connections number between these 3 regions was abruptly lower – which can be clearly seen in Figrue 58 d), number of connections is only 62.

4.2.9 Results summary

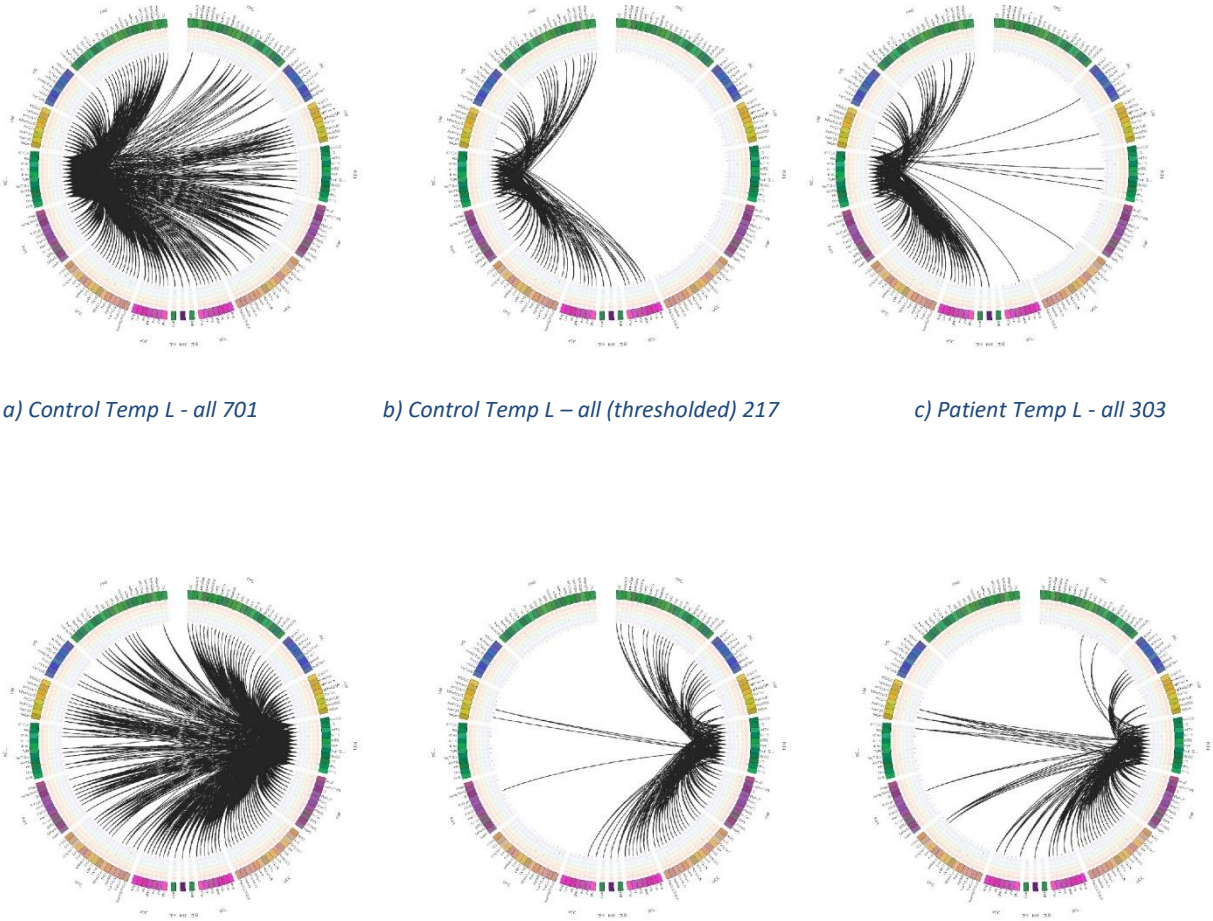
All the results from Chapter 4.2.9 are summed up in Table 12.

Connections	Control	Patient
Fro L – all	1352	487
Fro R – all	1051	279
Fro R – Fro R	129	82
Tem L – all	701	303
Tem R – all	780	246
Tem R – Tem R	43	34
Tem R – Fro R	113	3
Par L – all	812	306
Par R – all	820	251
Par R – Fro R	42	36
Par R – Par R	119	25
Par R – Tem R	90	36
Tem L – Par L – Fro L (intra-lobe included)	493	261
Tem R – Par R – Fro R (intra-lobe included)	563	216
Tem L – Par L – Fro L	274	101
Tem R – Par R – Fro R	322	62

Table 12. Number of edges in selected submatrix

4.2.10 Density-thresholding outcomes

Since the effects of density thresholding on computed network measures were focus of this work, lastly, it was explored how this thresholding method affects the connectivity representation given by connectograms. Results are displayed in Figure 56.



a) Control Temp L - all 701 *b) Control Temp L – all (thresholded) 217* *c) Patient Temp L - all 303*

d) Control Temp R - all 780 *e) Control Temp R - all (thresholded) 239* *f) Patient Temp R - all 246*

Figure 56. Thresholding in connectograms

A connectogram displaying the links from/to right and left Temporal lobe for the control and the patient was used for studying the effects of thresholding.

