Geometric and Optimization Methods for Diffusion Magnetic Resonance Imaging

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Hamza Farooq

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Tryphon T. Georgiou, advisor
Christophe Lenglet, co-advisor

August 2017

Acknowledgments

My gratitude and thanks to my advisor Professor Tryphon Georgiou, whose guidance and support during the last five years, has been pivotal to accomplish this work. The discussions I had with him were the main source of motivation for me to survive as a PhD student. I am thankful to him for describing subtle concepts in a very simple and interesting way. When I decided to embark on medical imaging research during second year of my PhD, he fully supported to pursue my passion. I feel grateful to have him as my PhD advisor.

I thank my co-advisor Professor Christophe Lenglet for accepting me at the Center for Magnetic Resonance Research (CMRR) and introducing me to diffusion MRI related research problems. It was his guidance and critical evaluation of the work, which helped in bringing more value to the study results. I am grateful to him for being so patient in mentoring this work. I am also thankful to Dr. Yongxin Chen for helping and collaborating in many interesting research problems. His support was invaluable for me in understanding and implementing geometrical ideas presented in this work. My sincere gratitude to Professor Allen Tannenbaum whose motivation and guidance was fundamental to graph curvature study for brain networks.

I wish to express my sincere thanks to committee members Professor Murti Salapaka and Professor Tay Netoff for providing useful feedback, which helped improving this thesis. I would also like to thank my collaborators Professors Essa Yacoub and Junqian Xu for providing valuable input regarding microstructure-imaging study. My special thanks to colleagues at CMRR, Dr. Pramod Pisharady, Shu-Hsien Chu and Dr. James Joers, for being so supportive. I am grateful to my labmates Dr. Lipeng Ning and Dr. Francesca Carli for supporting me in my early days at the department. My appreciations to Dr. Kaoru Yamamoto, Shreyas Bhaban, Saurav Talukdar, Dr. Armin Zare, Dr. Xiaofan Wu, Dr. Neil Dhingra, Sepideh Hassan Moghaddam, Wei Ran and Dongsheng Ding for all the help.

Last but not the least, my special thanks to my parents, my brother and my wife Kiran. My father encouraged me to realize my potential and pursue PhD and Kiran was always there to help. Your support is the most valuable thing I have.

This work was partly supported by NSF (ECCS-1509387), the AFOSR (FA9550-15-1-0045, FA9550-17-1-0435), NIH (P41 EB015894, P30 NS076408, P41 RR013218,

P41 EB015902, U24 CA18092401, R01AG048769, P30 CA008748), the Vincentine Hermes-Luh Endowment and the Fulbright Program.

Data used in this work were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Also in part by the Human Connectome Project, MGH-USC Consortium (Principal Investigators: Bruce R. Rosen, Arthur W. Toga and Van Wedeen; U01MH093765) funded by the NIH Blueprint Initiative for Neuroscience Research grant; the National Institutes of Health grant P41EB015896; and the Instrumentation Grants S10RR023043, 1S10RR023401, 1S10RR019307.

Dedication

To my parents and Kiran

Abstract

This thesis presents novel mathematical and computational methods aimed at enhancing and improving brain tissue structural imaging techniques that are based on diffusion Magnetic Resonance Imaging (dMRI). The most commonly used dMRI technique is Diffusion Tensor Imaging (DTI), which models water diffusion via a Gaussian pattern and estimates the corresponding covariance, also known as diffusion tensor. DTI forms the basis of brain structural connectivity methods like tractography and sub-cortical region parcellation, and thus provides useful markers for brain white matter integrity. Other, recently proposed dMRI techniques rely on modeling water diffusion in intra-axonal and extra-axonal spaces separately. Thereby, these so-called multi-compartment models hold the promise to provide detailed tissue microstructure information and to identify markers that may be specific to particular tissue development/diseases. In this thesis we address key mathematical challenges encountered by DTI, as well as by these newly proposed dMRI techniques, that pertain to recovering more detailed microstructure information.

We begin by focusing on DTI and present novel geometrical methods to improve DTI analysis (Chapters 3, 4, and 5). In particular, (i) we utilize the mathematical theory of Optimal Mass Transport to improve brain parcellation by comparing sub-cortical regions connectivity profiles and compute their corresponding geometric "average" connectivity profiles, (ii) we introduce Ricci flow applied to diffusion tensor fields to enhance feature extraction, and finally (iii) we introduce a notion of discrete Ricci curvature in brain connectivity networks as a novel nodal measure to detect critical regions (nodes) of the structural brain networks. This notion of node curvature can be used to identify changes in brain network structure due to disease/development as it supplements information that can be obtained by other conventional network nodal measures. We then study multi-compartment dMRI models, and present a novel model fitting method to such tissue models (Chapter 6). Our proposed method is generic to all multicompartment models and enables for the first time dMRI-imaging in multiple fiber orientations and fiber-crossings situations. In addition to potential improvements in imaging technology, we hope that the advances presented in this work will contribute to the diagnosis and treatment of neurological disorders.

Contents

List of Figures xi

Nomenclature xiv

Chapter 1 Introduction 1

Chapter 2 Basics of Diffusion MRI 6

2.1 Nuclear Magnetic Resonance (NMR) 7

2.1.1 Nuclear spin 7

2.1.2 Nuclear spin in an external magnetic field B_0 7

2.1.3 Magnetic Resonance 8

2.1.3.1 Coordinate system convention 8

2.1.3.2 Excitation Phase 8

2.1.3.3 Relaxation Phase 9

2.1.3.4 Spatial localization by applying magnetic gradient G 10

2.1.4 Bloch equations - preliminaries 11

2.1.4.1 Magnetic moment 11

2.1.4.2 Torque from Larmor precession 11

2.1.5 Bloch equations 12

2.2 Diffusion . . . 13
2.3 Diffusion MRI . . . 14
2.3.1 Sensitizing MR to water diffusion - Bloch Torrey equations . . . 14
2.3.2 Solution to the Bloch-Torrey equation . . . 15
2.4 Diffusion Tensor Imaging . . . 17
2.4.1 Imaging Sequence . . . 17
2.4.2 Diffusion Tensor . . . 17
2.5 Conclusion . . . 19

Chapter 3 Optimal Mass Transport for Brain Parcellation and Connectivity Mapping 20

3.1 Introduction . . . 21
3.2 Methods . . . 22
3.2.1 Transportation Cost and Average for Distributions . . . 22
3.2.1.1 Optimal mass transport cost for probability distributions: . . . 22
3.2.1.2 Wasserstein metric: . . . 23
3.2.1.3 Geometric mean of distributions (Wasserstein barycenter): . . . 24
3.2.2 Computing Geometric mean of distributions using Entropic Regularization . . . 25
3.2.3 Connectivity-based parcellation using Wasserstein metric . . . 27
3.3 Experiments and Results . . . 27
3.3.1 Experiments using synthetic data . . . 27
3.3.2 Experiments using HCP data . . . 29

3.4 Conclusion 31

Chapter 4 Ricci Flow for Feature Enhancement of Diffusion Tensor Fields 32

4.1 Notations & Preliminaries 34

4.2 Smoothing and Feature Enhancement Techniques 35

4.2.1 Ricci Flow for Tensor Fields 36

4.2.2 Anisotropic Smoothing for Tensor Fields 37

4.3 Numerical Implementation 37

4.4 Examples 39

4.4.1 Synthetic Data Experiments 39

4.4.1.1 Backward Ricci Flow - Feature Enhancement 39

4.4.1.2 Forward Ricci Flow - Tensor Smoothing 39

4.4.1.3 Anisotropic Smoothing of the Tensor Field 39

4.4.2 MRI Data Experiments 40

4.4.2.1 Ricci Flow 40

4.4.2.2 Anisotropic Smoothing 41

4.5 Conclusion and Discussion 41

Chapter 5 Ricci Curvature - Hallmark of Structural Brain Connectivity 47

5.1 Introduction 48

5.2 Methods 50

5.2.1 Curvature 50

5.2.2 Ollivier-Ricci and Scalar Curvature 52

5.2.2.1 Ollivier-Ricci Curvature . . . 52
5.2.2.2 Node Curvature . . . 53
5.2.3 Robustness and the Fluctuation Theorem . . . 53
5.2.4 Ollivier-Ricci Curvature and Graph Robustness . . . 53
5.2.5 Measures of Brain Networks Characteristics . . . 55
5.2.5.1 Node Strength s_i . . . 55
5.2.5.2 Betweenness Centrality g_i . . . 55
5.2.5.3 Clustering Coefficient C_i . . . 55
5.2.6 Diffusion MRI Datasets . . . 56
5.2.6.1 DSI Datasets from Hagmann et al. [89] . . . 57
5.2.6.2 DSI Datasets from the MGH-USC HCP Consortium 57
5.2.6.3 HARDI Datasets from the WU-Minn HCP Consortium Lifespan Pilot Project . . . 57
5.2.6.4 DTI Datasets from Rudie et al. [90] . . . 58
5.2.6.5 Generation of Connectivity Matrices for the HCP Datasets 58
5.3 Results . . . 58
5.3.1 Curvature as a Hallmark of Brain Areas Robustness . . . 58
5.3.2 Remarks on Graph Measures and Assessment of Brain Network Robustness . . . 60
5.3.2.1 Gaussian Transformation of the Connectivity Matrix Weights . . . 60
5.3.2.2 Comparison Across Nodal Scales . . . 61
5.3.2.3 Graph Measures to Assess Node Robustness . . . 62
5.3.3 Change in Curvature and Healthy Development/Aging of the Brain . . . 62

5.3.4 Change in Curvature and Autism Spectrum Disorders (ASD) . 63

5.4 Conclusion 64

Chapter 6 Microstructure Imaging of Crossing (MIX) White Matter Fibers from diffusion MRI 73

6.1 Introduction 74

6.2 Tissue Compartment Model Parameters and Functions 76

6.2.1 Parameters of the Compartment Model Functions 77

6.2.2 Biophysical Functions for Compartment Models 78

6.3 Results 79

6.3.1 Synthetic data experiments 79

6.3.1.1 ActiveAx 79

6.3.1.2 NODDI 80

6.3.1.3 Tensor-Stick-Dot (TSD) 82

6.3.1.4 ZCDx 84

6.3.2 Real data experiments 88

6.3.2.1 ActiveAx 88

6.3.2.2 NODDI 88

6.3.2.3 Zeppelin-Cylinder-Cylinder-Dot (ZCCD) 89

6.3.2.4 NODDIx 91

6.4 Methods 92

6.4.1 Problem Formulation 92

6.4.2 Solution Framework 94

6.5 Optimal Scanning Protocols 96

6.6 Conclusion and Discussion 96

Chapter 7 Conclusion and Future Work **98**

Bibliography **101**

List of Figures

2.1 Break down of Chapter 2. 6

2.2 Randomly oriented spinning protons. 7

2.3 Precession and alignment with B_0 8

2.4 Net magnetization vector and its components 9

2.5 Excitation Phase 9

2.6 T_1 relaxation. 10

2.7 T_2 relaxation. 11

2.8 PGSE sequence 18

2.9 Diffusion Tensor Imaging 19

3.1 Arithmetic and geometric means of distributions 25

3.2 Histograms of correlation coefficients and OMT cost 27

3.3 FiberCup replica parcellation 28

3.4 Comparison of arithmetic and geometric mean of connectivity profiles for FiberCup data 29

3.5 Parcellations of the Corpus Callosum 30

3.6 Comparison of arithmetic and geometric mean of connectivity profiles of Corpus Callosum 30

4.1 Feature extraction and smoothing - Synthetic data set 1 experiments 40

4.2 Feature extraction and smoothing - Synthetic data set 2 experiments 43

4.3 Feature smoothing - Synthetic data set 1 and 2 experiments 44

4.4 Feature enhancement experiment (with backward Ricci flow) using MRI data. Number of iterations $n = 40$. 45

4.5 Tensor field smoothing experiment (with anisotropic diffusion) using MRI data, (a) Original tensor field (b) Smoothed tensor field. 46

5.1 Ricci curvature example 51

5.2 Nodes distribution with high Curvature, Strength and Betweenness Centrality - High resolution matrices 66

5.3 Nodes distribution with high Curvature, Strength and Betweenness Centrality - Low resolution matrices 67

5.4 Robustness analysis on the basis of node deletion using high resolution average connectivity matrix 68

5.5 Robustness analysis on the basis of node deletion using low resolution average connectivity matrix from MGH-USC DSI data 69

5.6 Robustness analysis on the basis of node deletion with topological entropy as a measure 70

5.7 Nodes with significant changes in healthy aging 71

5.8 Nodes with significant changes in ASD cohort as compared to TD . . 72

6.1 Fibre Orientation 77

6.2 ActiveAx parameters estimation using synthetic data. 81

6.3 NODDI parameters estimation using synthetic data. 83

6.4 Tensor-Stick-Dot parameter estimation using synthetic data. 85

6.5 ZCDx - Axon radius index estimates in three orientations using synthetic data. 87

6.6 ActiveAx ex-vivo parameter estimation comparison using fixed monkey brain data. 89

6.7 NODDI parameters estimation comparison for in-vivo human brain data. 90

6.8 Zeppelin-Cylinder-Cylinder-Dot parameter estimation for in-vivo human brain data. 91

6.9 NODDI and NODDIx parameter estimation comparison using Human Connectome Project (HCP) data. 93

Nomenclature

Abbreviations and Acronyms

RHS	Right-Hand Side
OMT	Optimal Mass Transport
SNR	Signal to Noise Ratio
dMRI	Diffusion Magnetic Resonance Imaging
DTI	Diffusion Tensor Imaging
TD	Typically Developing (controls)
ASD	Autism Spectrum Disorder
FA	Fractional Anisotropy
MIX	Microstructure Imaging of Crossing Fibers

List of Symbols

∇	Gradient operator
$\times$	Outer product
$\lvert \cdot \rvert$	Absolute value
q	Displacement reciprocal vector
D	Diffusion coefficient / Diffusion Tensor
$\overrightarrow{\mu}$	Magnetic moment
$\overrightarrow{M}$	Macroscopic version of magnetic moment $\overrightarrow{\mu}$
$\langle\rangle$	Ensemble average
$\otimes$	Point wise multiplication
$\oslash$	Point wise division
μ	Initial distribution
ν	Terminal distribution
Γ	Christoffel symbols
κ	Ricci curvature
$\mathbb{R}^n$	Euclidean space of dimension n

Chapter 1

Introduction

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique with the potential to provide detailed brain structural and functional information. In fact, using an MR scanner, with different imaging modalities, one can acquire diverse information on brain structure and function. This includes information on anatomy (anatomical MRI), brain functional activity (functional MRI), metabolites distributions (Spectroscopy), and blood flow information (perfusion MRI).

When sensitized to diffusion of water molecules in brain tissues, MRI is referred to as diffusion-MRI (dMRI) [1, 2]. The technology of dMRI measures hindered/ anisotropic diffusion of water molecules, thereby probing anisotropic properties of tissue and providing invaluable information on the underlying tissue structure. Over the last three decades, MRI scanners aimed at human scanning, have evolved to a magnetic field strength of 10 Tesla. This, in conjunction with advances in acquisition techniques, such as parallel imaging, have lead to unprecedented availability of high-resolution data. However, in order to extract useful information from high-resolution MR data, new and powerful processing tools are needed. Efficient and reliable MR data processing techniques will be invaluable in the future for understanding brain tissue development and identifying markers of various conditions/diseases.

In this thesis, we focus on developing such novel processing tools for dMRI. In particular, we address several theoretical issues in dMRI-data processing by introducing into the subject, Optimal Mass Transport (OMT) techniques and the key geometrical concepts of Ricci curvature and Ricci flow. We also introduce a complex-

tissue microstructure model-fitting technique for dMRI data, that applies generically to all existing models, and allows detecting multiple axonal directions (and related features) for the first time. Thus, next, we provide historical highlights of MRI/dMRI and related technological challenges, we then discuss briefly the OMT problem that is central to certain geometrical methods introduced in this thesis, and we conclude with an overview of the contributions in this work.

Brief history of MRI: Gorter first predicted Nuclear Magnetic Resonance (NMR) phenomenon in 1936. However, Felix Bloch and Edward Purcell, simultaneously and independently demonstrated it for the first time in 1946. Both were awarded jointly Nobel Prize for Physics in 1952. Since then, NMR spectroscopy with uniform external magnetic field has been used to study composition of solids and liquids. In 1950, Herman Carr introduced gradient in magnetic field enabling acquisition of one-dimensional images and formed basis of NMR tomographic imaging. Hahn, in the same year, noted MR signal attenuation due to random thermal motion of the spins, laying foundation for dMRI. Paul Lauterbur, in 1971 demonstrated that using field gradients, two dimensional NMR images can be obtained. Peter Mansfield took first NMR medical image in 1974. For this work, Paul Lauterbur and Peter Mansfield jointly received Nobel Prize in 2003. In 1976, Raymond Damadian used T1 and T2 relaxation to image cancer tissue in a live animal. The imaging methodology has been persistently improved since then and the work is ongoing with MR as the most promising medical imaging modality for the future.

Diffusion Tensor Imaging and Multicompartment models: As noted earlier, dMRI probes hindered diffusion of water molecules and inherently provides information about tissue microstructure. Diffusion Tensor Imaging (DTI) assumes a Gaussian distribution for the diffusion of water molecules in a given structure. More specifically, diffusion of water molecules is modeled by the covariance matrix (referred to as the diffusion tensor) of Gaussian distributions in a volume element (voxel) [3,4]. Since anatomical MRI does not help in identifying any structural information of white matter, DTI has proved to be very useful in identifying tissue geometry. Anisotropic water diffusion captured by DTI helps identify axonal bundle orientations and thereby probe the connectivity of sub-cortical regions (tractography). DTI is being applied to a great variety of clinical studies including studies for brain development, the effects of aging, Alzheimer's, Parkinson's disease, and so on. DTI also provides markers like Fractional Anisotropy (FA) and Mean Diffusivity (MD) to study white matter

structure. However, the diffusion tensor model is often too coarse, and the markers turn out to be non-specific. To address this issue, more detailed multicompartment tissue models have been proposed (see for example [5, 6]). These complex models bring out information on water diffusion separately in intra and extra axonal spaces, thus enabling direct measurement of axonal features like axon bundle radius and density etc. in each voxel. This invaluable information, that can be drawn from multicompartment models comes at the cost of great difficulty in model fitting as well as sensitivity of the process to various factors. Indeed, such model fitting to dMRI data is a challenging task and requires substantially greater computational effort as well as efficient fitting algorithms [7, 8]. This is one of the issues that we address in this work by presenting a novel fitting algorithm for multicompartment models. The algorithm we present is generic, in that it applies to all models presented to date, and enables model-fitting in fiber crossings, a feat that was not possible using existing state-of-the-art methods.

Brief history of Optimal Mass Transport: Gaspar Monge first formulated OMT in 1781. He sought to find an optimal way to transport a pile of soil from one place to another. In other words, the problem amounts to seeking an optimal transportation plan so as to correspond mass between two end point marginal distributions. Monge explored certain geometric properties of the solution to this problem but was unable to provide a full solution. The problem attracted the interest of several great mathematicians of the 19th and 20th century but remained largely open until 1942 when Leonid Kantorovich modified the original formulation to what is now known as the Monge-Kantorovich problem. He introduced a relaxation of Monge's problem along with duality theory and linear programming to find a solution. Recent contributions in the field by Brenier, McCann, Cullen, Mather and others have brought about a new fast developing phase in OMT-theory with great many applications in a wide range of fields (probability theory, physics, image and biomedical image processing, spectral analysis, economics, stochastic control, galaxy dynamics, and so on, see e.g., [9, 10]). Our interest stems from the relevance of OMT in image analysis as well as the fact that OMT provides a natural metric to compare distributions. Those two aspects come together in a variety of instances in our effort to develop analytical tools for MRI technology.

Contributions in this thesis: The main contributions presented in this thesis are as follows. First, we introduce the theory of OMT and certain geometrical con-

cepts of Ricci curvature and Ricci flow to develop innovative signal processing tools aimed at DTI. We refer to these contributions collectively as geometrical methods for DTI and we present these in Chapters 3, 4 and 5 of the thesis. Second, we propose a novel fitting method of multicompartment tissue models to dMRI data. This contribution initiates the use of certain optimization methods for multicompartment model-parameters estimation and is detailed in Chapter 6. The outline of the thesis is as follows:

Chapter 1: General background on magnetic resonance and related technology is provided in Chapter 1. More specifically, in this chapter we explain the theoretical background behind MRI technique. We derive the Bloch-Torrey equations and discuss their solution so as to clarify the basic mathematics behind the MRI modality. The diffusion process of water molecules, and how MRI data are sensitized to this, leading to dMRI is also explained. In short, the purpose of the chapter to collect and explain basic facts that will offer the reader sufficient background to appreciate and follow the subsequent exposition of DTI and dMRI processing ideas in the body of the thesis.

Chapter 2: In this chapter we explain brain areas parcellation and their connectivity mapping. Generally speaking, connectivity profiles are lines/links which are formed by following principal eigen-vectors of diffusion tensors. These profiles point to connections between one part of the brain with the other. Brain areas are parcellated based on similarity or correlation of these connectivity profiles. Here we introduce OMT to quantify (using the Wasserstein metric) similarity between profiles treating them as distributions. In addition, we introduce the notion of a geometric barycenter for the connectivity profiles to capture mean connectivity of a sub-cortical region in a meaningful anatomical sense. To this end, we introduce a Wasserstein-barycenter with support only on the support of connectivity profiles. This contribution has been accepted for publication and will appear in Computational Diffusion MRI: MICCAI Workshop 2017.

Chapter 3: This chapter provides a feature enhancing method based on "backward" Ricci flow on tensor fields. DTI models diffusion of water molecules via a tensor corresponding to each voxel. Ricci flow represents an analogue of the heat equation and, in the forward directions, smooths out the tensor field. In the "backward" direction is accentuates structural features and provides a sharpening effect to the tensor field. Judicious choice of step size and number of steps can be used to a great advantage in highlighting important features in a manner akin to edge detection in image process-

ing. This contribution was presented in the 55th IEEE Conference on Decision and Control (CDC) 2016 [11].

Chapter 4: In this chapter we introduce graph curvature as a notion of robustness in brain networks. Specifically, we identify key areas of the brain (nodes) which contribute more towards the overall curvature. We explain how this concept potentially relates to robustness against physical disturbances as well (injury/disease). We show by way of case studies that Ricci curvature as node measure provides useful information about brain network structure and changes due to development (healthy aging) or disease (Autism Spectrum Disorder).

Chapter 5: This chapter deals with multicompartment model fitting methods. Specifically, we introduce optimization methods that allow fitting complicated models and, thereby, provide detailed tissue microstructure information based on dMRI data. Thus, here, instead of diffusion tensor, complex models are used to capture diffusion of water molecules in complex tissue geometry. Fitting these highly non-linear models to dMRI data is a nontrivial non-convex problem. Prior existing methods are limited to tissue parameter estimation in only one fiber/axonal orientation, thereby limiting the power of this imaging modality to areas of the brain with single fiber orientation like the Corpus Callosum. We explore a special structure in the fitting-problem, a form of separability, and accordingly introduced a suitable generic optimization method. This method enables fitting rather complex models in multiple orientations while providing enhanced accuracy in single orientation situations. Our early work on the subject has been published in [12] while detailed study is available at [7].

While possible extensions of the geometrical ideas discussed in this thesis are given at the end of each chapter, Chapter 6 provides a summary in the form of concluding thoughts on the subject. We also envision a framework for optimal scanning protocols in identifying multi-compartment tissue models based on the optimization methods of Chapter 6.

Chapter 2

Basics of Diffusion MRI

This chapter provides a brief theoretical background on diffusion MRI (dMRI). Basic ingredients of dMRI modality are the concepts of Nuclear Magnetic Resonance (NMR) and the sensitivity of Magnetic Resonance (MR) signal to diffusion of water molecules. To explain dMRI theory, this chapter has been divided into four sections (Figure 2.1). Section 2.1 provides basics of the NMR theory, Section 2.2 describes the diffusion, Section 2.3 combines MR and diffusion basics to explain dMRI while Section 2.4 explains MR imaging using the diffusion tensor model.

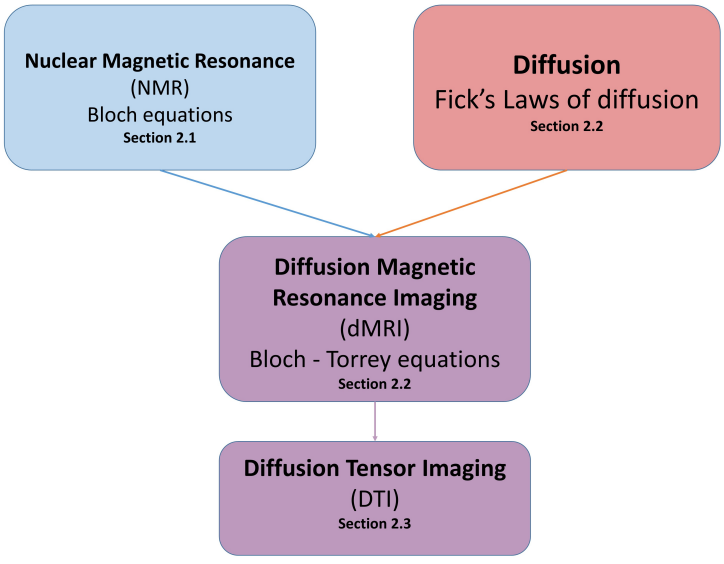


Figure 2.1: Break down of Chapter 2.

2.1 Nuclear Magnetic Resonance (NMR)

NMR is a spectroscopic technique to obtain microscopic chemical and physical information about molecules. It is based on the fact that atomic nuclei, when placed in an external field absorb electromagnetic radiations and then re-emit the radiations when relaxed back to the original state [13]. As it measures the re-emitted radiations, it is called the Magnetic Resonance.

2.1.1 Nuclear spin

Nuclear spin is an intrinsic property of the electrical charge or mass i.e, protons, electrons and neutrons, all possess spin. Also spinning protons act like little magnets. The faster it moves, the larger the magnetic field it produces. Normally, the direction that these tiny magnets point in, is randomly distributed as shown in Figure 2.2.

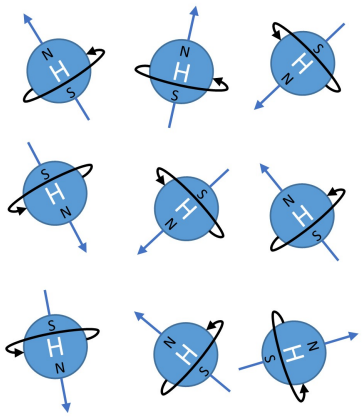
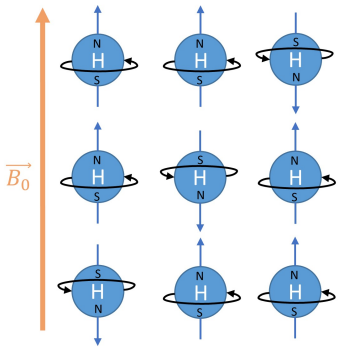


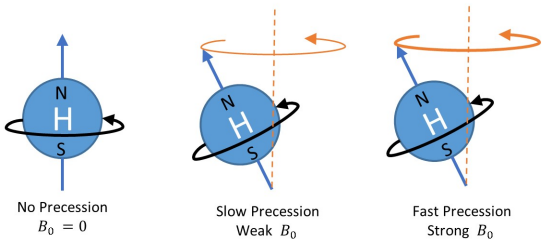
Figure 2.2: Randomly oriented spinning protons.

2.1.2 Nuclear spin in an external magnetic field B_0

When spinning protons are placed within an external magnetic field B_0 of strength Tesla (T), they align with it. However, at the atomic level, not all the protons align with the field. Some protons align against the field canceling out each other. Overall more will align with the field giving the net result as an alignment with the external field B_0 as shown in Figure 2.3a.



(a) Spinning protons aligned with B_0



(b) B_0 causing precession

Figure 2.3: Precession and alignment with B_0

2.1.3 Magnetic Resonance

In MR scanner, tissue is placed in an external magnetic field, B_0. After some time the magnetic moments (represented by net magnetization vector M_0) will reach an equilibrium favoring the direction of B_0. The idea of resonance is transmitting and receiving energy at certain specific frequency, known as the resonant frequency or Larmor frequency, ω (Hz).

$$\omega_0 = \gamma B_0 \tag{2.1}$$

Where γ is the magnetic moment of the target nucleus. Normally ^{1}H is the choice due to its natural (99.985%) and biological abundance (i.e. the fraction of one type of atom in the human body, 63%). γ for hydrogen nucleus is 42.58 $MHz.T^{-1}$. The Larmor or resonant frequency is proportional to B_0 (Figure 2.3b).

2.1.3.1 Coordinate system convention

By convention, the z-axis is aligned with external magnetic field, B_0. And a Cartesian basis $(\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$ is introduced such that $r = x\overrightarrow{i} + y\overrightarrow{j} + z\overrightarrow{k}$ and the z-axis is aligned with $\overrightarrow{k}$ (Figure 2.4).

2.1.3.2 Excitation Phase

When an RF pulse is applied at the Larmor frequency, the protons can absorb the energy. In the laboratory frame of reference, the magnetization vector M_0 spirals down towards the XY plane while in the rotating frame of reference M_0 would seem to tip down smoothly. The tip angle α depends on the strength and duration of the

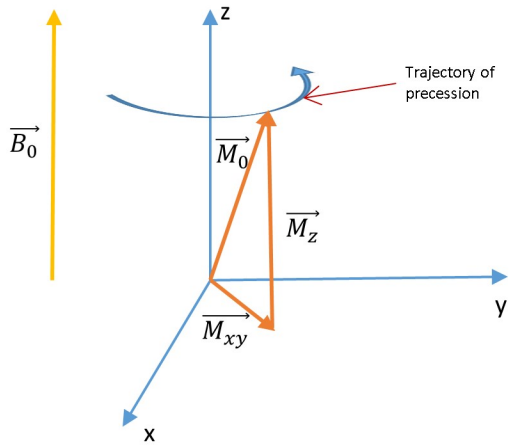


Figure 2.4: Net magnetization vector and its components.

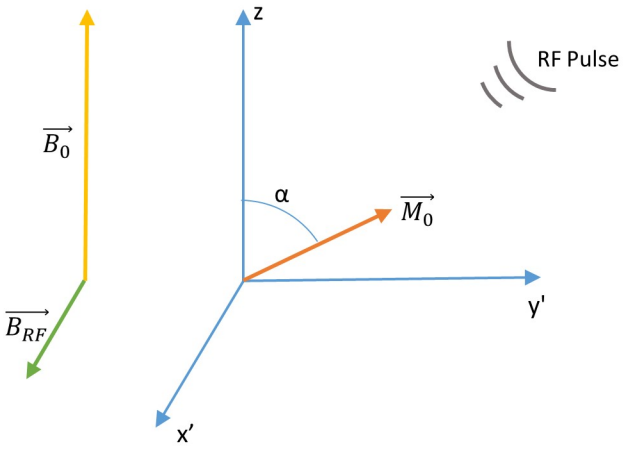


Figure 2.5: RF pulse tips M_0 to $X'\,Y'$ plane (of rotating frame) with an angle α.

RF pulse (Figure 2.5).

2.1.3.3 Relaxation Phase

After RF pulse transmission is removed, particles return to their original state. While doing so, the excess energy stored is transferred to the surrounding particles and is detected by MRI scanner coil. Following two types of relaxation processes are observed:

1. T_1 process (spin-lattice/ longitudinal component of magnetization relaxation): Each proton has to exchange its stored energy due to RF pulse, with the lattice

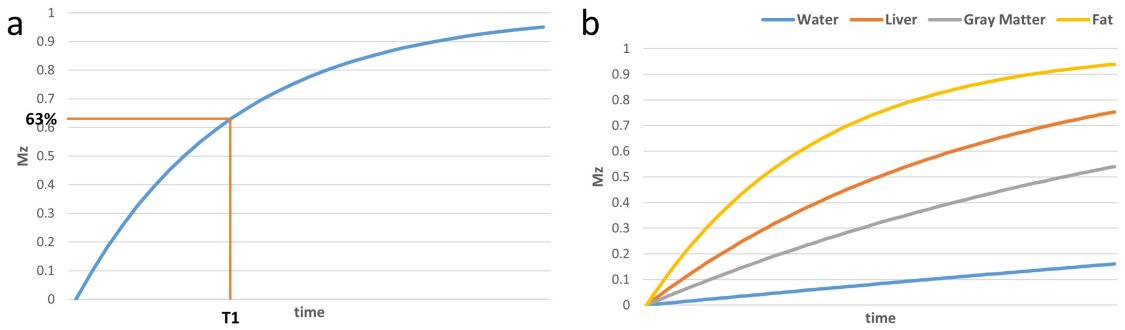


Figure 2.6: T_1 relaxation.

in order to move back parallel to B_0. The efficient energy exchange happens at the resonant (Larmor) frequency ω of the protons, similar to the RF pulse sent by the scanner. The system returns to equilibrium, or z-component of the magnetization vector M_z grows to M_0 as given in (2.2). Tissues with a fast T1 will have recovered the most signal in specified recovery time (TR), so they will have the brightest signal and vice versa, as shown in Figure 2.6b. The time constant T1 is the period for M_z to recover 63 % of its equilibrium value (Figure 2.6a).

$$M_z = M_0(1 - e^{-t/T_1}). \tag{2.2}$$

2. T_2 process (spin-spin/ transverse component of relaxation or de-phasing):
Spin-spin relaxation is the net magnetization loss due to protons dephasing. The spinning protons are all in phase after being tilted down to the XY plane by RF pulse. As some protons spin faster than the others, dephasing occurs and the signal starts decaying in the XY plane due to the loss of phase. Time constant T_2 describes the time required by M_{xy} to lose 63 percent of its value after the RF pulse is removed (Figure 2.7 a). T_2 values are related to chemical properties of the tissue and are not related to B_0 , as shown in Figure 2.7b..

$$M_{xy} = M_0\, e^{-t/T_2}. \tag{2.3}$$

2.1.3.4 Spatial localization by applying magnetic gradient G

If we apply a uniform magnetic field B_0, then all spins will oscillate at the same frequency. Therefore, no spatial localization of the frequencies will be achieved i.e.,

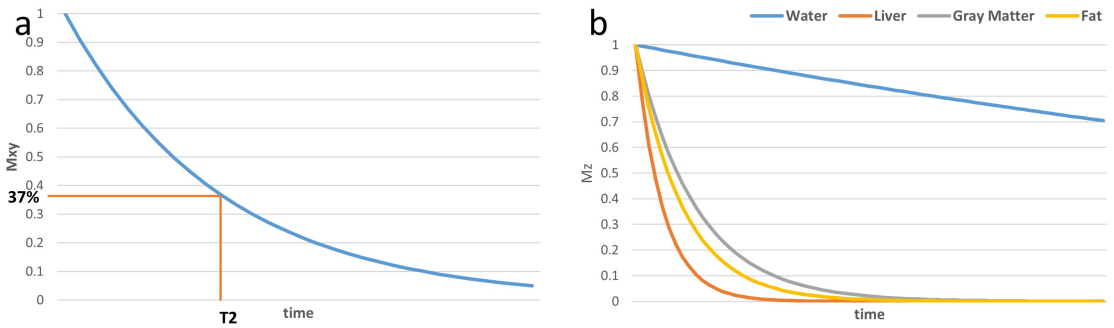


Figure 2.7: T_2 relaxation.

the RF receiver will respond to the volume as a whole, equally. To have spatial localization, we need to induce field gradient G, which should be varying with a constant gradient. This is achieved by augmenting smaller magnetic fields to B_0, with changing electrical fields. MR scanners have three sets of electrical coils to produce gradient in the magnetic fields in x, y, and z directions. Resultantly, the protons at each position in the body experience a slightly different magnetic field, as B_0 is now spatially varying i.e., $B(r)$ giving a gradient of precession frequencies.

