Whole-brain cortical parcellation:

A hierarchical method

based on dMRI tractography

Whole-brain cortical parcellation: A hierarchical method based on dMRI tractography

DISSERTATION

in partial fulfillment
of the requirements for the degree of
Doctor of Engineering
Doktoringenieur (Dr.-Ing.)

developed at:
Max Planck Institute for Human Cognitive and Brain Sciences Leipzig
Methods & Development Unit "Cortical Networks and Cognitive Functions"

submitted to:
Ilmenau University of Technology
Faculty of Computer Science and Automation
(*Fakultät für Informatik und Automatisierung
der Technischen Universität Ilmenau*)

by Dipl.-Ing. David Moreno-Dominguez
born on 16. November 1983 in Valladolid, Spain

Reviewers:
Prof. Dr.-Ing. habil Jens Haueisen
PD Dr.-Ing. habil Thomas R. Knösche
Jr. Prof. Dr. Mario Hlawitschka

Date of submission: 19. May 2014
Date of defense: 19. November 2014

Contents

Abstract

In modern neuroscience there is general agreement that brain function relies on networks and that connectivity is therefore of paramount importance for brain function. Accordingly, the delineation of functional brain areas on the basis of diffusion magnetic resonance imaging (dMRI) and tractography may lead to highly relevant brain maps.

Existing methods typically aim to find a predefined number of areas and/or are limited to small regions of grey matter. However, it is in general not likely that a single parcellation dividing the brain into a finite number of areas is an adequate representation of the function-anatomical organization of the brain.

In this work, we propose hierarchical clustering as a solution to overcome these limitations and achieve whole-brain parcellation. We demonstrate that this method encodes the information of the underlying structure at all granularity levels in a hierarchical tree or dendrogram. We develop an optimal tree building and processing pipeline that reduces the complexity of the tree with minimal information loss. We show how these trees can be used to compare the similarity structure of different subjects or recordings and how to extract parcellations from them.

Our novel approach yields a more exhaustive representation of the real underlying structure and successfully tackles the challenge of whole-brain parcellation.

Zusammenfassung

In den modernen Neurowissenschaften ist allgemein anerkannt, dass die Gehirnfunktionen auf dem Zusammenwirken von verschiedenen Regionen in Netzwerken beruhen und die strukturelle Konnektivität daher großer Bedeutung ist. Daher kann die Abgrenzung funktioneller Hirnbereiche auf der Grundlage der Diffusions-Magnet-Resonanz-Tomographie (dMRT) und der Traktografie zu wertvollen Hirnkarten führen.

Existierende Verfahren versuchen eine fest vorgegebene Anzahl von Regionen zu finden und/oder sind auf kleine Bereiche der grauen Substanz beschränkt. Im Allgemeinen ist es jedoch unwahrscheinlich, dass eine einzelne Parzellierung des Kortex, eine ausreichende Darstellung der funktio-anatomischen Organisation des Gehirns erlaubt.

In dieser Arbeit schlagen wir eine hierarchische Clusteranalyse vor um diese Einschränkungen zu überwinden und das gesamte Gehirn zu parzellieren. Wir zeigen, dass dieses Verfahren die Eigenschaften der zugrundeliegenden Struktur auf allen Granularitätstufen des hierarchischen Baums (Dendrogramm) kodieren kann. Weiterhin entwickeln wir eine optimale Verarbeitungspipeline zur Erstellung dieses Baums, die dessen Komplexität mit minimalem Informationsverlust reduziert. Wir zeigen wie diese Datenstrukturen verwendet werden können um die Ähnlichkeitstruktur von verschiedenen Probanden oder Messungen zu vergleichen und wie man daraus verschiedene Parzellierungen des Gehirns erhalten kann.

Unser neuer Ansatz liefert eine ausführlichere Analyse der anatomischen Strukturen und bietet eine Methode zur Parzellierung des ganzen Gehirns.

Abbreviations

ADC	apparent diffusion coefficient.
DSI	diffusion spectrum magnetic resonance imaging.
CNS	central nervous system.
CPCC	cophenetic correlation coefficient.
cXX	centroid method with XX neighborhood.
dMRI	diffusion magnetic resonance imaging.
FA	fractional anisotropy.
fMRI	functional magnetic resonance imaging.
fODF	fiber orientation density function.
FT	Fourier transform.
GPU	graphics processing unit.
HARDI	high angular resolution diffusion imaging.
IFG	inferior frontal gyrus.
IPCC	inferior parietal cortex convexity.
MRI	magnetic resonance imaging.
NMR	nuclear magnetic resonance.
PAS	persistent angular structure.
PICo	probabilistic index of connectivity.
PLS	polarized light imaging.
SMA	supplementary motor area.
SNR	signal to noise ratio.
SS	spread vs. separation
tCPCC	tree cophenetic correlation coefficient.
TE	echo time.
TR	repetition time.
wTriples	weighted triples similarity.

1. Introduction

1.1 Overview

This chapter contains an overview of prior knowledge needed to understand the framework of this thesis. In particular, basics of human brain anatomy and structure, the rationale behind brain parcellation, the role and organization of brain anatomical connectivity, and a short description of the techniques that can measure it. Finally, we highlight the possibilities and current limitations of *in-vivo* connectivity-based parcellation and the motivation of our work in order to overcome them, present the contribution of our developed method, and give an outlook on the contents of remaining chapters in this thesis.

Overall, this introductory chapter is inspired from atlases, books and thesis chapters from: Gray (1918), Brodmann (1909), Johansen-Berg and Behrens (2009), Descoteaux (2010), Jones (2011) and Sporns (2011a). They are great sources for the understanding of human brain anatomy, structure and connectivity.

1.2 The human brain

The brain is the main structure of the central nervous system (CNS), and regulates all human activity. From an embryonic development standpoint, it consists of hindbrain, midbrain and forebrain. This latter, which is the most recently evolved, can be subdivided into diencephalon (which contains among other structures the thalamus, a critical relay for sensory information, and the hypothalamus, the central organizing structure for the regulation of the body's many homeostatic functions such as feeding, and thermoregulation) and telencephalon or cerebrum (formed by the basal ganglia and cerebral hemispheres).

The cerebral hemispheres control all voluntary actions in the body and are responsible for higher cognitive functions. In humans, they are proportionally larger than in any other mammals, and have characteristic folds called gyri (singular, gyrus). The grooves dividing different gyri are the sulci (singular, sulcus). Although the gyral patterns may vary across individuals, there are some features

that consistently divide the hemispheres morphologically into four lobes, named after the cranial bones that overlie them (Figure 1.1). The two hemispheres of the brain are separated by the central fissure.

The frontal lobes extend from the central sulcus to the anterior part of the brain and are involved in numerous functions including conscious thought, planning, motor control and language. The parietal lobes, located between the central and the occipital sulci, are important for sensory information integration and processing of visual-spatial stimuli. The temporal lobes are the most lateral parts of the cortex, delimited by the lateral fissure, and are responsible for processing auditory stimuli and some language related functions. Lastly, the occipital lobes are in the posterior part of the cortex and process visual stimuli. It is notable that although the hemispheres exhibit strong bilateral symmetry (in both structure and function) it is not complete. As a structural example, the left hemisphere has generally a larger lateral sulcus than the right one. Functionally, Broca's area and Wernicke's areas (involved in the understanding and generation of speech; Broca, 1861; Wernicke, 1874) are present only in the left hemisphere in most of the population (Amunts et al., 1999).
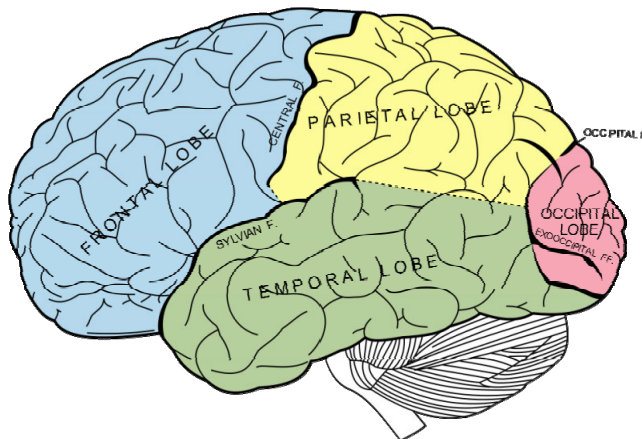
Figure 1.1:   *Reproduction from "Gray's Anatomy of the Human Body" showing the four brain lobes and the cerebellum below (belonging to the hindbrain). (vectorized image from (Gray, 1918), from*
`http://commons.wikimedia.org/wiki/File:Gray728.svg`*)*

At a microscopic scale, neurons constitute the basic building block of the nervous system. These are highly specialized cells capable of transmitting information through electrical and chemical signals. Within the cell, information is

transmitted along its surface via changes in the membrane potential. Between neurons, information is exchanged through specialized connections called synapses, mediated by chemical signals. A typical neuron consists of the cell body, dendrites, and an axon. The cell body contains the nucleus and most of the cell metabolic machinery. Dendrites are narrow structures elongating from the cell body, often branching in a tree-like shape, and are responsible for receiving incoming signals. The axon is another special neuron elongation that usually extends for longer distances, transmitting the signal to the next cell.

Neurons do not work in an isolated way, they are organized into circuits that process specific kinds of information and provide the foundation for more complex functions (Figure 1.2). Neural circuits are formed from three types of neuronal cells: afferent neurons carry sensory input towards the brain or spinal cord or bring information up to the brain cortex from lower structures; efferent neurons transmit information away from their neural circuit (or away from the CNS); interneurons regulate circuits locally, reinforcing or inhibiting certain signals.
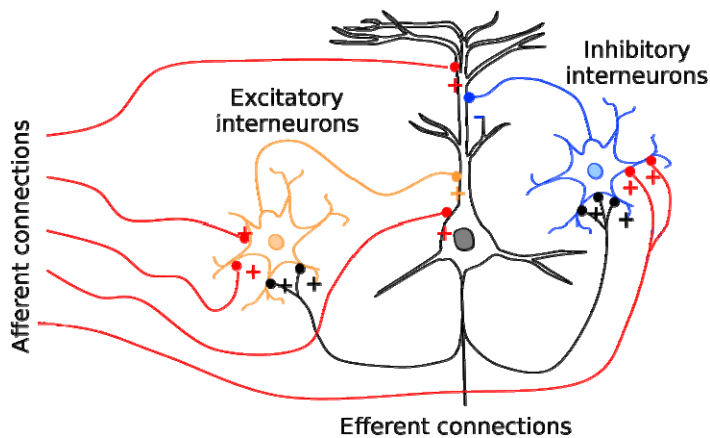
Figure 1.2:   Depiction of a basic neural circuit. The cell in the center represents a pyramidal neuron, the main type of information processing neuron in the cortex.

In the brain, neuronal cell bodies reside in what is called gray matter, and the axons form the white matter, named so for the color these different tissues have in preserved brains (they also produce a different signal contrast under magnetic resonance imaging [MRI]). Axons are often called nerve fibers, and bundles of these axons are called fiber tracts. It is in the gray matter where the processing of information takes place, and the white matter is responsible for transmitting the information inside/outside of the brain and between regions of the grey matter. In

the cerebral hemispheres, the grey matter is situated in the outermost layer, denominated the cortex.

There is evidence that in many areas, the cerebral cortex is organized in small patches of neuronal ensembles (Hubel and Wiesel, 1968; Mountcastle, 1957; for a review see Horton and Adams, 2005). These ensembles are coordinated through their local connections and work coherently, achieving functional integration (Varela et al., 2001). At the end of the scale, higher cognitive functions arise from the coordinated work of these neuronal populations, many times located in different parts of the brain and connected through the white matter, forming brain networks (Bullmore and Sporns, 2009; Dosenbach et al., 2007).

"The nervous system is organized on multiple scales, from synaptic connections between single cells, to the organization of cell populations within individual anatomical regions, and finally to the large-scale architecture of brain regions and their interconnecting pathways. Different techniques are sensitive to different levels of organization. This last point deserves to be emphasized. The multi-scale aspect of the nervous system is an essential feature of its organization and network architecture. Descriptions of the brain at large scales should not be regarded as poorly resolved approximations of an underlying microscopic order. Instead, brain connectivity at the large scale (among regions and systems) describes neural processes that are the outcome of dynamic coordination among smaller elements. (…) Perhaps the most fundamental distinction is between structural connectivity as a wiring diagram of physical links and functional connectivity as a web of dynamic interactions." (Sporns, 2011a).

1.3 Structural mapping and parcellation

(Summarized from the chapter "*Connectivity Fingerprinting of Gray Matter*" written by J. Klein, T. Behrens and H. Johansen Berg from the book on diffusion magnetic resonance imaging (dMRI) by Johansen-Berg and Behrens [2009]).

The concept that functions in the brain are localized in regions has important implications for neuroscience. The work of Korbinian Brodmann, published in 1909, represents one of the first comprehensive studies aiming to parcellate the cortex (Figure 1.3; Brodmann, 1909). In his atlas, he observed the cytoarchitecture (the layering of cells) and myeloarchitecture (properties of white matter) of the human cortex and identified 52 different areas. Other scientists such as Von Economo (1929) also produced detailed atlases of the brain, but these have not established themselves as broadly as the Brodmann parcellation scheme.

Brodmann's atlas still remains widely used in modern neuroscience, despite its age. Using Brodmann's parcellation as a guide, researchers have filled this map with other anatomical or functional results (as well as earlier findings that predated his atlas: Broca, 1861; Wernicke, 1874; Exner, 1881; Campbell, 1904).
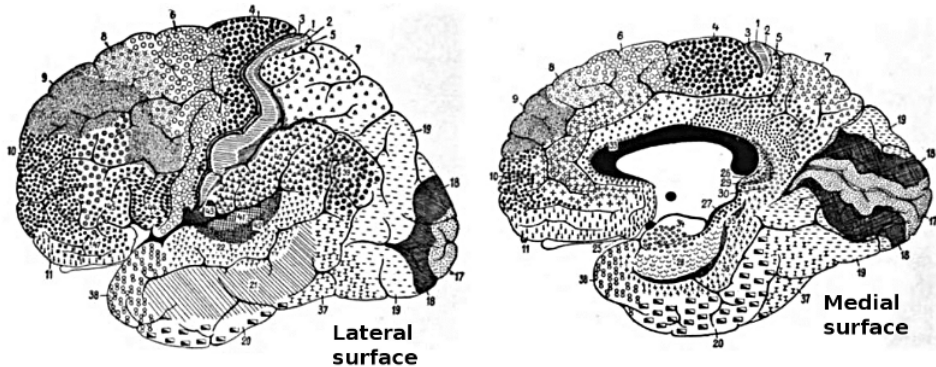
Figure 1.3:   Regions of the human cerebral cortex as delineated by Korbinian Brodmann on the basis of cytoarchitecture. (reprint from (Brodmann, 1909)).

On a microscopic level, boundaries between gray matter regions are commonly characterized, along with the already mentioned cyto- and myeloarchitecture, by patterns of neuronal chemistry and by the patterns of receptors present or absent on cells. Functional localization can be obtained using a range of techniques and experiments such as observation of symptoms in patients with specific lesions or localized epileptic seizures or electrode stimulation techniques in human subjects undergoing brain surgery. More recently, the non-invasive functional magnetic resonance imaging (fMRI) has become the most widely used tool for functional studies in the brain.

Sometimes, microscopically characterized borders are not well defined (e.g. Brodmann area 43 gradually transitioning into area 40). But even in some clearly defined areas, distinct functional regions can be observed within a single Brodmann area (Zilles et al., 1996). Choosing when two areas of the cortex are different enough to merit being considered separate areas is an arbitrary decision. It is not surprising then that some of Brodmann's areas have been proposed for further subdivision (Vogt and Vogt, 1919; Geyer et al., 2000a). Also, there is no formal proof that cytoarchitectonically distinct regions are also functionally different (however, functional specialization is evident in the layering of cells in

the examples of primary motor and visual cortices; Brodmann, 1909; Passingham, 2007).

In summary, there is a clear need for parcellation of the cortex for neuroanatomical studies. Cytoarchitecture is hypothesized to be strongly related to localized function, and has therefore been one of the most important tools to relate function between individuals and between species. However, as cytoarchitecture is usually not available for a given individual subject, other techniques are needed to provide *in-vivo* parcellations of the brain.

1.4 Anatomical connectivity

1.4.1   Connectivity as a structural trait

There are two important competing factors that describe a network in network theory: its efficiency and its cost (Latora and Marchiori, 2001, 2003). Efficiency relates to how well interconnected the elements of a network are (so that if information is to be exchanged between any two elements, the path will be as direct and fast as possible). Cost relates to how expensive the network layout is (as each direct connection between elements has an associated cost, and usually this cost is higher for faster connections). A fully interconnected network, where each element has a direct link to any other element will be very effective, but also very expensive. In the case of the human brain, the cost for each link is related to energy consumption, and mostly, to volume required: each axon needs space, and more the wider it is and the thicker the myelin sheath surrounding it (both of which enable faster signal transmission). In an organ as complex and compact as the human brain, this means that neurons cannot be fully interconnected with each other.

Given the brain organization, the only information a neuron can process is that directly received through its afferents, and this information can only affect neurons or circuits directly linked with it. Therefore it seems reasonable to assume that areas participating in the same function need to be connected with each other and have similar connectional patterns to the rest of the brain. Furthermore the specific patterns of this connectivity strongly influence the function of neural networks, enabling complex neuropsychological tasks and cognitive abilities (Mesulam, 1990, 1998). To be able to decode brain function we need not only a good understanding of its components, but also to comprehend how these components connect with each other, and on the reverse side, how these connection patterns might help shed light on their function.

1.4.2   The organization of white matter

As introduced before, there are several scales at which anatomical connectivity can be described: at the micro-scale, neurons communicate through synaptic connections forming neural circuits; at the meso-scale, neuronal populations interconnect into networks of ensembles; and at the macro-scale, large numbers of these ensembles interconnect through fiber pathways, constituting the white matter. It is on the macro-scale connectivity that we will focus on this thesis.

In the white matter, axons can be diffusely distributed or concentrated in bundles called fiber tracts. There are three different types of tracts.

Projection tracts: they extend vertically establishing connections between the cerebral cortex and subcortical structures, such as the basal ganglia and the thalamus. Afferent projection tracts carry information from different parts of the body to the cerebral cortex. Efferent projection tracts carry commands from the cortex down to the brainstem and the spinal cord.

Commissural tracts: bundles of axons connecting a region in one hemisphere to another region of the opposite hemisphere.

Association tracts: these connect different cortical areas within a given hemisphere (Figure 1.4), and can be divided into two categories. Long association fibers communicate between different cerebral lobes whereas short association tracts establish connections within a given lobe and adjacent gyri. The smallest of these are called U-fibers, as they link adjacent cortical zones separated by a sulcus, forming a characteristic "U" shape.
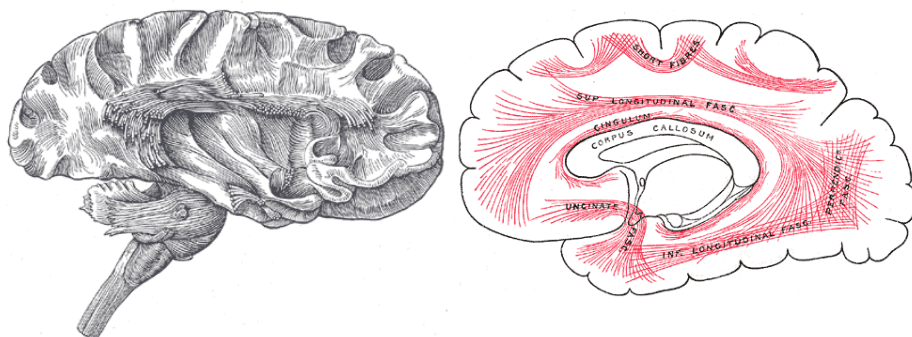

Figure 1.4:   *Dissection (left) and diagram (right) of the cerebral cortex showing principal systems of association fibers. (reprint from (Gray, 1918)).*

"Anatomical connections at all levels of scale are both specific and variable. Specificity is found in the arrangement of individual synaptic connections between morphologically and physiologically distinct neuronal types and in long-range connectivity between neural structures such as cell nuclei or brain regions. Variability is found in the shape of individual neurons and their processes, as well as in the size, placement and interconnection of large-scale structures. Variability may be measured between corresponding structures in brains of individuals of the same species. In addition, neural structures within the same individual vary across time, as a result of experiential and developmental processes of growth, plasticity and repair. It is likely that anatomical variability is one of the main sources for functional variability, expressed in neural dynamics and behavioral performance." (Sporns, 2007).

It is important then, to develop methods that can characterize variability of white matter patterns along the cortex in individual brains (*in-vivo*), and that allow systematic comparison of changes in these patterns across individuals, or across different time points.

1.4.3   Measuring anatomical connectivity

The following descriptions of invasive tracing methods are a summarized from the chapter "*Invasive methods from tracing white matter architecture*" written by H. Axer for the book on dMRI by Jones (2011).

- *Fiber dissection*
In 1935, Klingler (1935) described a technique to dissect major fiber tracts in the human brain. It allowed the fibers to be carefully separated using fine surgical tools, and through this method, the course of the major fiber tracts could be shown. Current knowledge about fiber tracts in the human brain is based mainly on such dissection studies. However, the process only allows for investigation of single tracts of fibers, as other tracts must be cut away. Although the method is not best suited for searching new unexpected neuroanatomy, it is useful for proving hypothesis obtained through other techniques.

- *Fiber degeneration*
Degeneration of the nerve fibers after transection of the axons was first described by Waller (1850) and first observed in the brain by Ramón y Cajal (1928). Degenerating fibers arising from local lesions in specific brain areas can be detected, and therefore, their tracts somehow traced. This principle can be applied

to animal brains in experimental settings, and human cadaver brains can be dissected after brain injury (Dhanarajan et al., 1977; Brodal, 1978). With MRI, Wallerian degeneration has also been detected in living brain (Axer et al., 2008; Uchino et al., 1990; Pierpaoli et al., 2001);

- *Tract tracing*

    In tract tracing a tracer substance is injected into a specific brain region and is transported via the axons into connected brain areas. Most of these techniques are based on active transport mechanisms in the living cell. Different substances can be used to achieve anterograde or retrograde transport (respectively, from the cell body to the synapse or the other way around). As working cell mechanisms are necessary, the method is limited to experiments in living animals or dead human brains with very short postmortem delays before injection (Haber, 1988). Some substances can potentially be applied to fixed brain tissue as they are transported passively along the myelin sheaths, but maximal distance of diffusion is limited and waiting time can be as long as half a year (Lukas et al., 1998).

- *Myelin staining*

    Different methods of myelin stains can be used to visualize part of the myelin sheaths and the axons, using compounds that bind with different target molecules. The basic method was initially described by Weigert (1897). Currently, the Luxol fast blue approach is the one most often used (Klüver and Barrera, 1953) because it is easy to apply. However, with this method intermingling fibers with distinct orientation cannot be clearly distinguished, so the fine architecture of white matter cannot be exactly analyzed. Myelin stains are preferably used in the gray matter, where the nerve fibers are separated from each other (Schmitt et al., 2004).

- *Confocal laser microscopy*

    Through the use of confocal laser scanning microscopy and a special fluorescent dye, it is possible to make serial optical sections through fiber bundles with high magnification. This technique collects information from well-defined tissue sections by sequentially illuminating only a narrow slide of tissue volume (Wright et al., 1993). Afterwards, a 3D image with high resolution information about fiber orientation can be reconstructed. The drawback of such high detail is a narrow field of view. It is therefore not suitable for analyzing large-scale architectural patterns.

- *Polarized light imaging*

    Polarized light imaging (PLI) allows the visualization of anisotropic fiber bundles with a lower magnification than confocal laser microscopy but a larger

field of view (Axer et al., 2000; Larsen et al., 2007). PLI can selectively visualize anisotropic structures with birefringent properties, such as nerve tissue. As with confocal laser microscopy, sections are analyzed separately and later 3D images can be reconstructed (Axer et al., 2002). However in this technique separate sections of 60 to 100 μm thick must first be carefully cut. Difficulties arise mainly from the elaborate histological processing of the tissue and the access to the large cryotome machines required.

- *Diffusion MRI*

dMRI was first developed in the 1980s (Le Bihan and Breton, 1985; Taylor and Bushell, 1985). It is based on the property that water does not freely diffuse in tissue, as it is hindered by obstacles such as membranes and macromolecules. Microscopic properties and architecture of tissues can then be obtained by observing how water diffuses in them. This is specially so in nerve tissue, where water can diffuse more freely along the direction of the axons and the myelin sheaths, than in the perpendicular plane, where it is restricted by the cell membrane. As will be explained in more detail in the next chapter, in dMRI the hydrogen atoms in the water molecules are spatially encoded in a specific direction by their spin phase through specialized electromagnetic pulses while in a strong magnetic field. The molecules then emit the energy absorbed through a process called relaxation and their signal is captured. If the water molecules diffuse spatially in the encoded direction, the spin phases of the hydrogen atoms at a given point will not all be in phase, and it will translate into a decrease of the detected signal. Repeating the process for different directions a 3D image encoding the water diffusion patterns at each point in the brain can be reconstructed. By application of models to this water diffusion pattern that relate to the underlying tissue microstructure, several types of informative data can be obtained, such as fractional anisotropy (FA) maps (thought to be influenced by fiber density, axonal diameter, and myelination in white matter; Basser and Pierpaoli, 1996) and fiber tractography (3D reconstruction the main nerve fiber paths, which will be the basis of the connectivity information that the methods developed in this thesis will analyze and characterize; Basser et al., 2000; Stieltjes et al., 2001; Koch et al., 2002).

However, diffusion tractography cannot provide the same level of evidence as invasive tracing: *"Essentially, we are looking at diffusion of water within the brain, not at the actual fibers we are really interested in. Diffusion tractography cannot decide whether a reconstructed pathway is a direct one or involves one or more synapses, as there is no signature of synapses in the diffusion signal. Also, as we are dealing with water diffusing along the paths of the axons, diffusion tractography*

cannot provide information about the polarity of connections." (Johansen-Berg and Behrens, 2009).

Diffusion tractography still remains, despite these shortcomings, the only technique available to assess brain anatomical connectivity *in-vivo* in humans, making it most valuable.

1.5   Connectivity based brain parcellation

Tractography enables us to obtain connectivity fingerprints or patterns from different points of the cortex. These fingerprints can then be compared in other to analyze how similar are the connectivity patterns of two given points. This information can be used by clustering algorithms in order to generate parcellations that group together points with similar patterns in the same parcel, while keeping points with distinct patterns in different ones. This is the basis for all methods for *in-vivo* anatomical connectivity based parcellation in humans (i.e.: Johansen-Berg et al., 2004; Anwander et al., 2007).

However, while all methods follow this basic scheme, there are many possible ways to implement tractography, compute similarity between tractograms, and define and perform the clustering and parcellations. In the following chapter, we will review in more detail the different choices available and being used today.

It is unlikely that a single parcellation dividing the brain into a finite number of functional areas would be an adequate representation of the functional organization of the brain, in the same way that a political map subdividing the earth's land surface is not a perfect representation of the cultural differences and kinships amongst its people. The measurable changes of properties on the cortical surface are often gradual rather than abrupt. In these cases, we might find different partitions depending on how we define the minimum structural difference that merits that these points belong to different regions, that is, on the level of granularity of the partition. Also, even in cases where these changes are sharp and a partition remains constant for a wide range of granularities, there can still exist nested divisions within the regions of this partition. This is exemplified by the cytoarchitecture work of Caspers and colleagues (2008) and the tractography work of Ruschel and colleagues (2013), where Brodmann's areas 39 and 40 were further subdivided. A partition should, therefore, be seen as an approximation of the similarity structure (i.e., expressed by a correlation matrix) of some structural property.

1.6 Main contribution and overview of this thesis

In this work, we propose hierarchical clustering as a solution to overcome the challenge of whole-brain high-resolution multi-granularity parcellation. We aim to demonstrate that hierarchical clustering is a promising means by which to characterize the connectivity similarity structure of the human brain, where the information of the underlying structure at all granularity levels is encoded in a hierarchical tree or dendrogram. The idea is that these trees can then be sampled to obtain partitions at different granularity levels, and are more suitable for whole-brain parcellation that other available methods.

We compared the performance of several classical hierarchical methods and implemented our own method specially tuned for the challenge of whole brain parcellation based from highly dimensional tractograms. Our method combines hierarchical centroid linkage clustering with a physical neighborhood restriction, and an initial homogeneous merging stage. It proved to be the best performing algorithm by data-fit and computational cost criteria.

Once trees are obtained, interpreting the large amount of data encoded and extracting the most relevant information is not an easy task. To aid this process, a dendrogram pre-processing pipeline was designed and implemented that reduces the complexity of the resulting trees, while keeping most of its information, to facilitate further analysis.

We then show how the trees can be used to compare the similarity structure of different subjects or time points, all while remaining in the subject space without the need to transform the data to a common space prior to partitioning.

Global comparison can be achieved using the full connectivity structure information through dendrogram comparison. For this purpose, tree leaves must first be matched: we devised and implemented a method to achieve leaf matching in connectivity-based trees and then applied dendrogram comparison methods present in the literature.

We also compare the trees at selected granularity levels through the use of partition finding algorithms. We apply the some common partition methods and propose a new partition-quality measure coupled with an effective tree search algorithm in order to find relevant partition ranges.

Original contributions of this thesis.

1 – Brought the concept of full multi-granularity clustering into brain anatomical connectivity parcellation.

2 – Built a framework that allows for whole-brain multi granularity parcellation from high resolution data (1mm) based on connectivity by bringing together existing fast probabilistic tractography with hierarchical clustering used in other fields such as genetics and parcellation of functional data.

3 – Replaced the traditional Pearsons correlation measure for tractogram similarity with a normalized dot-product more suitable for describing the similarity of non-negative anatomical tractograms.

4 – Optimized a hierarchical algorithm for probabilistic tractogram clustering by applying neighborhood restrictions to the centroid method and using an initial size-restricted stage in order to obtain an early set of homogenous-sized clusters. The resulting algorithm captures as much information as the best performing of the traditional methods and reduces the number of needed tractogram distance computations by two orders of magnitude compared to the traditional methods (therefore drastically reducing the time needed for building the tree). It also facilitates posterior processing and comparison of the resulting trees.

5 – Implemented a dendrogram preprocessing pipeline in order to reduce the complexity of the resulting trees while minimizing the information loss. The pipeline consists of the following steps: monotonicity correction, limiting of maximum granularity, and detecting and collapsing non-binary structures in the tree. This pipeline successfully reduces tree complexity by 90% while keeping the information loss below 0.05%.

6 – Introduced the idea of leaf-matching in order to enable the application of tree-comparison algorithms and through them the possibility of full connectivity similarity structure comparison (using the information from all levels of granularity) across subjects or measurements. Implemented a first proof-of-concept method by using greedy matching of mean tractograms for the leaf-matching and tree cophenetic correlation and triples methods for comparison of the matched trees.

7 – Implemented a Spread vs. Separation based partition quality measure that uses only information contained in the tree and is therefore very fast to evaluate the

quality of a partition (as opposed to using the original data, typically done in the literature). Combined with a hierarchical search algorithm this allows to search for the optimal quality partition for each possible granularity and to obtain a profile of most relevant granularities based on this quality measure.

8 – Developed a tool for interactive exploration of the hierarchical trees and real-time projection of results onto freesurfer brain surfaces (within the OpenWalnut developer framework). The implemented tree-processing and partition selection algorithms were included along with some of the most common ones from the literature.

9 – Realized a proof-of-concept pilot study with 4 healthy subjects applying all the developed techniques, analyzed the results and contrasted with some other techniques from the literature.

10 – Replicated the clustering divisions obtained by Ruschel and colleagues (2013) on the inferior parietal cortex convexity (IPCC) of 20 healthy subjects in 3 clusters through the application of our algorithm to their tractography data.

<u>This document is organized as follows:</u>

The first chapter has been an introduction to the human brain organization and function, along with the need for brain parcellation and specifically parcellation based on connectivity.

In Chapter 2, the technical foundations of dMRI, methods for diffusion-based tractography and a review of the main connectivity-based parcellation algorithms are presented. The rationale for using hierarchical clustering to overcome some of their limitations is introduced.

Chapter 3 details the methods used and implemented in this thesis, as well as the different methodological hypothesis and choices considered during the research. First we describe the traditional hierarchical clustering algorithms, and present a new modified method more suitable for clustering anatomical connectivity. A tree processing pipeline is then presented that maximizes information compression in the trees while minimizing information loss. Next tree-matching and tree-comparison algorithms are described that allow comparison of the full connectivity similarity structure across datasets. Finally, different schemes for partition selection within the tree are presented.

The methods described are then applied to datasets obtained from a small cohort of healthy subjects in Chapter 4. The fit to the data of the different hierarchical algorithms is tested and the best performing method is chosen. Then, it is shown how the tree processing pipelines successfully reduces tree complexity without loss of information. Next, the results of the tree comparison scheme are presented and discussed, and finally, the partitions obtained with the different partition-selection methods are explored and compared.

Chapter 5 discusses the possible approaches and challenges faced when validating clustering data from dMRI. Two small studies are proposed and carried out to increase the degree of confidence in the proposed method.

Chapters 6 and 7 discuss the findings of our work, and compare them with other methods in the current literature. The advantages and weaknesses of the method are pointed out, and possible ways of improvement. A summary of the work done on this thesis is then offered, and future development and potential applications are suggested.

The Appendix at the end of the document contains extra figures not included in the main chapters. Namely, results of the circumstantial validation study for all 20 subjects of the original study, and the partitions obtained with the four different partition-selection methods used and developed in both hemispheres of the five datasets acquired.

Part of the methods and results presented in this thesis have been published as an article in the scientific journal *Human Brain Mapping* (Moreno-Dominguez et al., 2014a) and in posters and talks at international conferences (Anwander et al., 2012; Knösche et al., 2012; Moreno-Dominguez et al., 2013, 2012a, 2012b, 2011a, 2011b, 2011c).

2. Brain analysis based on water diffusion measured by MRI

2.1 Overview

This chapter deepens in the technical concepts needed to understand the choices and challenges faced during the work developed in this thesis, and where the data it is based on comes from (in particular, dMRI, fiber orientation modeling, and tractography techniques). It also contains a review of the literature in dMRI-connectivity based parcellation algorithms, and points out the main limitations of these particular techniques when faced with a whole brain approach, which inspired the work of this thesis in an effort to overcome them.

This current methods chapter is inspired from materials, books and thesis chapters from: Hornak (1996), Koch (2000), Mori (2007), Johansen-Berg and Behrens (2009) and Jones (2011). They are great sources for the understanding of magnetic resonance and dMRI-based techniques.

2.2 dMRI imaging

2.2.1 Basics of MRI

Before explaining how it is possible to measure the amount of water diffusion in tissue, we must first understand the principles of conventional MRI. MRI makes use of a physical phenomenon called nuclear magnetic resonance (NMR) in order to detect the nuclei of atoms in a body placed within the MRI scanner. Due to NMR, when an atom nucleus is placed in a strong magnetic field, it can absorb and then re-emit electromagnetic energy at a frequency determined by the magnetic field strength. The strong magnetic field is generated by the great ring-shaped scanner magnet (Figure 2.1), and the excitation energy is transmitted to the nuclei and detected upon re-emission by specialized transmitter and receiving coils in the scanner. If the magnetic field strength is not uniform but changes in dependence with position in space through a field gradient, then the amount of energy re-

emitted by the nuclei at different positions can be disentangled based on the frequency. Applying a similar principle in order to vary the phase of the wave emitted based on position, two different dimensions can be encoded and a 2D image can be obtained. This 2D image is usually encoded perpendicularly to the direction of the strong magnetic field (to the magnet ring axis). Using a third magnetic field gradient in the third dimension before the excitatory pulse is applied, only incoming energy pulses at a certain frequency will be absorbed by the nuclei at each position (along that dimension). This way slice selection can be achieved, only a slice of nuclei will be excited by each pulse and generate a 2D image. Adding all the slices a full 3D image is reconstructed. Hydrogen in its single proton isotope is the nucleus most widely used when performing MRI, as it is the most widely present in tissue through water molecules. With this technique we can obtain what is called *proton density images* as the amount of signal received from each volume depends on the density of hydrogen nuclei present in them.

Figure 2.1:   *Siemens MT-Trio 3T MR scanner (used to acquire the data used in this thesis), Max Planck Institute for Human Cognitive and Brain Sciences.*

But there are also other types of image that can be obtained, producing other contrasts or, so called, *weightings*. This requires some further explanation on the magnetic properties of nuclei. Atom nuclei generate a magnetic field, and at any instant time, the magnetic field generated by atom nuclei in a certain volume (subjected to a homogenous outside magnetic field) can be represented by a magnetization vector. When in equilibrium, this vector is aligned with the applied magnetic field (the field created by the large MRI magnet), and its magnitude determined by the field strength. When an electromagnetic pulse is absorbed by the nuclei, it changes their magnetization vector, and through a process called

relaxation, the magnetization vector returns back to its equilibrium state, releasing energy in the process which will be detected (Figure 2.2).
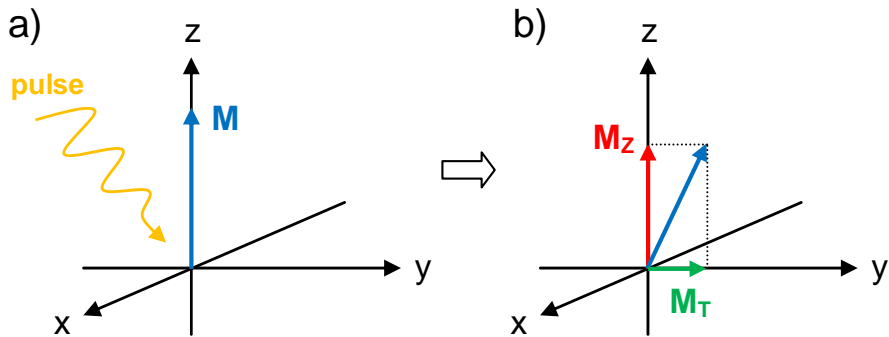

Figure 2.2:   *a) In equilibrium, the magnetization vector M is aligned with the scanner axis (usually z dimension). b) when a pulse is applied and absorbed by the nuclei, this magnetization vector changes, the component of the new vector in the xy plane is the transversal component M_T, and the one in the z direction is the longitudinal component M_Z.*

However, relaxation times in the longitudinal and transversal directions to the equilibrium one depend on different characteristics of the sample, and using specific pulses to change the magnetization vector in one or other direction can reveal different information of the tissue. The relaxation time in the longitudinal direction depends on the mobility of the lattice (the nuclei surroundings) and is denominated spin-lattice or T_1 relaxation. Exploiting this relaxation gives rise to *T_1-weighted* images (Figure 2.3 left), which in brain imaging show a good contrast between gray and white matter. Relaxation time in the transverse direction depends on molecular interactions, is denominated spin-spin or T_2 relaxation and gives rise to *T_2-weighted* images (Figure 2.3 right). In these images fluids show a very bright contrast and they are useful to detect pathological brain tissue.
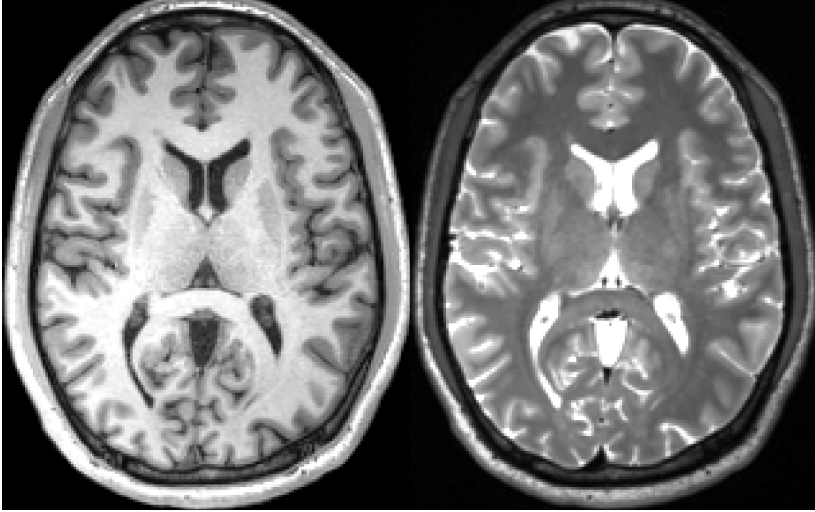
Figure 2.3:   *Sagittal view of a healthy young volunteer in a T_1 weighted image (left) and T_2*
                 weighted image (right), Max Planck Institute for Human Cognitive and Brain
                 Sciences.

2.2.2   Measuring diffusion

Molecules in a fluid in equilibrium do not stay fixed or static, but move respect to each other in a random pattern called Brownian motion. This property, which water molecules also share, is called diffusion. This can be exemplified by a drop of ink falling into a water container: at first the ink particles will remain localized close to the point where it dropped, but with time they will slowly spread randomly and evenly through the container. The distance that water molecules in tissue diffuse in a given time can be described by the equation (Einstein, 1956):

$$\left\langle x^2 \right\rangle = 2Dt \tag{2.1}$$

where x is the mean diffused distance in a time t and D is a constant called the *diffusion coefficient*.

This diffusion constant can be measured in MRI by applying special additional field gradients. The magnetization vectors of hydrogen nuclei are actually rotating around an axis aligned at the strong magnetic field direction, this vector rotation is called *precession*.