Previously employed thresholding strategy, called density thresholding, was applied to the connectivity matrix prior to extraction of the submatrix and visualization of connectogram. After thresholding process, density of control matrix was equal to the density of the matrix corresponding to the patient.

Then original matrix and the selected submatrix after the thresholding process are compared in Figure 56. Observing Figure 56 a) and b) it can be clearly seen that none of the connections between the two hemispheres have survived thresholding process. Similar features can be seen for the Figures d) and f), where only 3 connections are present between the hemispheres. By comparing these connectograms to the ones of the patient, it can be concluded that the density thresholding process is canceling out valuable differences between the two subjects.

Overall, connectivity pattern after thresholding is largely changed. Hence, it should be noted that density is an important feature of a specific network. Moreover, many network measures are dependent on density. Thus, metrics computed for thresholded network may not be fully representative of the connectivity in that network.

It can be concluded that thresholding control matrix in this way, and thus changing its density with the goal of making more suitable comparison between subjects of two groups, has the drawback of removing valuable knowledge for identifying abnormalities.

4.3 Borderline intellectual functioning study

4.3.1 T-test results

Two samples t-test was performed twice, firstly on all connections, following the test on connections with weights lower than 0.2 that are considered as weak. Results are displayed in Table 13.

Coefficient	All connections	Weak connections
Strength	h = 0; p = 0.0667	h = 1; p = 0.0071
Clustering coefficient	h = 0; p = 0.2089	h = 0; p = 0.4315
Characteristic path length	h = 0; p = 0.0577	h = 1; p = 0.0481
Global efficiency	h = 1; p = 0.0434	h = 1; p = 7.1157e-47
Density	h = 1; p = 0.0089	h = 1; p = 0.0076

Table 13. T-test results

Regarding the test results for all connections, it can be seen that null hypothesis was rejected only for global efficiency and density, hence these are only two network metrics based on which statistically significant difference between control and patient group can be drawn if all network edges are taken into consideration.

On the other hand, by making test solely on the weak connections (meaning all connections with weights higher than 0.2 are discarded), statistically significative difference between control and patient group can be drawn regarding global efficiency and density, but also the characteristic path length and strength.

Results displayed in Table 13 suggest that constructive discrimination can be made between the two groups based on the weak connections. Thus, these results were the motive for further exploration of weak connections utility and reliability.

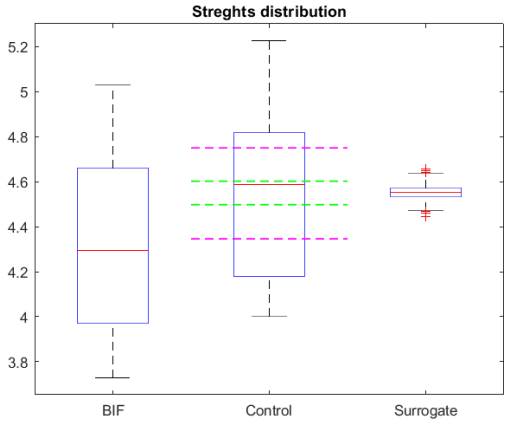
4.3.2 Bootstrapping

Bootstrapping method was employed to create surrogates of control subjects. Process of generating surrogate samples is described in Par. 3.2.3. It is expected that making 1000 new samples of controls will provide a certain narrower normality range where all the values of a network measure for control group are located. This means that would be possible to more precisely identify an interval to discriminate more reliably the two populations.

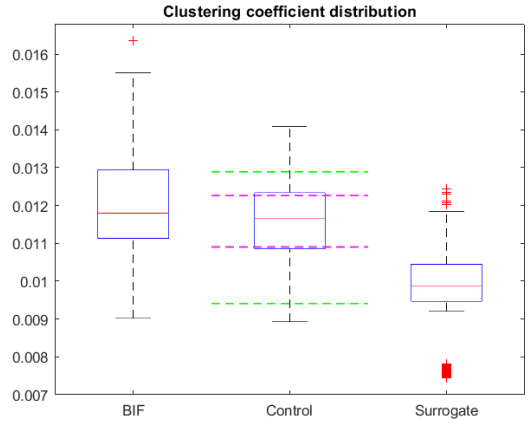
As an outcome of this analysis it should be verified that subjects from the BIF group will all form a cluster on the same side compared to the position of the surrogate confidence range. Depending on the network measure, more precisely on the fact that if it is lower or higher in case of disease, BIF cluster will be positioned higher or lower with the respect to the surrogate interval.

Distribution for all computed network measures for 32 subjects composing BIF group, 14 healthy controls, and 1000 surrogate samples is displayed in Figure 60. Moreover, confidence interval was calculated for control and surrogate groups, and it is overlapped to the plotted distribution of controls. Confidence interval for control group is outlined by pink, while for the surrogates lines are green. Next, all the data was normalized by mean value, however the obtained results did not valuably differentiate from the displayed ones.