2.1.4 Bloch equations - preliminaries

2.1.4.1 Magnetic moment

The magnetic moment is a quantity that determines the torque it will experience in an external magnetic field. Magnetic moment $\overrightarrow{\mu}$ ($N.mT^{-1}$ or $J.T^{-1}$) for spinning protons can be related to the angular momentum $\overrightarrow{J}$ ($J.sec$) as:

$$\overrightarrow{\mu} = \gamma \overrightarrow{J} \tag{2.4}$$

where γ is the is the gyromagnetic ratio ($MHz.T^{-1}$).

2.1.4.2 Torque from Larmor precession

It is the precession of the magnetic moment of any object about an external magnetic field. The external magnetic field $\overrightarrow{B}$ (T) exerts a torque τ ($N.m$) on the magnetic moment $\overrightarrow{\mu}$ ($N.mT^{-1}$ or $J.T^{-1}$) as:

$$\overrightarrow{\tau} = \overrightarrow{\mu} \times \overrightarrow{B} = \gamma \overrightarrow{J} \times \overrightarrow{B}. \tag{2.5}$$

Also torque τ is the rate of change of angular momentum i.e.,

$$\overrightarrow{\tau} = \frac{d\overrightarrow{J}}{dt}. \tag{2.6}$$

Combining (2.5) and (2.6), we get:

$$\frac{d\overrightarrow{J}}{dt} = \gamma \overrightarrow{J} \times \overrightarrow{B}. \tag{2.7}$$

Define macroscopic version ($\overrightarrow{M}$) of the microscopic magnetic moments ($\overrightarrow{\mu}$) as a vector sum of all the magnetic moments in the tissue:

$$\overrightarrow{M} = \sum_{i=1}^{N} \overrightarrow{J}_i \tag{2.8}$$

$$\frac{d\overrightarrow{M}}{dt} = \gamma \overrightarrow{M} \times \overrightarrow{B}. \tag{2.9}$$

2.1.5 Bloch equations

The Bloch equations describe macroscopic net magnetization of nuclear spin. Specifically, these calculate $M = [M_x, M_y, M_z]$ as a function of time when relaxation times T_1 and T_2 are present. Basically the equations model how the input B affects the tissue magnetization M. Let r be defined as in Section 2.1.3.1, then:

$$\overrightarrow{M}(r,t) = \begin{bmatrix} M_x(r,t) \\ M_y(r,t) \\ M_z(r,t) \end{bmatrix} \tag{2.10}$$

$$\frac{d\overrightarrow{M}(r,t)}{dt} = \overrightarrow{M} \times \gamma \overrightarrow{B} - \frac{M_x \overrightarrow{i} + M_y \overrightarrow{j}}{T_2} - \frac{M_z + M_{z0} \overrightarrow{k}}{T_1}. \tag{2.11}$$

Expanding (2.11), we get:

$$\overrightarrow{M} \times \overrightarrow{B} = \begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ M_x & M_y & M_z \\ B_x & B_y & B_z \end{vmatrix} \tag{2.12}$$

$$\begin{bmatrix} \frac{dM_x}{dt} \\ \frac{dM_y}{dt} \\ \frac{dM_z}{dt} \end{bmatrix} = \gamma \begin{bmatrix} (M_y B_z - M_z B_y) - \frac{M_x}{T_2} \\ (M_x B_z - M_z B_x) - \frac{M_y}{T_2} \\ (M_x B_y - M_y B_x) - \frac{M_z - M_{z0}}{T_1} \end{bmatrix}. \tag{2.13}$$

In matrix form

$$\begin{bmatrix} \frac{dM_x}{dt} \\ \frac{dM_y}{dt} \\ \frac{dM_z}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{1}{T_2} & \gamma B_z & -\gamma B_y \\ -\gamma B_z & -\frac{1}{T_2} & -\gamma B_x \\ \gamma B_y & -\gamma B_x & -\frac{1}{T_1} \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{bmatrix}. \tag{2.14}$$

Assume $M_{xy} = M_x + iM_y$ and $B_{xy} = B_x + iB_y$, (2.13) can be re-written as:

$$\frac{dM_{xy}}{dt} = -i\gamma(M_{xy}B_z - B_{xy}M_z) - \frac{M_{xy}}{T_2}, \tag{2.15}$$

$$\frac{dM_z}{dt} = \frac{i}{2}\gamma(M_{xy}\overline{B_{xy}} - \overline{M_{xy}}B_{xy}) - \frac{M_z - M_{z0}}{T_1}. \tag{2.16}$$

Where $\overline{B_{xy}}$ and $\overline{B_{xy}}$ are complex conjugates of B_{xy} and M_{xy} respectively. (2.15) represents transverse nuclear magnetization while (2.16) represents longitudinal nuclear magnetization.

2.2 Diffusion

Brownian motion is a rondom motion of the particles driven by thermal energy [14]. At a macroscopic level, the phenomenon is called diffusion. We know that the rate of change of concentration $\frac{\partial\psi}{\partial t}$ is equal to the net flux (J_f) across boundaries of the area of interest. That is:

$$\frac{\partial\psi}{\partial t} = -\nabla J_f. \tag{2.17}$$

From Fick's first law of diffusion along one dimension, say distance x:

$$J_f = -D\frac{\partial\psi}{\partial x}. \tag{2.18}$$

Where D is the diffusion co-efficient and related by Einstein [15,16], in one dimension as:

$$D = \frac{\Delta x^2}{2\Delta t}, \tag{2.19}$$

Δx is the mean squared displacement of the molecules in time Δt. For three dimensional case let $\Delta x = x - x_0$ where x_0 is the original position and x is the final position of a particle after diffusion for time Δt, (2.19) can be written as:

$$D = \frac{\langle \Delta x^T \Delta x \rangle}{6\Delta t}. \tag{2.20}$$

Where $\langle\rangle$ is the ensemble average. Finally from (2.17) and (2.18), we have:

$$\frac{\partial \psi}{\partial t} = -\nabla(-D\nabla\psi) = D\nabla^2\psi. \tag{2.21}$$

Note that the diffusion co-efficient D ($length^2.time^{-1}$) is a scalar constant which gives mobility of the molecules in an isotropic case and it depends on the molecule as well as the medium in which the molecule is diffusing, however, it cannot not give the direction of the diffusion.

2.3 Diffusion MRI

Effect of diffusion on spin echo was first noted by Hahn [17] in his NMR experiments. On the observations made by Hahn, Carr and Purcell [18] created spin echo sequence to measure diffusion. Torrey added diffusion term to Bloch equations i.e. using diffusion as a relaxation process [19]. Stejskal and Tanner provided solution to the Bloch Torrey equations and showed the relation of magnitude and phase of NMR signal to diffusivity [20]. In the following sections we recall the Bloch Torrey equations and derivation of the solution.

2.3.1 Sensitizing MR to water diffusion - Bloch Torrey equations

Bloch equations (2.11) with diffusion term (2.21) sensitize MR signal to water diffusion [19]. The relation is given as under:

$$\frac{d\overrightarrow{M}(r,t)}{dt} = \overrightarrow{M} \times \gamma \overrightarrow{B} - \frac{M_x \overrightarrow{i} + M_y \overrightarrow{j}}{T_2} - \frac{M_z + M_{z0} \overrightarrow{k}}{T_1} + D\nabla^2 M. \tag{2.22}$$

Where $M = M_x \overrightarrow{i} + M_y \overrightarrow{j} + M_z \overrightarrow{k}$, $M_{(r,t=0)} = M_0$.

2.3.2 Solution to the Bloch-Torrey equation

The solution to (2.22) will give expression of the MR signal attenuation due to diffusion after a 90° RF pulse [19]. To proceed, we need following assumptions:

1. A uniform magnetic field (B_0) is present in the z direction.
2. A superposed gradient field (G) is present, which is vanishing at the origin in the z direction and has axial symmetry about the z-axis (see Section 2.1.3.4). That is:
$$B_x = -\frac{1}{2}G_x, \quad B_y = -\frac{1}{2}G_y, \quad B_z = B_0 + G_z. \tag{2.23}$$

Since we are only interested in reading the transverse (XY) component of the magnetization, put (2.23) in (2.13) for M_x and M_y:

$$\frac{dM_x}{dt} = \gamma(M_y(B_0 + G_z) - M_z(-\frac{1}{2}G_y)) - \frac{M_x}{T_2} + \nabla D \nabla M_x. \tag{2.24}$$

$$\frac{dM_y}{dt} = \gamma(M_x(B_0 + G_z) - M_z(-\frac{1}{2}G_x)) - \frac{M_y}{T_2} + \nabla D \nabla M_y. \tag{2.25}$$

Let $M_{xy} = M_x + iM_y$ and $\omega_0 = \gamma B_0$ (2.1). To get expression for M_{xy}, multiply (2.25) with i and add in (2.24):

$$\frac{dM_{xy}}{dt} = \gamma\omega_0 M_y + \gamma G_z M_y + \frac{\gamma}{2}M_z G_y - \frac{M_x}{T_2} + i\omega_0 M_x + i\omega_0 M_x G_z + i\frac{\gamma}{2}M_z G_x - i\frac{M_y}{T_2} + \nabla D \nabla (M_x + iM_y). \tag{2.26}$$

$$\frac{dM_{xy}}{dt} = -i\omega_0 M_{xy} - i\gamma M_{xy} G_z - \frac{M_{xy}}{T_2} - i\frac{\gamma}{2}M_Z G(x + iy) + \nabla D \nabla (M_{xy}). \tag{2.27}$$

In the absence of diffusion, M_{xy} is exponentially damped with T_2, given by the following relation [19]:

$$M_{xy} = \varphi e^{-i\omega_0 t - t/T_2}, \tag{2.28}$$

where φ is the unattenuated amplitude of magnetization [19]. Put (2.28) in (2.27), for RHS:

$$\frac{d(\varphi e^{-i\omega_0 t - t/T_2})}{dt} = \frac{d\varphi}{dt} e^{-i\omega_0 t - t/T_2} + \varphi e^{-i\omega_0 t - t/T_2}(-i\omega_0 - \frac{1}{T_2}). \tag{2.29}$$

$$\frac{d(\varphi e^{-i\omega_0 t - t/T_2})}{dt} = \frac{d\varphi}{dt} e^{-i\omega_0 t - t/T_2} - i\omega_0 M_{xy} - \frac{M_{xy}}{T_2}. \tag{2.30}$$

(2.27) can be re-written as:

$$\frac{d\varphi}{dt} = -i\gamma G_z \varphi - i\frac{\gamma}{2} M_Z G(x + iy) e^{i\omega_0 t + t/T_2} + \nabla D \nabla(\varphi). \tag{2.31}$$

Middle term on the RHS oscillates fast and negligibly contributes to φ. Neglecting the term we can write:

$$\frac{d\varphi}{dt} = -i\gamma G_z \varphi + \nabla D \nabla(\varphi). \tag{2.32}$$

without diffusion (2.32) gives precession of φ in XY plane with phase being $-\gamma G_z t$, giving following relation (see [19] for details):

$$\varphi = M_0 A(t) e^{-i\gamma G_z t}. \tag{2.33}$$

Substituting (2.33) in (2.32), we get:

$$M_0 \frac{dA(t)}{dt} e^{-i\gamma G_z t} - (i\gamma G_z) M_0 A(t) e^{-i\gamma G_z t} = -i\gamma G_z M_0 A(t) e^{-i\gamma G_z t} + \nabla D \nabla (M_0 A(t) e^{-i\gamma G_z t}). \tag{2.34}$$

Simplifying

$$\frac{dA(t)}{dt} = A(t) e^{i\gamma G_z t} D \nabla^2 (e^{-i\gamma G_z t}). \tag{2.35}$$

$$\frac{dA(t)}{dt} = -A(t) D \gamma^2 G^2 t^2. \tag{2.36}$$

Since $A(0) = 1$, integrating above equation gives:

$$A = e^{-\frac{1}{3} D \gamma^2 G^2 t^3}. \tag{2.37}$$

Equation (2.37) gives attenuation by diffusion following a 90° pulse. Solution for pulse sequences when a 180° pulse is followed by a 90° pulse, is also given in [19]. For Pulse Gradient Spin Echo (Section 2.4.1), the solution term has been derived in [20].

2.4 Diffusion Tensor Imaging

In biological tissues, diffusion of water molecules is hindered by surrounding membranes and structures. Specifically in case of brain tissues, myelin sheath causes the aniostropic water diffusion. As described in the previous section, MR signal sensitized with diffusion can give information about intrinsic tissue structure. Using diffusion tensor model to capture the tissue structure is called Diffusion Tensor Imaging (DTI). Before proceeding further to explain MR signal attenuation model, it is important to introduce Pulse Gradient Spin Echo imaging sequence and its parameters, as given in the following section.

2.4.1 Imaging Sequence

Imaging sequence used to measure diffusion of water molecules was given by the Stejskal-Tanner imaging sequence called Pulse Gradient Spin Echo (PGSE) [20]. The sequence in shown in Figure 2.8 and the parameters of the sequence are given as under:

- G (Tesla.μm^{-1}) = Gradient strength as given in earlier sections.
- $g = [\ g_x; \quad g_y; \quad g_z\ ]$ Gradient direction (unit) vector.
- δ (seconds) = Pulse width.
- Δ (seconds) = Diffusion time.
- γ $(rad.(sec.Tesla)^{-1})$ = Gyro-metric ratio as described in Section 2.1.4.1.
- q $(rad.Tesla^{-1}) = \gamma\delta g$, displacement reciprocal vector.
- b (seconds. μm^{-2}) $= \Delta - (\delta/3)(\gamma\delta G)^2$, Diffusion weighting factor also called b-value.

2.4.2 Diffusion Tensor

Since the diffusion is anisotropic in the brain tissues, scalar diffusion coefficient as given in (2.19), which is for the isotropic case, has to be replaced by the co-variance matrix of the net displacement vectors Δx of diffusion as follows:

$$D = [D_{ij}]_{i,j=1}^{3} \ \in \ S_{++}^{3}. \tag{2.38}$$

The second order symmetric positive definite tensor as given in (2.38) can be used to model the anisotropic diffusion of water molecules in biological tissues [3]. Using Gaussian model for free diffusion of water molecule, the probability of finding a water molecule at position x, which was initially at position x_0 after time Δt can be given as under:

$$p(x|x_0, \Delta t) = \frac{1}{\sqrt{(4\pi\Delta t)^3|D|}}\, e^{\left(-\frac{(x-x_0)^T D^{-1}(x-x_0)}{4\Delta t}\right)}. \tag{2.39}$$

The MR signal attenuation $S(q, \Delta t)$ can be modeled by the following equation [21]:

$$S(q, \Delta t) = S_0 \int_{R^3} p(x|x_0, \Delta t)\; e^{(iq^T R)} dr. \tag{2.40}$$

Where S_0 is the reference signal without diffusion gradient, R denotes the spin displacement between the two pulses of PGSE and $p(x|x_0, \Delta t)$ is the ensemble-average diffusion propagator (EAP) [22, 23]. Using (2.39) as EAP in (2.40) gives expression for the MR signal [20] as under:

$$S(q, \Delta t) = S_0\; e^{(-bg^T Dg)}. \tag{2.41}$$

Equation (2.41) shows relation of MR signal sensitivity to water diffusion with b-values and D. Lower b-values can result in a low signal attenuation while higher b-values can give poor signal to noise ratio (SNR), therefore, b-values should be tuned properly. Since D has six unknown parameters, we need at least six MR measurements to estimate the diffusion tensor model. Figure 2.9 shows an example image (coronal slice) of in-vivo human brain dMRI data fitted with diffusion tensor model.

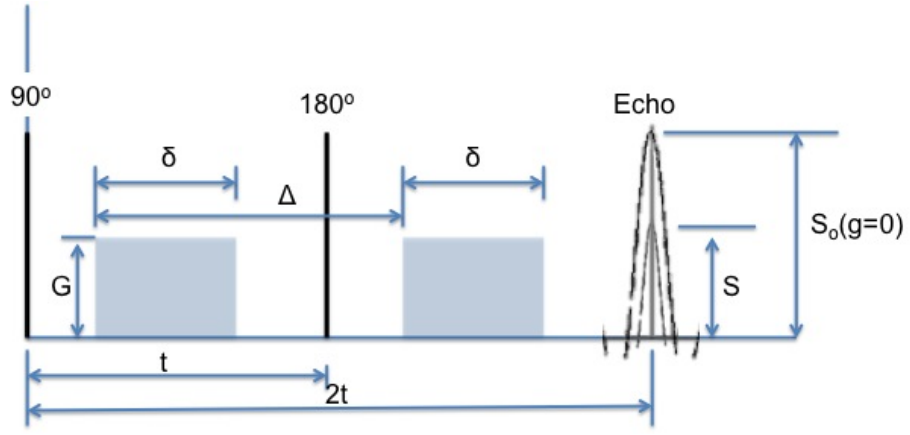


Figure 2.8: PGSE sequence

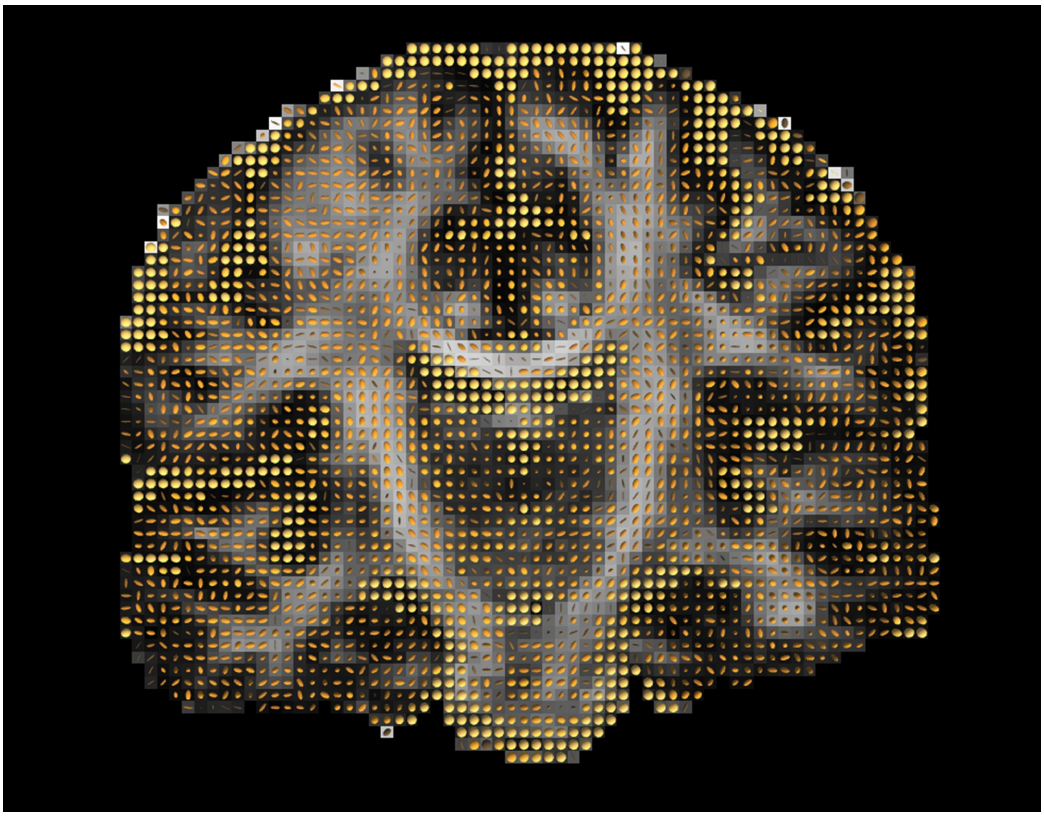


Figure 2.9: Diffusion Tensor Imaging (DTI) example image - coronal slice of in-vivo human brain data. In each voxel, a tensor (ellipsoid) represents diffusion pattern of water molecules.

2.5 Conclusion

We have explained briefly theory of dMRI by recalling basics of NMR, its sensitization to water diffusion in tissue structures and DTI. Based on DTI, many brain tissue structural analysis techniques have been found like tractography, connectivity based brain parcellations and average connectivity of brain sub-cortical region. Work presented in chapters 3 to 5 is based on developing geometrical methods to improve DTI based brain structural analysis techniques.

Chapter 3

Optimal Mass Transport for Brain Parcellation and Connectivity Mapping

Several studies have used structural connectivity information to parcellate brain areas like the corpus callosum, thalamus, substantia nigra or motor cortex, which is otherwise difficult to achieve using conventional MRI techniques. They typically employ diffusion MRI (dMRI) tractography and compare connectivity profiles from individual voxels using correlation. However, this is potentially limiting since the profile signals (e.g. probabilistic connectivity maps) have non-zero values only in restricted areas of the brain, and correlation coefficients do not fully capture differences between connectivity profiles. Our first contribution is to introduce the Wasserstein distance as a metric to compare connectivity profiles, viewed as distributions. The Wasserstein metric (also known as Optimal Mass Transport cost or, Earth Mover's distance) is natural as it allows a global comparison between probability distributions. Thereby, it relies not only on non-zero values but also takes into account their spatial pattern, which is crucial for the comparison of the brain connectivity profiles. Once a brain area is parcellated into anatomically relevant sub-regions, it is of interest to determine how voxels within each sub-region are collectively connected to the rest of the brain. The commonly used arithmetic mean of connectivity profiles fails to account for anatomical features and can easily over-emphasize spurious pathways. Therefore, our second contribution is to introduce the concept of Wasserstein barycenters of dis-

tributions, to estimate "average" connectivity profiles, and assess whether these are more representative of the neuroanatomy. We demonstrate the benefits of using the Wasserstein geometry to parcellate and "average" probabilistic tractography results from a realistic phantom dataset, as well as *in-vivo* data from the Human Connectome Project.

3.1 Introduction

The parcellation of white matter, cortical and sub-cortical brain areas, based on their connectivity information, has the potential to identify structures which are otherwise difficult to visualize e.g., parcellations of Broca's area, Substantia nigra etc. [24–27]. Connectivity-based parcellation has successfully demonstrated changes in the substantia nigra as a result of Parkinson's disease [27], and organization of Broca's area [26], and the lateral premotor cortex [25].

Previous studies employing connectivity-based parcellation can broadly be divided into two categories: The first category focuses on the clustering of streamlines generated from deterministic tractography. These methods can be supervised [28] or use unsupervised learning approaches such as spectral clustering [29]. Other unsupervised methods have addressed issues related to massive datasets [30] or metric choice [31] but, generally, streamline clustering approaches are only applicable to white matter areas. The second category leverages probabilistic information from connectivity profiles of individual voxels to aggregate them into sub-regions which share similar connectivity profiles. In this context, the choice of metric between connectivity profiles becomes crucial. Spectral and k-means clustering approaches have been used [26, 27] in combination with spatial correlation of connectivity profiles as a metric. In addition, data driven selection of the number of sub-regions has been demonstrated using independent component analysis [32] and Dirichlet processes [33]. Most of the above-mentioned studies rely on Pearson's correlation coefficients between connectivity profiles of voxels. Here, we view connectivity profiles as probability distributions over a space spanning the complete brain. This enables the use of the Wasserstein distance and techniques from the theory of Optimal Mass Transportation (OMT) to compare profiles in a mathematically rigorous fashion. Notable advantages of this new approach are: (i) The Wasserstein metric leads to improved parcellation results and enables the estimation of more anatomically meaningful "average" connectivity profiles for the estimated sub-regions. (ii) The Wasserstein metric naturally

leverages the spatial information of connectivity profiles while allowing to ignore locations with zero-value. For example, in case of high resolution data like the Human Connectome Project (HCP), the connectivity profile vector can be of the order of a million voxels while having non-zero values (or support) only at about 2 to 3 thousand voxels approximately (depending upon seed strength and threshold). OMT intrinsically allows to keep and compare only non-zero values in the profiles. Therefore, it provides improved "contrast" between the connectivity profiles, which enables outlier profiles rejection. We also introduce a Wasserstein-based Barycenter (WB) of connectivity profiles and demonstrate that this approach yields more representative results for the "average" anatomical connections. The type of barycenter that we use here is not exactly what one would consider a discrete Wasserstein barycenter, as such a choice may allocate mass at location that would be anatomically incorrect. Instead, we enforce the fact that our WB (which is a probability distribution) must have support only on the union of the support of the distributions (connectivity profiles) being averaged.

Remaining of the chapter is organized as follows. Section 3.2 introduces the Wasserstein geometry and parcellation strategy. Section 3.3 presents experiments using publicly available synthetic data (using Fiberfox [34] to replicate the structure of the FiberCup phantom[1] [35]), and Human Connectome Project (HCP) [36, 37] data for *in-vivo* experiments. Our results show that the Wasserstein geometry provides a rigorous framework for the analysis of structural connectivity profiles. In particular, we demonstrate that parcellations are better resolved in both synthetic and brain data experiments. We also demonstrate that more realistic and robust average connectivity profiles can be obtained for each estimated sub-region.

3.2 Methods

3.2.1 Transportation Cost and Average for Distributions

3.2.1.1 Optimal mass transport cost for probability distributions:

We briefly introduce key OMT concepts for brain connectivity profiles [38–40]. Let μ and ν be probability distributions (i.e. connectivity profiles from two different voxels) on the discrete space $\mathscr{X}$. The brain areas identified as possibly connected with each of the two voxels, from dMRI tractography. Denote the transportation

[1] https://www.nitrc.org/frs/shownotes.php?release_id=2341

cost of a unit mass from location $x_i \in \mathscr{X}$ to $x_j \in \mathscr{X}$ as $c_{i,j} \geq 0$. A transference plan $\pi_{i,j} \geq 0$ measures the amount of mass transferred from x_i to x_j. In the OMT problem, one seeks the optimal transference plan that minimizes the total cost. This can be formulated (due to Monge-Kantorovich) as

$$\begin{aligned} &\min_{\pi} \sum_{i,j} c_{i,j}\pi_{i,j} \\ \text{such that} \quad &\sum_{j} \pi_{i,j} = \mu_i, \quad \forall i = 1, \ldots, |\mathscr{X}| \\ &\sum_{i} \pi_{i,j} = \nu_j, \quad \forall j = 1, \ldots, |\mathscr{X}| \\ &\pi_{i,j} \geq 0 \quad \forall i, j. \end{aligned} \tag{3.1}$$

In matrix form, the above problem in (3.1) can be expressed as:

$$\min_{\Pi \in \mathscr{M}_{(\mu,\nu)}} trace(C^T \Pi) \tag{3.2}$$

with

$$\mathscr{M}_{(\mu,\nu)} := \left\{ \Pi \mid \Pi \mathbb{1} = \mu, \Pi^T \mathbb{1} = \nu, \Pi \geq 0 \right\}.$$

Here, Π and C are matrices associated with π and c, i.e., $\Pi = [\pi(i,j)]$ and $C = [c(i,j)]$, and $\mathbb{1}$ is the column matrix of ones with appropriate dimensions.

3.2.1.2 Wasserstein metric:

When the cost c is defined as $c_{i,j} = d(x_i, y_j)^p$, for any positive integer p, we can define p-Wasserstein distance [41, 42] as:

$$W_p(\mu, \nu) := \left(\min_{\Pi \in \mathscr{M}_{(\mu,\nu)}} \text{trace}(C^T \Pi) \right)^{1/p}. \tag{3.3}$$

When $p = 1$ this is also know as the Earth Mover's Distance (EMD). When $p = 2$, it is defined as 2-Wasserstein metric. Where:

$$C = [c(i,j)], \quad \Pi = [\pi(i,j)].$$

3.2.1.3 Geometric mean of distributions (Wasserstein barycenter):

To compute the average connectivity profile of a given sub-region (e.g. genu of corpus callosum or a thalamic nucleus), one can simply take the arithmetic mean of the connectivity profiles of all the voxels in the sub-region. Although the arithmetic mean will capture the approximate overall connectivity profile for that region, it can easily be influenced by outliers, and does not generate results guaranteed to be distributions, thereby rendering them potentially anatomically inaccurate. Given a set of N distributions μ_k, $k = 1 \dots N$, the problem of finding the geometric mean (Wasserstein barycenter) of the distributions is formulated as

$$\begin{aligned} &\min_{\pi} \sum_{k} \sum_{i,j} c_{i,j} \pi^k_{i,j} \\ \text{subject to} \quad &\sum_{i} \pi^k_{i,j} = \mu_{k,j}, \qquad k = 1 \dots N, \quad \forall j \\ &\sum_{j} \pi^1_{i,j} = \sum_{j} \pi^2_{i,j} = \dots = \sum_{j} \pi^N_{i,j}, \quad \forall i \\ &\pi^k_{i,j} \geq 0 \quad \forall i, j, k. \end{aligned} \tag{3.4}$$

The geometric mean is then defined as

$$\nu_i = \textstyle\sum_j \pi^1_{i,j}. \tag{3.5}$$

We should note that geometric mean given in (3.5) is different from the traditional Wasserstein barycenter (WB), as we force the support of ν to lie inside the union of the supports of the distributions μ_k. This is important for brain connectivity profiles to ensure that the resulting WB describes a white matter pathway which optimally overlaps with all the profiles being averaged. In addition, for distributions with large supports as in case of brain data ($|\mathscr{X}|$ in the order of $10,000$ voxels), the linear programing problem (3.4) can be efficiently solved using Entropic Regularization [43].

Figure 3.1 shows an example of the differences between arithmetic and geometric means of two arbitrary distributions. It can be observed that the arithmetic mean distribution (green) is bi-modal (even though the original distributions are not), while the geometric mean distribution (red) stays uni-modal, and its support is located geometrically half-way between the original distributions' supports.

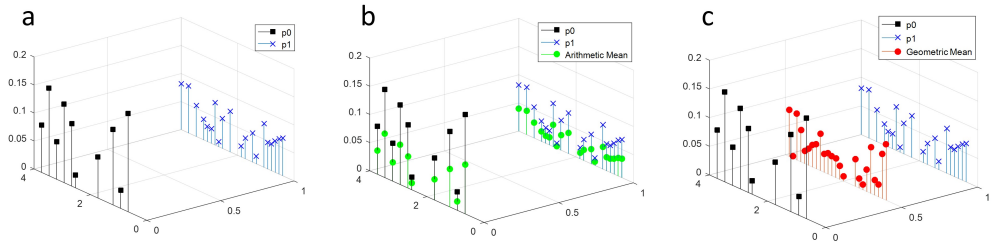


Figure 3.1: Arithmetic and geometric (Wasserstein barycenter) means of distributions. (a) Two distributions $p0$ and $p1$ (with support at $x = 0$ and $x = 1$ respectively). (b) Arithmetic mean (green) (c) Geometric mean (red).

3.2.2 Computing Geometric mean of distributions using Entropic Regularization

Consider an undirected graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ with n nodes and m edges. For each edge $e_k \in \mathscr{E}$, we attach a positive weight d_k to represent the transportation cost of moving unit mass across that edge. The cost of a path is the sum of the costs of each edge along the path and the cost of moving unit mass from node $v_i \in \mathscr{V}$ to v_j, denoted by c_{ij}, is the minimum of the cost of all the path connecting them. Given to distributions $\mu, \nu \in \mathbb{R}^n_+$ on this graph, the Wasserstein-1 distance between them is defined as

$$W_1(\mu, \nu) = \min_{\pi \geq 0} \sum_{i,j} c_{ij}\pi_{ij} \tag{3.6a}$$

$$\sum_j \pi_{ij} = \mu_i, \quad \sum_i \pi_{ij} = \nu_j. \tag{3.6b}$$

The barycenter of N distributions $\mu^\ell, 1 \leq \ell \leq N$ with respect to the weights $(\lambda_1, \cdots, \lambda_N)$ is the (unique) ν that minimizes the cost [1]

$$\sum_\ell \lambda_\ell W_1(\mu^\ell, \nu).$$

[1]Note: μ^ℓ represents ℓth distribution out of N distributions and does not depict power of μ.

It is straightforward to see that this can be formulated as a linear programming

$$\min_{\nu,\, \pi^\ell \geq 0} \sum_\ell \lambda_\ell \sum_{i,j} c_{ij} \pi_{ij}^\ell \tag{3.7a}$$

$$\sum_i \pi_{ij}^\ell = \mu_j^\ell, \quad \ell = 1, 2, \cdots, N \tag{3.7b}$$

$$\sum_j \pi_{ij}^\ell = \nu_i, \quad \ell = 1, 2, \cdots, N. \tag{3.7c}$$

Employing the entropic regularization scheme we arrive at the formulation

$$\min_{\nu,\, \pi^\ell \geq 0} \sum_\ell \lambda_\ell \left\{ \sum_{i,j} c_{ij} \pi_{ij}^\ell + \gamma H(\pi^\ell) \right\} \tag{3.8a}$$

$$\sum_i \pi_{ij}^\ell = \mu_j^\ell, \quad \ell = 1, 2, \cdots, N \tag{3.8b}$$

$$\sum_j \pi_{ij}^\ell = \nu_i, \quad \ell = 1, 2, \cdots, N, \tag{3.8c}$$

where γ is a scalar constant and the entropy term is

$$H(\pi) = \sum_{i,j} \pi_{ij} (\log \pi_{ij} - 1).$$

Let ξ be the matrix with entries

$$\xi_{ij} = \exp(-c_{ij}/\gamma),$$

then the solution of the regularized problem (3.8) has the structure

$$\pi^\ell = \operatorname{diag}(v^\ell) \xi \operatorname{diag}(w^\ell), \quad \nu = \prod_\ell (v^\ell \otimes \xi w^\ell)^{\lambda_\ell}.$$

And the algorithm for computing barycenter is:

$$\begin{aligned} w^\ell &= \mu^\ell \oslash (\xi^T v^\ell), \ \ell = 1, \ldots, N \\ \nu &= \prod_\ell (v^\ell \otimes \xi w^\ell)^{\lambda_\ell} \\ v^\ell &= \nu \oslash (\xi w^\ell), \ \ell = 1, \ldots, N. \end{aligned}$$

3.2.3 Connectivity-based parcellation using Wasserstein metric

We use probabilistic tractography to estimate connectivity profiles from voxels in a region of interest, and follow a standard parcellation procedure [44]:

- Connectivity profiles estimation: We generate connectivity profiles using *probtrackX* from fiber orientation estimates (*bedpostX*), using FMRIB's (FSL) Diffusion Toolbox [45]. For each voxel in the region of interest (e.g. corpus callosum, thalamus), a spatial probability distribution is computed.