If after the initial excitation (where all the magnetization vectors have the same orientation) special gradients are applied, vectors at different locations can be

made to precess at different speeds. If after a certain time T the direction of precession is suddenly reversed (with a specialized energy pulse), all the vectors will be at the same orientation again exactly after a time T since the inversion (that is, $2T$ since the original application of the special gradients, called diffusion gradients). To understand this with an analogy, if at a race each runner has a constant but different speed, and at a certain moment T all simultaneously turn around, they will all reach again the starting point at the same instant which will be $2T$ (as the slower runners were closer from the start when they turned around). If all the magnetization vectors within each measurable volume (called voxel) are aligned, the net magnetization will be the addition of the magnetization vectors of the contained nuclei, and upon relaxation the full signal will be received, as if the gradients had never been applied (Figure 2.4).
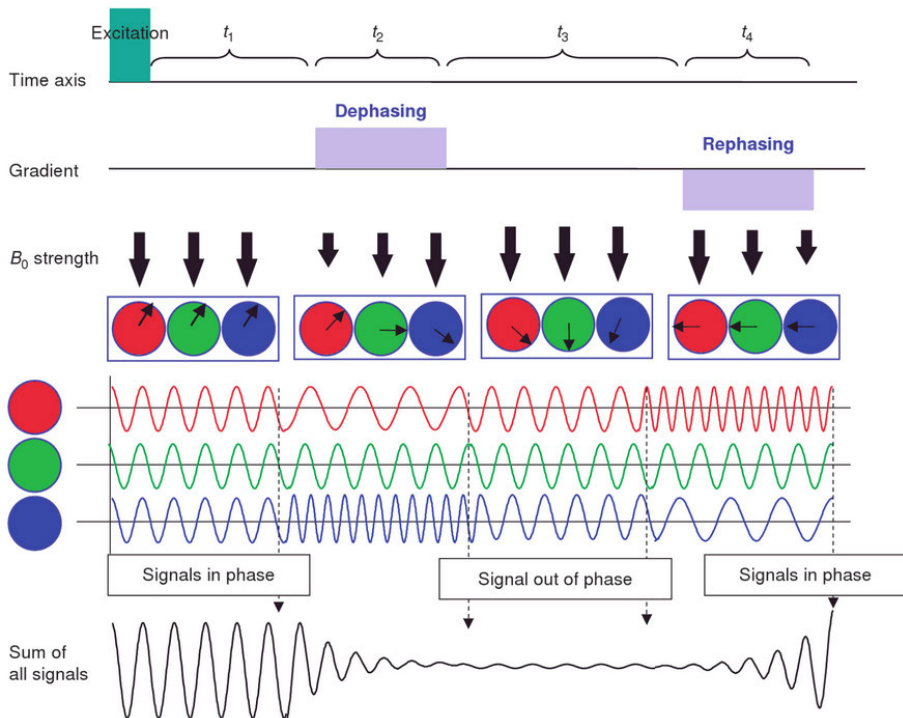

Figure 2.4:   *Example of the application of a dephase-rephase gradient sequence in the absence of water diffusion. The red, green and blue circles indicate three water molecules at different positions. Thick arrows indicate the strength of the magnetic field applied and narrow arrows within the circles indicate the magnetization vector orientation form each molecule. Image taken from "Introduction to Diffusion Tensor Imaging". (reprint from (Mori, 2001)).*

However, the scenario just portrayed does not take into account the diffusion of water molecules. If diffusion occurs, a nucleus from a position with precession at a certain speed might move to a different position were original precession speed was different. As a result, when the instant $2T$ arrives, the magnetization vectors in a voxel will not all be aligned, but there will be small directional difference in the vectors. Parts of these vectors will cancel each other, and the net magnetization vector will be smaller the greater the mixing of hydrogen atoms with different precession speeds (that is, the greater the mean diffusion distance was). As the net magnetization vector will have smaller magnitude than in absence of diffusion, the signal received upon relaxation in will also be smaller (Figure 2.5).
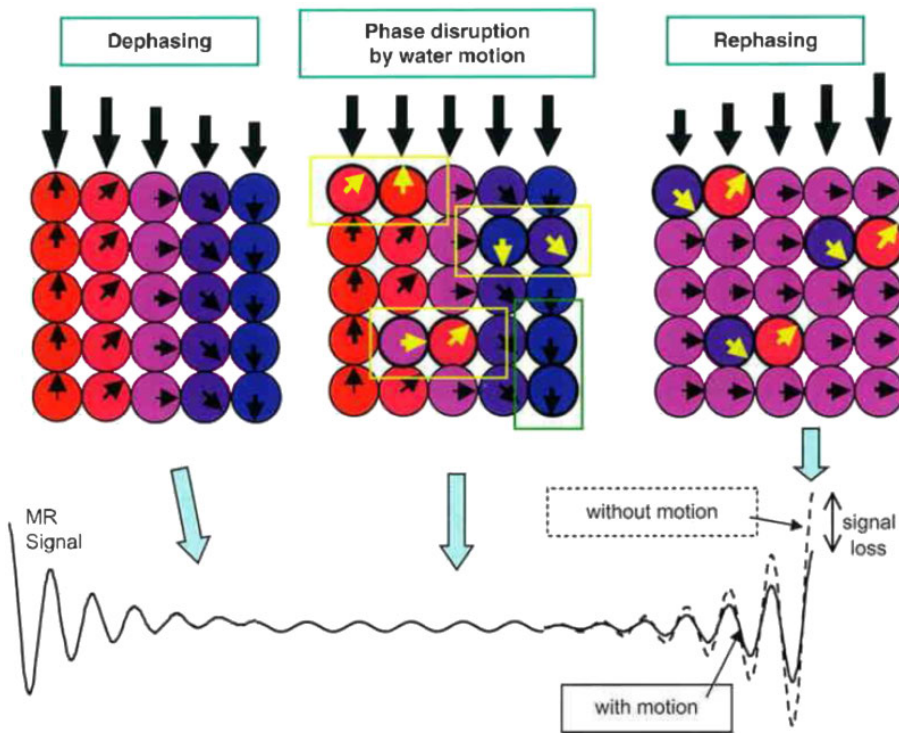

Figure 2.5: *Effect of water diffusion on the dephase-rephase sequence pictured in Figure 2.4. Thick arrows indicate the strength of the magnetic field applied and narrow arrows within the circles indicate the magnetization vector orientation form each molecule. This orientation is also indicated by gradation of colors. Water molecules that diffused away from their original positions are highlighted by boxes. Image taken from "Introduction to Diffusion Tensor Imaging". (reprint from (Mori, 2007)).*

When water can diffuse without restriction in all directions, the displacement of water molecules follows a Gaussian distribution, and a single diffusion coefficient D is enough to characterize it. In tissue, however, water diffusion can be hindered by big molecules, and differently in each direction (depending on the microstructure of the tissue at each point). In this case, what is measured for each direction is called the *apparent diffusion coefficient* (ADC). Approximating the displacement of the water molecules in tissue as a Gaussian distribution (as in unrestricted free diffusion), and following *Equation 2.1*, the attenuation of the measured signal due to diffusion can be described by:

$$\frac{I_b}{I_0} = \exp(-b, ADC) \tag{2.2}$$

where I_0 and I_b are the intensities of the signals received without diffusion weighting and with a diffusion weighting equal to b, respectively (Le Bihan, 1986; Stejskal and Tanner, 1965).

With this technique, using a diffusion gradient in a specific direction we can measure the apparent diffusion constant in that direction, translated in a decay of the signal respect to the full signal obtained without a gradient. Repeating the measurement in different gradient directions we can obtain a map of how water diffuses at each point in the brain and for each direction in space. The sensitivity of the measurement to diffusion (that is, how much the signal decays with the mean diffusion distance) depends on the strength of the diffusion gradients applied, on the time this gradients are active, and on the time T were the precession inversion is forced. These parameters are usually combined in a single one called the *b*-value.

But how many measured directions are enough to fully characterize water diffusion in tissue? As will be seen in the next section, this depends on the mathematical model used to represent water diffusion properties at each point in space, and in the angular resolution and accuracy desired. The number of acquired directions ranges from 6 (the minimum number of directions necessary for the simplest model) to 256 (very high angular resolution).

2.2.3   Modeling fiber orientation

2.2.3.1   The diffusion tensor

When diffusion is not equal in all directions it is said that the medium is anisotropic. White matter tissue is highly anisotropic given that water molecules within the neurons diffuse fairly unrestricted along the axons but cannot diffuse well across the cellular membrane. The water molecules in the extracellular matrix

are also less restricted in the direction of the axons than in the perpendicular one. dMRI data from white matter can therefore provide information about the orientation of the nerve fibers.

But in order to obtain measures that describe interpretable properties of the underlying microstructure of the tissue, and more importantly, to be able to use the information to reconstruct fiber paths (which will be explained in the next section) first we need to define a mathematical model that can integrate the information from the diffusion images into a more practical and usable form. The information of the diffusion coefficient in different directions in space is called the diffusion propagator. The simplest model to describe it is the diffusion tensor. It can be represented by a 3x3 matrix of numbers in the form:

$$D = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} \tag{2.3}$$

The elements on the diagonal (D_{xx}, D_{yy} and D_{zz}) correspond to the diffusivities on the x, y and z axes of the dMRI images, while the other elements describe the correlation between the diffusivity on those axes. The matrix is symmetrical, so the elements below the diagonal are equal to those elements above (inverting the order of the subscripts). A more intuitive way to understand the diffusion tensor is through its graphical ellipsoid representation (Figure 2.6).
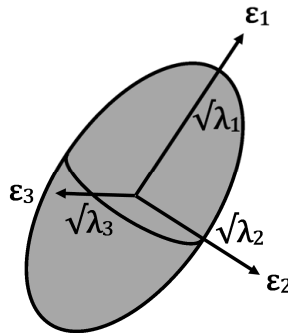

Figure 2.6:   *Schematic of the diffusion tensor ellipsoid. With ε_i indicating the tensor eigenvectors and λ_i the tensor eigenvalues.*

The surface of this ellipsoid represents the points where a water molecule situated at the origin would diffuse to with equal probability. Usually a reference frame aligned with the axes of the ellipsoid (rather than the image axes) is used for simplicity. This coordinate system is called eigensystem. The axes of the ellipsoid

are described by the eigenvectors (ε_1, ε_2 and ε_3), and the lengths of these axes by the squared root (as by *Equation 2.1*) of the eigenvalues (λ_1, λ_2 and λ_3), where λ_1 corresponds to the largest eigenvalue and λ_3 to the smallest.

If the axes of the ellipsoid coincide with the axes of the image acquisition, then all the off-diagonal elements in *Equation 2.3* equal zero and the diagonal elements correspond to the eigenvalues. As the tensor matrix is symmetrical, there are six unknown variables to characterize (or seen graphically, 3 variables for the main direction of the ellipsoid in space, plus 3 more for the length each axis). Six is then the minimum amount of diffusion encoded images (in different directions) that must be measured in order to obtain the diffusion tensor (plus an extra image without diffusion weighting). However, it is usual to acquire a greater number of images and later fit the data into the tensor in order to reduce the effects of noise and obtain more precise data (Jones, 2004).

Intuitively, for white matter the diffusion tensor for each voxel gives us the main direction of the fibers going through that voxel, and also how strong is the directionality of the diffusion. Several magnitude measures have been developed from the diffusion tensor in order to allow intuitive and easy exploration of diffusion data.

The measure most used in the clinical setup is the mean diffusivity (average diffusivity in a voxel regardless of the direction). It is computed as the averaged sum of the three diagonal elements (or averaged sum of the eigenvalues) and is represented by the symbol λ.

$$\lambda = \frac{1}{3}\left(D_{xx} + D_{yy} + D_{zz}\right) = \frac{1}{3}\left(\lambda_1 + \lambda_2 + \lambda_3\right) \tag{2.4}$$

Although it can be useful to detect abnormal or injured tissue, this measure does not give any information on how anisotropic is the tissue. Fractional anisotropy (FA) is the most widely used anisotropy index, and it is given by the equation (Basser and Pierpaoli, 1996):

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \lambda)^2 + (\lambda_2 - \lambda)^2 + (\lambda_3 - \lambda)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.5}$$

Where λ is the aforementioned mean diffusivity. FA is thought to be influenced by fiber density, axonal diameter, and myelination. A typical FA image can be seen in Figure 2.7. Other types of anisotropy indices and their properties can have been discussed by Papadakis and colleagues (1999).
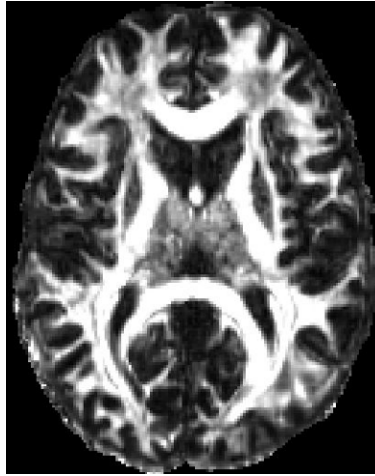
Figure 2.7:   FA image of a healthy young volunteer (sagittal view). The subject and the position of the image are the same as in those shown in Figure 2.3. Max Planck Institute for Human Cognitive and Brain Sciences.

Acquiring diffusion tensor images from dMRI data has thus become a very popular practice: it enables the quantification of diffusion anisotropy (an index of white matter integrity) and estimates the main direction of nerve fibers, needed for tractography.

2.2.3.2   Modeling multiple fibers

Although the diffusion tensor is a powerful tool, it has one key limitation: as it estimates only one main fiber direction per voxel, it cannot model fiber crossings. This should be carefully considered when performing tractography and connectivity analysis. There are other alternative models and algorithms that aim to overcome this limitation by extracting more exhaustive information about the fiber orientations.

-     *Multi-tensor*

The multi tensor model is a simple extension of the diffusion tensor. In this case the diffusion propagator is approximated by a number n of Gaussian density functions (the number of tensors), which at the same time model *n* different fiber populations in different orientations. This assumes that the water molecules stay within one population and do not diffuse into the others.

- *Ball and stick*

The ball and stick model (Behrens et al., 2003b; Hosey et al., 2005) can be considered a special case of the multi tensor model. It assumes that water molecules are contained in two different types of compartments: one with restricted diffusion within and around fibers (the "stick"), and one with free diffusion that does not interact with the restricted one (the "ball"). In its simplest version the sticks are modelled as tensors in which only one of the eigenvalues is non-zero, that is, water molecules have Gaussian diffusion and only in one direction. The ball is modelled as a spherical tensor with isotropic diffusion. Multiple fibers can be considered by including multiple sticks, same as in the multi-tensor model.

These methods have some disadvantages: the large number of parameters to be fitted can cause instability, and as they try to fit a certain number of orientations they are unable to differentiate fanning/bending fibers from fibers running in parallel. In order to avoid this, other methods try to obtain what is called fiber orientation density function (fODF) from the dMRI measurements, which contains more detailed information of the fiber configuration. They can be referred as non-parametric methods as they try to reconstruct the fODF directly from the data without constraining its shape to a particular parametric model.

- *Diffusion spectrum imaging*

Diffusion spectrum imaging relies on the fact that if infinitesimally short gradient pulses are used in the acquisition of the dMRI image, the diffusion propagator can be obtained by performing the Fourier transform (FT) of the measurement data (Wedeen et al., 2000; Tuch, 2002). This way no assumptions are made on the tissue microstructure or in the shape of the fODF. The main disadvantage of this method is the long acquisition time due to the high number of diffusion images needed (usually an order of magnitude higher than with other methods). Also, the Fourier relationship to the data is an approximation: in reality gradient pulses are not ideal and have a duration in the same order of magnitude as the diffusion time, which constitutes a significant deviation from the original assumptions.

- *QBall imaging*

This method approximates the dODF obtained in diffusion spectrum imaging by using less measurements through a special acquisition scheme. The approximation relies on using a transform function called the Funk-Radon transform instead of the Fourier one (Tuch, 2004; Tuch et al., 2003). This way fewer images need to be obtained and acquisition requirements are reduced. The approximation however,

translates into some blurring, which can affect the precision of the detected peak directions and reduce angular resolution.

- *Spherical deconvolution*

In spherical deconvolution, the measured image is considered as the convolution of the fODF and the signal that a single fiber population would create. In order to obtain the fODF a deconvolution of the measurements and the single population signal (which is required beforehand by the method) must be performed (Anderson and Ding, 2002; Tournier et al., 2004). In addition, some techniques have recently been developed to extract the properties of the different peaks in order to obtain measures analogous to the FA in the diffusion tensor for each separate direction and/or tract (Riffert et al., 2014). A limitation of spherical deconvolution is that noise may cause spurious peaks in the resulting fODF. Regularization techniques eliminate these spurious peaks but they also reduce angular resolution.

In summary, these presented techniques expand and improve the directional fiber density information provided by the diffusion tensor, but they still have limitations and also require further validation work. Also, there is an implicit trade-off between number of acquisitions and image resolution. In the future, it might be possible that with high enough resolution a basic model will be enough as each voxel would contain only one fiber population (although this is still not possible with current *in-vivo* technology). Finally, while these techniques do succeed in describing fiber crossings, they are still unable to distinguish them from bending fibers and cannot properly characterize fiber fannings.

2.3   Tractography

2.3.1   Deterministic tractography

Tractography is a method that lets us reconstruct the main fiber pathways through the white matter from diffusion data. Compared to techniques that measure connectivity directly in the brain (presented in section 1.4.3) tractography is indirect, more difficult to interpret and error-prone. It is however, the only method available to study anatomical connectivity in-vivo, and is therefore an important technique to understand function in normal and diseased brain.

In order to reconstruct fiber tracts, tractography methods rely on the fiber orientation descriptive models presented in section 2.2.3, and they all try to find paths of minimal hindrance to diffusion using this local voxel-wise orientation information, but there are different strategies to integrate local information into a path.

Deterministic tractography tries to recover the most likely single path from a chosen starting point, and it is based on the streamline concept: in the presence of a continuous vector field, a streamline is a curve that is always tangential to the direction of the vector field at each point. This can be applied to tractography using as vector field the calculated principal fiber directions. That is, choosing a starting seed point, we can guide the streamline looking at the principal fiber direction at each step. An example of this process is shown in Figure 2.8.
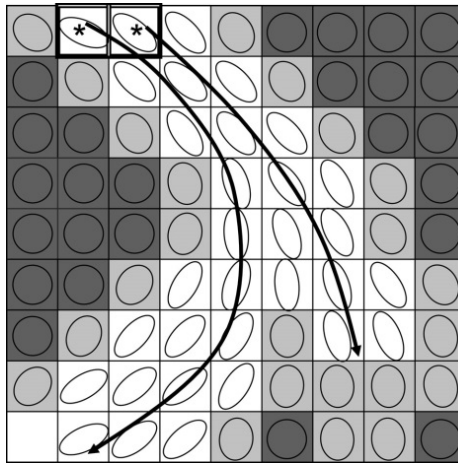
Figure 2.8:   *Example of streamline tractography following orientations of least hindrance to diffusion and being terminated at voxels of low anisotropy (in darker tone). (reprint from (Mori and Zhang, 2006)).*

The mathematical formula guiding the streamline can be written as:

$$\frac{dr(s)}{ds} \parallel \varepsilon(r(s)) \tag{2.6}$$

Where $\parallel$ is the *parallel* operator, ε is the tangent to the vector field at position $r(s)$ and $r(s)$ in turn represents the position of the streamline curve in 3D space at a distance s along the streamline from the starting point (Basser et al., 2000). If the local fiber orientation is described by a diffusion tensor, the tangent to the curve has to be parallel to this direction, i.e., it has to be parallel to the principal

eigenvector of the tensor at that point. It should be noted that as it is a differential equation, it is not possible to directly calculate the position of the streamline for a particular value of s without previously calculating the intermediate values sequentially from the beginning. In other words, the streamline propagates from the seed voxel. An important implication of this fact is that any errors that may occur during the tracking will propagate and compound through the rest of the streamline, there is however a lot of literature on differential equations about strategies to minimize this.

So far we have assumed a continuous vector field of principal fiber directions. In practice, we only possess one principal direction for each measured voxel and must therefore infer their values for a continuous field. Early methods simply assign the same value to the space covered by its respective voxel (Mori et al., 1999) but that leads to significant propagation errors (Lazar and Alexander, 2003). Improved approaches interpolate information from neighboring voxels in order to obtain a smooth vector field, either combining data directly from the diffusion image or the obtained diffusion propagators (Pajevic et al., 2002).

There are three possible error sources during tractography: noise in the diffusion image (causing wrong estimation of fiber directions), modeling errors (the particular local model fails to properly characterize the structure at that point) and errors caused by the approximation to the specific fiber direction at each continuous point through the interpolation scheme.

In order to reduce compounding propagation errors (and to finish the tractogram at some point) streamline termination criteria are applied. There are two common ones: minimum FA value (in tensor tractography) and maximum rate of direction change (minimum streamline curvature radius). A very low FA value indicates that there is a lot of uncertainty in the principal diffusion direction, and therefore a lot of potential error (that would propagate and compound if the streamline was continued). It can also mean that the streamline has reached grey matter, and therefore the pathway target. Maximum curvature follows the idea that major white-matter pathways are usually smooth and do not bend sharply. The presence of a sharp bend is highly likely due to an error source, rather than a representation of the real microstructure, and such a streamline should be stopped.

There are adaptations of the streamline concept to other multiple fiber methods in order to allow tracking through regions with crossings (Wedeen et al., 2008; Descoteaux et al., 2009; Malcolm et al., 2010; Tournier et al., 2012; for a review see Lenglet et al., 2009).

2.3.2   **Probabilistic tractography**

Deterministic tractography has some limitations on its capability to describe the fiber bundles as it only follows the central line of fanning fibers, or the "strongest" path of bifurcating ones. In points where multiple equally probable directions are possible, only one is selected and the rest neglected. Therefore this approach is not suitable to assess connectivity between arbitrary regions. Probabilistic tractography aims to overcome this limitation (Koch et al., 2002; Behrens et al., 2003b; Anwander et al., 2007; Kaden et al., 2007; Jeurissen et at., 2011).

In order to understand the principles behind it, we will focus on the algorithm developed by Koch (2000) to perform probabilistic tractography in a 2D slice.

"For the assessment of anatomical connectivity between arbitrary regions an algorithm is needed that differentiates between trajectories in highly aligned bundles and paths through almost isotropic matter. In order to find a numerical measure meeting this requirement, a Monte-Carlo type algorithm was implemented. Imagine a particle in one of the voxels of a cortical region 'A' that jumps in a random manner from voxel to voxel. It will perform a random walk through the set of voxels. Let us further make the probability of a jump to a neighboring voxel dependent on the diffusion tensor in the current (and in the neighboring) voxel, such that the probability is higher the larger the diffusion coefficient in the jump direction. Then our particle will move with a higher probability along a fiber direction than perpendicular to it. If we perform this "experiment" many times and count how often our particle starting in a region 'A' has reached region 'B', we obtain a (relative) measure of the anatomical connectivity between regions 'A' and 'B'. For each elementary jump the probabilities for the eight possible jump directions to a neighboring voxel (the particle motion was confined to the imaging slice) were calculated from the diffusion tensors in the start voxel (m) and the target voxel (n) according to

$$p(m \rightarrow n) = \frac{\left[d(r_{mn},m)+d(r_{mn},n)\right]^{a}}{\sum_{n}\left[d(r_{mn},m)+d(r_{mn},n)\right]^{a}} \tag{2.7}$$

where $p(m \rightarrow n)$ is the probability for a jump from voxel m to voxel n, $d(r_{mn},m)$ is the 'diffusion coefficient' defined in in voxel m for the direction from the center of voxel m to the center of voxel n, and a = 7. The sum of the probabilities over the eight possible n is 1.

The exponent *a* was introduced to make the probability distribution sufficiently localized at the directions corresponding to the fiber orientation. With *a* = 1 the particle path was not confined to the fiber direction. The objective of the simulation was to obtain a numerical measure of the subjective impression of thickness and coherence of a fiber tract. Thus the exponent *a* was adjusted to keep the majority of the particle paths in the voxels that constituted a fiber on the diffusion tensor fiber orientation map. For the same purpose only jumps in the "fiber direction" in the previous voxel and the two directions that deviated from it by ±45° were allowed. The fiber direction in a voxel was defined as that among the 8 directions with the largest in-plane diffusion coefficient. Among these two opposite directions that direction was chosen that did not include an acute angle with the direction of the jump to the current voxel. If the angle was 90⊡ then the choice was arbitrary.

A pseudo-random integer number between 0 and 7 with the calculated probability distribution was generated by the transformation method (Press et al., 1992), and used to select the jump direction. The particle path was terminated if a voxel with a FA < 0.2 was reached or a maximum number of jumps had been performed. The maximum number of elementary jumps was chosen to be sufficiently large to allow the particle to reach the neighboring gyri when starting on a gyral crown. The frequency with which each voxel was hit (as a result of any particle jump during a path or at its terminating point) was recorded. This number was normalized to the maximum over all considered pixels in the slice."
(Koch, 2000; Figure 2.9).

A 3D extension of this method was developed by Anwander and colleagues (2007). The resulting tractograms are often viewed as three dimensional images, and the values in each voxel (ranging from 0 to 1) indicate the plausibility of an anatomical connection between that voxel and the starting seed voxel. Tractograms can also be rearranged as a one dimensional vector with as many elements as tractography target voxels (the voxels in the white matter) in order to facilitate computations of tractogram similarity.
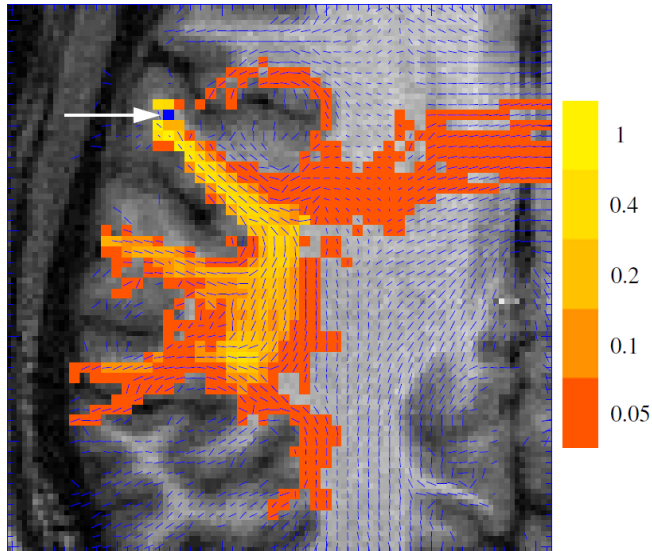
Figure 2.9:   *Example for the result of the particle-jump probabilistic tract algorithm from Koch. The start pixel is shown in blue and marked with an arrow. Normalized visitation values are indicated by red-to-yellow colors. Uncolored pixels were never reached by the particle. Blue lines represent main local fiber orientations at each voxel. (reprint from (Koch, 2002)).*

We have used this particular algorithm as example as it will be the one used to obtain the connectivity fingerprints used in this thesis (see chapter 3 and discussion section 6.1 for detailed arguments on that choice). However, there are also implementations of probabilistic tractography that use the same basic principles applied to the other local models of diffusion previously presented (Behrens et al., 2007; Descouteaux et al., 2009).

2.3.3   Global tractography

Global tractography methods aim to find an alternative to the sequential step scheme in order to avoid the compounding propagation errors. This has been pursued in several different ways. One type of approach uses self-organizational principles to join multiple particles modeled into each voxel and build all the tractograms at the same time (Kreher et al., 2008; Fillard et al., 2009; Reisert et al., 2011). Other approaches model pathways by preselecting target points/areas and obtaining smooth curves between them using cubic splines. These splines are optimized through the vector field of orientations, also adding or removing

intermediate control points during this process (Tuch, 2002). The resulting curves represent the most probable pathways between the predefined points assuming that a connection does exist. In the approach by Jbabdi and colleagues (2007) a Bayesian framework is used to obtain the most probable course using different parameters (local diffusion, fiber orientations, anisotropy...). A comparison between deterministic, probabilistic and global tractography can be found at (Bastiani et al., 2012).

A recently proposed Plausibility Tracking method (Schreiber et al., 2014), combines and extends some of the previous methods and introduces new approaches to quantify the directional alignment. It proposes a multi-stage approach: first, a close-to-optimal initialization of the spline parameters is obtained by probabilistic tractography. Second, the parameters of the spline describing the pathway are optimized in accordance with the relative local fODF derived from constrained spherical deconvolution (Tournier et al., 2007).

It was chosen not to use global tractography to obtain the tractograms used in this thesis, as in most algorithms both the start and finishing target areas of the pathways must be set, and in any case require significant computational power that would make our whole-brain approach unfeasible in the required timeline. Therefore the fast implementation of probabilistic tractography from Anwander and colleagues (2007) was chosen instead.

2.4   dMRI connectivity based parcellation

2.4.1   Basis of connectivity based parcellation

The tools presented in this chapter offer the possibility to characterize the changes in the long range connectivity patterns of different patches of gray matter, or in other words, for human anatomical connectivity based parcellation *in vivo*.

The process of connectivity based parcellation involves typically four main steps: first, the connectivity properties of a point or patch of gray matter are characterized in what is called the connectional fingerprint of that region; secondly, these fingerprints must be somehow compared or evaluated, the usual way is defining a measure between fingerprints, that can numerically account for their similarity; next, a clustering algorithm is applied on the properties of these fingerprints or on their similarities, which yields as result a partition that characterizes the main connectivity changes of the studied region; lastly, the partitioning obtained is projected back into the grey matter points from where they originated and their relevance, meaning and validity is discussed.

There are many different possible approaches when deciding what a fingerprint consists of, how to measure their similarities, and what clustering algorithms to use, each with their own advantages and limitations. Many times, the specific decision taken on one steps, has an influence in the choices available for the rest, making them not completely independent from one another. In order to explore the different possibilities, in the next section we will go through the main solutions that can be found in the literature up to date.

2.4.2   Review of dMRI-based parcellation methods

The term connectional fingerprint was first used by Passingham and colleagues (2002) where connectivity in the macaque brain was studied through tracer data (however the term fingerprint had already been used to define the properties of a point in the brain by Hudspeth and colleagues [1976] regarding cell density across cortical layers). In their work, Passingham and colleagues characterized the connection strengths between 12 different points of the prefrontal cortex. These strengths were coded with a number from 0 (non-existent) to 4 (strong), and the fingerprints represented through radial diagrams (Figure 2.10). The fingerprints obtained can also be viewed as 12-dimensional vectors (where each dimension is the connectivity strength to each of the other points) with 4 possible different values at each dimension. Once the fingerprint is represented as a vector in multidimensional space it is possible to define a vector-based similarity measure such as the Euclidean distance or a correlation coefficient, as done in the study. The resulting matrix of correlations (containing a correlation value for each possible pair of fingerprints) was subjected to clustering, showing areas sharing similar connectivity patterns, dubbed connectional ''families''.

This work opened the door to anatomical connectivity based parcellation. Behrens and colleagues (2003a) were the first ones to use diffusion generated tracts for this purpose. In their approach, they generated probabilistic tractograms from each voxel in the thalamus, and assigned the voxels to one or other cluster based on which cortical area they reached with a stronger connection. As pre-requisite then, a predefined division of cortical targets was needed. The number of these divisions would define the number of regions in which the thalamus would be clustered (Figure 2.11). This work was also extended to study the variability of the parcellations obtained across subjects (Johansen-Berg et al., 2005).
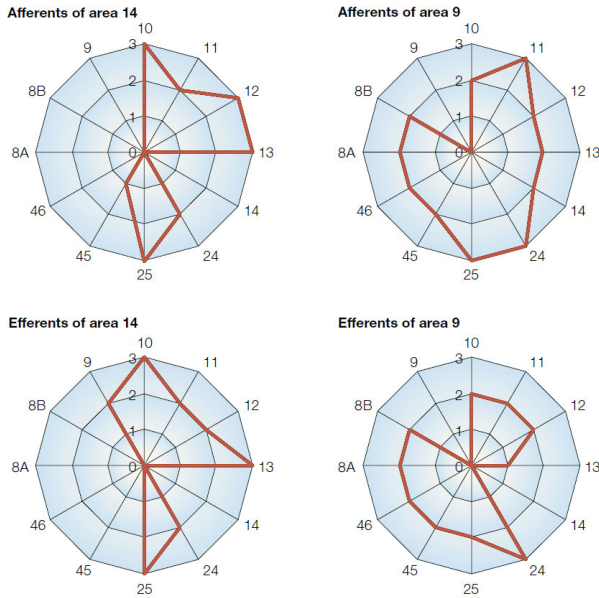
Figure 2.10:   *Dissimilar connectional fingerprints in two different areas of macaque prefrontal cortex (note that the distinction between afferent connections [carrying information input to that area] and efferent connections [transporting processed output to a different area] cannot be made from dMRI data). (reprint from (Passingham et al., 2002)).*
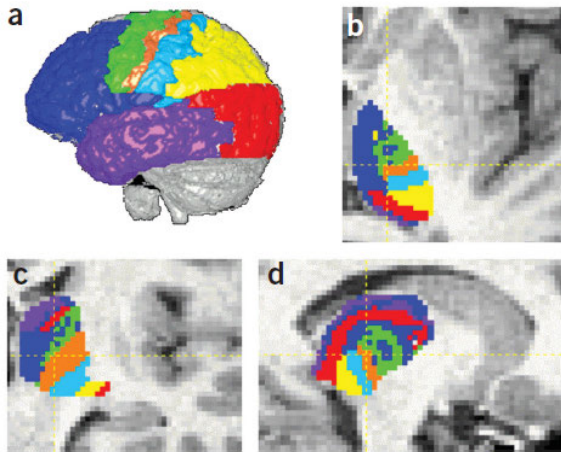
Figure 2.11:   *Cortical target atlas a) and connectivity target-based parcellation of the Thalamus b,c,d). (reprint from (Behrens et al., 2003a)).*

The first reported parcellation of cortical regions based on diffusion tractography corresponds to the work of Johansen-Berg and colleagues (2004). The whole white matter volume was chosen as target space for the tractography, meaning that the strength of connectivity at each voxel in the white matter was used to define the fingerprint, instead of only the values at the cortical ends. The tractograms were seeded from grey matter voxels of the medial frontal cortex and the connectivity values were binarized to reduce storage requirements. Correlation was used as similarity measure between fingerprints (by previously representing the connectivity values as a vector, this time with as many dimensions as voxels in the white matter). The resulting correlation matrix was processed through spectral reordering. This algorithm permutes the positions of the rows/columns and makes the existence of clusters apparent by visual inspection of the reordered matrix. Using this technique a change in connectivity profile and the corresponding boundary was found between the supplementary motor area (SMA) and pre-SMA (Figure 2.12)

In contrast to the approach used by Behrens and colleagues to parcellate the thalamus, this one does not require a-priori knowledge/assumptions on the connectivity patterns (such as each point connecting primarily to one area of a predefined atlas). This type of clustering is called 'blind' or 'free' clustering. Although in this case, the process was not automated as it required visual inspection to define the number of regions yielded and the exact boundaries of these regions.
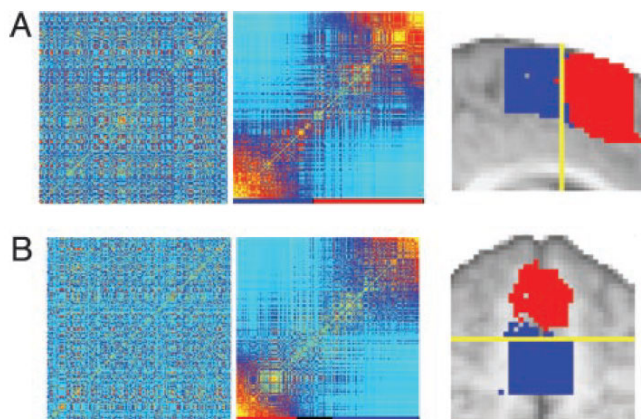
Figure 2.12:   Spectral-reordering parcellation of medial frontal cortex. (a and b) Result of parcellating a sagittal (a) and axial (b) slice in a single subject. Original (Left) and reordered (Center) cross-correlation matrices are shown. (reprint from (Johansen-Berg et al., 2004)).

An unsupervised method for defining boundaries between cortical regions was employed by Anwander and colleagues (2007) by using *k*-means clustering to divide Broca's area in different regions based on connectivity similarity (of fingerprints considering also whole white matter as target space and correlation as similarity).

In this type of clustering, the concept of distance between fingerprints is used (*distance=1-similarity*). After the distance matrix is obtained, several fingerprints are randomly chosen as cluster centers and the remaining ones are assigned to the cluster with the closest center fingerprint. Once all fingerprints are assigned, new cluster centers are computed for each cluster by finding the fingerprint with minimum total sum of squares distance to the other fingerprints in the same cluster. Then, the reassignment phase starts again and the process is iterated until the cluster centers do not change across integrations. While there is no more need for visual inspection to define boundaries, the number of initial centers, and therefore the number of clusters to be found, must be set in advance.

Some other interesting considerations were taken in this work: the connectivity probability values of the tractograms were not binarized, but this raised the issue of the intrinsic bias of probabilistic tractography to result in lower values of connectivity probability the further away a target voxel is from the original seed voxel. In order to compensate for this bias, a logarithm transform was applied to the connectivity values and these were then normalized.

k-means has become possibly the most popular clustering method for connectivity based parcellation. Tomassini and colleagues (2007) used it to identify dorsal and ventral sub-regions in the lateral premotor cortex (also compensating for the bias in connectivity probability values mentioned above by adding an Euclidean constraint matrix to the correlation matrix, thus reducing the effective similarity for voxels closer to the seed). Klein and colleagues (2007) followed up on the work by Johansen-Berg and colleagues (2004), studying the reproducibility of pre-SMA/SMA connectivity based boundary by applying *k*-means to the original tractography. Schubotz and colleagues (2010) applied the methods defined in (Anwander et al., 2007) to parcellate the lateral premotor cortex. Mars and colleagues (2011) studied the parietal cortex with *k*-means applied to tractograms obtained as per (Johansen-Berg et al., 2004), and Ruschel and colleagues (2013) used the methods defined in (Anwander et al., 2007) to subdivide the inferior parietal cortex further than the Brodmann divisions.

Despite its popularity, *k*-means suffers from the important limitation of needing the number of resulting clusters as a parameter (Hartigan, 1975). This means that either the number of expected regions to be found has to be previously known, or that a range of possible numbers have to be tested and the best result somehow decided.

Jbabdi and colleagues (2009) offer a quite different approach in order to automatically optimize the number of clusters directly from the data. Instead of defining a similarity measure between fingerprints and clustering according to these values, they assume that the different fingerprints originate from a mixture of Gaussian distributions (again we must interpret these fingerprints as points in multidimensional space). In this Bayesian setting, using a non-parametric model where the number of parameters (priors) can change adaptively depending on the data, a posterior distribution can be estimated directly from the data. This also means estimating the number of distributions that presumably generate the data observed, that is, the number of clusters. The model can also be extended to cluster several subjects simultaneously (through what is called a hierarchical mixture of Dirichlet processes), with the further advantage of direct correspondence of clusters across subjects and of using a population to estimate the optimal number of clusters rather than a single subject. They successfully applied this method to parcellate thalamus and SMA data from (Behrens et al., 2003a) and (Johansen-Berg et al., 2004).

The studies presented above focus on parcellation of localized areas of the brain, but when aiming for parcellation of the full cortex new challenges arise. Namely, that the sheer increment in data to be clustered and in expected number of clusters dramatically increases memory and computation requirements, making typical algorithms not directly scalable (there is a fair amount of whole-brain parcellation literature for resting-state fMRI studies, but the data volumes involved in this modality are far less restrictive).

Perrin and colleagues (2008) approach this issue through a double dimensionality reduction of the problem. Firstly, the cortical sheet is split into 36 large gyri using a sulcus recognition system (that then are subdivided using the connectivity information through *k*-means). Secondly, *q*-ball probabilistic tractography is computed from each voxel in the cortex, but only their overall connectivity to each of the 36 gyri is used as a fingerprint for clustering. The problem is then simplified to 36 smaller datasets where the connectional fingerprints have only 36 dimensions.

Roca and colleagues (2009) also tackle the connectivity-based parcellation but through a different set of dimension reductions. Firstly, the algorithm is applied only to connectivity profiles from the top of the cortical gyri. Secondly, the raw

connectivity matrix is smoothed to allow for enough overlap across profiles. Thirdly, the clustering is not performed on a whole brain basis but with an iterative patch by patch strategy. Lastly, the connectivity profiles are collapsed in an adaptive way for each patch to be parcellated: *"the segmentation for collapsing is based on the catchment basins of the watershed of the density of connection to the patch computed on the cortical surface"* (Roca et al., 2009). The clustering algorithm used on each patch is a variation of the *k*-means method (*k*-medoids) tested for several output cluster numbers and choosing the one that optimizes its silhouette (refer to paper for more details on the method).

All the clustering approaches so far reviewed tend to neglect the possibility of a hierarchical architecture underlying the cortex, but brain networks are more appropriately conceived of as forming nested modules (Bassett et al., 2010; Bassett and Gazzaniga, 2011). In the work of Gorbach and colleagues (2011) particular effort is made to characterize the hierarchical properties of cortical connectivity structure. The method proposed works in two steps. First, a global partition is made that defines the hierarchical level of maximum division. For that purpose it is assumed that in each cortical subunit a representative tractogram can be chosen that summarizes the connectivity pattern of the entire cortical subunit. The number of parcels in this global partition is determined by testing the robustness of the clustering solution against the uncertainty in the data, in order to let this uncertainty drive the choice of representative tractograms (and therefore the finest level of the hierarchy). Rate distortion theory is used to stochastically map remaining tractograms to exemplars allowing for a fuzzy partition between cortical areas (the amount of "fuzziness" can be controlled with a parameter called temperature, *T*). Secondly, to represent the nested structure of cortical subunits (making the prior assumption that such a structure does indeed exist) the so-called information bottleneck method is used. This method selects the mergers to be made on the basis of preserving as much information about the partitioning as possible in the new level with respect to representative tractogram (Figure 2.13).
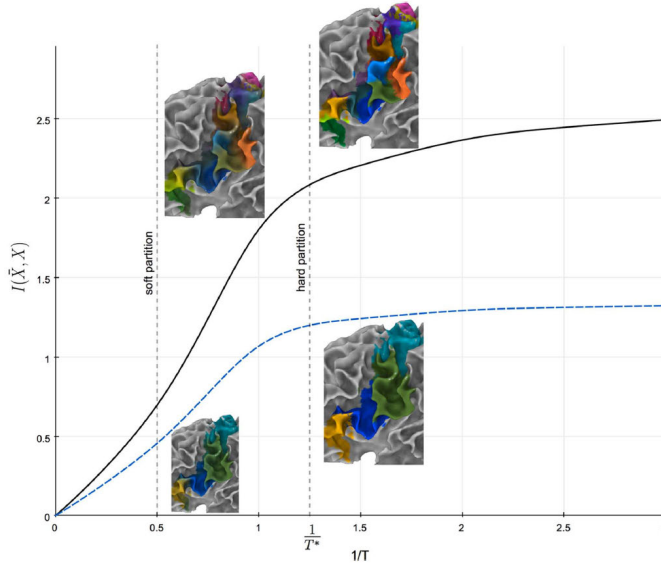
Figure 2.13: *Information rate plotted against inverse temperature. As the information rate increases finer structure is accounted for, and as temperature decreases, partitions are less fuzzy. (reprint from (Gorbach et al., 2011)).*

2.4.3 Limitations of the current methods and motivation for this work

In the previous section, we have reviewed the state of the art approaches on diffusion-based connectivity parcellation. However, for the purpose of whole brain characterization of connectivity similarity, as is our goal and motivation, they all suffer from limitations.

Target-based clustering (Behrens et al., 2003a) makes the strong assumption that each parcel should be mainly connected to one target area, and it requires the previous delineations of such targets. This is not applicable in the scenario of a whole brain parcellation, where the particular connectivity properties of less studied regions are not known and where considering the connectivity to only one target might not be enough to characterize differences between multiple regions. Furthermore the use of a prior delineation of main cortical targets defeats the purpose when the targets coincide with the seeds to be clustered, as would be this case.

On the other hand, free-clustering algorithms do not make this assumption, but the number of expected parcels, average size of clusters, or a similar parameter must be known in advance (Anwander et al., 2007; Tomassini et al., 2007; Klein et

al., 2007; Schubotz et al., 2010; Mars et al., 2011; Ruschel et al., 2013), posing a classical model selection problem. Furthermore, the most popular of these algorithms, *k*-means, has been shown not to work effectively on high dimensional data (Keim and Hinnenburg, 1999), and to be dependent on the particular initialization centers (Nanetti et al., 2009), it is possible to solve this issue via a Monte-Carlo approach, but the repetitions required for would grow with the expected number of clusters and render it impractical for the scales of a whole-brain scenario.

Also, most of the algorithms presented make the implicit assumption that there is a parcellation that can be considered a reasonably unique and complete representation of the connectivity similarity structure, which is rarely likely to be the case. Ultimately, a parcellation is an approximation to the connectivity properties of the data, and even when attempting to find an, in some sense, optimal parcellation purely from the data (Jbabdi et al., 2009) this might account for only a very small part of the underlying structure. When faced with a whole-brain approach, the challenge of not only having a high and unknown expected number of areas, but also that number being subject to the desired granularity of the partitioning, arises. Current whole brain-approaches only provide a two level hierarchy (Perrin et al., 2008; Roca et al., 2009). We would like to aim for a whole-brain partitioning that is based purely on connectivity information, and is capable of characterizing the connectivity similarity structure at multiple granularity levels.

In order to achieve this, we think it is important to use high resolution diffusion data, and to keep this resolution as high as possible in the connectivity fingerprints, as it might be a key element in obtaining high sensitivity to a change in connectivity pattern. This brings a further important challenge to overcome, as in high resolution both the fingerprint dimensionality and the number of points to cluster increase dramatically, with the corresponding strain in computation and memory requirements.
While the approach from Gorbach and colleagues (2011) has many interesting features (like possibility of fuzzy partitions and inclusion non-linear dependencies in tractogram similarity) and can account for hierarchical structure in the data, it is also computationally more expensive and it remains to be seen if it can scale up for a full-brain scenario of high resolution data.

In order to comply with our goals of a full cortical clustering that accounts for multiple granularity levels and is applied over high resolution and high dimensionality data, we propose the use of an agglomerative hierarchical clustering approach. This type of algorithms start by considering each point in the

data a single cluster and then iteratively merge these points until only one cluster remains including all the dataset. As output a hierarchical tree or dendrogram is obtained, which encodes the information of the similarities between datapoints in a much reduced dimensionality, and contains within parcellation information at all granularity levels, plus the hierarchical relations between these parcels.

While to our knowledge this solution has to date not yet been implemented for diffusion based connectivity cortical clustering, it has been already used for white matter clustering and for the parcellation of resting-state fMRI data. However, although the core clustering algorithm may be the same, the challenges and particular requirements for cortical based parcellation are very different than for these other modalities. Nevertheless it is interesting to briefly mention a few of these approaches.