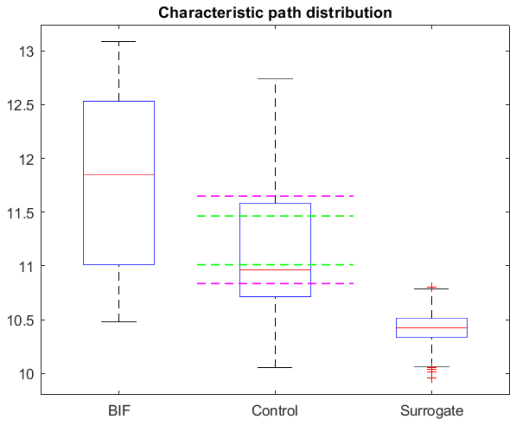
In this way the normality range of all three populations can be visually compared by obtaining their distribution properties.



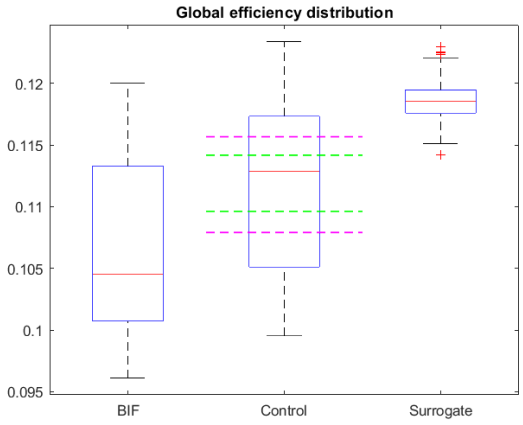
a) Average node strength



d) Global clustering coefficient



c) Characteristic path length



d) Global efficiency

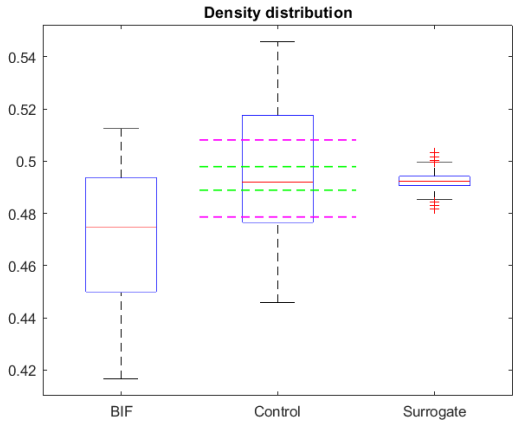


Figure 57. Distribution of values for three groups: BIF, Control and Surrogate for different network measures: a) Average node strength; b) Clustering coefficient; c) Characteristic path length; d) Global efficiency and e) Density. Confidence intervals of control and surrogate groups are denoted in pink and green, respectively.

From Figure 57, it can be seen that the mean values for clustering coefficient, characteristic path length and global efficiency are biased (Figure 57b, c and d)). Hence, it could be concluded that the method used for resampling should be revised and improved.

Furthermore, it would be expected that bootstrapping will narrow down the confidence interval for surrogates, compared to the one for controls. This is approved in all measures except for the clustering coefficient.

5 Conclusion

Research in the topic of the brain networks is growing rapidly in last decades. Accordingly, this growth of interest and research scope necessitates the new technical tools for conducting more efficient analysis. Priorly developed software tools have limited reach in terms of visualizing the brain connectivity. Namely, it is a growing issue that instrumentation equipment (e.g., MRI), and the next processing (e.g., connectivity analysis) is giving us more information than we can usefully visualize and handle. Hence, a bottleneck is envisaged to be tackled by suitable and agile software tools. It is indeed clear that even a non-dense fraction of order 100^2 links is hardly assessed without the support of suitable exploration tools giving meaningful visual portraits and quantitative indices.

The research and design presented in this thesis should provide a systematic approach for researchers interested in studying brain connectivity. This is based on the introduction of a novel software that we called SPIDER-Net. The overarching goal of the software design and development process was to create a tool that would provide a more convenient and advantageous means of exploring structural and functional brain connectivity. Focus in put on permitting the selection of connection subsets, thus facilitating a systematic exploration of brain networks.

Herby, visualization can be controlled by displaying connections emanating only from one hemisphere, one lobe, or from a single brain structure, thus allowing the user great leeway in the ability to modify the level of complexity being displayed in the connectogram. Taking into consideration the complexity and the number of connections between hemispheric parcels, the benefit of choosing the regions of interest and subsequent extraction of a corresponding subgraph is immediately clear and has been proven in present analysis.

More precisely, the software was tested on case study of stroke compared to a healthy control. Patient presented a lesion in the right brain hemisphere, which was further explored by employing new approach for visualization. Based solely on comparing connectograms between the patient and the control, major discrepancies can be observed in the connectivity patterns.

Moreover, not only the connections within the right hemisphere of the patient were compromised due to the presence of lesion, but also the inter-hemispheric connections. Number of connections together with plotted visualization were compared between the patient and the control for various selection of hemispheric parcels. This selection was based on the previously known information on brain areas that are particularly affected by the lesion.

Further test should also employ different atlases used for parcellations. Additionally, another possibility in employment of this software is checking the validity of connections in fiber pathways reconstructed by tractography.

Furthermore, as the software was developed and simultaneously tested, some of the functionalities remained uncovered by this work. For example, the ability to modulate the edge attributes (such as thickness and color) based on the degree of connectivity between nodes was not exploited. Also, the possibility to extract a subgraph based on functional attributes of parcels (e.g., a-priori functional knowledge or, conversely, the data-driven inclusion into a resting state network) was not explored, so far. This work validated the tool only upon anatomical attributes. In this regard, it is worth recalling that the fusion between functional and structural connectivity is a major open problem in current brain research.