- Connectivity profiles comparison: By using correlation coefficients or the Wasserstein metric, we construct a weighted undirected graph for the region of interest. In the graph, voxels become nodes, while edge weights are defined by the distance between connectivity profiles.

- Graph parcellation: We use graph spectral clustering methods (normalized graph cut [46, 47]) i.e., eigenvectors of graph Laplacian for identifying voxels sub-regions with similar connectivity profiles (with the rest of the brain).

3.3 Experiments and Results

3.3.1 Experiments using synthetic data

We used the digital replication of the FiberCup phantom [35] generated by Fiberfox [34]. This provides an ideal synthetic dataset, with complex fiber patterns, to objectively assess the performance and benefits of the Wasserstein geometry. Its dimensions are 64 x 64 x 3 with b-value 1500 s/mm^2. Following the procedure outlined

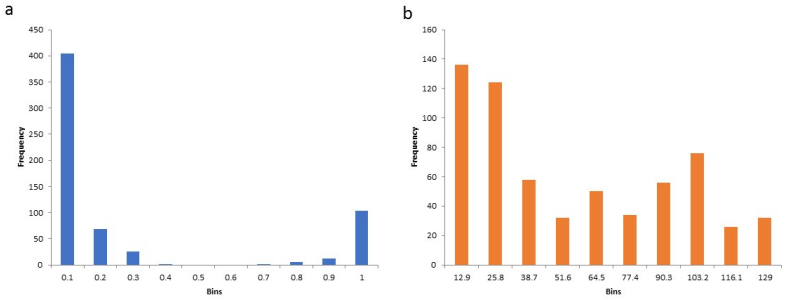


Figure 3.2: Histograms of correlation coefficients and OMT cost. (a) Absolute correlation coefficients. (b) OMT cost.

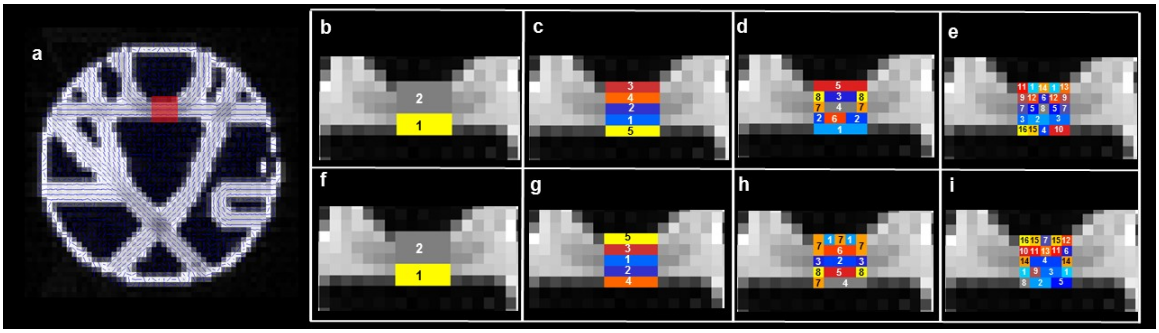


Figure 3.3: FiberCup replica parcellation in 2, 5, 8 and 16 sub-regions. (**a**) Region of interest in red, blue lines show primary fiber orientations. (**b-e**) Parcellations using Wasserstein metric. (**f-i**) Parcellations using correlation coefficients.

in Section (3.2.3), we generate graphs using correlation coefficients and Wasserstein metrics. The region of interest (ROI) is defined in the central slice at the intersection of the u-shaped bundle connecting label regions 13 and 15, and linear bundle between regions 11 and 17 (see [34] for region labels detail).

Histograms of absolute values of the correlation coefficients and the Wasserstein metric (from connectivity matrices) are shown in Figure 3.2. It can be seen that the Wasserstein metric provides increased contrast, which may help improve clustering results. We first evaluate parcellation results, as illustrated on Figure 3.3, with increasing number of sub-regions. With 2 sub-regions, both metrics identify the upper and lower parts of the ROI as separate clusters, since the gray area is primarily part of the u-shaped bundle, while the yellow area belongs to the linear bundle. When the number of sub-regions increases, Wasserstein parcellations preserve the symmetry of the ROI, in relation to the underlying bundles, while correlation parcellations identify labels more randomly. With 8 sub-regions, the top (red) and bottom (blue) clusters (panel d) are consistently labeled as two unique sub-regions with the Wasserstein metric, but randomly broken into smaller sub-regions (orange, blue and gray on panel h) with the correlation metric. A similar behavior is observed on panels e and i. Figure 3.4 illustrates the difference between geometric and arithmetic means of connectivity profiles for the two sub-regions (yellow, orange). Wasserstein barycenters are consistent with the ground truth bundles (panel b), while arithmetic means lead to spurious tracts into adjacent bundles.

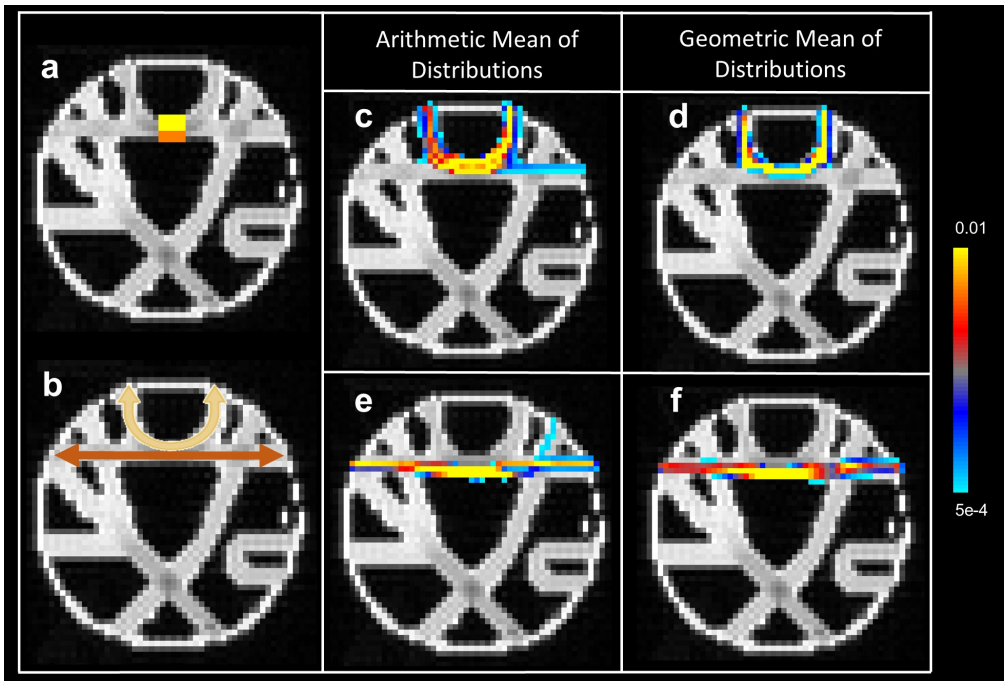


Figure 3.4: Comparison of arithmetic and geometric mean of connectivity profiles. (**a**) Sub-regions from Figure 3.3**b**. (**b**) Ground truth. (**c** and **e**) Arithmetic mean profile. (**d** and **f**) Geometric mean profile.

3.3.2 Experiments using HCP data

We used data from the Human Connectome Project [36,37] collected on a Siemens 3T Skyra system with voxel size 1.25 mm^3, three b-values (1000, 2000, 3000 s/mm^2), each with 90 directions, and 18 additional b=0 volumes.
We focus on the parcellation and connectivity mapping of the corpus callosum. Figure 3.5 illustrates parcellations results. Although Wasserstein metric and correlation lead to consistent results with 3 and 5 sub-regions, differences appear with 8 and 10 sub-regions. First we note that both approaches use one or two clusters to detect outliers voxels (mis-labeled during manual segmentation due to partial volume effects) next to the rostrum, genu and splenium. Wasserstein parcellation detects more outliers in the posterior splenium (pink area). More importantly, correlation parcellation over-estimates the size of the anterior midbody (brown on panel i), under-estimates the rostral body (light blue on panel i) and splits the genu in two sub-regions (purple and yellow on panel i). Wasserstein parcellation appears more in line with the known anatomy of the corpus callosum (see Figure 1 in [48]).

Finally, Figure 3.6 illustrates differences between arithmetic and geometric means

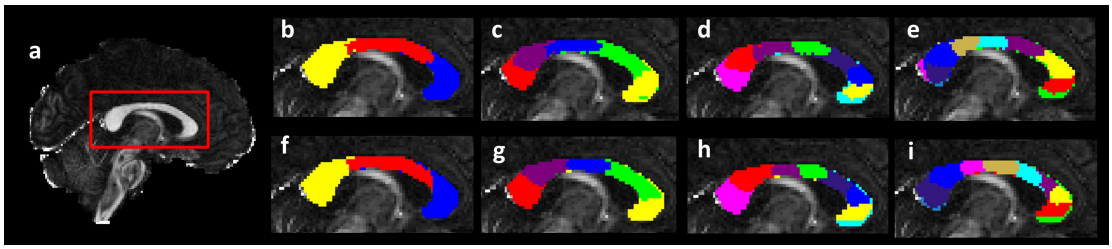


Figure 3.5: Parcellations of the Corpus Callosum. (**a**) ROI (red box). Parcellations in 3, 5, 8 and 10 sub-regions (**b-e**) using Wasserstein metric and (**f-i**) correlation coefficients.

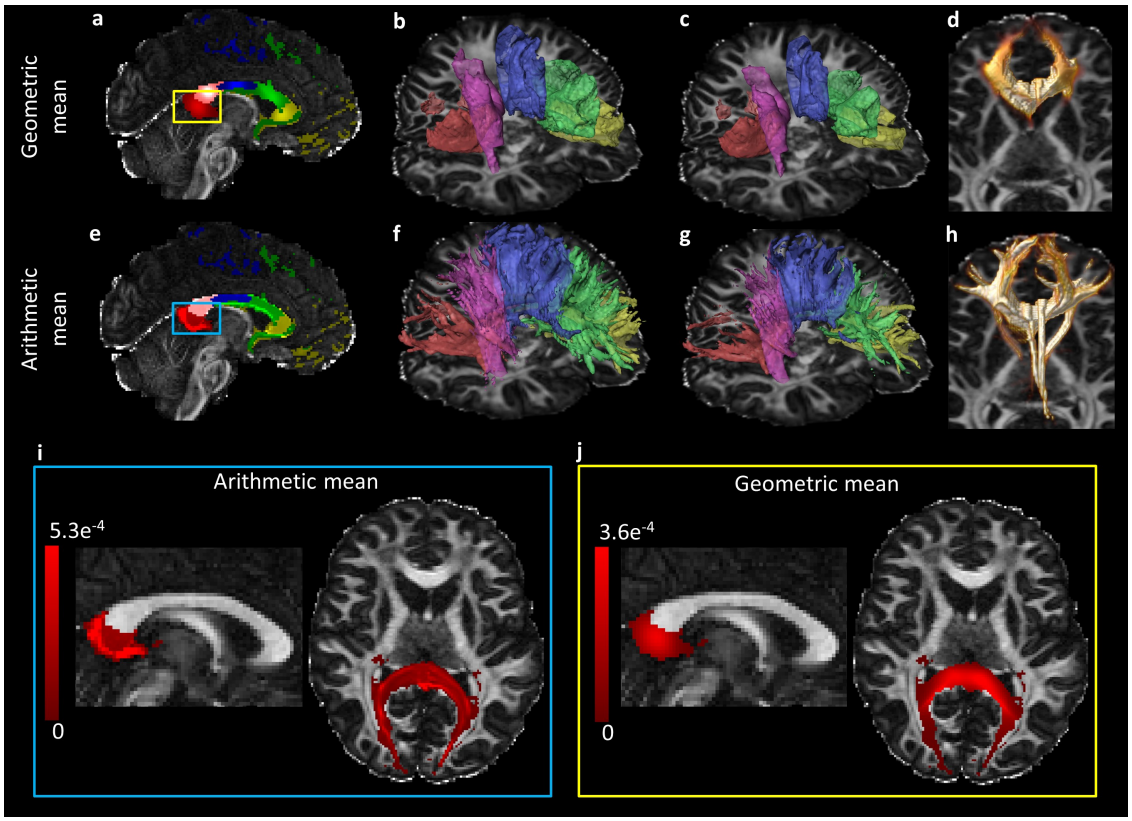


Figure 3.6: Comparison of arithmetic and geometric mean of connectivity profiles for 5 sub-regions (from Figure 3.5**g**) of the corpus callosum. Panels **a** and **e** show mean connectivity profiles in the mid-sagittal plane. Panels **b, f** and **c, g** show iso-surfaces of the connectivity distributions respectively with low/high thresholds. Volume rendering in panels d, h illustrates differences in mean connectivity profiles from the genu/rostrum. Panels **i** and **j** show detailed (zoomed in) differences between the arithmetic and geometric means in the splenium of the corpus callosum.

of connectivity profiles for five sub-regions of the corpus callosum. It can be observed that Wasserstein barycenters are much more robust to noise and spurious connections in the distributions to be averaged (see panels d, h where tracts from the cingulum

bundle and internal capsule are incorrectly visible for the arithmetic mean). Panels a and e show sagittal views of the average connectivity profiles. It can be seen that the geometric mean of all the sub-regions has higher values towards the center of each area while the arithmetic mean tends to be more noisy since it does not enforce any spatial consistency in the averaging process. For instance, panels i and j focus on the splenium of the corpus callosum. In this region, the highest values of the arithmetic average are localized in the very posterior and ventral parts. This does not capture the anatomy of this brain region as well as the geometric center of the region, which should have higher values around the center of mass of this region. Additionally, panels b, c, f, g illustrate the ability of the Wasserstein metric to maintain a consistent spatial support between the distributions and their average, thereby avoiding the creation of multi-modal mean distributions (visible as numerous "thin" and noisy tracts in panels f, g), which results in spurious pathways.

3.4 Conclusion

We have shown that the Wasserstein geometry provides a novel and powerful computational framework for the parcellation and connectivity mapping of brain areas. The Wasserstein metric is a natural similarity measure between structural connectivity profiles obtained from probabilistic tractography, which we leveraged within a graph spectral clustering algorithm to identify sub-regions of the corpus callosum and demonstrate improvements over results using correlation coefficients. Finally, this metric can be used and adapted to define the notion of Wasserstein barycenter, and estimate the geometric mean of a set of connectivity profiles, thereby characterizing the average connectivity of sub-regions of the corpus callosum, while avoiding the generation of spurious tracts, as typically observed when using the arithmetic mean.

Chapter 4

Ricci Flow for Feature Enhancement of Diffusion Tensor Fields

Diffusion Tensor Imaging (DTI) generates a 3-dimensional 2-tensor field that encapsulates properties of diffusing water molecules. We present two complementing ideas that may be used to enhance and highlight geometric features that are present. The first is based on Ricci flow and can be understood as a *nonlinear bandpass filtering technique* that takes into account directionality of the spectral content. More specifically, we view the data as a Riemannian metric and, in manner reminiscent to reversing the heat equation, we regularize the Ricci flow so as to taper off the growth of the higher-frequency speckle-type of irregularities. The second approach, in which we again view data as defining a Riemannian structure, relies on averaging nearby values of the tensor field by weighing the summands in a manner which is inversely proportional to their corresponding distances of the tensors. The effect of this particular averaging is to enhance *consensus* among neighboring cells, regarding the principle directions and the values of the corresponding eigenvalues of the tensor field. This consensus is amplified along directions where distances in the Riemannian metric are short.

DTI is a magnetic resonance imaging technique that allows mapping the diffusion properties of water molecules (see Section 2.4.2 for details). Thereby, it provides valuable information about pathways and other microstructures that restrict or hinder

diffusion. It also has a great potential to study neurological disorders which affect the brain microstructure and connectivity. The data produced encodes the intensity and directionality of diffusion in the form of an ellipsoid at every unit volume element (voxel). DTI data can be viewed as either a manifold of multivariable normal distributions or, alternatively, as a Riemannian metric. The latter representation allows geometric techniques and insights on how to process such data and extract features. The work presented here is along the lines of series of earlier studies with a similar point of view, see e.g., [49–51] and the references therein.

In this chapter, we consider the data as a position-dependent positive-definite quadratic form, i.e., as a Riemannian metric. The local alignment of principle directions of this matrix, across voxels, may signify the presence of a fiber pathway whereas abrupt discrepancies may characterize the presence of pathways with different orientations. Thus, information about tissue microstructures is encoded in the spatial correlation between the values of the metric. On the other hand, speckle fluctuations of the parameters that specify principle directions and their corresponding eigenvalues, are the inevitable consequence of measurement noise. Naturally, at places, noise masks key features and limits resolution. Thus, the problem we wish to address is on how to suppress noise while maintaining structural information in the tensorial field that is available to us. We pursue and study to complementing approaches that we outline next.

The first approach is based on the Ricci flow [52]. The Ricci flow is a geometric (non-linear) evolution on a Riemannian manifold that deforms the metric and, for manifolds with positive Ricci curvature, smooths out irregularities. Thus, formally, it is analogous to the heat equation that smooths out irregularities of, e.g., a temperature distribution. In this sense, it represents a low-pass filtering for the 3-dimensional 2-tensorial field (in continuous or a discrete space). Depending on the quality of the data, our aim may be to either smooth out irregularities, or to amplify and highlight structural features. Here, we focus on the latter. To this end, we reverse time in the Ricci flow. In a way that echoes solving backwards heat equation, the flow rapidly amplifies "high frequency" spacial irregularities. The backwards heat equation, studied for inverse heat conduction problems, requires some form of regularization [53]. Likewise, for the backwards Ricci flow we introduce a suitable regularization that tapers the fastest modes. The result appears to produce a higher quality contrast and accentuate structures that are dominant in the data.

The second approach is again based on interpreting the data as a Riemannian metric. In this, we estimate distance between neighboring voxels in the induced geometry (by the data/ metric itself). We then average local values of the metric with a weighing coefficient which is inversely proportional to the distance between voxels. The process represents a geometric diffusion that allows nearby voxels to reach some form of consensus more easily than ones that are far away. As a result, dominant structures such as fibers, where lines of voxels share similar values for the metric, and where their corresponding dominant eigenvectors are lined up with the fiber, are made consistent in that small irregularities quickly dissipate. Thus, again, the approach represents a nonlinear averaging/ filtering that is designed to accentuate correlations between nearby and similar voxels. The method is exceptionally fast. The result will also be exemplified on case studies using DTI brain data, and appears to consistently produce the desired effect.

Either of the methods we present, can be seen as an anisotropic diffusion of the tensorial field itself. The anisotropy is intended to preserve the key features of the (already anisotropic) tensorial field. Thus, the flow (in the first method) and averaging (in the second) are intended to respect the underlying anisotropy of the field.

The remaining chapter is structured as follows. In Section 4.1 we describe notations and basic concepts. Section 4.2 gives theoretical details of techniques for feature enhancement and smoothing in the tensor fields. Section 4.3 gives algorithmic form of the techniques discussed in Section 4.2. While in Section 4.4, we show examples by implementing the suggested methods to synthetically generated tensor fields as well as on MRI data.

4.1 Notations & Preliminaries

Our methods are based on the concepts from Riemannian geometry. We shall follow standard notations in [54–56] and give a brief overview.

A *Riemannian metric* on a smooth manifold is an inner product g on the tangent space at each point, varying smoothly from point to point. Riemannian manifolds are smooth manifolds equipped with a Riemannian metric. Locally, in a neighborhood of any point, it can be viewed as a weighted Euclidean space with weight given by g.

To define derivatives of vector fields on a Riemannian manifold, the concept of

connection is introduced. Intuitively, it provides a way to "connect" nearby tangent spaces. The Levi-Civita [57] connection is the special one which is symmetric and compatible with g. It is specified by the Christoffel symbols Γ.

Another key ingredient of Riemannian geometry is curvature. There are several different notions of curvature (like Sectional, Gaussian, Ricci and Scalar). The one we need for this work is *Ricci curvature* [58], which we denote by κ. It is a 2-tensor field on the manifold uniquely defined by the Riemannian metric g.

For our purpose, we are interested in a 3 dimensional Riemannian manifold. More specifically, we view the DTI as a 3 dimensional manifold where the diffusion tensor (please see section 2.4.2)

$$D = [D_{ij}]_{i,j=1}^{3} \ \in \ S_{++}^{3}$$

at each point $x = (x_1, x_2, x_3)$ (or its inverse) is treated as the Riemannian metric at that point. Here S_{++}^{3} denotes the space of 3×3 positive definite matrices.

In DTI, one important quantity is *Fractional Anisotropy (FA)*. For a positive definite 3×3 symmetric matrix, with $\lambda_1, \lambda_2, \lambda_3$ as its eigenvalues arranged in descending order, FA is defined by [5]

$$\sqrt{\frac{3((\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2)}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}}. \tag{4.1}$$

FA captures some structural information of a diffusion tensor. A large FA indicates anisotropic diffusion, which can be related to a locally linear and well-organized structure. Conversely, a low FA characterizes a less organized structure, or the presence of multiple fiber pathways.

4.2 Smoothing and Feature Enhancement Techniques

We view DTI data D as a 3 dimensional Riemannian manifold with metric specified by the diffusion tensors. From this point of view, smoothing the diffusion tensors is equivalent to smoothing the metric of the underlying manifold. Two related smoothing techniques are presented here, first based on the idea of Ricci flow and second based on anisotropic smoothing of the tensor field.

4.2.1 Ricci Flow for Tensor Fields

Given the Riemannian manifold with metric tensor $g = [g_{ij}]$ specified by the diffusion tensors, we can compute the Ricci curvature

$$\kappa = [\kappa_{ij}]_{i,j=1}^{3}.$$

Both the Riemannian metric g and the Ricci curvature κ are 2-tensor fields on the manifold. The idea of Ricci flow [58] is to adjust the metric g using the Ricci curvature κ to smooth out irregularities in the metric. Formally, the Ricci flow is defined by the nonlinear evolution equation [58, 59]:

$$\partial_t g_{ij} = -2\kappa_{ij}. \tag{4.2}$$

Now we outline the steps to compute κ from g. First we compute the inverse of the metric g, this is given by

$$[g^{ij}] = [g_{ij}]^{-1}. \tag{4.3}$$

With $[g^{ij}]$ and $[g_{ij}]$ we can the calculate the Christoffel symbols [60] as:

$$\Gamma_{ij}^{k} = \sum_{\ell=1}^{3} \frac{1}{2} g^{kl}(\partial_i g_{j\ell} + \partial_j g_{i\ell} - \partial_\ell g_{ij}) \tag{4.4}$$

for all $i, j, k = 1, 2, 3$. Here ∂_i denotes the directional derivative in x_i direction. In matrix form, (4.4) can be written as:

$$\begin{bmatrix} \Gamma_{ij}^{1} \\ \Gamma_{ij}^{2} \\ \Gamma_{ij}^{3} \end{bmatrix} = \begin{bmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{bmatrix} \begin{bmatrix} \partial_i g_{j1} + \partial_j g_{i1} - \partial_1 g_{ij} \\ \partial_i g_{j2} + \partial_j g_{i2} - \partial_2 g_{ij} \\ \partial_i g_{j3} + \partial_j g_{i3} - \partial_3 g_{ij} \end{bmatrix} \tag{4.5}$$

for all $i, j = 1, 2, 3$. After obtaining the Christoffel symbols, we can readily compute the Ricci curvature tensor κ as

$$\kappa_{\alpha\beta} = \sum_{\rho,\lambda=1}^{3} \partial_\rho \Gamma_{\beta\alpha}^{\rho} - \partial_\beta \Gamma_{\rho\alpha}^{\rho} + \Gamma_{\rho\lambda}^{\rho} \Gamma_{\beta\alpha}^{\lambda} - \Gamma_{\beta\lambda}^{\rho} \Gamma_{\rho\alpha}^{\lambda} \tag{4.6}$$

for all $\alpha, \beta = 1, 2, 3$.

In this approach, the Riemannian metric g is defined by DTI data, namely, $g = D$.

Applying the Ricci flow framework, as explained above, on this manifold will result in an evolution of the underlying metric, therefore the diffusion tensor D.

4.2.2 Anisotropic Smoothing for Tensor Fields

In this approach, unlike the previous case, we treat DTI as a 3 dimensional Riemannian manifold with metric given by the inverse of the diffusion tensor, i.e., $g = D^{-1}$, instead of the diffusion tensor itself [51]. And we smooth the tensor field by averaging nearby values with weights inversely proportional to their distances. The intuition behind this approach is that the useful structures in DTI datasets are usually captured by the principle directions of the diffusion tensors. By taking weights inversely proportional to distances, the effect in the principle direction is emphasized. To implement this framework, we first compute the distance between any point $x = (x_1, x_2, x_3)$ and its m neighbors as:

$$d_k = \frac{1}{v_k^T D_x^{-1} v_k}, \qquad 1 \leq k \leq m, \tag{4.7}$$

where v_k is the coordinate vector between x and its kth neighbor, and D_x is the diffusion tensor at x. More specifically, at any point x, v_k gives the direction vector for the respective neighboring D_x^k. After obtaining these distances, we can then update the diffusion tensor at any point based on the weighted average of its neighbors as follows:

$$D_x = \alpha D_x + (1-\alpha)\Sigma_{k=1}^{m} \frac{d_k}{\Sigma_{l=1}^{m} d_l} D_x^k. \tag{4.8}$$

Here D_x^k is the value of diffusion tensor of the kth neighbor of x and $\alpha \in [\frac{1}{2}, 1]$ is an adjustable parameter specifying weight of the average (weighted) of the neighboring tensors.

4.3 Numerical Implementation

In this section, we provide detailed implementation of the Ricci flow and the anisotropic smoothing methods in Algorithm 1 and 2 respectively. Both the algorithms were implemented in MATLAB to solve example problems presented in Section 4.4.

Algorithm 1 Ricci Flow Smoothing for Tensor Fields

Require: g_{ij} (Riemannian metric tensor, diffusion tensor). n (number of iterations).

1: **for** l:=1 **to** n **do**
2: Compute inverse Riemannian curvature tensor g^{ij} by taking inverse of g_{ij}.
3: Compute three Christoffel symbol matrices by (4.4).
4: Calculate Ricci curvature matrix κ by (4.6).
5: Choose step size Δ.
6: Compute g_{ij} using (4.9) for backward Ricci flow or by (4.10) for forward Ricci flow.

$$g_{ij}(t + \Delta t) = g_{ij}(t) + 2\Delta\kappa_{ij}(t) \tag{4.9}$$

$$g_{ij}(t + \Delta t) = g_{ij}(t) - 2\Delta\kappa_{ij}(t) \tag{4.10}$$

7: update $g = g_n$
8: **end for**
9: Noise Suppression

$$g(\frac{n}{2}) = g(\frac{n}{2}) - \lambda_{max(g(n))}g(n) \tag{4.11}$$

Algorithm 2 Anisotropic Smoothing for Tensor Fields

Require: D (Diffusion tensor).
n (number of iterations).
$\alpha \in [0.5,\ 1]$
v (unique distance vectors from D to each neighboring tensor).

1: **for** l:=1 **to** n **do**
2: Compute inverse tensor matrix D^{-1}.
3: Compute d by (4.7).
4: Update D using (4.8).
5: **end for**

4.4 Examples

4.4.1 Synthetic Data Experiments

Two substrates of synthetic tensor fields were generated to experiment smoothing and feature enhancement. For simplicity, results have been shown for only two dimensions. Figure 4.1 (panel a) and Figure 4.2 (panel a) show the generated synthetic tensor fields. Both the plots were generated using MATLAB function plotDTI [61].

4.4.1.1 Backward Ricci Flow - Feature Enhancement

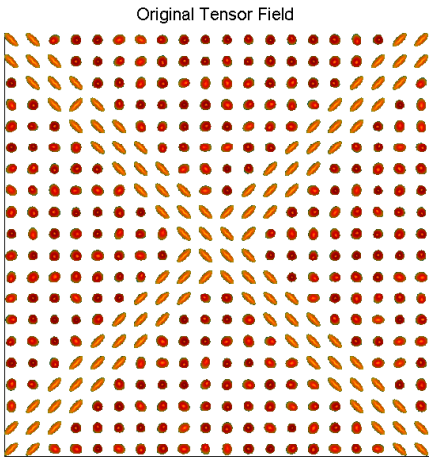
Performing as a non-linear inverse "heat like" operation using (4.9), underlying geometric features were enhanced in the field. Figure 4.1 (panel b) and Figure 4.2 (panel b) show the enhanced features as compared to original tensor fields in Figure 4.1 (panel a) and Figure 4.2 (panel a) respectively.

4.4.1.2 Forward Ricci Flow - Tensor Smoothing

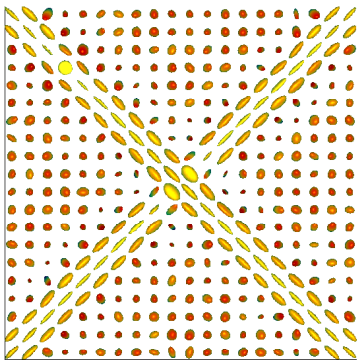
Forward Ricci flow as given in (4.10) performs tensor field evolution resulting in overall smoothing of the field. Figure 4.1 (panel c) and Figure 4.2 (panel c) were obtained after smoothing Figure 4.1 (panel b) and Figure 4.2 (panel b) respectively. Both forward and backward Ricci flow operations are reversible in this setting under suitable constraints on the Ricci tensor. This can be seen as we obtain Figure 4.1 (panel c) and Figure 4.2 (panel c) which are very close to the original tensor fields in Figure 4.1 (panel a) and Figure 4.2 (panel a), after backward and then forward Ricci flow.

4.4.1.3 Anisotropic Smoothing of the Tensor Field

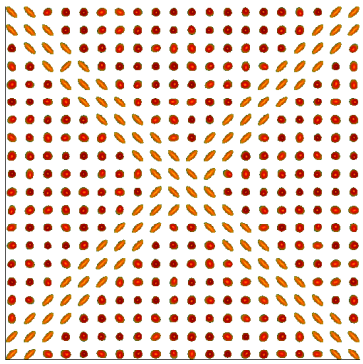
Anisotropic smoothing method suggested as Algorithm 2 requires very few iterations to smooth a given tensor field. Figure 4.3 shows smoothing results after four iterations.



(a) Original Tensor Field



(b) Backward Ricci Flow



(c) Forward Ricci Flow

Figure 4.1: Experiment with synthetic data set 1, (a) Original tensor field (b) Feature extraction after backward Ricci flow (c) Smoothing after forward Ricci flow.

4.4.2 MRI Data Experiments

4.4.2.1 Ricci Flow

A healthy volunteer was scanned at the Center for Magnetic Resonance Research (CMRR), University of Minnesota, to obtain DTI data. Figure 4.4 (panel a) shows an FA map for a coronal view of the brain. Region of interest (including the thalamus and part of the internal capsule) has been highlighted in red. Figure 4.4 (panel b) shows g_n which enhances features but overall leads to a more noisy image (given by Eq. 4.11). Figure 4.4 (panel c) shows $g(n/2)$ and Figure 4.4 (panel d) shows $g(3n/4)$.

It can be seen that both Figure 4.4 (panel c) and Figure 4.4 (panel d) show higher FA in the white matter (which appears in white on these images) while FA remains almost unchanged in the other areas. This demonstrates the possibility to perform noise suppression and feature enhancement, as expected. For this example, forty iterations (n) were used.

4.4.2.2 Anisotropic Smoothing

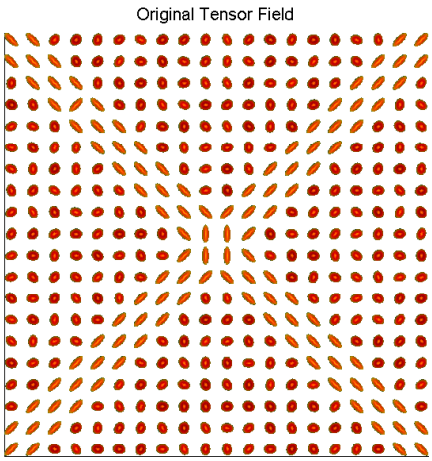
Using the same MRI data as in the previous experiment, anisotropic smoothing was performed. Figure 4.4 (panel a) shows the original tensor field in a coronal slice, while Figure 4.4 (panel b) shows the smoothed tensor field after only four iterations of the algorithm.

4.5 Conclusion and Discussion

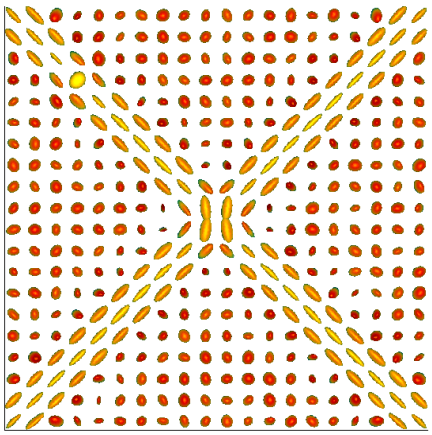
The intend of our study has been to develop tools for de-noising as well as for feature enhancement of tensorial fields, as these arise in DTI. As noted, DTI is based on modeling the diffusion of water molecules by a Gaussian process inside voxels. The covariance matrix for the anisotropic diffusion of water molecules is represented by the diffusion tensor (symmetric positive-definite 3×3 matrix). These tensor fields are often pre-processed to reduce the amount of noise arising in the acquisition process. With our presented schemes, de-noising of tensor fields can be done quite efficiently, without the need for any delicate tuning. Thus, it holds the potential of providing a useful tool for DTI data analysis.