As a curiosity note the work of Passingham and colleagues (2002; the first reported cortical parcellation using tracer anatomical connectivity) already featured a hierarchical clustering algorithm in order to study relationships between connectivities, although this was very low dimensionality data. In dMRI, agglomerative hierarchical clustering has been used to perform parcellation of white matter pathways (Wassermann et al., 2010). These were computed from deterministic tracts smoothed via Gaussian processes to allow for overlap, and once the tree was built a partition was chosen through best fit to a predefined cortical target atlas.

Guevara and colleagues (2011) also used hierarchical clustering for two of the steps in their multi-stage white matter clustering solution (decomposition of fibers according to presence in left/right hemisphere; length-based segmentation; agglomerative hierarchical voxel-based clustering; extremity-based clustering; hierarchical fascicle merging).

In the field of resting-state fMRI, hierarchical clustering has been used more widely. The first appearance corresponds to the work of Cordes and colleagues (2002) where functional time courses were clustered using their correlation as similarity measure. In their work however a single partition is selected from the tree out as final solution. This work was followed up by Stanberry and colleagues (2003) were the trees were processed *a posteriori* in order to improve the resulting parcellation, an interesting concept that we also explore in our work.

As a final example, the recent work of Blumensath and colleagues (2013) also features a whole brain hierarchical clustering of resting state data. Although the main focus of this work lies on a region growing algorithm to find a maximum granularity parcellation, upon which the tree is built.

In the following chapter, we will take some of the principles already used in these works and apply and specialize them for the objective of high resolution tract-based cortical parcellation.

3. A hierarchical method for whole-brain connectivity-based parcellation

3.1 Overview

In this work, the concepts of parcellation and different scales of brain structure are brought together, and *hierarchical clustering* is proposed as a solution to overcome the limitations exposed in the previous chapter. We aim to demonstrate that hierarchical clustering is a promising means by which to characterize the connectivity similarity structure of the whole human brain, where the information of the underlying structure at all granularity levels is encoded in a hierarchical tree or dendrogram. The idea is that these trees can then be sampled to obtain partitions at different granularity levels. This chapter will introduce the methods proposed and implemented along with their mathematical formulation. Part of the methods presented in this section have been published as an article in the scientific journal *Human Brain Mapping* (Moreno-Dominguez et al., 2014a) and in posters and talks at international conferences (Moreno-Dominguez et al., 2012a, 2012b, 2011b, 2011c).

3.2 Data acquisition and preprocessing

High resolution dMRI images as well as T_1 and T_2 weighted images were acquired for 4 young and healthy participants (3 males and a female) on a Siemens TimTrio scanner with a 32-channel array head coil and maximum gradient strength of 40 mT/m. For one of the participants, a second set of images was acquired after a one-week interval. Written informed consent was obtained from the subjects in accordance with the ethical approval from the *University of Leipzig*.

The dMRI data was acquired using spin-echo echo-planar imaging, with time repetition (TR) = 11s, echo time (TE) = 90ms, 85 axial slices, resolution 1.5 mm

isotropic, GRAPPA/3 (generalized auto-calibrating partially parallel acquisitions), and 3 acquisitions. We used 60 diffusion gradient directions, which were evenly distributed over the half-sphere (b-value = 1000 s/mm^2). The diffusion-weighted volumes were interspersed by acquisitions with no diffusion weighting (b0 images) at the beginning and after each block of 10 volumes (7 volumes). The total scan time for the dMRI protocol was approximately 45 min.

As a first preprocessing step, the 3D T$_1$-weighted (magnetization prepared-rapid gradient echo, TR = 1300 ms, time to inversion = 650 ms, TE = 3.93 ms, resolution 1.0 x 1.0 x 1.5 mm, 2 acquisitions, reconstructed to 1mm isotropic resolution) images were reoriented to the mid-sagittal plane through the anterior and posterior commissures and the brain volume was segmented using the *Lipsia* software package (Lohmann et al. 2001). The 21 images without diffusion weighting were used to estimate motion correction parameters using rigid-body transformations (Jenkinson et al., 2002), implemented in *FSL* (*FMRIB Software Library, Oxford, UK*). Motion correction parameters were interpolated for all 201 volumes and combined with a global registration to the T$_1$ anatomy using a mutual information registration algorithm. The diffusion gradient direction for each volume was corrected using the rotation parameters. The registered images were linearly interpolated to the new reference frame with an isotropic voxel resolution of 1 mm and the three corresponding acquisitions and gradient directions were averaged. Next, the diffusion tensor was calculated for each voxel after logarithmic transformation of the signal intensities (Basser et al., 1994). Finally, the fractional anisotropy of the tensor in each voxel was subsequently determined, and a multi-slice FA image (Basser and Pierpaoli, 1996) was created. The combined motion correction and registration to the individual T$_1$ anatomy provided some advantages. A simple motion correction to the first image in the diffusion weighted sequence would have introduced a variable amount of smoothing caused by the interpolation of the images to the reference image. E.g. the first images in the sequence would have needed less interpolation and the reduced smoothing would have caused a directional bias. Using the independent orientation of the T$_1$ image as reference removed this potential bias. Additionally, the sampling of the data with a higher spatial resolution (1mm instead of 1.5mm) allowed keeping more details of the data compared to a resampling with the original resolution. In this way, interpolation of the raw data provided some methodological advantages in the following tractography step.

It is possible to acquire fieldmaps in order to unwarp the diffusion images and correct for distortion (Jenkinson, 2004; Jezzard and Balaban, 1995) (although in this study they were not available for all datasets). In these cases, the transformation matrices for motion correction, distortion unwarping, and

registration to T_1 anatomy are obtained independently and combined into a single transformation using the *fugue* and *convertwarp* utilities of *FSL* (in addition to those introduced in the previous paragraph). The dMRI data was then transformed using this single transformation in order to avoid multiple interpolation steps and unnecessary smoothing of the diffusion data.

The brain volume was segmented into white and gray matter compartments by means of FA thresholding (white matter: FA ≥ 0.15) and interactive corrections for deep white matter imperfections. Using an FA based mask allows to define seed voxels at a clearly defined white matter boundary. This precession would not have been possible using the white matter mask from the segmented T_1 image, since the diffusion image shows small non-linear distortions. Each white matter voxel that neighbored a cortical gray matter voxel was used as a seed voxel for the probabilistic dMRI tractography (that is, each single grey matter/white matter boundary voxel at 1 mm resolution, between 130,000 and 200,000 seed voxels per brain depending on size), as proposed by Anwander and colleagues (2007). The tractography algorithm computed a transition probability of a simulated particle jumping from one voxel to the next from the diffusion data. Next, the probabilistic tractography started 100.000 particles in each seed voxel. The particles propagated in the white matter as guided by the local transition probabilities, defined by the probability density function from the diffusion tensor model. The target space was the whole white matter volume with a resolution of 1 mm³. The diffusion data was not interpolated in this step and used the interpolation of the raw diffusion data as computed in the preprocessing steps. Finally, a visitation map was computed from the number of particles which cross each voxel. The tractography algorithm was parallelized and implemented on a consumer PC graphic board (GPU) and took only a few seconds per seed point.

The 3D distribution of the connectivity values (visitation map) of a particular seed voxel with all voxels in the brain is called a tractogram. In these tractograms, which we use as connectivity fingerprints, the value associated with a particular white matter voxel represents the visitation fraction, that is, what proportion of all particles started at the seed voxel went through that particular voxel. The visitation values ranging between 0 and 100,000 were *log* transformed to reduce the dynamic range (in order to palliate the intrinsic bias that visitation-based connectivity values have towards favoring short connections against longer distance ones, which are especially problematic for the computation of similarities between tractograms) and scaled between 0 and 1 (1 means all, 0 means none of the started streamlines touched the voxel). These values are taken as a correlate for the anatomical connectivity between that voxel and the seed voxel of the tractogram. Although based on a simple local model (diffusion tensor), this

probabilistic tractography can, to a certain extent, account for fanning fibers and fiber crossings. This provides tractograms with enough overlap area to detect connectivity pattern differences between voxels at the discrimination level required for successful parcellation.

The tractograms obtained constitute the data points for our clustering method. At the resolution of 1mm³ which we will use, a typical dataset will consist of around 100,000 points (tractograms from seed voxels within one hemisphere) in *n*-dimensional space with *n* having a value between 600,000 and 800,000 (number of white matter voxels in the brain).

To analyze the effects of a reduced signal-to-noise ratio (SNR) onto the developed analysis methods, a second set of tractograms was obtained for the first three subjects using just a single acquisition of the diffusion data (in contrast to averaging the 3 available acquisitions).

Despite this particular choice of tractography, the effort of this thesis is on the characterization of the connectivity similarity structure based on tractography, not on the tractography itself. The methods developed to that objective aim to be valid for any particular tractography technique chosen (tuning possibly the distance measure used tone that properly captures the changes between the particular tractograms yielded by that technique). Therefore a change to different tractography could be easily introduced in the method pipeline. If the description of connectivity similarity is enhanced by the new technique it would only further improve the results of our analysis.

3.3 Tractogram distance measure

In order to perform any kind of clustering a distance measure between the object points must first be defined. This distance quantifies the similarity between the connectivity patterns of two seed points. It must satisfy the properties of symmetry, non-negativity and identity of indiscernibles:

$$\text{Symmetry:} \quad d(x,y) = d(x,y) \text{ for any } x,y. \tag{3.1}$$

$$\text{Non-negativity:} \quad d(x,y) \geq 0 \text{ for any } x,y. \tag{3.2}$$

$$\text{Identity of indiscernibles:} \quad d(x,y) = 0, \text{ if } x = y. \tag{3.3}$$

If the triangle inequality is also satisfied the distance measure is also a metric.

$$\text{Triangle inequality:} \quad d(x,y) \leq d(x,z) + d(y,z) \text{ for any } x,y,z. \tag{3.4}$$

While the Euclidean distance is one of the most commonly used ones for low-dimensional data, it does not score well for scaling patterns or very high dimensionality (Wang et al., 2002, Beyer et al., 1999).

The correlation coefficient is a convenient way to measure the linear dependency between two variables and it has been previously used as a similarity measure between tractograms with successful results (Anwander et al., 2007).

$$\rho(X,Y) = \frac{\Sigma_i(x_i - \eta_X)(y_i - \eta_Y)}{\sqrt{\Sigma_i(x_i - \eta_X)^2}\sqrt{\Sigma_i(y_i - \eta_Y)^2}} \text{ with } \eta_X = \frac{1}{n}\Sigma_i x_i \tag{3.5}$$

where x_i is the ith element of tractogram X, and n is the number of elements in the tractogram.

Correlation as such can also produce negative values, which cannot be sensibly interpreted for spatial connectivity patterns (two uncorrelated patterns are just as dissimilar as two negatively correlated ones). That is why we modified the measure by omitting the centering. That is, the similarity between tractograms would be calculated as their dot product normalized by the multiplication of the norms, following the equation:

$$NDP(X,Y) = \frac{\Sigma_i(x_i \cdot y_i)}{\sqrt{\Sigma_i(x_i^2) \cdot \Sigma_i(y_i^2)}}, \tag{3.6}$$

The working principle in this measure is the same as in the traditional correlation, widely established, with the difference that negative correlations are disregarded and the discerning power is focused in positive correlations that have no negative linear dependencies, which is better suited for comparison of anatomical tracts. Geometrically speaking, the proposed measure relates to the scaled projection of one vector on the other, while the correlation relates to the cosine of the angle between the vectors. Both measures are closely related. As our tractograms contain very many zeroes, this causes the mean values to be very small. In consequence the differences between our measure and classical Pearson's correlation are minimal.

A third option for calculating tractogram similarity is to use the measure known as mutual information. This measure has the advantage of also capturing not only linear dependencies, but also higher order dependencies between the vectors. It has also been successfully used for measuring anatomical connectivity similarity (Gorbach et al., 2011) and is defined as:

$$I(X,Y) = \sum_{x \in X}\sum_{y \in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \tag{3.7}$$

This measure is also computationally more expensive than the previous two. Given that we aim to achieve whole-brain parcellation in our work, where computation load is a great challenge, that the distance measure is the main contributor to the method complexity, and that linear measures have given good results in the past for the specific tractograms that we will be using, a normalized dot product based distance is chosen. The tractogram distance would then be defined as:

$$d(X,Y) = 1 - NDP(X,Y)$$

(3.8)

Notwithstanding, we would like to point out that, as with the choice of tractography, the hierarchical method developed in this thesis should remain valid for different choices of distance measure, as long as the combination of tractography and distance is able to capture the connectivity pattern dissimilarity between different points of the cortex. If with higher availability of computational and/or time resources a study wished to be made using mutual information, it would be possible to implement this change without changing the basic workings of the method proposed.

A distance based on any of the above proposed similarities would not be a complete metric since it would not be guaranteed to satisfy the triangle inequality, but this does not lead to any shortcomings for clustering purposes.

In order to render the similarity measure robust to random artifacts in the probabilistic tractography, connectivity values smaller than 0.4 (less than 100 out of 100,000 seeded particles, as visitation values are *log* transformed and normalized) are set to 0 prior to computing the similarity (Anwander et al., 2007). This value was chosen in order to eliminate only minimal noise and remain conservative (as any target voxel visited by more than 0.1% of the seeded particles will be considered), but the best threshold for probabilistic tractography is still an open question in literature (Jones, 2010).

3.4 Hierarchical clustering

3.4.1 Agglomerative vs. divisive hierarchical clustering

Hard clustering methods (methods where a point either belongs or doesn't belong to a certain cluster) can be classified into partitional algorithms (dividing the data into a single partition) and hierarchical algorithms (which obtain a series of nested partitions). It is our hypothesis that a series of nested partitions would be better suited to characterize the connectivity similarity information of a whole brain, being able to capture it a different granularity levels.

Hierarchical algorithms are further subdivided into agglomerative and divisive algorithms. Agglomerative (bottom-up) hierarchical clustering starts by considering every object in the dataset as a separate cluster, then it merges the closest (i.e., most similar) pair of clusters, according to some similarity criterion, and iterates until all of the data points belong to one single cluster. The result is essentially a binary tree. An outline of the clustering process applied to anatomical connectivity can be seen in Figure 3.1.
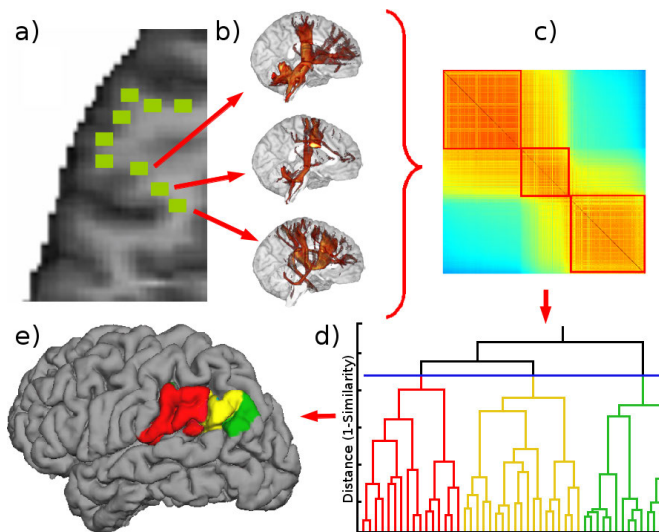
Figure 3.1:   Schema of the hierarchical clustering process: a) Select gray-matter/white-matter interface voxels; b) Generate probabilistic tractograms of seed voxels; c) Compute similarities between tractograms; d) Build-up connectivity tree; e) Select partitions within the tree and map back to the cortex.

On the other hand, divisive (top-down) hierarchical clustering works in the opposite fashion. Starting with a single cluster covering the whole dataset, further subdivisions are obtained iteratively using a secondary partitional algorithm. Divisive clustering is conceptually more complex than agglomerative clustering since a second clustering algorithm is needed as a "subroutine". Also, although both types of hierarchical clustering suffer from the disadvantage of any merging/division decision being irreversible and any errors being dragged through the rest of the hierarchy, this is critical in the top-down case, as boundary errors in an early stages span across a big portion of the dataset and will more dramatically affect the following partition decisions. In some cases where only the coarser granularity levels are desired, the top-down method can have the advantage of being more efficient, as instead of obtaining the hierarchy all the way down to individual datapoints some stopping criterion is implemented and unwanted computations are avoided. However for very large datasets (as is our case) the secondary partitioning algorithm is more likely to have significant errors in the early stages, or require more complex computations reducing or nullifying the previously mentioned advantage. For this reason a divisive approach is discarded for our particular objective, in favor of an agglomerative one.

But even within the bottom-up approach, several modalities exist; differing in the way new distances are computed to a third element once two elements have been joined (*linked*; Murtagh, 1983). Several possible solutions were implemented and tested, in order to find the method that best performs when clustering whole brain connectivity data.

3.4.2   Graph methods

There are two main groups of agglomerative hierarchical clustering methods. The first of these is integrated by the *graph methods*, stemming from graph theory, where a cluster can be represented by a subgraph of interconnected points. There are four possible linkages within the graph methods:

In the *single linkage* method (Florek et al., 1951; Sneath, 1957), the new distance to a third cluster will be the smaller of the two distances from any of the joining elements to that cluster before the merge:

$$d_{SL}(xy,z) = \min(\, d(x,z), d(y,z)\,) \tag{3.9}$$

where x and y are the clusters being merged, xy is the resulting new cluster and z is a cluster not being merged at that particular step.

In the *complete linkage* method (Johnson, 1967), the new distance to a third cluster will be the greater of the two distances from any of the joining elements to that cluster before the merge:

$$d_{CL}(xy,z) = \max(\, d(x,z), d(y,z)\,) \qquad (3.10)$$

In the *average linkage* method (Jain and Dubes, 1988), the new distance will be the mean of the distances from the joining clusters weighted by the number of data points each cluster holds, in other words, the new distance will be the average of all the pairwise distances between de points contained in clusters x and y with the points of cluster z .

$$d_{AL}(xy,z) = (\, S_x{\cdot}d(x,z) + S_y{\cdot}d(y,z)\,) \,/\, (S_x + S_y) \qquad (3.11)$$

where S_i is the size or number of data points contained in cluster i.

In the *weighted linkage* method (Jain and Dubes, 1988), the new distance will be the unweighted mean of the distances from the joining clusters.

$$d_{WL}(xy,z) = (\, d(x,z) + d(y,z)\,) \,/2 \qquad (3.12)$$

In all four of these methods, it is necessary to calculate the pairwise distances between all data points. After obtaining the full connectivity distance matrix, the two most similar elements are joined and the corresponding rows and columns are combined into a new row and column as per the desired linkage equation. The process is then iterated until the tree is completed. The main disadvantage of the graph methods is that computing the full distance matrix can prove costly when there is a large number of elements and the points are in a very high dimensional space, as is the case considered here. Additionally, once obtained, the matrix should be fully loaded in the random access memory in order for the algorithm to be efficient, which can amount to a considerably large volume typically only available in computing servers.

3.4.3   Centroid Method

3.4.3.1   Basic method

The second type of agglomerative clustering methods is integrated by the *geometric methods*, where a cluster is represented by a center point. In order to reduce the computation and memory requirements, we elected to use a fifth method based on one of the geometric methods: the *centroid linkage* method (Jain and Dubes, 1988). In this method, each cluster is defined by its centroid: a data point that represents all the points included in the cluster. In the study presented here, the centroid was computed as the average of the non-thresholded tractograms in natural space of the contained data points. In this scenario, when two clusters merge, the mean tractogram of the new cluster is computed, and the new distances to the rest of the clusters are recalculated. A newly merged point is then defined as:

$$xy = \Sigma_i \left(S_x \cdot x_i + S_y \cdot y_i \right) / \left(S_x + S_y \right) \tag{3.13}$$

3.4.3.2   Neighborhood restriction

In principle, the centroid method involves an extra computing effort, as the new distances that must be calculated in every merging step involve normalized dot products high-dimensional mean tractograms (*Equation 3.13*), opposed to *max*, *min*, or averaging operations of distance values in the graph methods. However, it can also be used to avoid the necessity of calculating the whole pairwise distance matrix by means of applying a neighborhood restriction. If the assumption is made that a connectivity defined region in the brain must always be a connected patch of gray matter, then only mergers between neighboring clusters are allowed, and only those distances have to be computed, drastically reducing the cost of the algorithm (a neighborhood restriction may also be used in the graph methods in order to force morphologically continuous clusters, but the whole distance matrix must still be calculated and thus it yields no computational advantage). The concept of spatially constrained hierarchical clustering has also been exploited for fMRI data (Blumensath et al., 2013), a modality where this particular restriction has proved of advantage for parcellation in the past (Craddock et al., 2012).

Several neighborhood levels may be chosen. The following neighborhoods were implemented in this study: 18 ($d_v = \sqrt{2}$), 26 ($d_v = \sqrt{3}$), 32 ($d_v = 2$), 92 ($d_v = 2\sqrt{2}$) and 124 (not defined by a value of d_v) where d_v stands for the maximum distance (in voxel units) of a neighbor voxel from the seed voxel (Figure 3.2).
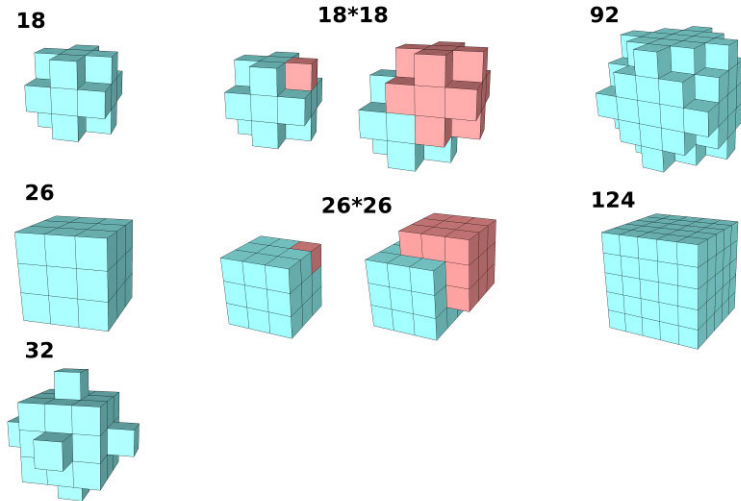
Figure 3.2:   *Neighborhoods model implemented: 18, 26, 32, 92, 124. 92 and 124 neighborhoods are obtained through the convolution of two 18 or 26 neighborhood kernels, respectively.*

As we worked with high resolution 1mm images, there is no risk of adjacent voxels corresponding to the grey matter – white matter interface of opposite gyri. In the case of the 92 and 124 neighborhoods, however, which expanded to non-adjacent voxels, there was a risk of considering an element as a neighbor that resides in a different gyrus. To avoid this, the algorithm was implemented as the convolution of two smaller neighborhoods kernels: 18*18 yields a 92 neighborhood, while 26*26 leads to 124 neighbors. In this sense, the smaller neighborhood was scanned, and if neighbors were detected, the respective neighborhood of each one of them was considered as well. This way, neighbors are considered as such only if they form a continuous sheet around the seed voxels. The results are analogous to what would be obtained through surface analysis with only a fraction of the cost.

3.4.3.3   Homogeneous merging

One of the advantages of using a hierarchical method is the possibility of comparing the full connectivity structure across datasets through tree comparison, which we will further develop in a later subsection. In order to do this, the terminal elements or leaves of the trees to be compared must first be matched. With this in mind, an extra restriction was applied to the centroid method during the initial

iterations of the tree building process. The objective was to ensure that at the lower levels of the tree (that is, the ones with highest granularity) the clusters were joined in a homogenous way, with roughly equal sizes, until a certain number of clusters had been reached. As will be explained later, this allowed for easier matching. There is, however, no restriction upon the shape of these clusters, and their merging is still guided by connectivity pattern similarity. Although this restriction mathematically involves a certain loss of information in comparison with the restriction-free algorithm, it was expected that the information contained at these very high granularity levels would be very low due to the sensitivity limits of the tractogram distance measure (which is shown in chapter 4). This step will improve matching as the homogeneous clusters produce wider and smoother mean tractograms that overlap better across datasets. A conceptually similar 2-stage clustering approach has also successfully been used by Gorbach and colleagues (2011) and Blumensath and colleagues (2013) to partition dMRI and fMRI data, respectively (where first a maximum granularity partition is obtained from which to build the hierarchical tree, although their particular implementations differ greatly with our solution. See section 2.4.2).

```
smallestSize = 1;
activeSize = 1;
// homogeneous merging stage
while ( cluster# < N )
{
    Find most similar neighboring cluster pair;
    Verify size1 == smallestSize and size2 <= activeSize, otherwise choose next;
    if ( no pair matches the conditions )
    {
        if ( there are no more clusters with size == smallestSize )
        {
            ++smallestSize;
            activeSize = smallestSize;
        }
        else
        {
            ++ activeSize;
        }
    }
    else
    {
        Merge clusters and recalculate distances to neighbors;
    }
}
// unrestricted stage
while (active clusters > 1 )
{
    Find most similar neighboring cluster pair;
    Merge clusters and recalculate distances to neighbors;
}
```

Pseudocode 3.1:   Pseudocode for the final modified centroid method devised and implemented, including an initial homogeneous merging stage and neighborhood restriction.

For simplicity, from this point on, this modified algorithm including neighborhood restriction and initial homogeneous merging stage will just be referred to as the *centroid method* or cXX where *XX* indicates the neighborhood level used (*Pseudocode 3.1*).

3.4.4   The hierarchical tree or dendrogram

As an output of these algorithms (both graph and centroid) a binary tree (also called bifurcating rooted tree or fully resolved dendrogram) is obtained (like the one shown in Figure 3.1d). This tree encodes the connectivity similarity structure of the dataset at all granularity levels, transforming into a much-reduced dimensionality ($2n$, n being the number of seed elements) the information of the distance matrix (dimension n^2) obtained from the tractogram space (dimension $n \cdot m$, m being the number of white matter voxels).

The terminal elements of the tree (those at $y = 0$) represent the original data points (single seed voxel tractograms) and are called *leaves* of the tree (therefore in our plotting convention the bottom of the tree represents the highest granularity). The intermediate points of the tree, where two elements join, are called *nodes* of the tree (and in the centroid method are also identified by the mean tractogram of their contained leaves). In a binary tree, like the ones obtained with the methods described, a node is always formed by joining two elements. Trees can also be non-binary (also called partially-resolved dendrograms) where a node can be formed by the simultaneous joining of any number of elements.

The *x* axis of the tree identifies a specific data point ($y = 0$). However, data point positions are not absolute, that is, a tree obtained from the same data set but with a different algorithm might not have the same leaves at the same *x* locations, as these may have been reorganized in order to be able to plot the tree without line crossings. The y axis of the tree indicates the distance value between the elements joining at a node at that particular level. It also encodes the distance between any pair of leaves that meet in a node at that level (tracking their path upwards through any number of intermediate nodes).

3.5 Assessing the quality of the trees

In order to choose the hierarchical clustering method that best represents the information in the original dataset, a validity measure has to be defined to assess the goodness of fit of the obtained dendrogram. The cophenetic correlation coefficient (*CPCC*) (Farris, 1969) serves this purpose by calculating the degree of agreement between the distances encoded in the tree (named cophenetic distances, obtained by looking at the distance value of the merger where the desired elements are found in the same cluster for the first time) and the pairwise distances obtained from the original tractograms. It is defined as

$$CPCC = \frac{\frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_{ij}c_{ij} - \mu_D\mu_C}{\sigma_D\sigma_C} \qquad (3.14)$$

with $\sigma_X = \sqrt{\frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij}^{2} - \mu_X^{2}}$, $M = \frac{n(n-1)}{2}$ and $\mu_X = \frac{1}{M}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij}$,

where *n* is the total number of elements and d_{ij} and c_{ij} are the distance values between elements *i* and *j*, as computed from the tractograms or obtained from the tree, respectively. The range of *CPCC* is [−1, 1]. The higher the value, the better the fit between the tree and the data, a value of 1 indicating that the matrix and the tree contain exactly the same information (there is a linear dependence between both, which is not possible unless the distances between all the tractograms are equal) and a value of 0 meaning that the tree contains none of the original information (due to the nature of the hierarchical agglomerative method, negative *CPCC* values will not occur).

3.6 Tree processing

3.6.1 Confounds and challenges for dendrogram interpretation

The resulting dendrograms serve two purposes: On the one hand they are a compression of the pairwise similarities between connectional fingerprints, and on the other hand they also hold information on the similarities between clusters at every possible granularity and the hierarchical relationships between them, allowing for easy and quick partition generations. They are, however, complex structures and their interpretation and partition selection are not always straightforward. In addition, several factors might add confounds and complicate the analysis.

- *Artefactual datapoints*

As in most types of clusterings, these can produce unwanted outliers that obscure the data and introduce errors in the analysis. In our particular case errors and spatial discontinuities in the mask of seed voxels might result in unusable tractograms characterized by a very limited number of target voxels reached. This results in a very low similarity of these tractograms to the rest.

- *Non-monotonicity*

In the most widely used linkage methods, the distance between a newly merged group of elements and the rest of the set are computed as a weighted average of the distance between elements (as in the graph methods, where the type of weighting defines the type of linkage). This means that this distance is always equal or greater than the distance between the groups that existed prior to the merge, resulting in a monotonic tree. In the centroid method, however, this is not always the case. As each group of elements is represented by a new representative centroid, this centroid could be closer to other elements than any of its components were before the merging (Morgan and Ray, 1995), which is called an inversion. In other words, it can happen that the intra-cluster distance exceeds the inter-cluster distance. These inversions or non-monotonic steps can appear when more than two points in the data have very similar distances to each other, and indicate areas with no clear binary cluster structure (Gower, 1990). As a toy example, if we consider points in 2D space positioned like vertices of a roughly equilateral triangle and use Euclidean distance, the centroid of 2 merging points will be closer to the third point than any of them were before (Figure 3.3). While these inversions do contain information about the distances encoded (when the

tree is seen as a compression of the similarity matrix) they do not provide any additional information on the hierarchy structure, and they make interpretation of the hierarchy and tree analysis difficult and inconvenient (Murtagh, 1985).
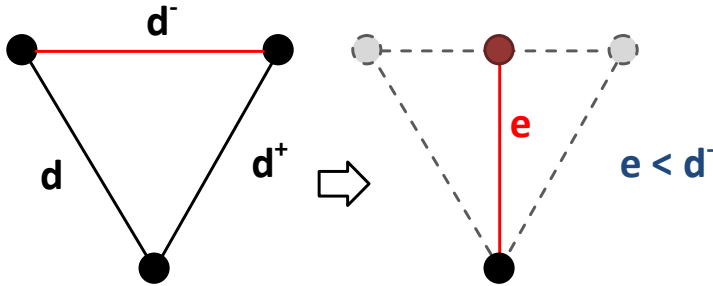
Figure 3.3:   *Toy example of a centroid linkage merging leading to a non-monotonic step in a hierarchical tree. The first two point would merge first as their distance is the smallest, and the distance of their centroid to the third point is even smaller than any of the initial distances.*

- *Hierarchy-resolution limitation at highest granularities*

The proposed method produces connectivity profiles with a very high spatial sampling resulting from seeding tractography at the white matter boundary with a voxel resolution of 1mm. This produces an oversampling of the diffusion profiles compared to the limited spatial resolution of the diffusion acquisition and the uncertainty of the tractogram computation. As a result seed points with a very high similarity cannot be distinguished (for this reason, neighboring seed points with very high similarity are grouped together to base-areas as part of the proposed tree-building algorithm). The hierarchical relationships within these base-areas are characterized by several consecutive mergers with very small distance change indicating the non-separability of these regions and the irrelevance their internal structure for the hierarchical tree, while adding to the complexity of the tree.

Forced binary structure

As mentioned before, the iterative nature of the clustering process forces the dendrogram to always have binary bifurcations, whereas in reality the dataset may have structures nested in a non-binary way. This means that some of the nodes in the tree do not contribute to any real information about the similarity structure and are merely a by-product of the pair-wise agglomerative method.

3.6.2   Dendrogram preprocessing pipeline

In order to address these problems and ease the information extraction, four dendrogram preprocessing steps were developed: outlier elimination, monotonicity correction, limiting the maximum encoded granularity and collapse of non-binary structures. These steps, detailed below and exemplified in Figure 3.4, effectively reduce the number of branchings, and in turn reduce the tree complexity and possible confounds in the dendrogram, while still maintaining maximum usable information (shown quantitatively in section 4.3).
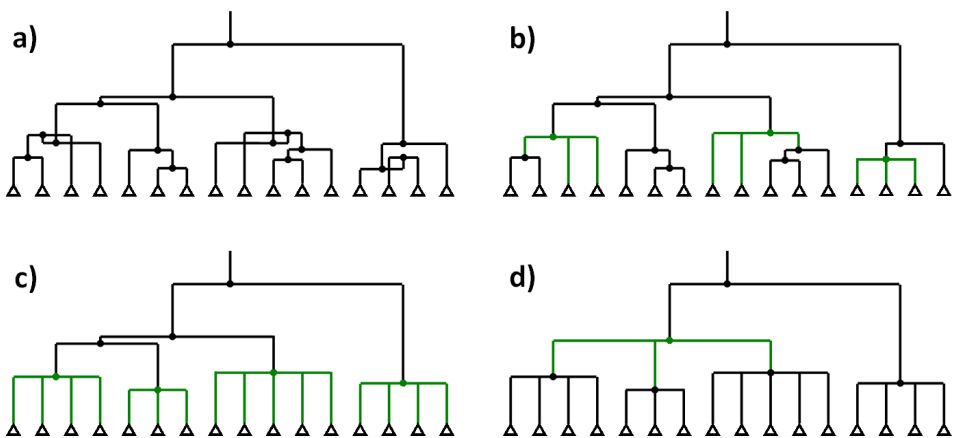
Figure 3.4:   Dendrogram pre-processing: example raw tree (a), monotonicity correction (b), limiting the highest granularity encoded (c) and collapse of non-binary structures (d).

- *Outlier elimination*

Isolated leaves resulting from faulty tractograms can easily be detected and eliminated without negative influence on the whole brain coverage. Data points with a distance value compared to their most similar neighbor higher than a threshold were discarded and removed from the analysis. This step was actually implemented as part of the tree building algorithm, in order to prevent the outliers from affecting the value of the centroids. Removing these outliers in general stabilizes the tree and the clustering result and simplifies its interpretation.

- *Monotonicity correction*

As inversions occur when more than two elements are at similar distances from each other, it is possible to transform the non-monotonic trees of the centroid method into monotonic ones with little information loss. This is accomplished by merging every two nodes where an inversion occurs, creating a non-binary branching with more than two nodes joining simultaneously into one (Figure 3.4b). This non-binary structure more parsimoniously describes the original information present in the data. For each correction, the level value of the simplified node is calculated as the mean of the levels of the original nodes, weighted by their respective sizes in terms of number of leaves. Corrections are applied starting at the root node and working through the tree down to the leaf level.

- *Limiting maximum granularity*

In terms of tree processing, the small differences between the leaves in the base-areas are ignored and the tree is transformed in a so-called rose tree, where the meta-leaves branch into single voxels (leaves, Figure 3.4c). The partition defined by these meta-leaves would then represent the maximum effective granularity achievable from the data. While rose-trees can be computed directly from data (Blundell et al., 2010), the computation costs are far greater than with the method proposed here.
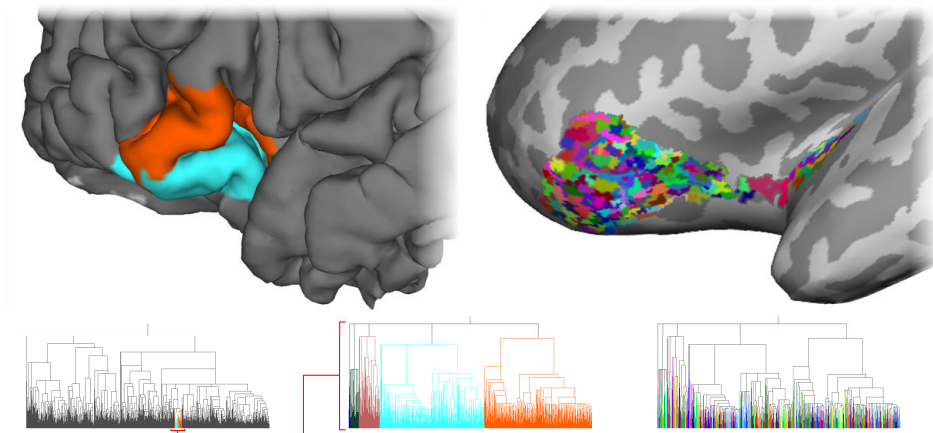
Figure 3.5: *Detail of the clusters contained by the sub-tree covering the IFG region of the left hemisphere of subject A (upper left) and its position in the complete tree (lower left) as well as a view of the zoomed-in sub-tree (lower center). The meta-leaves contained in the mentioned sub-tree have been projected onto the inflated surface (upper right) and the zoomed-in sub-tree (lower right).*

In our implementation, the meta-leaves are the homogenous clusters obtained during the first stage of the proposed centroid algorithm. All branchings within those nodes are then eliminated and their contained data points joined simultaneously at the original node level. Additionally, this grouping sharpens the connectivity profiles of the meta-leaves and allows for a better identification of connectivity similarities and differences between neighboring regions. A detail showing the meta-leaves obtained for part of the tree of one of the subjects studied in this thesis is presented in Figure 3.5.

- *Collapse of non-binary structures*

Cases where non-binary structures are present in the data are generally characterized in the tree by merges where the distance change is much smaller than the absolute distance level of the nodes being merged (when not resulting in an inversion). The dependency on the distance level accounts for the fact that the significance of distance change is the lower the higher a node stands in the tree hierarchy. A similar leveling concept to the one used with the non-monotonic steps was used here, flattening any merging with a distance change smaller than a certain proportion of the absolute distance value of the node considered. Constant and square dependencies were also considered, but the linear solution proved the best trade-off between complexity reduction and information loss. The resulting tree will be a better representation of the original data and will have a considerably reduced number of internal nodes, making it easier to identify natural divisions in the data (Figure 3.4d).

The preprocessing methods described in this section effectively reduce the number of branchings, which in turn reduces the tree complexity and possible confounds in the dendrogram, while still maintaining maximum usable information (shown quantitatively in the *Results* section). This also facilitates the task of the information extraction algorithms, introduced in the following sections.

3.7 Dendrogram comparison

3.7.1   Introduction

Once the connectivity similarity structure of a brain is encoded in a dendrogram, there is the possibility of using the information from the whole tree to assess the structural differences in brain connectivity between different subjects or measurements.

Dendrogram comparison techniques are already in use in other fields, with most efforts being dedicated in the field of genetics (Critchlow et al., 1996; Restrepo et al., 2007). However, these techniques are used to compare different trees built over the same dataset, relying on a perfect match between the leaf elements of both trees. In the scenario of brain connectivity trees from different measurements, this would only be the case if the dendrograms being compared originate from the same brain, and only if there have not been significant changes in morphology nor the data acquisition method.

3.7.2   Leaf matching across trees

In order to be able to apply these comparison methods when assessing connectivity structure variability across subjects, the problem of leaf identification had to be tackled. Potentially, there are different possible criteria for the identification of associated pairs of leafs in two dendrograms, for example the amount of spatial proximity after a more or less sophisticated co-registration of the images or cortical surfaces derived from these images. However, as the dendrograms to be compared are based on the similarity of tractograms, it seems appropriate to use the same criterion for finding matched pairs of leafs. The solution provided involved several steps:

- First, the trees are pre-processed with the techniques previously introduced, in order to reduce the number of leaves and provide a maximum granularity partition. These maximum granularity partitions are fine-tuned so that all the trees have the same number of meta-leaves. This number is chosen to obtain meta-leaves big enough to produce wider and smoother mean tractograms that will overlap better across datasets and make matching easier, and also provide an acceptable complexity reduction (the lower the number of meta leaves in each tree, the smaller the matching distance matrix that needs to be computed will be) while incurring minimal information loss due to the elimination of highest granularity levels.

- Mean tractograms corresponding to each of the meta-leaves are obtained for all subjects. The mean tractogram of any given node is calculated as the log-transformed average of the raw (not log-transformed) seed tractograms contained in the respective node.

- The subject FA images are non-linearly registered to a common space, and this transformation was applied to the mean tractograms. The registration is performed through the *ANTS* package (SyN registration algorithm; Avants et al., 2008; Klein et al., 2009). The mean tractograms are then transformed to the same common space using the deformation fields obtained from the FA image registration.

- For each pair of trees being compared, a tractogram distance matrix between their corresponding meta-leaves is obtained.

- Matching of the meta-leaves of the trees is done by applying a greedy algorithm to the distance matrix: The two tractograms with the highest similarity are matched and their entries are eliminated from the data. This step is iterated until there are no more entries in the matrix. In order to avoid poor matches, restrictions on minimum tractogram similarity and maximum Euclidean distance between cluster morphological centers are applied (minimum mean-tractogram similarity: 0.1 and minimum spatial distance between cluster centers: 2 cm) Clusters for which no suitable correspondence can be made are discarded and not considered in the comparison. There are other matching algorithms available, such as the Hungarian method (Kuhn, 1955), which tries to optimize the matching in terms of global rather than local distance between matched elements. However, this also means higher computation time and resources. For a first implementation and proof of the method, we chose the simpler greedy matching with reduced computation time.

The leaf matching process is outlined in Figure 3.6.
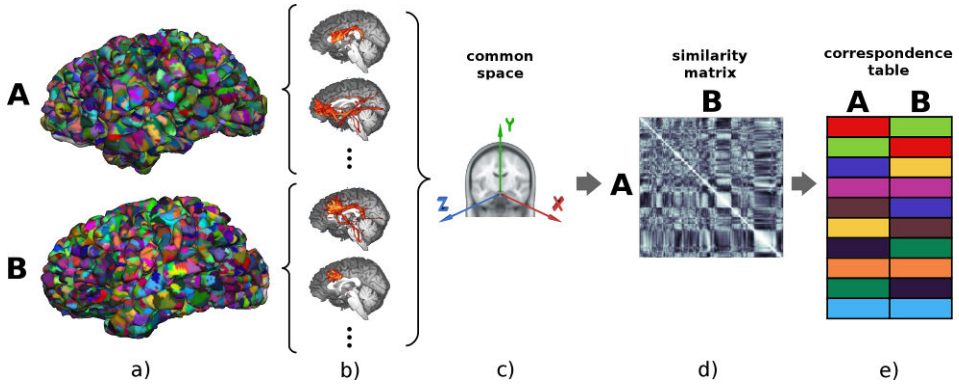
Figure 3.6: *Leaf-identification pipeline: maximum effective granularity partitions are obtained for each subject, i.e., Subjects A and B (a); a mean tractogram is computed for each cluster (b); all tracts are registered to a common space (c); pairwise tract similarity matrix is computed between the subjects (d); a greedy algorithm is used to extract the cluster correspondence table from the matrix (e). These clusters will become the new leaves of the trees.*

3.7.3   Matching quality

The smaller the distance between matched clusters, the most similar they are. These distances can be used to obtain an overall quality value for the matching of the trees. By obtaining the mean distance of the matched clusters, weighted by their size, we obtain a normalized average distance between seed voxels of both trees.

$$MatchQ = 1 - \sum_{i}^{M} \frac{d(T_{1i}, T_{2i})}{S_{1i} + S_{2i}} \tag{3.15}$$

were M is the number of matched pairs across the trees, T_{1i} is the mean tractogram of the matched meta-leaf i in tree 1, and S_{1i} is the size of that meta-leaf (the number of data points it represents). We will use this value as our matching quality coefficient for that certain tree pair. The higher the matching quality, the more reliable the tree comparison results will be.