Thresholding and Borderline Intellectual Functioning study.

Even though the binarization process was widely used in early applications of graph theory to brain data, it is inherently insensitive to architectural principles embodied in edge weights. This realization has more generally motivated the development of methods that can be used to concur sensitivity to the patters of the weights on the edges, and topologies present in weak versus strong weights. However, the role of edge weights in brain computations and higher-level cognition still has received less attention.

According to neuroanatomical data, the weights of structural connections may be dictated by developmental growth regulations, energy and metabolic constraints, and physical limitations on the volume of neural systems, particularly brains encased by bone.

The choice for weighted approaches was the main challenge in this thesis. Particularly in structural connectivity studies, weighing is generally considered as problematic, since DTI is inherently noisy and prone to artifacts, which causes derived weights to have limited reliability. Moreover, this work has highlighted the large impact of the chosen metrics, whether fractional anisotropy (FA) or streamline number (SN) or other. The most common approach in dealing with this uncertainty is thresholding, which should remove lower weight connections that are considered as false connections that are consequence of noise. Moreover, there is no universal standard and agreement over generalization of network measures from binary to weighted case, and also on the thresholding method.

Thus, accent in this work was put on the thresholding process, more precisely one of the most common thresholding methods, i.e., density thresholding. This analysis was performed on case study of stroke, since matrices corresponding to control and patient were available in two variants - connection strength was expressed by using FA or SN. From the obtained results it can be firmly concluded that thresholding is strongly dependent on the connection-strength

distribution in the connectivity matrix which is a significant determining factor in the behavior of weighted graph-theoretical metrics. Moreover, major discrepancies in thresholding outputs can be observed between the two mentioned matrix variants.

Generally, it can be seen that connectivity matrices constructed by SN as a measure for connectivity strength are more invariant to thresholding process since most commonly calculated network measures stay stable under this process.

The analysis of weighted connection was motivated by proofs that individual differences in strength of weak connections display individual differences in fluid intelligence and are strongly correlated with full scale, verbal, and performance IQ.

Hence, the dataset of Border Intellectual Functioning was considered as fitting for the further analysis made in this thesis.

Moreover, by performing simple statistical test this work shows valuable insights about the potential utility of the weak connections. However, the work is still in progress and further investigations are needed. Regardless of this, some considerations can still be made relevant to the hypothesis tested and in proposing experimental procedures, which are laid out in the following paragraphs.

In order to address the issue of unreliableness of connections in tractography data bootstrap method was applied. This method is substantially based on the law of large numbers, which in short says that with enough data the empirical distribution will be a good approximation of the true distribution. Hence, this approach could provide extracting more stable normality range of values of various network measures. Thus, further comparison between patient population and large size sample of simulated controls could provide well-grounded analysis of the differences between the groups.

However, the bootstrap protocol applied in this work introduced unexpected biases in the surrogate data compared to the original control group. Afterwards, data was normalized by the mean value, but results did not significantly change. Hence, as a future prospect, this method should be reviewed and enhanced in order to make the analysis more robust and valid. The advantage of bootstrap analysis is to provide data-driven confidence intervals, without the need of a-priori hypotheses, and yield a solid ground for between group comparisons. However, our first guess of randomly and independently extracting from controls matrix elements, gave unexpectedly shrank confidence intervals, smaller than those of the original set. This might be related to subject specific biases and should be further investigated.

Furthermore, sensitivity analysis can be performed for weighted global measures in the control group. From this analysis, links that contribute the most to the value of certain measure can be extracted. This particular link can be searched for in patient group, and if noticed as missing or

having largely higher or lower weight value, this can be connected to the underlying condition of the patient.

Summing up the findings from the study on borderline children, it may be stated that the approaches sensitive to the strength (or weakness) of the individual connections are critical for the advancement in understanding individual differences in cognitive ability and their modification in psychiatric disorder. Hence, the overall conclusion that can be drawn is that adjustments to the simulation procedures, for the purpose of studying the weighted networks, will most certainly be necessary.

Passing to an overall conclusion about the whole thesis, the analysis of neuroimaging derived graphs was touched both in its weaknesses and strength, showing that it cannot be undertaken by a simple implementation of graph theory. This is prevented by large inherent uncertainties and artifacts, as well as by the peculiarity of damage in pathological subjects which may range from large focal losses (here observed in a stroke case) to subtle changes involving also weak connections (here observed in the borderline children). Given also the complexity of brain, this imposes suitable analyses which should include specific subgraph studies also sizing the uncertainty levels by appropriate methods.

Bibliography

[1] A. Irimia, M. C. Chambers, C. M. Torgerson, and J. D. Van Horn, "Circular representation of human cortical networks for subject and population-level connectomic visualization," *NeuroImage*, vol. 60, no. 2, pp. 1340–1351, Apr. 2012, doi: 10.1016/j.neuroimage.2012.01.107.

[2] O. Sporns, "Brain connectivity," *Scholarpedia*, vol. 2, no. 10, p. 4695, 2007, doi: 10.4249/scholarpedia.4695.