In analogy to the stochastic framework of *Schrödinger bridges* [62–65] where a path ρ_t is constructed to interpolate two end-point in time density functions–a stochastic control problem of steering the Fokker-Planck equation, it will be interesting to control the flow of the Ricci curvature so as to interpolate, in a similar manner, tensorial distributions. That is, given two DTI tensor fields, construct a suitable homotopy/deformation linking the two. Such a problem represents a far reaching generalization of the analogous problems of controllability involving the heat equation. The value of such interpolation schemes would be to integrate DTI records across time in a way that reflects changes in the underlying tissue structure. Likewise, it will be interesting to consider the geometry of Ricci flow for the purpose of extrapolating DTI data. In this direction, it would be important to investigate

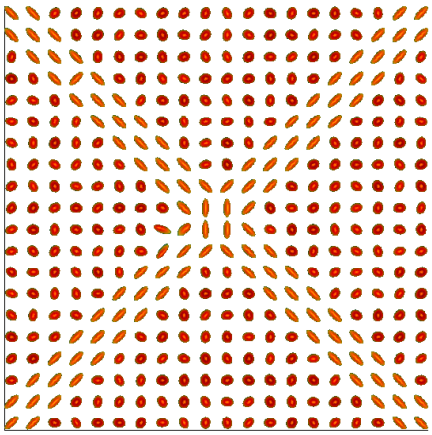
possible regularization terms that may preserve and capture structural features as these evolve (e.g., in the spirit of [66]). This extension may help in learning tensor field evolution (tissue structural development) in time, which can be of significance in clinical/neuroscience applications.



(a) Original Tensor Field

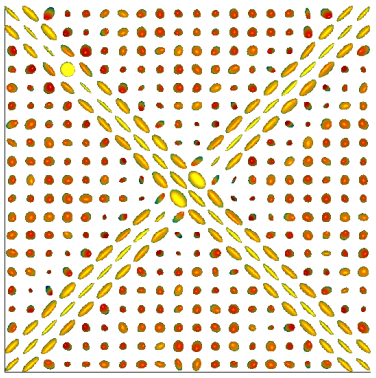


(b) Backward Ricci Flow

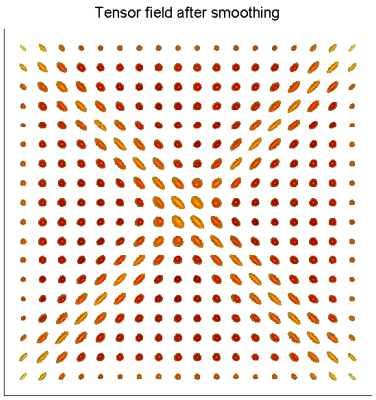


(c) Forward Ricci Flow

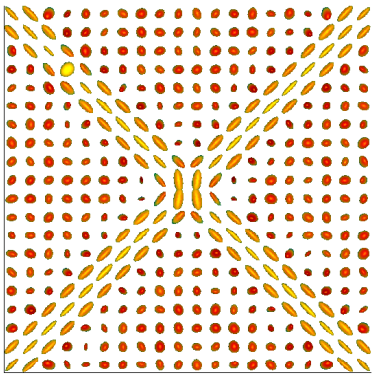
Figure 4.2: Experiment with synthetic data set 2, (a) Original tensor field (b) Feature extraction after backward Ricci flow (c) Smoothing after forward Ricci flow.



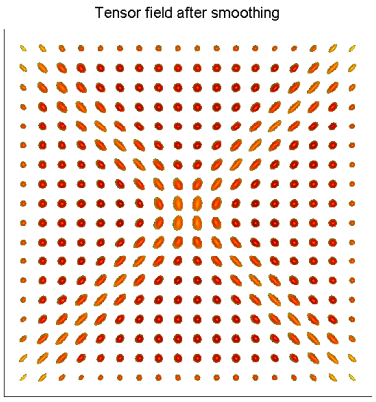
(a) Original Tensor Field



(b) After Smoothing



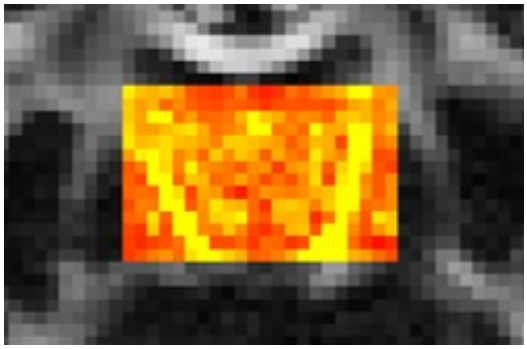
(c) Original Tensor Field



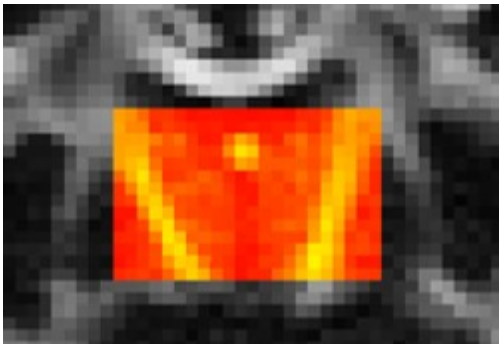
(d) After Smoothing

Figure 4.3: Smoothing experiment with synthetic data set 1 and 2, (a) and (c) show original tensor field, while (b) and (d) show tensor fields after smoothing.

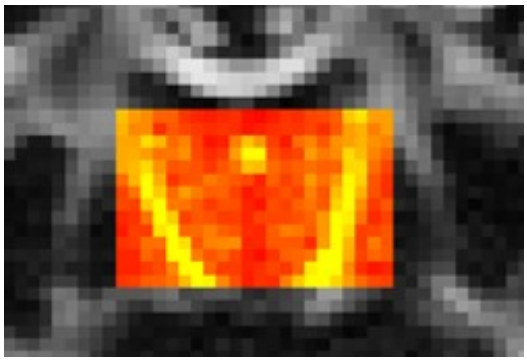
(a) Coronal slice showing ROI (in red)



(b) Feature enhancement after n iterations g_n

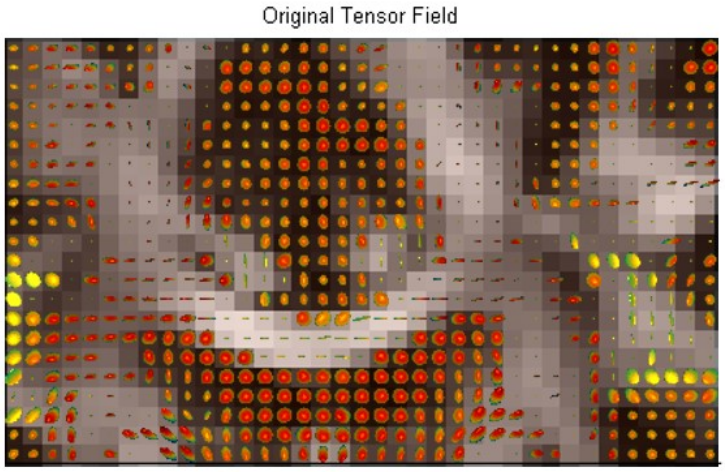


(c) $g(\frac{n}{2})$

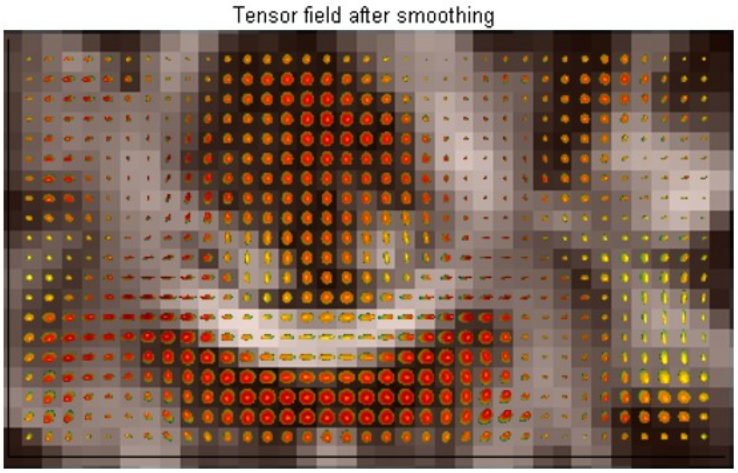


(d) $g(\frac{3n}{4})$

Figure 4.4: Feature enhancement experiment (with backward Ricci flow) using MRI data. Number of iterations $n = 40$.



(a) Original tensor field



(b) Tensor Field after smoothing

Figure 4.5: Tensor field smoothing experiment (with anisotropic diffusion) using MRI data, (a) Original tensor field (b) Smoothed tensor field.

Chapter 5

Ricci Curvature - Hallmark of Structural Brain Connectivity

Studies show that while brain networks are remarkably robust to a variety of adverse events, such as injuries and lesions due to accidents or disease, they may be fragile when the disturbance takes place in specific locations. This seems to be the case for diseases in which accumulated changes in network topology dramatically affect certain sensitive areas. To this end, previous attempts have been made to quantify robustness and fragility of brain functionality in two broadly defined ways: (i) utilizing model-based techniques to predict lesion effects, and (ii) studying empirical effects from brain lesions due to injury or disease. Both directions aim at assessing functional connectivity changes resulting from structural network variations. In the present work, we follow a more geometric viewpoint that is based on a notion of curvature of networks, the so-called Ollivier-Ricci curvature. A similar approach has been used in recent studies to quantify financial market robustness as well as to differentiate biological networks corresponding to cancer cells from normal cells. The same notion of curvature, defined at the node level for brain networks obtained from MRI data, may help identify and characterize the effects of diseases on specific brain regions. In the present work, we apply the Ollivier-Ricci curvature to brain structural networks to: i) Demonstrate its unique ability to identify robust (or fragile) brain regions in healthy subjects. We compare our results to previously published work which identified a unique set of regions (called *structural core*) of the human cerebral cortex. This novel characterization of brain networks, complementary to measures such as

degree, strength, clustering or efficiency, may be particularly useful to detect and monitor candidate areas for targeting by surgery (e.g. deep brain stimulation) or pharmaco-therapeutic agents; ii) Illustrate the power our curvature-derived measures to track changes in brain connectivity with healthy development/aging and; iii) Detect changes in brain structural connectivity in people with Autism Spectrum Disorders (ASD) which are in agreement with previous morphometric MRI studies.

5.1 Introduction

This chapter describes a geometric network-theoretic approach to study brain structural connectivity. As is well-known, comprehensive structural connectivity of the human brain, the so-called connectome [67, 68], is not available to date despite recent progress using neuroimaging such as the Human Connectome Project (HCP) [36,69–71]. Powerful imaging techniques such as diffusion MRI (dMRI) can be used to map the structural connectivity between different brain regions [37,72,73]. At the macroscopic level, brain regions are delineated and defined as nodes of a network, while edges describe connectivity (structural or functional) between them. Consequently, the overall structure may be mathematically represented as a graph [68,74]. Depending on the method used to construct the edges, brain networks can be divided into three types [68, 74, 75]: (i) structural networks that have edge weights based on strength of anatomical links between the nodes; (ii) functional networks in which the edges are given by statistical inter-dependence; (iii) functional networks whose edges are based on the causal influence of nodes. Clearly, the method employed to spatially parcellate the brain and consequently construct nodes will also affect the network parameters [76, 77]. The connectome has the potential to explain brain functions such as cognition, language, vision or audition, and their relation to specific brain regions since there exists a structure-to-function relationship between brain structural and functional networks [78].

In [79], robustness of the brain networks has been defined as the "degree to which the topological properties of a network are resilient to lesions such as the removal of nodes or edges." In particular, robustness measures to what extent the brain can withstand damage from, or be affected by, lesions arising e.g. from tumor, trauma or stroke. Reduced robustness not only suggests potential for dysfunction due to the lesion, but may also indicate candidate target locations for treatment.

Brain resilience has been studied previously by considering the effect of deleting network nodes or edges from brain structural and functional networks, both computationally and empirically; see [80] for a comprehensive review. Brain robustness studies can broadly be divided into two categories. In the first, one attempts to predict lesion effects by computational models, i.e., virtually lesioning nodes and edges of the structural connectivity matrix and applying computational models to predict functional connectivity changes [81–83]. Subsequently, the predicted functional connectivity matrix can be compared with the empirical one and lesioning effects can be quantified using various graph measures. In the second category, one employs the empirical effects from brain lesions due to injury or disease. Studies using this approach focus on examining brain networks of patients with, e.g. traumatic brain injury (TBI), stroke or tumors, and quantify the effect of lesion location on the brain [84, 85], by comparison with data from age- and gender-matched healthy control subjects. Regardless of the approach, the structure-to-function network relation is utilized to predict the amount of damage which the brain can withstand due to lesions in a given location.

We apply the geometric notion of graph curvature to brain structural networks, and leverage this novel concept to analyze brain robustness. Previous studies have shown that network curvature can be used to differentiate cancer from normal tissue using gene co-expression networks [86], and to indicate market fragility in economic/ financial networks [87]. It is important to note that, since network robustness can be viewed as the rate function at which a network returns to its original state after a perturbation, it has a positive correlation with entropy [88]. Consequently, network robustness and curvature are positively correlated through entropy [9].

In this chapter, we introduce the concept of graph curvature for studying brain structural connectivity networks. We use Ricci curvature and its contraction, the scalar curvature, on brain networks to introduce curvature as a nodal measure. In this study, using node curvature, we make two contributions to brain structural connectivity analysis: First, we identify areas of the brain that significantly contribute to the overall brain robustness, and hence we identify "important" nodes in brain networks. Previous studies have shown that hub nodes are critical for brain networks, but identifying such nodes is not straightforward. Node measures such as degree or strength do not identify all the hub nodes, and typically a combination of those measures, with centrality measures, is required [80, 89]. We show that node curvature not only corroborates findings based on strength and centrality measures, but

additionally finds other key areas (e.g., inferior-frontal gyrus, middle-frontal gyrus, and inferior-temporal gyrus), which are not identified by any other node measure, and are important parts of the brain network. Second, by looking at the difference in node curvature, one can identify areas of the brain that are related to healthy development/aging, or abnormal neurodevelopment disorders such as autism spectrum disorders (ASD). In particular, we show that node curvature uniquely enables the identification of certain brain areas, with significantly affected structural connectivity in people with ASD.

The remainder of the chapter is organized as follows. Section 5.2 provides some brief theoretical background on Ricci curvature, its relation to network robustness and curvature as a node measure for graphs. Details of the diffusion MRI data sets used in this study are also provided in this section. Section 5.3 presents results illustrating how brain curvature identifies: critical nodes (i.e. supporting robustness) in normal brain networks, nodes undergoing significant changes in healthy development and aging, and nodes involved in abnormal development in ASD. Specifically, Section 5.3.1 shows critical nodes found using two different datasets of Diffusion Spectrum Imaging (DSI), which is a variant of dMRI, used by Hagmann et al. [89] to identify nodes constituting the structural core of the human cerebral cortex and by the MGH-USC HCP Consortium [73]. In Sections 5.3.3 and 5.3.4, we respectively use the WU-Minn HCP Consortium [37] Lifespan Pilot Project high angular resolution diffusion imaging (HARDI) data, and diffusion tensor imaging (DTI) data from the ASD study by Rudie et al. [90]. We describe changes in structural connectivity due to healthy development/aging, as well as ASD. We conclude by summarizing the results and future research directions in Section 5.4.

5.2 Methods

Detailed description of OMT and Wasserstein distance is given in Sections 3.2.1 and 3.2.1.2.

5.2.1 Curvature

In this section, we introdcue the key notion of curvature from Riemannian geometry. Full details may be found in [55]. For X an n-dimensional Riemannian manifold, $x \in X$, let T_x denote the tangent space at x, and $v_1, v_2 \in T_x$ orthonormal

tangent vectors. Then, for geodesics (local curves of shortest length) $\alpha_i(t) := \exp(tv_i)$, $i = 1, 2$, the sectional curvature $K(v_1, v_2)$ measures the deviation of geodesics relative to Euclidean geometry, i.e.,

$$d(\alpha_1(t), \alpha_2(t)) = \sqrt{2}t(1 - \frac{K(v_1, v_2)}{12}t^2 + O(t^4)). \tag{5.1}$$

The Ricci curvature is the average sectional curvature. In other words, given a (unit) vector $u \in T_x$, we complete an orthonormal basis $v, v_2, \ldots, v_n$, and define $Ric(v) := \frac{1}{n-1}\sum_{i=2}^{n} K(v, v_i)$. The Ricci curvature may be extended to a quadratic form known as the Ricci curvature tensor [55]. The scalar curvature is then defined to be the trace of this quadratic form.

There are a number of alternative characterizations of Ricci curvature [55]. In this chapter, we employ the following definition: Referring to Figure 5.1, consider two very close points x and y in X and associated tangent vectors w and w', where w' is obtained by parallel transport of w along a geodesic (in the direction v) connecting the two points. Now, geodesics drawn from x, y along w, w' will get closer when the curvature is positive (positively curved space). This is also reflected in the fact that the distance between two small (geodesic balls) is less than the distance of their centers. The Ricci curvature $Ric(v)$ along the direction v connecting x, y quantifies this contraction. Similar considerations apply to negative and zero curvature.

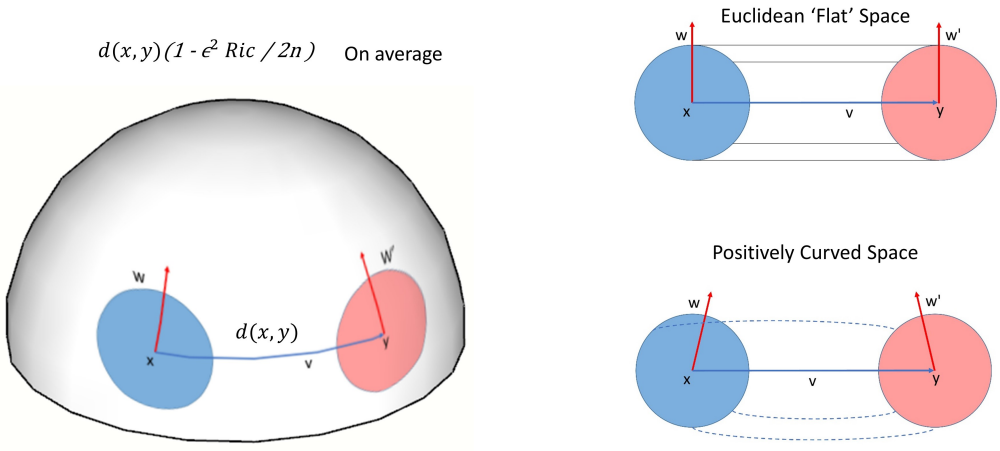


Figure 5.1: In a space with positive Ricci curvature, parallel geodesics emanating from points x and y, e.g., in directions along tangent vectors w (at x) and w' (at y), are drawn closer. In a Euclidean space, the distance between points moving along parallel geodesics at constant speed remains constant.

5.2.2 Ollivier-Ricci and Scalar Curvature

An analogue of the Ricci and of the scalar curvature exists for discrete spaces. Here, the place of the earlier geodesic (in a direction of a tangent vector v) connecting two points x, y, is taken simply by the pair (x, y). The so-called Ollivier-Ricci curvature is now denoted by $\kappa(x, y)$ and the scalar curvature at x is the average $\kappa(x, y)$ over neighboring nodes y. We now define these ideas in greater details.

5.2.2.1 Ollivier-Ricci Curvature

Ollivier-Ricci curvature or coarse Ricci curvature is the discrete generalization of the Ricci curvature [91–93]. Let (X, d) be a geodesic metric space equipped with a family of probability measures $\{p_x : x \in X\}$. We define the Ollivier-Ricci curvature $\kappa(x, y)$ along the geodesic connecting x and y as

$$W_1(p_x, p_y) = (1 - \kappa(x, y))d(x, y), \tag{5.2}$$

where W_1 denotes the Earth Mover's Distance (Wasserstein 1-metric), and d the geodesic distance on the space.

For the case of an undirected weighted graph (e.g. a brain structural connectivity network) $G = (V, E, W)$, where V is the set of vertices (nodes), E the set of edges, and W the set of edge weights, we let

$$\begin{aligned} d_x &= \sum_{y \in \mathcal{N}(x)} w_{xy} \\ p_x(y) &:= \frac{w_{xy}}{d_x}, \end{aligned}$$

where $\mathcal{N}(x)$ denotes the set of nodes that are adjacent to x; throughout we assume that all the edge weights $w_{xy} = w_{yx} \geq 0$ and that $w_{xy} = 0$ if $d(x, y) \geq 2$, or equivalently, if $y \notin \mathcal{N}(x)$. Note here that the geodesic distance $d(x, y)$ is taken to be the hop distance between node x and y, i.e., the minimum number of steps it takes to go from x to y.

5.2.2.2 Node Curvature

The (scalar) node curvature for node x (κ_x) in the graph is computed by summing the curvature between node x and all its neighboring nodes, i.e.,

$$\kappa_x := \sum_{y \in \mathcal{N}(x)} \kappa(x, y). \tag{5.3}$$

We also note that an alternative "weighted" version of the node curvature may be defined as

$$\tilde{\kappa}_x := \sum_{y} p_x(y)\kappa(x, y). \tag{5.4}$$

5.2.3 Robustness and the Fluctuation Theorem

We now turn to the notion of robustness which we will employ in this work, and subsequently make the link between robustness and curvature. It is based on ideas from statistical mechanics and, in particular, the Fluctuation Theorem formulated in [88]. The Fluctuation Theorem measures the ability of a network to maintain its functionality in the face of perturbations (internal or external).

Let $p_\delta(t)$ be the probability that the mean value of an observation (for a given network) deviates from its original value, by more that δ at time t, due to some perturbation. The rate R at which the system returns back to its original state is defined as

$$R := \lim_{t \to \infty} \ (-\frac{1}{t} \log p_\delta(t)). \tag{5.5}$$

Note that large R means not much deviation and small R implies a large deviation. In statistical mechanics, it is well-known that entropy and rate functions from large deviations are very closely related [88,94]. The Fluctuation Theorem is a mathematical statement relating the positive correlation of changes in system entropy ΔS to changes in robustness ΔR:

$$\Delta S \times \Delta R \geq 0. \tag{5.6}$$

5.2.4 Ollivier-Ricci Curvature and Graph Robustness

In this section, following [86,87], we establish the relationship between curvature and robustness. For ease of explanation, we consider X to be a Riemannian manifold;

the relations drawn can be extended to discrete spaces.

It turns out that one can endow the space of probability densities on X (taken with respect to the volume measure) with a natural Riemannian structure as well; see [95] for details. We denote the space of probability densities by $\mathscr{P}(X)$. Then, one defines the Boltzmann entropy of $\rho \in \mathscr{P}(X)$ as

$$S(\rho) := -\int_X \rho \log \rho \, dm \tag{5.7}$$

where dm denotes the volume measure on X. (There are several technical assumptions that should be made to ensure the existence of S, see [9, 96].)

We can then express the following remarkable result from [9, 96]: Let Ric denote the Ricci curvature defined on X, and suppose the $Ric \geq k$ for any tangent vector on X. Then, for every $\rho_0, \rho_1 \in \mathscr{P}(X)$, the constant speed geodesic ρ_t with respect to the Wasserstein 2-metric connecting ρ_0 and ρ_1 satisfies

$$S(\rho_t) \geq (1-t)S(\rho_0) + tS(\rho_1) + \frac{kt(1-t)}{2} W(\rho_0, \rho_1)^2, \quad 0 \leq t \leq 1. \tag{5.8}$$

This relationship indicates the positive correlation of changes in entropy and curvature [86], namely

$$\Delta S \times \Delta Ric \geq 0. \tag{5.9}$$

Similar arguments can be made on discrete spaces. Following the arguments of [86, 87], and the Fluctuation Theorem, one can express the positive correlation of graph Ricci curvature (5.2) and robustness, expressed as:

$$\Delta R \times \Delta Ric \geq 0. \tag{5.10}$$

The main conclusion is that one can use network curvature as a measure of functional robustness. Since curvature can easily be computed via linear programming [10, 97], it provides an attractive and novel tool to characterize the robustness of networks represented as weighted graphs, such as brain connectivity networks. In the next section, we briefly summarize existing measures to characterize complex brain networks and provide information about the datasets which will be used next to demonstrate the benefits of curvature.

5.2.5 Measures of Brain Networks Characteristics

We hereafter briefly summarize important graph-theoretical measures, which have been introduced to characterize brain networks [98], and will be used in our experiments.

5.2.5.1 Node Strength s_i

The strength of a node i is the sum of the weights w_{ij} of the node's adjacent edges [99], i.e.,

$$s_i = \sum_{j \in \mathcal{N}(i)} w_{i,j}. \tag{5.11}$$

Then dMRI data may be employed, in combination with deterministic or probabilistic integration/propagation methods of vector fields (called *tractography*), to assess the likelihood of connections between cortical and sub-cortical areas [100]. Such likelihood can be obtained by the number of three-dimensional curves generated by these integration or propagation methods and used, in the context of brain structural networks, to define the weight $w_{i,j}$ of an edge between two nodes i and j. This summarizes how strongly connected those nodes are to each others, and to the rest of the brain.

5.2.5.2 Betweenness Centrality g_i

The betweenness centrality of a given node i is defined as the number of shortest paths between pairs of nodes that pass through the node i [101], i.e.,

$$g_i = \sum_{i \neq j \neq k} \frac{\sigma_{j,k}(i)}{\sigma_{j,k}}. \tag{5.12}$$

Where $\sigma_{j,k}$ is the total number of shortest paths from node j to node k and $\sigma_{j,k}(i)$ is the number of those paths that pass through i.

5.2.5.3 Clustering Coefficient C_i

The Clustering Coefficient of node i is a measure of the density of connections among the node's topological neighbors [102, 103]. This is defined as follows: Take

$i \in V$, the vertex set of a graph $G = (V, E, W)$, and assume unit weights $e_{ij} \in W$ for all existing edges. Suppose that the node i has k_i neighbors. For an un-directed graph (which is usually the case for brain structural connectivity networks), there can be at the most $k_i(k_i - 1)/2$ edges in the local neighborhood. Then, C_i is defined as the fraction of the edges that actually exist between the immediate neighbors of i over the maximal number of edges, i.e.,

$$C_i = \frac{2|\mathscr{N}(i)|}{k_i(k_i - 1)}. \tag{5.13}$$

As before, $\mathscr{N}(i) = \{j : e_{ij} \in E\}$ is the immediate neighborhood of nodes if i, and $|\mathscr{N}|$ denotes the cardinality of this set.

5.2.6 Diffusion MRI Datasets

As briefly described in the introduction, we used four different public/open access datasets in our experiments, and we now provide more details about this data. First, we assess the ability of node curvature to capture novel information, which is complementary to existing measures of brain networks characteristics (e.g. strength). We analyze the high-resolution connectivity matrices created and analyzed by Hagmann et al [89], using DSI data from 5 healthy subjects. These matrices are available from the USC Multimodal Connectivity Database [104], which enables the reproduction of the original results [89], and evaluation of our method with the exact same datasets. We also analyze 33 new DSI datasets obtained from the MGH-USC HCP Consortium [73], to demonstrate the consistency of our findings on critical brain areas. Next, we illustrate the ability of node curvature to capture changes in certain brain areas, which are related to healthy development and aging. These experiments use HARDI datasets obtained from the WU-Minn HCP Consortium Lifespan Pilot Project [37]. Finally, we show that node curvature is uniquely capable of detecting changes in brain structural connectivity in ASD, which are in agreement with previous morphometric MRI studies. These last experiments are performed with DTI datasets from 51 subjects with ASD and 43 typically developing subjects, as previously published by Rudie et al. [90]. The connectivity matrices from this study are also available from the USC Multimodal Connectivity Database [104], which again enabled straightforward reproduction of those results and comparison with our method.

5.2.6.1 DSI Datasets from Hagmann et al. [89]

Data were acquired in 5 healthy right-handed male subjects, on a Philips Achieva 3T scanner with voxel size $2 \times 2 \times 3$ mm^3, TR/TE=4200/89 ms and 129 diffusion gradients with a maximum b-value of 9000 s/mm^2, for a total acquisition time of 18 min. After segmentation of the white and gray matter, 998 cortical regions-of-interest were created, with an average size of 1.5 cm^2. Tractography was then performed, and structural connectivity matrices created by defining the weight of each edge as the number of streamlines per unit surface (i.e. density). Additional details can be found in the original paper [89].

5.2.6.2 DSI Datasets from the MGH-USC HCP Consortium

Data were acquired in 35 healthy adults (age range 20 to 59) scanned on the customized Siemens 3T Connectom scanner and are available at https://db.humanconnectome.org. Two of the datasets were not included in our experiments because of pre-processing errors in our analysis pipeline. Acquisition parameters included voxel size of $1.5 \times 1.5 \times 1.5$ mm^3, TR/TE=8800/57 ms and four b-values (with corresponding number of diffusion gradients in parenthesis): 1000 s/mm^2 (64), 3000 s/mm^2 (64), 5000 s/mm^2 (128), 10000 s/mm^2 (256), for a total acquisition time of about 89 min. Connectivity matrices were generated using DSI Studio (http://dsi-studio.labsolver.org) as described in Section 5.2.6.5.

5.2.6.3 HARDI Datasets from the WU-Minn HCP Consortium Lifespan Pilot Project

Data were acquired from healthy subjects across the lifespan in 6 age groups: 4-6, 8-9, 14-15, 25-35, 45-55 and 65-75 years and are available at https://db.humanconnectome.org. We analyzed the data acquired on the UMinn Siemens 3T Prisma scanner (Phase 1a), which include 5 participants per age group (ages 25-35, 45-55 and 65-75) or 6 participants per age group (ages 8-9 and 14-15). Acquisition parameters included voxel size of $1.5 \times 1.5 \times 1.5$ mm^3, TR/TE=3222/89 ms and two b-values, 1000 s/mm^2 and 2500 s/mm^2, each with 75 diffusion gradients acquired twice with opposite phase-encoding polarity, for a total acquisition time of about 22 min. Connectivity matrices were also generated using DSI Studio (http://dsi-studio.labsolver.org) as described in Section 5.2.6.5.

5.2.6.4 DTI Datasets from Rudie et al. [90]

Data were acquired in 51 subjects with ASD and 43 typically developing subjects, on a Siemens 3T Trio scanner with voxel size $2 \times 2 \times 2\ mm^3$, TR/TE=9500/87 ms and one b-value of 1000 s/mm^2 with 32 diffusion gradients, for a total acquisition time of about 8 min. Deterministic tractography was used to create structural connectivity matrices, with weights defined as the number of streamlines connecting each pair of 264 cortical areas (nodes). Additional details can be found in the original paper [90].

5.2.6.5 Generation of Connectivity Matrices for the HCP Datasets

We used DSI Studio (`http://dsi-studio.labsolver.org`) [105] to process the HCP DSI and HARDI datasets.

To run tractography and generate connectivity matrices for the DSI data, seeds were placed randomly in the whole brain with the following settings: normalized quantitative anisotropy (NQA) threshold: 0.08, angular threshold: 60°, tractography method: Runge-Kutta [106], total number of streamlines: 5 million (Although similar results were obtained with $500,000$ streamlines, we used 5 millions to ensure consistency with [89]). 116 cortical areas (nodes) were automatically segmented via non-linear registration of the Automated Anatomical Labeling (AAL) template available in DSI Studio.

For the HARDI data, diffusion tensors were estimated to perform tractography. Seeds were also placed randomly in the whole brain with the following settings: fractional anisotropy (FA) threshold: 0.1, angular threshold: 60°, tractography method: Runge-Kutta [106], total number of streamlines: $500,000$. 116 cortical areas (nodes) were automatically segmented via non-linear registration of the Automated Anatomical Labeling (AAL) template available in DSI Studio.

5.3 Results

5.3.1 Curvature as a Hallmark of Brain Areas Robustness

Individual node curvature (as defined in Section 5.2.2.2) of brain regions contributes to the overall (average) curvature of the brain network. This measure not only helps identify changes in the network, as discussed in the previous sections, but also can help identify relatively more important parts of the brain structural net-

work. As explained in Section 5.2.4, equation (5.5), curvature is directly correlated to network robustness. Therefore, nodes with higher curvature contribute more to the overall structural robustness of the network.

To demonstrate this, we performed experiments using two different Diffusion Spectrum Imaging datasets: First, the DSI data for 5 participants, as presented in [81, 89], was considered, to enable comparison of our results to previous studies. High-resolution connectivity matrices (998×998) were obtained from the USC Multimodal Connectivity Database [104] . Second, the DSI data for 33 participants from the MGH-USC HCP Consortium was also employed, and lower resolution connectivity matrices (116×116) were generated as described in Section 5.2.6.2.

As previously described, we emphasize here that comparing properties across brain networks with different resolutions (i.e. number of nodes) should be done only with great care [76], as brain network properties can differ significantly with nodal parcellations [76, 77]. Nonetheless, it is worth studying curvature as a measure which may provide information across different network resolutions: High resolution parcellations, also known as *dense connectomes* [107] will ultimately provide greater insights into the structure of brain networks, while lower resolution parcellations are more easily manageable, since they requires less computational resources.

First, we present results based on the high resolution connectivity matrices [104] in Figure 5.2. Here, we show the top 25% of nodes with the highest node curvature, strength and betweenness centrality, appearing consistently across 5 participants. Figures in panels B and C follow the same convention as Figure 2 (C) and Figure 7 (A) of [89] respectively, and are presented here for comparison purposes. In previous study [89], using several network analysis methods, eight anatomical sub-regions were identified as belonging to the so-called *structural core network* of the human brain. These include the posterior cingulate cortex, precuneus, cuneus, paracentral lobule, isthmus of the cingulate, banks of the superior temporal sulcus, and inferior/superior parietal cortex, all of them in both hemispheres. Additionally, [81] showed that lesions in the temporo-parietal junction, cortical midline and frontal cortex have the most extensive effects on brain functionality. Also, we note that the medial prefrontal cortex forms part of the default mode network of the human brain [108]. Panel A of Figure 5.2 shows that curvature identifies areas in the inferior-frontal gyrus, middle-frontal gyrus and inferior-temporal gyrus, consistent with [81, 89], and thus providing

very interesting information based on network structure, which is not captured by strength or betweenness centrality.

Second, following the same organization as Figure 5.2, Figure 5.3 shows results for the low resolution matrices generated from the MGH-USC HCP Consortium datasets. As expected, distinct areas are identified with all three measures (since cortical parcellation is different from the one used in Figure 5.2) [76]. We should also note that the high resolution data did not include the cerebellum. Nodes with high strength and betweenness centrality are found more towards the frontal, precentral, superior parietal areas and in the cerebellum. Once again, curvature supplements the information provided by other measures and identify areas in the inferior-frontal gyrus and transverse temporal gyrus (Heschl's gyrus) in both hemispheres, where lesions are known to induce pronounced effects in loss of brain functionality [81].