3.7.4   Tree cophenetic correlation coefficient (*tCPCC*)

Once the leaves of two trees have been matched, it is possible to apply tree comparison algorithms. The first of these we implemented is the *tCPCC*. The principle and equation used for the tree quality assessment (*CPCC*; Farris, 1969, *Equation 3.14*) can be adapted for tree comparison. In this case, instead of evaluating the distance values encoded by the tree against the ones obtained by the original tractograms for each pair of seed voxels, the distance values encoded by each of the trees for each pair of corresponding meta-leaves were compared. However, as different meta-leaves may have different sizes (in the sense of containing a different number of seed voxels), the *CPCC* factor was modified in order to include a weighting with cluster size. This way the relevance of the distance value between two meta-leaves was proportional to the fraction of the total seed voxels contained in them. The mathematical formula for the *tCPCC* resulted as follows:

$$tCPCC = \frac{\displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij} S_{x_{ij}} y_{ij} S_{y_{ij}} - \mu_{X'}\mu_{Y'}}{\sigma_X \sigma_Y \displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} (S_{x_{ij}} S_{y_{ij}})} \qquad (3.16)$$

$$\text{with } \sigma_X = \sqrt{\frac{\displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} (x_{ij} S_{x_{ij}})^2 - \mu_{X'}^{2}}{\displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} S_{x_{ij}}^2}} \text{ and } \mu_{X'} = \frac{\displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_{ij} S_{x_{ij}}}{\displaystyle\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} S_{x_{ij}}}$$

where x_{ij} is the distance between meta-leaves i and j as encoded in tree X and Sx_{ij} is the sum of the sizes of meta-leaves i and j for tree X.

As with the *CPCC*, a value of 1 would indicate that the distance values between single meta-leaves encoded by both trees are linearly dependent (meaning that both trees contain the same information encoded in their distance values), and a value of 0 means that the trees do not share any common information.

3.7.5   Weighted triples similarity (*wTriples*)

An alternative tree comparison method, described in detail by Bansal and colleagues (2011), consists of comparing the joining order of all possible triples of leaves of each tree. The number of triples for which the joining order is exactly the same is divided by the total number of possible triples; obtaining a value ranging between 0 and 1. As with the *tCPCC*, a weighting was included to account for meta-leaf size, and the final formula was expressed as:

$$wTriples = \frac{\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} f_{top}(i,j,k)}{\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} S_{x_{ijk}} + S_{y_{ijk}}}$$

(3.17)

$$\text{with} \quad f_{top}(i,j,k) = \begin{cases} S_{x_{ijk}} + S_{y_{ijk}} & \text{if topologies match} \\ 0 & \text{if topologies do not match} \end{cases}$$

The *tCPCC* and *wTriples* comparison methods are (partially) complementary: while the former stresses the similarity of the distance values encoded by both trees, the latter focuses on the similarity of the hierarchical topologies of both trees, regardless of the numerical values encoded.

3.7.6   Restricting compared structure based on matching quality

When using the tree comparison algorithms mentioned above, we are assuming that the leaf matching is correct and optimal, and contrasting the tree structures across all granularities and levels. However, it might be argued that if the matching distance between two leaves is not optimal, this means that we can only identify these leaves within a granularity lower than the maximal, and therefore we should not use these leaves for comparing structures in the trees that are above such granularity.

In order to address this concern, and to test how the results could vary from it, a modification was built into the *tCPCC* as an extra method, that we will call matching quality restricted comparison, or simply restricted comparison.

In this modification, a binary function is embedded into the formula of *Equation 3.16*, at both numerator and denominator sums. This function follows the formula:

$$r(i,j) = \begin{cases} 1 & if \quad \min\left(d_{T_1}(i,j), d_{T_2}(i,j)\right) > \max\left(d_M(i), d_M(j)\right) \\ 0 & otherwise \end{cases} \quad (3.18)$$

where α is the scaling factor, $d_{T1}(i,j)$ is the distance between meta-leaves i and j as encoded in tree 1, and $d_M(i)$ is the matching distance of leaf i across trees. Note that all four conditions must be met, otherwise the function will be 0, and that particular leaf-pair will have no influence in the value of the restricted *tCPCC*.

As only a subset of all possible pairs will contribute to the index value, this could be understood as comparing trees with smaller number of leaves, and therefore with lower highest granularity. In order to characterize this we can define the effective granularity of a particular comparison as:

$$effGran = \frac{N^2}{N^2 - 2P} \quad (3.19)$$

where N is the number of leaves of each of the trees being compared and P is the number of pairs that contributed to the *tCPCC*.

3.8   Interpretation of hierarchical information

3.8.1   Introduction

As argued above, the tree is suitable for assessing the structural map of the cortical sheet as a whole. However, in order to fully appreciate the function-anatomical organization of the cortex, we also need to map this information back onto the cortical surface. Because the tree is a multidimensional structure, it cannot be fully projected directly onto this 2-dimensional space. Some strategies have been proposed that allow including some degree of multi-granularity information into surface mapping, such as using similar color hues for subclusters of a bigger cluster (for example using reddish, greenish and bluish hues for subclusters of 3 main divisions) or hierarchical "space-blobs" (Cachia et al., 2003). These approaches, however, are not suitable for the very high range of granularities and high number of nodes present in our trees. As an alternative, representative parcellations (being equivalent to a complete cut of the tree that severs all connections between the top node and any leaf) may be found that best approximate the information encoded in the tree. It is very unlikely that a single partition can represent the entire similarity structure of the data. Using a series of

partitions at different granularity levels, which in this case would also be hierarchically nested, might be a better way to achieve it.

3.8.2   Partition extraction algorithms

3.8.2.1   Dendrogram partitioning

Many different methods for comparing and assessing partitions can be found in the literature (i.e: Halkidi et al., 2002; Rand, 1971; Theodoridis and Koutroumbas, 1999). However, these methods usually refer to the original data, which in our case would involve operations with high dimensional tractograms, making them computationally expensive and slow. Limiting the data used to that contained in the tree allows fast partition assessment algorithms to be implemented. There is also literature available on tree partitioning algorithms (Jain and Dubes, 1988; Langfelder et al., 2008; Zahn, 1971), but these methods did not translate into meaningful partitions in the case of the brain connectivity trees studied here. The most traditional case for tree partitioning does, however, deserve introduction.

3.8.2.2   Minimum guaranteed intra-cluster similarity (Horizontal cut)

By definition, if a horizontal cut is made through a dendrogram the partition obtained is the one that guarantees, for a given number of clusters, a maximum lower bound for the intra-cluster similarity. Therefore, this cut yields regions with high consistency. In order to select a partition, either a number of desired clusters, or the distance level where the horizontal cut shall be made must be chosen.

3.8.2.3   Cluster spread vs. separation (*SS*) index

The horizontal cut method only takes into account the distance level of the clusters involved in the partition, that is, the encoded distance between the elements contained in those clusters (which relates to spread or scatter of the clusters). A more complete partition selection method should also factor in the distance between such clusters related to their spread. Furthermore the horizontal cut may only be used with a pre-defined granularity level and is unable to assess the quality of a partition. In order to tackle these shortcomings, we introduced a second algorithm presented below.

The overall spread of the clusters in a partition can be quantified through the formula:

$$Spread = \frac{1}{S_T} \sum_{i}^{N} S_i d_i \qquad (3.20)$$

where d_i and S_i are the distance level and size of cluster i, respectively. N is the number of clusters in the partition, and ST the sum of all clusters sizes in the partition.

On the other hand, the distance level of the parent of a given node in the tree encodes the separation between the center of that cluster and that of its closest neighbor. The average separation between neighboring clusters for a given partition can then be expressed as:

$$Separation = \frac{1}{N} \sum_{i}^{N} d_{p(i)} \qquad (3.21)$$

where $d_{p(i)}$ is the distance level of the parent of node i.

Using these two formulas, a partition quality measure is obtained by calculating the ratio between the mean spread of clusters in the partition and the mean separation between neighboring clusters: the *SS* index.

$$SS = \frac{S_T \sum_{i}^{N} d_{p(i)}}{N \sum_{i}^{N} d_i S_i} \qquad (3.22)$$

A higher value will indicate that, for that partition, the mean separation of clusters is high compared to the separation of elements within the clusters. This index can be coupled with a required number of clusters, in order to find the best possible partition of a given desired granularity. Alternatively, it can also be used to find global or local maxima in the *SS* values across the tree, thus revealing partitions of particular significance.

3.8.2.4   Minimum cluster size difference

Due to the nature of the tractogram similarity measure used, areas that share long common pathways (like, for example, the longitudinal fasciculus) will tend to be more similar to their surrounding areas sharing these large connections than to those with shorter pathways or more local connectivity fingerprints (such as the superior frontal lobe). Such highly cohesive areas tend to remain less partitioned by the spread-separation scheme than areas with local connectivity.

Depending on the purpose of the partitioning, it may be useful to circumvent this side-effect by obtaining partitions guided by the connectivity structure encoded in the tree but with an emphasis on clusters of similar sizes. This can be accomplished by finding partitions that minimize the mean square size difference for a given number of clusters. The objective function to be minimized is expressed as:

$$SizeDiff = \frac{2}{N(N-1)} \sum_{i}^{N-1} \sum_{j=i+1}^{N} (S_i - S_j)^2 \qquad (3.23)$$

3.8.2.5   Efficient hierarchical search of partitions

The data size and the extremely high number of possible partitions contained in the trees means that an exhaustive assessment of the SS or minimum cluster difference values for all possible cuts (even if only those with a certain number of clusters) would result in a very slow algorithm not usable in real time. In order to obtain partition selection methods fast enough to be integrated into an interactive tree exploration tool, a top-down hierarchical search algorithm was implemented here. This means that, starting at the partition defined by the first branching of the tree (or sub-tree), all possible subdivisions of each cluster going down up to four branching levels are considered, and the resulting partitions are evaluated. The best performing partition is identified, and the corresponding cluster from which division it had derived is subdivided down one level. The process is iterated until the desired number of clusters has been obtained or until the maximum granularity partition has been reached.

3.8.2.6   Stable boundaries

The methods described above focus on finding complete partitions in the tree, but information can also be extracted from the tree that does not involve a complete partition of the dataset. As each bifurcation in the tree represents the separation between two clusters (i.e., a boundary), a technique could aim at finding the most relevant or persistent boundaries rather than entire parcellations. An idea would be to look at the branch lengths of the nodes involved. The longer the branch, the wider the range of granularities at which the boundary for that region remain stable. This way, important boundaries would be mapped on the cortex, rather than entire parcellations.

Procedurally, the branch lengths of all nodes are checked (the distance level difference between a node and its parent node), and those above a given value entered as parameter are selected. Their corresponding nodes will not necessarily constitute a partition (and they might be nested) and their edges will delimit more stable boundaries than those not detected for that given parameter. Also, the branch length could be normalized by the node height from which it stems (and multiplied by a tuning coefficient), as at higher granularities the node branches are usually much shorter, so the same absolute value might have different meaning at different granularity ranges. However, this principle requires more study, and the algorithm and results presented in this study are only preliminary.

3.8.3   Visualization of clustering results

In order to interpret clustering results via sets of partitions (or selected clusters), these must be visualized in the hierarchical tree and in the brain surface. A simple dendrogram drawing algorithm allowing an input of cluster identifiers and desired colors was implemented for partition projection onto the tree. To project our results onto the brain, we desired to use the surface generated from the T_1 images by the *Freesurfer* package (`surfer.nmr.mgh.harvard.edu`), however, our clustering results refer to seed voxel coordinates from the diffusion image.

For the datasets where a diffusion field map had been acquired and the diffusion image unwarped, we used a nearest neighbor interpolation to project our results onto the *original Freesurfer* image (corresponding to the gray matter/white matter interface, same as our seed voxels). Each vertex in the *original* surface is assigned to the cluster the nearest data point belongs to. *Freesurfer* generates several surfaces from each brain, *original* (grew matter/white matter surface used to generate the rest), *pial* (cortical surface), *inflated* (allowing clearer view of the sulci and insula), *sphere*, etc. All these surfaces have 1-to-1 correspondence between their vertices, so once the results are projected onto the *original* surface, they can also be projected into any of the other ones, most commonly *pial* surface (Figure 3.7).

As all the clustering analysis are done in the diffusion image, in those cases where diffusion field maps are unavailable (and therefore unwarping cannot not be performed) the results would have some distortion with respect to the T_1. In order to minimize projection errors, an extra processing step is included. After generating the *Freesurfer* surfaces, the T1 image histogram is normalized to the FA image, and non-linearly registered to it using *ANTS*. The transformation matrix obtained is then used to unwarp the surface with unwarping tool from the *CBS*

toolset within *JIST* (`www.cbs.mpg.de/institute/software/cbs-hrt`; Landman et al., 2013). The nearest neighbor interpolation is then used with this surface, and the results are finally shown in the unchanged ones (as there is still 1-to-1 correspondence). Still it must be noted that in some cases small projection errors might still occur, these can be seen as irregular boundaries, or as small "spots" near the boundaries with the color of an adjacent cluster (see for example the boundaries between purple and cyan clusters and purple and green clusters in the rostral part of Figure 3.7d).
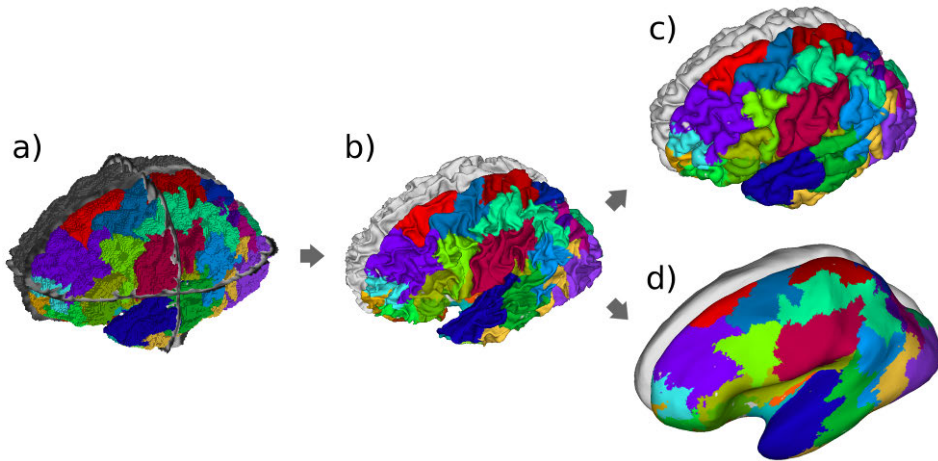
Figure 3.7:   *Visual example of the intermediate steps of the partitional results before being presented in the final surfaces: cluster results in dMRI seed voxels (a); nearest-neighbor projection onto Freesurfer original surface (b); visualization on pial (c) or inflated (d) surfaces.*

3.8.4   Interactive hierarchical exploration module

The visualization algorithms described above, along with the partition selection methods of section 3.8.2 and the dendrogram processing tools of section 3.6 were included as part of a fast interactive exploration and visualization tool for hierarchical characterization of brain connectivity. With a tree file and the brain surfaces as input our tool allows for real-time exploration of the tree and partition selection, while visualizing the results simultaneously on the 3D brain surface and in the dendrogram plot (Figure 3.8). This tool was implemented as modules of the open-source *OpenWalnut* framework (`www.openwalnut.org`) and can be freely downloaded from the website.

An implementation of the same tools is also planned for the visualization tool *BrainGL* (`code.google.com/p/braingl`).
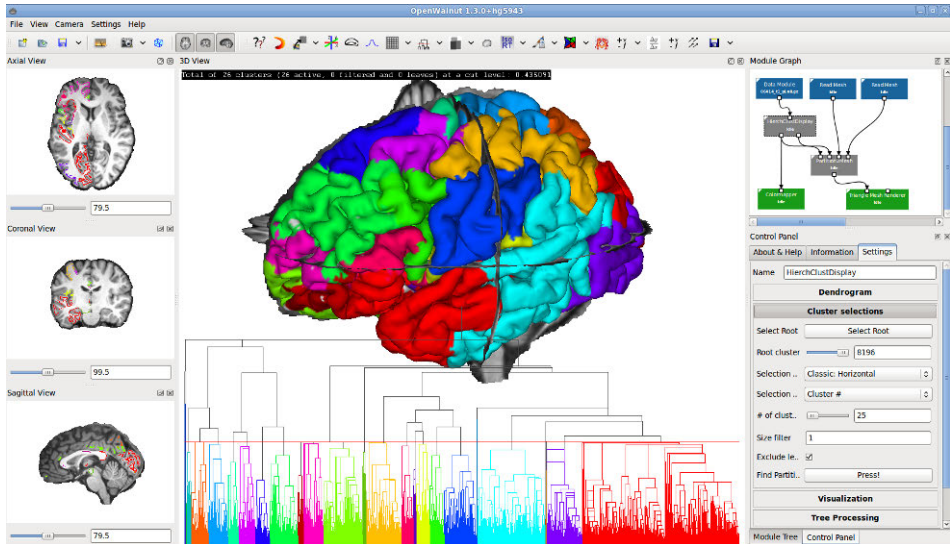
Figure 3.8:   *Snapshot of the interactive hierarchical exploration tool implemented in OpenWalnut.*

The tree building method (including also tree processing), *CPCC* algorithm, and tree comparison methods were implemented using C++ as standalone programs to be run through bash command line in a linux system.

3.8.5   Partition color matching across datasets

As our trees potentially contain very high granularity partitions, it was not viable to implement a morphological or atlas based coloring scheme. Instead, in the *OpenWalnut* module each node of the tree is assigned a specific color. Starting at the top of the tree (lowest granularity) a primary color is assigned to the first node, at every branching, the node color will be inherited by its biggest children node (larger number of contained data points), and the other nodes will be given new colors trying to maximize color distinctiveness. This process is iterated until every node in the tree has been assigned a color.

Unfortunately, this means that when comparing partitions across datasets, cluster colors of similar areas would not necessarily match, which would makes

visual comparison less intuitive. However, if the trees have been matched using the method described in section 3.7.2, this information can be used to bring a certain degree of color matching to the partitions.

   The following algorithm was implemented and used when comparing partitions across subjects and datasets. Separately for each tree, each cluster of the desired partition to be color-matched is identified with the cluster of the partition from the other tree with which the overlap of leaves is maximal. In cases when identification between two clusters agrees bi-directionally, the cluster of the second partition is colored using the first one as template. In cases where clear multiple-to-one identification is obtained, the clusters are assigned different colors but of a similar hue to the color of the one cluster matched. In the remaining cases the original colors are kept.

   When contrasting partitions from different datasets, the information obtained when matching leafs on different trees can be exploited to give the same color to overlapping clusters across matched trees and ease visual interpretation.

4. A proof-of-principle study on multi-granularity dMRI-based whole-brain characterization

4.1 Overview

In this section we will test the developed method on a small group of healthy young subjects. The datasets consist of four young healthy participants and a repeated measurement of the fourth participant (denominated A, B, C, D1 and D2 respectively, see section 3.2 for details on acquisition and preprocessing). The results obtained with each of the methods proposed in the last section will be presented and contrasted. Part of the results presented in this section have been published as an article in the scientific journal *Human Brain Mapping* (Moreno-Dominguez et al., 2014a) and in posters and talks at international conferences (Anwander et al., 2012; Knösche et al., 2012; Moreno-Dominguez et al., 2013, 2011a).

4.2 Selecting a linkage method

Pairwise tractogram distance matrices were obtained for both hemispheres of subjects A, B and C (i.e., 6 datasets). Hierarchical trees were built over these matrices using each of the graph methods proposed in section 3.4.2. Trees were also built directly from the tractograms using the centroid method presented in section 3.4.3. for each of the different neighborhood levels implemented, with and without the initial homogeneous merging stage. For the centroid trees, several values for the number of target clusters where to stop the homogeneous merging phase (top of the sub-trees being grown) were tested on one of the datasets. 5000 was selected as the best value that incurring in minimal information loss while facilitating many steps of the tree processing. As will be shown below, the

information loss was also minimal for the rest of the datasets, and the parameter was accepted as valid.

In order to test the outcome of the tree building algorithms over unstructured data, a set of artificial tractograms (equal in number to those obtained from the real datasets) was generated in a way that they would yield a distance matrix of random values uniformly distributed between 0 and 1 (that is, a dataset absent any hierarchical structure). This was achieved by creating tractograms representing points uniformly distributed over the surface of a sphere in n-dimensional space. However, in order to ensure this uniformity in a reasonable generation time, the dimension of the random tractograms was limited to $n = 10$. When testing the centroid method (which requires physical neighborhood information), each of the three random tractogram sets was assigned coordinates from a different real dataset.

The computation time for obtaining one full distance matrix needed for the graph methods amounted to two full weeks on a dual-core computer. After the matrix is obtained, the tree building processes are quick (within 15 minutes) but require a large amount of available RAM memory (in excess of 20 Gb) for loading the full matrix. With the centroid method these requirements are reduced to an average of 48 hours of computing using 4 Gb of RAM. We would like to note that in our study we have emphasized working with high resolution images in order to obtain the maximum quality and granularity possible while maintaining a reasonable computing time. However, if faster processing would have priority, working with lower resolution images would drastically cut down the required computation times and memory requirements, taking days to obtain a full matrix instead of two weeks, and only a few hours for a centroid tree.

It is not possible to detect significant differences in the overall topology of the trees obtained with the different methods by mere visual inspection, except perhaps that the distance values for the single and complete linkage methods tend to be much lower and much higher, respectively, than the ones for the other methods (Figure 4.1). Numerical analysis is, therefore, necessary to assess their fit to the data. For this purpose, *CPCC* values were computed for all obtained dendrograms. In order to set a baseline level for the *CPCC* values, trees were also built from unstructured datasets (using artificially generated tractograms that yield random uniformly distributed distance matrices, as explained in the previous subsection), and their *CPCC* values computed. The results are shown in Figure 4.2.

Single

Complete

Average

Weighted
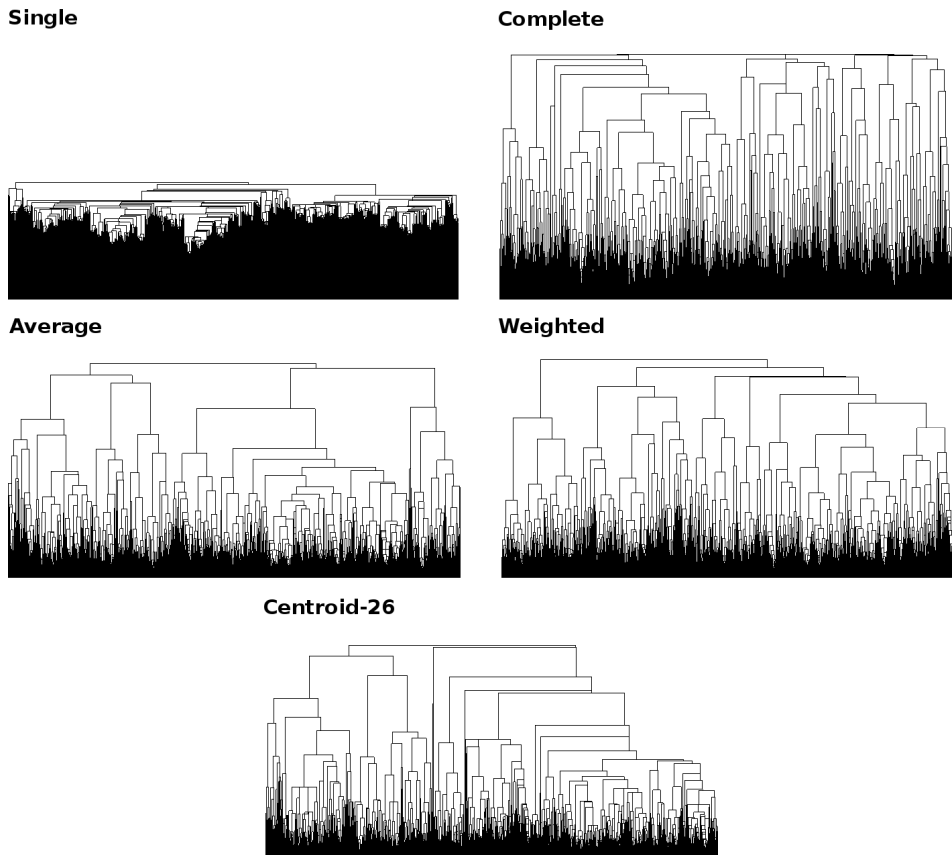
Centroid-26

Figure 4.1:   Trees obtained from the left hemisphere data of subject A for each of the graph methods plus the centroid method with a 26 neighborhood. Note that a particular position in the x-axis does not identify a particular seed voxel, this may change in order to allow for the representation of the structure in tree form without any line crossings.
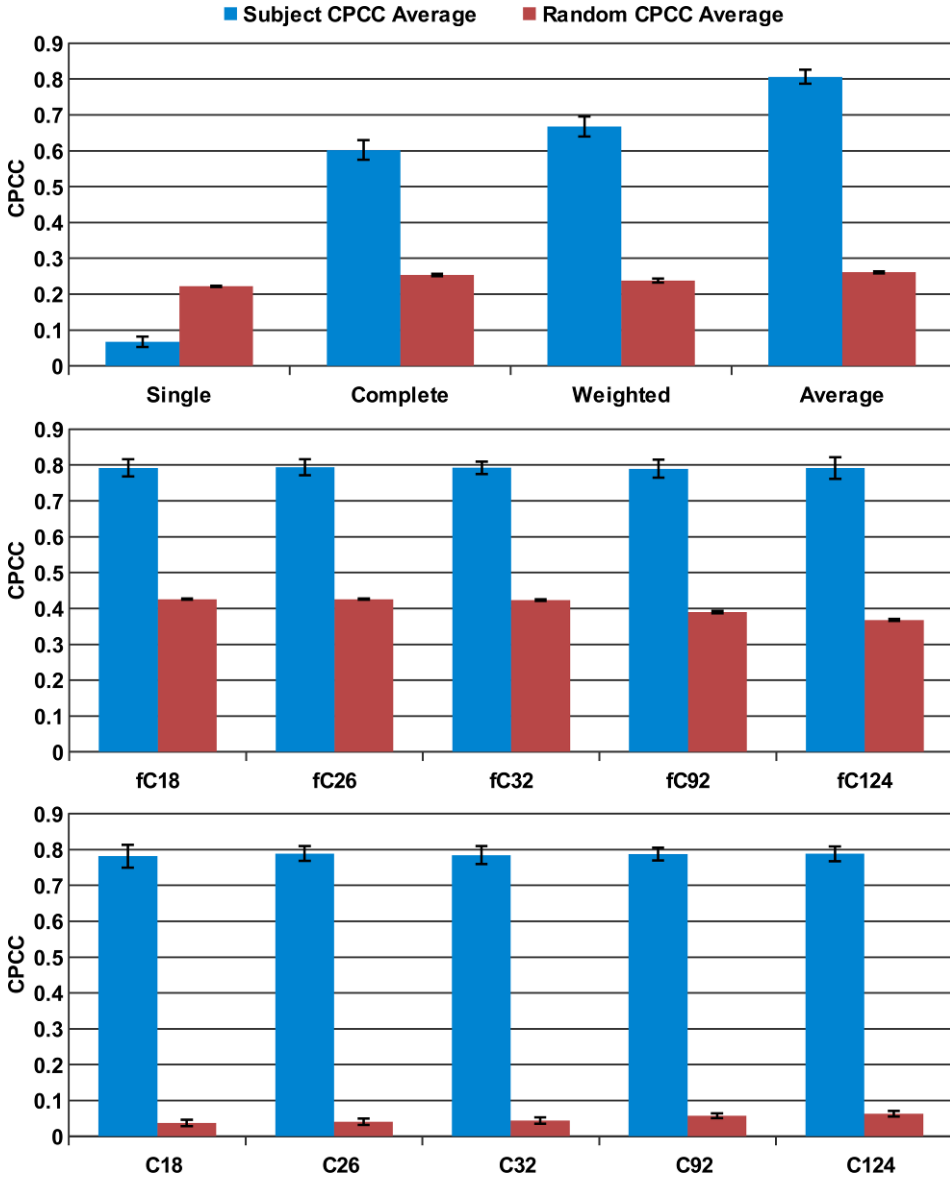
Figure 4.2: *Average CPCC values for trees obtained from each hemisphere of subjects A, B, and C, and from the three random tractogram sets using the graph methods (top), restriction-free centroid method (middle), and centroid method with initial homogeneous merging phase (bottom). The cyphers in the centroid method labels indicate the different degrees of neighborhood (see section 3.4.3.2 for more details). The error bars indicate the standard deviations.*

The results show that, for the real datasets, the single linkage method performs worst, the complete and weighted linkage methods are not a very good match to the data either. The average and centroid methods provide the best fit to the original data, obtaining high and very similar *CPCC* scores, with no statistically significance difference between them. Moreover, there was no significant improvement in quality using wider neighborhoods in the centroid method. Also, there was no significant change observed in the *CPCC* value between restriction-free centroid methods and those with initial homogeneous merging stage, indicating this process (carried out with the selected parameter) did not deteriorate the quality of the obtained trees (overall average *CPCC* difference was of 0.75% with a standard deviation of 0.65%). In all cases the values obtained were well above their baseline levels, especially in the case of the full centroid method. Given that there was no significant information loss by applying the initial homogeneous merging stage to the centroid method, from this point on we discard the restriction-free algorithm, in favor of the one including all the features.

The computational load incurred for obtaining each tree was empirically derived as the number of tractogram similarities computed, and the results are plotted in Figure 4.3.
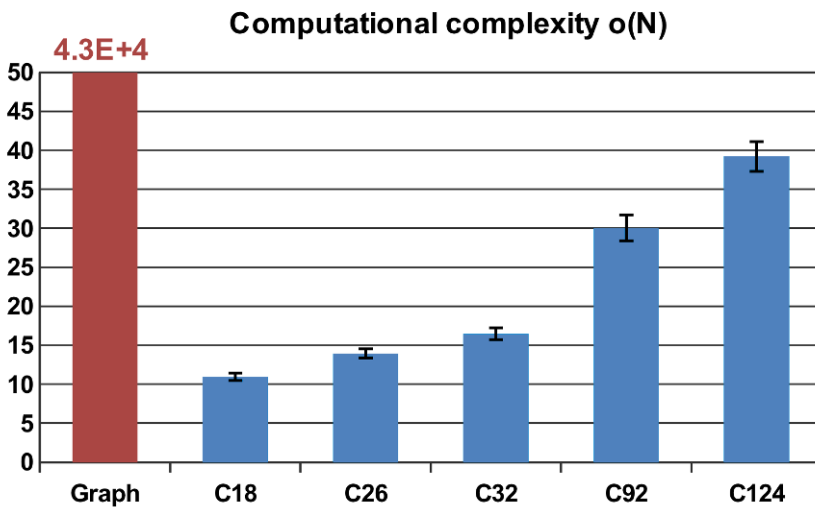
Figure 4.3:   *Average computational complexity (expressed as the number of tractogram similarity operations performed normalized by the size of the dataset N) of the tree building methods applied to the real datasets (graph linkage in red, centroid method with different neighborhood levels, 18 to 124, in blue). For interpretability, the bar for the graph linkage methods is truncated and the numerical value is indicated. Error bars show the standard deviation.*

As can be seen, an average of $4.3 \cdot 10^4 \times N$ tractogram similarity operations were necessary to build up the graph linkage trees (value out of axis range), with N being the size of the dataset from which the trees were computed (the complexity of the graph methods is $N(N\text{-}1)/2$ and the datasets used are in the range of $6.5 \cdot 10^4$ to $10 \cdot 10^4$ points). On the other hand, centroid methods required only $15N$ to $50N$ operations, three orders of magnitude less than the graph methods.

It is clear from these results that from the methods considered, average and centroid linkages are the best fit to the data, with the latter having the further advantage of incurring far less computational load. Within the centroid methods, the computational load increased almost linearly with the number of neighboring voxels considered.

The 26 neighborhood centroid method with initial homogeneous merging stage ($c26$) was chosen as the optimal trade-off between the quality of the tree and the computational cost, and was the only method used for the remainder of the study (while the $c18$ method showed slightly lower complexity, it also showed slightly lower $CPCC$ mean scores and higher $CPCC$ standard deviation, so $c26$ was deemed the most conservative choice among the two).

4.3 Cleaning the dendrograms

The tree preprocessing steps described in section 3.6 were applied to the $c26$ dendrograms of each hemisphere from subjects A, B, and C.

Firstly, those data points with distances greater than 0.1 to their nearest neighbor were considered as outliers and excluded, resulting in a rejection of an average of 0.5% of the data points (as indicated in the methods section, this step was actually taken during the tree building process, but for clarity of structure we present it in this section) .

Following, non-monotonicity was corrected and the maximum granularity was limited by merging all inner nodes of the 5000 homogeneous sub-trees obtained during the first phase of tree construction, effectively transforming these nodes into non-binary meta-leaves; non-binary structures at all levels of the tree were detected and flattened using a parameter of $l = 0.05$ (nodes with branches shorter than 5% of the node height were eliminated). These parameters were selected empirically in order to obtain some added complexity reduction at higher levels of the tree (measuring complexity as the number of branchings or inner nodes in the tree) while keeping the total information loss in the same order range (<1%).

As with the number of homogeneous sub-trees, these parameter values were optimized for one of the datasets by testing multiple values and selecting those who performed best, achieving further complexity reduction without significantly adding any information loss. As the performance of these parameter values was similar for the rest of the datasets, the values were accepted. Although a wider and more strict testing could be performed to better optimize parameter values, and would make for an interesting future study, significant changes would not expected and a considerable computational effort and time resources would be necessary, which were deemed more worthwhile spending in the following sections of this study.

In order to quantitatively assess the complexity reduction and the information loss caused by the pre-processing, inner node count and *CPCC* values were obtained for the trees at each processing step, and their relative changes in relation to the previous states were evaluated (Figure 4.4). The results show that neither of the first two steps (monotonicity correction and limiting of maximum granularity) significantly reduced the amount of information contained in the trees, while the second step achieved a complexity reduction of almost 90%. The third step (flattening of non-binary structures) further reduced the complexity by 5%, while introducing an average of 0.2 % of information loss (without statistical significance). Overall, the whole pre-processing pipeline achieved a complexity reduction of more than 90% with a loss of information of less than 0.5% (0.15% on average), making it a remarkably efficient and useful tool for improving the performance of partition finding and tree comparison algorithms. It can also ease interpretation of the trees through visual inspection, although this still remains a challenging task. Visual changes on tree structure caused by the pre-processing are exemplified in Figure 4.5.
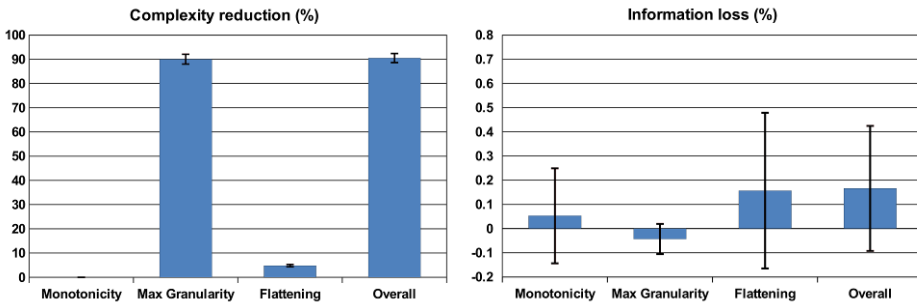
Figure 4.4:   *Average Tree Complexity Reduction (left) and Information Loss (right) of each step in the pre-processing pipeline, relative to the previous step. The last column of each chart represents the overall added effect of the complete pipeline. Complexity reduction is measured as being the relative number of inner nodes eliminated in each step. Information loss is measured as being the relative decrease in CPCC index value. Error bars show the standard deviation.*
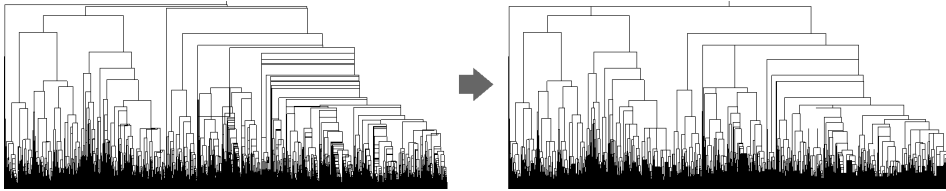
Figure 4.5:   *Tree corresponding to the connectivity structure of the left hemisphere of subject A before (left) and after (right) preprocessing.*

4.4   Comparing whole connectivity structure across datasets

The connectivity structure information encoded in the cleaned trees can be used as a whole, to compare changes in this structure across datasets. Following the steps described in section 3.7.2, mean tractograms were obtained for each meta-leaf of the processed trees and non-linearly transformed to a common space, guided by FA registration. In this instance the data of subjects A and B was registered into the space of subject C. For the within-subject comparisons across hemispheres, the tractograms of the right hemisphere were flipped and transformed into the left hemisphere; also guided by a previous FA registration, as in the inter-subject case. Next, the tractogram-distance matrices were obtained and the greedy leaf-matching algorithm was applied, restricted to a maximum

Euclidean distance between matched cluster centers of 2 cm. Using the resulting leaf-matching tables, the *tCPCC* and the *wTriples* similarity values were obtained. As a relative indication of the matching quality, the mean tractogram distances between matched clusters as well as the mean Euclidean distances between the matched clusters' geometrical centers were computed.

In order to test the reliability of the method and its robustness against noise, the whole process (starting at tractogram computation and tree building) was repeated with a noisier version of the same dataset, using only one, instead of three repetitions of the dMRI acquisition. Test-retest performance was also assessed using two datasets obtained from a fourth subject within a short period of time (1 week), referred to as D1 and D2.

In order to establish a baseline level for the matching values, a random matching scheme was set up, in which each meta-leaf of the first tree was matched at random to a meta-leaf of the second tree whose cluster center was not further than 2cm away. Afterwards, *tCPCC* and *wTriples* values were obtained. This process was repeated 100 times for each possible subject combination and the average value obtained. Distinct baseline values from both *tCPCC* and *wTriples* were computed for inter-subject comparisons, left *vs.* right hemisphere comparisons and high *vs.* low SNR comparisons.
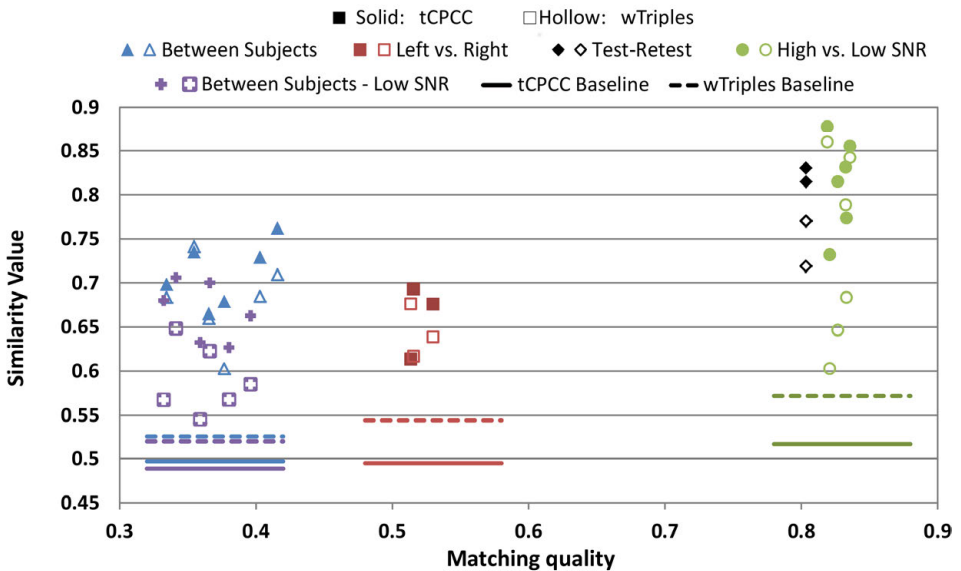
Figure 4.6:   *Tree similarity values plotted against matching quality for tree comparisons. Baseline levels for the corresponding matchings are shown below their datapoints in the same color, solid lines correspond to tCPCC baselines, and dotted lines to wTriples ones.*

The results are shown in Figure 4.6 where *tCPCC* and *wTriples* are plotted against the leaf matching quality between the compared data sets. Several observations can be made:

- All tree comparison values obtained are well above the baseline levels, rejecting the hypothesis that the leaf-matching obtained might have completely failed, making the comparison values meaningless. This confirms that there are indeed non-random structural similarities between the trees that can be detected.

- For *tCPCC*, the information loss by lower SNR and the variability between separate measurements of the same subject are smaller than the differences between different hemispheres or subjects (i.e. the values are higher and their spread smaller for the full green circles than the red squares or blue triangles). This indicates that differences in leaf similarity, as encoded in the trees, are not generally obscured by noise and can be interpreted. In contrast, *wTriples*, which only measures tree topology (joining orders), seems to be much more susceptible to noise, as the values are always lower and the spread, especially in the intra-subject high vs. low SNR measurements is increased.

- The similarities between the same hemispheres in different subjects (high SNR) and those between different hemispheres in the same subjects are within the same order of magnitude (between-hemispheres slightly lower, but not significant).

- Same-subject comparisons features much better leaf-matching quality compared to between-hemispheres comparisons, which in turn match better than between-subject comparisons.

- When comparing across subjects using lower SNR datasets, comparison values slightly decrease. Spread and matching quality however, remain almost unchanged.
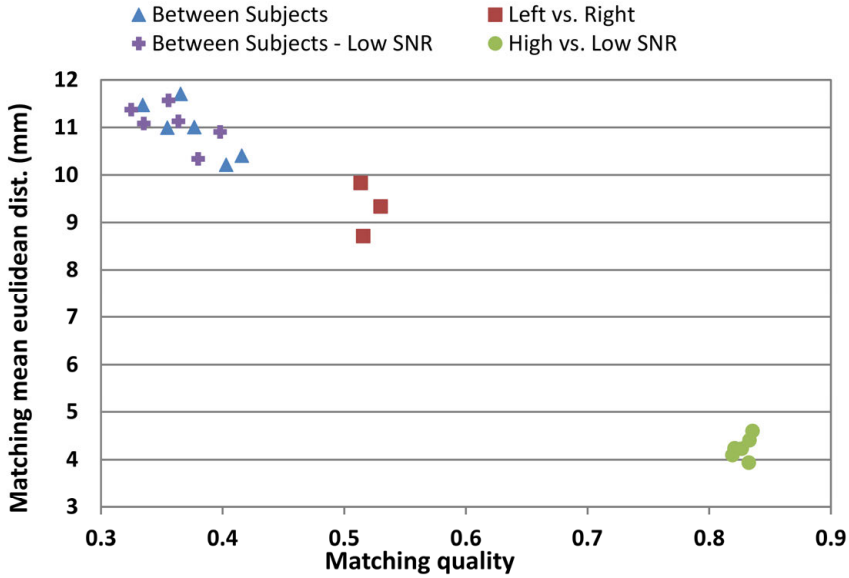
Figure 4.7:   *Relationship between matching quality and Euclidean distance between geometric cluster centers.*

In Figure 4.7 the matching quality is shown against the mean Euclidean distance (in the morphology) between matched cluster centers, and an inverse linear relationship can be observed between the two. This is to be expected as a higher morphological variability (such as across subjects or hemispheres) will cause distance between cluster centers to be greater and at the same time affect the maximum similarity that the tracts of matching cluster can have (which is what we have called matching quality). It is interesting to notice though, that a lower matching quality does not necessarily translate into a lower tree similarity index, as can be observed in Figure 4.6, where inter-subject comparisons (blue triangles) have lower quality matching as inter-hemispheric comparisons (red squares) while staying in the same index value range.