[3] D. S. Bassett and E. T. Bullmore, "Human Brain Networks in Health and Disease," *Curr. Opin. Neurol.*, vol. 22, no. 4, pp. 340–347, Aug. 2009, doi: 10.1097/WCO.0b013e32832d93dd.

[4] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010, doi: 10.1016/j.neuroimage.2009.10.003.

[5] P. Rakic, "Specification of cerebral cortical areas," *Science*, vol. 241, no. 4862, pp. 170–176, Jul. 1988, doi: 10.1126/science.3291116.

[6] C. Destrieux, B. Fischl, A. Dale, and E. Halgren, "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature," *NeuroImage*, vol. 53, no. 1, pp. 1–15, Oct. 2010, doi: 10.1016/j.neuroimage.2010.06.010.

[7] A. Fornito, A. Zalesky, and E. T. Bullmore, *Fundamentals of brain network analysis*. Amsterdam ; Boston: Elsevier/Academic Press, 2016.

[8] S. Ramón y Cajal, *Histology of the nervous system of man and vertebrates*. New York: Oxford University Press, 1995.

[9] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proc. Natl. Acad. Sci.*, vol. 104, no. 24, pp. 10240–10245, Jun. 2007, doi: 10.1073/pnas.0701519104.

[10] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003, doi: 10.1016/S1053-8119(03)00202-7.

[11] O. Sporns, D. Chialvo, M. Kaiser, and C. Hilgetag, "Organization, development and function of complex brain networks," *Trends Cogn. Sci.*, vol. 8, no. 9, pp. 418–425, Sep. 2004, doi: 10.1016/j.tics.2004.07.008.

[12] P. Hagmann *et al.*, "Mapping Human Whole-Brain Structural Networks with Diffusion MRI," *PLoS ONE*, vol. 2, no. 7, p. e597, Jul. 2007, doi: 10.1371/journal.pone.0000597.

[13] M. Kaiser, "A tutorial in Connectome Analysis: Topological and Spatial Features of Brain Networks. Running title: Principles of Connectome Analysis," p. 29.

[14] Y. Iturria-Medina and A. C. Evans, "On the central role of brain connectivity in neurodegenerative disease progression," *Front. Aging Neurosci.*, vol. 7, May 2015, doi: 10.3389/fnagi.2015.00090.

[15] O. Sporns, G. Tononi, and R. Kötter, "The Human Connectome: A Structural Description of the Human Brain," *PLoS Comput. Biol.*, vol. 1, no. 4, p. e42, 2005, doi: 10.1371/journal.pcbi.0010042.

[16] B. Jie, C.-Y. Wee, D. Shen, and D. Zhang, “Hyper-connectivity of functional networks for brain disease diagnosis,” *Med. Image Anal.*, vol. 32, pp. 84–100, Aug. 2016, doi: 10.1016/j.media.2016.03.003.
[17] E. Bullmore and O. Sporns, “Complex brain networks: graph theoretical analysis of structural and functional systems,” *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009, doi: 10.1038/nrn2575.
[18] C. Bordier, C. Nicolini, and A. Bifone, “Graph Analysis and Modularity of Brain Functional Connectivity Networks: Searching for the Optimal Threshold,” *Front. Neurosci.*, vol. 11, 2017, doi: 10.3389/fnins.2017.00441.
[19] K. Supekar, V. Menon, D. Rubin, M. Musen, and M. D. Greicius, “Network Analysis of Intrinsic Functional Brain Connectivity in Alzheimer’s Disease,” *PLoS Comput. Biol.*, vol. 4, no. 6, p. e1000100, Jun. 2008, doi: 10.1371/journal.pcbi.1000100.
[20] C. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, “Small-World Networks and Functional Connectivity in Alzheimer’s Disease,” *Cereb. Cortex*, vol. 17, no. 1, pp. 92–99, Feb. 2006, doi: 10.1093/cercor/bhj127.
[21] W. de Haan, W. M. van der Flier, T. Koene, L. L. Smits, P. Scheltens, and C. J. Stam, “Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer’s disease,” *NeuroImage*, vol. 59, no. 4, pp. 3085–3093, Feb. 2012, doi: 10.1016/j.neuroimage.2011.11.055.
[22] K. Konrad and S. B. Eickhoff, “Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder,” *Hum. Brain Mapp.*, vol. 31, no. 6, pp. 904–916, Jun. 2010, doi: 10.1002/hbm.21058.
[23] H.-C. Baggio *et al.*, “Functional brain networks and cognitive deficits in Parkinson’s disease: Functional Network Analysis in PD,” *Hum. Brain Mapp.*, vol. 35, no. 9, pp. 4620–4634, Sep. 2014, doi: 10.1002/hbm.22499.
[24] B. M. Dale, M. A. Brown, and R. C. Semelka, *MRI: Basic Principles and Applications*. Wiley, 2015. [Online]. Available: https://books.google.it/books?id=2D1UCgAAQBAJ
[25] V. P. B. Grover, J. M. Tognarelli, M. M. E. Crossey, I. J. Cox, S. D. Taylor-Robinson, and M. J. W. McPhail, “Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians,” *J. Clin. Exp. Hepatol.*, vol. 5, no. 3, pp. 246–255, Sep. 2015, doi: 10.1016/j.jceh.2015.08.001.
[26] P. M. Matthews and P. Jezzard, “Functional magnetic resonance imaging,” *Funct. Magn. Reson. Imaging*, p. 7.
[27] O. Dietrich, A. Biffar, A. Baur-Melnyk, and M. F. Reiser, “Technical aspects of MR diffusion imaging of the body,” *Eur. J. Radiol.*, vol. 76, no. 3, pp. 314–322, Dec. 2010, doi: 10.1016/j.ejrad.2010.02.018.
[28] P. B. Kingsley, “Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors,” *Concepts Magn. Reson. Part A*, vol. 28A, no. 2, pp. 101–122, Mar. 2006, doi: 10.1002/cmr.a.20048.
[29] R. Woodhams *et al.*, “Diffusion-weighted Imaging of the Breast: Principles and Clinical Applications,” *RadioGraphics*, vol. 31, no. 4, pp. 1059–1084, Jul. 2011, doi: 10.1148/rg.314105160.