5.3.2 Remarks on Graph Measures and Assessment of Brain Network Robustness

In order to quantify the importance of a given node in a brain network, the effect of node(s) deletion on graph measures can be considered [81]. Based on node measures such as strength, betweenness centrality, and curvature, given node(s) and all related edges are removed (i.e., by deleting the row and the column) from the connectivity matrix. Independent graph measures such as connectedness or global efficiency are then computed on the remaining connectivity matrix, and the effect of node(s) and edges removal is quantified by the faster or slower decay of those measures. For instance, if a network becomes disconnected (i.e. size of largest connected component becomes small) quickly, when removing nodes based on their importance, as quantified by a particular graph measure, we can conclude that this graph measure provides an effective way to identify important nodes (because the graph breaks down quickly after removing only a few nodes). If, however, the graph becomes disconnected slowly when removing nodes based on a different graph measures, we can conclude that this particular is not as good as identifying important nodes.

5.3.2.1 Gaussian Transformation of the Connectivity Matrix Weights

Streamlines numbers (i.e., the weights of structural connectivity graphs) produced by tractography algorithms are exponentially distributed [76, 89]. Without altering

the rank-ordering of strong to weak pathways, those distributions can be transformed to Gaussian distributions [81] for ease of analysis. However, we would like to emphasize that such transformation may lead to changes in edge weights (and consequently node strengths) that may affect graph measures differently, as illustrated next and thereby possibly biasing the definition of important nodes, as well as the identification of graph measures particularly adequate to quantify this importance. For instance, the weights, while preserving order, can be mapped to a completely different range of values, thereby increasing or diminishing the relative importance of nodes. (e.g. mapping weights 1 and 1000 to 10 and 11). We therefore recommend to apply such transformation(s) with care.

Using the size of largest graph component and global efficiency as metrics (i.e., the independent graph measures), as in [81] (see e.g., Figure 3 in the paper), Figures 5.4 and 5.5 illustrate the fact that if such a Gaussian transformation is performed [81], betweenness centrality can be seen as a better measure of node importance compared to strength (see Figures 5.4 and 5.5, panels C-D). Both the largest graph component and betweenness centrality decrease rapidly. However, without this transformation (i.e, with original connectivity matrices), although the largest graph component preserves its behavior with deletion of nodes based on either strength or betweenness centrality, global efficiency does not: strength appears to be a better measure of node importance (see Figures 5.4 and 5.5, panels A-B). We note that, despite having very different numbers of nodes, both DSI Datasets from Hagmann et al. [89] and the MGH-USC HCP Consortium exhibit the same results. This illustrates the potential effect of transforming connectivity matrix weights, and provides justification for our choice to generally not transform those weights.

5.3.2.2 Comparison Across Nodal Scales

For each step of a robustness analysis as presented in Figures 5.4 and 5.5, a node is deleted and the size of the network is reduced. Thus, we end up comparing parameters across brain networks with different parcellations (i.e. different number of nodes). It is important to note that network topological measures show strong dependence on the nodal scale [76], therefore, discrepancy in the measures may be attributed, to some extent, to disparate nodal scales.

5.3.2.3 Graph Measures to Assess Node Robustness

Traditionally, integration and centrality graph metrics like global efficiency, betweenness centrality, degree centrality, characteristic path length, and clustering coefficient have been used to study the robustness of brain networks [80]. While these measures certainly provide very useful information about the networks, we argue, based on the discussion provided in Section 5.2.4, that graph curvature and entropy may be more relevant metrics for robustness. To explore this, Figure 5.6 displays the changes in topological entropy [1] when nodes are deleted based on decreasing strength, curvature and betweenness centrality. For this experiment, the average of the 5 MGH-USC HCP Consortium DSI connectivity matrices (as described in Section 5.2.6.2) was used. Nodes with least importance, i.e., lower curvature, strength and betweenness centrality were deleted first, therefore node measures maintaining the entropy after the most number of nodes removal should give a better indication of robustness. We observe that node curvature and strength show similar results, as entropy remains high as more nodes (with low strength or curvature) are deleted, as compared to betweenness centrality. Thus in both cases, i.e., the original connectivity matrix (Figure 5.6, panel A) and the matrix after the Gaussian transformation (Figure 5.6, panel B), curvature and strength identify nodes contributing more towards the overall graph robustness.

5.3.3 Change in Curvature and Healthy Development/Aging of the Brain

We used datasets from the WU-Minn HCP Consortium Lifespan Pilot Project to study structural changes in brain networks related to healthy aging. Details about the data and construction of connectivity matrices are given in Section 5.2.6.3. Although node measures such as strength, betweenness centrality, and clustering coefficient can provide useful information about areas involved in healthy aging, we focus mainly on information provided by curvature in the present discussion. In Figure 5.7, we summarize the most pronounced changes related to healthy aging, and show results for the differences in node measures between the Lifespan group 1 (age 8-9) and group 5 (age 65-75). Figure 5.7 (panel A) illustrates that curvature finds significant bilateral differences in the parahippocampal gyrus, precuneus, paracentral lobule, thalamus and cerebellum. Additionally, curvature can detect differences in the hippocampus,

[1] The ***topological entropy*** of a graph G is the logarithm of the spectral radius of the adjacency matrix A, i.e., the logarithm of the maximum of the absolute values of the eigenvalues of A [109].

Heschl's gyrus, and temporal areas in the left hemisphere, as well as the caudate nucleus, gyrus rectus, frontal inferior gyrus, pars triangularis in the right hemisphere. Those findings are in agreement with previous studies demonstrating age related structural changes in the thalamus and hippocampus [110, 111], cerebellum [111], precuneus [112], para-hippocampus and precentral gyrus [113].

Exploratory statistical analysis: We note that in order to evaluate the statistical significance of differences for node measures (e.g. curvature) between two groups (e.g., different age groups, or controls and patients below), which can be of different sizes, unpaired two-sample t-tests were conducted. The null hypothesis of equal means between groups, and for each node, was tested and rejected for p-values (uncorrected for our exploratory results) less than 0.05, without assuming equal variances.

5.3.4 Change in Curvature and Autism Spectrum Disorders (ASD)

The aim of this experiment is to test whether various measures of node importance/robustness (curvature, strength, centrality and clustering) can detect differences in structural connectivity between individuals with Autism Spectrum Disorders (ASD) and typically developing subjects (TD). We utilized DTI data from 51 individuals with ASD and 43 TD participants, from a previously published study [90], so as to compare ASD-related differences in the brain network organization. Details about the data are given in Section 5.2.6.4. The DTI connectivity matrices capturing the brain structural connectivity of each participant were generated using a set of 264 coordinates [114] in MNI space. Additional details about the scanning protocol, diffusion data pre-processing and generation of connectivity matrices may be found in [90].

The original study [90] identified both functional and structural networks differences related to autism. For structural networks, changes in the white matter integrity and streamlines count were noted. Further, age-related atypical changes in the balance of local and global efficiency in structural and functional networks were observed. However, six global graph-theoretic metrics, namely the clustering coefficient, normalized clustering coefficient (λ), characteristic path length (CPL), normalized CPL (γ), small-worldness and modularity did not show statistically significant differences between subject the ASD and TD groups [90]. Interestingly, we found that strength, betweenness centrality, clustering coefficient and node curvature

are able to identify areas with significant differences between groups. Therefore, we focus in these last four measures and summarize our findings next.

Figure 5.8 shows nodes with statistically significant (see remark above) differences between the ASD and TD groups, using node measures: curvature, strength, betweenness centrality and clustering coefficient. While node strength, betweenness centrality and clustering coefficient provide important information about affected areas, we should note that curvature identify differences in areas that are not detected by the other measures. In particular, curvature discovered differences in the following areas of the right hemisphere: inferior temporal gyrus (temporo-occipital), temporal occipital fusiform cortex, para-hippocampal gyrus (posterior), cuneus cortex, supramarginal gyrus (posterior), insular cortex, frontal orbital cortex, and frontal pole (Figure 5.8, Panel A), which is in line with previous studies [115–119]. Interestingly, a morphometric analysis of asymmetry in volume/structure of cortical areas in individuals with ASD found a pronounced rightward bias [115], which appears to agree with our findings based on curvature. Indeed, node curvature detected all the affected areas (in case of both ASD and developmental language disorder (DLD)) in the right hemisphere, except for the lingual gyrus, as also reported in Table 4 of [115].

5.4 Conclusion

In this work, we have introduced the concept of graph curvature to quantify the importance of nodes (meaning that their disruption leads to large changes in the overall graph) in brain networks. We have shown that curvature can indeed help identify important areas/nodes and points to changes in the brain network structure that may be attributed to development/aging or diseases. The close relation between curvature and network robustness points to the significance of the detected nodes/areas in ensuring robust operation. Thus, this study lays the foundation for a new approach to assess brain network robustness at the nodal level. We argue that the information provided by curvature may be used in combination with other nodal measures for identifying/studying global changes in the brain.

Curvature (averaged across the network) can also provide a global graph measure for brain network robustness quantification. A similar viewpoint has recently been proposed in the context of financial networks and of gene regulatory networks [86,87,120]. It is indeed quite interesting that the connection between curvature

and the ability of a dynamical process on a network to return to equilibrium after a perturbation (robustness) is observed in such disparate problems (economy, thermodynamics gene regulation and cancer, brain networks). Several other directions may be worthy of investigating in a similar spirit. In particular, studying curvature changes between nodes at the edge level may prove profitable as, in that case, critical changes in interactions between areas in the brain may be easier to detect. We offer these pointers with the caveat that curvature is sensitive to the way connectivity matrices are generated, i.e., curvature is affected by the choice of parcellation scale, tractography algorithms as well as the particular type of diffusion data, e.g., DTI, HARDI, DSI, etc. Therefore, care must be exercised to minimize such possible effects. The present work focused mainly on exploring the concept of node curvature as a measure of robustness of brain structural networks and comparing with alternative measures.

In future studies, it will be worthwhile to explore the relevance of curvature to identify areas belonging to the so-called (functional) default mode network, which can be derived from resting-state functional MRI data, as well as to improve the detection of changes related to development/aging or diseases.

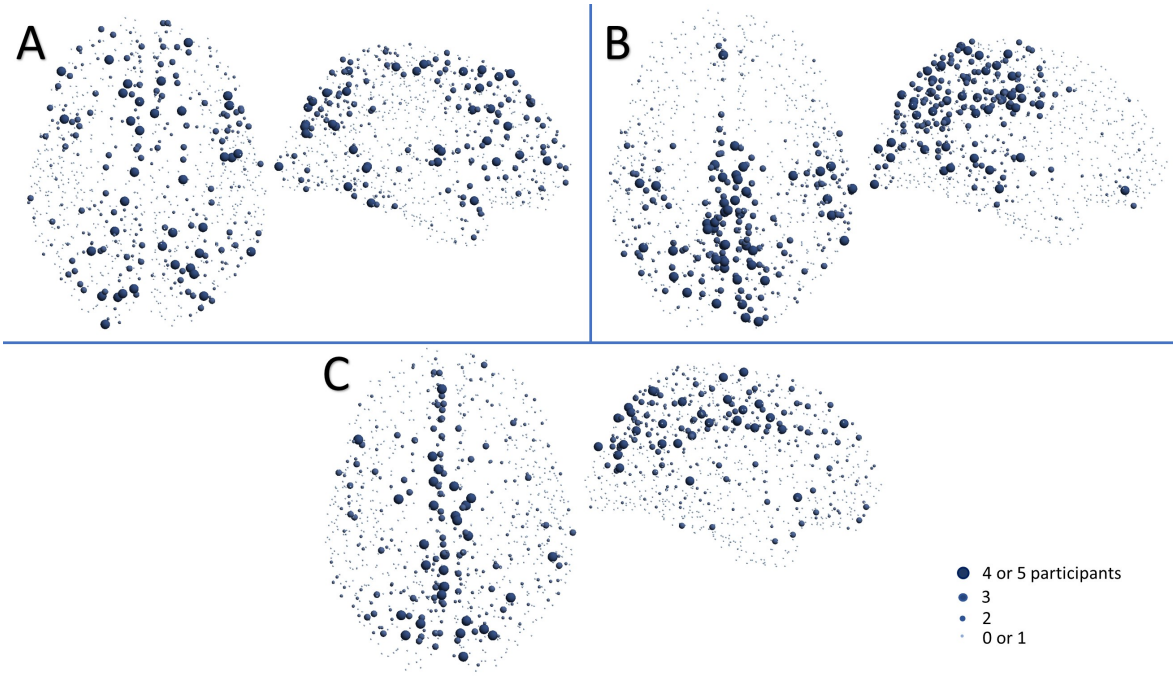


Figure 5.2: Nodes (top 25%) with highest curvature (A), strength (B) or betweenness centrality (C) consistently present across subjects. For instance, the largest spheres indicate nodes with high values in 4 or 5 out of the 5 subjects. Results from high resolution connectivity matrices (998×998) from Hagmann et al [89].

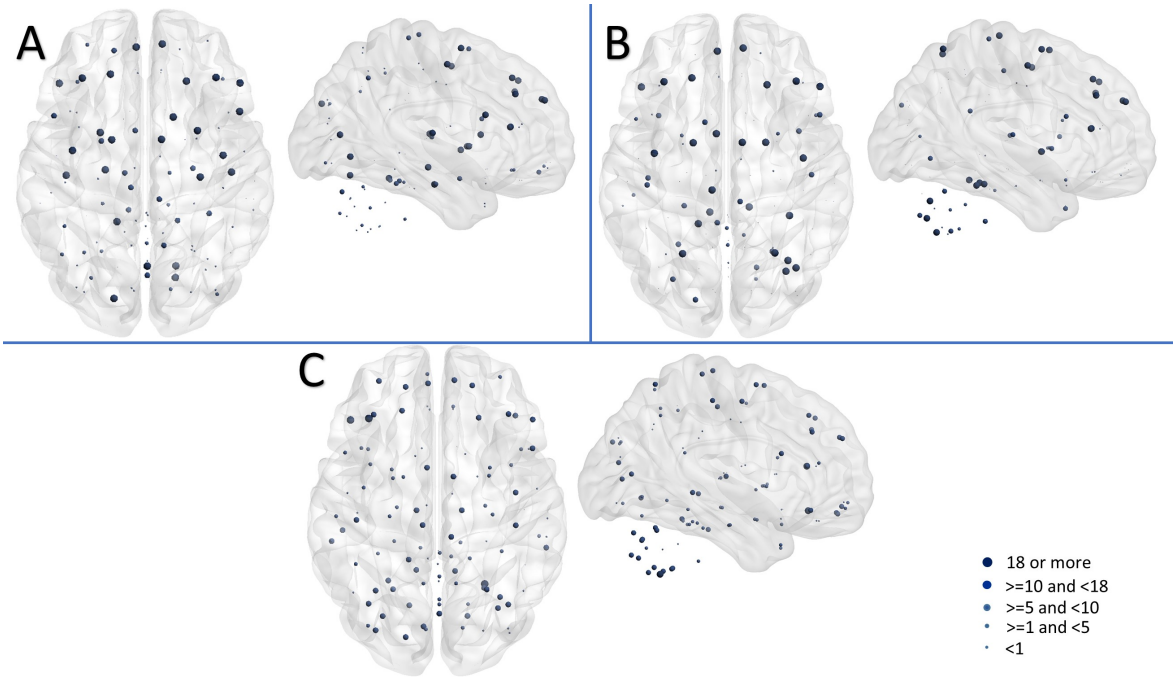


Figure 5.3: Nodes (top 25%) with highest curvature (A), strength (B) or betweenness centrality (C) consistently present across subjects. For instance, the largest spheres indicate nodes with high values in 18 out of the 33 subjects. Results from lower resolution connectivity matrices (116×116) generated using the AAL atlas and MGH-USC DSI datasets.

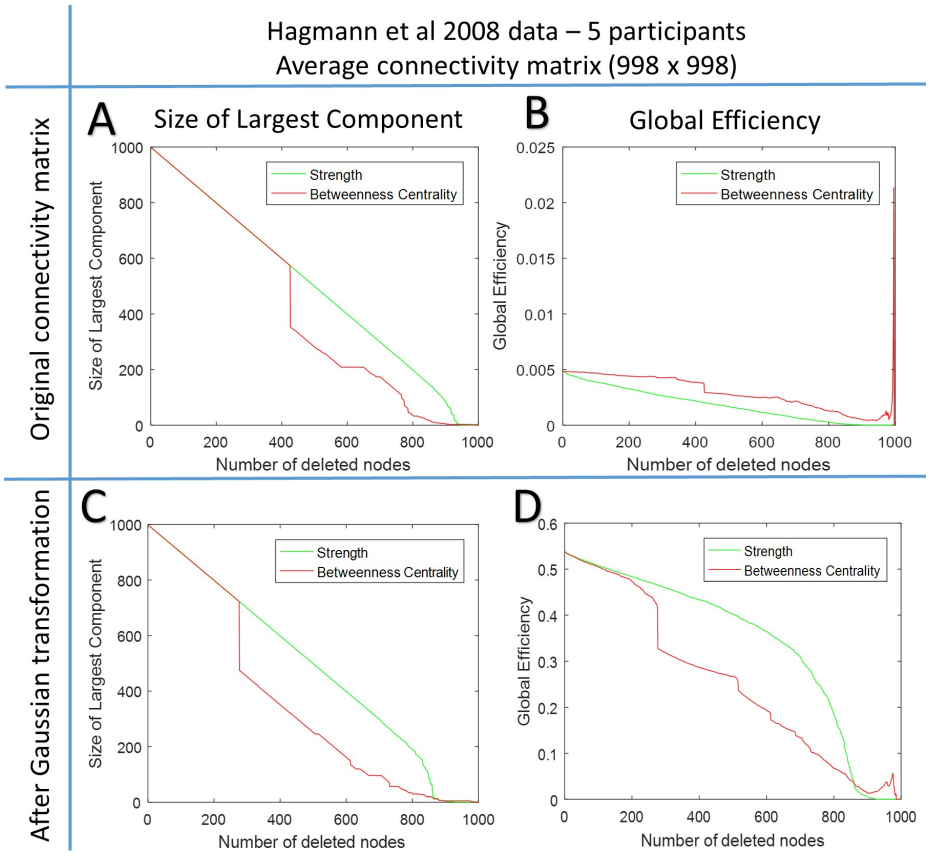


Figure 5.4: Robustness analysis using node deletion based on their strength or betweenness centrality, for the high resolution connectivity matrices (998×998) from Hagmann et al [89]. The size of the largest component and global efficiency are computed (with or without transformation of the connectivity matrix weights) after targeted removal of nodes with high strength or betweenness centrality. The top row shows results for the original connectivity matrix while the bottom row shows results after Gaussian transformation of its weights. Note that results shown in panel C and D are similar to Fig. 3 in [81].

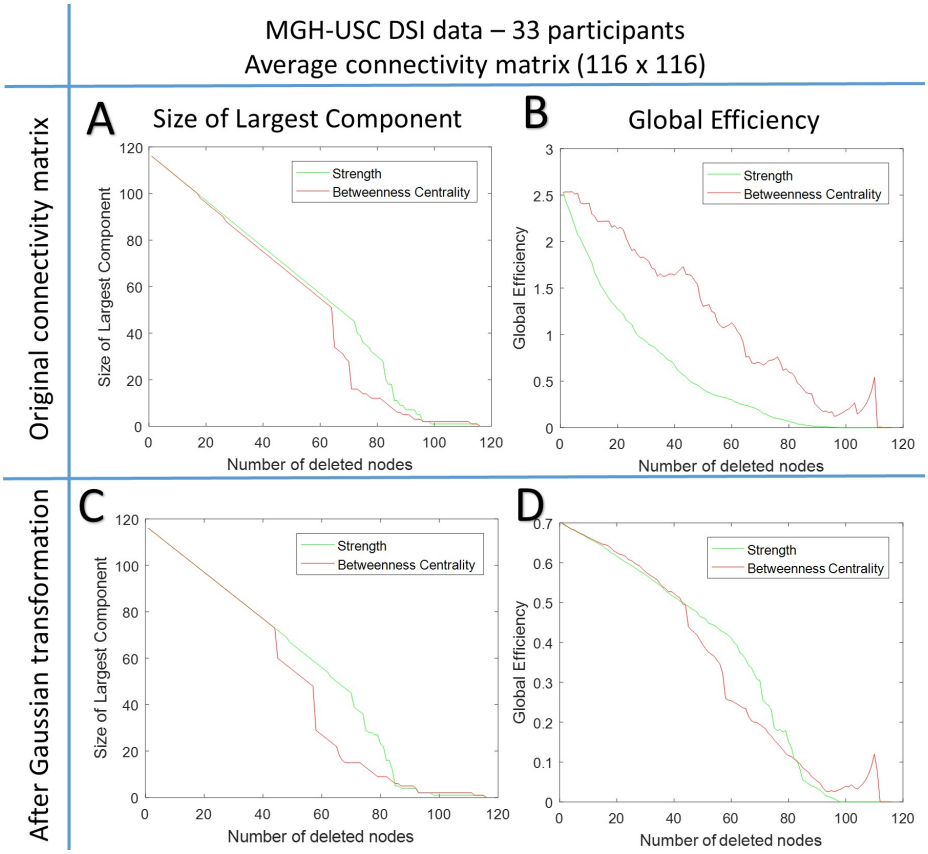


Figure 5.5: Robustness analysis using node deletion based on their strength or betweenness centrality, for the lower resolution connectivity matrices (116×116) generated using the AAL atlas and MGH-USC DSI datasets. The size of the largest component and global efficiency are computed (with or without transformation of the connectivity matrix weights) after targeted removal of nodes with high strength or betweenness centrality. The top row shows results for the original connectivity matrix while the bottom row shows results after Gaussian transformation of its weights. Note that results are consistent with those presented in Figure 5.4.

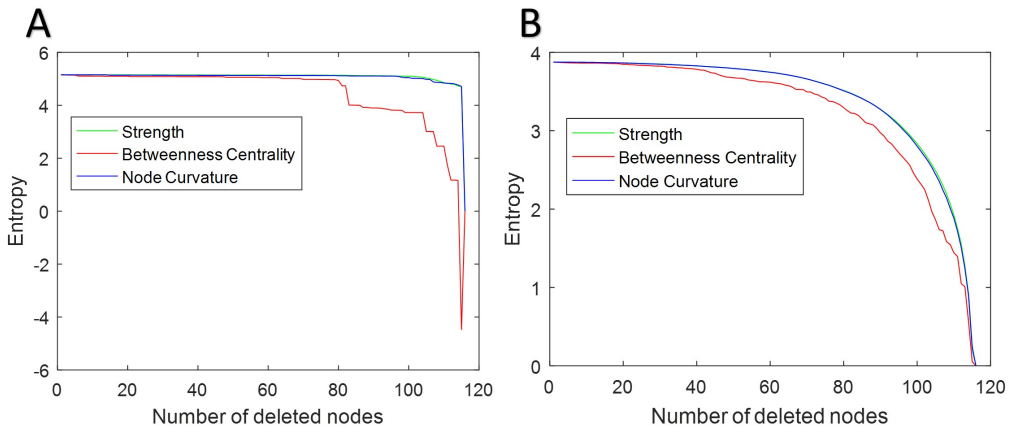


Figure 5.6: Robustness analysis on the basis of node deletion with topological entropy as a measure, using low resolution average connectivity matrix from MGH-USC DSI data. (A) Original average connectivity matrix (B) After Gaussian transformation of the connectivity matrix. Note that Node Curvature and Strength show similar trend and are better measures of robustness as compared to Betweenness Centrality.

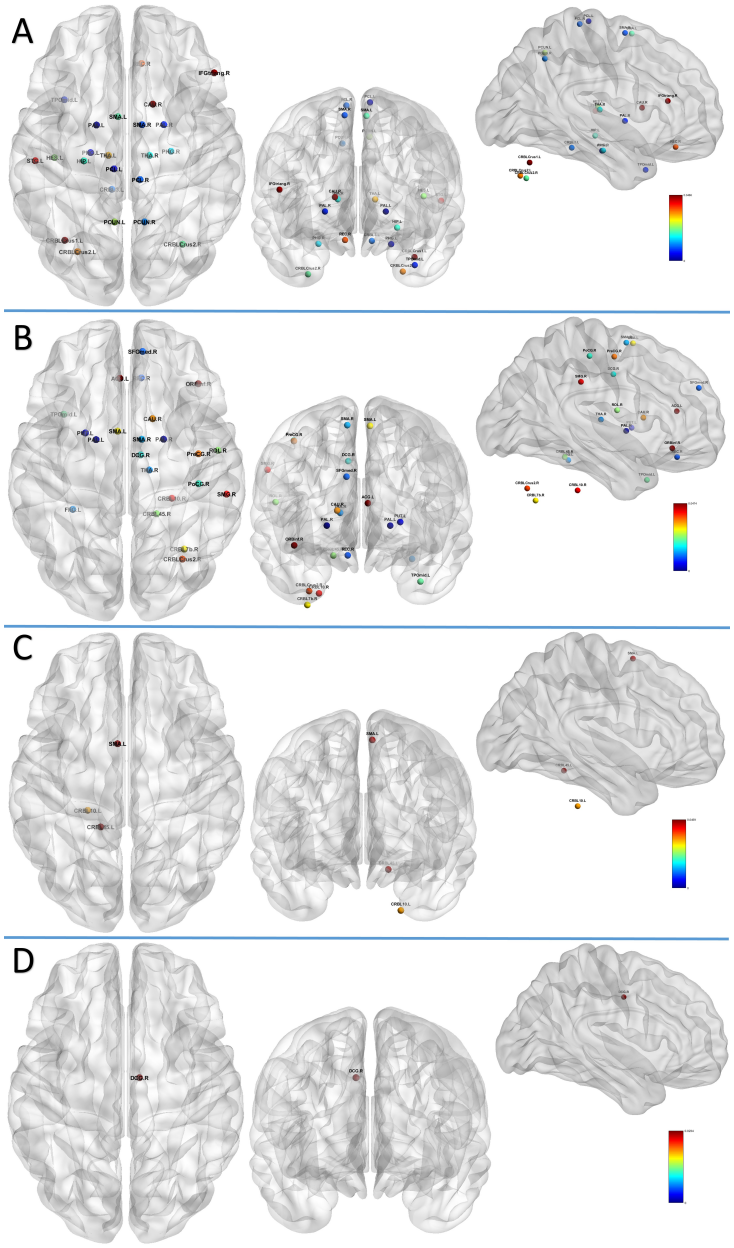


Figure 5.7: Nodes with significant difference in structural connectivity related to healthy aging. Node color correspond to the p-value. (A) Node curvature. (B) Node strength. (C) Betweenness centrality. (D) Clustering coefficient. **Node Name Abbreviations**: Suffix R = Right hemisphere, L = Left hemisphere; ACG = Anterior cingulate gyrus; CAU = Caudate nucleus; CRBL = Cerebellum; CRBLCrus = Crus of cerebellar hemisphere; DCG = Mid-cingulate area; FFG = Fusiform gyrus; HES = Heschl's gyrus; HIP = Hippocampus; IFGtriang = Pars triangularis; ORBinf = Pars orbitalis; PAL = Pallidum; PCL = Paracentral lobule; PCUN = Precuneus; PHG = Parahippocampus; PoCG = Postcentral gyrus; PreCG = Precentral gyrus; PUT = Putamen; REC = Rectus gyrus; ROL = Rolandic operculum; SFGmed = Medial frontal gyrus; SMA = Supplementary motor area; SMG = Supramarginal gyrus; STG = Superior temporal gyrus; THA = Thalamus; TPOmid = Middle temporal pole.

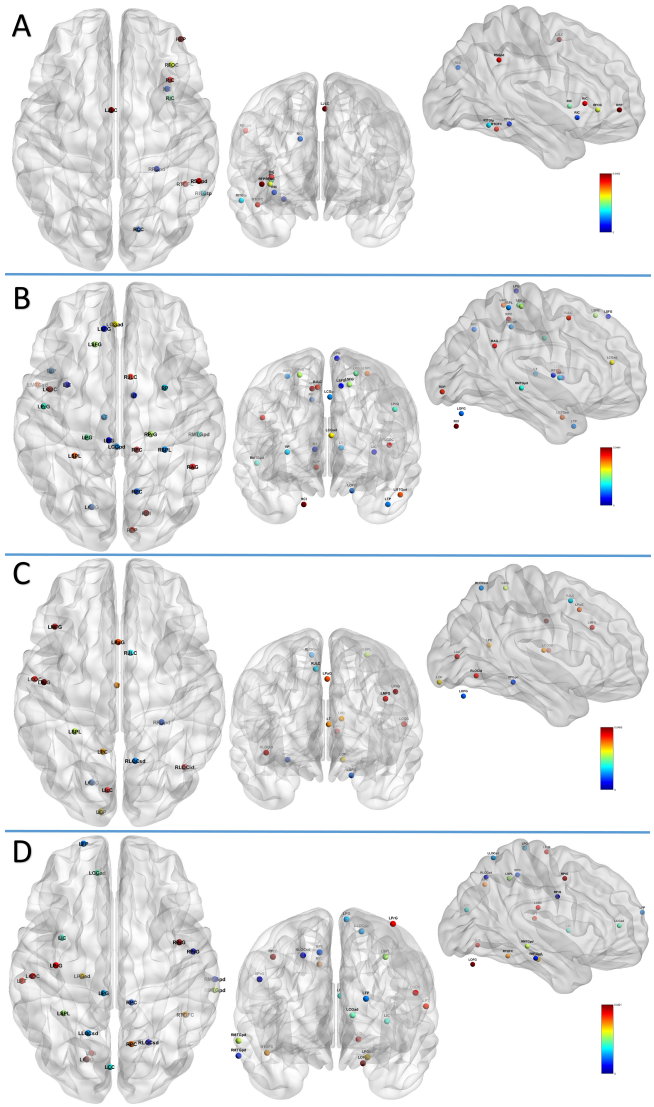


Figure 5.8: Nodes with significant difference in structural connectivity between individuals with Autism Spectrum Disorders and typically developing subjects. Node color correspond to the p-value. (A) Node curvature. (B) Node strength. (C) Betweenness centrality. (D) Clustering coefficient. **Node Name Abbreviations**: Suffix R = Right hemisphere, L = Left hemisphere; AG = Angular gyrus; CC = Cuneus cortex; CGad = Cingulate gyrus, anterior division; CGpd = Left cingulate gyrus, posterior division; COC = Central opercular cortex; FOC = Frontal orbital cortex; FP = Frontal pole; IC = Insular cortex; IcC = Intracalcarine cortex; ITGtp = Inferior temporal gyrus, temporo-occipital part; JLC = Juxtapositional lobule cortex; LG = Lingual gyrus; LOCid = Lateral occipital cortex, inferior division; LOCsd = Lateral occipital cortex, superior division; MTGad = Middle temporal gyrus, anterior division; MTGpd = Middle temporal gyrus, posterior division; OFG = Occipital fusiform gyrus; OP = Occipital pole; P = Putamen; PaG = Paracingulate gyrus; PC = Precuneus cortex; PG = Postcentral gyrus; PGad = Parahippocampal gyrus, anterior division; PGpd = Parahippocampal gyrus, posterior division; PrG = Precentral gyrus; SFG = Superior frontal gyrus; SGpd = Supramarginal gyrus, posterior division; SPL = Superior parietal lobule; T = Thalamus; TOF = Temporal occipital fusiform cortex.

Chapter 6

Microstructure Imaging of Crossing (MIX) White Matter Fibers from diffusion MRI

Diffusion MRI (dMRI) reveals microstructural features of the brain white matter by quantifying the anisotropic diffusion of water molecules within axonal bundles. Yet, identifying features such as axonal orientation dispersion, density, diameter, etc., in complex white matter fiber configurations (e.g. crossings) has proved challenging. Besides optimized data acquisition and advanced biophysical models, computational procedures to fit such models to the data are critical. However, these procedures have been largely overlooked by the dMRI microstructure community and new, more versatile, approaches are needed to solve complex biophysical model fitting problems. Existing methods are limited to models assuming single fiber orientation, relevant to limited brain areas like the corpus callosum, or multiple orientations but without the ability to extract detailed microstructural features. Here, we introduce a new and versatile optimization technique (MIX), which enables microstructure imaging of crossing white matter fibers. We provide a MATLAB implementation of MIX, and demonstrate its applicability to general microstructure models in fiber crossings using synthetic as well as ex-vivo and in-vivo brain data.

6.1 Introduction

This work presents a method to characterize tissue microstructure in the brain white matter from Diffusion MRI (dMRI) data. The method, named Microstructure Imaging of Crossing (MIX) [121, 122] white matter fibers, is unique in that it allows estimating detailed axonal features in the presence of multiple fiber orientations. This is achieved by exploiting a separable structure in the relevant optimization/ fitting problem, which enables the combination of stochastic search algorithms and gradient-based methods. Experiments are reported, that demonstrate the broad applicability of the method to existing microstructure models, as well as to extensions of these models to multiple fiber orientations. The experiments have been carried out using synthetic, ex-vivo, and in-vivo brain data. Results can be reproduced with a MATLAB implementation of MIX which has been made available online (http://www.cmrr.umn.edu/downloads/mix/).

Diffusion Tensor Imaging (DTI) [3] is the most widely used technique applicable to dMRI. It assumes a single compartment and is unable to distinguish diffusion patterns in heterogeneous biological compartments within a voxel. It only provides non-specific markers such as fractional anisotropy (FA) and mean diffusivity (MD) [4] that cannot resolve microstructure features [123]. During the past decade, a variety of multi-compartment models have been proposed, that aim to capture more accurately the variability of diffusion in sub-voxel regions, such as intra-axonal and extra-axonal compartments. In particular, these models directly provide information about the white matter microstructure [124,125]. They include ball-and-stick [24,45], CHARMED [126,127] (composite hindered and restricted model of diffusion), AxCaliber [128,129], ActiveAx [130,131] or the Minimal Model of White Matter Diffusion (MMWMD), NODDI [6] (Neurite Orientation Dispersion and Density Index) and DIAMOND [132]. The CONNECT [133] project recently leveraged the CHARMED, AxCaliber and ActiveAx models, in combination with tractography methods, to improve structural connectivity mapping methods.

Recent taxonomy studies have extensively compared existing multi-compartment analytical models using ex-vivo rat brain [134] and in-vivo human brain [135] dMRI data, and provided a ranking using the Bayesian information criterion (BIC). They rank three-compartment (intra-axonal, extra-axonal and isotropic) models consistently higher than models with fewer compartments. We note however that these

studies only consider models that are based on single fiber orientation, although 60 to 90 percent of the brain white matter have fiber crossings [136] (i.e., have several fiber bundles with different orientation in each voxel). The apparent obstacle in considering additional compartments and multiple fiber orientations can be traced back to the complex mathematical nature of the model functions and required fitting techniques.