However, when we focus on the values for specific inter-subject comparisons (Figure 4.8, corresponding to the blue triangles and purple crosses in the previous charts), we see that relative relationship between different pairs of datasets are not kept across different levels of SNR. This means that although promising, the comparison method is in its current state not stable enough to interpret differences of absolute tree similarity values across subject pairs.
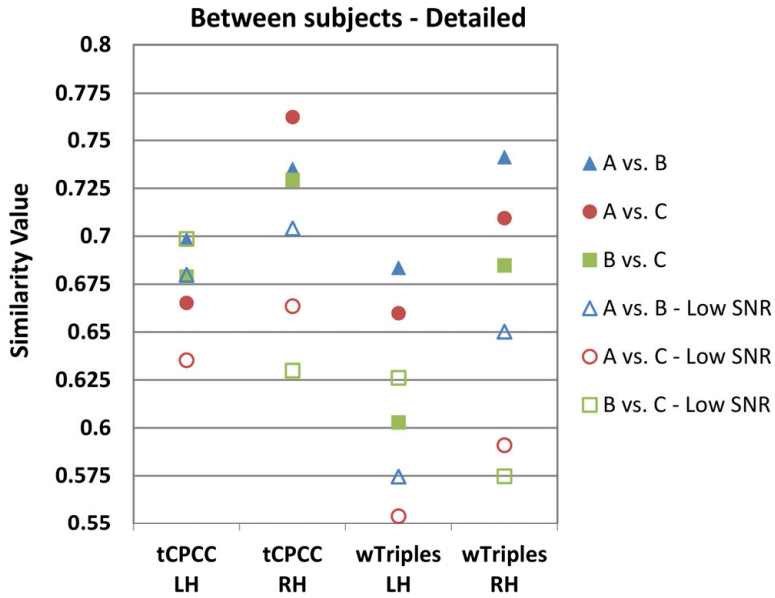
Figure 4.8:   *Detail of tree similarity values between different subject pairs for high and low SNR levels.*

Restrictive tree comparison (as presented in section 3.7.6) was also computed between the same datasets, and results are shown in Figure 4.9. Comparisons with high values of overall matching quality remain unchanged, across hemisphere comparisons are slightly reduced, while across subject comparisons remain in the same range for high SNR but decrease significantly when low SNR datasets are used. This means that this method is more susceptible to noise, although only in cases where matching quality is reduced.

If we have a look at the percentage of used pairs (Figure 4.10), it increases linearly with the matching quality. But as the effective granularity decreases rapidly with the amount of pairs used, the effective granularity of across subject comparisons is extremely low. It appears then, that with the current matching quality obtained, this restrictive method is unfortunately not applicable.
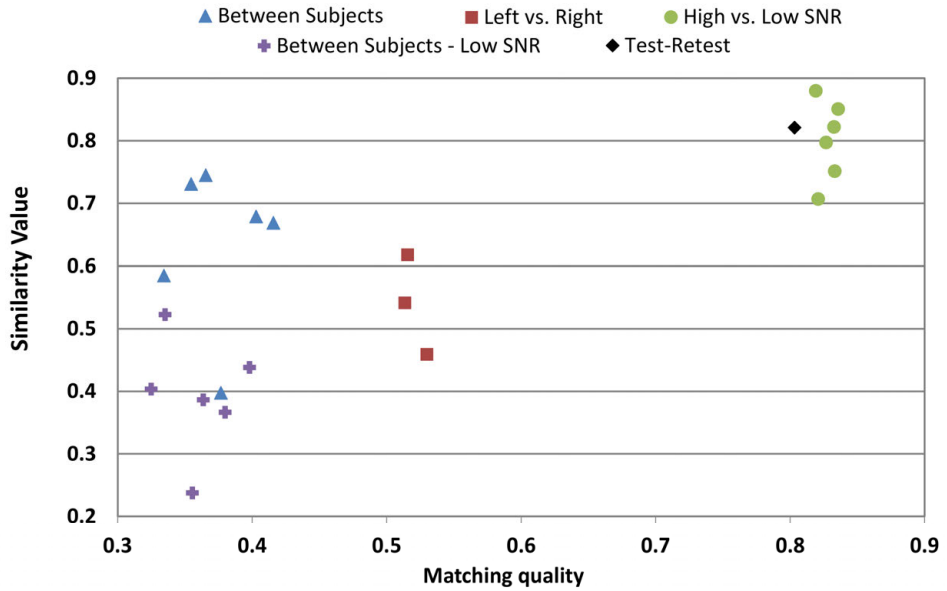
Figure 4.9:   *Tree tCPCC similarity values plotted against matching-quality-restricted tree comparisons.*
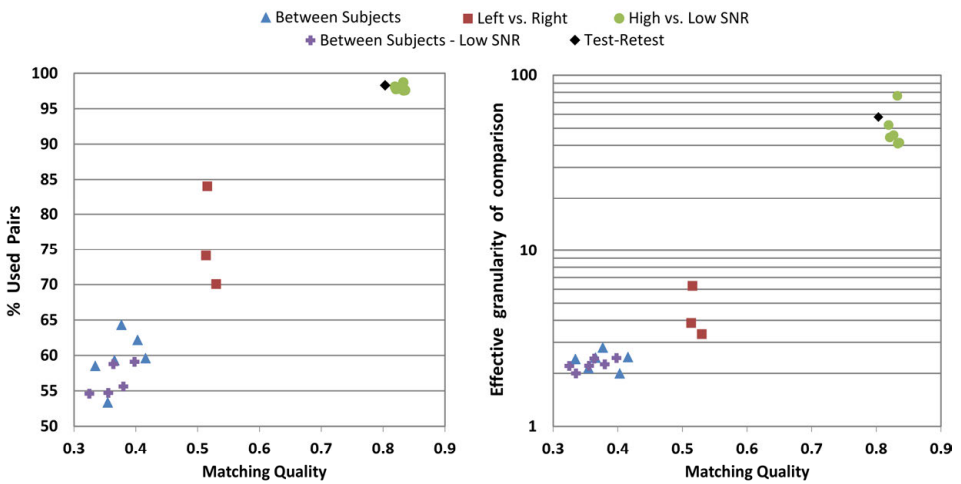
Figure 4.10:   *Percentage of used pairs and effective granularity of the comparisons performed with the restrictive method, plotted against the matching quality.*

4.5 Exploring the hierarchy: single subject partitioning

As we have shown, the hierarchical tree comprises a fairly complete account of the similarity structure of the cortex's connectivity as measured by dMRI, even after simplification. It constitutes a compressed representation of a very large number of possible parcellations at all possible levels of granularity. Each parcellation is equivalent to some complete cut of the tree that severs all connections between the top node and any leaf. Therefore, as demonstrated above, the tree is suitable for assessing the structural map of the cortical sheet as a whole. However, in order to fully appreciate the function-anatomical organization of the cortex, we also need to map this information back onto the cortical surface. Because the tree is a multidimensional structure, it cannot be projected directly onto this 2-dimensional space. As an alternative, a set of representative parcellations may be found that best approximates the information encoded in the tree.

Using the simplest partition selection method, the horizontal cut, we were able to obtain nested parcellations at different granularity levels by selecting a particular number of clusters for each level. This is exemplified in Figure 4.11 where we show the left hemisphere of subject A cut at 4 different granularity levels, exploring a wide range of hierarchical boundaries.

At very low granularity (15 clusters) the parcellation seemed to reflect the rough course of major fiber bundles (e.g., red for the fronto-occipital fascicle, green for the arcuate fascicle, purple for the cingulum bundle, and cyan for the cortico-spinal tract). Increasing the granularity to 50 clusters caused further subdivisions, most especially in the dorsolateral and dorsomedial frontal and parietal cortices, and also in the inferior frontal cortex and around the auditory cortex, reaching area sizes similar to Brodmann areas. Meanwhile, the cortex near the fronto-occipital fascicle, the superior part of the arcuate fascicle, and the cingulum bundle remained largely undivided. To obtain more fine-grained subdivisions in these regions, the threshold of the clustering criterion had to be lowered further, allowing for 100 clusters. Further increase of granularity continued changing details, for example by further subdividing the inferior frontal gyrus.
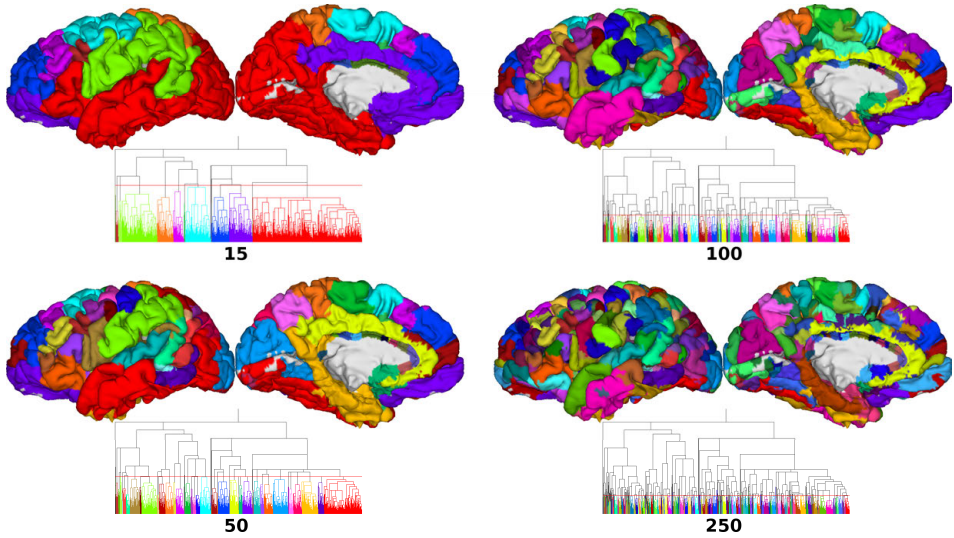
Figure 4.11:   Parcellation extracted from the hierarchical tree of left hemisphere of subject A using the horizontal cut algorithm. The numbers indicate the predefined number of clusters. The red horizontal lines in the trees denote the cutting level. The spread-separation method yields almost identical results. See text for further explanation.

Ultimately, the spread-separation method yielded highly similar results. Hence, for any pre-selected number of clusters, the horizontal cut seems to be a good approximation to spread-separation partitioning.

In Figure 4.12, we focused on the subdivision of the left inferior frontal gyrus (IFG) using the *SS* method. At relatively low granularity (50 clusters), only some of the major boundaries between the opercular and triangular parts (subject A, B) and between the triangular and orbital parts of the IFG were revealed. At higher granularity, more subdivisions appeared, including those that are not covered by the classical tripartition (into opercular, triangular, and orbital parts). For the repetitive acquisitions in the same subject (D1 and D2), the subdivision was highly reproducible.
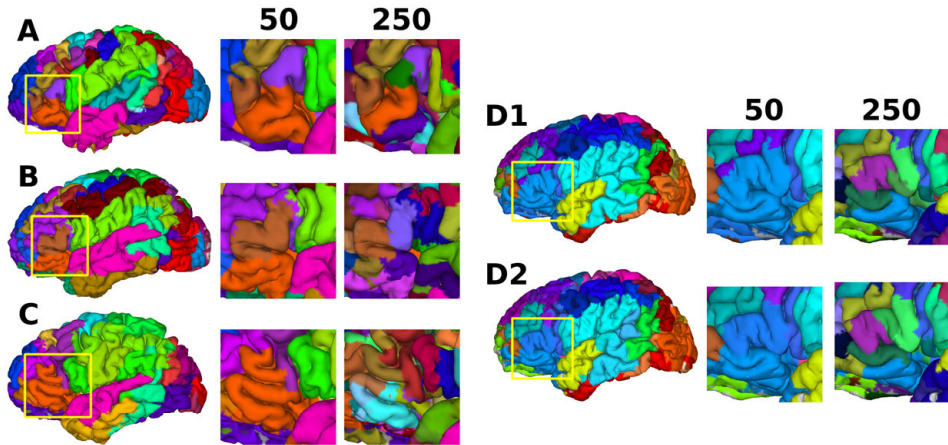
Figure 4.12:   Subdivision of the inferior frontal gyrus at two different levels of granularity, for the left hemispheres of subjects A, B, and C (left), as well as the two acquisitions of subject D (right). See text for further explanation.

It appears that, if cutting the tree was done according to the internal coherence of the clusters (i.e., the value on the vertical axis in the tree) or the maximum ratio between spread and separation (*SS* index), uninteresting "background" connectivity by large fiber tracts caused, at any given level of granularity, some regions of the brain to remain largely undivided, while others were split into small sub-areas. This supported the use of the minimized cluster size difference method (see *Methods* section), which gives more weight to area size and strives to obtain a more homogeneous parcellation. In Figure 4.13, the result was depicted for the same subject featured in Figure 4.11. When comparing the results of the two partitioning methods, some clear differences are apparent. At low granularity (15 clusters), the large temporal-occipital-frontal cluster (in red, see Figure 4.11) broke up into smaller areas, especially on the medial brain surface, while in frontal and prefrontal cortex fewer clusters were formed. This trend is also evident at higher granularities. For example, at 250 clusters the occipital lobe was more subdivided and the frontal one was less subdivided than with the horizontal cut method.
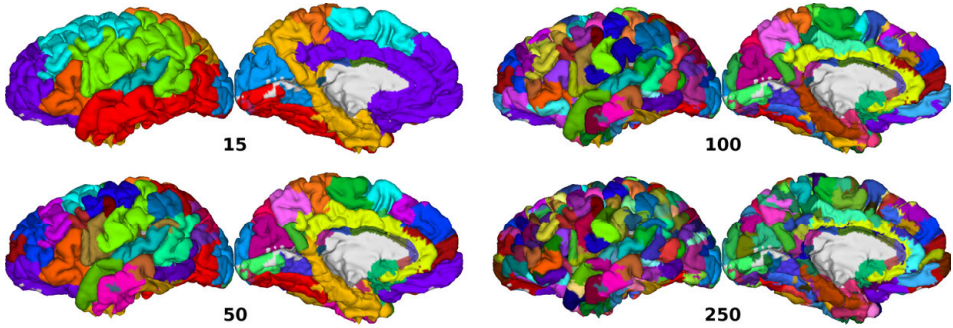
Figure 4.13:   Equal size partitioning for the same subject as depicted in Figure 4.11.

Thus far, we have explored the partitioning methods that required the input of a global granularity level (here expressed as number of clusters, but it could also be the average size of the clusters, or similar). However, the question remains: Which granularity levels might be the most representative ones for the tree? In order to reduce this arbitrariness, one can use the *SS* index (see *Methods* section) to select partitions. Using the *SS* partitioning (or the horizontal cut method, as it is a good approximation to *SS* for a given granularity), a series of parcellations can be obtained with maximum *SS* indices for each granularity level. In Figure 4.14 the *SS* indices were plotted as function of granularitiy for all data sets.

It can be seen that for small numbers of clusters the index rises steeply, meaning that in this range further subdivision usually leads to much better parcellations. In many data sets, this is followed by a shoulder (at about 50–200 clusters), where further subdivision does not greatly improve, or even slightly reduces, the quality of the parcellation (as measured by the *SS* index). Next, there follows a moderate increase, where subdivisions tend to (slightly to moderately) improve the *SS* index, until a maximum value is reached at about 200–600 clusters. From there, the curve steadily decreases, meaning that further subdivisions always lead to worse partitions. Consequently, the relevant range of partitions seems to start at the edge of the first shoulder (where mergings of clusters cause a rapid decrease and subdivisions of clusters cause no or only a small increase of the *SS* index) and end at the maximum (where both mergings and subdivisions cause a moderate decrease of the *SS* index). Ultimately, the interesting range of partitions based on the diffusion data seems to be roughly 20–600 clusters.
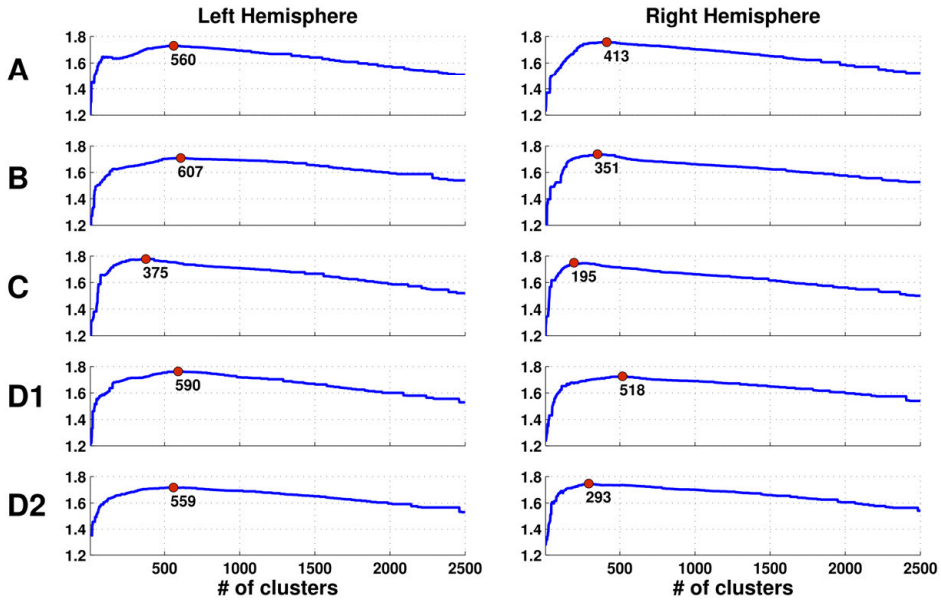
Figure 4.14:   SS indices obtained by the hierarchy search method, plotted against number of clusters. The red circle denotes the maximum of the curve.

Figure 4.15 shows the maximum *SS* index partitions for all subjects and hemispheres. These partitions have the maximum distinctness for the respective data sets, that is, the best ratio between intra-cluster inhomogeneity and between-cluster separation. In the event, these parcellations feature small parcels with an extension roughly comparable to the width of a major gyrus. It is evident that, at this level of granularity, the partitions of the two data sets from subject D are quite similar, while the partitions belonging to different hemispheres and/or subjects appear very different.

A different way to extract relevant information from the tree can be to detect those boundaries that are stable across a wider range of granularities than the rest, as proposed in section 3.8.2.6. Some preliminary results using absolute branch lengths as threshold values are shown in Figure 4.16.
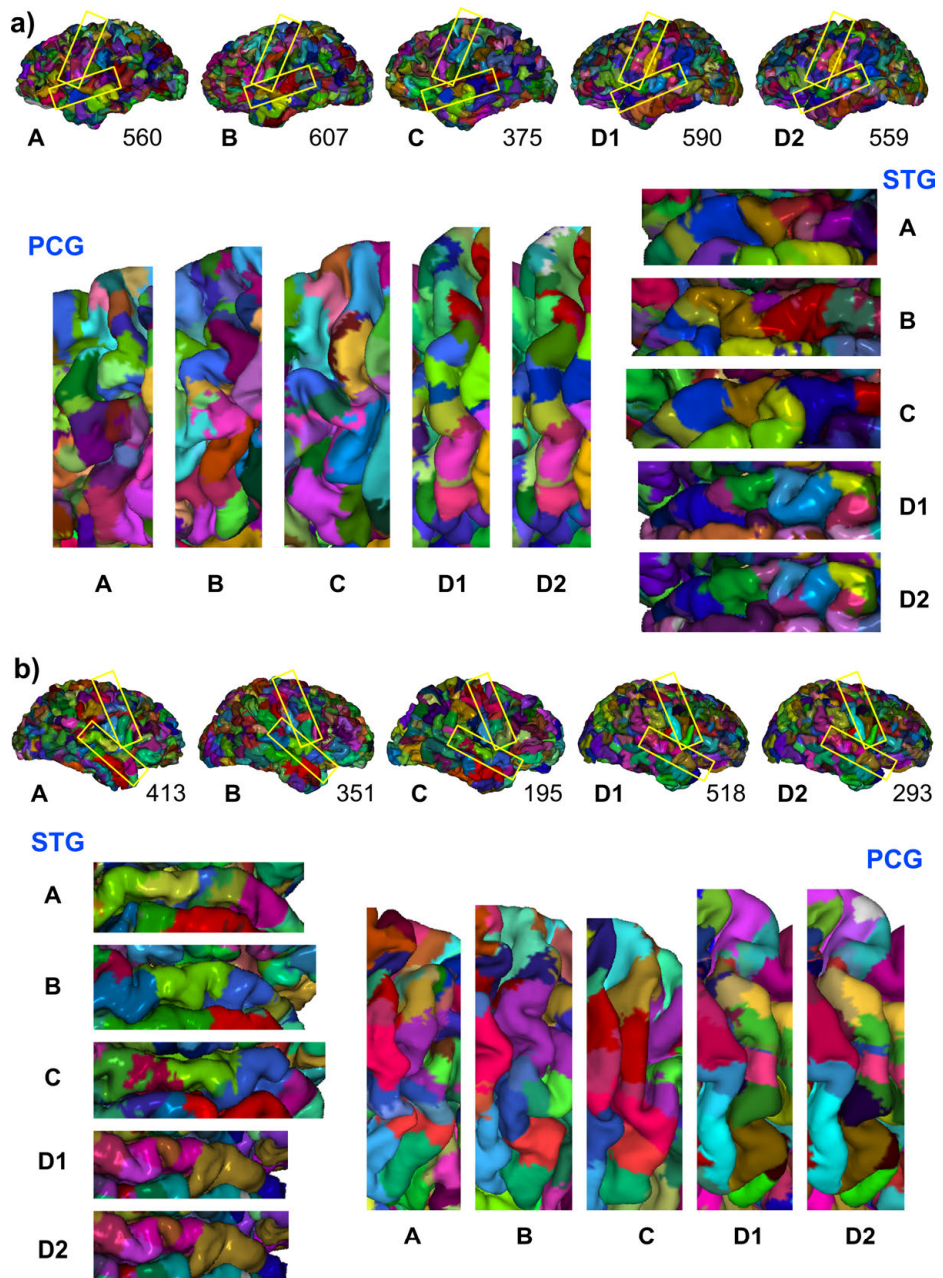
Figure 4.15:   Partitions with maximum spread-separation index for all subjects' left (a) and right (b) hemispheres. The top subpanels show the whole brain parcellation, the bottom subpanels zoom into the superior temporal gyrus area and the precentral gyrus.

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface   Inflated surface
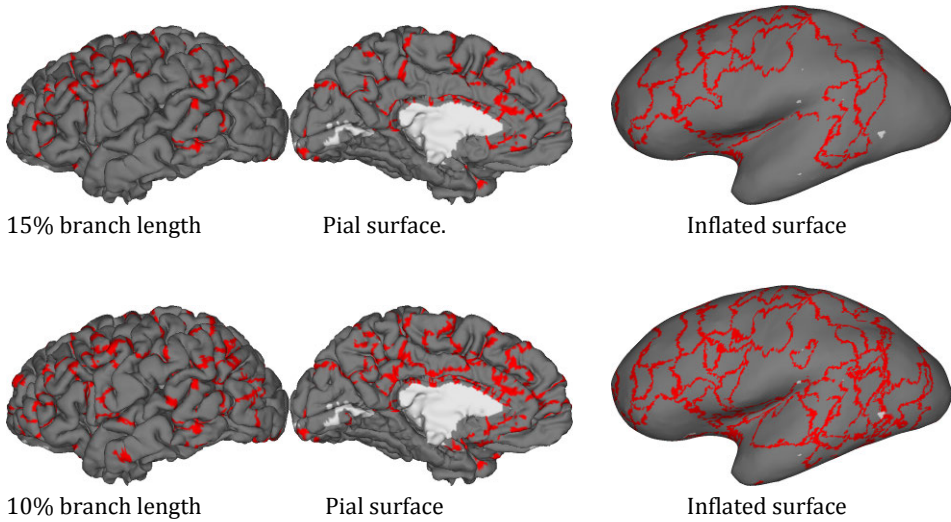
Figure 4.16:   *Stable boundaries at different granularity range widths for the left hemisphere of subject A. For branch lengths equal or greater than 15% of the tree height (top) and 10% (bottom).*

We can see that the boundaries between regions corresponding to main connections through the longitudinal, arcuate and temporal bundles are the first to be detected, as well as regions in the superior frontal cortex. The information in this maps should be interpreted differently as in our previous parcellations: regions across boundaries have significantly different connectivity patterns (relative to the value of the threshold parameter selected), but it does not give any information on wether the region within a boundary has a cohesive pattern or wether that pattern does change but with less distinctiveness. Changing the threshold value to less restrictive levels shows the additional boundaries that now meet the criteria.

However, this principle requires more study, and the algorithm, results and visualization method presented here are only preliminary.

5. Approaches and challenges in validation of tractography-based clustering

5.1 Overview

For a clustering method to be truly useful and reliable, verification tests and validation of its results must be carried out. In the case of in-vivo brain clustering based on tractography though, this is a very challenging task, mainly due to the lack of a ground truth to compare the results to, and partly due to the difficulty to verify and validate tractography itself. In this chapter, we will firstly discuss the challenges and possibilities for verification and validation of tractography and clustering based on tractography (see Johansen-Berg and Behrens [2009] and Knösche and Tittgemeyer [2011] for a more detailed discussion on tractography and tractography based clustering validity, respectively, from where part of the information here presented was extracted), and secondly present a pair of circumstantial verification and validation tests performed in order to give a higher degree of confidence to our method.

5.2 The challenge of verification and validation in brain dMRI based methods

5.2.1 Introduction

Verification aims to confirm that the designed algorithm is really working as it is expected and supposed to. It is an internal process and it is more commonly carried out by studying the results of the algorithm over synthetic data were the correct outcome is known by design. On the other hand, validation aims to prove that the results obtain by the algorithm on its real intended target datasets is correct, and for this purpose a comparison with a ground truth or gold standard is commonly used.

In the context of tractography based parcellation this can be done at two levels: assessing how faithfully the tractography algorithm (and its underlying model fitting) uncovers the structural connections in the brain; and assessing how faithfully the clustering algorithm represent the natural divisions and connectivity-pattern differences in the cortex. Both of them mainly suffer from a lack of ground truth.

5.2.2   Verification and validation in tractography

Tractography methods embed many assumptions that can potentially lead to errors and diminish the fidelity of the resulting brain pathway representations. It is therefore important to ensure their validation in order to safely interpret the obtained results and be aware of their limitations. The main factors influencing tractography results are the fiber-orientation inference step (relevant neuronal structures are much smaller than the available resolution of dMRI, usually in the millimeter range, where thousands of axons traverse the volume of a voxel), data acquisition parameters (voxel resolution, angular resolution, SNR, number of diffusion weighting, etc.) and the choice of the particular tractography algorithm and its parameters (local model, deterministic, probabilistic, maximum angle of turns allowed, stopping criteria, etc.). Without understanding the sources of error and their influence in the obtained tracts, tractography could fail to characterize physical brain connections and even yield misleading results.

As mentioned before, verification can be carried out by testing different tractography methods on synthetic datasets. These might be artificially constructed structures simulating the diffusivity properties of neural tissue, called physical phantoms, or artificially generated datasets, called software phantoms. Testing with these phantoms can help find out how a given algorithm performs in certain conditions.

Physical phantoms are a straightforward verification method where tractography is over data obtained from these artificially constructed structures, providing a well-defined ground truth against which to compare the output of the algorithms (Li et al., 2012). Such phantoms have been constructed from synthetic fibers such as rayon, cotton (Scifo et al., 2004), polyester (Watanabe et al., 2006), ultra-molecular weight polyethylene (Fieremans et al., 2005), and hemodialysis fibers (Perrin et al., 2005a). However, the unavoidable differences in microstructure and diffusivity properties of synthetic fibers and neural tissue, limits the degree of verification that they can offer.

Software phantoms have the advantage of high flexibility, changing relatively easily the conditions where the tractography algorithm is tested, compared to their physical counterparts. As examples, Lazar and Alexander (2003) generated sets of artificial datasets with different curvatures to compare deterministic approaches and Lori and colleagues (2002) studied the effects of noise and encoding strategies using Monte Carlo methods. An important limitation apparent from the literature however, is that simpler phantoms favor simpler tracking algorithms, and that it is usually possible to generate a phantom biased towards a given algorithm by using the same fundamental assumptions in its construction as were made in the tracking approach. Therefore in these phantoms there is a trade-off between the flexibility to alter the conditions tested and the degree of ground truth provided.

Both software and physical phantoms are useful to determine if given tractography implementations work as expected under different predefined circumstances. However they are both ultimately gross approximations of the real in-vivo scenario they simulate and not neurobiologically realistic, not able to fully characterize the full complexity of the microanatomical neural structure that affects the dMRI signal and the tractography.

"The search for a gold standard for tractography can be thought of as a search for an ideal model that possesses true axonal characteristics, combined with an ideal model of signal generation from a population of such fibers" (Johansen-Berg and Behrens, 2009). As of yet no such software or physical phantom has been created.

In order to perform validation of a method, it needs to be tested on real target data. The best way to approximate a "ground truth" in this case is using data from histological tracers (see section 1.4.3). These allow to map connections through the injection of a visible tracer substance that then travels through the neuronal axons. These techniques are very invasive, so the closest data available to in-vivo human brain are post-mortem tracing of human brain or *in-vivo* tracing of macaque brain (through databases that are characterized by their own biases and limitations, see Bakker et al., [2012] for a detailed discussion). Having a "ground truth" approximation, different tractography algorithms can be used to assess the best performing one (Iturria-Medina et al., 2011; Li et al., 2012). The main obvious limitation of these methods is the inter-species differences in the case of the macaque, and the limited availability and variation of properties of post-mortem tissue. In addition, these techniques can be used only in a reduced area or a limited number of times per individual/sample, so for the case of mapping whole brain connectivity, data from many individuals or samples must be compiled, not being able to obtain a full tracing connectome for an individual brain.

A last alternative is to have "empirical indexes" for assessing the quality of certain algorithms (Bastiani et al., 2012). However this approach requires demarcation of areas so the case of a full brain connectome becomes rather unfeasible.

We have presented above the main options for verifying and validating tractography algorithms, along with their limitations, as it is highly relevant to understand verification and validation of clustering based on this data (which, along with its challenges in a whole brain approach, which will be covered next). However, in our study we have strived to design our algorithm as independent of the underlying tractography as possible, as long as the end result can be presented in the same data structure (ultimately the similarity between connectivity fingerprints). The verification and validation of tractography itself is out of the scope of this thesis, as we focus on providing a proof of principle for a whole-brain high-resolution multi-granularity connectivity characterization and clustering framework. Therefore, a tractography method with good published results (Anwander et al., 2007), proved in agreement with prior knowledge and other methodologies (Ruschel et al., 2013), and with very fast computation time, important for our high resolution whole-brain approach, was chosen (see discussion chapter 6 for further justification of this choice).

5.2.3   Verification and validation in tractography-based clustering

In this thesis, a framework for characterizing and clustering the whole human brain cortex based on tractography information was developed. In order to rule out methodological errors in the algorithms, and to ensure that the results our method yields are neurobiologically relevant, verification and validation must also be performed. Following a similar logic as for tractography in the previous section, we will consider and discuss possible options to achieve this goal.

Also in this case, an initial choice for verification would be the use of physical phantoms. But in order to be useful for clustering based on tractography, such a phantom would need to have to have the same qualities as for tractography validation, plus a much higher degree of fiber complexity. It would not be sufficient to simulate the diffusivity properties of fiber bundles, but their layout and distributions must provide scenarios relevant and challenging enough for testing the outcome of clustering algorithms. This is, if not entirely unfeasible, extremely challenging. No such phantom has been created up to date to the authors' knowledge.

A more approachable alternative is the use of software generated data, in this instance synthetic tractography datasets, and run clustering simulations on them. However, as with physical phantoms, in order to provide truly informative verification, these datasets must simulate real data to a sufficient degree (although in this case not anymore the diffusivity but directly fiber tracts shape and trajectories) and at the same time their layout and distribution must be clearly defined so that it can be assessed whether the clustering algorithm classified them appropriately or not. For our algorithm this would preferably also include presence of clear hierarchical relations.

Such a test dataset would indeed be a great tool for the verification of not only our method, but also most other tractography clustering methods in the literature. However, such a synthetic dataset is not available to the author's knowledge, and generating one of enough quality would be, while feasible, quite a laborious task (a dataset too simplistic would not be able to shed any clear verification conclusions), requiring a considerable amount of invested time (maybe an investment worth of a project on its own).

Given the scope of work aimed to be included in this thesis, building a framework of algorithms and concepts wide enough to test the potential benefits of whole brain hierarchical parcellation (from choice of agglomerative method to partition selection and tree comparison) and with the objective to serve as a proof of principle, a faster and simpler option for basic verification was favored. A synthetic dataset of randomized values was generated in order to test the performance of the algorithm against a set containing no relevant information, setting a baseline for quality control, with favorable results (see section 4.3). We consider that this provides a certain level of confidence enough for a proof of principle, but do however acknowledge that it does not replace an exhaustive verification, which should be covered by future work.

Contemplating validation clustering based on tractography the question arises: What is the "ground truth"? How many areas/clusters should constitute the human brain? Where do their boundaries lie? We simply do not know, especially when it comes to the whole brain. Focusing in smaller regions, such as boundaries between SMA/preSMA might seem more manageable but yet again which can be considered the "true borders" of these areas? Some validation can be obtained by comparing clustering results in brain areas where white matter anatomy is well understood and connectivity differences are clear, but the number of such areas is limited (see Figure 4.12 in section 4.5 for focus of our clustering results in the IFG). Macroscopic landmarks may also be used, but they are linked to cortical areas to a questionable degree (there is for instance no firm protocol to distinguish the prefrontal Brodmann area 9/46v from area 8B through macroscopic landmarks

with certainty, although it is easier in areas like the primary and secondary visual areas (V1/V2).

A valid solution would be the use of post-mortem dMRI tractography and invasive tracing on the same sample of post-mortem brain (which would also eliminate the confound of inter-subject variability). However, this would only be feasible in a limited area of the cortex and not for in a whole-brain approach, though it still would be an interesting validation scenario. Regrettably, such resources were not available to our group during the realization of this thesis.

Another interesting option is the use of circumstantial validation/verification by comparing the results obtained for a certain dataset with those obtained from other complementary modalities, or from a different algorithm of the same modality that has given good results. While this may not constitute a "true" validation as results are not being compared with a ground truth, and in case of different modalities results may not be necessarily overlapping, a good agreement between different approaches provides a level of confidence that the clustering is yielding reasonable results. With this objective, two comparisons within and across modalities were carried out, laid out in the following section.

5.3 Circumstantial validation of our clustering algorithm

Given the challenges exposed in the previous section, we opted to bring a certain degree of verification to our method by comparing it to some other widely used and tested method of clustering based on tractography. While there is a lack of an established method for whole brain dMRI based parcellation (which motivated this thesis), other methods are available that, while not well suited for whole brain tractography, have given good results for smaller regions (see section 2.4.2 and 6.2).

We chose to replicate the results of Ruschel and colleagues (2013) who carried out k-means clustering based on probabilistic tractography of the inferior parietal cortex convexity (IPCC) of 20 healthy participants (10 females). In their study they found 3 clusters to be the number that best stabilized results across subjects and were in agreement with previous dMRI-based findings (Rushworth et al., 2006; Mars et al., 2011) and macaque data (Gregoriou et al., 2006; Rozzi et al., 2006, 2008). Their results for the 4 subjects featured in their published figure are shown in Figure 5.1.

Subject 1   Subject 2

Subject 3   Subject 4
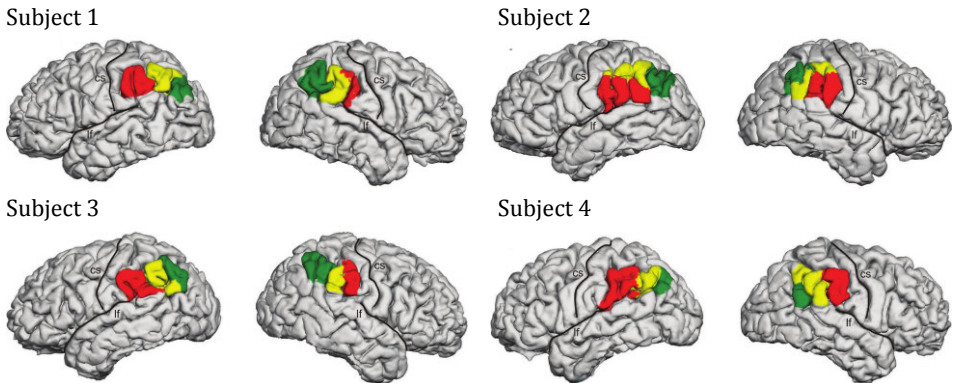
Figure 5.1:   Parcellations of the left and right IPCC of 4 representative participants into anterior (red), middle (yellow) and posterior (green) IPCC, superimposed on a freesurfer reconstruction of the pial surface; cs, central sulcus; lf, lateral fissure. (reprint from (Ruschel et al., 2013)).

Using the tractography data from their study, we built hierarchical trees with our centroid-neighborhood method for the same IPCC areas, and partitioned the tree in 3 clusters and a higher 5-cluster granularity for each individual using the *SS* method. Results are shown in Figure 5.2.

It can be observed that for more than half of the datasets presented the clusters obtained are virtually identical to those suggested by Ruschel and colleagues in their work (left hemisphere of subject 1, right hemispheres of subjects 2 and 3, and both hemispheres of subject 4). The other three cases, while having small differences in the boundary between the middle and posterior clusters at the 3 clusters partition, yield completely overlapping boundaries when granularity is increased to 5 clusters.

This shows a remarkable level of agreement between the solutions from both algorithms, and hints further at the idea already put forward that in our hierarchical parcellation, the boundaries might be of singular relevance themselves rather than partitions at specific granularities. It also comes to show how hierarchical clustering is an extension of the narrower concept of partitional clustering such as *k*-means (complementing, rather than invalidating, the results obtained with the latter), containing within the tree much more additional information about finer granularity boundaries. Comparative results for all the subjects used in the study and presented in (Ruschel, 2013), showing the same trend, are included in the Appendix.

Subject 1

Subject 2

Subject 3

Subject 4

3 clusters   5 clusters
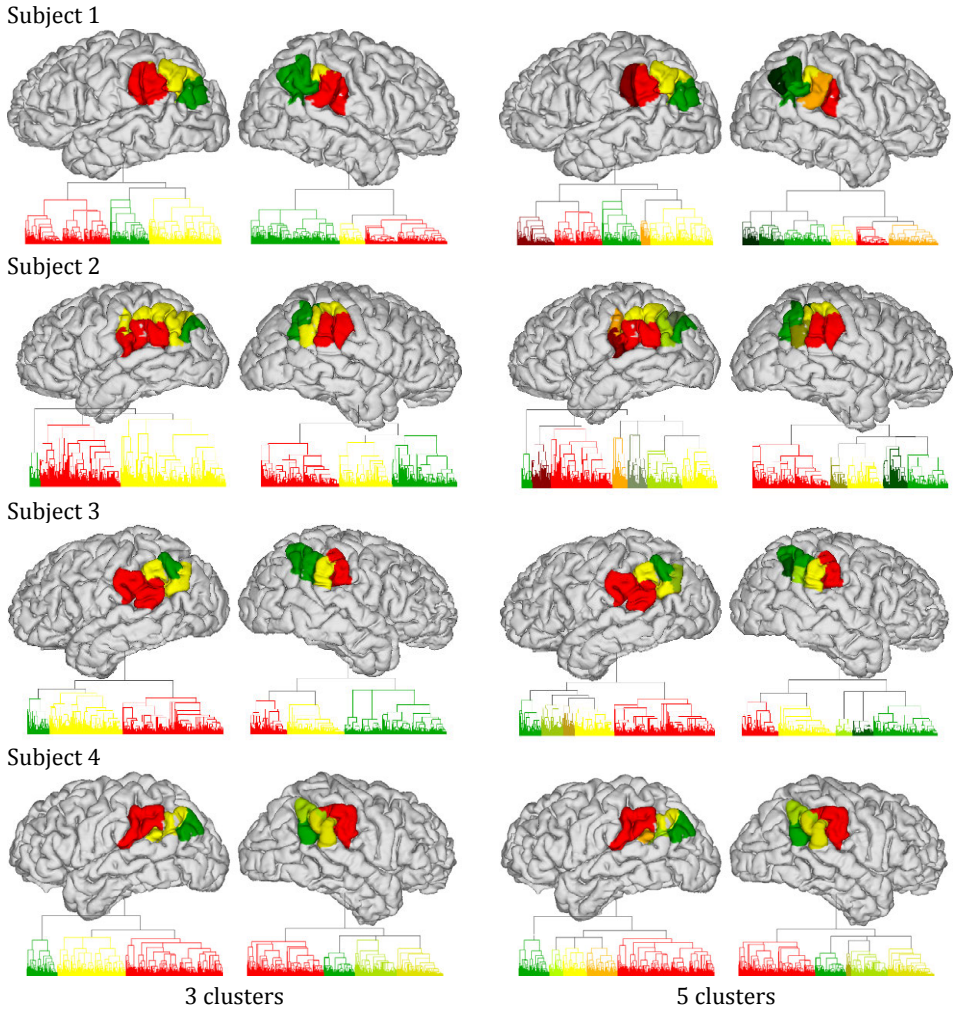
Figure 5.2:   Centroid hierarchical trees and SS partitioning at 3 clusters (left) and 5 clusters
(right) of the same subjects on Figure 5.1, computed on tractography data from
Ruschel and colleagues (2003).

   Other possible suggestion would be to have the probabilistic maps of areas
derived from histology as a ground truth, but problems exist there as well: These
maps might have "residual variability" due to non-optimal cutting of the cortex for
subsequent histology (all brains are cut coronally – this is not optimal for detecting
borders between all cortical areas); Connectional subdivisions might exist within
an area that appears as unitary and defined on cytoarchitectonic features (i.e.
dorsal premotor area F2 can be divided to sectors F2d and F2vr based on

pronounced connection differences; Luppino et al., 2005), thus, cortical areas previously identified as homogenous might harbor further connectional subdivisions. Moreover, available probabilistic maps are mainly derived from *Nissl* stained sections. It can be the case that *Nissl* stained sections are not optimal to unveil borders between areas whereas immunohistochemistry like staining for somatostatin might give rise to a different picture (Geyer et al., 2000b).

Notwithstanding, we compared the output of our algorithm to the cytoarchitectonic parcellation available from *Jülich Research Center* (`https://www.jubrain.fz-juelich.de/apps/cytoviewer/cytoviewer-main.php`).

The similarity to our parcellation is exemplified in Figure 5.3. In order to carry out the comparison, the partitioning through the *SS* method at 100 clusters from subject A was taken. Clusters corresponding to the *Jülich* parcellation were manually color-matched, and in the cases of the precentral gyrus and postcentral gyrus, its constituting clusters merged, as in our partitioning these gyri were further subdivided. Areas not covered by the Jülich map were greyed out.
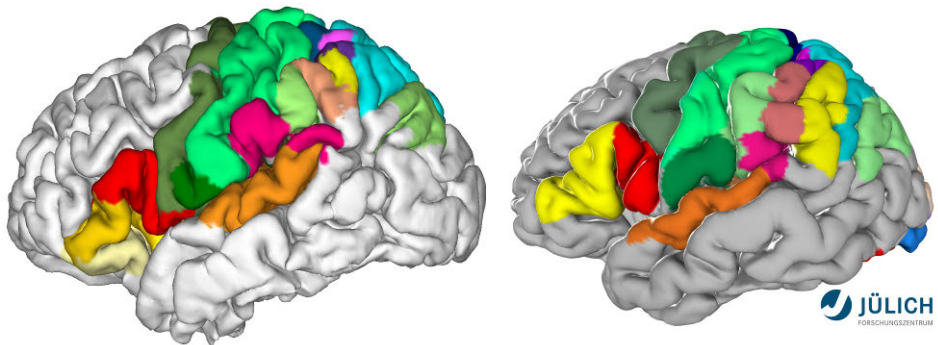
Figure 5.3:   *Cytoarchitectonic parcellation provided by Jülich Research Center (right), compared to the corresponding subtree of the left hemisphere of Subject A at a global horizontal partition for 100 clusters (left; two clusters, one in the IFG, and other in the parietal cortex over the STG, have been further subdivided once to better show the corresponding matching).*

We believe that both these comparison studies which show a high level of agreement, do by no means replace a full and exhaustive verification and validation. However, they bring a degree of confidence in the method proposed to show its value and potential for whole brain connectivity characterization. Other potential sources for circumstantial validation are discussed in section 6.8 at the end of the Discussion chapter.

6. Discussion

6.1 Tractography-based parcellation

As argued before, connectivity is among the most relevant structural cues for the characterization of the function-anatomical identity of cortical tissue. Being the only method that can be applied to healthy human subjects, diffusion tractography is the method of choice for the reconstruction of these connectivity patterns (Anwander et al., 2007; Johansen-Berg et al., 2004). For a thorough discussion of this issue, see Knösche and Tittgemeyer (2011).