[30] R. Bammer, “Basic principles of diffusion-weighted imaging,” *Eur. J. Radiol.*, vol. 45, no. 3, pp. 169–184, Mar. 2003, doi: 10.1016/S0720-048X(02)00303-0.
[31] M. E. Moseley, J. Kucharczyk, J. Mintorovitch, and Y. Cohen, “Diffusion-Weighted MR Imaging of Acute Stroke: Correlation with T2- Weighted and Magnetic Susceptibility-Enhanced MR Imaging in Cats,” p. 7.
[32] E. O. Stejskal and J. E. Tanner, “Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient,” *J. Chem. Phys.*, vol. 42, no. 1, pp. 288–292, Jan. 1965, doi: 10.1063/1.1695690.
[33] C. Beaulieu, “The basis of anisotropic water diffusion in the nervous system - a technical review,” *NMR Biomed.*, vol. 15, no. 7–8, pp. 435–455, Nov. 2002, doi: 10.1002/nbm.782.
[34] L. J. O’Donnell and C.-F. Westin, “An Introduction to Diffusion Tensor Image Analysis,” *Neurosurg. Clin. N. Am.*, vol. 22, no. 2, pp. 185–196, Apr. 2011, doi: 10.1016/j.nec.2010.12.004.
[35] P. G. P. Nucifora, R. Verma, S.-K. Lee, and E. R. Melhem, “Diffusion-Tensor MR Imaging and Tractography: Exploring Brain Microstructure and Connectivity,” *Radiology*, vol. 245, no. 2, pp. 367–384, Nov. 2007, doi: 10.1148/radiol.2452060445.
[36] K. G. Schilling *et al.*, “Limits to anatomical accuracy of diffusion tractography using modern approaches,” *NeuroImage*, vol. 185, pp. 1–11, Jan. 2019, doi: 10.1016/j.neuroimage.2018.10.029.
[37] T. E. J. Behrens, S. N. Sotiropoulos, and S. Jbabdi, “MR Diffusion Tractography,” in *Diffusion MRI*, Elsevier, 2014, pp. 429–451. doi: 10.1016/B978-0-12-396460-1.00019-6.
[38] P. Mukherjee, S. W. Chung, J. I. Berman, C. P. Hess, and R. G. Henry, “Diffusion Tensor MR Imaging and Fiber Tractography: Technical Considerations,” *Am. J. Neuroradiol.*, vol. 29, no. 5, pp. 843–852, May 2008, doi: 10.3174/ajnr.A1052.
[39] J. S. W. Campbell and G. B. Pike, “Potential and limitations of diffusion MRI tractography for the study of language,” *Brain Lang.*, vol. 131, pp. 65–73, Apr. 2014, doi: 10.1016/j.bandl.2013.06.007.
[40] C. A. Baron and C. Beaulieu, “Acquisition strategy to reduce cerebrospinal fluid partial volume effects for improved DTI tractography: Reduction of CSF Partial Volume Effects for DTI,” *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1075–1084, Mar. 2015, doi: 10.1002/mrm.25226.
[41] T. E. Conturo *et al.*, “Tracking neuronal fiber pathways in the living human brain,” *Proc. Natl. Acad. Sci.*, vol. 96, no. 18, pp. 10422–10427, Aug. 1999, doi: 10.1073/pnas.96.18.10422.
[42] D. K. Jones, “Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI,” *Magn. Reson. Med.*, vol. 49, no. 1, pp. 7–12, Jan. 2003, doi: 10.1002/mrm.10331.
[43] O. Ciccarelli, M. Catani, H. Johansen-Berg, C. Clark, and A. Thompson, “Diffusion-based tractography in neurological disorders: concepts, applications, and future developments,” *Lancet Neurol.*, vol. 7, no. 8, pp. 715–727, Aug. 2008, doi: 10.1016/S1474-4422(08)70163-7.
[44] D. K. Jones, T. R. Knösche, and R. Turner, “White matter integrity, fiber count, and other fallacies: The do’s and don’ts of diffusion MRI,” *NeuroImage*, vol. 73, pp. 239–254, Jun. 2013, doi: 10.1016/j.neuroimage.2012.06.081.