Parameter estimation in multi-compartment models requires non-convex optimization even for a single axonal orientation. The task is further complicated by (i) the type of nonlinear relations between model parameters and dMRI data that render the problem ill-posed, and by (ii) the effect of noise, particularly in clinical (shorter) scans, which is significant.

Several methods and software packages have recently been published to estimate microstructural features from dMRI. Here, we focus on three toolkits which are specific to fitting complex (three-compartment) microstructure models: CAMINO [137] (Open-Source Diffusion-MRI Reconstruction and Processing), AMICO [8] (Accelerated Microstructure Imaging via Convex Optimization) and, NODDI [6]. Other toolkits include e.g. DSI Studio (`http://dsi-studio.labsolver.org/`), which implements the Restricted Diffusion Imaging (RDI) [138] method and CHARMED [126,127] (`http://neuroimaging.tau.ac.il/ya/charmed.html`). In CAMINO, AMICO and NODDI, fiber orientation is obtained using Diffusion Tensor (DT) estimation. The remaining parameters (specifying volume fractions, radius indices, density indices, orientation dispersion, etc.) are determined using nonlinear regression. The limitation of CAMINO, AMICO and NODDI in their ability to scale and handle multiple fiber orientations with complex biophysical models can be attributed to the fact they do not exploit the separable structure of the problem.

The need to reliably identify multiple fiber orientations per voxel, and associated microstructural parameters, has motivated recent attempts that are based on simple biophysical models [139–141]. For instance, AMICOx [139] (AMICO in fiber crossing) estimates axon diameter indices in two fiber orientations (synthetic data only, using ActiveAx model in two orientations [142]). However, the approach has not yet been shown to scale so as to incorporate additional fiber orientations, generic tissue models, and real dMRI data. A second approach [140] introduced the "spherical mean technique", capable of factoring out the effects of fiber crossing to estimate "per-axon"

parallel and perpendicular effective diffusion coefficients, and subsequently extract fiber orientation using spherical deconvolution. Similarly, estimation of NODDI in two directions [141] for tractography uses fiber orientation estimates from neighboring voxels. Despite these efforts, which highlight the critical need for novel computational techniques, a method to fit complex biophysical models to dMRI data (e.g. with ten or more parameters) would significantly improve our ability to quantify intrinsic white matter properties of the human brain.

We introduce a novel regression method, which is robust and versatile. It enables to fit existing biophysical models with improved accuracy, and open the possibility to propose and test new models that were impossible to solve until now. It utilizes the Variable Separation Method [143] (VSM) to distinguish parameters that enter in both, linear and non-linear manner, in the model (Methods). The estimation of non-linear parameters is a non-convex problem and is handled first. This is done by stochastic search that utilizes Genetic Algorithms (GA); GAs have been shown to be effective in approximating exponential time series models [144]. The task to estimate linear parameters amounts to a convex problem and can be solved using standard least squares techniques. These parameter estimates provide a starting point for a Trust Region method in search for a refined solution. A detailed description of the algorithm is given in the Methods section.

Specific attributes of the new method (MIX) are: (i) ability to estimate generic model parameters with multiple fiber orientations, (ii) ease in initializing the parameter search, and (iii) ability to cope with complex/realistic models that may have more than four compartments (an apparent limit for all existing methods). With MIX, it is only the amount of available dMRI data that may impact the variance of parameter estimates and, thereby, indirectly limit the achievable number of compartments.

6.2 Tissue Compartment Model Parameters and Functions

This section briefly presents compartment model functions, describing the dMRI signal attenuation due to diffusion in a certain tissue geometry. We limit the description only to the functions used in this work, please see [134] for a comprehensive list of available tissue compartment models. We describe the parameters of the model functions in Section 6.2.1 followed by the model functions in Section 6.2.2. Paremeters of the PGSE have already been described in Section 2.4.1.

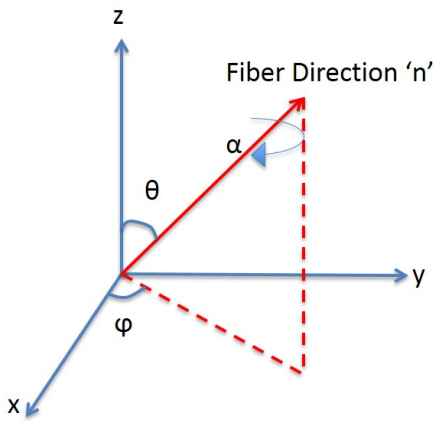


Figure 6.1: Fibre Orientation

6.2.1 Parameters of the Compartment Model Functions

- *Fiber orientation vector 'n'*: Fiber orientation vector is given by (6.1) where θ, ϕ and α are angles in radians as shown in Fig. 6.1.

$$n = \left[\cos\phi \sin\theta; \quad \sin\phi \cos\theta; \quad cos\theta \right] \tag{6.1}$$

- $d_{\parallel}$ $(\mu m^2.sec^{-1})$ = Diffusivity in direction parallel to 'n' (Intrinsic free diffusivity).
- $d_{\perp 1}$ $(\mu m^2.sec^{-1})$ = Diffusivity in direction perpendicular to 'n'.
- $d_{\perp 2}$ $(\mu m^2.sec^{-1})$ = Diffusivity in direction perpendicular to 'n' and $d_{\perp 1}$.
- d_{iso} $(\mu m^2.sec^{-1})$ = Isotropic diffusivity (fixed to $2e^3$ $\mu m^2.sec^{-1}$ for ex-vivo and $3e^3$ $\mu m^2.sec^{-1}$ for in-vivo).
- R (μ m) = Axon radius index [130].
- v_{ic} = Intra-cellular volume fraction [6].
- v_{iso} = Isotropic volume fraction [6].
- OD = Orientation dispersion index [6].

6.2.2 Biophysical Functions for Compartment Models

- ***Cylinder model function*** ($S_{cylinder}$):

$$S_{cylinder} = \Delta - (\delta/3)(\gamma\delta G)^2 d_{\parallel} 2\gamma^2(\hat{G}^2 - (\hat{G}.n)^2) \sum_{i=1}^{\infty} \frac{2d_{\parallel}\beta_i^2\delta - 2 + 2e^{-d_{\parallel}\beta_i^2\Delta} - e^{-d_{\parallel}\beta_i^2(\Delta-\delta)} - e^{-d_{\parallel}\beta_i^2(\Delta+\delta)} + 2e^{-d_{\parallel}\beta_i^2\delta}}{d_{\parallel}^2\beta_i^6(\beta_i^2R^2 - 1)} \quad (6.2)$$

  Where $\hat{G} = G.g$ and β_i is the i th root of the equation, $J'(\beta_i R) = 0$, and J' is the derivative of Bessel function of the first kind, order one.

- ***Zeppelin model function*** ($S_{zeppelin}$):

$$S_{zeppelin} = b(d_{\parallel} - d_{\perp})(g.n)^2 + d_{\perp} \quad (6.3)$$

- ***Tensor model function*** (S_{tensor}):

$$S_{tensor} = e^{-bg'Dg} \quad (6.4)$$

  Where $D \rightarrow S^3_{++}$ (symmetric positive definite 3 x 3 matrix).

- ***Stick model function*** (S_{stick}):

$$S_{stick} = e^{-bd_{\parallel}(nG)^2} \quad (6.5)$$

- ***Ball model function*** (S_{ball}):

$$S_{ball} = e^{-bd_{iso}} \quad (6.6)$$

- ***Dot model function*** (S_{dot}):

$$S_{dot} = e^{b_0=0} = 1 \quad (6.7)$$

6.3 Results

6.3.1 Synthetic data experiments

We have evaluated MIX on synthetic data and compared the results with the CAMINO, AMICO and NODDI toolboxes. Four different tissue models were considered and are described in following sub-sections with details on specific compartment model functions provided in Section 6.2.2. Not all of these methods support all four tissue models, with the exception of MIX. We follow the model naming convention used in the dMRI microstructure literature [8, 134, 135]. Detailed description of the experiments is as follows:

6.3.1.1 ActiveAx

ActiveAx is a four compartment model i.e, Zeppelin-Cylinder-Ball-Dot [130, 134]. The model function showing only the parameters to be estimated (unknown parameters) is given below:

- *Model function*: The estimated dMRI normalized signal $\widehat{S}_{ActiveAx}$ is assumed to be coming from the following four compartments:

$$\widehat{S}_{ActiveAx} = f_1 S_{cylinder}(R, \theta, \phi) + f_2 S_{zeppelin}(d_\perp, \theta, \phi) + f_3 S_{ball} + f_4 S_{dot}. \quad (6.8)$$

  Where f_is are the volume fractions of each compartment such that $\sum_{i=1}^{4} f_i = 1$.

- *Parameters to be estimated*: Eight $(f_1, f_2, f_3, f_4, R, d_\perp, \theta, \phi)$.

- *Generating synthetic data*: True values were generated randomly and uniformly spanning complete parameters space (given for each parameter as below) using MATLAB for 100 different substrates. Detail for true parameters range is as under:

$$f_i \in [0.01, 1] \quad \text{such that} \sum_{i=1}^{4} f_i = 1$$
$$R \in [0.1, 10] \ \ (\mu m), \quad \theta, \phi \in [0.01, \pi] \ \ (rad)$$
$$d_\perp = (1 - f_1) d_\parallel \ \ (\mu m^2.sec^{-1}) \quad \text{where} \ \ d_\parallel = 0.6e3 \ (\mu m^2.sec^{-1}) \ \text{fixed for ex-vivo.}$$

- *Noise*: Rician noise was added to signal for each substrate with SNR=1000,

20 and 8 with ten different realizations for each SNR value, i.e., we have 3000 different noise realizations.

- *Protocol*: An optimized protocol for ActiveAx given in the original study [130] was used for synthetic data generation and subsequently for parameter estimation. The protocol has four shells and a total of 372 measurements (including b_0 non-diffusion weighted measurements).

- *Results and discussion*: We used MIX to estimate the parameters for the ActiveAx model. Although the model has only eight unknown parameters, we find that the absolute error between estimated and actual parameters is significantly lower for MIX, for SNR values of 1000, 20 and 8, as compared to CAMINO, AMICO and NODDI toolboxes (Figure 6.2). In addition, we find that AMICO tends to over-estimate low axonal density values (below 0.01) with ActiveAx, and to provide limited contrast in radius values (which are limited to the range 4 to 6 microns). We hypothesize that this may be due to the required regularization in AMICO. MIX is not affected by such requirement and provides estimates that are closer to the ground truth and CAMINOs estimates.

6.3.1.2 NODDI

- *Model function*: NODDI is a three compartment model [6] and estimated dMRI normalized signal $\widehat{S}_{Noddi}$ is given by the following relation:

$$\widehat{S}_{Noddi} = (1 - v_{iso})(v_{ic}S_{ic}(OD, \theta, \phi) + (1 - v_{ic})S_{ec}(d_{\perp}, \theta, \phi) + v_{iso}S_{dot}. \quad (6.9)$$

  Where S_{ic} and S_{ec} are the normalized signals from intra-cellular and extra-cellular compartments respectively.

- *Parameters to be estimated*: Six $(v_{ic}, v_{iso}, OD, d_{\perp}, \theta, \phi)$.

- *Generating synthetic data*:
  Synthetic data was generated using "WatsonSHStickTortIsoV-B0" model i.e. using NODDI MATLAB model functions "SynthMeasWatsonSHCylNeuman-PGSE" as intra-cellular model, "SthMeasWatsonHinderedDiffusion-PGSE" as extra-cellular model and S_{dot} as isotropic compartment model. Ground truth values for the six parameters were generated randomly and uniformly over the

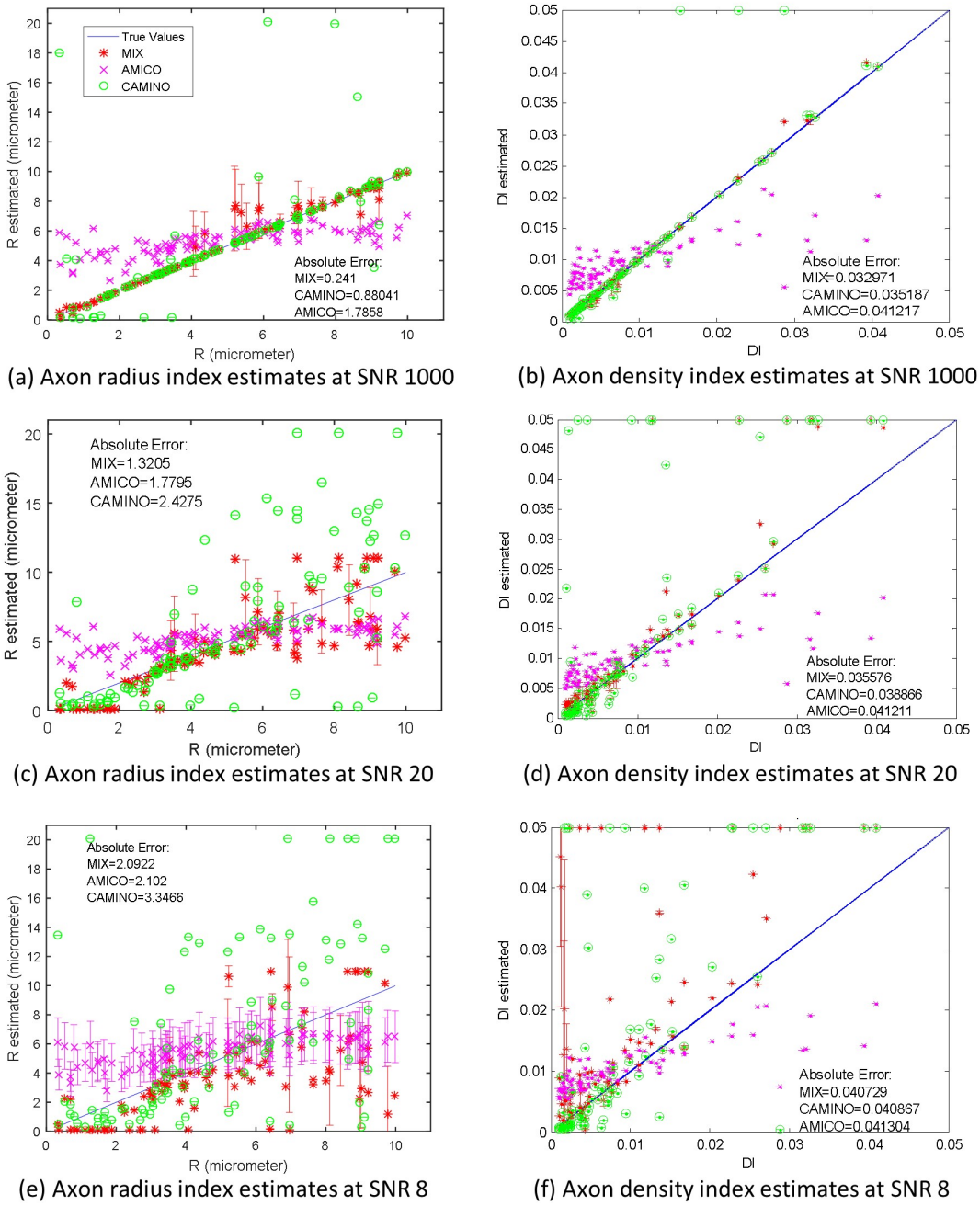


Figure 6.2: Axon radius index and axon density index estimates comparison at SNR=1000, 20 and 8. The plots are 'errorbars' showing mean and stanfard deviation of the ten estimation results, each with different noise realization for every substrate (i.e., 1000 noise realizations for each noise level). Further, for CAMINO, each estimate is already a mean of 100 samples using MCMC fitting algorithm, as described at CAMINO website (http://camino.cs.ucl.ac.uk/index.php?n=Tutorials.ActiveAx).

following parameter space:

$$v_{ic}, v_{iso} \in [0.01, 0.9],$$
$$OD \in [0.1, 1], \quad \theta, \phi \in [0.01, \pi] \ \ (rad)$$
$$d_\perp = (1 - v_{ic}) d_\parallel \ \ (\mu m^2.sec^{-1}) \quad \text{where} \ \ d_\parallel = 1.7e3 \ (\mu m^2.sec^{-1}) \ \text{fixed for in-vivo.}$$

- *Noise*: Rician noise was added to each substrate with SNR = 1000, 20 and 8. Ten different noise realizations for each SNR value were added to the substrates giving 3000 different noise realizations.

- *Protocol*: We used the optimized protocol for NODDI given in original study [6] for synthetic data generation and parameter estimation. The protocol has two b-values and a total of 81 measurements (including b_0 measurements).

- *Results and discussion*: Results for the three algorithms are shown in Figure 6.3. We note that estimation of orientation dispersion and intra-cellular volume fractions are particularly challenging for all methods, respectively for high dispersion and low volume fractions. MIX outperforms the other methods for orientation dispersion estimation while AMICO performs better for intra-cellular volume fractions.

6.3.1.3 Tensor-Stick-Dot (TSD)

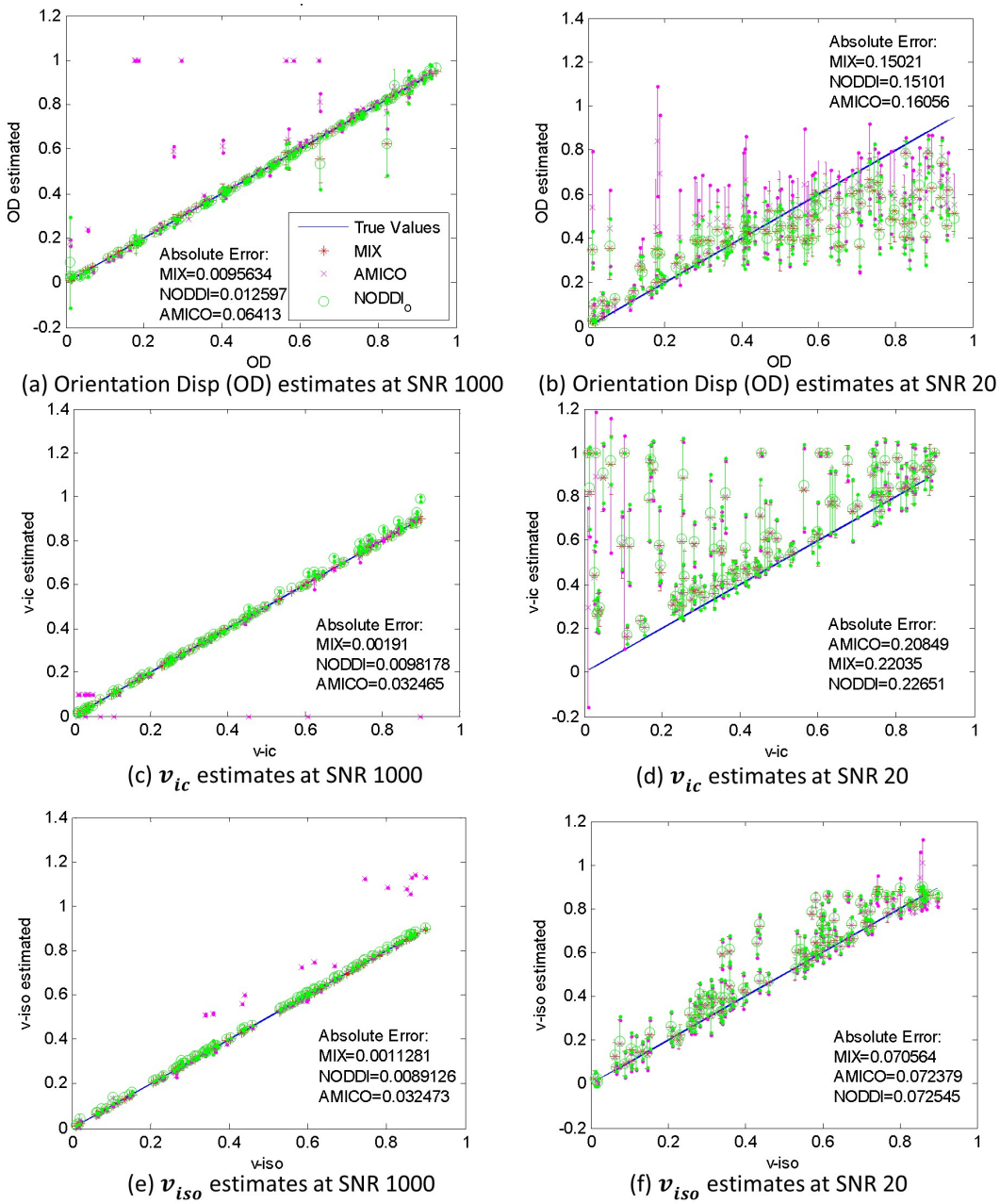
- *Model function*: The estimated dMRI signal $\widehat{S}_{TSD}$ function is:

$$\widehat{S}_{TSD} = f_1 S_{stick}(d_\parallel, \theta, \phi) + f_2 S_{tensor}(d_\parallel, d_{\perp 1}, d_{\perp 2}, \theta, \phi, \alpha) + f_3 S_{dot}. \tag{6.10}$$

  Here we do not fix values for $d_\parallel, d_{\perp 1}, d_{\perp 2}$.

- *Parameters to be estimated*: Nine $(f_1, f_2, f_3, d_\parallel, d_{\perp 1}, d_{\perp 2}, \theta, \phi, \alpha)$.

- *Generating synthetic data*: Similar to previous experiments, synthetic data was generated using CAMINO datasynth command for 100 substrates. Since no optimized protocol is available for the model, we used the ActiveAx protocol (as in ActiveAx experiment). Using MATLAB, ground truth values of the parameters were randomly and uniformly generated over the parameter space,



(a) Orientation Disp (OD) estimates at SNR 1000

(b) Orientation Disp (OD) estimates at SNR 20

(c) $\boldsymbol{v_{ic}}$ estimates at SNR 1000

(d) $\boldsymbol{v_{ic}}$ estimates at SNR 20

(e) $\boldsymbol{v_{iso}}$ estimates at SNR 1000

(f) $\boldsymbol{v_{iso}}$ estimates at SNR 20

Figure 6.3: NODDI parameters estimation comparison for the three algorithms. The plots are 'errorbars' showing mean and standard deviation of ten estimation results, each with different noise realization for every substrate (i.e., 1000 noise realizations). Axonal orientation dispersion OD and Intracellular volume fraction v_{ic} estimation results are in-line with the results reported in Figure 7. of [8], i.e., with the increase in noise, higher values of OD are under-estimated while lower values of v_{ic} are over-estimated. Absolute errors of each algorithm have been reported in each sub-figure for comparison. It can be seen that MIX has the least error except for v_{ic} at SNR = 20 estimation only.

which is given below:

$$\theta, \phi, \alpha \in [0.01, \pi] \;\; (rad)$$
$$d_{\parallel} \in [0.55e3, 1.8e3] \; (\mu m^2.sec^{-1}) \quad \text{while} \quad d_{\parallel} > d_{\perp 1} > d_{\perp 2}$$
$$f_i \in [0.01, 1] \quad \text{such that} \sum_{i=1}^{3} f_i = 1$$

- *Noise*: Ten different noise realizations (at SNR=1000) were generated and added to each substrate to have 1000 different noise realizations in total.

- *Results and discussion*: Tensor-Stick-Dot model has nine unknown parameters, we only compare MIX with CAMINO as AMICO has not yet been implemented for such a model. For the experiment, MIX was vastly superior to CAMINO, even when using 100 Levenberg-Marquardt runs, both in terms of accuracy as well as computational efficiency (Figure. 6.4), when estimating perpendicular and parallel diffusivities.

6.3.1.4 ZCDx

- *Model function*: In order to demonstrate the ability of our method to recover axonal microstructure information in complex white matter configurations, we hypothesize three fiber orientations in a voxel. Specifically, we assume that a zeppelin and a cylinder representing extra-axonal and intra-axonal signal respectively, in each of the three orientations. This can be seen as an extension of ActiveAx [130] model for three different fiber orientations in a voxel. The estimated dMRI signal $\widehat{S}$ is assumed to be composed of normalized signals from the following seven compartments:

$$\begin{aligned}\widehat{S}_{ZCDx} = {} & f_{11} S_{cylinder1}(R_1, \theta_1, \phi_1) + f_{12} S_{zeppelin1}(d_{\perp 1}, \theta_1, \phi_1) \\ & + f_{21} S_{cylinder2}(R_2, \theta_2, \phi_2) + f_{22} S_{zeppelin1}(d_{\perp 1}, \theta_2, \phi_2) \\ & + f_{31} S_{cylinder3}(R_3, \theta_3, \phi_3) + f_{32} S_{zeppelin3}(d_{\perp 3}, \theta_3, \phi_3) + f_3 S_{dot}. \end{aligned} \quad (6.11)$$

- *Parameters to be estimated*: Nineteen, give as under:
$f_{11}, f_{12}, f_{21}, f_{22}, f_{31}, f_{32}, f_3, R_1, d_{\perp 1}, \theta_1, \phi_1, R_2, d_{\perp 2}, \theta_2, \phi_2, R_3, d_{\perp 3}, \theta_3, \phi_3$.

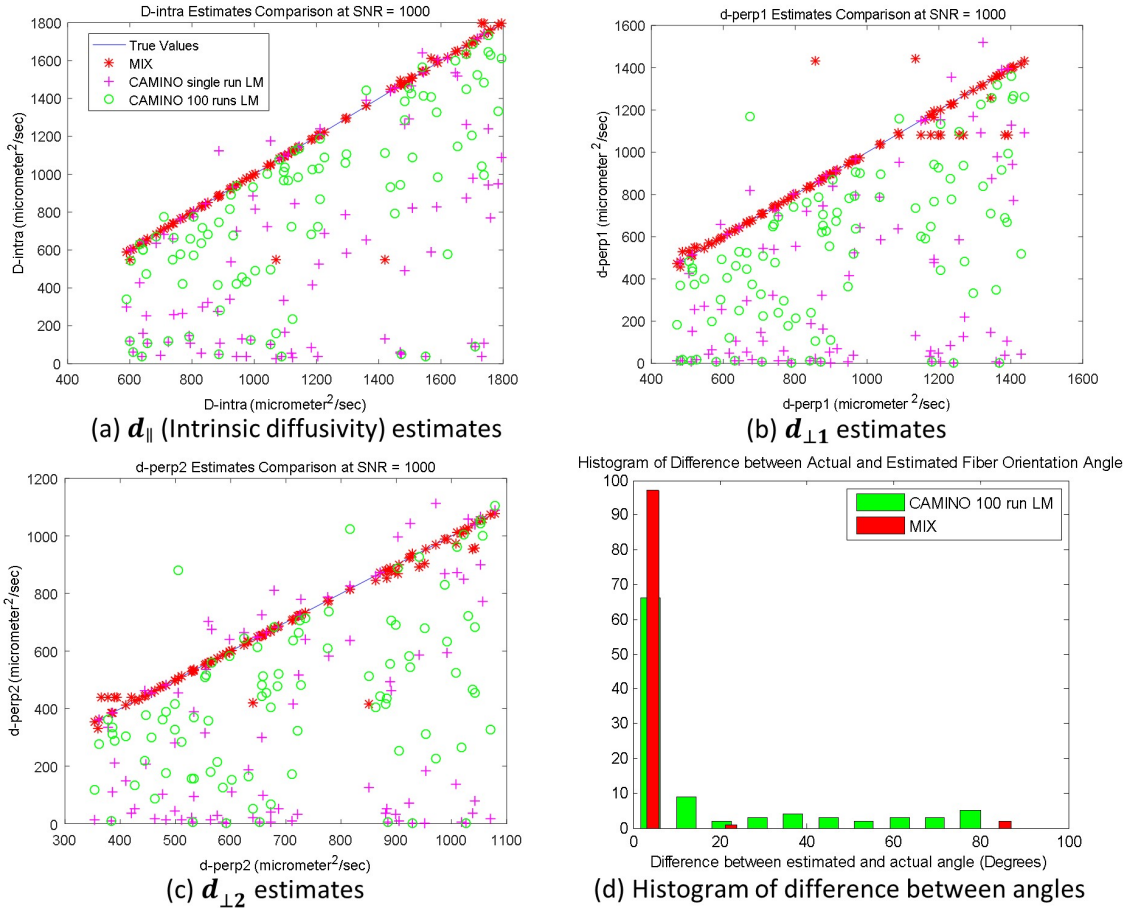


Figure 6.4: Tensor-Stick-Dot parameter estimation comparison betweeen MIX and CAMINO at SNR=1000. MIX results shown are for single estimation run. For CAMINO, results are shown for single run LevenbergMarquardt (LM) and 100 runs LM (using CAMINO's "MULTIRUNLM"). Subfigures (a), (b) and (c) show estimates of $d_{\parallel}$, $d_{\perp 1}$ and $d_{\perp 2}$ ($\mu m^2.sec^{-1}$) respectively while (d) shows fiber orientation estimation comparison (in degrees). It can be seen that all parameters are more accurately by MIX, as compared to CAMINO.

- *Generating synthetic data*: Synthetic data was generated using CAMINO datasynth command for 100 substrates. True values (ground truth values) of the parameters were randomly and uniformly generated over the parameter space (as given below), using MATLAB:

$$f_{ik} \in [0.01, 1] \quad \text{such that} \sum_{i=1}^{3} \sum_{k=1}^{4} f_{ik} + f_3 = 1$$
$$R_k \in [0.1, 10] \ \ (\mu m), \quad \theta_k, \phi_k \in [0.01, \pi] \ \ (rad),$$
$$d_{\parallel k} = 1.7e3 \ \ (\mu m^2.sec^{-1}) \ \text{ for in-vivo}, \quad d_{\perp k} = (1 - f_{1k}) d_{\parallel k},$$

where k = 1,2,3 represent number of fiber orientations inside a voxel.

- *Noise*: Similar to all the experiments above, Rician noise was added to each substrate with SNR=1000 and SNR=20 i.e., we have 2000 different noise realizations.

- *Protocol*: Protocol used for this experiment was made available for the White Matter Modeling Challenge (http://cmic.cs.ucl.ac.uk/wmmchallenge/) for predicting dMRI data. The protocol has 36 shells and a total of 3612 measurements (including b_0 measurements).

- *Results and discussion*: Figure 6.5 shows that MIX provides an accurate and robust parameter estimates even at a reasonably low signal to noise ratio (SNR = 20). Since no other method is capable of estimating parameters for ZCDx, a comparison could not be carried out. However, this illustrates the ability of MIX to reliably estimate the parameters of a complex model, which might provide novel insight into the brain microstructure.

Computational Efficiency: From a computational standpoint, throughout our experiments AMICO was the most efficient for ActiveAx and NODDI. However, for more complex biophysical models, AMICO has not yet been implemented, as it relies on search over a dictionary of nonlinear dependences, which appears prohibitively expensive. For problems amenable to CAMINO, MIX is in general twice as fast.

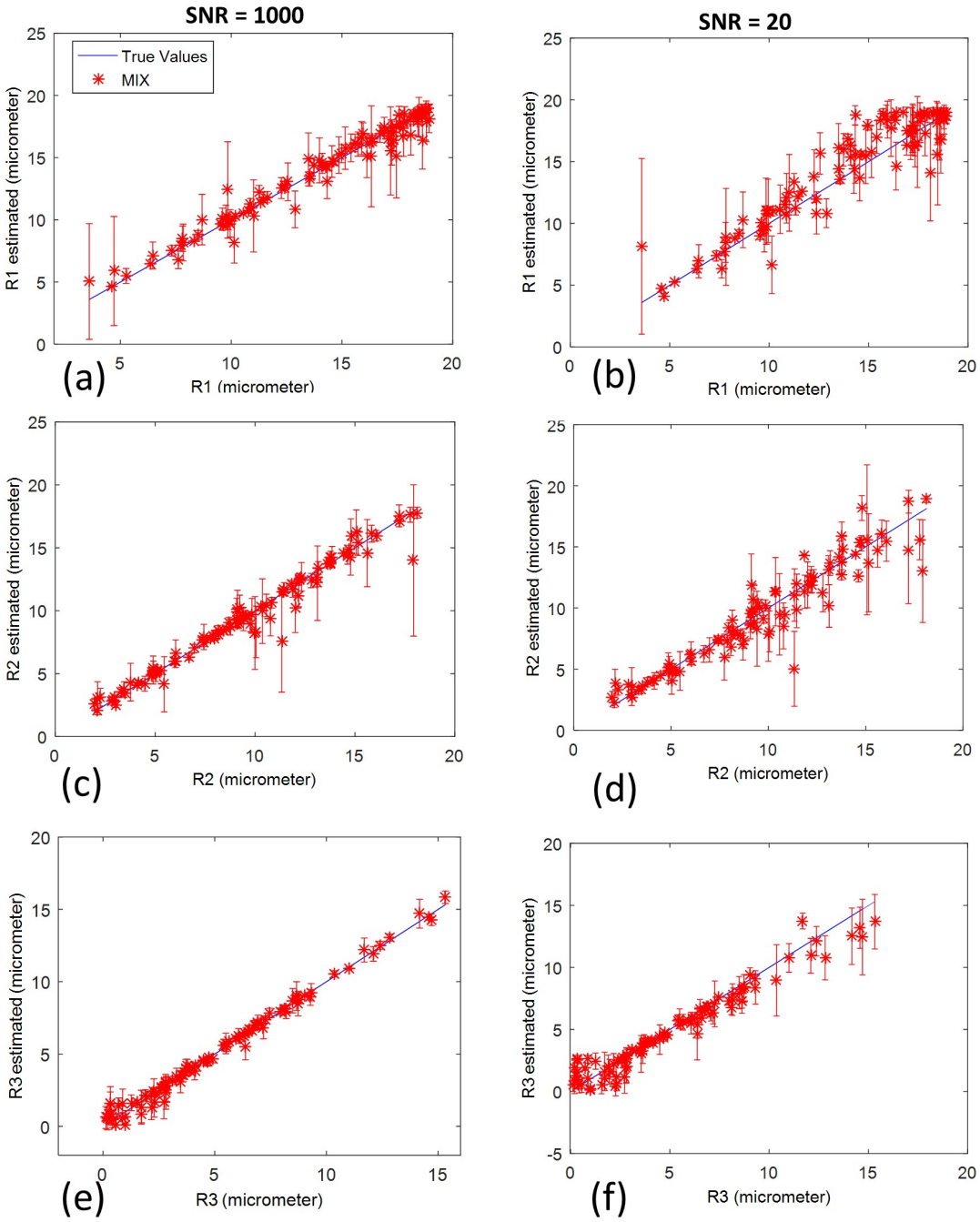


Figure 6.5: Axon radius index estimates in three orientations using ZCDx model. R_1, R_2 and R_3 estimates are associated with the three different fiber orientations in a voxel. Results shown are "errorbar" plots of 10 fitting results for both SNR= 1000 (left column) and SNR = 20 (right column) data-sets. It can be seen that radius index can be estimated with good accuracy even at a reasonably low signal to noise ratio. With increased noise in the signal, there is a trend of over-estimating low radii and under-estimating larger radii.