The tractography based parcellation requires a robust tractography method. The local tensor model based on High Angular Resolution Diffusion Images (HARDI) allows a reproducible computation of the connectivity profile. The method is sensitive to small changes in connectivity between two voxels and is robust to noise which could affect the local model. Other tractography methods like the Probabilistic Index of Connectivity (PICo) based on the Persistent Angular Structure (PAS) (Parker and Alexander, 2005) or probabilistic tractography based on spherical deconvolution (Descoteaux et al, 2009) had shown to better resolve crossing fiber structures. The more complex local model might have been less robust to remaining noise in the diffusion data, which might have affected the local estimation of the fiber orientations (Yo et al, 2009). While comparing a tensor based tractography with fiber tracking using spherical deconvolution Kristo and colleagues (2013) showed a higher reproducibility for the tensor based tractography. In this initial study we choose to use the more robust local model. The fact that probabilistic tractography is employed ensures that, to a certain degree, fiber crossings and branching are taken into account. The parcellation method we proposed could be applied on any other tractography method. The comparison of the result using different local models and tractography algorithms will be subject of future investigations. In addition, all tractography algorithms including the one used here have a number of adjustable parameters, which potentially can affect the tractography result and the parcellation. For example, here we had to make choices on the number of streamlines, the scaling and the

thresholding of the tractograms and the sharpening of the local diffusion profile (Anwander et al., 2007). While a systematic parameter study on this and other tractography algorithms would certainly be very useful, the previous use of our approach in a number of parcellation studies yielding neuroanatomically plausible results provides some confidence (e.g., Anwander et al., 2007; Ruschel et al. 2013; Schubotz et al., 2010; Gorbach et al. 2011, 2012).

In most implementations of tractography based parcellation the target space comprises the entire rest of the brain, including white matter (Anwander et al., 2007; Johansen-Berg et al., 2004; Mars et al., 2011; Schubotz et al., 2010; Tomassini et al., 2007). A possible alternative is to restrict the target space to grey matter (or, for technical reasons, the white matter voxels just adjacent to grey matter) (e.g., Bach et al., 2011). It is, however, not clear whether this really improves the situation. Most tractography methods are iterative algorithms that, especially over long distances, tend to accumulate errors and hence are subject to substantial blurring (Jones, 2010). So, it is likely that differences between tracts, which are still quite evident in the intermediate white matter, become smoothed out at the distant cortical targets. On the other hand, using the entire brain as target space might also introduce biases of its own, as the tracts starting from two spatially distinct cortical elements are different by definition in their initial sections, even if they finally reach the same targets. This is especially true, if the tracts start in different gyri. How much this effect influences the result depends on the overall extent of the tractogram, that is, the relative weight of short and long range connections. So, the fact that parcellations often seem to reflect, to some degree, sulcal patterns (see Figures 37 to 41), might have a methodological background. On the other hand, it is well known that in many cases macroanatomical landmarks, such as sulcal lines, are indeed likely to play a role as function-anatomical boundaries (Hasnain et al., 2001; Tahmasebi et al., 2012). To what extent correlation between gyrification and tractography based parcellation is a product of methodological peculiarities or reflects neuroanatomical reality remains to be investigated.

6.2 Advantages and limitations of hierarchical clustering

In this work, we propose a hierarchical clustering method for the analysis of high-resolution, whole-brain anatomical connectivity data that provides an optimal data compression with minimal information loss. The method uses

differences in connectivity patterns for drawing a function-anatomical map of the cortex without the need to choose a particular granularity level. This way, almost all of the information on the connectivity pattern similarities is retained and all possible parcellations of the cortical sheet are not only stored, but also related to each other in a meaningful way. While this concept is not entirely new (Blumensath et al., 2013; Guevara et al., 2011), it is the first time that is applied to whole-brain diffusion based anatomical connectivity data. Compared to classical single-partition connectivity-based brain parcellation methods (for a review, see, Knösche and Tittgemeyer, 2011), it offers a number of advantages.

First, it is important to compare function-anatomical maps between subjects or between different datasets of the same subject (e.g., at different ages). With single-partition parcellation, one has to choose a particular level of granularity in order to obtain a parcellation. This level of granularity can be expressed, for example, by the number of desired clusters, by the differences between or the homogeneity within clusters, or by the sizes of the clusters. All these criteria can require different values in different datasets for defining the same function-anatomical subdivision. It is therefore difficult to obtain comparable parcellations. Moreover, there might be more than one level of granularity relevant for the comparison. Using the whole information encoded in hierarchical trees, connectivity similarity (and therefore function-anatomical organization of the cortex) can be compared efficiently without any explicit choices on granularities. Such comparisons can be potentially used to show changes or differences in the function-anatomical organization of the brain in a great number of settings, including disease, development, aging and cognitive abilities. The particular advantage is that one can start at a general comparison (i.e., comparing the entire trees) without making any choices or assumptions, and then gradual zoom into certain parts of the trees (i.e., comparing subtrees) and/or particular levels of detail (i.e., pruning the lower level nodes).

Second, if larger parts of the cortex or the entire brain are to be parcellated, the definition of a granularity level, as required by non-hierarchical methods, becomes quite arbitrary. Even if comparison is not the goal, it is not easy to say, how many clusters are to be expected or how big they are. Also, the magnitude of difference between parcels depends on the brain region. For example, regions near large fiber tracts, such as the arcuate fascicle, tend to exhibit higher similarity in terms of their connectivity pattern, requiring lower thresholds for parcellation. Hierarchical parcellation circumvents the granularity choice. The obtained trees can be explored interactively in order to discover the function-anatomical organization in different brain regions. Of course, it remains an important issue to extract actual partitions of the cortex from the tree (see below).

Third, the hierarchical trees encode the interrelation between different levels of description of the function-anatomical cortex organization, from relatively local to very global. In fact, using very high resolution MRI data one could even imagine bridging the gap between microscopic and macroscopic levels (see Heidemann et al., 2012, for an intermediate step into that direction). This is of particular importance, if the parcellation is used as a basis for building a connectome. If the connectome is truly, as defined by Sporns (2011b), *"a comprehensive structural description of the network of elements and connections forming the human brain"*, it essentially has to span multiple levels of detail. Using the parcels of a hierarchical parcellation as the elements of the connectome could lead to a hierarchical connectome that not only describes the brain network at different levels of detail, but also encodes the relations between these levels. Note, however, that the construction of a true connectome relies on adequacy of the employed connectivity measures in terms of the true function-anatomical structure of the brain. Certainly, non-invasive measures based on MRI, valuable as they may be, bear significant limitations in that respect. The parcellation resulting from hierarchical clustering could also be used as initial regions for global tractography methods like the recently proposed plausibility tracking method (Schreiber et al., 2014).

Nevertheless, hierarchical clustering also suffers from some principled limitations. Given its iterative agglomerative nature, established mergers cannot be undone. The procedure therefore has some sensitivity to local effects and errors may propagate, missing on the global optimum, when considering specific partitions. For these reasons, in scenarios dealing with small datasets or when only a single optimal partition is desired, optimization based methods such as *k*-means or model-based methods might be more adequate. However, for large datasets and a large number of expected clusters, these other methods may lead to exploding complexity and computation power requirements in order to achieve acceptable reliability (by design in the case of model-based methods, in order to maintain stability against local effects due to initial conditions in the case of *k*-means: Kuncheva and Vetrov, 2006; Pham et al., 2005). For these reasons, we strongly believe that in our scenario of whole brain parcellation, the advantages that hierarchical clustering offers (namely: multiple-nested-granularity, possibility for whole-structure comparison, and scalability with dataset size) greatly compensate for its limitations.

6.3 Meta-leaf matching

For the comparison of any cortical map between datasets, hierarchical parcellations being no exception, it is necessary to establish a correspondence between the cortical elements. In other words, we need to decide for each element (e.g., voxel) in one dataset, what is the function-anatomically equivalent element in the other dataset. This is not a big issue when comparing repeated measurements of the same subject, but due to the natural anatomical variability (Thompson et al., 1996) it poses quite a challenge if we want to compare across subjects or hemispheres. Attempts to obtain such a mapping on the basis of structural MRI have resulted in numerous linear and non-linear registration algorithms (e.g., matching of *freesurfer* surfaces nodes; Roca et al., 2010), but the results are not always satisfactory, in particular if the surfaces differ in terms of number and orientation of gyri and sulci (Ono et al., 1990). Here, this problem concerns the meta-leaf identification between trees, which was achieved by maximizing mean-tractogram similarities using a greedy algorithm. This approach relies on the assumption that the connectivity pattern is a good reflection of the function-anatomical identity of a cortical element - the same assumption that underlies the entire connectivity-based parcellation idea. For a more detailed discussion of the justification of this assumption, see Knösche and Tittgemeyer (2011). Our analysis showed that the meta-leaf similarity method yields meaningful comparisons between trees. However, at this stage, inter-subject matching is not always stable enough to quantitatively interpret small variations in them. The leaf matching is certainly one of the current challenges of the method. It remains to be investigated whether other matching strategies, like the "Hungarian" method (Kuhn, 1955), yield an improvement. In general, however, it is not likely that by improved mathematical algorithms alone this issue is going to be resolved in a satisfactory way. Instead, the very notion of function-anatomical equivalence needs to be refined. A comprehensive and reproducible definition of the equivalence of elements in two brains would provide solid ground from which to gauge any difference in structural properties or functional organization. Such a mapping would have to be unique, that is, each element in one brain must be assigned to exactly one element in the other brain, and vice versa. Furthermore, as the leaf matching criterion has of course a profound influence of the resulting tree comparison results, it has to be biologically meaningful. In other words, only if we have good reason to compare an element in one brain to just a particular element in the other brain (and not to any other), it makes sense to interpret their differences in, for example, connectivity or cytoarchitecture. Similar connectivity

to the rest of the brain is certainly a good starting point for such an equivalence criterion, but it is surely not the ultimate solution. An interesting option might be guiding mesh matching with connectivity properties, as proposed (using much smaller pattern vectors) by Cathier and Mangin (2006) or Petrovic and Zollei (2011).

6.4 Tree comparison

The hierarchical tree allows for comparison of the whole connectivity similarity structure across measurements, and not just particular partitions, which is not possible with the other methods. Note that the tree does actually contain all possible partitions together with their mutual relationships.

This comparison measure gives us the degree by which the structure of the connectivity similarity organization varies across different measurements. More specifically, the *tCPCC* measure focuses on the actual degree of similarity between connectivity patterns, while wTriples measures topological similarity (for example if the region most similar to a given selected area is the same in both measurements).

Unfortunately, compared to repeated measurements, the quality of meta-leaf matching across subjects or hemispheres inevitably decreases (see above), and so does the reliability of the comparison. There might be two possible solutions to this problem: either improving the quality of the matching by using more sophisticated methods, like combining surface topology information with connectivity pattern information (although this is unlikely to boost the quality to the same level as repeated measurements), or accepting that, due to the inter-subject variability, a perfect matching at high granularities is not possible, and trying to establish suitable levels at which the matching may be done with sufficient quality (one would have to be aware that the matching results obtained are only valid at those granularities).

6.5 Extraction of partitions

Although a hierarchical tree in its entirety comprises the joint information of all possible partitions and their mutual relations, concrete anatomical interpretation requires the generation of actual partitions. As a compromise between single partitions and the entire tree, we characterized the hierarchical structure of the

trees through series of partitions at different levels of granularity. Several partition schemes were implemented. Horizontal partitioning was shown to be a good approximation of the more sophisticated spread separation (*SS*) partitioning for a given granularity level. These partitions are very stable against noise and the boundaries have a high degree of reproducibility across subjects. In order to palliate the tendency of regions of the cortex that share large common tracts to remain in a single cluster across a higher range of granularities, a minimum size-difference clustering was implemented. This method effectively extracts more homogeneous parcellations.

Calculating the *SS* index for every granularity level, we showed that for each data set there is an entire range of similarly good partitions (approx. between 50 and 200 clusters). This fact raises general concerns about the search for a single optimal partition or even a series of a few partitions. Although one is able to single out one partition with the highest information content (in some sense) of all partitions, this information might still be completely insufficient to describe the entire structure. Hence, one has to try to find ways to (approximately) represent entire classes of parcellations in an effective manner. As each bifurcation in the tree represents the separation between two clusters (i.e., a boundary), such a technique could aim at finding the most relevant or persistent boundaries rather than entire parcellations. An idea would be to look at the branch lengths of the nodes involved. The longer the branch (in absolute value or in relation to the node height), the more stable that region is in comparison to its neighboring ones. This way, important boundaries would be mapped on the cortex, rather than entire parcellations. However, this principle needs further investigation.

The extracted partitions could be used to do a connectome-based analysis of connectivity (Hagmann et al., 2008) or as *a priori* partition for white matter fiber analysis (Wassermann et al., 2010). Within each method, partitions are always fully nested. This eases the interpretation of the boundary changes from one granularity level to the next. On the other hand, in an agglomerative method the information about the fuzzyness of the changes in connectivity similarity is not as well captured as in other approaches (Cerliani et al., 2012; Gorbach et al., 2011), although it might be extracted to a limited degree from the tree topology.

6.6 Relationship to other multi-granularity methods

As explained above, multi-granularity methods like the one proposed here offer several general advantages over single-partition methods: they yield a more exhaustive representation of the real connectivity similarity structure; they are preferable for the analysis of larger regions (up to entire hemispheres or brains), due to the expectation that different boundaries may be relevant at different levels of granularity; they facilitate comparisons between data sets; and they allow for adaptive parcellation depending on the features that we would like to emphasize.

Other researchers have approached multi-granularity in different ways. For example, Kahnt and colleagues (2012) generated a series of *k*-means based parcellations from resting-state fMRI data of the orbito-frontal cortex using different numbers of expected clusters. The fundamental difference between their approach and the one proposed in the current work lies in the fact that the hierarchical tree imposes a constraint on the relationship between the different parcellations, in that finer parcellations are nested in the coarser ones. Hence, in our method any finer subdivision complements, rather than competes with, the previous parcellation. Moreover, the embedding of the parcellations into a tree structure yields immediate clues about the distinctness and stability of certain boundaries, as well as to the topological relationship between different parcellations. An effort to bring multiple *k*-means parcellations at different granularities into a hierarchy has been presented for fMRI co-activation data by Clos and colleagues (2013), where hierarchically inconsistent voxels from the clusters obtained are removed resulting in nested partitions.

The work of Gorbach and colleagues (2011) takes a different approach to multi-granularity by obtaining a "space" of optimal parcellations from dMRI data through an information bottleneck method, minimizing the tradeoff between data compression and information preservation. For each desired granularity, the number of clusters is determined by a Lagrange multiplier parameter and an upper boundary for the number of clusters. In their approach, while boundaries are not necessarily nested across granularities, they seem more stable. The method may have an advantage over agglomerative methods at granularity levels where changes are gradual and boundaries fuzzy. It offers a solution between nested partitions and single partitioning at multiple levels. However, computational costs also escalate for growing datasets and granularities.

In comparison, our approach tries to characterize the whole connectivity similarity information in a compact tree, which is then easy to process. As demonstrated by the high *CPCC* values, most information of the connectivity similarity matrix (N^2 floating points, with N being the number of tractogram seed voxels) is successfully encoded with only a fraction of the size ($2N$ floating points plus $2N$ integers, easily stored as an ASCII text file). Furthermore, the number of tractogram similarities that must be computed in order to obtain the tree is 3 orders of magnitude lower than that needed to compute the matrix. This is an important advantage, given that tractogram similarity computation is a costly operation, if, like in our case, all the white matter is used as target space and high resolution (1 mm) is used (amounting to more than $15 \cdot 10^5$ floating point operations).

However, the use of multi-granularity methods does not yet solve the problem of selecting relevant partitions. Cluster number selection remains an open problem in connectivity-based clustering literature. Various solutions have been proposed to solve it, such as visual inspection of reordered connectivity matrices (Johansen-Berg et al., 2004), consistency across subjects (Ruschel et al., 2013), correspondence with cytoarchitectonic maps (Anwander et al., 2007), hierarchical consistency (when using optimization methods for different numbers of expected clusters; Clos et al., 2013), variation of information (Kahnt et al., 2012; Clos et al., 2013), information-based model selection (Gorbach et al., 2012), consistency across modalities (Kelly et al., 2012) and the tree-based methods we propose here, which are especially suitable for whole brain parcellation. The hierarchical tree method is actually open to all these approaches, while offering a much richer stock of available partitions, among which to select.

6.7   Fine-tuning of parameters

Our clustering includes a number of parameters whose values influence the results obtained (e.g.,. neighborhood used and maximum cluster size restriction while building the tree; minimum tract dissimilarity to be considered an outlier; coefficients of the tree preprocessing pipeline etc.). Due to the small size of the subject sample used in the study, this parameters were optimized for a single subject (in terms of best trade-off between information loss and tree complexity reduction), and the values applied for all other subjects (controlling that for the other subjects this trade-off stayed within the same range). A future study with a larger number of subjects should also include further optimization of these

parameters via a bigger training sample. However, it is not expected that results change dramatically.

But even before clustering is applied, many decisions are taken that can greatly affect the outcome of the parcellation: the choice of the connectivity fingerprints (in the case of anatomical connectivity: full tracts through white matter (this thesis; Anwander et al., 2007; Johansen-Berg et al., 2004) or grey matter end-points endpoints (Bach et al., 2011); the application or not of data smoothing, the choice of similarity measure between fingerprints (correlation, mutual information, independence of components...).

The consequences of these choices can be as important as the choice of clustering algorithm itself and its parameters and finding an optimal solution is not a simple problem. In most studies these matters are usually not investigated and the choices that are deemed more convenient for the specific objective at hand are taken. No study up to date has systematically characterized the effects and meaning of all these choices in the final parcellations/connectivity values. There is an urgent need for systematic evaluation studies that meticulously scrutinize the influence all these choice have onto the results.

6.8 Biological validity

Here we made a proposal how to account for the structural organization of the cortex based on anatomical connectivity measures. A key question that remains is the one for the biological relevance of the obtained results. First of all, our method is primarily a way to represent given information in a convenient way. Hence the validity and relevance of the parcellations hinges on the appropriateness of the underlying diffusion tractography. However, on top of this, also the construction of the tree and the selection of partitions need to be evaluated.

As this is a proof-of-principle study we only offer some preliminary evaluation of the neurobiological significance of the results, for example by comparing the inferior frontal gyrus parcellation with cytoarchitectonic maps. Much remains to be done in future studies. In particular, within-subject validation will be crucial as it avoids the inevitable uncertainties of comparing different brains. For example, functional localizer tasks in fMRI experiments could be used to gauge the functional significance of parcellations (Schubotz et al., 2010; Johansen-Berg et al., 2004). Alternatively, resting-state functional connectivity (Kelly et al., 2012) and meta-analytic co-activation studies (Clos et al., 2013) also offer promising comparison possibilities. For example, one might apply the same method to

structural and functional connectivity measurements. In-vivo Brodmann mapping (Bazin et al., 2013) based on quantitative T1 imaging might offer another option. In this study we do not, and cannot, aim at the construction of a connectome or even a function-anatomical atlas. This will indeed require much more work, in particular involving many more subjects that in some sense are representative for a population.

7. Summary and outlook

It is commonly accepted among neuroscientists that the cerebral cortex can be subdivided into areas according to various structural criteria, such as cytoarchitecture, or receptorarchitecture (Amunts et al., 1999; Brodmann, 1909; Zilles et al., 1996). It is also generally agreed that brain structure is closely related to brain function and, therefore, structurally defined cortical areas tend to carry functional meaning. Consequently, many studies have aimed to find the boundaries between these areas, using a variety of techniques based on local structural tissue properties. However, the brain is not only a collection of isolated functional units; the different parts communicate and interact in a complex network ultimately resulting in higher cognitive capabilities. The connectivity pattern of a specific point in the cortex is, therefore, a major source of information about its function and an important parameter for the description and distinction of cortical areas (Knösche and Tittgemeyer, 2011; Passingham et al., 2002). The subdivision of the brain into function-anatomically defined areas is also a necessary step for the connectome, characterized by elements (the regions being connected) and the connections between them (Sporns, 2011b).

In this thesis, we have presented a framework to carry out whole-brain characterization and parcellation based on anatomical connectivity information from high-resolution dMRI images. A whole brain approach is faced with particular challenges not present in traditional brain-clustering scenarios. Namely, a high volume of data and an unknown distribution and number of clusters (these being subject to the desired granularity). In order to overcome these difficulties, a multi-granularity approach was chosen by using agglomerative-hierarchical clustering. The datapoints to be clustered were the connectivity fingerprints from single voxels in the gray matter / white matter cortical boundary. These fingerprints were computed via probabilistic tractography based on tensors with the whole white matter as target space (Anwander et al., 2007), but we have strived to build our framework as independent from the tractography choice as possible. A normalized dot product between tractogram vectors was used as a similarity measure, closely related to Pearson's correlation. The output of the algorithm is a hierarchical tree encoding not only the pair-wise connectivity similarity

information between datapoints but also and the similarity between areas of different granularities and their hierarchical relationships with one another.

Several traditional hierarchical methods were implemented, and a new method especially suited to our problem was developed by combining the centroid linkage method (Jain and Dubes, 1988) with neighborhood restriction (in order to produce contiguous parcels) and with an initial size-restricted stage (in order to obtain an early set of homogenous-sized clusters to use as tree meta-leaves). Cophenetic correlation coefficient (*CPCC*; Farris, 1969) was used to test the fit to the data of the generated trees using the aforementioned methods over real and random datasets. The centroid method with a 26-voxel neighborhood proved to be superior in performance by having a fit to the data equal to the best scoring traditional method but requiring much less computation and memory resources. Values obtained were also well above baseline levels obtained from the random datasets. This algorithm was adopted for the remainder of the study.

In order to make tree interpretation and partition extraction easier, a tree-processing pipeline was developed and implemented in order to reduce tree complexity while incurring in minimal information loss. Pre-processing steps included outlier pruning, monotonicity correction, limiting maximum granularity and detection of non-binary structures and consequent tree de-binarization. The combined effect of these steps achieved a complexity reduction of more than 90% with a loss of information of less than 0.5%, making it a remarkably efficient and useful tool for improving the performance of partition finding and tree comparison algorithms.

The information encoded in the cleaned trees was used as a whole in order to detect structural differences between datasets. For this purpose, a leaf matching strategy was set in place by warping the mean-tractograms of each subject's meta-leaves to a common space (guided by a prior FA registration) and applying a greedy matching algorithm over the pair-wise similarity matrix of the meta-leaves tractograms (with a maximum anatomical distance restriction in order to avoid poor matches at the last stages of the greedy algorithm). This way each meta-leaf from the first subject is matched to the meta-leaf of the second subject with most similar connectivity fingerprint. Once tree meta-leaves were matched, tree comparison algorithms were applied (*tCPCC*, based on the correlation of the similarity matrices encoded by the trees, and *wTriples*, based on the triples method: Bansal et al., 2011). A baseline level for the tree comparison values was obtained by applying the algorithms over randomly matched trees (but subject to the same maximum anatomical distance restrictions). Results on real datasets showed values well above baseline levels and good test-retest reliability. *tCPCC* performed better against noise. Inter-subject comparison however, proved not

robust enough for quantitative comparisons at this stage, maybe due to challenges in leaf-matching (same-subject comparisons features much better leaf-matching quality compared to between-hemispheres comparisons, which in turn match better than between-subject comparisons).

The information in the hierarchical tree can be mapped back to cortex by means of projecting full partitions (that is, cuts through the hierarchical tree) onto the individual cortical surface. The most widely used tree partition selection method is the horizontal cut, which guarantees, for a given number of clusters, a lower bound for the intra-cluster spread. This method however, does not take into account inter-cluster distance. We implemented a new partition method using both intra-cluster spread and inter-cluster separation (*SS*). Results on healthy participants showed that for a given granularity level (defined by number of clusters), partitions obtained from the horizontal cut are a good approximation to those obtained by the *SS* method. Partitions at low granularities (~15 clusters per hemisphere) reflect the rough course of major fiber bundles. Increasing the granularity to 50 clusters provides cluster area sizes similar to Brodmann areas. Finer subdivisions could be achieved increasing granularity to 100 clusters. Focusing on the IFG shows partitioning consistent with known literature: while at relatively low granularities only some of the major boundaries between the opercular, triangular and orbital parts of the IFG were revealed, at higher granularity, more subdivisions appeared, including those that are not covered by the classical tripartition. For repetitive acquisitions in the same subject, the subdivisions were highly reproducible. The spread-separation partitioning method also allows to evaluate the quality of the partition (defined by a higher separation to spread ratio index) and find for each defined granularity the one maximizing this value. Plotting this index value of each optimized partition against its number of clusters can help establish a range of higher quality partitions. This range was found to be between 20 and 600 clusters per hemisphere on our high resolution datasets.

In order to provide a higher degree of confidence in the methodological framework developed, two small validation studies were carried out.

The first study looked to reproduce the results of Ruschel and colleagues (2013) where they divided the inferior parietal cortex convexity (IPCC) of 20 healthy participants, finding 3 clusters to be the number that best stabilized results across subjects. Applying our methods on their tractography data and using *SS* partitioning to obtain 3 clusters provided virtually identical partitions for more than half the datasets. The remaining datasets showed slight differences in one of

the boundaries, but completely overlapping boundaries could be obtained if granularity was increased to 5 clusters.

In a second validation test the partitions obtained in an individual dataset were matched to the cytoarchitectonic parcellation provided by *Jülich Research Center*. The boundaries obtained by our methods at mid-level granularities appeared in agreement with the cytoarchitectonic atlas results.

This study aims at proposing a novel technology for parcellating the brain and offering initial proof-of-principle validation. Obviously, much remains to be done.

There are a number of methodological issues that require further attention. These especially involve the partition extraction method and the tree comparison technique (especially the leaf matching). An alternative way to match leaves could be based on surface registration (Moreno-Dominguez et al., 2014b). A mixed approach between both also sounds a promising avenue.

Also, in order to produce partitions representative of a population, a larger study with a greater pool of subjects is essential. Such a study should also include a more thorough verification and validation. Some ideas in that respect could involve software phantoms, systematic parameter optimization, and comparison to functional data parcellations.

Using a more numerous and representative cohort of brains, we believe that this technology can be used to build a hierarchical function-anatomical atlas or a hierarchical connectome of the brain. A whole-brain parcellation derived from anatomical connectivity would serve as an ideal starting point to characterize structural and functional connections between anatomically meaningful parcels. Furthermore, it would be a very interesting option to study the connectome at different granularities with the hierarchical partitions provided, and produce a hierarchical connectome. Here, the issue of neurobiological validation requires substantial attention. For example, it has to be investigated to what extent features that are not easily captured by agglomerative trees, such as gradation or non-nested hierarchies, are present in the brain and how our method reacts to them.

Prospectively, the proposed individual parcellation can also be used as *a priori* anatomical knowledge in the localization of functional data such as fMRI or EEG/MEG. This uses the assumption of local functional homogeneity of cortical activations within the parcel and could potentially increase the precision and statistical power of the localization, compared to the currently used approaches resulting in a spatially smoothed activation without respecting function-anatomical boundaries.

Although we have conceived and used our methods for the analysis of diffusion based anatomical connectivity, they should also be useful for the study of other kinds of multidimensional data, like resting-state functional connectivity. Whole brain parcellation methods have already been successfully used for the study of resting-state fMRI signals (Blumensath et al., 2013) and our framework might also bring new insights and possibilities to these approaches.

Publications of the author arising from this work

Journal Papers

- Moreno-Dominguez D, Anwander A, Knösche TR (2014). A hierarchical method for whole-brain connectivity-based parcellation. *Human Brain Mapping* 35, 5000-5025. doi: 10.1002/hbm.22528.

Conference Abstracts

- Moreno-Dominguez D, Anwander A, Knösche TR (2013). Multi-subject comparison of whole-brain connectivity-based hierarchical parcellations. *Proceedings of the 19th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2013)*.

- Moreno-Dominguez D, Anwander A, Knösche TR (2012). Matching and comparison of hierarchical trees for whole-brain connectivity-based cortex parcellation. *Proceedings of the 18th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2012)*.

- Moreno-Dominguez D, Anwander A, Knösche TR (2012). Dendrogram processing for whole-brain connectivity-based hierarchical parcellation. *Proceedings of the 20th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM 2012)* 688.

- Knösche TR, Moreno-Dominguez D, Anwander A (2012). Human connectome: Hierarchical clustering. *Proceedings of 29th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB 2012)*.

- Anwander A, Moreno-Dominguez D, Knösche, TR (2012). Diffusion MRI based parcellation as prior knowledge in EEG/MEG source reconstruction. *18th International Conference on Biomagnetism*.

- Moreno-Dominguez D, Anwander A, Knösche TR (2011). Evaluation of hierarchical methods for connectivity-based whole-brain parcellation. *Proceedings of 28th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB 2011)* e-Poster 492. Awarded Certificate of Merit.

- Moreno-Dominguez D, Anwander A, Schurade R, Knösche TR (2011). Connectivity-based whole-brain hierarchical parcellation of the human brain. *Proceedings of the 17th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2011)*.

- Moreno-Dominguez D, Anwander A, Schurade R, Knösche TR (2011). Connectivity-based whole-brain hierarchical parcellation of the human brain. *CONNECT MEETING: MRI of Brain Micro-structure and Connectivity*.

References

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings H, Zilles K (1999). Broca's region revisited: cytoarchitecture and intersubject variability. *Journal of Comparative Neurology* 412, 319-341.

Anderson A, Ding Z (2002). Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging. *Proc. 10th Annual Meeting of the ISMRM Honolulu, 2002*, Berkeley, USA: ISMRM, p. 440.

Anwander A, Tittgemeyer M, von Cramon DY, Friederici AD, Knösche TR (2007). Connectivity-based parcellation of Broca's area. *Cerebral Cortex* 17, 816-825.

Anwander A, Moreno-Dominguez D, Knösche, TR (2012). Diffusion MRI based parcellation as prior knowledge in EEG/MEG source reconstruction. *18th International Conference on Biomagnetism*.

Avants B, Epstein C, Grossman M, Gee J (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* 12, 16-41.

Axer H, Berks G, Keyserlingk DGV (2000). Visualization of nerve fiber orientation in gross histological sections of the human brain. *Microscopy Research and Technique* 51, 481-492.

Axer H, Leunert M, Mürköster M, Gräßel D, Larsen L, Griffin LD, Keyserlingk DGV (2002). A 3D fiber model of the human brainstem. *Computerized medical imaging and graphics* 26, 439-444.

Axer H, Axer M, Grasharpel D, Witte OW (2008). Imaging of Wallerian degeneration in the brain. *Current Medical Imaging Reviews* 4, 71-76.

Bach D, Behren ST, Garrido L, Weiskopf N, Dolan R (2011). Deep and superficial amygdala nuclei projections revealed in vivo by probabilistic tractography. *The Journal of Neuroscience* 31, 618-623.

Bakker R, Wachtler T, Diesmann M (2012). CoCoMac 2.0 and the future of tract-tracing databases. *Frontiers in Neuroinformatics* 6.

Bansal MS, Dong J, Fernandez-Baca D (2011). Comparing and aggregating partially resolved trees. *Theoretical Computer Science* 412, 6634-6652.

Basser PJ, Mattiello J, LeBihan D (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance Series B* 103, 247-254.

Basser PJ, Pierpaoli C (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance Series B* 111, 209-219.

Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A (2000). In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine* 44, 625-632.

Bassett DS, Greenfield DL, Meyer-Lindenberg A, Weinberger DR, Moore SW, Bullmore ET (2010). Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Computational Biology* 6, e1000748. doi: 10.1371/journal.pcbi. 1000748.

Bassett DS, Gazzaniga MS (2011). Understanding complexity in the human brain. *Trends in Cognitive Sciences* 15, 200–209.

Bastiani M, Shah NJ, Goebel R, Roebroeck A (2012). Human cortical connectome reconstruction from diffusion weighted MRI: The effect of tractography algorithm. *Neuroimage* 62, 1732-1749.

Bazin PL, Weiss M, Dinse J, Schäfer A, Trampel R, Turner R (2013). A computational framework for ultra-high resolution cortical segmentation at 7 Tesla. *Neuroimage* doi: 10.1016/j.neuroimage.2013.03.077.

Behrens TEJ, Johansen-Berg HJ, Woolrich MW, Smith, SM, Wheeler-Kingshott CAM, Boulby PA, Barker GJ, Sillery EL, Sheehan K, Ciccarelli O, Thompson AJ, Brady M, Matthews PM (2003a). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* 6, 750-757.

Behrens TEJ, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, Matthews PM, Brady JM, Smith SM (2003b). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine* 50, 1077-1088.

Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34, 144-155.

Beyer K, Goldstein J, Ramakrishnan R, Shaft U (1999). When is "nearest neighbor" meaningful? *Database Theory ICDT'99*. Springer Berlin Heidelberg 217-235.

Blumensath T, Jbabdi S, Glasser MF, Van Essen DC, Ugurbil K, Behrens TE, Smith SM (2013). Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. *Neuroimage* 76, 313–324.

Blundell C, Teh YW, Heller KA (2010). Bayesian rose trees. *Proceedings of 26th Conference of Uncertainty in Artificial Intelligence*, UAI 2010, 65-72.

Broca P (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin de la Société Anatomique* 6, 330-357.

Brodal P (1978). The corticopontine projection in the rhesus monkey Origin and principles of organization. *Brain* 101, 251-283.

Brodmann K (1909). Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellaufbaues. Barth, Leipzig.

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

Cachia A, Mangin JF, Riviere D, Kherif F, Boddaert N, Andrade A, Papadopoulos-Orfanos D, Poline JB, Bloch I, Zilbovicus M, Sonigo P, Brunelle F, Regis J (2003). A primal sketch of the cortex mean curvature: a morphogenesis based approach to study the variability of the folding patterns. *IEEE Transanctions in Medical Imaging* 22, 754-765.

Campbell AW (1904) Histological studies on the localisation of cerebral function. *Journal of Mental Science* 50, 651–662.

Cathier P, Mangin JF (2006). Registration of cortical connectivity matrices. *IEEE International Conference in Computer Vision and Pattern Recognition Workshop CVPRW'06*, 66-73.

Caspers S, Eickhoff SB, Geyer S, Scheperjans F, Mohlberg H, Zilles K, Amunts K (2008). The human inferior parietal lobule in stereotaxic space. *Brain Structure and Function* 212, 481-495.

Cerliani L, Thomas RM, Jbabdi S, Siero JCW, Nanetti L, Crippa A, Gazzola V, D'Arceuil H, Keysers C (2012). Probabilistic tractography recovers a rostrocaudal trajectory of connectivity variability in the human insular cortex. *Human Brain Mapping* 33, 2005-2034.

Clos M, Amunts K, Laird AR, Fox PT, Eickhoff SB (2013). Tackling the multifunctional nature of Broca's region meta-analytically: Co-activation-based parcellation of area 44. *Neuroimage* 83, 174188.

Cordes D, Haughton V, Carew J, Arfanakis K, Maravilla K (2002). Hierarchical clustering to measure the connectivity in fMRI resting-state data. *Magnetic Resonance Imaging* 20, 305-317.

Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping* 33, 1914-1928.

Critchlow DE, Pearl DK, Qian CL (1996). The triples distance for rooted bifurcating phylogenetic trees. *Systematic Biology* 45, 323-334.

Descoteaux M, Deriche R, Knösche TR, Anwander A (2009). Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transanctions in Medical Imaging* 28, 269-286.

Descoteaux M (2010). High angular resolution diffusion MRI: from local estimation to segmentation and tractography. Doctoral dissertation, *Max Planck Institute for Human Cognitive and Brain Sciences*, Germany.

Dhanarajan P, Rüegg DG, Wiesendanger M (1977). An anatomical investigation of the corticopontine projection in the primate Saimiri sciureus. The projection from motor and somatosensory areas. *Neuroscience* 2, 913-922.

Dosenbach NUF, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RAT, Fox MD, Snyder AZ, Vincent JL, Raichle ME, Schlaggar BL, Petersen, SE (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences* 104, 11073-11078.

Einstein A (1956). Investigations on the Theory of the Brownian Movement. *Dover Publications Inc.*

Exner S (1881) Untersuchungen über die Localisation der Functionen in der Großhirnrinde des Menschen. Wien, Braumüller.

Farris JS (1969). On the cophenetic correlation coefficient. *Systematic Biology* 18, 279-285.

Fieremans E, Delputte S, Deblaere K, De Deene Y, Truyens B, D'Asseler Y, Achten E, Lemahieu I, Van de Walle R (2005). A flexible hardware phantom for validation of diffusion imaging sequences. *Proceeding of the International Society of Magnetic Resonance Medicine (ISMRM)* 13.

Fillard P, Poupon C, Mangin JF (2009). A novel global tractography algorithm based on an adaptive spin glass model. *Proceedings of Medical Image Computing and Computer-Assisted Intervention 2009* 927-934.

Florek K, Lukaszewicz J, Steinhaus H, Zubrzycki S (1951). Sur la liaison et la division des points d'un ensemble fini. *Colloquium Mathematicum* 2, 282-285.

Geyer S, Matelli M, Luppino G, Zilles K (2000a). Functional neuroanatomy of the primate isocortical motor system. *Anatomy and Embryology* 202, 443–474.

Geyer S, Zilles K, Luppino G, Matelli M (2000b). Neurofilament protein distribution in the macaque monkey dorsolateral premotor cortex. *European The Journal of Neuroscience* 12, 1554-1566.

Gorbach NS, Schütte C, Melzer C, Goldau M, Sujazow O, Jitsev J, Douglas T, Tittgemeyer M (2011). Hierarchical information-based clustering for connectivity-based cortex parcellation. *Frontiers in Neuroinformatics* 5, 18.

Gorbach NS, Siep S, Jitsev J, Melzer C, Tittgemeyer M (2012). Information-theoretic connectivity-based cortex parcellation. *Machine Learning and Interpretation in Neuroimaging*, NIPS. Springer Berlin Heidelberg 186-193.

Gower JC (1990). Clustering Axioms. *Classification Society of North America Newsletter* 15, 2-3.

Gray H (1918). Anatomy of the Human Body. *Lea and Febiger*.

Gregoriou GG, Borra E, Matelli M, Luppino G (2006). Architectonic organization of the inferior parietal convexity of the macaque monkey. *Journal of Computational Neurology* 496, 422–451.

Guevara P, Poupon C, Rivière D, Cointepas Y, Descoteaux M, Thirion B, Mangin JF (2011). Robust clustering of massive tractography datasets. *Neuroimage* 54, 1975-1993.

Haber S (1988). Tracing intrinsic fiber connections in postmortem human brain with WGA-HRP. *Journal of Neuroscience Methods* 23, 15-22.

Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology* 6, e159.

Halkidi M, Batistakis Y, Vazirgiannis M (2002). Clustering Validity checking methods: part II. *ACM Sigmod Record* 31, 19-27.

Hartigan JA (1975). Clustering Algorithms. New York: *John Wiley & Sons*.

Hasnain MK, Fox PT, Woldorff, MG (2001). Structure-function spatial covariance in the human visual cortex. *Cerebral Cortex* 11, 702-716.

Heidemann RM, Anwander A, Feiweier T, Knösche TR, Turner R (2012). *k*-space and *q*-space: combining ultra-high spatial and angular resolution in diffusion imaging using ZOOPPA at 7T. *Neuroimage* 60, 967-978.

Hornak JP (1996). The Basics of MRI. A hypertext book on magnetic resonance imaging. `www.cis.rit.edu/htbooks/mri`. *Interactive Learning Software*, Henietta, NY.

Horton JC, Adams DL (2005). The cortical column: A structure without a function. *Philosophical Transactions of the Royal Society B* 360, 837–62.

Hosey T, Williams G, Ansorge R (2005). Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine* 54, 1480-1489.

Hudspeth AJ, Ruark JE, Kelly JP (1976). Cytoarchitectonic mapping by microdensitometry. *Proceedings of the National Academy of Sciences* 73, 2928-2931.

Hubel DH, Wiesel TN (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* 195, 215–243.

IIturria-Medina Y, Fernández AP, Hernández PV, Pentón LG, Canales-Rodríguez EJ, Melie-Garcia L, Castellanos AL, Ortega MO (2011). Automated discrimination of brain pathological state attending to complex structural brain network properties: the shiverer mutant mouse case. *PloS one*, 6.

Jain A, Dubes R (1988). Algorithms for Clustering Data. *Prentice Hall*, Upper Saddle River, USA.

Jbabdi S, Woolrich MW, Andersson JLR, Behrens TEJ (2007). A Bayesian framework for global tractography. *Neuroimage* 37, 116-129.

Jbabdi S, Woolrich MW, Behrens TEJ (2009). Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models. *Neuroimage* 44, 373-384.

Jenkinson M, Bannister P, Brady M, Smith S (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825-841.

Jenkinson M (2004). Improving the registration of B0-distorted EPI images using calculated cost function weights. *10th International Conference on Functional Mapping of the Human Brain*.

Jeurissen B, Leemans A, Jones DK, Tournier JD, Sijbers J (2011). Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. *Human Brain Mapping* 32, 461-479.

Jezzard P, Balaban RS (1995). Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic Resonance in Medicine* 34, 65-73.

Johansen-Berg H, Behrens TE, Robson MD, Drobnjak I, Rushworth MF, Brady JM, Smith SM, Higham DJ, Matthews PM (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proceedings of the National Academy of Sciences USA 101, 13335-13340. Copyright (2004) National Academy of Sciences, U.S.A.

Johansen-Berg H, Behrens TEJ, Sillery E, Ciccarelli O, Thompson AJ, Smith SM, Matthews PM (2005). Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. *Cerebral Cortex* 15, 31–39.

Johansen-Berg H, Behrens, TE (2009). Diffusion MRI: From quantitative measurement to in-vivo neuroanatomy. *Academic Press*.

Johnson S (1967). Hierarchical clustering schemes. Psychometrik, 32, 241-254.

Jones DK (2004). The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. Magnetic Resonance in Medicine 51, 807-815.

Jones DK (2010). Challenges and limitations of quantifying brain connectivity in vivo with diffusion MRI. Imaging in Medicine 2, 341–355.

Jones DK (2011). Diffusion MRI: Theory, methods, and applications. *Oxford University Press*.

Kaden E, Knösche TR, Anwander A (2007). Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *Neuroimage* 37, 474-488.

Kahnt T, Chang LJ, Park SQ, Heinzle J, Haynes JD (2012). Connectivity-based parcellation of the human orbitofrontal cortex. *The Journal of Neuroscience* 32, 6240-6250.

Kelly C, Toro R, Di Martino A, Cox CL, Bellec P, Castellanos FX, Milham MP (2012). A convergent functional architecture of the insula emerges across imaging modalities. *Neuroimage* 61, 1129-1142.

Keim DA, Hinneburg A (1999). Optimal grid-clustering: Towards breaking the curse of dimensionality in high-dimensional clustering. *VLDB* 99, 506-517.

Klein JC, Behrens TE, Robson MD, Mackay CE, Higham DJ, Johansen-Berg H (2007). Connectivity-based parcellation of human cortex using diffusion MRI: establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. *Neuroimage* 34, 204-211.

Klein A, Andersson J, Ardekani B, Ashburner J, Avants B, Chiang M, Christensen G, Collins D, Gee J, Hellier P, Song J, Jenkinson M, Lepage C, Rueckert D, Thompson P, Vercauteren T, Woods R, Mann J, Parsey R (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46, 786-802.

Klingler J (1935). Erleichterung der makrokopischen Präparation des Gehirns durch den Gefrierprozess. *Orell Füssli*.

Klüver H, Barrera E (1953). A Method for the Combined Staining of Cells and Fibers in the Nervous System. *Journal of Neuropathology and Experimental Neurology* 12, 400-403.

Knösche TR, Tittgemeyer M (2011). The role of long-range connectivity for the characterization of the functional-anatomical organization of the cortex. *Frontiers in Systems Neuroscience* 5, 58.

Knösche TR, Moreno-Dominguez D, Anwander A (2012). Human connectome: Hierarchical clustering. *Proceedings of 29th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB 2012)*.

Koch M (2000). Measurement of the Self-Diffusion Tensor of Water in the Human Brain. Doctoral dissertation, *Max Planck Institute for Human Cognitive and Brain Sciences*, Germany. *MPI Series in Cognitive Neuroscience* 14.

Koch M, Norris DG, Hund-Georgiadis M (2002). An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage* 16, 241--250.

Kreher BW, Mader I, Kiselev VG (2008). Gibbs tracking: A novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine* 60, 953-963.