[45] C.-H. Yeh, R. E. Smith, T. Dhollander, F. Calamante, and A. Connelly, “Connectomes from streamlines tractography: Assigning streamlines to brain parcellations is not trivial but highly consequential,” *NeuroImage*, vol. 199, pp. 160–171, Oct. 2019, doi: 10.1016/j.neuroimage.2019.05.005.
[46] J. P. Yuan *et al.*, “Test–Retest Reliability of Graph Theoretic Metrics in Adolescent Brains,” *Brain Connect.*, vol. 9, no. 2, pp. 144–154, Mar. 2019, doi: 10.1089/brain.2018.0580.
[47] M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. T. Yeo, and R. Schmidt, “Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations,” *NeuroImage*, vol. 152, pp. 437–449, May 2017, doi: 10.1016/j.neuroimage.2017.02.005.
[48] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998, doi: 10.1038/30918.
[49] D. S. Bassett and E. T. Bullmore, “Small-World Brain Networks Revisited,” *The Neuroscientist*, vol. 23, no. 5, pp. 499–516, Oct. 2017, doi: 10.1177/1073858416667720.
[50] O. Sporns, *Networks of the brain*. Cambridge, Massachusetts ; London: The MIT Press, 2011.
[51] M. D. Humphries and K. Gurney, “Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence,” *PLoS ONE*, vol. 3, no. 4, p. e0002051, Apr. 2008, doi: 10.1371/journal.pone.0002051.
[52] A.-L. Barabási and R. Albert, “Emergence of Scaling in Random Networks,” *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999, doi: 10.1126/science.286.5439.509.
[53] L. M. Colon-Perez, M. Couret, W. Triplett, C. C. Price, and T. H. Mareci, “Small Worldness in Dense and Weighted Connectomes,” *Front. Phys.*, vol. 4, 2016, doi: 10.3389/fphy.2016.00014.
[54] N. T. Markov *et al.*, “The role of long-range connections on the specificity of the macaque interareal cortical network,” *Proc. Natl. Acad. Sci.*, vol. 110, no. 13, pp. 5187–5192, Mar. 2013, doi: 10.1073/pnas.1218972110.
[55] M. E. J. Newman, “Clustering and preferential attachment in growing networks,” *Phys. Rev. E*, vol. 64, no. 2, p. 025102, Jul. 2001, doi: 10.1103/PhysRevE.64.025102.
[56] L. da F. Costa, M. Kaiser, and C. C. Hilgetag, “Predicting the connectivity of primate cortical networks from topological and spatial node properties,” *BMC Syst. Biol.*, vol. 1, no. 1, p. 16, 2007, doi: 10.1186/1752-0509-1-16.
[57] A. Barrat, M. Barthélemy, and A. Vespignani, “Weighted Evolving Networks: Coupling Topology and Weight Dynamics,” *Phys. Rev. Lett.*, vol. 92, no. 22, p. 228701, Jun. 2004, doi: 10.1103/PhysRevLett.92.228701.
[58] J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski, “Intensity and coherence of motifs in weighted complex networks,” *Phys. Rev. E*, vol. 71, no. 6, p. 065103, Jun. 2005, doi: 10.1103/PhysRevE.71.065103.
[59] B. Zhang and S. Horvath, “A General Framework for Weighted Gene Co-Expression Network Analysis,” *Stat. Appl. Genet. Mol. Biol.*, vol. 4, no. 1, Jan. 2005, doi: 10.2202/1544-6115.1128.
[60] C. J. Stam and J. C. Reijneveld, “Graph theoretical analysis of complex networks in the brain,” *Nonlinear Biomed. Phys.*, vol. 1, no. 1, p. 3, Dec. 2007, doi: 10.1186/1753-4631-1-3.

[61] V. Latora and M. Marchiori, “Efficient Behavior of Small-World Networks,” *Phys. Rev. Lett.*, vol. 87, no. 19, p. 198701, Oct. 2001, doi: 10.1103/PhysRevLett.87.198701.
[62] M. Cao, H. Huang, Y. Peng, Q. Dong, and Y. He, “Toward Developmental Connectomics of the Human Brain,” *Front. Neuroanat.*, vol. 10, Mar. 2016, doi: 10.3389/fnana.2016.00025.
[63] S. P. Borgatti and M. G. Everett, “Models of corerperiphery structures,” p. 21, 1999.
[64] O. Civier, R. E. Smith, C.-H. Yeh, A. Connelly, and F. Calamante, “Is removal of weak connections necessary for graph-theoretical analysis of dense weighted structural connectomes from diffusion MRI?,” *NeuroImage*, vol. 194, pp. 68–81, Jul. 2019, doi: 10.1016/j.neuroimage.2019.02.039.
[65] C. R. Buchanan *et al.*, “The effect of network thresholding and weighting on structural brain networks in the UK Biobank,” *NeuroImage*, vol. 211, p. 116443, May 2020, doi: 10.1016/j.neuroimage.2019.116443.
[66] B. C. M. van Wijk, C. J. Stam, and A. Daffertshofer, “Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory,” *PLoS ONE*, vol. 5, no. 10, p. e13701, Oct. 2010, doi: 10.1371/journal.pone.0013701.
[67] S. Achard and E. Bullmore, “Efficiency and Cost of Economical Brain Functional Networks,” *PLoS Comput. Biol.*, vol. 3, no. 2, p. e17, Feb. 2007, doi: 10.1371/journal.pcbi.0030017.
[68] D. S. Bassett, E. T. Bullmore, A. Meyer-Lindenberg, J. A. Apud, D. R. Weinberger, and R. Coppola, “Cognitive fitness of cost-efficient brain functional networks,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 28, pp. 11747–11752, Jul. 2009, doi: 10.1073/pnas.0903641106.
[69] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. E. Hulshoff Pol, “Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain,” *NeuroImage*, vol. 43, no. 3, pp. 528–539, Nov. 2008, doi: 10.1016/j.neuroimage.2008.08.010.
[70] C. E. Ginestet, T. E. Nichols, E. T. Bullmore, and A. Simmons, “Brain Network Analysis: Separating Cost from Topology Using Cost-Integration,” *PLoS ONE*, vol. 6, no. 7, p. e21570, Jul. 2011, doi: 10.1371/journal.pone.0021570.
[71] M. A. de Reus and M. P. van den Heuvel, “Estimating false positives and negatives in brain networks,” *NeuroImage*, vol. 70, pp. 402–409, Apr. 2013, doi: 10.1016/j.neuroimage.2012.12.066.
[72] R. F. Betzel, A. Griffa, P. Hagmann, and B. Mišić, “Distance-dependent consensus thresholds for generating group-representative structural brain networks,” *Netw. Neurosci.*, vol. 3, no. 2, pp. 475–496, Jan. 2019, doi: 10.1162/netn_a_00075.
[73] J. A. Roberts, A. Perry, G. Roberts, P. B. Mitchell, and M. Breakspear, “Consistency-based thresholding of the human connectome,” *NeuroImage*, vol. 145, pp. 118–129, Jan. 2017, doi: 10.1016/j.neuroimage.2016.09.053.
[74] M. W. Cole, T. Yarkoni, G. Repovš, A. Anticevic, and T. S. Braver, “Global Connectivity of Prefrontal Cortex Predicts Cognitive Control and Intelligence,” *J. Neurosci.*, vol. 32, no. 26, pp. 8988–8999, Jun. 2012, doi: 10.1523/JNEUROSCI.0536-12.2012.
[75] E. Santarnecchi, G. Galli, N. R. Polizzotto, A. Rossi, and S. Rossi, “Efficiency of weak brain connections support general cognitive functioning,” p. 17.

[76] D. S. Bassett, B. G. Nelson, B. A. Mueller, J. Camchong, and K. O. Lim, “Altered resting state complexity in schizophrenia,” *NeuroImage*, vol. 59, no. 3, pp. 2196–2207, Feb. 2012, doi: 10.1016/j.neuroimage.2011.10.002.
[77] M. Novkovic *et al.*, “Topological Small-World Organization of the Fibroblastic Reticular Cell Network Determines Lymph Node Functionality,” *PLOS Biol.*, vol. 14, no. 7, p. e1002515, Jul. 2016, doi: 10.1371/journal.pbio.1002515.
[78] M. Krzywinski *et al.*, “Circos: An information aesthetic for comparative genomics,” *Genome Res.*, vol. 19, no. 9, pp. 1639–1645, Sep. 2009, doi: 10.1101/gr.092759.109.
[79] B. Fischl *et al.*, “Whole Brain Segmentation,” *Neuron*, vol. 33, no. 3, pp. 341–355, Jan. 2002, doi: 10.1016/S0896-6273(02)00569-X.
[80] K. Shen *et al.*, “Exploring the limits of network topology estimation using diffusion-based tractography and tracer studies in the macaque cortex,” *NeuroImage*, vol. 191, pp. 81–92, May 2019, doi: 10.1016/j.neuroimage.2019.02.018.
[81] V. Blasi *et al.*, “Early Life Adversities and Borderline Intellectual Functioning Negatively Impact Limbic System Connectivity in Childhood: A Connectomics-Based Study,” *Front. Psychiatry*, vol. 11, 2020, doi: 10.3389/fpsyt.2020.497116.
[82] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertész, “Generalizations of the clustering coefficient to weighted complex networks,” *Phys. Rev. E*, vol. 75, no. 2, p. 027105, Feb. 2007, doi: 10.1103/PhysRevE.75.027105.
[83] T. Taoka *et al.*, “Fractional Anisotropy–Threshold Dependence in Tract-Based Diffusion Tensor Analysis: Evaluation of the Uncinate Fasciculus in Alzheimer Disease,” *Am. J. Neuroradiol.*, vol. 30, no. 9, pp. 1700–1703, Oct. 2009, doi: 10.3174/ajnr.A1698.
[84] A. Kunimatsu *et al.*, “The Optimal Trackability Threshold of Fractional Anisotropy for DiŠusion Tensor Tractography of the Corticospinal Tract,” *Magn. Reson. Med. Sci.*, vol. 3, no. 1, p. 7, 2004.
[85] M. P. van den Heuvel and O. Sporns, “Network hubs in the human brain,” *Trends Cogn. Sci.*, vol. 17, no. 12, pp. 683–696, Dec. 2013, doi: 10.1016/j.tics.2013.09.012.