6.3.2 Real data experiments

6.3.2.1 ActiveAx

In the first experiment, we used ex-vivo monkey brain data [145] from the original ActiveAx study [130]. Briefly, data (http://dig.drcmr.dk/activeax-dataset/) was obtained using a Varian 4.7T system with voxel size 0.4 mm3, four b-values each with 90 directions and with the two lowest identical, and 12 additional b=0 volumes. b-values and corresponding parameters are as follows: $b_1 = 1930\ s.mm^{-2}$ ($\Delta/\delta = 16$ / 10 ms, $|G| = 140\ mT.m^{-1}$), $b_2 = b_1$, $b_3 = 3090\ s.mm^{-2}$ ($\Delta/\delta = 45$ / 7 ms, $|G| = 131\ mT.m^{-1}$), $b_4 = 13190\ s.mm^{-2}$ ($\Delta/\delta = 35$ / 17 ms, $|G| = 140\ mT.m^{-1}$) [145]. This experiment focused on fitting the ActiveAx model to a mid-sagittal slice of the corpus callosum (CC) using CAMINO, AMICO and MIX toolboxes. Axonal diameter index estimates show that all the algorithms exhibit similar patterns throughout the CC (Figure. 6.6 d, f). More specifically, smaller axonal diameter indices and densely packed axons were identified in the genu and splenium, with larger axonal diameter indices and lower axonal densities in the mid-body. MIX, however, consistently provides smaller axonal diameter indices throughout the CC (Figure. 6.6 c). This appears to be more realistic and in line with recent studies [146] showing that axonal radii are always over-estimated using dMRI and existing scanning and fitting methods. The improvement appears to have been gained by our proposed data fitting method (MIX).

6.3.2.2 NODDI

In the second experiment, we considered the NODDI model for in-vivo human brain data [6] from the original NODDI study. The data was obtained on a 3T Philips Achieva system with voxel size 2 mm^3 and two b-values of 711 $s.mm^{-2}$ with 30 directions, and 2855 $s.mm^{-2}$ with 60 directions. Nine b=0 volumes are included in the protocol and Δ/δ are fixed to 37.8 / 17.5 ms. $|G|$ is varied to achieve the desired b-values.

We compared MIX with AMICO and NODDI and found that the three algorithms show similar trends for the estimated parameters in the white matter (Figure 6.7). Largely, MIX and NODDI are in very close agreement with regard to fiber Orientation Dispersion index (OD), while MIX offers better contrast for intracellular volume fraction and isotropic volume fraction (Figure 6.7, lower right four figures, see

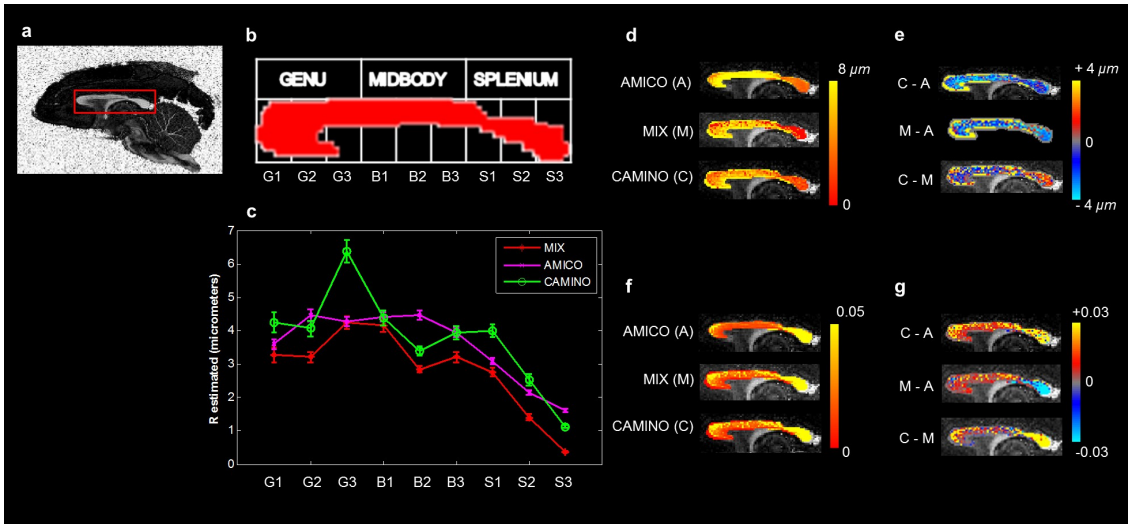


Figure 6.6: ActiveAx ex-vivo parameter estimation comparison using fixed monkey brain data. (a) Region of interest (in red box). (b) Partitions of the corpus callosum for this study. (c) Error plots (showing mean and standard deviation) for the axon radius index estimates in each partition. (d) Axon Diameter Index (D) estimates. (e) Differences between D estimates (f) Axon Density Index (DI) estimates. (g) Differences between DI estimates.

e.g. frontal white matter) as compared to the other two. In the cerebro-spinal fluid (CSF), MIX gives high values for intra-cellular volume fraction, whereas the other two predict the exact opposite. The discrepancy can be attributed to the fact that MIX explicitly constrains the sum of the volume fractions to be one. Therefore, estimates by MIX are relative to extracellular volume fraction only, and not to isotropic volume fraction. Nonetheless, estimated isotropic volume fraction is high in CSF areas, as expected, and in agreement with AMICO and NODDI.

6.3.2.3 Zeppelin-Cylinder-Cylinder-Dot (ZCCD)

In the third experiment, we used the ZCCD model to demonstrate the performance of MIX in multiple axonal orientations. The choice of model is intended to allow improved modeling of complex tissue geometry of in-vivo human brain. Cylinders model water diffusion in two intra-axonal bundles (if they exist in a given voxel) and the zeppelin represents extra-axonal diffusion. Both the cylinders and zeppelin can have arbitrary orientations. Additionally, if the data does not support more than one or two fiber orientations, the corresponding volume fractions will be automatically driven to zero.

Data was acquired on a healthy volunteer using a Siemens 3T Skyra system with

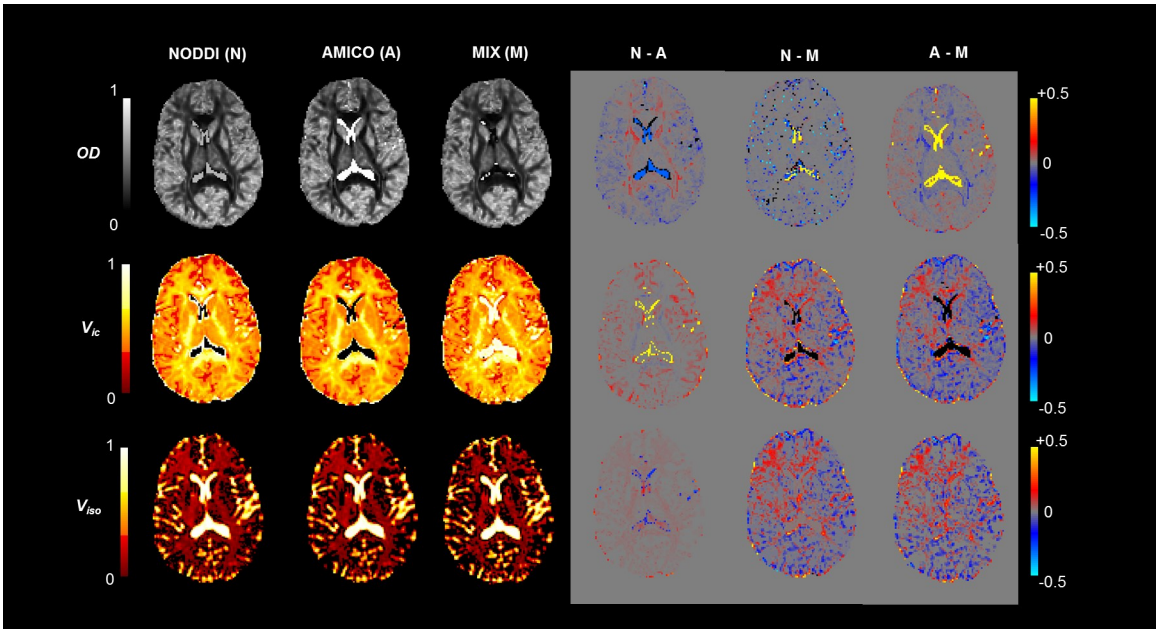


Figure 6.7: NODDI parameters estimation comparison for in-vivo human brain data. Fiber Orientation Dispersion (OD), intra-cellular volume fraction (v_{ic}) and isotropic volume fraction (v_{iso}) estimates are shown in first, second, and third row respectively. Left most column shows estimates by NODDI toolbox (N), second from the left by AMICO (A), third column from the left by MIX (M). The three columns on the right side show differences between the estimates of the three algorithms.

voxel size 2 mm^3, four b-values, each with 119 directions, and 18 additional b=0 volumes. b-values and the corresponding parameters were chosen as follows: $b_1 = 820\ s.mm^{-2}$ ($\Delta/\delta = 17.6\ /\ 9\ ms$, $|G| = 98.5\ mT.m^{-1}$), $b_2 = 980\ s.mm^{-2}$ ($\Delta/\delta = 55.5\ /\ 5.2\ ms$, $|G| = 97.1\ mT.m^{-1}$), $b_3 = 3010\ s.mm^{-2}$ ($\Delta/\delta = 38.5\ /\ 22.2\ ms$, $|G| = 52.4\ mT.m^{-1}$) and $b_4 = 7600\ s.mm^{-2}$ ($\Delta/\delta = 37.8\ /\ 29.3\ ms$, $|G| = 66.6\ mT.m^{-1}$). The study was approved by the local Institutional Review Board and informed consent was obtained from the research participant.

Results show that in the corpus callosum (CC), both cylinders and zeppelin are aligned. In the area, one of the cylinders and the zeppelin have very small densities, therefore their associated axonal radii estimates and perpendicular diffusivity are meaningless and axonal radii estimates can be inferred from the cylinder with non-zero volume fraction (e.g. density). These findings conform to the well-known single dominant fiber orientation in the CC area (Figure. 6.8 b-e). In contrast, in the centrum semi-ovale (CSO), where callosal, cortico-spinal and superior longitudinal fasciculus fibers cross, cylinders and zeppelin have distinct orientations and

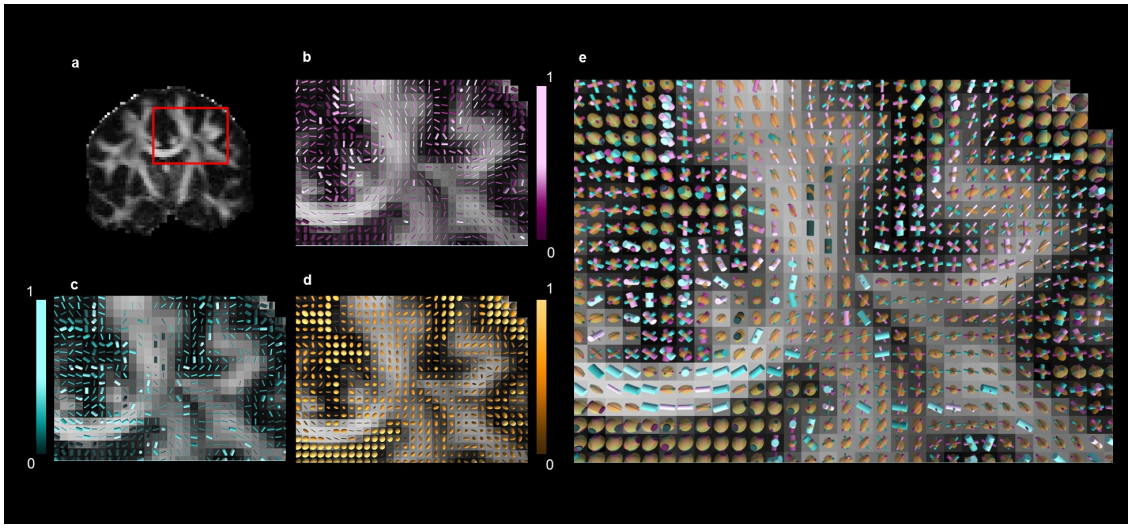


Figure 6.8: Zeppelin-Cylinder-Cylinder-Dot parameter estimation for in-vivo human brain data. (a) FA map for coronal slice and region of interest marked in red. (b) Cylinder 1. (c) Cylinder 2. (d) Zeppelin. (e) Zeppelin-Cylinder-Cylinder super-imposed. Color intensities show volume fractions for each compartment respectively. Both cylinders show estimated radii (μm) while zeppelin glyph thickness shows perpendicular diffusivity $d_\perp$ ($\mu m^2.sec^{-1}$).

significant volume fractions, thereby reflecting the existence of multiple white matter pathways. We note that, for some voxels where these three major pathways are known to simultaneously exist and cross, the zeppelin tend to align with the third orientation. This is due to the fact that the ZCCD model only has two cylinders to model intra-axonal compartments and that, at these particular locations, the data more strongly supports a third intra-axonal, rather than the extra-axonal compartment. In the CSF, zeppelins have the highest volume fraction and large perpendicular diffusivity (Figure 6.8 d) while cylinders have very low volume fraction, capturing the largely free and isotropic diffusion occurring at this location. These findings can also be related to FA values (Figure 6.8 a) in CSF, which are very small since there is no dominant fiber orientation in this region. Although the results offer more comprehensive information (Figure. 6.8 e), as compared to any single orientation biophysical model, further improvements in estimates accuracy could be made with model specific scanning protocol optimization.

6.3.2.4 NODDIx

Finally, in the fourth experiment, we use our proposed NODDIx model (Supplementary Note 6) to estimate in-vivo orientation dispersion (OD), intra-cellular (v_{ic})

and isotropic (v_{iso}) volume fractions in multiple fiber pathways. NODDIx is an extension of the NODDI model, which uses thirteen parameters to capture OD, v_{ic}, v_{iso} and fiber orientations in up to two fiber pathways, and which can be accurately estimated with MIX. We use data from the HCP [37] collected on a Siemens 3T Skyra system with voxel size 1.25 mm^3, three b-values (1000, 2000, 3000 $s.mm^{-2}$), each with 90 directions, and 18 additional b=0 volumes. Δ/δ are fixed to 43.1 / 10.6 ms, while $|G|$ ($G_{max} = 97.4\ mT.m^{-1}$) is varied to achieve the desired b-values.

As shown in Figure 6.9, NODDIx interestingly enables the differentiation of cortical/ sub-cortical areas (i.e. mostly dendrites) or complex white matter areas like the centrum semi-ovale (i.e. crossings) with high orientation dispersion, from simpler white matter areas like the corpus callosum, cortico-spinal tract, etc. with low orientation dispersion. Contrary to NODDI, v_{ic} stays high in areas of high OD, presumably because the dMRI signal is better explained by the NODDIx model. This improved fit can also be concluded from the much lower residual values for v_{iso} throughout the brain, while still correctly identifying CSF outside the cortex and in the ventricles.

6.4 Methods

6.4.1 Problem Formulation

Define $\widehat{S}(x, f)$ as the model predicted normalized dMRI signal from 'n' different tissue compartments i.e.

$$\widehat{S} = f_1 S_1(x_1, \ldots, x_{m1}) + f_2 S_2(x_{m1+1}, \ldots, x_{m2}) + ... + f_n S_n(x_{m2+1}, \ldots, x_m) \quad (6.12)$$

where $S_1, S_2 \ldots S_n$ are selected bio-physical models for intra-axonal, extra-axonal, cerebrospinal fluid (CSF) and glial cells compartments etc. [134, 135]. These models depend upon 'm' different parameters given in vector $x = [x_1\ x_2\ \ldots x_m]^T$. While $f = [f_1\ f_2\ \ldots f_n]^T$ is the vector containing volume fractions of the n tissue compartments. Volume fractions are positive and they all sum up to one. Objective function for the model parameter estimation from dMRI data, with implicit assumption of offset Gaussian noise as used in the literature [8, 134, 135] is as follows:

$$\min_{x,\, f} \parallel S - (\widehat{S}(x, f) + \sigma) \parallel_2^2 \quad (6.13)$$

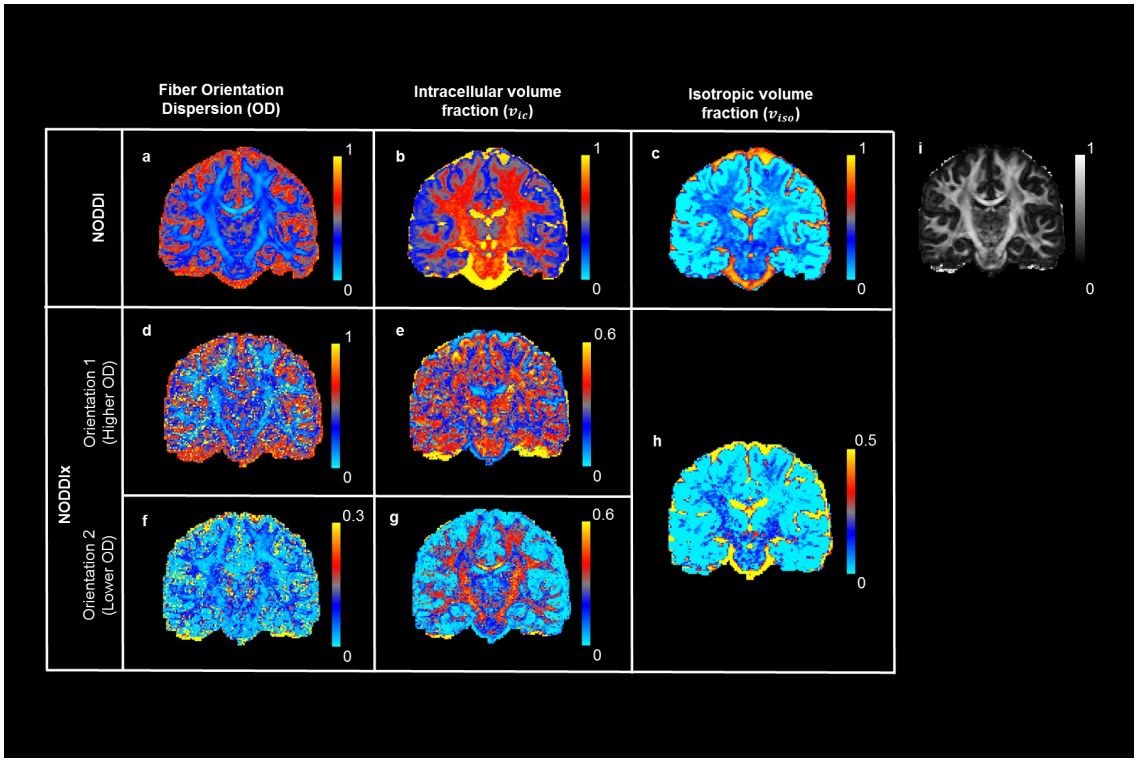


Figure 6.9: NODDI and NODDIx parameter estimation comparison using Human Connectome Project (HCP) data. First row shows results for NODDI model fitting while lower rows show results for NODDIx. Estimated model parameters in the two orientations for NODDIx have been sorted for OD values. It can be seen that v_{ic} for higher OD in (e) shows mostly grey matter while v_{ic} for lower OD in (h) shows white matter. (a) Fiber Orientation Dispersion (OD) (b) v_{ic} (c) v_{iso} (d) OD (higher values) along orientation 1. (e) v_{ic} for higher OD (orientation 1). (f) OD (lower values) along orientation 2. (g) v_{ic} for lower OD (orientation 2). (h) v_{iso}. (i) FA map.

$$\text{such that} \quad \sum_{i=1}^{n} f_i = 1, \qquad f_i \geq 0 \quad i = 1, 2 \ldots n$$
$$x_j^{min} \leq x_j \leq x_j^{max} \quad j = 1, 2 \ldots m$$

where 'S' represents the normalized dMRI measurements and 'σ' is the standard deviation of noise (assumed to be a constant, calculated a priori, and without loss of generality can be added to S). As defined in (6.12), $\widehat{S}(x, f)$ represents the estimated signal from multicompartment tissue model, while f_i s are volume fractions of 'n' compartments. x_j^{min} and x_j^{max} represent lower and upper bounds for unknown deterministic variable vector 'x' respectively. Diffusion weighted MRI data inherently has Rician noise [123]. However, with an assumption of offset Gaussian noise model, the objective function, as given in (6.13), becomes simple and more stable numerically

than Rician log-likelihood function [8, 134, 135].

6.4.2 Solution Framework

MIX can be divided into four main steps. First separate linear parameters 'f' from non-linear parameters 'x' by projecting variables. The parameters 'x' estimation problem is a non-linear one and can be solved by stochastic search procedures, here we use Genetic Algorithms (GA). Once estimates of 'x' are obtained, searching for 'f' is linear least squares estimation. Once estimates of 'x' are obtained, searching for 'f' is a linear least-squares estimation problem. Finally, using estimates from step 2 and 3, as starting point, we perform search for both 'x' and 'f' using the Trust Region Method. Below is the detailed description of each step:

- **Step 1.** ***Variable Separation***. The objective function described in (6.13) has a separable structure. Which can be exploited to separate the variables by variable separation method [144]. We can re-write our objective function in (6.13) in the following form:

$$\min_{x,f} \parallel S - \Phi(x)f \parallel_2^2 \tag{6.14}$$

$$\text{where } \Phi(x) = [S_1(x_1, \ldots, x_{m1}) \quad S_2(x_{m1+1}, \ldots, x_{m2}) \ldots S_n(x_{m2+1}, \ldots, x_m)]$$

$$f = \Phi^{\dagger}(x)S \tag{6.15}$$

$$\text{where } \Phi^{\dagger}(x) \text{ is Moore-Penrose inverse of } \Phi(x)$$

$$\text{i.e.,} \quad \Phi^{\dagger}(x) = (\Phi(x)^T\Phi(x))^{-1}\Phi(x)^T \tag{6.16}$$

By substituting (6.15) in (6.14), our objective function takes the following form:

$$\min_{x} \parallel S - \Phi(x)\Phi^{\dagger}(x)S \parallel_2^2 \tag{6.17}$$

(6.17) is called the variable projection functional. Assuming that $\Phi(x)$ has a locally constant rank, it has been proven [143] that the global minimum of (6.17) remains the same as the global minimum of (6.14).

Rank Constancy of $\Phi(x)$: The matrix $\Phi(x)$ has much larger number of rows (measurements) than the number of columns (compartments). Furthermore,

the measurements are always noisy. Thus generically, it is safe to assume that $\Phi(x)$ will always have full column rank.

- **Step 2.** ***Stochastic search for non-linear parameters 'x'***. The objective function given in (6.17) is non-convex, particularly of non-linear least-square form. Any gradient based method employed to estimate the parameters will have critical dependence on a good starting point, which is unknown. Alternative approach can be regular grid search, which is time consuming and adds computational burden. This particular type of problem therefore points towards considering stochastic search methods like GA. In case of time series analysis, it has been shown [144] that GA can be used efficiently for sum of exponentials functions. GA parameters can be varied for each selected biophysical model and time complexity may change with each choice. However, we found that these do not have substantial effect on convergence both in terms of accuracy and time complexity. For all the experiments and results reported in this study, we keep GA parameters in the following range (i) GA method: Elitism based, (ii) Population size: 24 to 48, (iii) Stopping criteria: fixed to 90 iterations (generations), (iv) Generation gap: fixed to 0.7.

- **Step 3.** ***Constrained search for linear parameters 'f'*** . After estimating parameters 'x', estimation of linear parameters 'f' is a constrained linear least squares estimation problem as shown below:

$$\min_{f} \parallel S - \Phi(x) f^T \parallel_2^2 \tag{6.18}$$

$$\text{such that} \quad \sum_{i=1}^{n} f_i = 1, \quad f_i \geq 0 \quad i = 1, 2 \ldots n$$

We used CVX (http://cvxr.com/cvx) for solving the optimization problem in (6.18).

- **Step 4.** ***Non Linear Least Squares Estimation using Trust Region Method***. Step 2 and step 3 give a reliable initial guess of both 'x' and 'f' to solve (6.13) by applying Trust Region method. Particularly, MATLAB's 'lsqcurvefit' was used to solve the constrained NLLS estimation problem in (6.13). This guarantees that stopping point of the algorithm is at least a stationary point. Also, it reduces the number of iterations (generations) and population

size used in the GA (step 2). For example, in time series analysis to have precise results, GA requires population size of 250 with 150 iterations [144]. With the type of functions describing white matter multi-compartment biophysical models, such a stochastic search will become computationally prohibitive.

6.5 Optimal Scanning Protocols

Our suggested dMRI data fitting method MIX enables finding parameters of the complex tissue models in multiple fiber orientations. However, for parameter estimation methods to work properly, dMRI data has to be information rich. In other words, for each choice of multicompartment model, we can find an optimized PGSE parameters. Below is the problem formulation:

Let $\beta = [G;\ \Delta;\ \delta]$ be a vector of parameters of the PGSE as described in Section 2.4.1. Define α as vector of all the remaining parameters of the model function (known/ fixed parameters of models in this case) e.g., $\alpha = [d_{\parallel};\ d_{\perp};\ g;\ n;\ f_i; R]$, as defined in Section 6.2.1. Please note that parameters in α will change with each choice of model function. The problem of finding optimum scanning parameters of PGSE for a model function $S(\alpha, \beta)$ can be written as under:

$$\max_{\beta} \frac{\partial\ S(\alpha, \beta)}{\partial \alpha} \tag{6.19}$$

For example, in case of ActiveAx (zeppelin-cylinder-dot), the optimization problem is as under:

$$\max_{G,\Delta,\delta} \begin{bmatrix} \frac{\partial S_{zcd}}{\partial R} \\ \frac{\partial S_{zcd}}{\partial \theta} \\ \frac{\partial S_{zcd}}{\partial \phi} \\ \frac{\partial S_{zcd}}{\partial f_i} \end{bmatrix}$$

6.6 Conclusion and Discussion

We have presented a technique (MIX) to fit complex biophysical models (e.g. ActiveAx, NODDI, NODDIx, ZCDx) to dMRI data of the brain, and estimate quantities such as axonal diameter and density. Unlike existing methods, MIX is versatile and thus suitable to a broad range of generic multi-compartment models, in particular for

brain areas where axonal pathways cross. Since the prevalence of such complex brain areas is estimated to 60 to 90 percent of the white matter, at the current imaging resolution (1 mm^3 approximately), MIX will enable the neuroimaging community to investigate the microstructure of the brain white matter in ways which are currently not possible.

We have shown that MIX is a superior alternative to methods that rely on grid search or other common optimization methods. MIX exploits the structure of the problem and is more robust than all other methods. Furthermore, it does not assume single fiber orientations and can directly estimate microstructure parameters in brain areas with complex geometry. Nonetheless, in brain areas with one primary fiber pathway, we have also shown that MIX outperforms other techniques. MIX is broadly applicable to many already available dMRI datasets, such as the Human Connectome Project (HCP). Using HCP data, we have demonstrated its ability to fit a generalized version of the NODDI model (called NODDIx) to estimate axonal orientation dispersions, volume fractions and orientations in up to two pathways. Finally, MIX has a simple and self-contained implementation in MATLAB which has not yet been optimized for computational efficiency, but has been designed for ease in dealing with realistic and generic biophysical models.

We note that recent works have focused on models describing the diffusion signal directly, which can subsequently be used to identify microstructural features such as those detailed in our work. These techniques include Q-ball imaging [147, 148], DSI [149] (Diffusion Spectrum MR Imaging), DKI [150] (Diffusion Kurtosis Imaging), SHORE [151] (Simple Harmonic Oscillator Based Reconstruction and Estimation), MAP [152, 153] (Mean Average Propagator) and MAPL [154] (Laplacian- regularized MAP). Although this chapter focuses on the direct estimation of microstructural features from biophysical models, we would like to emphasize that MIX could be applied to the diffusion signal models which rely on weighted sums of exponential functions.

We hope that our proposed method might provide a new, more flexible computational framework to revisit and expand studies on biophysical models evaluation and classification [134, 135], to improve current models [155, 156], and to develop future contests such as the White Matter Modeling Challenge [157] (`http://cmic.cs.ucl.ac.uk/wmmchallenge/`).

Chapter 7

Conclusion and Future Work

This thesis deals with geometrical and optimization methods to extract brain structural information from dMRI data. To start with, we focus on several theoretical challenges posed by DTI, where local water diffusion in brain tissue is modeled by diffusion tensors and, thereby, presenting brain white matter as a tensor field. We provid geometrical techniques for DTI-based analysis to assist in parcellation of brain areas based on their connectivity to rest of the brain. To this end, we introduce a notion of average connectivity of sub-cortical regions that depends on the geometric mean of connectivity profiles. By exploring a notion of Ricci flow on tensor fields we introduce a technique to enhance white matter structural features in a given diffusion tensor field. After developing suitable analysis tools for DTI, we focus on the analysis of brain structural networks. We introduce a notion of Ricci curvature to propose a novel node measure for identifying critical areas of such brain networks. Lastly, we address the problem to fit complex tissue models to dMRI data. To this end, we introduce an optimization method which is not only generic, applying to all existing multi-compartment tissue models, but it also allows modeling and fitting models in situations where there are multiple fiber orientations. Below we briefly summarize each of these contributions, highlight advantages and limitations, and suggest possible future extensions.

We introduce an optimal mass transport (OMT) distance as a similarity measure for brain connectivity profiles, considered as distributions of connecting fibres. We argue that OMT provides a better metric than the more commonly used correlation coefficient and show that improved resolution can be achieved using the OMT approach

in parcellating brain areas. From a computational standpoint OMT-metrics are substantially more demanding than evaluating correlations. However, when dealing with high resolution HCP data, the OMT approach leverages the fact that only non-zero values of connectivity profiles are relevant, to become computationally tractable. Further, geometric OMT-means of connectivity profiles in sub-cortical regions provides suitable notions of average connectivity of an area. A salient feature of our approach is to require that the geometric mean of connectivity profiles, a distribution in its own right, have support that coincides with the union of the support of the corresponding connectivity profiles, thereby ensuring anatomically meaningful results. Our approach can potentially be improved from a computational viewpoint by improving graph cut methods. Another topic for future work is the quantitative assessment of robustness and sensitivity of the results across different data sets and individuals.

We introduced a version of Ricci flow for processing of diffusion tensor fields. We explored a limited "backward" flow, analogous to backward heat flow, to enhance features and accentuate structures in white matter. The technique should prove helpful in increasing resolution and achieving more accurate tractography. However, the method is sensitive and depends critically on careful selection of step size and the number of steps. In our experience, in longitudinal brain studies for disease/development using DTI, the method can provide critical information on changes in corresponding tensor fields.

Brain structural connectivity networks describe connections between subcortical regions (nodes) based on their anatomical connectivity (number of connections). Graph theoretical measures can provide quantitative assessment of node centrality and help understand structural and functional brain characteristics. To this end, we have introduced a novel graph measure, a notion of Ricci node-curvature for brain networks, to identify their structural importance. We point out that curvature is a measure of node robustness and that it provides information that is not brought out by other node centrality measures. Experiments show that curvature can be used to detect changes due to healthy aging of the brain and also, in case of diseases like Autism Spectrum Disorder, to identify potentially important markers. It is of great interest to extend our findings to functional brain networks and to estimate curvature of functionally critical nodes, as for instance, nodes included in the Default Mode Network of resting state functional networks. Both, in the case of structural and functional brain networks, it should be worthwhile to study curvature changes at

edge level as well. Such an endeavor may point to new ways to detect critical changes in interactions between sub-cortical regions/nodes.

Lastly, we have introduced a method for fitting complex multicompartment tissue models to dMRI data. Our method is generic in that it applies to all existing models and allows tissue features estimation in cases of multiple fiber orientation. The current implementation of our algorithm, although not yet optimized for time efficiency, is twice as fast as CAMINO (a concurrent popular method that is applicable to a limited range of tissue models). In general, the tissue-model fitting depends on how rich the information content of the dMRI data is. Specifically, model-fitting is sensitive to shells number of shells and measurements available. In the absence of optimal scanning protocols, for particular choices of model, it may be challenging for a fitting algorithm to yield meaningful results.

We expect that the contributions made in this thesis will be benefit future dMRI technologies. More specifically, we expect that our methods for tensor field enhancement will lead to improved tractography techniques. Further, we expect that our approach in comparing and averaging connectivity profiles will point to anatomically meaningful parcellations and connectivity of sub-cortical regions, and will bring out vital brain structural information. Node curvature and Brain curvature (average of node curvature) can both serve as hallmarks for structural brain network connectivity. Optimized scanning protocols can unleash the ability of our proposed method (referred to as MIX) to provide reliable and detailed tissue microstructure information in complex tissue geometries.

Bibliography

[1] Le Bihan, D. and Breton, E., "Imagerie de diffusion in-vivo par résonance magnétique nucléaire," *Comptes-Rendus de l'Académie des Sciences*, Vol. 93, No. 5, Dec. 1985, pp. 27–34.

[2] Merboldt, K.-D., Hanicke, W., and Frahm, J., "Self-diffusion NMR imaging using stimulated echoes," *Journal of Magnetic Resonance (1969)*, Vol. 64, No. 3, 1985, pp. 479 – 486.

[3] Basser, P., Mattiello, J., and LeBihan, D., "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal*, Vol. 66, No. 1, 1994, pp. 259 – 267.

[4] Basser, P. J. and Pierpaoli, C., "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *Journal of Magnetic Resonance Series B*, Vol. 111, No. 3, 1996, pp. 209–219.

[5] Alexander, A. L., Hasan, K., Kindlmann, G., Parker, D. L., and Tsuruda, J. S., "A geometric analysis of diffusion tensor measurements of the human brain," *Magnetic Resonance in Medicine*, Vol. 44, No. 2, 2000, pp. 283–291.

[6] Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C., "NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain," *NeuroImage*, Vol. 61, No. 4, 2012, pp. 1000 – 1016.

[7] Farooq, H., Xu, J., Nam, J. W., Keefe, D. F., Yacoub, E., Georgiou, T., and Lenglet, C., "Microstructure Imaging of Crossing (MIX) White Matter Fibers from diffusion MRI," *Scientific reports*, Vol. 6, 2016.

[8] Daducci, A., Canales-Rodrguez, E. J., Zhang, H., Dyrby, T. B., Alexander, D. C., and Thiran, J.-P., "Accelerated Microstructure Imaging via Convex Op-

timization (AMICO) from diffusion MRI data," *NeuroImage*, Vol. 105, 2015, pp. 32 – 44.