Kristo G, Leemans A, Raemaekers M, Rutten GJ, Gelder B, Ramsey NF (2013). Reliability of two clinically relevant fiber pathways reconstructed with constrained spherical deconvolution. *Magnetic Resonance in Medicine* 70, 1544-1556.

Kuhn HW (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly* 2, 83-97.

Kuncheva LI, Vetrov DP (2006). Evaluation of stability of k-means cluster ensembles with respect to random initialization. *IEEE Transactions in Pattern Analysis and Machine Intelligence* 28, 1798-1808.

Landman BA, Bogovic JA, Carass A, Chen M, Roy S, Shiee N, Yang Z, Kishore B, Pham D, Bazin PL, Resnick SM, Prince JL (2013). System for Integrated Neuroimaging Analysis and Processing of Structure. *Neuroinformatics* 11, 91-103.

Langfelder P, Zhang B, Horvath S (2008). Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for R. *Bioinformatics* 24, 719-720.

Larsen L, Griffin LD, Gräßel D, Witte OW, Axer H (2007). Polarized light imaging of white matter architecture. *Microscopy Research and Technique* 70, 851-863.

Latora V, Marchiori M (2001). Efficient behavior of small-world networks. *Physical Review Letters* 87, 198701.

Latora V, Marchiori M (2003). Economic small-world behavior in weighted networks. *The European Physical Journal B-Condensed Matter and Complex Systems* 32, 249-263.

Lazar M, Alexander AL (2003). An error analysis of white matter tractography methods: synthetic diffusion tensor field simulations. *Neuroimage* 20, 1140–1153.

Le Bihan D, Breton E (1985). Imagerie de diffusion in-vivo par résonance. *Comptes Rendus de l'Académie des Sciences* (Paris) 301, 1109–1112.

Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161, 401-407.

Lenglet C, Campbell JSW, Descoteaux M, Haro G, Savadjiev P, Wassermann D, Anwander A, Deriche R, Pike GB, Sapiro G, Siddiqi K, Thompson PM (2009). Mathematical methods for diffusion MRI processing. *Neuroimage* 45, 111-122.

Li L, Rilling JK, Preuss TM, Glasser MF, Hu X (2012). The effects of connection reconstruction method on the interregional connectivity of brain networks via diffusion tractography. *Human Brain Mapping* 33, 1894-1913.

Lohmann G, Mueller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001). Lipsia - A new software system for the evaluation of functional magnetic resonance images of the human brain. *Computerized Medical Imaging and Graphics* 25, 449-457.

Lori NF, Akbudak E, Shimony JS, Cull TS, Snyder AZ, Guillory RK, Conturo TE (2002) Diffusion tensor fiber tracking of human brain connectivity: acquisition methods, reliability analysis and biological results. *NMR in Biomedicine* 15, 493–515.

Lukas JR, Aigner M, Denk M, Heinzl H, Burian M, Mayr R (1998). Carbocyanine postmortem neuronal tracing: Influence of different parameters on tracing distance and combination with immunocytochemistry. *Journal of Histochemistry and Cytochemistry* 46, 901-910.

Luppino G (2005). 12 Organization of the Posterior Parietal Lobe and of Parietofrental Connections. *From Monkey Brain to Human Brain: A Fyssen Foundation Symposium* p. 235. MIT Press.

Malcolm JG, Shenton ME, Rathi Y (2010). Filtered multitensor tractography. *IEEE Transactions in Medical Imaging* 29, 1664-1675.

Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Croxson PL, Olivier E, Noonan MP, Bergmann C, Mitchell AS, Baxter MG, Behrens TEJ, Johansen-Berg H, Tomassini V, Miller KL, Rushworth MF (2011). Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *The Journal of Neuroscience* 31, 4087-4100.

McQuitty L (1957). Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies. *Educational and Psychological Measurement* 17, 207-222.

Mesulam M (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology* 28, 597-613.

Mesulam M (1998). From sensation to cognition. *Brain* 121, 1013-1052.

Moreno-Dominguez D, Anwander A, Knösche TR (2014a). A hierarchical method for whole-brain connectivity-based parcellation. *Human Brain Mapping* 35, 5000-5025 doi:10.1002/hbm.22528.

Moreno-Dominguez D, Watanabe A, Gorgolewski KJ, Schaefer A, Goulas A, Kipping J, Kanaan A, Anwander A, Toro R, Margulies DS (2014b). Multimodal Parcellation of the Frontal Lobe. *Proceedings of the 20th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2014)*.

Moreno-Dominguez D, Anwander A, Knösche TR (2013). Multi-subject comparison of whole-brain connectivity-based hierarchical parcellations. *Proceedings of the 19th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2013)*.

Moreno-Dominguez D, Anwander A, Knösche TR (2012a). Matching and comparison of hierarchical trees for whole-brain connectivity-based cortex parcellation. *Proceedings of the 18th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2012)*.

Moreno-Dominguez D, Anwander A, Knösche TR (2012b). Dendrogram processing for whole-brain connectivity-based hierarchical parcellation. *Proceedings of the 20th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM 2012*) p. 688.

Moreno-Dominguez D, Anwander A, Knösche TR (2011a). Evaluation of hierarchical methods for connectivity-based whole-brain parcellation. *Proceedings of 28th Annual Scientific Meeting of European Society for Magnetic Resonance in Medicine and Biology (ESMRMB 2011)* e-Poster 492.

Moreno-Dominguez D, Anwander A, Schurade R, Knösche TR (2011b). Connectivity-based whole-brain hierarchical parcellation of the human brain. *Proceedings of the 17th Annual Meeting of the Organization for Human Brain Mapping (OHBM 2011)*.

Moreno-Dominguez D, Anwander A, Schurade R, Knösche TR (2011c). Connectivity-based whole-brain hierarchical parcellation of the human brain. *CONNECT MEETING: MRI of Brain Micro-structure and Connectivity*.

Morgan BJ, Ray AP (1995). Non-uniqueness and inversions in cluster analysis. *Applied Statistics* 117-134.

Mori S, Crain BJ, Chacko VP, Van Zijl P (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* 45, 265-269.

Mori S, Zhang J (2006). Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron* 51, 527-539.

Mori S (2007). Introduction to diffusion tensor imaging. *Elsevier*.

Mountcastle VB (1957). Modality and topographic properties of single neurons of cats somatic sensory cortex. *Journal of Neurophysioly* 20, 408–34.

Murtagh F (1983). A survey of recent advances in hierarchical-clustering algorithms. *The Computer Journal* 26, 354-359.

Murtagh F (1985). Multidimensional Clustering Algorithms. *Vienna: Physica*.

Nanetti L, Cerliani L, Gazzola V, Renken R, Keysers C (2009). Group analyses of connectivity-based cortical parcellation using repeated k-means clustering. *Neuroimage* 47, 1666-1677.

Ono M, Kubick S, Albernathey CD (1990). Atlas of the cerebral sulci. *Thieme*, New York.

Pajevic S, Aldroubi A, Basser PJ (2002). A continuous tensor field approximation of discrete DT-MRI data for extracting microstructural and architectural features of tissue. *Journal of Magnetic Resonance* 154, 85–100.

Papadakis NG, Xing D, Houston GC, Smith JM, Smith MI, James MF, Parsons AA, Huang CLH, Hall LD, Carpenter TA (1999). A study of rotationally invariant and symmetric indices of diffusion anisotropy. *Magnetic Resonance Imaging* 17, 881-892.

Parker GJ, Alexander DC (2005). Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transaction of the Royal Society B* 360, 893-902.

Passingham RE, Stephan KE, Kötter R (2002). The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience* 3, 606-616.

Passingham RE (2007). Commentary on Devlin and Poldrack. *Neuroimage* 37, 1055–1056; discussion 1066–1068.

Perrin M, Poupon C, Cointepas Y, Rieul B, Golestani N, Pallier C, Rivière D, Constantinesco A, LeBihan D, Mangin JF (2005). Fiber tracking in q-ball fields using regularized particle trajectories. *Lecture Notes in Computer Science*, Springer Berlin/ Heidelberg 3565:52–63.

Perrin M, Cointepas Y, Cachia A, Poupon C, Thirion B, Riviere D, Cathier P, Kouby VE, Constantinesco A, Le Bihan D, Mangin JF (2008). Connectivity-based parcellation of the cortical mantle using q-ball diffusion imaging. *Journal of Biomedical Imaging* 4.

Petrovic A, Zollei L, 2011. Evaluating volumetric brain registration performance using structural connectivity information. *Medical Image Computing and Computer-Assisted Intervention 2011* 524-531.

Pham DT, Dimov SS, Nguyen CD (2005). Selection of K in K-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 219, 103-119.

Pierpaoli C, Barnett A, Pajevic S, Chen R, Penix L, Virta A, Basser P (2001). Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *Neuroimage* 13, 1174–1185.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992). Numerical recipes in C: the art of scientific computing. *Cambridge University Press.* pp 994

Ramón y Cajal S (1928). Degeneration and regeneration of the nervous system (Vol. 1). *Oxford University Press*, Humphrey Milford.

Rand W (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association* 66, 846-850.

Reisert M, Mader I, Anastasopoulos C, Weigel M, Schnell S, Kiselev V (2011). Global fiber reconstruction becomes practical. *Neuroimage* 54, 955-962.

Restrepo G, Mesa H, Llanos EJ (2007). Three dissimilarity measures to contrast dendrograms. *Journal of Chemical Information and Modeling* 47, 761-770.

Riffert T, Schreiber J, Anwander A, Knösche TR (2014). Beyond fractional anisotropy: Extraction of bundle-specific structural metrics from crossing-fiber models. (*under review*).

Roca P, Rivière D, Guevara P, Poupon C, Mangin J (2009). Tractography-based parcellation of the cortex using a spatially-informed dimension reduction of the connectivity matrix. *Medical Image Computing and Computer-Assisted Intervention 2009* 935-942.

Roca P, Tucholka A, Rivière D, Guevara P, Poupon C, Mangin JF (2010). Inter-subject connectivity-based parcellation of a patch of cerebral cortex. *Medical Image Computing and Computer-Assisted Intervention 2010* 347-354.

Rozzi S, Calzavara R, Belmalih A, Borra E, Gregoriou GG, Matelli M, Luppino G (2006). Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cerebral Cortex* 16, 1389–1417.

Rozzi S, Ferrari PF, Bonini L, Rizzolatti G, Fogassi L. (2008). Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *European Journal of Neuroscience* 28, 1569–1588.

Ruschel M (2013). Konnektivitätsbasierte parzellierung des humanen inferioren parietalkortex – eine experimentelle DTI-analyse. Doctoral dissertation, *Max Planck Institute for Human Cognitive and Brain Sciences*, Germany.

Ruschel M, Knösche TR, Friederici A, Turner R, Geyer S, Anwander A (2013). Connectivity architecture and subdivision of the human inferior parietal cortex revealed by diffusion MRI. *Cerebral Cortex*. doi: 10.1093/cercor/bht098.

Rushworth MFS, Behrens TEJ, Johansen-Berg H (2006). Connection patterns distinguish 3 regions of human parietal cortex. *Cerebral Cortex* 16, 1418–1430.

Schmitt O, Pakura M, Aach T, Hömke L, Böhme M, Bock S, Preusse S (2004). Analysis of nerve fibers and their distribution in histologic sections of the human brain. *Microscopy Research and Technique* 63, 220-243.

Schreiber J, Riffert T, Anwander A, Knösche TR (2014). Plausibility Tracking: A method to evaluate anatomical connectivity and microstructural properties along fiber pathways. *Neuroimage* 90C, 163-178.

Schubotz RI, Anwander A, Knösche TR, von Cramon DY, Tittgemeyer M (2010). Anatomical and functional parcellation of the human lateral premotor cortex. *Neuroimage* 50, 396-408.

Sneath P (1957). The applications of computers to taxonomy. *Journal of General Microbiology* 17, 201-226.

Sporns O (2007). *Scholarpedia*, 2, 4695.

Sporns O (2011a). Networks of the Brain. *MIT press*.

Sporns O (2011b). The human connectome: A complex network. *Annals of the New York Academy of Sciences* 1224, 109-125.

Stanberry L, Nandy R, Cordes D (2003). Cluster analysis of fMRI data using dendrogram sharpening. *Human Brain Mapping* 20, 201-219.

Stieltjes B, Kaufmann WE, Van Zijl PCM, Fredericksen K, Pearlson GD, Solaiyappan M, Mori S (2001). Diffusion tensor imaging and axonal tracking in the human brainstem. *NeuroImage* 14, 723–735.

Stejskal EO, Tanner JE (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics* 42, 288.

Tahmasebi AM, Davis MH, Wild CJ, Rodd JM, Hakyemez H, Abolmaesumi P, Johnsrude IS (2012). Is the link between anatomical structure and function equally strong at all cognitive levels of processing? *Cerebral Cortex* 22, 1593-1603.

Taylor DG, Bushell MC (1985). The spatial mapping of translational diffusion coefficients by the NMR imaging technique. *Physics in Medicine and Biology* 30, 345–349.

Theodoridis S, Koutroumbas K (1999). Pattern Recognition. *Academic Press*, New York.

Thompson P, Schwartz C, Lin R, Khan A, Toga A (1996). Three-dimensional statistical analysis of sulcal variability in the human brain. *The Journal of Neuroscience* 16, 4261-4274.

Tomassini V, Jbabdi S, Klein JC, Behrens TEJ, Pozzilli C, Matthews PM, Rushworth MFS, Johansen-Berg H (2007). Diffusion-weighted imaging tractography- based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *The Journal of Neuroscience* 27, 10259–10269.

Tournier J, Calamante F, Gadian DG, Connelly A (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176-1185.

Tournier JD, Calamante F, Connelly A (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35, 1459-1472.

Tournier JD, Calamante F, Connelly A (2012). MRtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology* 22, 53-66.

Tuch DS (2002). Diffusion MRI of complex tissue structure. Doctoral dissertation, *Massachusetts Institute of Technology, USA*.

Tuch DS, Reese TG, Wiegell MR, Wedeen VJ (2003). Diffusion MRI of complex neural architecture. *Neuron* 40, 885-895.

Tuch DS (2004). Q-Ball Imaging. *Magnetic Resonance in Medicine* 52, 1358–1372.

Uchino A, Imada H, Ohno M (1990). MR imaging of wallerian degeneration in the human brain stem after ictus. *Neuroradiology* 32, 191-195.

Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience* 2, 229–239.

Vogt C, Vogt M (1919). Allgemeinere Ergebnisse unserer Hirnforschung. *Journal für Psychologie und Neurologie* 25, 277–462.

Von Economo C (1929). The Cytoarchitectonics of the Human Cerebral Cortex. *Oxford University Press* London, UK.

Waller A (1850). Experiments on the section of the glossopharyngeal and hypoglossal nerves of the frog, and observations of the alterations produced thereby in the structure of their primitive fibres. *Philosophical Transactions of the Royal Society of London* 140, 423-429.

Wang H, Wang W, Yang J, Yu P (2002). Clustering by pattern similarity in large data sets. *Proceedings of the ACM SIGMOD International Conference on Management of Data* 394-405.

Wassermann D, Bloy L, Kanterakis E, Verma R, Deriche R (2010). Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers. *Neuroimage* 51, 228.

Watanabe M, Aoki S, Masutani Y, Abe O, Hayashi N, Masumoto T, Mori H, Kabasawa H, Ohtomo K (2006). Flexible ex vivo phantoms for validation of diffusion tensor tractography on a clinical scanner. *Radiation Medicine* 24, 605–609.

Wedeen VJ, Reese TG, Tuch DS, Weigel MR, Dou JG, Weiskoff RM, Chessler D (2000). Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI. *Proceedings of the 8th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM 2000)*.

Wedeen VJ, Wang RP, Schmahmann JD, Benner T, Tseng, WYI, Dai G, Pandya DN, Hagmann P, D'Arceuil H, de Crespigny AJ (2008). Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage* 41, 1267-1277.

Weigert C (1897). Die Markscheidenfärbung. *Ergebnisse der Anatomie und Entwicklung Geschichte* 3, 1-23.

Wernicke C (1874). Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer *Basis*. Cohn & Weigert.

Wright SJ, Centonze VE, Stricker SA, DeVries PJ, Paddock SW, Schatten G (1993). Introduction to confocal microscopy and three-dimensional reconstruction. Methods in Cell Biology 38, 1-45.

Yo TS, Anwander A, Descoteaux M, Fillard P, Poupon C, Knösche TR (2009). Quantifying brain connectivity: A comparative tractography study. *Medical Image Computing and Computer-Assisted Intervention 2009* 12, 886-93

Zahn CT (1971). Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions on Computers* 20, 68-86.

Zilles K, Schlaug G, Geyer S, Luppino G, Matelli M, Qü M, Schleicher A, Schormann T (1996) Anatomy and transmitter receptors of the supplementary motor areas in the human and nonhuman primate brain. Advances in Neurology 70, 29–43.

Acknowledgements

Firstly I would like to thank Thomas Knösche for giving me the opportunity to carry out this work in his methods development unit at the Max Planck Institute in Leipzig and to Jens Haueisen for enabling to pursue with it a doctoral degree through his biomedical engineering group at the University of Ilmenau. Special thanks must go to Alfred Anwander for his co-supervision and guidance in the development of this work; to Jan Schreiber for a wealth of knowledge, fruitful discussions and advice; to Ralph Schurade for his programming tips and first implementation of the interactive visualization software; to Alexandros Goulas for his insights on clustering validation and to Daniel Margulies for kindly enabling me to finish my doctoral work while working in parallel on our current project. Valuable input was also received from many colleagues and peers in the *Max Planck Institute*.

I would also like to extend my gratitude to the *Max Planck Institute for Human Cognitive and Brain Sciences, Fundacion Caja Madrid*, *Fazit-Stiftung* and the *FET* project *CONNECT* from the EU, for their resources and financial support to my doctoral work and doctoral education. Without them this research would not have been possible.

Finally, I would like to thank my family and friends for their understanding and unyielding moral support.

List of figures

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise untentgeltlich geholfen:

1. Der traktografie Algorithmus und die dazu gehörende vorverarbeitungskette zur Berechnung der Konnektivitätsmuster und die MRT Daten von Proband A, B und C wurde von Alfred Anwander vom Max Planck Institut für Kognitions- und Neurowissenschaft bereitgestellt.
2. Eine initiale Implementierung des Moduls in OpenWalnut zur interaktiven Analyse der hierarchischen Struktur wurde von Ralph Schurade vom Max Planck Institut für Kognitions- und Neurowissenschaft implementiert. Dieses Modul wurde weiterentwickelt und verbessert um komplexere Strukturen zu verarbeiten und um neue Algorithmen zu integrieren.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungsbzw. Beratungsdiensten (Promotionsberater oder andere Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt.

Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 8 der Promotionsordnung den Abbruch des Promotionsverfahrens zur Folge hat.

Appendix

A.1 Comparison to *k*-means (Ruschel 2013)

Following, results for the circumstantial validation comparison between the parcellation of the IPCC done with *k*-means and with our hierarchical clustering and tree partition algorithms are shown (see chapter 5 for further details on the experiment). While in the main document only the results of the four subjects shown on the journal article by Ruschel and colleagues (2013) are presented, here the results obtained for all 19 subjects included in the doctoral dissertation (Ruschel 2013) are compared (original results include 20 subjects, but tractography data for one of them could not be recovered for replication with our method and it was left out of the comparison).

In his work Ruschel found that three partitions gave the most stable results across subjects when using *k*-means. Using our method we obtained trees from the same tractography data and using the *SS* method obtained partitions at three clusters and five clusters solutions (however, due to the non-binary nature of our processed trees, sometimes these specific numbers could not be selected, and the closest higher granularity was chosen).

As with the results obtained for the four subjects considered in chapter 5, it can be observed that for about half of the datasets the clusters obtained are virtually identical to those suggested by Ruschel in his work (i.e: subjects 8 and 10). Most of the remaining cases, while having small differences in the boundaries between the middle and posterior clusters or the middle and anterior clusters at the three clusters partition, yield completely overlapping boundaries when granularity is increased to five clusters (i.e: subjects 1 and 6).

Only in six cases out of thirty-eight there is no agreement found between the k-means and hierarchical solutions (in both hemispheres of subjects 16 and 17, right hemisphere of subject 18 and left hemisphere of subject 19). But in these cases, the hierarchical solutions seem more plausible than their k-means counterparts, which present disjoint clusters that are not continuous along the cortical surface. This is probably due to the neighborhood constraint present in our method, which produces continuous parcels along the surface, and seems to stabilize the solution. This is an interesting point to be considered for cortical parcellations using dMRI data.

k-means (Ruschel 2013). 3 clusters. Subjects 1 to 6
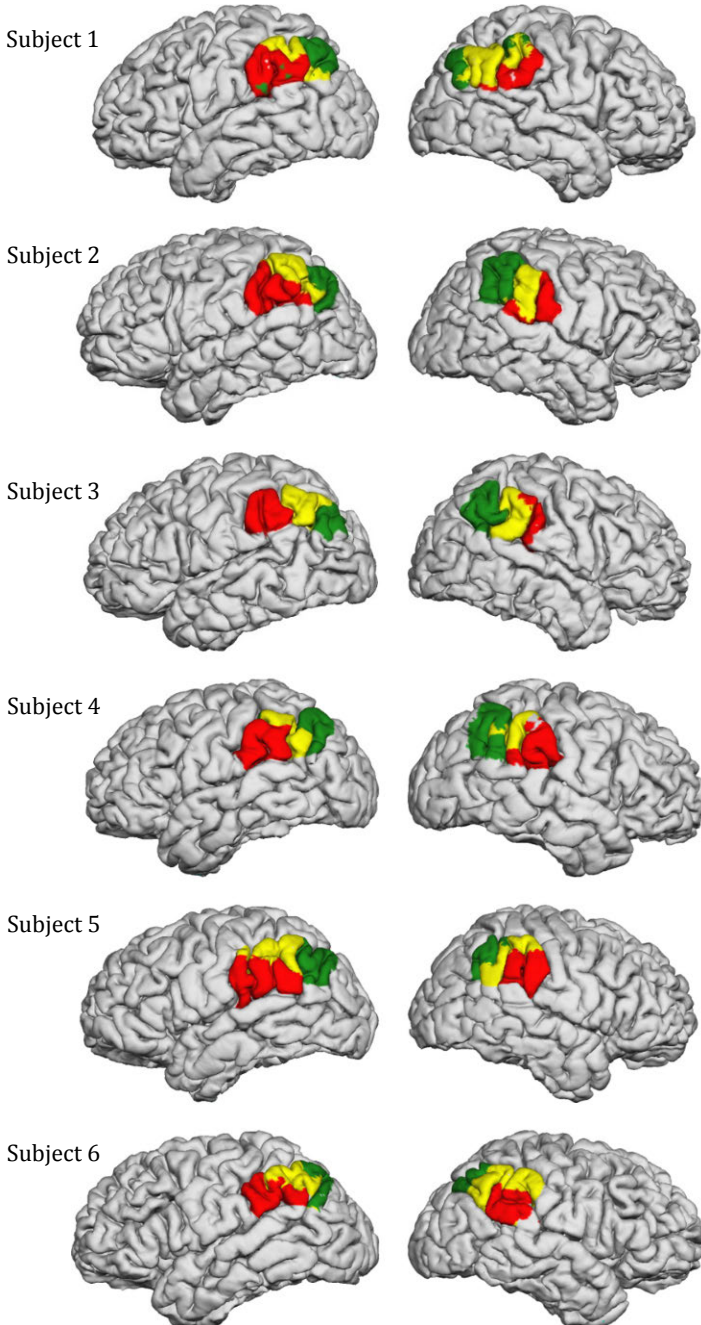
Subject 1

Subject 2

Subject 3

Subject 4

Subject 5

Subject 6

Centroid hierarchical clustering. *SS* partitions. Subjects 1 to 6

Sbj. 1

Sbj. 2

Sbj. 3

Sbj. 4

Sbj. 5

Sbj. 6

~3 clusters   ~5 clusters

_k_-means (Ruschel 2013). 3 clusters. Subjects 7 to 13

Subject 7
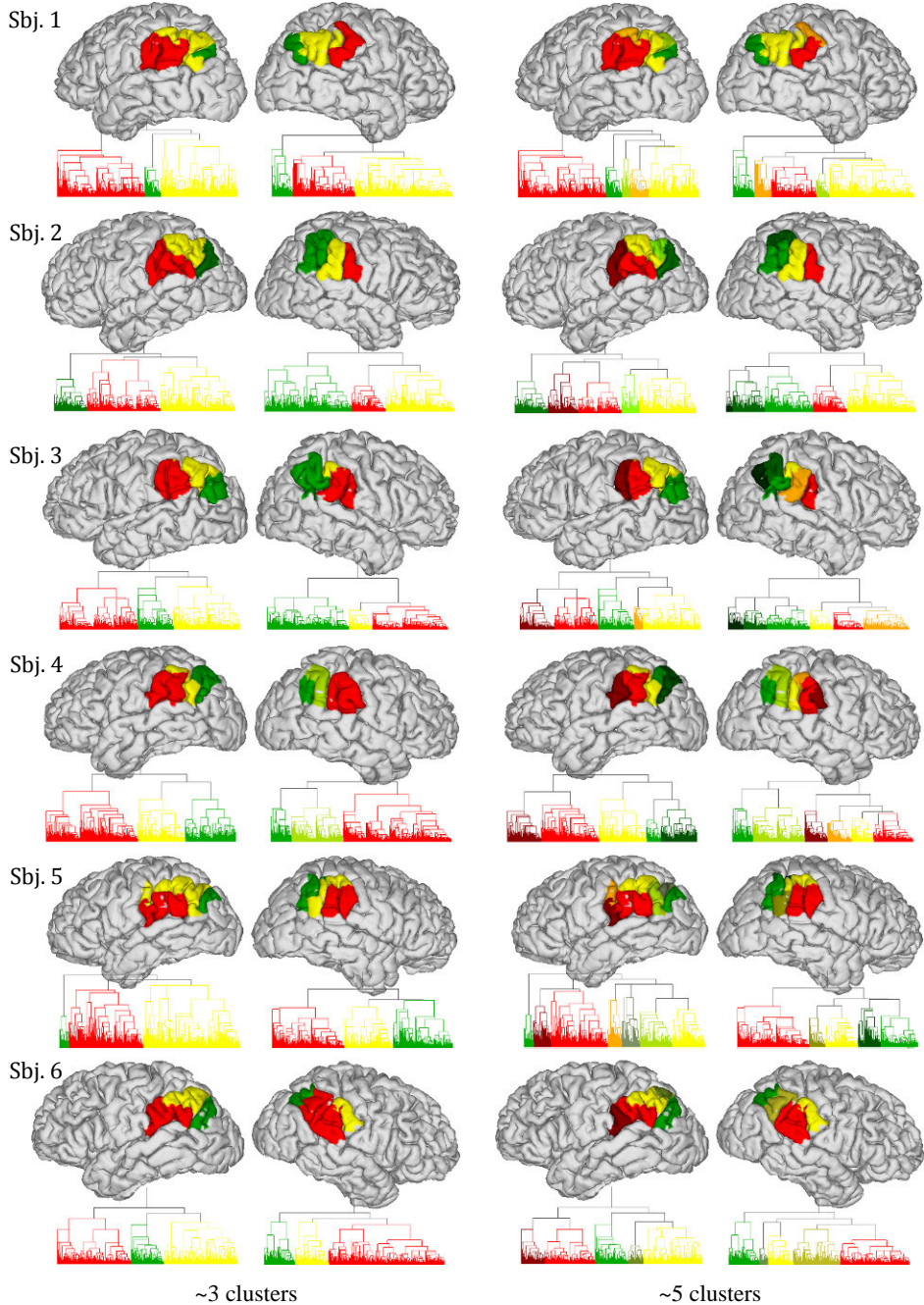
Subject 8

Subject 9

Subject 10

Subject 11

Subject 12

Subject 13

Centroid hierarchical clustering. *SS* partitions. Subjects 7 to 13

Sbj. 7

Sbj. 8

Sbj. 9

Sbj. 10

Sbj. 11

Sbj. 12

Sbj. 13
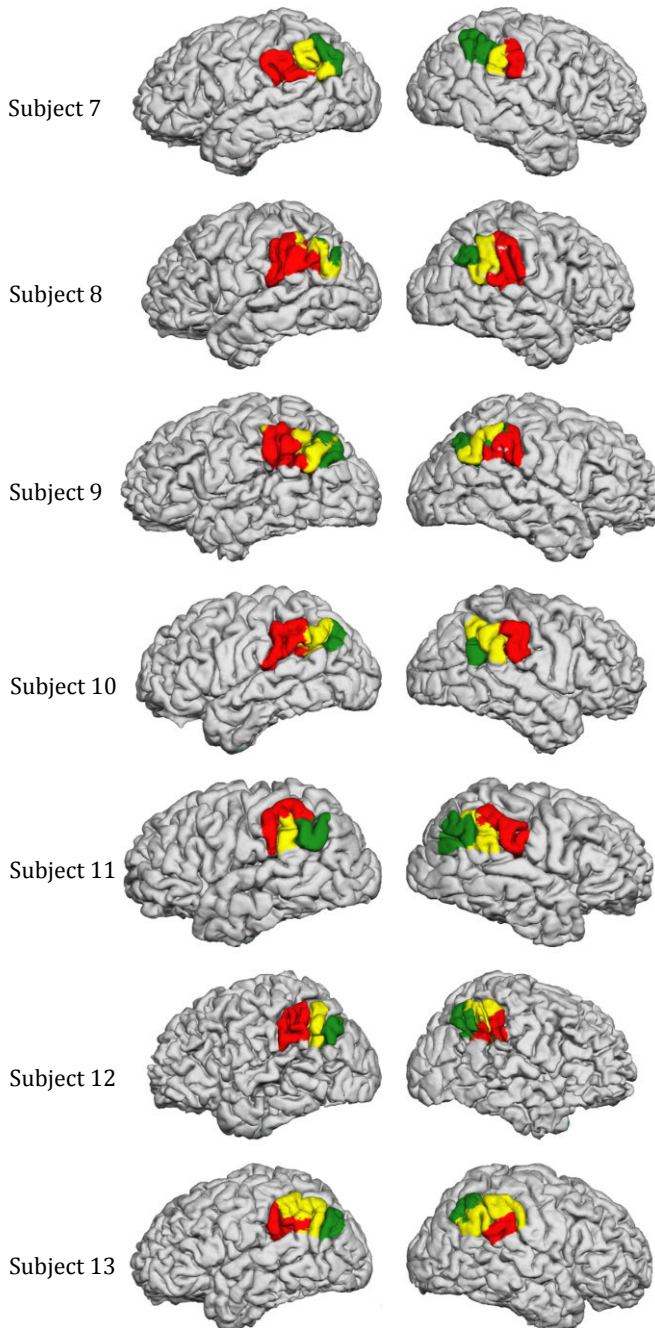
~3 clusters
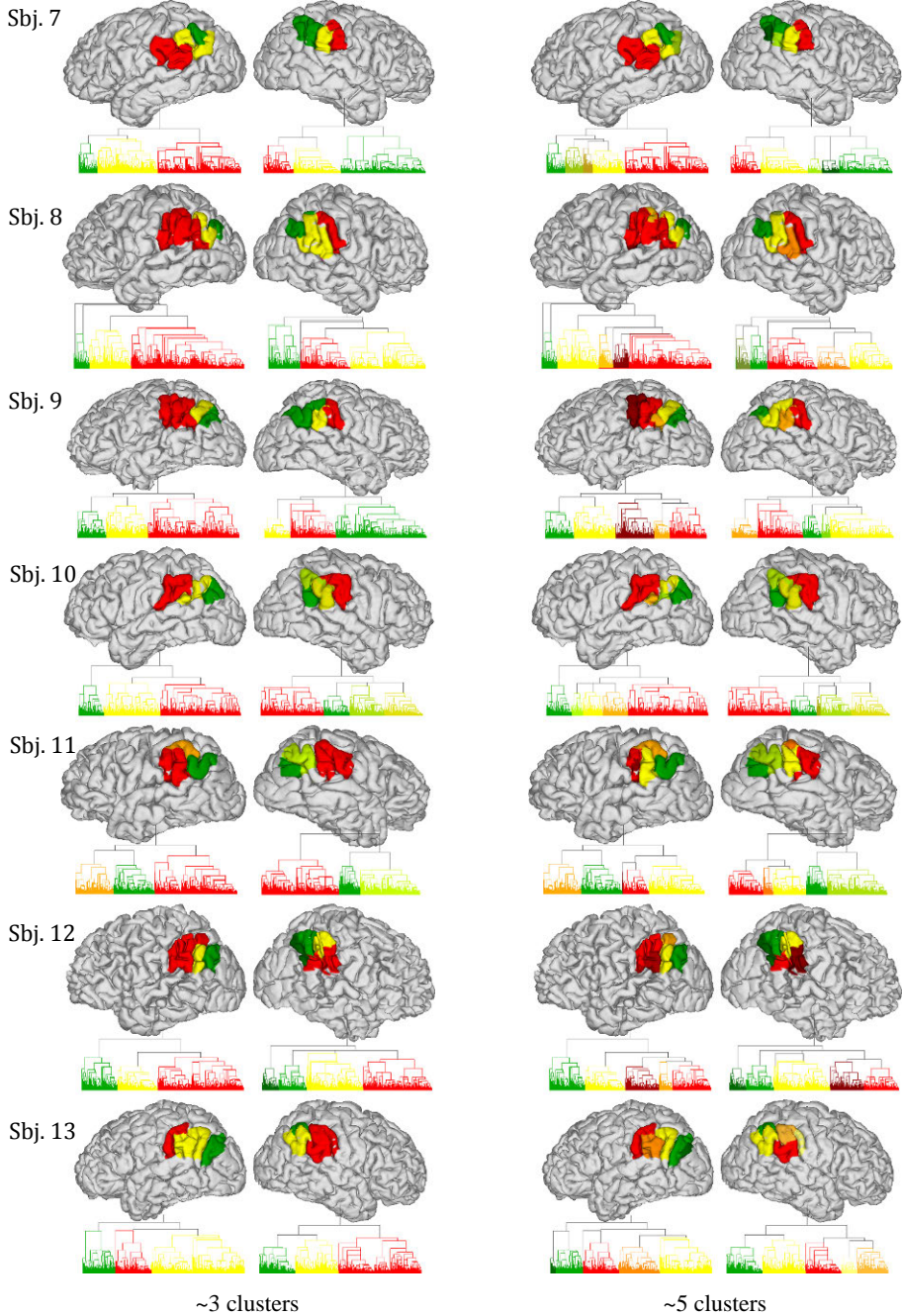
~5 clusters

_k_-means (Ruschel 2013). 3 clusters. Subjects 14 to 19

Centroid hierarchical clustering. *SS* partitions. Subjects 14 to 19

Sbj. 14

Sbj. 15

Sbj. 16

Sbj. 17

Sbj. 18

Sbj. 19

~3 clusters   ~5 clusters
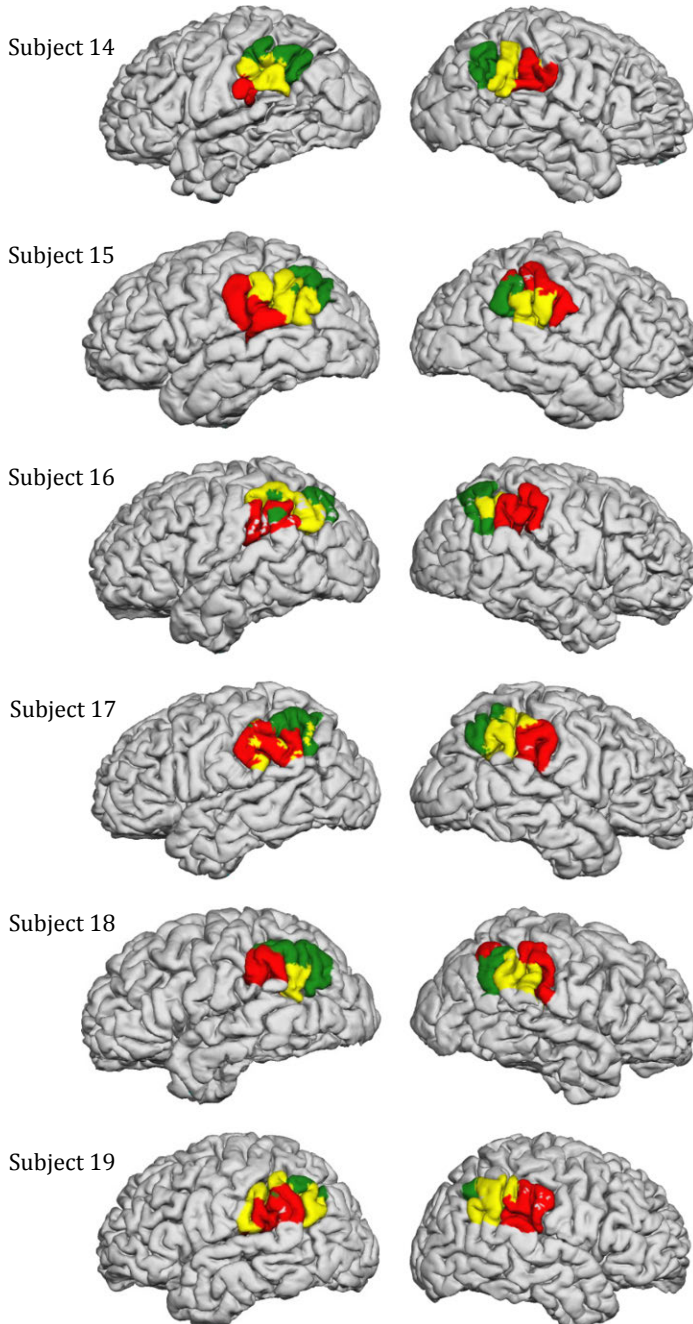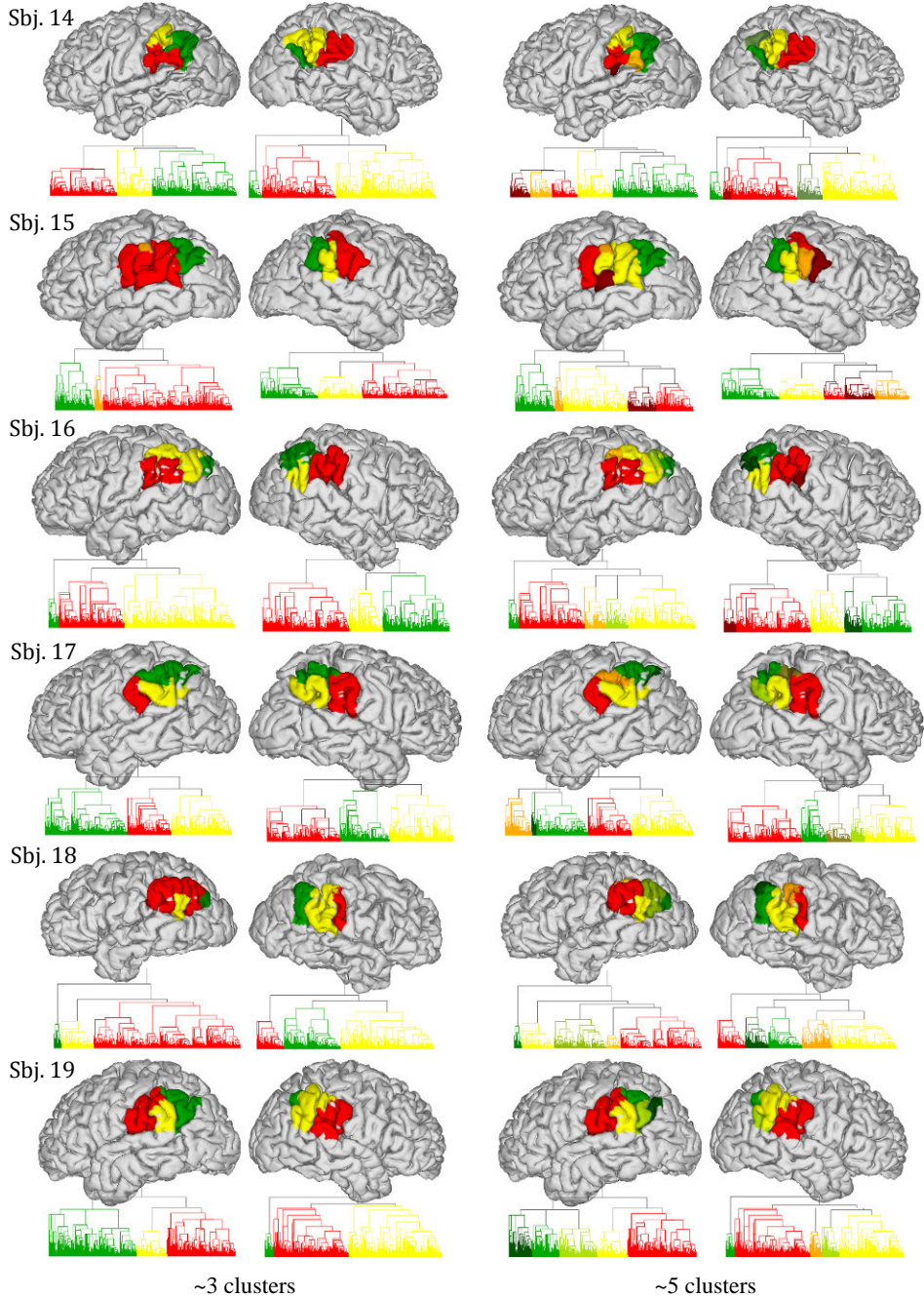
A.2 Horizontal cut partitions for subjects A, B & C

Horizontal cut, subject A: left hemisphere
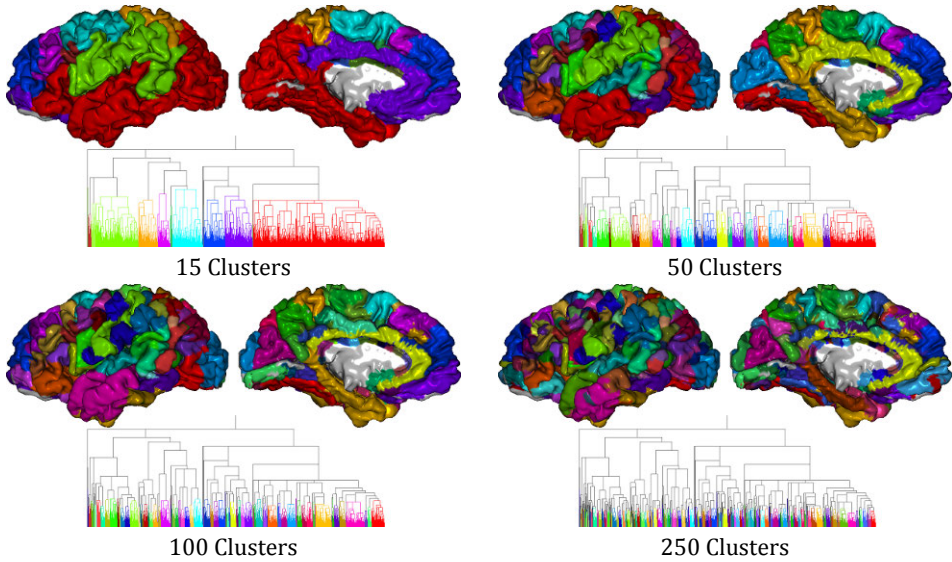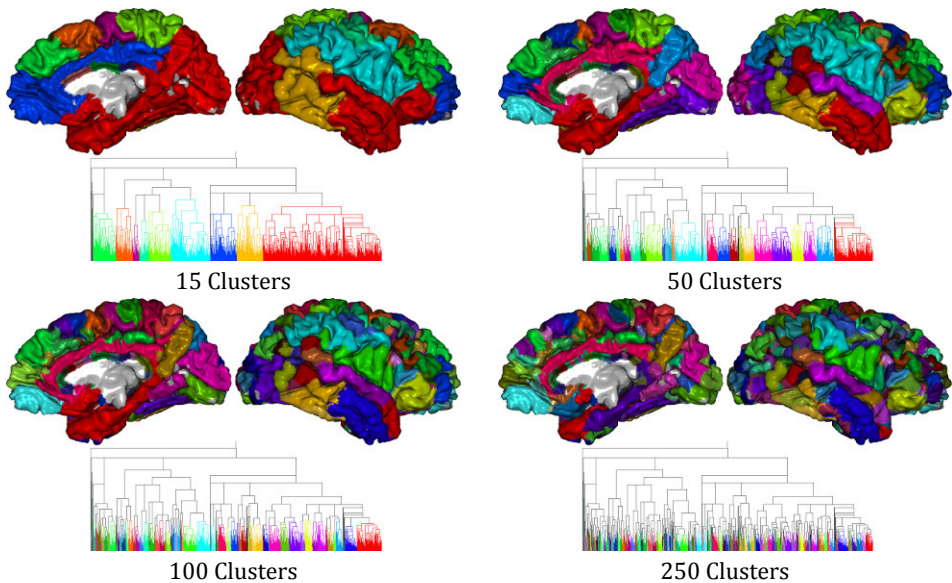
15 Clusters