[9] Lott, J. and Villani, C., "Ricci curvature for metric-measure spaces via optimal transport," *Annals of Mathematics*, Vol. 169, No. 3, 2009, pp. 903–991.

[10] Villani, C., *Topics in Optimal Transportation*, American Mathematical Society Publications, 2003.

[11] Farooq, H., Chen, Y., Georgiou, T. T., and Lenglet, C., "Some geometric ideas for feature enhancement of diffusion tensor fields," *55th IEEE Conference on Decision and Control, CDC 2016, Las Vegas,NV, USA, December 12-14, 2016*, 2016, pp. 3856–3861.

[12] Farooq, H., Xu, J., Yacoub, E., Georgiou, T., and Lenglet, C., *Brain Tissue Micro-Structure Imaging from Diffusion MRI Using Least Squares Variable Separation*, Springer International Publishing, Cham, 2016, pp. 55–64.

[13] Bloch, F., "Nuclear Induction," *Phys. Rev.*, Vol. 70, Oct 1946, pp. 460–474.

[14] Brown, R., "A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies," *Philosophical Magazine*, Vol. 4, No. 21, 1828, pp. 161–173.

[15] Einstein, A., "ber die von der molekularkinetischen Theorie der Wrme geforderte Bewegung von in ruhenden Flssigkeiten suspendierten Teilchen," *Annalen der Physik*, Vol. 322, No. 8, 1905, pp. 549–560.

[16] Einstein, A., "Investigation on the Theory of the Brownian Movement," *Ann. der Physik*, 1905.

[17] Hahn, E. L., "Spin Echoes," *Phys. Rev.*, Vol. 80, Nov 1950, pp. 580–594.

[18] Carr, H. Y. and Purcell, E. M., "Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments," *Phys. Rev.*, Vol. 94, May 1954, pp. 630–638.

[19] Torrey, H. C., "Bloch Equations with Diffusion Terms," *Phys. Rev.*, Vol. 104, Nov 1956, pp. 563–565.

[20] Stejskal, E. O. and Tanner, J. E., "Spin Diffusion Measurements: Spin Echoes in the Presence of a TimeDependent Field Gradient," *The Journal of Chemical Physics*, Vol. 42, No. 1, 1965, pp. 288–292.

[21] Hashemi, R., Bradley, W., and Lisanti, C., *MRI: The Basics*, LWW medical book collection, Lippincott Williams & Wilkins, 2004.

[22] Kärger, J. and Heink, W., "The propagator representation of molecular transport in microporous crystallites," *Journal of Magnetic Resonance (1969)*, Vol. 51, No. 1, 1983, pp. 1 – 7.

[23] Tuch, D. S., "Mapping cortical connectivity with diffusion MRI," *Proceedings IEEE International Symposium on Biomedical Imaging*, 2002, pp. 392–394.

[24] Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J. M., and Matthews, P. M., "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging," *Nature Neuroscience*, Vol. 6, No. 7, 2003, pp. 750–757.

[25] Tomassini, V., Jbabdi, S., Klein, J. C., Behrens, T. E. J., Pozzilli, C., Matthews, P. M., Rushworth, M. F. S., and Johansen-Berg, H., "Diffusion-weighted Imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations," *Journal of Neuroscience*, Vol. 27, No. 38, 2007, pp. 10259–10269.

[26] Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., and Knosche, T. R., "Connectivity-based parcellation of Broca's area," *Cerebral Cortex*, Vol. 17, No. 4, 2007, pp. 816–825.

[27] Menke, R. A., Jbabdi, S., Miller, K. L., Matthews, P. M., and Zarei, M., "Connectivity-based segmentation of the substantia nigra in human and its implications in Parkinson's disease," *Neuroimage*, Vol. 52, No. 4, 2010, pp. 1175–1180.

[28] Ros, C., Gullmar, D., Stenzel, M., Mentzel, H. J., and Reichenbach, J. R., "Atlas-Guided Cluster Analysis of Large Tractography Datasets," *Plos One*, Vol. 8, No. 12, 2013.

[29] O'Donnell, L. J., Golby, A. J., and Westin, C. F., "Fiber clustering versus the parcellation-based connectome," *Neuroimage*, Vol. 80, 2013, pp. 283–289.

[30] Guevara, P., Poupon, C., Riviere, D., Cointepas, Y., Descoteaux, M., Thirion, B., and Mangin, J. F., "Robust clustering of massive tractography datasets," *Neuroimage*, Vol. 54, No. 3, 2011, pp. 1975–1993.

[31] Wassermann, D., Bloy, L., Kanterakis, E., Verma, R., and Deriche, R., "Unsupervised white matter fiber clustering and tract probability map generation: Applications of a Gaussian process framework for white matter fibers," *Neuroimage*, Vol. 51, No. 1, 2010, pp. 228–241.

[32] O'Muircheartaigh, J., Vollmar, C., Traynor, C., Barker, G. J., Kumari, V., Symms, M. R., Thompson, P., Duncan, J. S., Koepp, M. J., and Richardson, M. P., "Clustering probabilistic tractograms using independent component analysis applied to the thalamus," *Neuroimage*, Vol. 54, No. 3, 2011, pp. 2020–2032.

[33] Jbabdi, S., Woolrich, M. W., and Behrens, T. E. J., "Multiple-subjects connectivity-based parcellation using hierarchical Dirichlet process mixture models," *Neuroimage*, Vol. 44, No. 2, 2009, pp. 373–384.

[34] Neher, P. F., Laun, F. B., Stieltjes, B., and Maier-Hein, K. H., "Fiberfox: Facilitating the Creation of Realistic White Matter Software Phantoms," *Magnetic Resonance in Medicine*, Vol. 72, No. 5, 2014, pp. 1460–1470.

[35] Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., and Mangin, J. F., "New Diffusion Phantoms Dedicated to the Study and Validation of High-Angular-Resolution Diffusion Imaging (HARDI) Models," *Magnetic Resonance in Medicine*, Vol. 60, No. 6, 2008, pp. 1276–1283.

[36] Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., Della Penna, S., Feinberg, D., Glasser, M. F., Harel, N., Heath, A. C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., Oostenveld, R., Petersen, S. E., Prior, F., Schlaggar, B. L., Smith, S. M., Snyder, A. Z., Xu, J., Yacoub, E., and Wu-Minn, H. C. P. C., "The Human Connectome Project: A data acquisition perspective," *Neuroimage*, Vol. 62, No. 4, 2012, pp. 2222–2231.

[37] Sotiropoulos, S. N., Jbabdi, S., Xu, J. Q., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., Van Essen, D. C., Ugurbil, K., Behrens, T. E. J., and Consortium, W. U.-M. H., "Advances in diffusion MRI acquisition and processing in the Human Connectome Project," *Neuroimage*, Vol. 80, 2013, pp. 125–143.

[38] Monge, G., *Mémoire sur la théorie des déblais et des remblais*, De l'Imprimerie Royale, 1781.

[39] Kantorovich, L. V., "On the transfer of masses," *Dokl. Akad. Nauk. SSSR*, Vol. 37, 1942, pp. 227–229.

[40] Villani, C., *Optimal Transport: Old and New*, Vol. 338, Springer, 2008.

[41] Rachev, S. and Rschendorf, L., *Mass Transportation Problems*, Vol. 1 of *Probability and Its Applications*, Springer-Verlag, New York, 1st ed., 1998.

[42] Rubner, Y., Tomasi, C., and Guibas, L. J., "The Earth Mover's Distance as a metric for image retrieval," *International Journal of Computer Vision*, Vol. 40, No. 2, 2000, pp. 99–121.

[43] Cuturi, M., "Sinkhorn distances: Lightspeed computation of optimal transport," *Advances in Neural Information Processing Systems*, 2013, pp. 2292–2300.

[44] Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., Smith, S. M., Higham, D. J., and Matthews, P. M., "Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 36, 2004, pp. 13335–13340.

[45] Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., and Woolrich, M. W., "Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?" *Neuroimage*, Vol. 34, No. 1, 2007, pp. 144–155.

[46] Shi, J. B. and Malik, J., "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, 2000, pp. 888–905.

[47] Ng, A. Y., Jordan, M. I., and Weiss, Y., "On spectral clustering: Analysis and an algorithm," *Advances in Neural Information Processing Systems 14, Vols 1 and 2*, Vol. 14, 2002, pp. 849–856.

[48] Witelson, S. F., "Hand and Sex-differences in the Isthmus and Genu of the Human Corpus-Callosum - A Postmortem Morphological-Study," *Brain*, Vol. 112, 1989, pp. 799–835.

[49] Castano-Moraga, C. A., Lenglet, C., Deriche, R., and Ruiz-Alzola, J., "A fast and rigorous anisotropic smoothing method for DT-MRI," *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.*, April 2006, pp. 93–96.

[50] Lenglet, C., Rousson, M., Deriche, R., and Faugeras, O., "Statistics on the Manifold of Multivariate Normal Distributions: Theory and Application to Diffusion Tensor MRI Processing," *Journal of Mathematical Imaging and Vision*, Vol. 25, No. 3, Oct 2006, pp. 423–444.

[51] Lenglet, C., Deriche, R., and Faugeras, O., *Inferring White Matter Geometry from Diffusion Tensor MRI: Application to Connectivity Mapping*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, pp. 127–140.

[52] Chow, I. B. and Knopf, I. D., "The Ricci Flow: An Introduction," 2013.

[53] Muniz, W., Ramos, F., and de Campos Velho, H., "Entropy- and tikhonov-based regularization techniques applied to the backwards heat equation," *Computers & Mathematics with Applications*, Vol. 40, No. 8, 2000, pp. 1071 – 1084.

[54] Lee, J. M., *Riemannian Manifolds*, Springer New York, 1993.

[55] DoCarmo, M., *Riemannian Geometry*, Birkhäuser, Boston, 1993.

[56] Jost, J., "Riemannian Geometry and Geometric Analysis, Fifth Edition," *Riemannian Geometry and Geometric Analysis, Fifth Edition*, 2008, pp. 1–583.

[57] Levi-Civita, M. d. T., "Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana," *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, Vol. 42, No. 1, Dec 1916, pp. 173–204.

[58] Hamilton, R. S., "Three-manifolds with positive Ricci curvature," *J. Differential Geom.*, Vol. 17, No. 2, 1982, pp. 255–306.

[59] Friedan, D., "Nonlinear Models in $2+\epsilon$ Dimensions," *Phys. Rev. Lett.*, Vol. 45, Sep 1980, pp. 1057–1060.

[60] Christoffel, E., "Ueber die Transformation der homogenen Differentialausdrcke zweiten Grades." *Journal fr die reine und angewandte Mathematik*, Vol. 70, 1869, pp. 46–70.

[61] Barmpoutis, A., Vemuri, B. C., Shepherd, T. M., and Forder, J. R., "Tensor Splines for Interpolation and Approximation of DT-MRI With Applications to Segmentation of Isolated Rat Hippocampi," *Ieee Transactions on Medical Imaging*, Vol. 26, No. 11, 2007, pp. 1537–1546.

[62] Föllmer, H., *Random fields and diffusion processes*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1988, pp. 101–203.

[63] Chen, Y. X., Georgiou, T. T., and Pavon, M., "On the Relation Between Optimal Transport and Schrodinger Bridges: A Stochastic Control Viewpoint," *Journal of Optimization Theory and Applications*, Vol. 169, No. 2, 2016, pp. 671–691.

[64] Chen, Y. X., Georgiou, T. T., and Pavon, M., "Optimal Steering of a Linear Stochastic System to a Final Probability Distribution, Part I," *Ieee Transactions on Automatic Control*, Vol. 61, No. 5, 2016, pp. 1158–1169.

[65] Chen, Y. X., Georgiou, T. T., and Pavon, M., "Optimal Steering of a Linear Stochastic System to a Final Probability Distribution, Part II," *Ieee Transactions on Automatic Control*, Vol. 61, No. 5, 2016, pp. 1170–1180.

[66] Perona, P. and Malik, J., "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 7, Jul 1990, pp. 629–639.

[67] Sporns, O., Tononi, G., and Ktter, R., "The Human Connectome: A Structural Description of the Human Brain," *PLOS Computational Biology*, Vol. 1, No. 4, 09 2005.

[68] Sporns, O., "The human connectome: a complex network," *Annals of the New York Academy of Sciences*, Vol. 1224, No. 1, 2011, pp. 109–125.

[69] Glasser, M. F., Smith, S. M., Marcus, D. S., Andersson, J. L. R., Auerbach, E. J., Behrens, T. E. J., Coalson, T. S., Harms, M. P., Jenkinson, M., Moeller, S., Robinson, E. C., Sotiropoulos, S. N., Xu, J., Yacoub, E., Ugurbil, K., and Van Essen, D. C., "The Human Connectome Project's neuroimaging approach," *Nature Neuroscience*, Vol. 19, No. 9, 2016, pp. 1175–1187.

[70] McNab, J. A., Edlow, B. L., Witzel, T., Huang, S. Y., Bhat, H., Heberlein, K., Feiweier, T., Liu, K., Keil, B., Cohen-Adad, J., Tisdall, M. D., Folkerth, R. D., Kinney, H. C., and Wald, L. L., "The Human Connectome Project and beyond: Initial applications of 300mT/m gradients," *NeuroImage*, Vol. 80, 2013, pp. 234 – 245, Mapping the Connectome.

[71] Fan, Q., Witzel, T., Nummenmaa, A., Dijk, K. R. V., Horn, J. D. V., Drews, M. K., Somerville, L. H., Sheridan, M. A., Santillana, R. M., Snyder, J., Hedden, T., Shaw, E. E., Hollinshead, M. O., Renvall, V., Zanzonico, R., Keil, B., Cauley, S., Polimeni, J. R., Tisdall, D., Buckner, R. L., Wedeen, V. J., Wald, L. L., Toga, A. W., and Rosen, B. R., "MGHUSC Human Connectome Project datasets with ultra-high b-value diffusion MRI," *NeuroImage*, Vol. 124, 2016, pp. 1108 – 1114, Sharing the wealth: Brain Imaging Repositories in 2015.

[72] Vu, A., Auerbach, E., Lenglet, C., Moeller, S., Sotiropoulos, S., Jbabdi, S., Andersson, J., Yacoub, E., and Ugurbil, K., "High resolution whole brain diffusion imaging at 7T for the Human Connectome Project," *NeuroImage*, Vol. 122, 2015, pp. 318 – 331.

[73] Setsompop, K., Kimmlingen, R., Eberlein, E., Witzel, T., Cohen-Adad, J., McNab, J., Keil, B., Tisdall, M., Hoecht, P., Dietz, P., Cauley, S., Tountcheva, V., Matschl, V., Lenz, V., Heberlein, K., Potthast, A., Thein, H., Horn, J. V., Toga, A., Schmitt, F., Lehne, D., Rosen, B., Wedeen, V., and Wald, L., "Pushing the limits of in vivo diffusion MRI for the Human Connectome Project," *NeuroImage*, Vol. 80, 2013, pp. 220 – 233, Mapping the Connectome.

[74] Sporns, O., *Networks of the Brain*, The MIT Press, 1st ed., 2010.

[75] Friston, K. J., "Functional and effective connectivity in neuroimaging: A synthesis," *Human Brain Mapping*, Vol. 2, No. 1-2, 1994, pp. 56–78.

[76] Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Ycel, M., Pantelis, C., and Bullmore, E. T., "Whole-brain anatomical networks: Does the choice of nodes matter?" *NeuroImage*, Vol. 50, No. 3, 2010, pp. 970 – 983.

[77] Yao, Z., Hu, B., Xie, Y., Moore, P., and Zheng, J., "A review of structural and functional brain networks: small world and atlas," *Brain Informatics*, Vol. 2, No. 1, 2015, pp. 45–52.

[78] Meier, J., Tewarie, P., Hillebrand, A., Douw, L., van Dijk, B. W., Stufflebeam, S. M., and Van Mieghem, P., "A Mapping Between Structural and Functional Brain Networks," *Brain Connectivity*, Vol. 6, No. 4, 2016, pp. 298–311.

[79] Bullmore, E. T. and Sporns, O., "The economy of brain network organization," *Nature Reviews Neuroscience*, Vol. 13, No. 5, 2012, pp. 336–349.

[80] Aerts, H., Fias, W., Caeyenberghs, K., and Marinazzo, D., "Brain networks under attack: robustness properties and the impact of lesions," *Brain*, Vol. 139, No. 12, 2016, pp. 3063–3083.

[81] Alstott, J., Breakspear, M., Hagmann, P., Cammoun, L., and Sporns, O., "Modeling the Impact of Lesions in the Human Brain," *PLOS Computational Biology*, Vol. 5, No. 6, 06 2009, pp. 1–12.

[82] Sporns, O., Honey, C. J., and Ktter, R., "Identification and Classification of Hubs in Brain Networks," *PLOS ONE*, Vol. 2, No. 10, 10 2007, pp. 1–14.

[83] Young, M. P., Hilgetag, C.-C., and Scannell, J. W., "On Imputing Function to Structure from the Behavioural Effects of Brain Lesions," *Philosophical Transactions: Biological Sciences*, Vol. 355, No. 1393, 2000, pp. 147–161.

[84] Warren, D. E., Power, J. D., Bruss, J., Denburg, N. L., Waldron, E. J., Sun, H., Petersen, S. E., and Tranel, D., "Network measures predict neuropsychological outcome after brain injury," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 111, No. 39, 2014, pp. 14247–14252.

[85] Gratton, C., Nomura, E. M., Prez, F., and D'Esposito, M., "Focal Brain Lesions to Critical Locations Cause Widespread Disruption of the Modular Organization of the Brain," *Journal of Cognitive Neuroscience*, Vol. 24, No. 6, 2012, pp. 1275–1285, PMID: 22401285.

[86] Sandhu, R. S., Georgiou, T. T., Reznik, E., Zhu, L., Kolesov, I., Senbabaoglu, Y., and Tannenbaum, A. R., "Graph Curvature for Differentiating Cancer Networks," *Scientific Reports*, Vol. 5, 2015, pp. 12323.

[87] Sandhu, R. S., Georgiou, T. T., and Tannenbaum, A. R., "Ricci curvature: An economic indicator for market fragility and systemic risk," *Science Advances*, Vol. 2, No. 5, 2016.

[88] Demetrius, L. A., "Boltzmann, Darwin and directionality theory," *Physics Reports-Review Section of Physics Letters*, Vol. 530, No. 1, 2013, pp. 1–85.

[89] Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O., "Mapping the Structural Core of Human Cerebral Cortex," *PLOS Biology*, Vol. 6, No. 7, 07 2008, pp. 1–15.

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

[91] Ollivier, Y., "A Visual Introduction to Riemannian Curvatures and Some Discrete Generalizations," *50th Seminar of Mathematics Superior*, Vol. 56 of *CRM Proceedings and Lecture Notes*, 2013, pp. 197–220.

[92] Ollivier, Y., "Ricci curvature of Markov chains on metric spaces," *Journal of Functional Analysis*, Vol. 256, No. 3, 2009, pp. 810–864.

[93] Ollivier, Y., "Ricci curvature of metric spaces," *Comptes Rendus Mathematique*, Vol. 345, No. 11, 2007, pp. 643–646.

[94] Varadhan, S. R. S., *Large Deviations and Applications*, SIAM, 1984.

[95] Otto, F., "The geometry of dissipative evolution equation: the porous medium equation," *Comm. Partial Differential Equations*, Vol. 26, 2001.

[96] Sturm, K. T., "On the geometry of metric measure spaces, I and II," *Acta Mathematica*, Vol. 196, No. 1, 2006, pp. 65–177.

[97] Evans, L. C., "Partial differential equations and Monge–Kantorovich mass transfer," *Current Developments in Mathematics*, 1999, pp. 65–126.

[98] Rubinov, M. and Sporns, O., "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, Vol. 52, No. 3, 2010, pp. 1059 – 1069, Computational Models of the Brain.

[99] Barrat, A., Barthlemy, M., Pastor-Satorras, R., and Vespignani, A., "The architecture of complex weighted networks," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. 11, 2004, pp. 3747–3752.

[100] Lenglet, C., Campbell, J., Descoteaux, M., Haro, G., Savadjiev, P., Wassermann, D., Anwander, A., Deriche, R., Pike, G., Sapiro, G., Siddiqi, K., and Thompson, P., "Mathematical methods for diffusion MRI processing," *NeuroImage*, Vol. 45, No. 1, 2009, pp. S111 – S122, Mathematics in Brain Imaging.

[101] Newman, M. E. J., "Scientific collaboration networks. I. Network construction and fundamental results," *Physical Review E*, Vol. 64, No. 1, 2001.

[102] Sporns, O., "Structure and function of complex brain networks," *Dialogues in Clinical Neuroscience*, Vol. 15, No. 3, 2013, pp. 247–262.

[103] Watts, D. J. and Strogatz, S. H., "Collective dynamics of 'small-world' networks," *Nature*, Vol. 393, No. 6684, 1998, pp. 440–442.

[104] Brown, J., Rudie, J., Bandrowski, A., Van Horn, J., and Bookheimer, S., "The UCLA multimodal connectivity database: a web-based platform for brain connectivity matrix sharing and analysis," *Frontiers in Neuroinformatics*, Vol. 6, 2012, pp. 28.

[105] Yeh, F.-C., Verstynen, T. D., Wang, Y., Fernndez-Miranda, J. C., and Tseng, W.-Y. I., "Deterministic Diffusion Fiber Tracking Improved by Quantitative Anisotropy," *PLOS ONE*, Vol. 8, No. 11, 11 2013.

[106] Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., and Aldroubi, A., "In vivo fiber tractography using DT-MRI data," *Magnetic Resonance in Medicine*, Vol. 44, No. 4, 2000, pp. 625–632.

[107] Essen, D. V. and Ugurbil, K., "The future of the human connectome," *NeuroImage*, Vol. 62, No. 2, 2012, pp. 1299 – 1310, 20 YEARS OF fMRI.

[108] Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L., "The Brain's Default Network," *Annals of the New York Academy of Sciences*, Vol. 1124, No. 1, 2008, pp. 1–38.

[109] Chen, Y., Georgiou, T., Pavon, M., and Tannenbaum, A., "A new approach to robust transportation over networks," *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 7623–7628.

[110] Fjell, A. M. and Walhovd, K. B., "Structural Brain Changes in Aging: Courses, Causes and Cognitive Consequences," *Reviews in the Neurosciences*, Vol. 21, No. 3, 2010, pp. 187–221.

[111] Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., and Acker, J. D., "Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers," *Cerebral Cortex*, Vol. 15, No. 11, 2005, pp. 1676.

[112] Ohnishi, T., Matsuda, H., Tabira, T., Asada, T., and Uno, M., "Changes in Brain Morphology in Alzheimer Disease and Normal Aging: Is Alzheimer Disease an Exaggerated Aging Process?" *American Journal of Neuroradiology*, Vol. 22, No. 9, 2001, pp. 1680–1685.

[113] Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., Morris, J. C., Dale, A. M., and Fischl, B., "Thinning of the Cerebral Cortex in Aging," *Cerebral Cortex*, Vol. 14, No. 7, 2004, pp. 721.

[114] Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., and Petersen, S. E., "Functional Network Organization of the Human Brain," *Neuron*, Vol. 72, No. 4, 2011, pp. 665 – 678.

[115] Herbert, M. R., Ziegler, D. A., Deutsch, C. K., O'Brien, L. M., Kennedy, D. N., Filipek, P. A., Bakardjiev, A. I., Hodgson, J., Takeoka, M., Makris, N., and Caviness, Jr, V. S., "Brain asymmetries in autism and developmental language disorder: a nested whole-brain analysis," *Brain*, Vol. 128, No. 1, 2005, pp. 213.

[116] Brieber, S., Neufang, S., Bruning, N., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., Fink, G. R., and Konrad, K., "Structural brain abnormalities in adolescents with autism spectrum disorder and patients with attention deficit/hyperactivity disorder," *Journal of Child Psychology and Psychiatry*, Vol. 48, No. 12, 2007, pp. 1251–1258.

[117] Caria, A. and de Falco, S., "Anterior insular cortex regulation in autism spectrum disorders," *Frontiers in Behavioral Neuroscience*, Vol. 9, 2015, pp. 38.

[118] Tan, G. C., Doke, T. F., Ashburner, J., Wood, N. W., and Frackowiak, R. S., "Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of CNTNAP2," *NeuroImage*, Vol. 53, No. 3, 2010, pp. 1030 – 1042.

[119] Bennett, R. H., Somandepalli, K., Roy, A. K., and Di Martino, A., "The Neural Correlates of Emotional Lability in Children with Autism Spectrum Disorder," *Brain Connectivity*, Vol. 7, No. 5, 2017, pp. 281–288.

[120] Tannenbaum, A. R., Sander, C., Sandhu, R. S., Reznik, E., Zhu, L., Kolesov, I., Senbabaoglu, Y., and Georgiou, T. T., "Graph Curvature and the Robustness of Cancer Networks," 2015.

[121] Farooq, H., Xu, J. Q., Nam, J. W., Keefe, D. F., Yacoub, E., Georgiou, T., and Lenglet, C., "Microstructure Imaging of Crossing (MIX) White Matter Fibers from diffusion MRI," *Scientific Reports*, Vol. 6, 2016.

[122] Farooq, H., Xu, J., Yacoub, E., Georgiou, T., and Lenglet, C., "Brain Tissue Micro-Structure Imaging from Diffusion MRI Using Least Squares Variable Separation," *Computational Diffusion Mri*, 2016, pp. 55–64.

[123] Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., and Chiro, G. D., "Diffusion tensor MR imaging of the human brain." *Radiology*, Vol. 201, No. 3, 1996, pp. 637–648, PMID: 8939209.

[124] Nilsson, M., Van Westen, D., Ståhlberg, F., Sundgren, P. C., and Lätt, J., "The role of tissue microstructure and water exchange in biophysical modelling of diffusion in white matter," *Magnetic Resonance Materials in Physics, Biology and Medicine*, Vol. 26, No. 4, Aug 2013, pp. 345–370.

[125] Stanisz, G. J., Szafer, A., Wright, G. A., and Henkelman, R. M., "An analytical model of restricted diffusion in bovine optic nerve," *Magnetic Resonance in Medicine*, Vol. 37, No. 1, 1997, pp. 103–111.

[126] Assaf, Y. and Basser, P. J., "Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain," *Neuroimage*, Vol. 27, No. 1, 2005, pp. 48–58.

[127] Assaf, Y., Freidlin, R. Z., Rohde, G. K., and Basser, P. J., "New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter," *Magnetic Resonance in Medicine*, Vol. 52, No. 5, 2004, pp. 965–978.

[128] Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., and Basser, P. J., "AxCaliber: A method for measuring axon diameter distribution from diffusion MRI," *Magnetic Resonance in Medicine*, Vol. 59, No. 6, 2008, pp. 1347–1354.

[129] Barazany, D., Basser, P. J., and Assaf, Y., "In vivo measurement of axon diameter distribution in the corpus callosum of rat brain," *Brain*, Vol. 132, 2009, pp. 1210–1220.

[130] Alexander, D. C., Hubbard, P. L., Hall, M. G., Moore, E. A., Ptito, M., Parker, G. J., and Dyrby, T. B., "Orientationally invariant indices of axon diameter and density from diffusion MRI," *NeuroImage*, Vol. 52, No. 4, 2010, pp. 1374 – 1389.

[131] Dyrby, T. B., Sogaard, L. V., Hall, M. G., Ptito, M., and Alexander, D. C., "Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI," *Magnetic Resonance in Medicine*, Vol. 70, No. 3, 2013, pp. 711–721.

[132] Scherrer, B., Schwartzman, A., Taquet, M., Sahin, M., Prabhu, S. P., and Warfield, S. K., "Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND)," *Magnetic Resonance in Medicine*, Vol. 76, No. 3, 2016, pp. 963–977.

[133] Assaf, Y., Alexander, D. C., Jones, D. K., Bizzi, A., Behrens, T. E. J., Clark, C. A., Cohen, Y., Dyrby, T. B., Huppi, P. S., Knoesche, T. R., LeBihan, D., Parker, G. J. M., Poupon, C., and Consortium, C., "The CONNECT project: Combining macro- and micro-structure," *Neuroimage*, Vol. 80, 2013, pp. 273–282.

[134] Panagiotaki, E., Schneider, T., Siow, B., Hall, M. G., Lythgoe, M. F., and Alexander, D. C., "Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison," *NeuroImage*, Vol. 59, No. 3, 2012, pp. 2241 – 2254.

[135] Ferizi, U., Schneider, T., Panagiotaki, E., Nedjati-Gilani, G., Zhang, H., Wheeler-Kingshott, C. A. M., and Alexander, D. C., "A ranking of diffusion MRI compartment models with in vivo human brain data," *Magnetic Resonance in Medicine*, Vol. 72, No. 6, 2014, pp. 1785–1792.

[136] Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K., and Sijbers, J., "Investigating the Prevalence of Complex Fiber Configurations in White Matter Tissue with Diffusion Magnetic Resonance Imaging," *Human Brain Mapping*, Vol. 34, No. 11, 2013, pp. 2747–2766.

[137] Cook, P. A., Bai, Y., Gilani, S. N., Seunarine, K. K., Hall, M. G., Parker, G. J., and Alexander, D. C., "Camino: Open-Source Diffusion-MRI Reconstruction and Processing," *14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, May 2006, p. 2759.

[138] Yeh, F.-C., Liu, L., Hitchens, T. K., and Wu, Y. L., "Mapping immune cell infiltration using restricted diffusion MRI," *Magnetic Resonance in Medicine*, Vol. 77, No. 2, 2017, pp. 603–612.

[139] Auria, A., Romanasco, D., Canales-Rodriguez, E., Wiaux, Y., Dirby, T., Alexander, D., Thiran, J.-P., and Daducci, A., "Accelerated Microstructure Imaging via Convex Optimization for regions with multiple fibres (AMICOx)," 2015.

[140] Kaden, E., Kruggel, F., and Alexander, D. C., "Quantitative mapping of the per-axon diffusion coefficients in brain white matter," *Magnetic Resonance in Medicine*, Vol. 75, No. 4, 2016, pp. 1752–1763.

[141] Reddy, C. P. and Rathi, Y., "Joint Multi-Fiber NODDI Parameter Estimation and Tractography Using the Unscented Information Filter," *Frontiers in Neuroscience*, Vol. 10, 2016, pp. 166.

[142] Zhang, H., Dyrby, T. B., and Alexander, D. C., *Axon Diameter Mapping in Crossing Fibers with Diffusion MRI", bookTitle="Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011: 14th International Conference, Toronto, Canada, September 18-22, 2011, Proceedings, Part II*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 82–89.

[143] Golub, G. H. and Pereyra, V., "The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate," *SIAM Journal on Numerical Analysis*, Vol. 10, No. 2, 1973, pp. 413–432.

[144] Mitra, S. and Mitra, A., "A Genetic Algorithms Based Technique for Computing the Nonlinear Least Squares Estimates of the Parameters of Sum of Exponentials Model," *Expert Syst. Appl.*, Vol. 39, No. 7, June 2012, pp. 6370–6379.

[145] Dyrby, T. B., Baar, W. F., Alexander, D. C., Jelsing, J., Garde, E., and Sgaard, L. V., "An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets," *Human Brain Mapping*, Vol. 32, No. 4, 2011, pp. 544–563.

[146] Innocenti, G. M., Carln, M., and Dyrby, T. B., "Chapter 15 - The Diameters of Cortical Axons and Their Relevance to Neural Computing," *Axons and Brain Architecture*, edited by K. S. Rockland, Academic Press, San Diego, 2016, pp. 317 – 335.

[147] Tuch, D. S., "Q-Ball imaging," *Magnetic Resonance in Medicine*, Vol. 52, No. 6, 2004, pp. 1358–1372.

[148] Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R., "Regularized, fast, and robust analytical Q-Ball imaging," *Magnetic Resonance in Medicine*, Vol. 58, No. 3, 2007, pp. 497–510.

[149] Wedeen, V. J., Wang, R. P., Schmahmann, J. D., Benner, T., Tseng, W. Y. I., Dai, G., Pandya, D. N., Hagmann, P., D'Arceuil, H., and de Crespignya, A. J., "Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers," *Neuroimage*, Vol. 41, No. 4, 2008, pp. 1267–1277.

[150] Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. Z., and Kaczynski, K., "Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging," *Magnetic Resonance in Medicine*, Vol. 53, No. 6, 2005, pp. 1432–1440.

[151] Özarslan, E., Koay, C. G., and Basser, P. J., *Simple Harmonic Oscillator Based Reconstruction and Estimation for One-Dimensional q-Space Magnetic Resonance (1D-SHORE)*, Birkhäuser Boston, Boston, 2013, pp. 373–399.

[152] Özarslan, E., Koay, C. G., Shepherd, T. M., Komlosh, M. E., rfanolu, M. O., Pierpaoli, C., and Basser, P. J., "Mean apparent propagator (MAP) MRI: A novel diffusion imaging method for mapping tissue microstructure," *NeuroImage*, Vol. 78, 2013, pp. 16 – 32.

[153] Avram, A. V., Sarlls, J. E., Barnett, A. S., Özarslan, E., Thomas, C., Irfanoglu, M. O., Hutchinson, E., Pierpaoli, C., and Basser, P. J., "Clinical feasibility of using mean apparent propagator (MAP) MRI to characterize brain tissue microstructure," *NeuroImage*, Vol. 127, 2016, pp. 422 – 434.

[154] Fick, R. H., Wassermann, D., Caruyer, E., and Deriche, R., "MAPL: Tissue microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP data," *NeuroImage*, Vol. 134, 2016, pp. 365 – 385.

[155] Jelescu, I. O., Veraart, J., Adisetiyo, V., Milla, S. S., Novikov, D. S., and Fieremans, E., "One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI?" *Neuroimage*, Vol. 107, 2015, pp. 242–256.

[156] Jelescu, I. O., Veraart, J., Fieremans, E., and Novikov, D. S., "Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue," *Nmr in Biomedicine*, Vol. 29, No. 1, 2016, pp. 33–47.

[157] Ferizi, U., Scherrer, B., Schneider, T., Alipoor, M., Eufracio, O., Fick, R. H., Deriche, R., Nilsson, M., Loya-Olivas, A. K., Rivera, M., Poot, D. H., Ramirez-Manzanares, A., Marroquin, J. L., Rokem, A., Ptter, C., Dougherty, R. F., Sakaie, K., Wheeler-Kingshott, C., Warfield, S. K., Witzel, T., Wald, L. L., Raya, J. G., and Alexander, D. C., "Diffusion MRI microstructure models with in vivo human brain Connectome data: results from a multi-group comparison," *NMR in Biomedicine*, 2017, pp. e3734–n/a, e3734 nbm.3734.