50 Clusters

100 Clusters

250 Clusters

Horizontal cut, subject A: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

Horizontal cut, subject B: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

Horizontal cut, subject B: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

Horizontal cut, subject C: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

Horizontal cut, subject C: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters
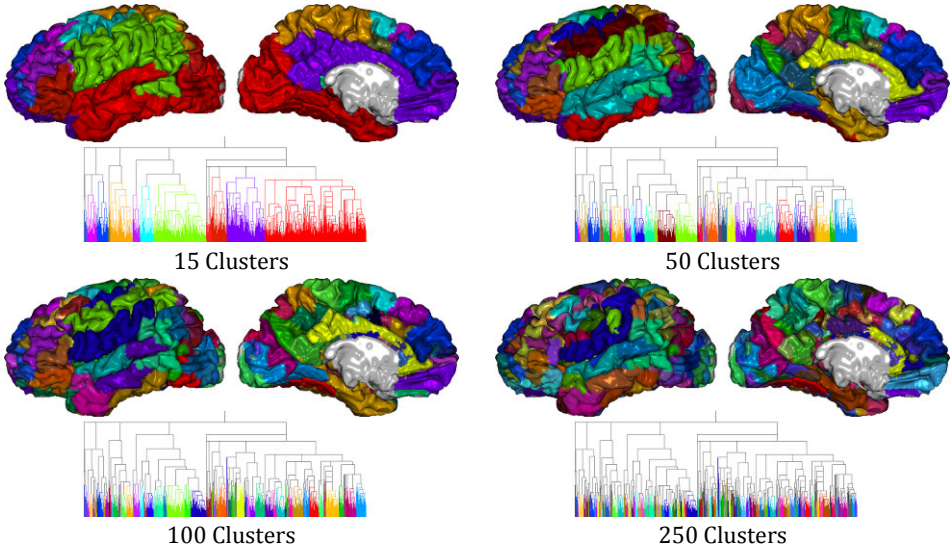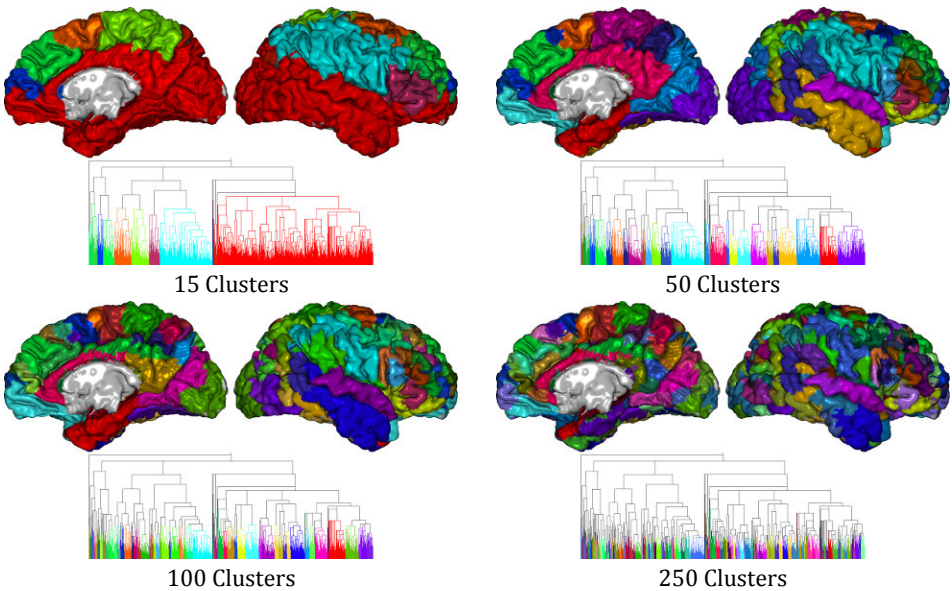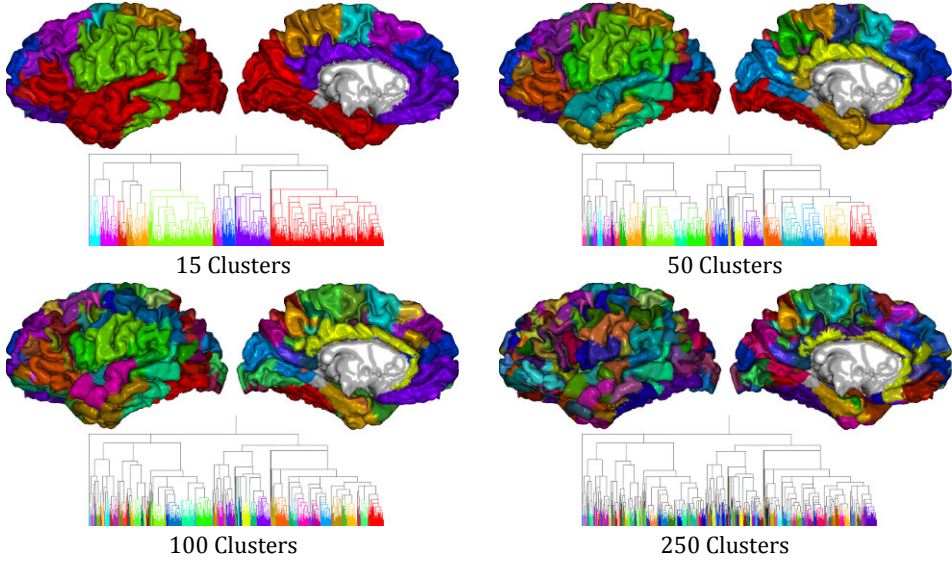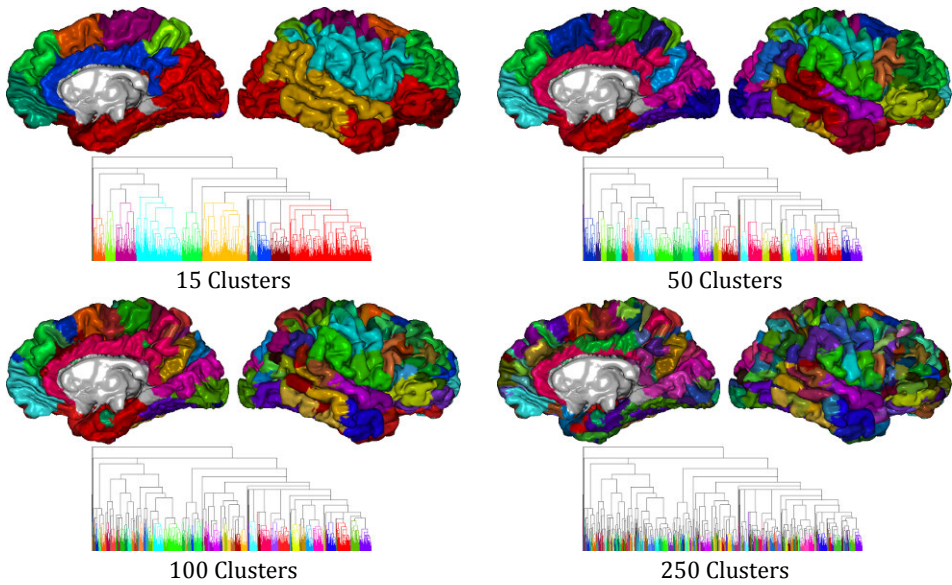
A.3 *SS* method partitions for subjects A, B, C & D

SS cut, subject A: left hemisphere
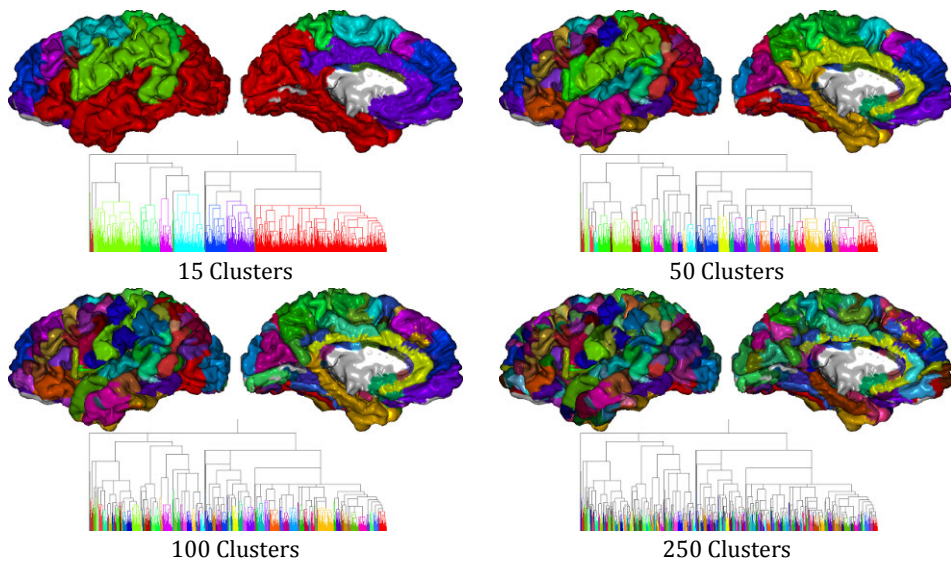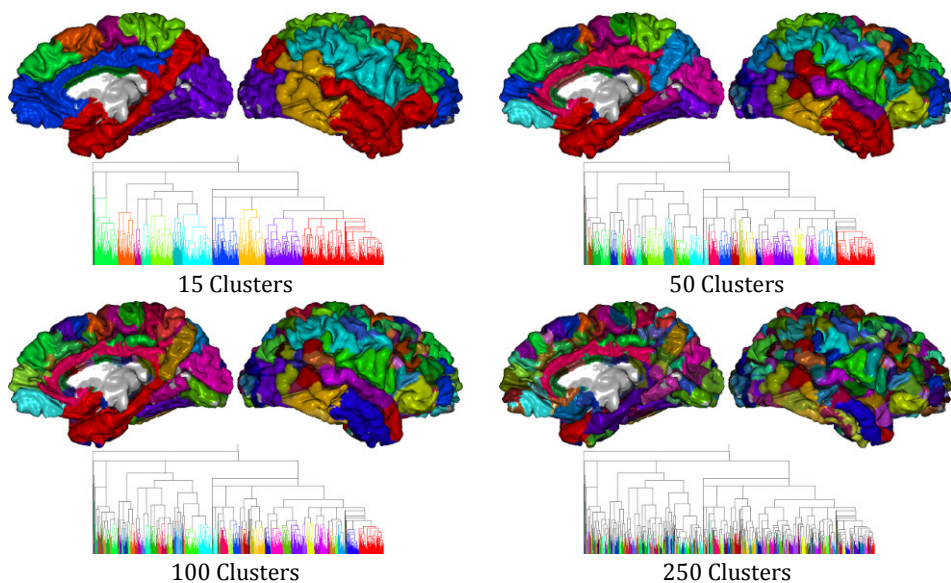
15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject A: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject B: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject B: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject C: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject C: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject D, measurement 1: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject D, measurement 1: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject D, measurement 2: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

SS cut, subject D, measurement 2: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters
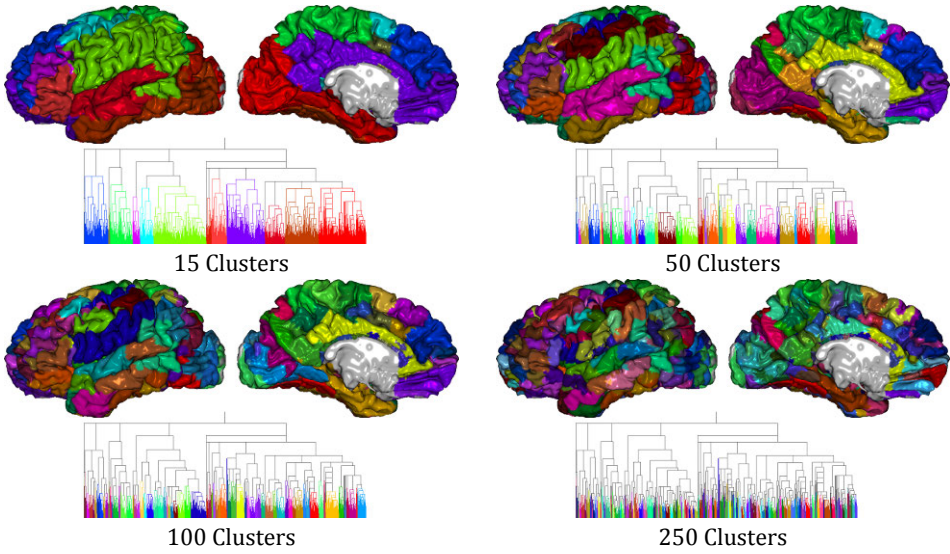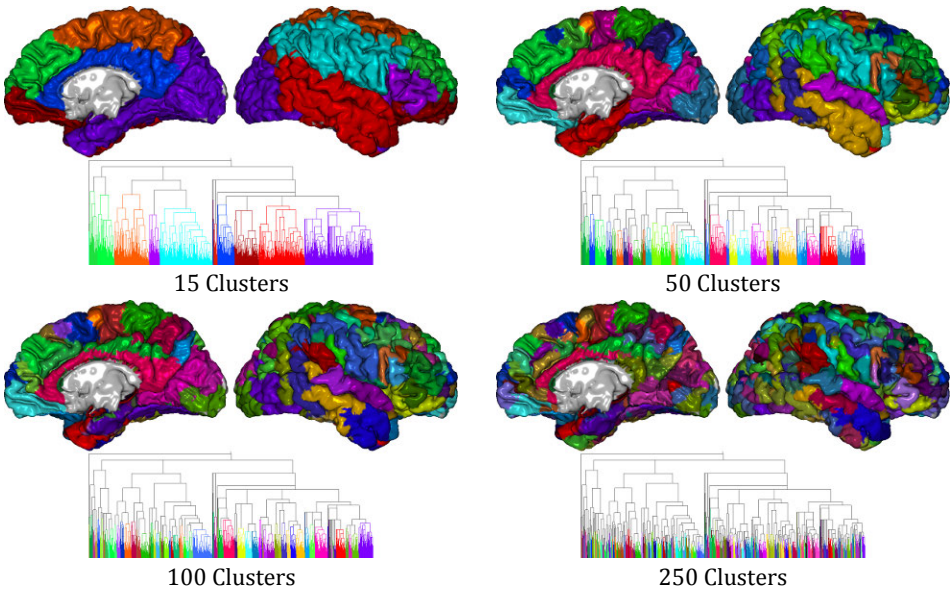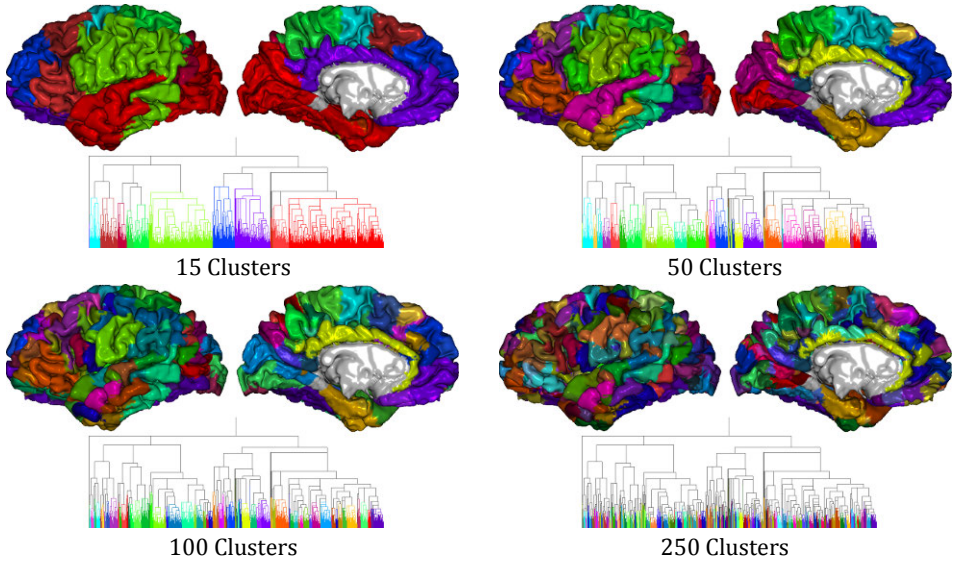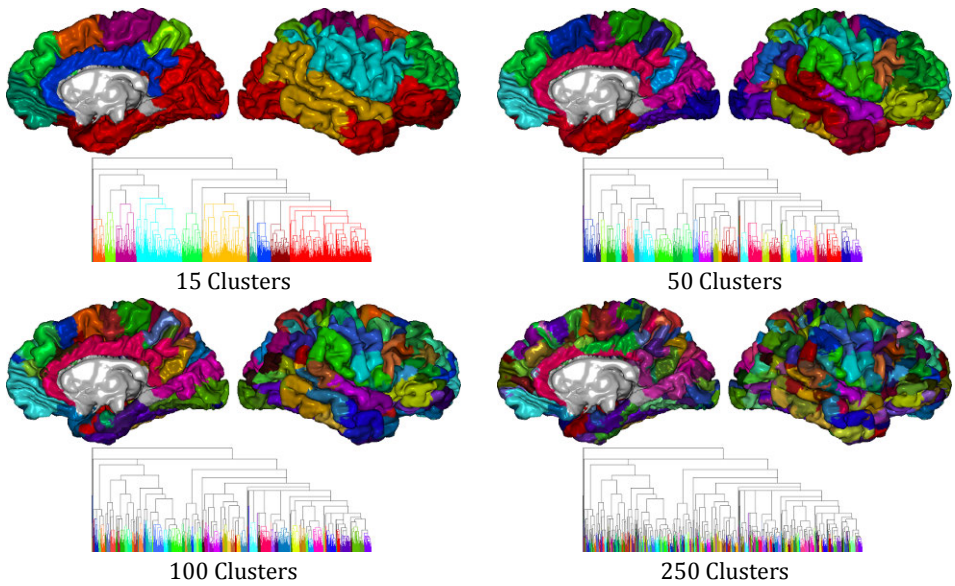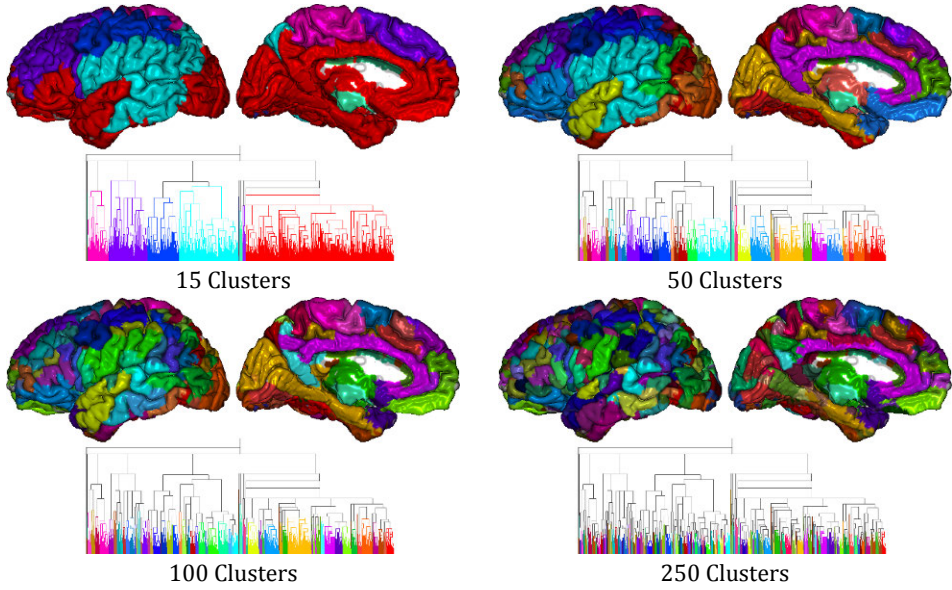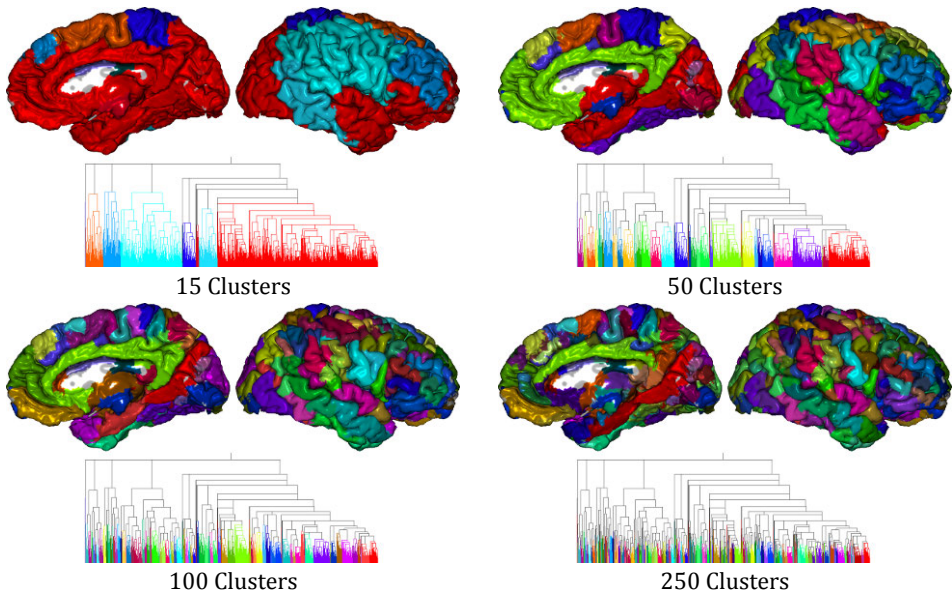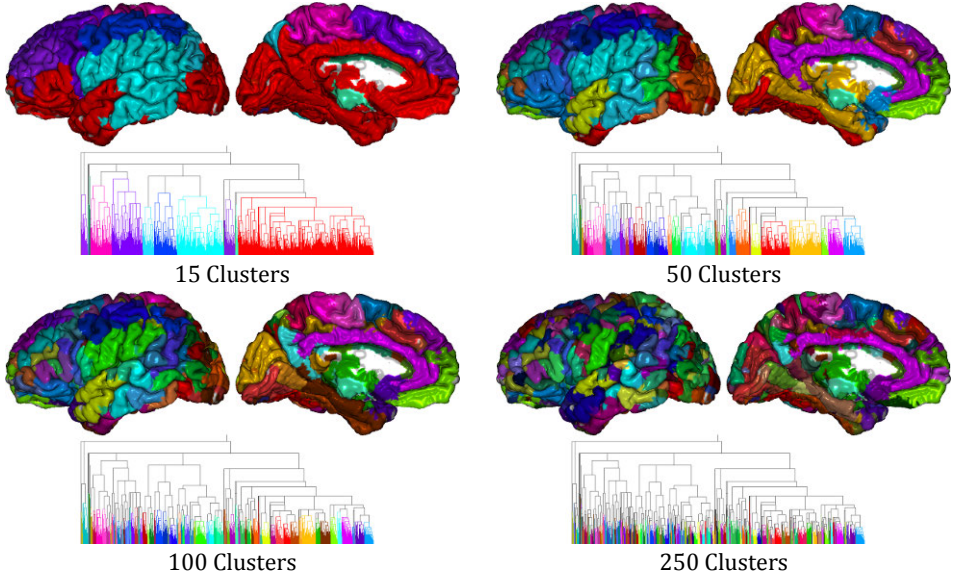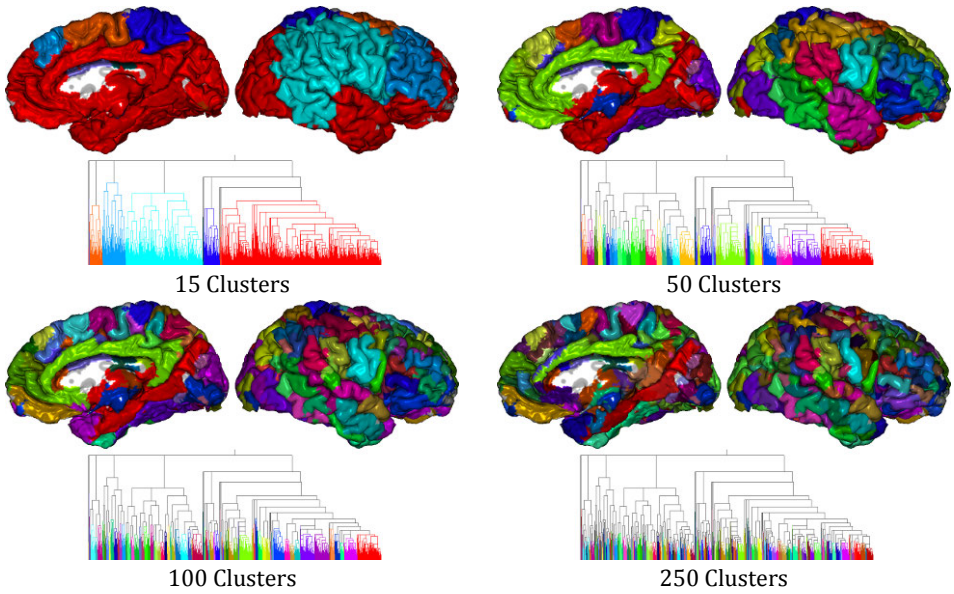
A.4 min. size difference partitions for subjects A, B & C

min. size cut, subject A: left hemisphere
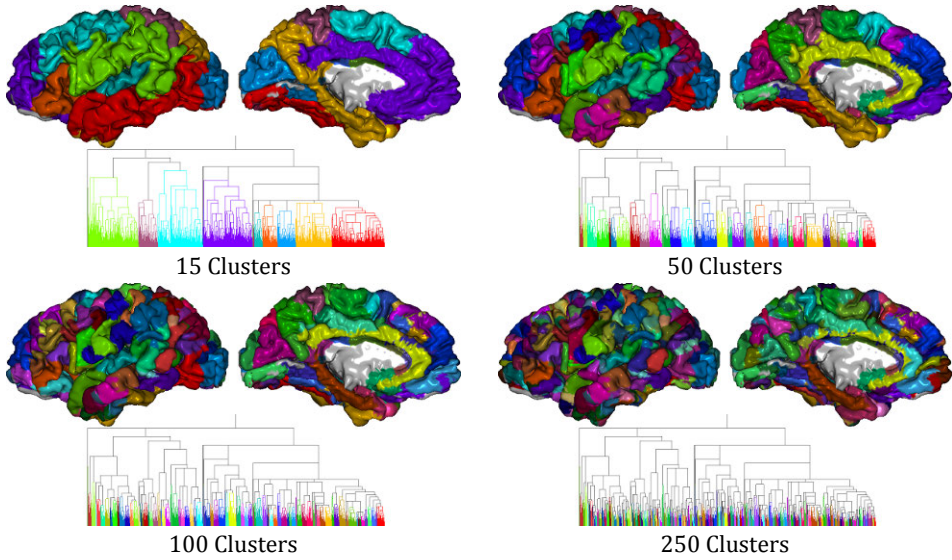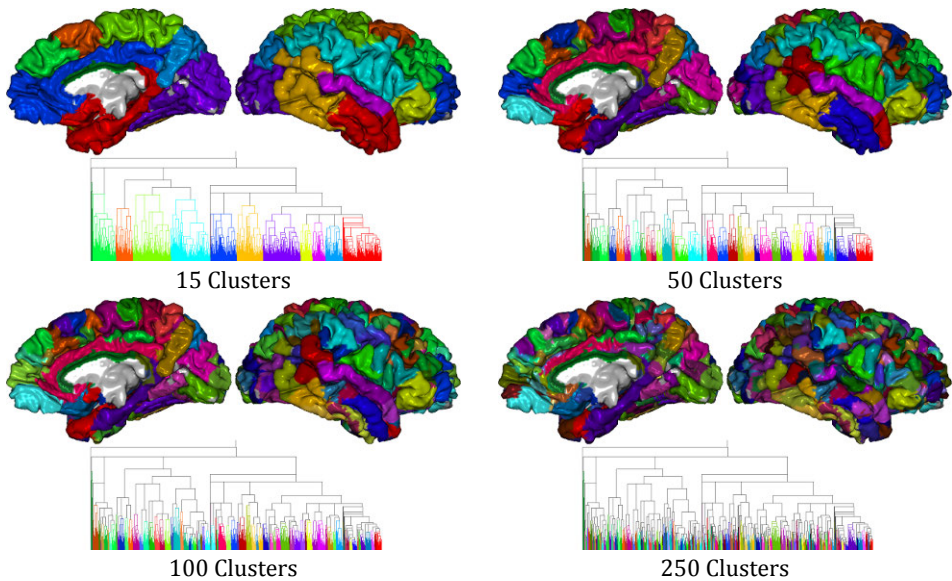
15 Clusters

50 Clusters

100 Clusters

250 Clusters

min. size cut, subject A: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

min. size cut, subject B: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

min. size cut, subject B: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

min. size cut, subject C: left hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters

min. size cut, subject C: right hemisphere

15 Clusters

50 Clusters

100 Clusters

250 Clusters
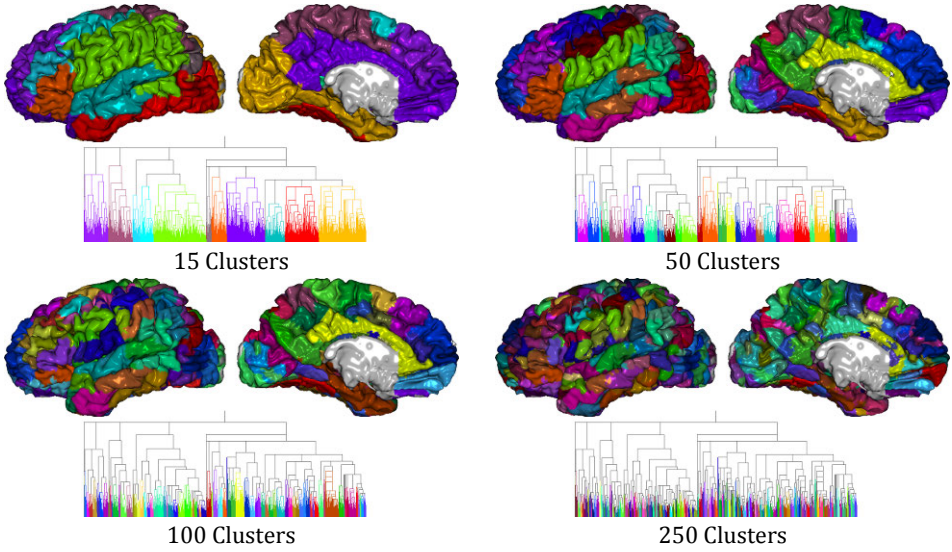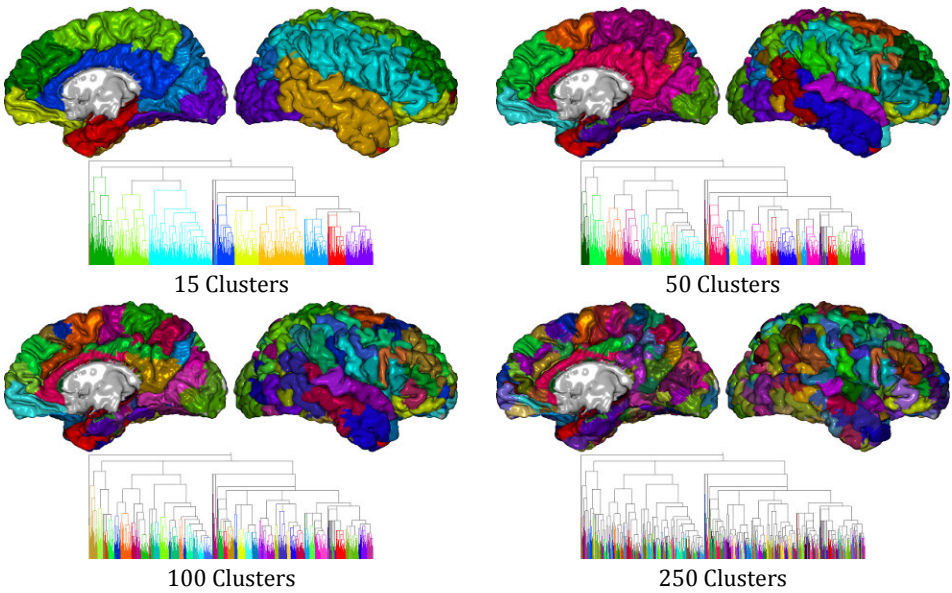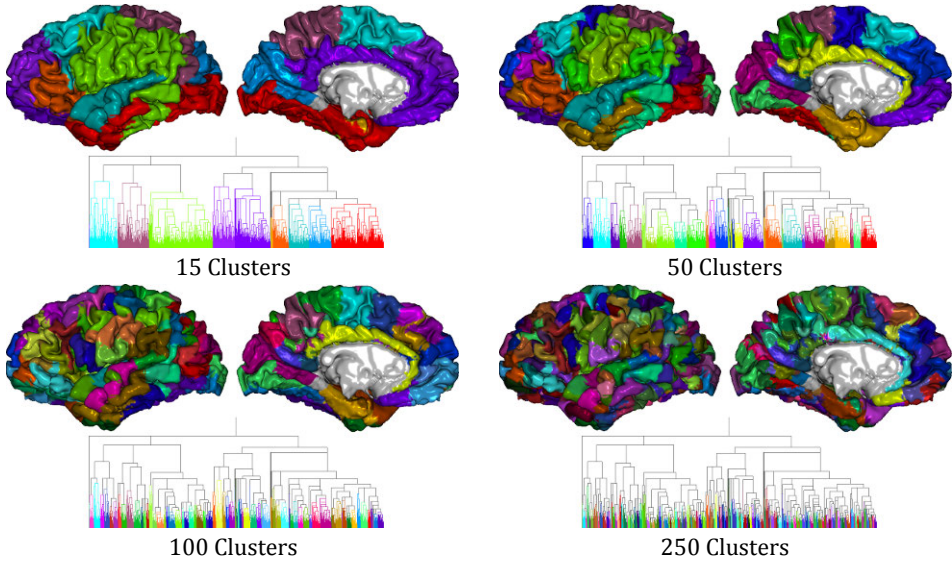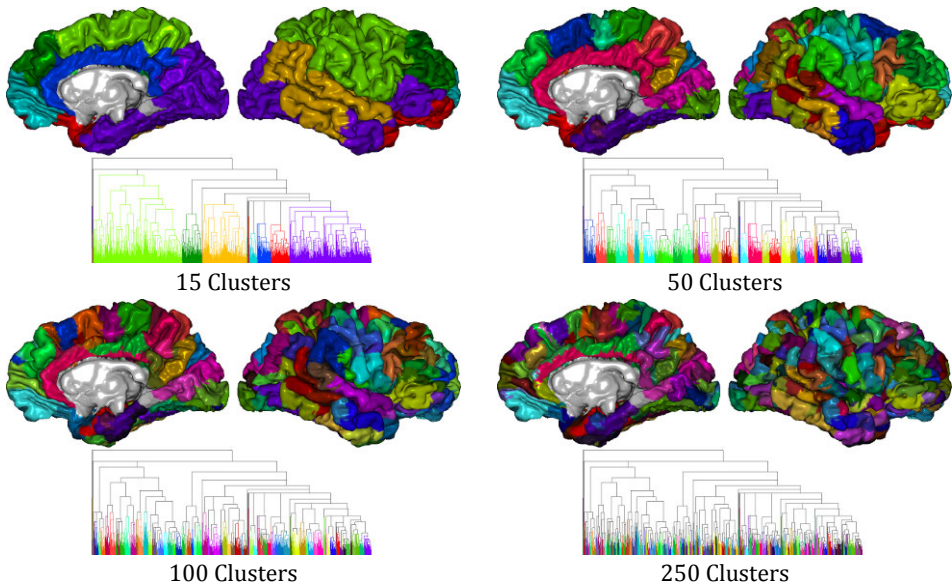
A.5 "Stable boundaries" solution in subjects A, B, C & D

Stable boundaries for subject A: left hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface
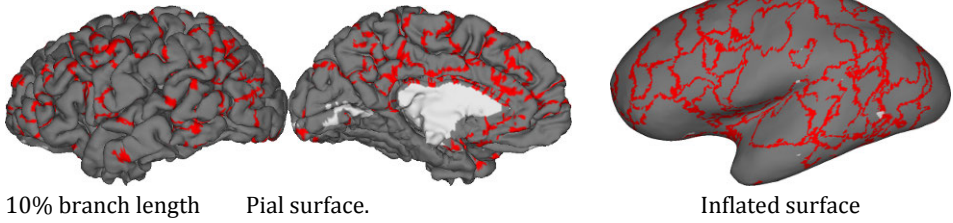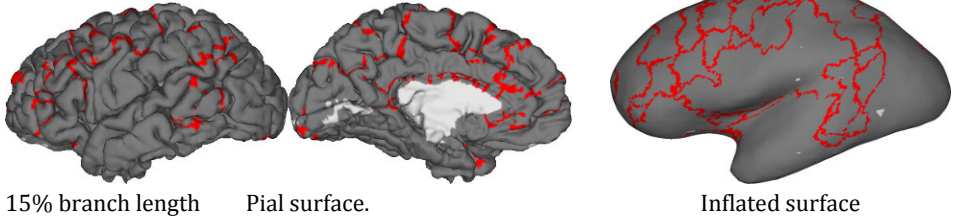
Stable boundaries for subject A: right hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface
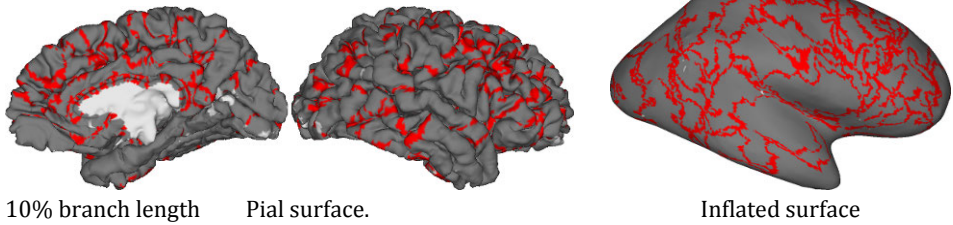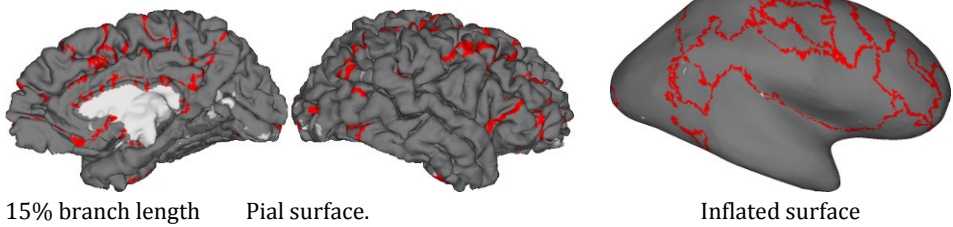
Stable boundaries for subject B: left hemisphere
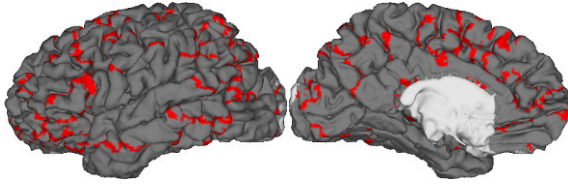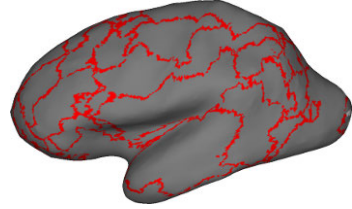
15% branch length   Pial surface.   Inflated surface
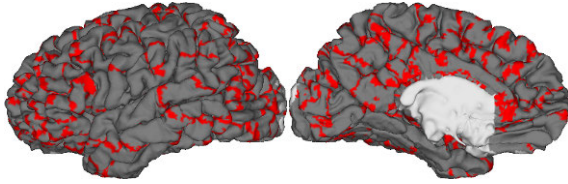
10% branch length   Pial surface.   Inflated surface
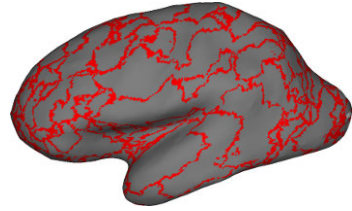
Stable boundaries for subject B: right hemisphere
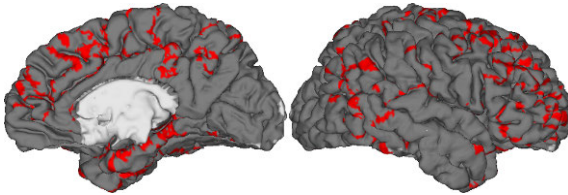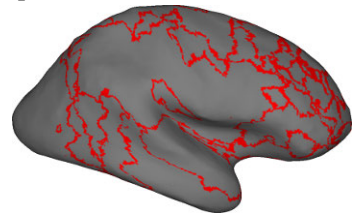
15% branch length   Pial surface.   Inflated surface
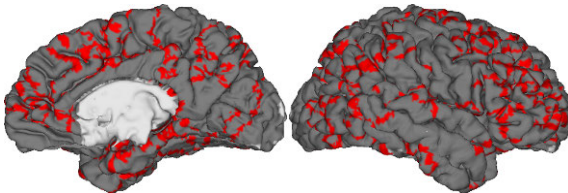
10% branch length   Pial surface.   Inflated surface
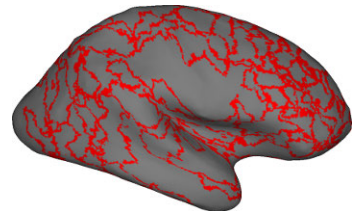
Stable boundaries for subject C: left hemisphere
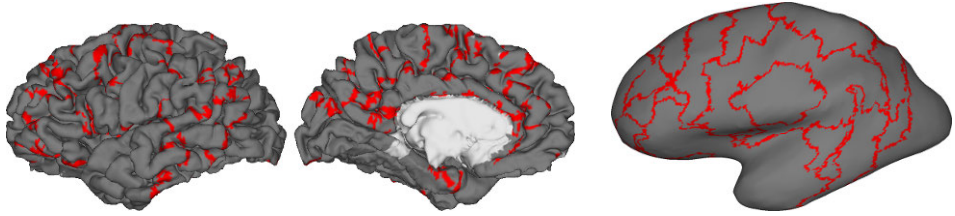
15% branch length   Pial surface.   Inflated surface
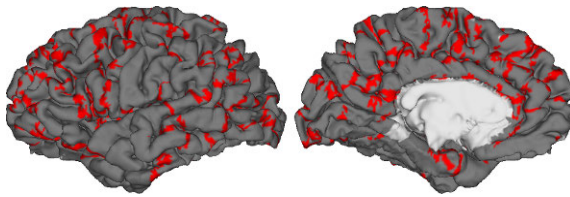
10% branch length   Pial surface.   Inflated surface

Stable boundaries for subject C: right hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface

Stable boundaries for subject D, measurement 1: left hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface

Stable boundaries for subject D, measurement 1: right hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface

Stable boundaries for subject D, measurement 2: left hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface

Stable boundaries for subject D, measurement 2: right hemisphere

15% branch length   Pial surface.   Inflated surface

10% branch length   Pial surface.   Inflated surface

MPI Series in Human Cognitive and Brain Sciences: