

**Diffusion MRI for Well-posed and Optimal  
White Matter Microstructure Characterisation:  
Beyond Single Diffusion Encoding**



Santiago Coelho  
The University of Leeds  
School of Computing

Submitted in accordance with the requirements for the degree of  
*Doctor of Philosophy*  
September 2019



## Intellectual Property Statement

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 3 of the thesis has appeared in journal publications as follows:

*Resolving degeneracy in diffusion MRI biophysical model parameter estimation using double diffusion encoding. S Coelho, JM Pozo, SN Jespersen, DK Jones, AF Frangi - Magnetic Resonance in Medicine, 2019. 82:395-410.*

I am the main author of the publication. I led the design of the study, and wrote the manuscript, including the analysis and discussion of the results. I also wrote the analysis code and carried out the data processing, statistical analysis, and experiments.

The contribution of the other authors was in helping me designing the study, discussion of the results and writing of the manuscript.

The work in Chapter 4 of the thesis has appeared in journal and conferences publications as follows:

*Optimal experimental design for biophysical modelling in multidimensional diffusion MRI. S Coelho, JM Pozo, SN Jespersen, AF Frangi - International conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2019.*

I am the main author of the publication. I led the design of the study, and wrote the manuscript, including the analysis and discussion of the results. I also wrote the analysis code and carried out the data processing, statistical

analysis, and experiments.

The contribution of the other authors was in helping me designing the study, discussion of the results and writing of the manuscript.

*Optimal experimental design in multidimensional diffusion MRI for parameter estimation of biophysical tissue models. S Coelho, JM Pozo, SN Jespersen, E Fieremans, DS Novikov, AF Frangi - Manuscript, submitted to NeuroImage, 2019*

I am the main author of the publication. I led the design of the study, and wrote the manuscript, including the analysis and discussion of the results. I also wrote the analysis code and carried out the data processing, statistical analysis, and experiments.

The contribution of the other authors was in helping me designing the study, discussion of the results and writing of the manuscript. E Fieremans led the data acquisition.

The work in Chapter 5 of the thesis has appeared in journal and conferences publications as follows:

*Local volume fraction distributions of axons, astrocytes, and myelin in deep subcortical white matter. S Coelho, JM Pozo, M Costantini, JR Highley, M Mozumder, JE Simpson, PG Ince, AF Frangi - NeuroImage, 2018. 179: 275-287.*

I am the main author of the publication. I led the design of the study, and wrote the manuscript, including the analysis and discussion of the results. I also wrote the analysis code and carried out the data processing, statistical analysis, and experiments.

The contribution of the other authors was in helping me designing the study, discussion of the results and writing of the manuscript. JR Highley led the histology data acquisition.

*Histological data of axons, astrocytes, and myelin in deep subcortical white matter populations. S Coelho, JM Pozo, M Costantini, JR Highley, M Mozumder, JE Simpson, PG Ince, AF Frangi - Data in brief, 2019. 23:103762.*

I am the main author of the publication. I led the design of the study, and wrote the manuscript, including the analysis and discussion of the results. I also wrote the analysis code and carried out the data processing, statistical

analysis, and experiments.

The contribution of the other authors was in helping me designing the study, discussion of the results and writing of the manuscript. JR Highley led the histology data acquisition.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Santiago Coelho to be identified as Author of this work has been asserted by Santiago Coelho in accordance with the Copyright, Designs and Patents Act 1988.

©2019 The University of Leeds and Santiago Coelho.



## Acknowledgements

There are many people that deserve my gratitude for their support during my PhD. Some of them gave valuable technical advice while others provided encouraging words when these were needed. Without them, I would not have enjoyed this journey and probably this thesis would not have been possible.

Firstly, I would like to thank my advisor Professor Alejandro Frangi, for his guidance, his motivation, and for many lessons about the academic world. Alex was always up for new ideas and gave me the freedom to choose my own way to explore this exciting research area. I also want to thank my second advisor, Dr Jose Pozo, for the support, the numerous and rich discussions far beyond my research area, and for sharing with me many new tools that I have enjoyed learning. Jose always found time to investigate with me where I was stuck and stimulated me to ask the right questions. I consider myself lucky to have had such a pair of complementary mentors during my PhD.

Besides my advisors, I would also like to thank Professors Sune Jespersen, Derek Jones, Dmitry Novikov and Els Fieremans, you were all excellent scientific collaborators and great people to work with. To Els and Dmitry (Sune as well), thank you for the opportunity to visit your team at New York University for a few months, I had a wonderful time there. To the rest of my co-authors, thank you for sharing your knowledge and skills with me. I hope to collaborate with all of you in the future.

I want to thank my fellow labmates at CISTIB, for the stimulating discussions over coffee and jokes over beer. Thanks to Marina, for countless debates with mate at the kitchen's whiteboard and for always challenging my knowledge extrapolation skills. Also, special thanks to my 'foodie' friends and my football friends in Sheffield, it would have been way less fun without you guys.

Last but not least, I would like to thank those who have always been there for me. To my family and friends in Argentina, who made sure that every time I came back it felt as if I had never left, I hope that feeling never ceases. To my uncle Milton, for his wise advice every time I was struggling to see things clearly and to my aunt Pato for her visits. To my brother, for our daily random political/scientific/meme discussions. Finally, my deepest gratitude goes to my parents. For so many visits that made me forget I was living far away. Thank you for believing in me, sometimes more than I did, and for your endless support.



## Abstract

The human brain hosts a colossal number of water molecules which are constantly moving due to Brownian motion. Their movement, random by nature, is restricted by the brain tissue walls. Magnetic Resonance Imaging (MRI) provides macroscopic measurements of the diffusion process in a non-invasive manner, *i.e.* diffusion MRI. Hidden in these measurements lies information about the underlying architecture. The ability to unravel tissue microstructure from the coarse-grained diffusion measurements is extremely valuable since this information is 2-3 orders of magnitude below typical MRI resolution. This makes diffusion MRI sensitive to pathological and developmental processes occurring at the mesoscopic scale, in the order of microns. Accessing this level of detail can lead to clinical biomarkers specific to early stages of neurodegenerative diseases or brain development.

Computational models of biophysical tissue properties have been widely used in diffusion MRI research to elucidate the link between microstructural properties and MR signal formation. The potential increase in sensitivity and specificity in detecting brain microstructural changes is their major driving force. However, these models establish complex relationships between biophysical properties and the MR signal, making the inverse problem of recovering model parameters from noisy measurements ill-conditioned with conventional diffusion MRI acquisitions.

This thesis explores ways to make diffusion MRI biophysical modelling more robust while maintaining time and hardware requirements that are feasible in clinical conditions. Firstly, we explore theoretically the benefits of incorporating functionally independent measurements, such as double diffusion encoding. Secondly, we propose an optimal experiment design framework that gives us, after exploring the whole multidimensional diffusion MRI measurement space, the acquisition that maximises accuracy and precision in the parameter estimation. Finally, we extract relevant information from histology images that can be used to feed or benchmark diffusion MRI models.



# CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.1.1	Why diffusion MRI? . . . . .	2
1.1.2	Future clinical applications for diffusion MRI . . . . .	2
1.1.3	MRI role in health systems . . . . .	3
1.2	Thesis aims . . . . .	4
<b>2</b>	<b>Diffusion MRI</b>	<b>5</b>
2.1	Magnetic Resonance Imaging (MRI) . . . . .	5
2.2	Measuring diffusion with MRI . . . . .	7
2.2.1	Spin Echo . . . . .	7
2.2.2	The diffusion phenomenon . . . . .	8
2.2.3	Single Diffusion Encoding (SDE) . . . . .	8
2.3	Analysing Diffusion Weighted Images . . . . .	10
2.3.1	Signal representations . . . . .	11
2.3.2	Biophysical tissue models: The Standard Model . . . . .	14
2.4	Beyond Single Diffusion Encoding . . . . .	17
2.4.1	Double Diffusion Encoding (DDE) . . . . .	17
2.4.2	Multidimensional dMRI . . . . .	20
<b>3</b>	<b>Resolving the Watson Standard Model degeneracy</b>	<b>23</b>
3.1	Degeneracy in SM parameter estimation . . . . .	24
3.2	SDE cumulant expansion and solutions to NODDIDA . . . . .	25
3.3	Solving the degeneracy with orthogonal information . . . . .	28
3.3.1	Model extension to DDE . . . . .	28
3.3.2	DDE information gain and unique solution . . . . .	29
3.3.3	<i>In silico</i> experiments . . . . .	32

## CONTENTS

---

3.4	Results . . . . .	35
3.5	Discussion . . . . .	37
<b>4</b>	<b>Optimal experiment design in Multidimensional dMRI</b>	<b>43</b>
4.1	Optimal experiment design literature in dMRI . . . . .	44
4.2	Rationale of our approach . . . . .	45
4.2.1	Problem statement . . . . .	45
4.2.2	Cramér-Rao Bounds . . . . .	46
4.2.3	Metric definition . . . . .	46
4.2.4	Optimisation strategy and constraints . . . . .	48
4.3	Experiments . . . . .	50
4.3.1	MRI data acquisition . . . . .	51
4.3.2	Data processing . . . . .	51
4.4	Results . . . . .	52
4.4.1	<i>In silico</i> experiments . . . . .	52
4.4.2	Human brain experiments . . . . .	54
4.5	Discussion . . . . .	54
<b>5</b>	<b>Tissue statistics from histological images</b>	<b>61</b>
5.1	Histology to inform diffusion MRI . . . . .	61
5.1.1	Methodological overview . . . . .	64
5.2	Image acquisition . . . . .	64
5.2.1	Tissue sample selection . . . . .	64
5.2.2	Histology acquisition . . . . .	65
5.3	Image analysis framework . . . . .	66
5.3.1	Histology segmentation . . . . .	66
5.3.2	Scale dependency of the area fraction . . . . .	67
5.3.3	Volume fraction from histology images . . . . .	70
5.4	Results . . . . .	73
5.4.1	Validation of automated segmentation . . . . .	73
5.4.2	Optimal scale selection . . . . .	74
5.4.3	Volume fraction distributions . . . . .	75
5.5	Discussion . . . . .	76
<b>6</b>	<b>Conclusions</b>	<b>81</b>
6.1	Achievements . . . . .	81
6.2	Future work . . . . .	83

---

<b>Co-authored published work not included in the thesis</b>	<b>85</b>
<b>A Appendix</b>	<b>87</b>
A.1 Invariant decomposition of an axially symmetric 4th order tensor . . . . .	87
A.2 Single solution for NODDIDA with DDE . . . . .	89
A.3 Tensorial formulation of the SM for a general ODF . . . . .	90
A.4 Degeneracy in the SM with LTE and LTE+STE . . . . .	91
A.5 Correction factor for cylinders' size dispersion . . . . .	93
A.6 Match between projected and true size statistics for spheres . . . . .	94
A.7 <i>In silico</i> validation of volume fraction computation . . . . .	95
A.8 Experiments on <i>corpus callosum</i> data . . . . .	98
<b>References</b>	<b>101</b>



# LIST OF FIGURES

2.1	K-space and image of MRI T1-weighted scan in a healthy brain . . . . .	7
2.2	Spin Echo sequence developed by Hahn . . . . .	8
2.3	PGSE sequence developed by Stejskal and Tanner . . . . .	9
2.4	FA map and whole brain tractography derived from DTI . . . . .	12
2.5	Representation of the two-compartment Standard Model . . . . .	15
2.6	Double Diffusion Encoding sequence based on PGSE . . . . .	17
2.7	Different microscopic environments and their global diffusion tensors . . . . .	19
2.8	B-tensor shapes available from multidimensional dMRI . . . . .	21
3.1	Diagram with SDE and DDE directions used in both experiments . . . . .	33
3.2	Histograms of the estimated NODDIDA parameters for SDE and DDE . . . . .	36
3.3	1-dimensional cuts of 5-dimensional NODDIDA objective function . . . . .	37
3.4	Violin plots of RMSE for different SDE and DDE combinations . . . . .	37
3.5	3-dimensional plots showing the projected RMSE for $f$ and $D_a$ . . . . .	38
4.1	Multi-modal likelihoods, Fisher information and Cramér-Rao Bounds . . . . .	47
4.2	Representation of the optimisation constraints . . . . .	50
4.3	Optimisation results for all constraints . . . . .	53
4.4	Standard Model parametric maps for a healthy brain from four B-sets . . . . .	55
4.5	Standard Model error maps for analysed B-sets . . . . .	56
5.1	Diagram showing the cuts of the tissue blocks and their coronal slices . . . . .	66
5.2	Histology: RGB, grayscale, and segmented images . . . . .	68
5.3	Window patch arrangement for the scale dependence analysis . . . . .	69
5.4	Synthetic 3D volume of straight cylinders and its 2D projection . . . . .	71
5.5	Microscopy image of an astrocyte with our volume model . . . . .	71
5.6	Segmentation validation - agreement scores . . . . .	73

## LIST OF FIGURES

---

5.7	Volume fraction scale dependency analysis . . . . .	74
5.8	Volume fraction estimated distributions . . . . .	75
5.9	Group differences in mean volume fraction . . . . .	77
A.1	<i>In silico</i> blocks of cylinders and spheres with their 2D projections . . . . .	95
A.2	Relative error in $V_V$ estimates for randomly placed spheres . . . . .	96
A.3	Relative error in $V_V$ estimates for randomly placed cylinders . . . . .	96
A.4	Scatter plot of $A_{A,sph}$ vs $V_{V,sph}$ in a simulated block of spheres . . . . .	97
A.5	Scatter plot of $A_{A,sph}$ vs $V_{V,sph}$ in a simulated block of cylinders . . . . .	97
A.6	Histological image of axons in the <i>corpus callosum</i> . . . . .	98
A.7	Scatter plot of $A_{A,cyl}$ vs $V_{V,cyl}$ for randomly placed parallel cylinders . . . . .	99

# LIST OF TABLES

3.1	NODDIDA parameter sets with equal diffusion–kurtosis parameters . . . . .	28
3.2	Ground truth NODDIDA parameters used in experiment 1. . . . .	34
3.3	RMSE over the full grid for all acquisition protocols . . . . .	39
4.1	RMSE for the <i>in silico</i> experiment comparing optimal B-set vs standard combinations of LTE, PTE and STE. . . . .	53
4.2	RMSE for the <i>in vivo</i> experiment comparing optimal B-set vs standard combinations of LTE, PTE and STE . . . . .	54
5.1	Patient cohort details for histology analysis . . . . .	65
5.2	Volume fraction distributions: mean and variance . . . . .	76

## Abbreviations

$A_A$	area fraction
ADC	apparent diffusion coefficient
CFAS	cognitive function and ageing study
CRB	Cramér-Rao Bound
DDE	double diffusion encoding
DKI	diffusion kurtosis imaging
dMRI	diffusion magnetic resonance imaging
DSCL	deep subcortical lesion
DTI	diffusion tensor imaging
DWI	diffusion weighted image
GM	grey matter
LTE	linear tensor encoding
MGCs	multiple Gaussian compartments
MRI	magnetic resonance imaging
NAWM	normal appearing white matter
NODDI(DA)	neurite orientation dispersion and density imaging (with diffusivity assessment)
ODF	orientation distribution function
PGSE	pulsed gradient spin echo
PLIC	posterior limb of the internal capsule
PTE	planar tensor encoding
(R)MSE	(root) mean squared error
SDE	single diffusion encoding
SM	standard model
SNR	signal-to-noise ratio
STE	spherical tensor encoding
TE	echo time
$V_V$	volume fraction
WM	white matter

---

# CHAPTER 1

---

## Introduction

### 1.1 Motivation

Neurons are the units through which the nervous system is built. About 100 billion of them conform the human brain [1]. Normally neurons do not reproduce or replace themselves, making difficult to undo the damage that our nervous system has suffered. Neurodegenerative diseases (*e.g.* amyotrophic lateral sclerosis, Parkinson's disease, Alzheimer's disease) affect neurons, thus, the life quality of millions of people [2]. These conditions produce progressive degeneration or death of nerve cells, leading to alterations in the brain microstructure. Moreover, when they evolve, the fidelity of the information transfer underlying normal and cognitive brain function is affected [3]. Although many treatments have been and are currently being developed to counter these effects at different stages, early diagnose has shown to be a key factor to pursue in everyday clinic [2]. To achieve it, we must develop devices that enable early detection of structural tissue changes. Recent years have seen a shift in the MRI research community towards developing tools aimed at quantifying tissue microstructure. An MRI modality dubbed diffusion MRI has gained attention due to its high sensitivity to microstructural changes. A promising approach to analyse these images involves modelling the biophysics relating the main tissue components and the diffusion signal formation. However, these models are complex, hindering the inverse problem of recovering model parameters from the images. This thesis focuses on advancing biophysical models of diffusion MRI proposing new ways to acquire.

## 1. INTRODUCTION

---

### 1.1.1 Why diffusion MRI?

Since its invention in the 1970s [4, 5], Magnetic Resonance Imaging (MRI) has been established as an invaluable tool for diagnosing brain diseases *in vivo* and non-invasively. This technique produces images by measuring a signal that arises from the water molecules inside the tissue. This signal, related to the evolution of the magnetisation density of the hydrogen nucleus spins, generates the contrast between tissues with different characteristics in an MR image. For example, it allows us to discriminate different regions of the brain like white and grey matter (WM, GM), and to detect abnormal areas such as tumours or lesions. There is a general interest in the MRI community in increasing the level of detail accessible by an MRI scan, since detecting small tissue structural changes is linked to early diagnosis. This can be seen in the number of papers mentioning both ‘MRI’ and ‘microstructure’ which has doubled almost every three years in the last 25 years [6].

Current clinical scanners can generate images with voxel sizes in the order of the millimetre. Thus, any qualitative diagnosis is practically not sensitive to alterations below this scale. Increasing the resolution, *i.e.* reducing the voxel size, in an MRI scanner has some inherent hardware limitations due to the high magnetic field needed to achieve sufficient signal-to-noise ratio (SNR) in a reduced volume. This, together with the confirmation that the diffusion of water molecules in brain WM is anisotropic [7], encouraged the development of a modality within MRI termed ‘Diffusion MRI’ (dMRI). The peculiarity of this technique is that it sensitises the signal measured from a voxel to the random displacement of water molecules within it. These displacements are in the order of the tens of  $\mu\text{m}$  for clinical experimental settings. Hence, these signals hold encoded information about the tissue architecture restricting the diffusion of water molecules at the mesoscopic scale where many disease processes occur. Thus, dMRI is able to probe tissue on a scale that is much smaller than the image resolution. Moreover, it has the potential to extract more specific information than other MRI modalities. In the brain, the main clinical applications that current dMRI methods have encountered are acute ischemic stroke, characterised by a decrease in the Apparent Diffusion Coefficient (ADC) [8] values [9], something poorly assessed with structural MRI. Another technique with significant clinical translation is tractography [10], which has been shown valuable for pre-surgical planning [11].

### 1.1.2 Future clinical applications for diffusion MRI

The diffusion times that we can probe with MRI range from a few to tens of  $ms$ . This, added to the diffusivity of free water ( $3\mu\text{m}^2/ms$  at body temperature), make experimental diffusion lengths lie between  $1 - 50\mu\text{m}$  [6], *i.e.* at the mesoscopic scale. It is at this

particular scale, orders of magnitude above molecular level and orders of magnitude below MR resolution, not only where cells arrange themselves forming different tissues but also where pathology produces structural alterations. This makes dMRI sensitive to tissue changes at scales that are undetectable by other *in vivo* imaging modalities.

These factors are the main reason behind the intense growth of this research area. Although signal representations, *e.g.* Diffusion Tensor Imaging, have proven to be sensitive to micro-anatomical changes in tissue, they lack specificity [12]. In the quest for specificity to developmental and pathological changes, many biophysical models have been proposed, with the Standard Model (SM) being the most prominent for brain white matter [13, 14]. This model captures the most relevant microstructural tissue properties present in the dMRI signal, to develop disease-specific biomarkers. Some disease processes that may be accessible with biophysical modelling are demyelination [15], beading [16], and oedema [17]. However, robust estimation of SM parameters is a challenging task that may require advanced imaging hardware as well as long acquisitions, something difficult to have even in pre-clinical settings.

### 1.1.3 MRI role in health systems

After its early developments in the 1970s [4, 5], MRI has rapidly become a fundamental part of any health system. For example, around 3.46 million MRI scans were performed by the National Health Service (NHS) in England during 2017/2018 [18]. This number has shown a steady rise of roughly 8% annually over the last six years. This is due to an increase in the availability of facilities that perform these tests, and their potential for early diagnosis of conditions such as brain diseases or cancer. Furthermore, Alzheimer's disease and other dementias are projected to show a 66% increase from 2015 to 2030 [19]. This growth means that research focusing on developing new MRI techniques or enhancing current ones has the potential to cause a high impact on society. Optimising the extraction of information from an MRI scan can lead to a reduction in the time needed for the acquisition or reduce hardware requirements and thus, decrease its cost. Furthermore, increasing the specificity and accuracy of these techniques can lead to earlier diagnosis of many diseases. These two factors can mean huge savings for the health system by avoiding costly treatments, and also help to improve the life quality of people with a neurological condition.

## 1. INTRODUCTION

---

### 1.2 Thesis aims

The aim of this thesis is to improve biophysical modelling by extending it outside conventional dMRI acquisition sequences to foster their clinical translation. First, I will explore non-standard diffusion encodings such as double diffusion encoding. The objective is to analyse the complementary information that these encodings provide and see if it is sufficient to solve the degeneracy in parameter estimation of the Standard Model. Secondly, I will develop the criteria to optimise the data acquisition process in this new extended acquisition space. Although there are many papers discussing optimal experiment designs for single diffusion encoding, currently there is no criteria for efficient data acquisition when we consider generalised encodings. Finally, I will extract relevant tissue properties from histological images that have the potential to be incorporated into the diffusion modelling as prior information. These goals and ongoing model validations add to the current efforts in the dMRI community to bridge the gap between research and clinic in biophysical modelling. Robust methods to acquire data and solve the inverse problem of fitting models to noisy measurements will enable accurate *in vivo* mapping of brain tissue microstructure. Thus, our main objective, has the potential to provide early diagnosis in multiple brain conditions [20] as well as increase our understanding of white matter microstructure.

---

# CHAPTER 2

---

## Diffusion MRI

### 2.1 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging is based on the interaction between the intrinsic magnetic moments (i.e. spins) of the hydrogen nuclei and spatially varying magnetic fields. Spins belonging to a specific tissue region produce a net magnetisation  $\mathbf{M}(t)$ , that can be manipulated to produce an MR signal. Every MR scanner applies a strong and constant magnetic field dubbed  $B_0$  or *main field*. If no perturbations are performed, the net magnetisation will have an equilibrium value of  $M_0$  which will be aligned with  $B_0$ . Generally, this direction is termed *longitudinal* or  $z$  direction, and the plane perpendicular to it is called *transverse plane* ( $xy$ ), usually the imaging plane. When a patient is located inside the bore of the scanner, its hydrogen spins start to align with the main field favouring the configuration with the lowest energy (where the majority of spins are aligned with the main field). This occurs with a precession motion around the  $z$  direction at the *Larmor frequency*,  $\omega_0 = \gamma B_0$ , where  $\gamma$  is the gyromagnetic ratio of the water molecules' spins. When the system is in equilibrium (long after the relaxation time has passed), spins are still precessing but most are practically parallel to  $B_0$ . The macroscopic temporal evolution of  $\mathbf{M}(t)$  at each spatial location is described by Bloch equations [21]

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B} - \frac{M_x \vec{i} + M_y \vec{j}}{T_2} - \frac{(M_z - M_0) \vec{k}}{T_1}, \quad (2.1)$$

## 2. DIFFUSION MRI

---

where  $\vec{i}, \vec{j}, \vec{k}$  are unit vectors defining the coordinate system,  $T_1$  is the spin-lattice relaxation time governing the evolution of longitudinal component of the magnetisation  $M_z$  and  $T_2$  is the spin-spin relaxation time governing the evolution of the transverse component of the magnetisation  $M_{xy}$ .

To acquire an MR image with contrast from different tissues, this equilibrium needs to be broken. This is done by flipping the net magnetisation into the transverse plane. The toss of the magnetization is done by the application of a radio frequency (RF) pulse in resonance with the precession velocity of the spins (*i.e.* the Larmor frequency). After this perturbation, the spins will try to realign with  $B_0$  with a precession movement. During this realignment, their precession generates a variable magnetic field that induces a voltage in the receiver coils of the scanner. This voltage is the signal we measure that will later be converted into an image.

To image a 3D volume, *e.g.* a full brain, the image acquisition is divided into slices. To image a specific slice, only the spins within that slice must be excited. This is achieved by applying a magnetic field gradient perpendicular to the plane of the slices. This generates a range of precession frequencies for each of the slices, and thus, it is dubbed *slice selection* gradient. In this way, when a certain slice needs to be imaged, the applied RF pulse has a frequency that matches the angular velocity of the spins in the desired slice,  $\mathbf{B}_0^{\text{slice}}(z) = \mathbf{B}_0 + \mathbf{G}_{\text{ss}} \cdot z$ , where  $\mathbf{B}_0^{\text{slice}}$  is the actual main field at the slice in the position  $z$ , and  $\mathbf{G}_{\text{ss}}$  is the slice selection magnetic gradient. Then, to recover the spatial positions of each voxel across the excited slice, the frequency and phase of the magnetisation across the whole slice is manipulated. For this, frequency and phase encoding magnetic gradients are applied along the  $x$  and  $y$  directions. These modifications of the frequency and phase of the spins enable measuring the spatial frequency components of the slice image in these two dimensions. The acquired complex signals contain the spatial frequencies of the image. Thus, these can be accommodated in a 2D array (k-space), and by applying a 2D inverse Fourier transform to this array the slice image is obtained with the original spatial locations preserved (see Fig. 2.1). Depending on the MRI modality, the sequence of magnetic fields and gradients may change, generating contrasts that emphasize different tissue properties.

Bloch equations describe the evolution of the net magnetisation,  $\mathbf{M}$ , of a continuum of spins in a field  $\mathbf{B}$ . However, it assumes that spins are stationary. Torrey [22] included an additional diffusion term, generating the well-known Bloch-Torrey equations

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B} - \frac{M_x \vec{i} + M_y \vec{j}}{T_2} - \frac{(M_z - M_0) \vec{k}}{T_1} - \nabla \cdot (D \nabla \mathbf{M}), \quad (2.2)$$

where  $D$  is a constant, the medium's diffusivity. This equation considers the effects of

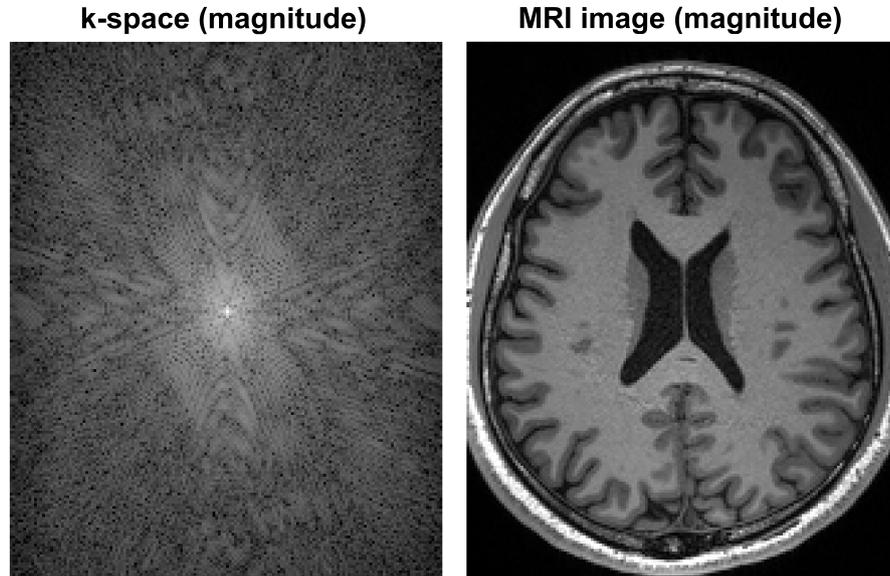


Figure 2.1: K-space magnitude data and the corresponding MRI image (T1-weighted) for a transverse slice of a healthy brain.

diffusive motion on spin magnetisation and sets the foundations for diffusion MRI methods.

## 2.2 Measuring diffusion with MRI

### 2.2.1 Spin Echo

One of the most relevant discoveries in Nuclear Magnetic Resonance (NMR) was the spin echo (SE) effect, first observed by Hahn [23]. He realised that while a single radio-frequency (RF) pulse generates a free induction decay (FID), the application of two successive RF pulses may produce an echo, *i.e.* spin echo (SE) (see Fig. 2.2).

In SE, a 90 degrees RF pulse is first applied to flip the net magnetisation to the plane transverse to the main field. Due to  $T_2^*$  processes, individual spins lose their phase coherence and the overall signal gets attenuated generating a FID. This disorganisation of spin phases is mostly symmetrically reversible. Thus, by applying a 180 degrees RF pulse after a time  $\tau = TE/2$  the phases of the spins take the same time to realign and get mostly refocused, forming an echo after a time  $2\tau = TE$ .

## 2. DIFFUSION MRI

---

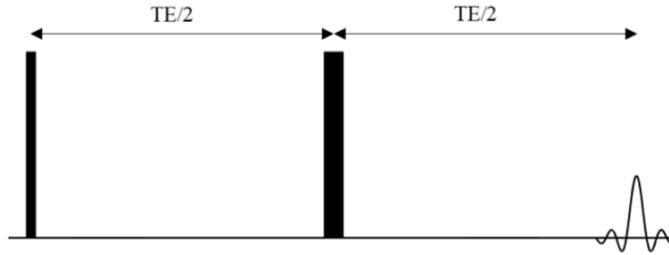


Figure 2.2: The Spin Echo sequence developed by Hahn [23]. Black rectangles denote RF pulses (90 and 180 degrees from left to right) and time is represented in the horizontal axis. TE is the echo time, the time between the 90 degrees RF pulse and the formation of the echo signal.

### 2.2.2 The diffusion phenomenon

Diffusion refers to the physical process that involves the random movement of molecules due to thermal collisions, sometimes called Brownian motion. This process is described by the diffusion propagator  $\mathcal{G}(t, \mathbf{r}, \mathbf{r}_0)$ , which is the probability density of a molecule that started at a location  $\mathbf{r}_0$  to finish at position  $\mathbf{r}$  after a time  $t$ . In a fluid, if we consider times much longer than  $1ps$  we can think of net displacements as a sum of multiple smaller independent displacements [24]. This implies that the net displacement of every molecule follows a Gaussian distribution. Here, the variance becomes proportional to the time  $\langle(\mathbf{r} - \mathbf{r}_0)^2\rangle = 2NDt$ , where  $N$  is the number of dimensions and  $D$  is the fluid's diffusivity. Note that in the absence of flow, the expected displacement due to diffusion is zero.

In the case of free water, the resulting probability of any molecule moving in any direction is the same, *i.e.* isotropic diffusion. Biological tissues contain barriers that restrict and hinder the mobility of water molecules. In this more general case, the propagator will deviate from the Gaussian isotropic behaviour of free fluids and will contain information about tissue barriers. Thus, any diffusion measurement we perform on a certain tissue will have the signature of the underlying architecture restricting the mobility of water molecules. The aim of diffusion MRI research is to develop techniques that enable the extraction of information about tissue from the diffusion-weighted images (DWIs).

### 2.2.3 Single Diffusion Encoding (SDE)

Diffusion weighted images are sensitive to the random displacement of water molecules within a voxel [25], probing tissue at scales considerably lower than the image resolution [24]. The signal obtained from a dMRI experiment depends not only on the diffusion

process but also on the applied pulse sequence. The simplest dMRI experiments use the Pulsed Gradient Spin Echo (PGSE) sequence [26] (Fig.2.3), which is based on the spin echo (SE) phenomenon. The PGSE sequence has two main modules, the excitation and refocusing of the spins, and the acquisition of the echo signal.

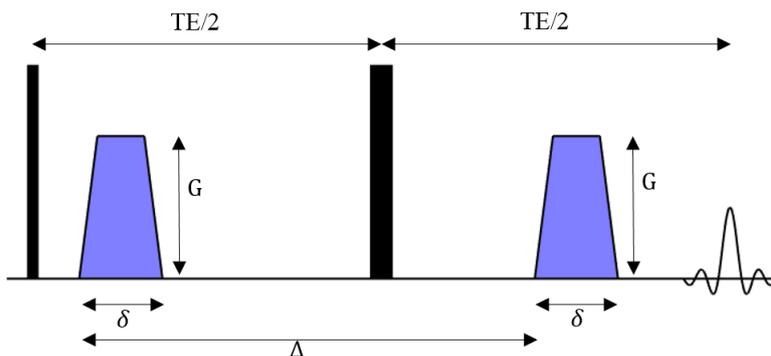


Figure 2.3: The PGSE sequence developed by Stejskal and Tanner [26]. Gradient pulses, shown in blue, are added to a SE sequence.  $G$  is the gradient amplitude,  $\delta$  the gradients' duration,  $\Delta$  the gradients' separation, and  $TE$  the total experimental time after which the echo signal is measured.

In the PGSE sequence a pair of magnetic gradients is added in the refocusing part of a SE. The magnetic gradient field points into the direction of the main static field, but varies linearly in three orthogonal directions. This generates a dephasing for all the spins that move along this direction between the application of the two gradients (i.e. during the diffusion time  $\Delta$ ). As the magnetic field varies linearly through space, each spin will end up with a phase gain that depends on their Larmor frequency and the projection of their random displacement into the diffusion gradient's direction. The phase gain for a spin when a certain gradient waveform  $\mathbf{g}(t)$  is applied during an Echo Time ( $TE$ ) is

$$\phi(TE) = \gamma \int_0^{TE} \mathbf{g}(t') \cdot \mathbf{r}(t') dt', \quad (2.3)$$

where  $\mathbf{r}(t)$  is the random trajectory of the spin. The measured signal is proportional to the net transverse magnetisation in each voxel. The total magnetisation can be computed by summing the contribution of all spins, which is analogue to the phase average over the whole voxel. Since each of them contain huge amounts of spins, we can express it as the expectation ( $\mathbb{E}\{\cdot\}$ ) over the ensemble of spins that exist in the sample [27]:

$$S(TE) = S_0 \mathbb{E}\{e^{i\phi(TE)}\}, \quad (2.4)$$

## 2. DIFFUSION MRI

---

where  $S_0$  is the measurement without diffusion weighting. The signal is attenuated due to the loss of coherence in the phase of the ensemble of spins. This happens in the regions in which spins were able to diffuse along the direction of the gradients. Since this sequence contains only one diffusion encoding period it is dubbed a single diffusion encoding (SDE) scheme [28].

If the gradient pulses in the PGSE sequence are sufficiently narrow, *i.e.*  $\delta \ll \Delta$ , for us to neglect spin motion during their duration, *i.e.* narrow pulse approximation, the area under a single pulse is dubbed  $\mathbf{q}$  and is referred as a *wave vector* [24]. This vector points in the same direction of the diffusion gradient and is given by:

$$\mathbf{q}(t) \equiv \int_0^t \mathbf{g}(t') dt'. \quad (2.5)$$

In this regime, we can write  $\mathbf{q} = \gamma \delta \mathbf{g}$  and the diffusion signal  $S$  as

$$S(\Delta, \mathbf{q}) = \int \rho(\mathbf{r}) \int \mathcal{G}(\Delta, \mathbf{r}', \mathbf{r}) e^{-i\mathbf{q} \cdot (\mathbf{r}' - \mathbf{r})} d\mathbf{r}' d\mathbf{r} = \int G(t, \mathbf{R}) e^{-i\mathbf{q} \cdot \mathbf{R}} d\mathbf{R}, \quad (2.6)$$

from which it can be seen that the signal is the equivalent to the Fourier transform of the voxel-averaged diffusion propagator  $G(t, \mathbf{R})$  [29]. The latter contains information about the medium's architecture but unlike the propagator, it does not depend on starting or ending positions but on the net displacement  $\mathbf{R} = \mathbf{r}' - \mathbf{r}$ . We may think of conventional SDE measurements as a point-sampling of this  $\mathbf{q}$ -space profile.

A typical way of reporting experimental diffusion weightings is done with the well-known *b-value*:  $b \equiv q^2 \Delta$ . In the case of Gaussian diffusion within the voxel's compartments, the signal is fully determined by  $b$  which has SI units of  $ms/\mu m^2$ . However, in more complex environments like those with restricted diffusion, there will be a dependence on both  $\Delta$  and  $q$ .

Performing measurements with different diffusion weightings while pointing the diffusion gradients along different orientations let us hold in the set of acquired dMRI signals information about the tissue architecture restricting the diffusion process.

### 2.3 Analysing Diffusion Weighted Images

We can classify existing techniques that deal with the extraction of information from the DWIs in two main groups, signal representations and tissue models [6]. Signal representations, *i.e.* phenomenological approaches, describe the dMRI signal in a mathematical basis of interest that allows us to compress the information in the measurements to a reduced number of parameters. This is done without specific assumptions about tissue structure. On the other hand, tissue models describe the dMRI signal based on various assumptions

on tissue structure, *e.g.* geometry and diffusivities, with the aim of capturing the most relevant physical properties affecting the signal. Signal representations and tissue models coexist in the dMRI literature and although this thesis focuses on the latter, this section provides a brief introduction to both groups.

### 2.3.1 Signal representations

#### Apparent Diffusion Coefficient

Le Bihan et al. [8] proposed one of the first clinical applications of dMRI as we know it. They applied a PGSE sequence and fitted the free isotropic diffusion expression to the measurements. They dubbed the fitted diffusivity as Apparent Diffusion Coefficient (ADC) to account for restricted diffusion (Eq. 2.7). They also suggested to group all the terms related to the sequence in the signal expression under a factor termed ‘b-value’, which for wide pulses is given by  $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$ . Then, the signal expression is

$$S(b) = S_0 e^{-\gamma^2 G^2 \delta^2 (\Delta - \delta/3) ADC} = S_0 e^{-b ADC}, \quad (2.7)$$

where  $S_0 = S(b = 0)$  is the unweighted diffusion signal. In that work, the estimation of the ADC was done on every voxel of the tissue by performing two measurements at two different b-values. Then, the ADCs from each voxel were used to create an image in which the contrast purely depended on diffusion.

This method provided a new source of contrast for *in vivo* imaging, which proved to have many clinical applications such as detecting ischemic stroke [9]. Although the ADC is currently used for diagnosis of multiple conditions, the information it provides is limited by the main assumptions: considering the diffusion process to be free and isotropic, implying that the diffusion should be independent of the direction of the applied gradients.

#### Diffusion Tensor Imaging

Diffusion tensor imaging (DTI) is a natural extension of the ADC approach, it accounts for anisotropic diffusion. Knowing that the scalar diffusion obtained from the measurements would depend on the direction of the applied magnetic gradient, Basser et al.[30] proposed to replace the ADC with an apparent diffusion tensor (Eq. 2.8) that contains the directional dependence of the medium:

$$\log(S(b, \hat{\mathbf{n}})) = \ln(S_0) - b n_i n_j D_{ij} = -b_{ij} D_{ij}, \quad (2.8)$$

where  $D_{ij}$  is the rank-2 apparent diffusion tensor,  $\hat{\mathbf{n}}$  the diffusion gradient direction, and  $b_{ij}$  the rank-2 b-matrix. From now on, we will use the Einstein summation convention, i.e. summation over repeated indices is implied.

## 2. DIFFUSION MRI

---

This extension considered the contribution from the cross terms of the diffusion gradients and  $D_{ij}$  off-diagonal elements. As the diffusion tensor is symmetric, it only has 6 independent components that need to be fitted to the measurements. In addition,  $S_0$  needs to be computed as well, leaving 7 unknown parameters to be fitted from the acquired data. Therefore, to perform DTI the dataset should consist of at least one non-DWI and at least six noncollinear DWIs. However, due to the low SNR of the DWIs, most DTI protocols use around 30 different directions [31].

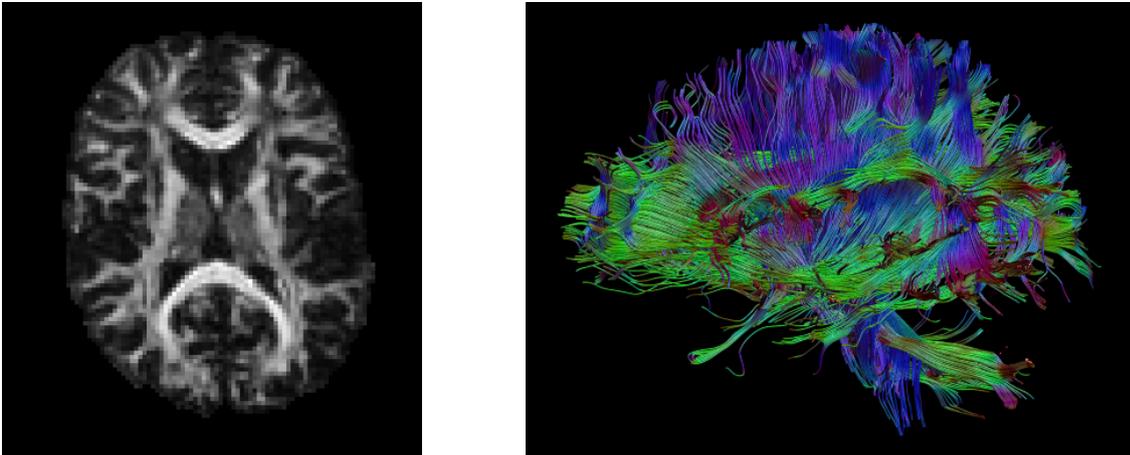


Figure 2.4: Whole brain FA map and deterministic tractography derived from a diffusion tensor analysis.

From the diffusion tensor, eigenvalues and eigenvectors are computed because they hold meaningful information about tissue. Scalar metrics such as the fractional anisotropy (FA) or the mean diffusivity (MD) can be computed from the eigenvalues (Eq. 2.9). These provide a quantitative way to represent tissue and create scalar maps of tissue anisotropy (Fig. 2.4) and diffusivity. These are defined as

$$\text{MD} = \bar{\lambda} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}; \quad \text{FA} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \quad (2.9)$$

where  $\lambda_i, i = 1, 2, 3$ , are the diffusion tensor eigenvalues. The principal eigenvector can be associated with the main orientation of the fibre bundles inside each voxel. Combined with FA, such information can then be used to track fibres within brain WM and perform the so called *fibre tractography* [10, 32], (see Fig. 2.4). There exist many papers addressing this tracking problem, however, its main objective is not to characterise tissue microstructural properties but to reproduce the connections between brain regions.

For optimal precision, DTI acquisitions sample uniformly distributed gradient directions with only one diffusion weighting [33], *i.e.* a shell of measurements, of around  $1 \text{ ms}/\mu\text{m}^2$  since typical values for MD in the brain WM are around  $1\mu\text{m}^2/\text{ms}$ .

### Diffusion Kurtosis Imaging and the Cumulant Expansion

A popular series representation for the dMRI signal is the cumulant expansion [34]. This is a Taylor expansion of the logarithm of the signal attenuation in powers of the diffusion-weighting factor, *i.e.* b-value. If we consider the apparent diffusion and apparent kurtosis as scalars, this expansion up to second order can be written in the following way [35]:

$$\log(S(b)) = \log(S_0) - bD + \frac{1}{6}b^2D^2K + O(b^3). \quad (2.10)$$

With this equation, the measured diffusional kurtosis will depend on the direction of the applied diffusion gradients. Similar to the case of the diffusion tensor, this dependence can be described by:

$$\log(S(b, \hat{\mathbf{n}})) = \ln(S_0) - b n_i n_j D_{ij} + \frac{1}{6}b^2 \bar{D}^2 n_i n_j n_k n_\ell W_{ijkl} + O(b^3), \quad (2.11)$$

where  $W_{ijkl}$  is the aparent kurtosis as defined in [36], a rank-4 fully-symmetric tensor that has 15 independent components, and  $\bar{D}$  is the mean diffusivity.

From Eq. (2.11), it is clear that DKI consists in adding to DTI another term in the cumulant expansion. This increases the complexity of the signal expression, and also raises the number of free parameters from 7 to 22 (including  $S_0$  in the parameter count since it is fitted from the data). Thus, increasing the demands in the data acquisition which in this case usually has two non-zero shells. The kurtosis tensor provides new information about tissue that is not available through DTI. It captures information from restricted diffusion across tissue (*i.e.* non-Gaussian diffusion). Some scalar metrics can be computed from the kurtosis tensor to generate maps of the brain highlighting different sources of contrast. These include the Mean Kurtosis (MK), Radial Kurtosis (RK), Axial Kurtosis (AK), and Kurtosis Anisotropy (KA).

### Diffusion coefficient/tensor distributions

Yablonskiy *et al.* [37] proposed to use a distribution of diffusion coefficients to represent the dMRI signal. Considering the dMRI signal as coming from a large number of non-exchanging Gaussian isotropic compartments we can write

$$S = S_0 \int_0^\infty \rho(D) e^{-bD} dD, \quad (2.12)$$

## 2. DIFFUSION MRI

---

where  $\rho(D)$  is the distribution of diffusion coefficients in the voxel. Following this same concept, it was later proposed that the signal from a voxel can be represented through a distribution of diffusion tensors [38]. Note that we consider these approaches to be representations rather than models.

A comprehensive discussion about signal representations and models is not the scope of this thesis, for that the reader can refer to [39]. Signal representations are very sensitive to microstructure, and have proven their potential in some clinical applications. However, their main limitation is the lack of specificity [12], because different alterations in tissue architecture can produce the same variations in, for example, DTI derived metrics.

### 2.3.2 Biophysical tissue models: The Standard Model

As signal representations have limited specificity, modelling the dMRI signal has become more and more popular since the seminal papers by Van Gelderen *et al.* and Stanisz *et al.* in the 1990s [40, 41]. The microstructural information models aim to extract may lead to biomarkers that are able to detect small tissue alterations corresponding to the early presence of neurodegenerative diseases. This group of approaches is complementary to signal representations since it uses information about what is expected to be measured to unravel dMRI data. These assume brain tissue can be described by a biophysical model containing the most relevant microstructural features. There are plenty of models in the literature, for a detailed discussion see [39]. Most models share common aspects, *e.g.* voxels contain several compartments with negligible water exchange for typical experimental times. Therefore, the total signal is then computed as the sum of the contributions from each pool. Generally, the intracellular compartment is considered to be composed of axons and neurites, while the extracellular space encloses the rest of the structures. Most tissue models focus on WM but there are some that try to describe GM.

In the early 2000s, Behrens *et al.* [42] proposed to model the dMRI signal with two compartments, one isotropic (ball) and one anisotropic (stick). Extending this concept, the *composite hindered and restricted water diffusion model* (CHARMED) [43, 44] aimed to represent the intracellular compartment as impermeable parallel cylinders and the extracellular compartment with a diffusion tensor. Later on, an extension of the CHARMED model dubbed AxCaliber [45] was proposed to estimate axonal diameter distributions. Jespersen *et al.* [13] proposed to add orientational dispersion to these models, by incorporating a spherical harmonics decomposition of the fibre ODF. Zhang *et al.* proposed the *neurite orientation dispersion and density imaging* (NODDI) model to consider axonal orientation dispersion but proposed a Watson distribution to reduce the complexity of the

ODF and enable clinical application [46, 47]. Panagiotaki *et al.* [48] studied the minimum requirements, in terms of modelling compartments, for a precise model of diffusion in white matter.

For brain white matter (WM), the overarching multiple Gaussian compartment framework, unifying a multitude of previous models, is the so-called *Standard Model* (SM) of diffusion in neural tissue. In this thesis, we focus on this model since it is the most general version of WM models (many others are constrained versions of the SM, see recent review by [39]).

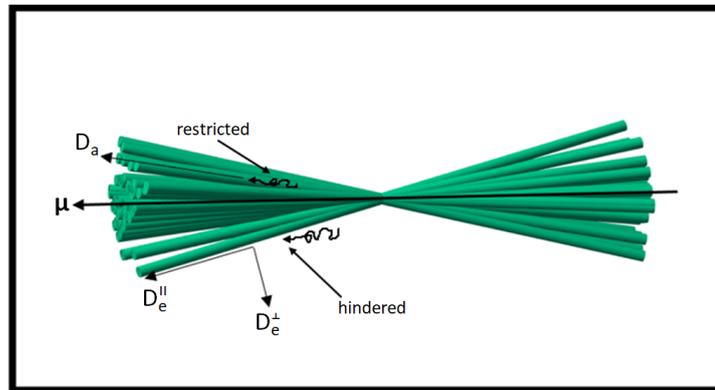


Figure 2.5: Representation of the two compartments present in the Standard Model with their corresponding diffusivities. Here the ODF is given by a Watson distribution.

The stick compartment (sometimes referred as intra-neurite) represents axons, which are expected to be the main contributors to the restricted diffusion signal, and, possibly, dendrites and glial processes [49]. The inclusion of dendrites and glial processes is open to discussion [39] and implies the assumption that in certain regimes (depending on *e.g.* diffusion time) they have similar diffusivity and  $T_2$  relaxation properties and directional distribution, a question which still has not been fully addressed (see discussion in [50]). Sticks are zero-radius cylinders that model fibres, where diffusion occurs only along the fibre's main direction, as it was first proposed for for NAA by [51] and for water in neurites by [13, 42]. Later, Nilsson *et al.* [52] showed theoretically that typical axonal diameters cannot be resolved with SDE and gradient amplitudes available on clinical scanners and thus, are indistinguishable from sticks. This was also confirmed experimentally in [53]. The second compartment represents the extra-neurite space where diffusion is hindered and is modelled as Gaussian anisotropic [49] (zeppelin compartment). A *fibre segment* is defined as the local bundle of aligned sticks with the extra-neurite space surrounding them. Voxels are composed of a large number of fibre segments. The SM consists of the

## 2. DIFFUSION MRI

---

fibre segment signal model (*i.e.* kernel) with the diffusivities and water fraction as free parameters, together with a general fibre orientation distribution function (ODF), which can be represented by its spherical harmonics decomposition

$$\mathcal{P}(\hat{\mathbf{n}}) \approx \frac{1}{4\pi} + \sum_{l=2,4,\dots}^{l_{\max}} \sum_{m=-l}^l p_{lm} Y_{lm}(\hat{\mathbf{n}}). \quad (2.13)$$

Some other works consider a third compartment that represents the contribution from stationary water in fixed tissue [41, 54]. However, recent works [55] have concluded that the signal arising from this compartment can be neglected in most structures for the diffusion times used in the clinic and should only be considered in the cerebellum [56]. Additionally, some versions of the SM include an isotropic diffusion compartment to account for the presence of cerebrospinal fluid (CSF) [39].

Considering a general fibre ODF involves a large set of parameters, which can hinder their unambiguous estimation from the dMRI signal. The NODDI model proposed to use the Watson distribution to model orientation dispersion [47]. However, it added constraints between model parameters to simplify parameter estimation. These were removed in the neurite orientation dispersion and density imaging with diffusivity assessment (NODDIDA) model [57], which is essentially the SM with the only constraint that the fibre ODF must be a Watson spherical distribution  $\mathcal{P}(\hat{\mathbf{u}}) = f(\hat{\mathbf{u}} \mid \hat{\boldsymbol{\mu}}, \kappa)$ , with concentration parameter  $\kappa$  and main direction  $\hat{\boldsymbol{\mu}}$  (see Fig. 2.5). This cylindrically symmetric ODF is usually considered a sufficiently good and parsimonious model for white matter regions without crossing fibres [46].

For the general SM, the signal from a SDE experiment, where the diffusion weighting  $b$  is applied in the direction  $\hat{\mathbf{n}}$ , is given by the convolution over the unit sphere [14]

$$S_{\text{SDE}}(b, \hat{\mathbf{n}}) = S_0 \int_{\mathbb{S}^2} \mathcal{P}(\hat{\mathbf{u}}) \mathcal{K}(b, \hat{\mathbf{n}} \cdot \hat{\mathbf{u}}) dS_{\hat{\mathbf{u}}}, \quad (2.14)$$

where

$$\mathcal{K}(b, \hat{\mathbf{n}} \cdot \hat{\mathbf{u}}) = f \exp[-bD_a(\hat{\mathbf{n}} \cdot \hat{\mathbf{u}})^2] + (1 - f) \exp[-bD_e^\perp - b\Delta_e(\hat{\mathbf{n}} \cdot \hat{\mathbf{u}})^2] \quad (2.15)$$

is the response signal (kernel) from a fibre segment oriented along direction  $\hat{\mathbf{u}}$ . Here,  $f$  is the  $T_2$ -weighted stick volume fraction,  $D_a$  the intra-neurite axial diffusivity, and  $\Delta_e = D_e^\parallel - D_e^\perp$ , with  $D_e^\parallel$ ,  $D_e^\perp$  the extra-neurite diffusivities parallel and perpendicular to the fibre-segment axis [39]. These kernel parameters ( $f$ ,  $D_a$ ,  $D_e^\parallel$ , and  $D_e^\perp$ ) provide important tissue microstructural information, and have shown potential clinical relevance as they are sensitive to specific disease processes such as demyelination, axonal loss or inflammation [15, 16, 58]. Recent works have tried to extend this modelling approach to

grey matter, considering the contribution from cell bodies [59, 60]. One limitation of this model is that within a voxel each fascicle is assumed to have identical diffusion properties, leading to identical microstructural parameters.

## 2.4 Beyond Single Diffusion Encoding

### 2.4.1 Double Diffusion Encoding (DDE)

Most of the dMRI post-processing techniques have been developed for an acquisition performed within an SDE framework. This means that whichever the sequence is, there is only one diffusion encoding period. Since Stejskal and Tanner created the PGSE sequence, there have been many works aimed to maximise the information that can be obtained from a dMRI experiment. Not only from the post-processing perspective, but also, from the data acquisition viewpoint. One of the many modifications that have been proposed to the gradient waveforms involves the addition of multiple pairs of gradients. Cory [61] was the first to propose this, while Mitra [62] generalised this concept for multiple wavevectors. Particularly, a scheme that has recently gained popularity is one termed Double Diffusion Encoding (DDE). Similarly to SDE, this terminology is used for all the experiments that aim to produce two consecutive diffusion encodings [28], irrespective if they are based on a SE, stimulated echo (STE) or PGSE.

This methodology employs two diffusion sensitizing gradient pairs,  $\mathbf{G}_1$  and  $\mathbf{G}_2$ , with durations  $\delta_1$  and  $\delta_2$ , and separations  $\Delta_1$  and  $\Delta_2$  respectively (Fig. 2.6). These diffusion encodings are separated by a mixing time  $\tau_m$ , and can even have different directions. Since

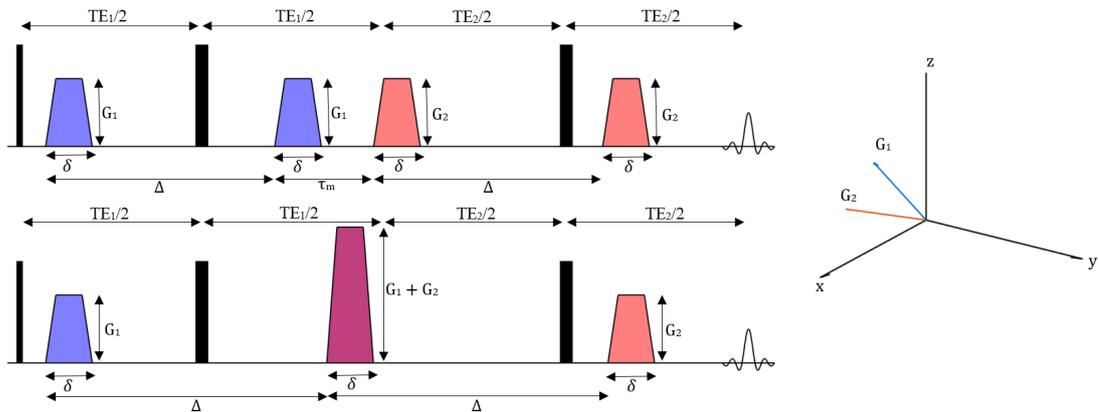


Figure 2.6: A general DDE sequence within a framework of a PGSE sequence [28]. Finite mixing time (top) and zero mixing time (bottom)

## 2. DIFFUSION MRI

---

two diffusion encodings are being used, this sequence offers more degrees of freedom. Even in its simplest (and most common) version, where we choose both wave-vectors with equal amplitude and diffusion times, we can still vary their orientations and mixing time. In DDE the resulting signal loss is also related to the correlation between the displacements during the two diffusion times rather than only the net displacements themselves [28]. The resulting Spin-Echo measured appears at an Echo Time that is usually higher than one of a SDE experiment, what leaves us with a more attenuated signal by the  $T_2$  relaxation. Thus, the SNR in DDE is slightly lower than the one we can achieve in SDE experiments.

Mitra [62] predicted that at long mixing times the signal arising from spherical pores would not present any dependence on the angle between the applied gradients. However, he stated that ellipsoidal pores would present an angular dependence in the fourth term of the signal Taylor expansion irrespective of their packing. This suggested that DDE is able to distinguish spherical pores from randomly oriented cylinders, something that SDE DTI is not able to do. Thus, in the long  $\tau_m$  regime, microscopic anisotropy ( $\mu$ FA) can be revealed. The  $\mu$ FA is a scalar parameter that provides information on the anisotropy of the individual structures that are present in the voxel. Unlike the FA, the  $\mu$ FA is not affected by the relative orientation of the individual pores. Fig. 2.7 shows schematic examples of the complementary information provided by the FA and  $\mu$ FA. The latter is not sensitive to the macroscopic arrangement of the pores while the former is. Lawrenz and Finterbusch [63] were the first to measure microscopic diffusion anisotropy in the living human brain. Furthermore, Jespersen *et al.* [64] developed a model-independent framework to extract the  $\mu$ FA in an orientationally invariant way using only parallel and perpendicular gradients.

Another fact Mitra found was that, for  $\tau_m \rightarrow 0$ , the angular dependence of the signal decay would hold for any pore shape (including spheres), thus giving us a measure of the size of the restricted compartment in the second order of  $q$ . In addition, Özarslan [65] extended the Multiple Correlation Function (MCF) formalism [66] to be applied in DDE sequences where the two gradient directions are different. This approach let us compute analytically the dMRI signal arising from certain geometries. With these results he derived the full angular form of the signal profile for  $\tau_m \rightarrow 0$ , in randomly oriented spheroids and capped cylinders [67].

Although there are no popular tissue models considering a DDE acquisition, many groups have used the extended MCF approach combined with some assumptions about tissue to extract the axon diameter in coherent WM tracts. Komlosh *et al.* [68] were able to estimate the average axon diameter and volume fraction on a pig spinal cord. In a recent work [69], Benjamini *et al.* tried to estimate a non-parametric axon diameter

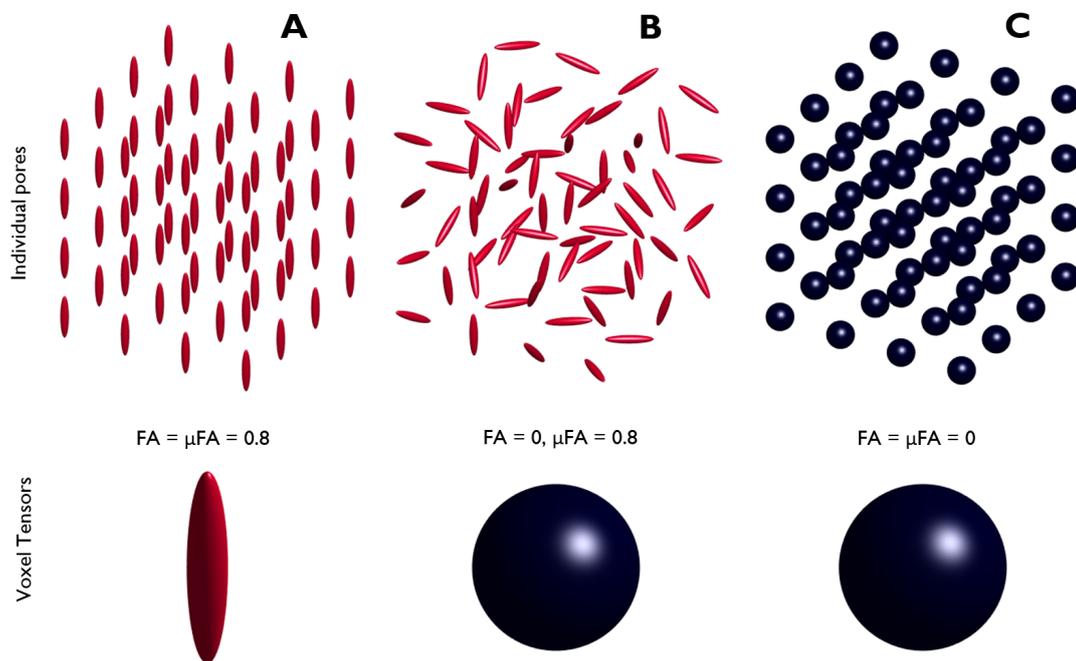


Figure 2.7: Tissue configurations with different FA and  $\mu$ FA values. The top row shows the individual domain structures and the bottom row shows their corresponding voxel-level tensors. In voxels A and B, the domain pores are the same but they differ in their arrangement and thus, have different FA and equal  $\mu$ FA values. The FA cannot distinguish between randomly oriented anisotropic pores (B) and isotropic pores (C).

## 2. DIFFUSION MRI

---

distribution in a ferret’s spinal cord.

*These results are quite impressive, due to the impact this information could have on clinical diagnosis.* However, these methods’ main limitation is the need of an extremely coherent bundle of axons. These approaches assume that the substrate is made of parallel cylinders, and that the diffusion gradients are applied in the plane perpendicular to the main orientation. This makes the signal very sensitive to the axon diameter but, restricts its application to highly coherent structures such as the spinal cord or the corpus callosum.

Jespersen [70] demonstrated that for the low  $q$  regime one could extract the same information using single-PGSE with a varying diffusion time than with multiple-PGSE. Nonetheless, he made emphasis that for intermediate  $b$ -values, DDE contains information which is not immediately accessible using SDE experiments. DDE sequences have the potential to extract novel microstructural information such as the  $\mu$ FA, not reachable through SDE. Nonetheless, it has not been fully addressed if modelling tissue considering single or multiple encodings can provide us with more precise parameter estimations. In Chapter 3, we investigate the advantages of combining biophysical modelling with DDE.

### 2.4.2 Multidimensional dMRI

Unlike conventional dMRI acquisitions, *i.e.* SDE, a single multidimensional dMRI measurement does not probe a point but a trajectory in  $q$ -space [38, 71, 72]. This generalises the concept of diffusion weighting along a direction ( $b$ -vector), to more complex scenarios, *viz.* simultaneously sensitising the MR signal to diffusion along multiple directions (multidimensional dMRI). If we consider each voxel as composed of multiple Gaussian compartments (MGCs), then the signal from any given  $q$ -space trajectory is fully specified by a rank-2 symmetric  $b$ -tensor defined by

$$\mathbf{B} = \int_0^{\text{TE}} \mathbf{q}(t') \otimes \mathbf{q}(t') dt', \quad \text{with } b = \text{Tr}(\mathbf{B}) = B_{ii}, \quad \mathbf{q}(t) = \gamma \int_0^t \mathbf{g}(t') dt', \quad (2.16)$$

where  $b$  is the conventional  $b$ -value or diffusion weighting,  $\mathbf{g}(t)$  is the diffusion gradient waveform and  $\mathbf{q}(t)$  the  $q$ -space trajectory.

The number of non-zero eigenvalues in  $\mathbf{B}$  reflects how many dimensions are being probed simultaneously. In SDE,  $\mathbf{B} = b \hat{\mathbf{n}} \otimes \hat{\mathbf{n}}$ , there is one non-zero eigenvalue, *viz.* linear tensor encoding (LTE). The  $b$ -tensor of a DDE acquisition is  $\mathbf{B} = b_1 \hat{\mathbf{n}}_1 \otimes \hat{\mathbf{n}}_1 + b_2 \hat{\mathbf{n}}_2 \otimes \hat{\mathbf{n}}_2$ , defined from the pair of gradient directions,  $\hat{\mathbf{n}}_1$ ,  $\hat{\mathbf{n}}_2$ , and their individual diffusion weightings,  $b_1$ ,  $b_2$ . It has two non-zero eigenvalues if  $\hat{\mathbf{n}}_1$  and  $\hat{\mathbf{n}}_2$  are not parallel, *viz.* planar tensor encoding (PTE). For multiple Gaussian compartments a SDE acquisition is a subset of the DDE acquisitions (SDE = DDE<sub>||</sub>  $\subset$  DDE), for which  $\hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2$  (parallel

direction pair). When  $\mathbf{B}$  has 3 equal non-zero eigenvalues it is dubbed spherical tensor encoding (STE) or isotropic encoding, since the signal is attenuated by displacements in any direction. Figure 2.8 shows different b-tensors arranged in a triangular diagram with the standard ones (linear, planar, and spherical) at the vertices [73]. The tensor eigenvalues define the shape and diffusion weighting (*i.e.* size) and eigenvectors the orientation.

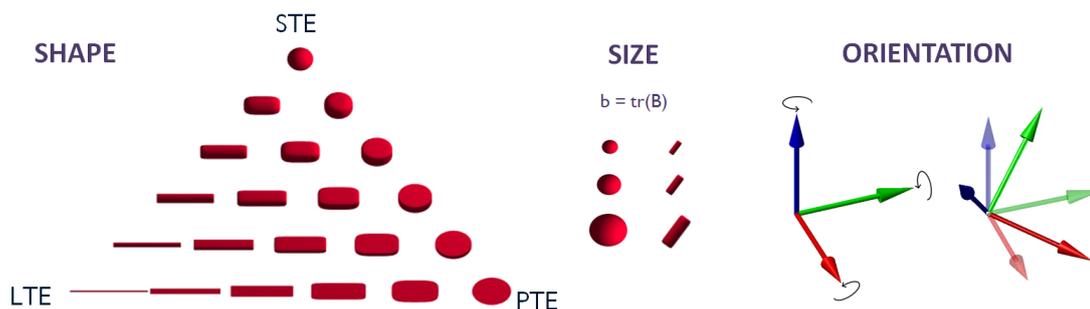


Figure 2.8: Superquadric tensor glyphs representing b-tensor shapes available with multidimensional dMRI. These are arranged in a barycentric ternary diagram [73] according to their linear, planar, and spherical components (LTE, PTE, and STE). Two degrees of freedom define the tensor shape, and an extra one is needed for its size. Three extra degrees of freedom define the tensor orientation.

So far, the scope of multidimensional dMRI has been to disentangle orientation dispersion and microstructural anisotropy by assuming voxels are made of multiple Gaussian compartments represented by an underlying diffusion tensor distribution. Thus, by measuring the diffusion tensor distribution isotropic and anisotropic variance like in [74], information about the microscopic anisotropy can be accessed, *i.e.*  $\mu\text{FA}$ , which may be a biomarker of microstructural degeneration. In Chapter 4, we focus on exploiting the combination of a multidimensional dMRI acquisition with biophysical modelling.

## 2. DIFFUSION MRI

---

---

# CHAPTER 3

---

## Resolving the Watson Standard Model degeneracy

Biophysical tissue models are increasingly used in the interpretation of diffusion MRI data, with the potential to provide specific biomarkers of brain microstructural changes. However, it has been shown recently that, in the general Standard Model, parameter estimation from dMRI data is ill-conditioned even when high b-values are applied.

In this chapter, we analyse this issue for the Neurite Orientation Dispersion and Density Imaging with Diffusivity Assessment (NODDIDA) model and demonstrate that its extension from SDE to DDE resolves the ill-posedness for intermediate diffusion weightings (*e.g.*  $b < 2.5\text{ms}/\mu\text{m}^2$ ), producing an increase in accuracy and precision of the parameter estimation. We analyse theoretically the cumulant expansion up to fourth order in b of SDE and DDE signals. Additionally, we perform *in silico* experiments to compare SDE and DDE capabilities under similar noise conditions. We prove analytically that DDE provides invariant information non-accessible from SDE, which makes the NODDIDA parameter estimation injective. *In silico* experiments show that DDE reduces the bias and mean square error of the estimation along the whole feasible region of 5D model parameter space. DDE adds additional information for estimating the model parameters, unexplored by SDE. We show, as an example, that this is sufficient to solve the previously-reported degeneracies in the NODDIDA model parameter estimation.

#### 3.1 Degeneracy in SM parameter estimation

This thesis focuses on the Standard Model of diffusion in neural tissue. The widely used Neurite Orientation Dispersion and Density Imaging (NODDI) [47] model is a constrained version of the SM that fixes the diffusivity values of the compartments present in the voxel to specific values. This was done to relax the data acquisition requirements and make parameter estimation more stable. However, some of NODDI’s assumptions have been shown to be incompatible with data from spherical tensor encoding (STE) in [75]. It has also been argued that fixing diffusivities introduces bias in the estimation of the remaining model parameters [76]. Jelescu *et al.* [57] extended NODDI by adding the diffusivities to the estimation routine and dubbed it NODDIDA (NODDI with Diffusivity Assessment). This approach eliminated some flawed assumptions made in the original version but led to multiple possible solutions that describe the signal equally well. This made the parameter estimation problem ill-posed or, at least, ill-conditioned, and is usually stated as the existence of degenerated model parameter sets. Recent work by Novikov *et al.* showed that this degeneracy is intrinsic to the SM [14], of which NODDIDA is a special case, independently of the fibre ODF. They show that choosing the correct solution is challenging even with the use of high b-value data, although Jespersen *et al.* [49] obtained stable estimations in ex-vivo brain tissue using extremely high b-values ( $15\text{ms}/\mu\text{m}^2$ ). Kaden *et al.* [77] proposed to use averages in each shell to remove one of the constraints in NODDI. Reisert *et al.* [78] proposed a supervised machine learning approach trained with the expected value of the Bayesian posterior, which, by definition, disregards the possible multimodality of the distribution. Furthermore, it was trained on simulated data with the prior assumption of similar traces for the intra- and extra-axonal diffusivities.

Recently, Lampinen *et al.* [75] have analysed the advantages of a multidimensional encoding over SDE NODDI. They proved that extending the acquisition to incorporate STE data increases the accuracy in quantifying microscopic anisotropy. However, it has not been fully explored, from the point of view of fitting a biophysical model to noisy measurements, if multiple encodings can provide us with more precise model parameter estimates than a single encoding (*cf.* [79, 80]). The advantages of combining linear with planar or spherical tensor encoding to address the degeneracy and increase the precision of parameter estimation have been investigated [81–83] in both *in silico* and/or *in vivo* experiments. Their results show that the estimation precision is increased by the addition of these orthogonal measurements. However, a theoretical background of why this happens is still missing. This chapter extends NODDIDA to a DDE scheme and assesses the accuracy of estimators based on SDE and DDE measurements. We investigate analytically

the information provided by SDE and DDE through their 4th order cumulant expansions.

### 3.2 SDE cumulant expansion and solutions to NODDIDA

It has been recently shown that NODDIDA parameter estimation is challenging in normal experimental conditions [57]. There are two issues here. The first one is that fitting these models to noisy measurements is generally a non-convex optimisation problem, potentially having several local minima of the objective function, requiring appropriate optimisation algorithms. However, the existence of multiple local minima opens the door to a second, more serious, issue: the objective function can present multiple minima with equal or very similar values. In the presence of noise, these minima are perturbed, changing which of the minima is the global minimum [84]. Jelescu *et al.* [57] evidenced this ill-posedness issue for clinically feasible dMRI acquisitions in two particular cases. They showed that the estimated parameters from a collection of independently simulated dMRI measurements follow a bi-modal distribution, despite being simulated from a single ground truth, and the presence of practically indistinguishable spurious minima in the objective function. A recent work by Novikov *et al.* [14] analysed in detail this inverse problem for the unconstrained SM by reparametrising it into its rotational invariants. They concluded that without any constraints on the ODF shape, it was not possible to estimate the kernel parameters with an acquisition sensitive up to order  $\mathcal{O}(b^2)$ .

For intermediate diffusion weightings (*i.e.*  $b < 2.5\text{ms}/\mu\text{m}^2$ ) the dMRI signal is accurately represented by its fourth-order cumulant expansion [85] (sensitive up to  $\mathcal{O}(b^2)$  contributions). For SDE, this expansion is shown in Eq. 2.11:

$$\log(S(b, \hat{\mathbf{n}})) = \ln(S_0) - b n_i n_j D_{ij} + \frac{1}{6} b^2 \bar{D}^2 n_i n_j n_k n_\ell W_{ijkl} + \mathcal{O}(b^3). \quad (3.1)$$

Let us consider a voxel with fibres oriented according to  $\mathcal{P}(\hat{\mathbf{n}}) = c(\kappa) \exp(\kappa(\hat{\boldsymbol{\mu}} \cdot \hat{\mathbf{n}})^2)$ , a Watson ODF with concentration parameter  $\kappa$ . Following an analogous procedure as in [14], we can expand the signal  $S(b, \hat{\mathbf{n}})$  in Eq. 2.14 up to order  $\mathcal{O}(b^2)$ . This gives a mapping between the biophysical (BP) parameter space and the diffusion kurtosis (DK) space, removing the dependence with the acquisition settings and simplifying the analysis of whether different sets of model parameters produce the same signal profile.

Due to the axial symmetry of the Watson distribution, the corresponding diffusion and kurtosis tensors can be expressed in terms of the projection,  $\xi = \hat{\mathbf{n}} \cdot \hat{\boldsymbol{\mu}}$ , of the gradient direction to the main direction  $\hat{\boldsymbol{\mu}}$  [86]:

$$\begin{aligned} D(\xi) &= (f D_a + (1-f) \Delta_e) h_2(\xi, \kappa) + (1-f) D_e^\perp, \\ \frac{1}{3} W(\xi) \bar{D}^2 &= (f D_a^2 + (1-f) \Delta_e^2) h_4(\xi, \kappa) + 2(1-f) \Delta_e D_e^\perp h_2(\xi, \kappa) + (1-f) D_e^{\perp 2} - D(\xi)^2, \end{aligned} \quad (3.2)$$

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

where  $h_2(\xi, \kappa) = \frac{1}{3} + \frac{2}{3}p_2P_2(\xi)$  and  $h_4(\xi, \kappa) = \frac{1}{5} + \frac{4}{7}p_2P_2(\xi) + \frac{8}{35}p_4P_4(\xi)$  are defined as in [86].  $P_2(\xi)$  and  $P_4(\xi)$  are the second and fourth order Legendre polynomials, and  $p_2, p_4$  the non-zero second and fourth order coefficients of the spherical harmonics expansion of the Watson distribution:

$$\begin{aligned} p_2 &= \frac{1}{4} \left[ \frac{3}{\sqrt{\kappa}F(\sqrt{\kappa})} - 2 - \frac{3}{\kappa} \right], \\ p_4 &= \frac{1}{32\kappa^2} \left[ 105 + 12\kappa(5 + \kappa) + \frac{5\sqrt{\kappa}(2\kappa - 21)}{F(\sqrt{\kappa})} \right], \end{aligned} \quad (3.3)$$

where  $F$  denotes the Dawson function [87]. Using these equations, we can derive the relations between BP and DK parameters that fully describe this axially symmetric environment, like in [88] for fully aligned fibres, but here for an arbitrary value of  $\kappa$ :

$$\begin{aligned} D_{\parallel} &= (fD_a + (1-f)\Delta_e) h_2(1, \kappa) + (1-f)D_e^{\perp}, \\ D_{\perp} &= (fD_a + (1-f)\Delta_e) h_2(0, \kappa) + (1-f)D_e^{\perp}, \\ \frac{1}{3}W_{\parallel}\bar{D}^2 + D_{\parallel}^2 &= (fD_a^2 + (1-f)\Delta_e^2) h_4(1, \kappa) \\ &\quad + 2(1-f)\Delta_e D_e^{\perp} h_2(1, \kappa) + (1-f)D_e^{\perp 2}, \\ \frac{1}{3}W_{\perp}\bar{D}^2 + D_{\perp}^2 &= (fD_a^2 + (1-f)\Delta_e^2) h_4(0, \kappa) \\ &\quad + 2(1-f)\Delta_e D_e^{\perp} h_2(0, \kappa) + (1-f)D_e^{\perp 2}, \\ \frac{1}{4}\bar{D}^2(\frac{5}{2}\bar{W} - W_{\perp} - \frac{1}{6}W_{\parallel}) + \frac{1}{4}(D_{\perp} + D_{\parallel})^2 &= (fD_a^2 + (1-f)\Delta_e^2) h_4(\frac{1}{\sqrt{2}}, \kappa) \\ &\quad + 2(1-f)\Delta_e D_e^{\perp} h_2(\frac{1}{\sqrt{2}}, \kappa) + (1-f)D_e^{\perp 2}, \end{aligned} \quad (3.4)$$

where  $\bar{D} = (2D_{\perp} + D_{\parallel})/3$ . Taking the limit for  $\kappa \rightarrow \infty$  we recover the system of equations for parallel fibres presented in [88] (Eq. 12).

In Jespersen *et al.* [86], the equivalent to the system in Eq. 3.4 is solved reaching two alternative equations for  $\kappa$ ,  $\mathcal{F}_{\pm}(\kappa) = 0$ , each giving possible solutions. This suggested that, in general, there should be two solutions, one for each branch. However, this is not always the case, as illustrated in Table 3.1. We derive here an alternative expression of the solution in one equation only. First, we can reparametrize the kernel parameters as:

$$\begin{aligned} \alpha &= fD_a + (1-f)\Delta_e, & \beta &= (1-f)D_e^{\perp}, \\ \gamma &= fD_a^2 + (1-f)\Delta_e^2, & \delta &= (1-f)\Delta_e D_e^{\perp}, & \epsilon &= (1-f)D_e^{\perp 2}. \end{aligned} \quad (3.5)$$

After this substitution, Eq. 3.4 can be expressed as a linear system of five equations for

### 3.2 SDE cumulant expansion and solutions to NODDIDA

the 5 unknowns  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$ , decoupled into two independent smaller systems:

$$\begin{aligned} \begin{bmatrix} D_{\parallel} \\ D_{\perp} \end{bmatrix} &= \begin{bmatrix} h_2(1, \kappa) & 1 \\ h_2(0, \kappa) & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \mathbf{L} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \\ \begin{bmatrix} \frac{1}{3}W_{\parallel}\bar{D}^2 + D_{\parallel}^2 \\ \frac{1}{3}W_{\perp}\bar{D}^2 + D_{\perp}^2 \\ \frac{5\bar{W}\bar{D}^2}{8} - \frac{W_{\perp}\bar{D}^2}{4} - \frac{W_{\parallel}\bar{D}^2}{24} + \frac{(D_{\perp}+D_{\parallel})^2}{4} \end{bmatrix} &= \begin{bmatrix} h_4(1, \kappa) & 2h_2(1, \kappa) & 1 \\ h_4(0, \kappa) & 2h_2(0, \kappa) & 1 \\ h_4(\frac{1}{\sqrt{2}}, \kappa) & 2h_2(\frac{1}{\sqrt{2}}, \kappa) & 1 \end{bmatrix} \begin{bmatrix} \gamma \\ \delta \\ \epsilon \end{bmatrix} = \mathbf{M} \begin{bmatrix} \gamma \\ \delta \\ \epsilon \end{bmatrix}. \end{aligned} \quad (3.6)$$

Observe that the coefficients of matrices  $\mathbf{L}$  and  $\mathbf{M}$  depend on  $\kappa$ . We will ignore for the moment that the 5 unknowns are not independent. The solution is unique as long as matrices  $\mathbf{L}$  and  $\mathbf{M}$  are invertible. This is the case when  $\kappa \neq 0$ , since  $\det \mathbf{L} = p_2$  and  $\det \mathbf{M} = -\frac{1}{2}p_2p_4$ . In the limit of a fully isotropic medium ( $\kappa = 0$ ) the system has only two independent equations, not allowing the recovering of the kernel parameters without additional information. By solving the two systems in Eq. 3.6 we find expressions for  $\alpha, \beta, \gamma, \delta$  and  $\epsilon$  that only depend on  $\kappa$  and the DK parameters. Those variables are actually defined from only 4 kernel parameters (Eq. 3.5), resulting in the coupling equation

$$\gamma(\epsilon - \beta^2) = \alpha^2\epsilon + \delta^2 - 2\alpha\beta\delta. \quad (3.7)$$

By plugging the expressions for  $\alpha, \beta, \gamma, \delta$  and  $\epsilon$  as functions of  $\kappa$  into Eq. 3.7, we obtain a nonlinear equation for  $\kappa$  with potentially multiple solutions. Each solution for  $\kappa$  gives a single solution for  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$ , which in turn, gives a single solution for the kernel parameters:

$$f = 1 - \frac{\beta^2}{\epsilon}, \quad D_a = \frac{\alpha\epsilon - \beta\delta}{\epsilon - \beta^2}, \quad \Delta_e = \frac{\delta}{\beta}, \quad D_e^{\perp} = \frac{\epsilon}{\beta}. \quad (3.8)$$

Thus, the number of solutions to Eq. 3.7 corresponds to the number of BP sets that have the same DK parameters. Table 3.1 presents cases with up to 4 solutions. We computed the number of solutions for 10k random points in the BP space. Most present 2 solutions (70.2%), some only 1 (29.3%), and only a small proportion have 4 solutions (0.5%). This gives rise to the previously discussed degeneracy in model parameter estimation from noisy measurements [57]. In contrast to the claim in Hansen *et al.* [88], even in the extreme case of parallel fibres leaving only four unknowns, the five equations in Eq. 3.4 are independent. This is possible due to the nonlinear nature of the system. If  $\kappa$  is known and not zero (including the limiting case  $\kappa \rightarrow \infty$  of parallel fibres), the full-system is invertible as long as  $f$  is not 0 or 1, and  $D_e^{\perp}$  is not null. In that case, each point in the DK parameter space (signal profile) corresponds to a single BP set. However, this is not the case for an arbitrary unknown  $\kappa$ . Here, the full-system has 5 independent equations with 5 unknowns,

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

but, depending on the parameter values, it can have only one or multiple solutions. This latter case makes the inverse mapping an ill-posed problem.

Using very high b-values might be considered an option to solve this problem, as it will add higher order terms in Eq. 3.1. However, it is still challenging due to very low associated signal-to-noise ratio and is also not feasible in most clinical scanners, although bespoke systems with ultra-strong gradients may provide leverage in this regard [89]. Another solution that does not require powerful gradients is to seek for functionally independent measurements providing new information.

Table 3.1: Illustration of sets of biophysical (BP) parameter values resulting in the same diffusion–kurtosis (DK) parameters. Each plus or minus branch can correspond to a single, multiple, or none BP parameters. Some sets of BP parameters fall outside the region of plausible parameters. The invariants of the not fully symmetric part of  $\mathbf{C}$ , incorporated by DDE, discriminate between the BP parameter sets having the same exact DK representation. All diffusivities are in  $\mu\text{m}^2/\text{ms}$  and  $C$  invariants in  $\mu\text{m}^4/\text{ms}^2$ .

DK parameters [ $D_{\parallel}$ , $D_{\perp}$ , $W_{\parallel}$ , $W_{\perp}$ , $\bar{W}$ ]	Branch	BP parameters [ $f$ , $D_a$ , $D_e^{\parallel}$ , $D_e^{\perp}$ , $\kappa$ ]	C new invariants	
			$\zeta_1$	$\zeta_2$
[1.50, 0.20, 1.46, 0.29, 0.93]	+	[0.73, 2.00, 1.00, 0.30, 8.00]	-0.01	0.21
	-	[0.61, 1.29, 2.19, 0.32, 11.5]	0.023	0.053
[1.56, 1.05, 0.40, 0.71, 0.33]	+	[0.25, 2.37, 1.30, 1.39, 50.0]	0.35	0.62
	-	-	-	-
[0.46, 0.41, 2.90, 2.70, 2.77]	+	[0.88, 1.32, 1.40, -0.23, 0.27]	-0.19	0.02
	-	[0.87, 0.95, 2.00, 0.72, 0.36]	-0.02	0.01
	-	[0.55, 0.18, 1.07, 0.77, 1.41]	0.15	0.00
	-	[0.51, 0.08, 0.93, 0.79, 3.19]	0.16	-0.01
[1.56, 1.26, 0.42, 0.54, 0.51]	+	-	-	-
	-	[0.24, 1.45, 2.10, 1.40, 2.33]	0.24	0.13
	-	[0.19, 0.67, 1.89, 1.49, 5.44]	0.33	0.06

### 3.3 Solving the degeneracy with orthogonal information

#### 3.3.1 Model extension to DDE

DDE adds an extra dimension to the dMRI acquisition, unexplored by SDE experiments. For a general multidimensional acquisition [38, 90], due to the assumption of impermeable compartments, within each of which the diffusion displacement profile is assumed to be

### 3.3 Solving the degeneracy with orthogonal information

---

Gaussian, the signal can be written as:

$$S_{\text{NODDIDA}}(\mathbf{B}) = S_0 \int_{\mathbb{S}^2} \mathcal{P}(\hat{\mathbf{u}}) \mathcal{K}(\mathbf{B}, \hat{\mathbf{u}}) dS_{\hat{\mathbf{u}}}, \quad (3.9)$$

with the kernel

$$\mathcal{K}(\mathbf{B}, \hat{\mathbf{u}}) = f \exp[-D_a B_{ij} u_i u_j] + (1 - f) \exp[-b D_e^\perp - \Delta_e B_{ij} u_i u_j]. \quad (3.10)$$

As discussed in Section 2.4.2, LTE, PTE and STE b-tensors differ in the amount of non-zero eigenvalues, thus, the information we can get from each type of measurement differs.

#### 3.3.2 DDE information gain and unique solution

We show that complementary information from DDE is sufficient to uniquely recover SM parameters from  $\mathcal{O}(b^2)$  measurements. Incorporating DDE measurements transforms the inverse mapping of recovering BP parameters from diffusion-weighted measurements into a well-posed problem. The fourth order cumulant expansion for the dMRI signal arising from a DDE experiment is

$$\begin{aligned} \log(S/S_0) &= -B_{ij} D_{ij} + \frac{1}{2} B_{ij} B_{kl} C_{ijkl} \\ &= -(b_1 n_{1i} n_{1j} + b_2 n_{2i} n_{2j}) D_{ij} + \frac{\bar{D}^2}{6} (b_1^2 n_{1i} n_{1j} n_{1k} n_{1l} + b_2^2 n_{2i} n_{2j} n_{2k} n_{2l}) W_{ijkl} \\ &\quad + b_1 b_2 n_{1i} n_{1j} n_{2k} n_{2l} C_{ijkl}. \end{aligned} \quad (3.11)$$

Here,  $\mathbf{C}$  is the second cumulant tensor of the dMRI signal expansion in terms of the b-tensor and satisfies minor and major symmetries:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{kl ij}, \quad (3.12)$$

but it is not totally symmetric. Its totally symmetric part is proportional to the kurtosis tensor:

$$\bar{D}^2 W_{ijkl} = 3C_{(ijkl)} = C_{ijkl} + C_{iljk} + C_{iklj}. \quad (3.13)$$

For multiple Gaussian compartments or DDE with long mixing times[70],  $\mathbf{D}$  and  $\mathbf{C}$  can be written as

$$\begin{aligned} D_{ij} &= \langle D_{ij} \rangle = \sum_{\alpha} f_{\alpha} D_{ij}^{(\alpha)}, \\ C_{ijkl} &= \left\langle (D_{ij} - \langle D_{ij} \rangle)(D_{kl} - \langle D_{kl} \rangle) \right\rangle = \sum_{\alpha} f_{\alpha} D_{ij}^{(\alpha)} D_{kl}^{(\alpha)} - D_{ij} D_{kl}, \end{aligned} \quad (3.14)$$

where  $f_{\alpha}$  and  $D_{ij}^{(\alpha)}$  denote the fraction and diffusion tensor of compartment  $\alpha$ , including in this summation the integral over the unit sphere with the ODF (*cf.* Eq. 2.14). This

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

motivated naming  $\mathbf{C}$  as the diffusion tensor covariance [38, 70]. Our definition of  $\mathbf{C}$  coincides with the one in [38], and for long mixing times it is also proportional to the  $\mathbf{Z}$  tensor ( $\mathbf{C} = \mathbf{Z}/(4\Delta^2)$ ), earlier introduced in [70]. The  $\mathbf{Z}$  tensor is defined more generally, *i.e.* not restricted to multiple Gaussian compartments, as a cumulant of the DDE signal.

In the case of a Watson ODF,  $\mathbf{W}$  and  $\mathbf{C}$  are transversely isotropic 4th order tensors, *i.e.* they have cylindrical symmetry. Hence, instead of having 15 and 21 independent components they only have 3 and 5, respectively. We can write both tensors as a function of coordinate independent tensor forms (for full derivation see Appendix A.1), like it is done for  $\mathbf{W}$  in [88] (Eq. 6):

$$\mathbf{W} = \omega_1 \mathbf{P} + \omega_2 \mathbf{Q} + \omega_3 \mathbf{I} \quad \text{and} \quad \mathbf{C} = \frac{1}{3} \bar{D}^2 \mathbf{W} + \zeta_1 \mathbf{R} + \zeta_2 \mathbf{J}, \quad (3.15)$$

where  $\mathbf{C}$  was written separating its fully symmetric part from the remaining part [91], and

$$\begin{aligned} P_{ijkl} &= \mu_i \mu_j \mu_k \mu_l, \\ Q_{ijkl} &= \frac{1}{6} \left( \mu_i \mu_j \delta_{kl} + \mu_k \mu_l \delta_{ij} + \mu_i \mu_k \delta_{jl} + \mu_j \mu_k \delta_{il} + \mu_i \mu_l \delta_{jk} + \mu_j \mu_l \delta_{ik} \right), \\ I_{ijkl} &= \frac{1}{3} \left( \delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right), \\ R_{ijkl} &= \frac{1}{2} \left( \mu_i \mu_j \delta_{kl} + \mu_k \mu_l \delta_{ij} \right) - \frac{1}{4} \left( \mu_i \mu_k \delta_{jl} + \mu_j \mu_k \delta_{il} + \mu_i \mu_l \delta_{jk} + \mu_j \mu_l \delta_{ik} \right), \\ J_{ijkl} &= \delta_{ij} \delta_{kl} - \frac{1}{2} \left( \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right), \end{aligned} \quad (3.16)$$

where  $\delta_{ij}$  is the Kronecker delta and  $\hat{\boldsymbol{\mu}}$  the Watson distribution main direction. Eq. 3.15 shows explicitly that  $\mathbf{C}$  contains two extra degrees of freedom independent of  $\mathbf{W}$ . Observe that the fully symmetric part of  $\mathbf{R}$  and  $\mathbf{J}$  vanishes, so that the information encoded in  $\zeta_1$  and  $\zeta_2$  is not accessible from a SDE experiment [64]. We can isolate the new non-symmetric components by the antisymmetrization

$$C_{ijkl} - C_{ikjl} = \zeta_1 (R_{ijkl} - R_{ikjl}) + \zeta_2 (J_{ijkl} - J_{ikjl}). \quad (3.17)$$

Considering a coordinate frame with the  $z$ -axis parallel to the fibers main direction  $\hat{\boldsymbol{\mu}}$ , we can identify

$$C_{xxyy} - C_{xyxy} = \frac{3}{2} \zeta_2 \quad \text{and} \quad C_{xxzz} - C_{zzxx} - C_{xxyy} + C_{xyxy} = \frac{3}{4} \zeta_1. \quad (3.18)$$

Similarly to Eq. 3.4 we can relate the elements of  $\mathbf{C}$  to the biophysical parameters like it

### 3.3 Solving the degeneracy with orthogonal information

was done for  $\mathbf{W}$ . For the SM, including NODDIDA,  $\mathbf{D}$  and  $\mathbf{C}$  are given by

$$\begin{aligned} D_{ij} &= [fD_a + (1-f)\Delta_e]H_{(2)ij} + (1-f)D_e^\perp \delta_{ij}, \\ C_{ijkl} &= [fD_a^2 + (1-f)\Delta_e^2]H_{(4)ijkl} + (1-f)D_e^\perp \Delta_e \left( \delta_{ij}H_{(2)kl} + \delta_{kl}H_{(2)ij} \right) \\ &\quad + (1-f)D_e^{\perp 2} \delta_{ij} \delta_{kl} - D_{ij}D_{kl}, \end{aligned} \quad (3.19)$$

where

$$H_{(2)ij} = \int_{\mathbb{S}^2} \mathcal{P}(\hat{\mathbf{u}}) u_i u_j dS_{\hat{\mathbf{u}}} \quad \text{and} \quad H_{(4)ijkl} = \int_{\mathbb{S}^2} \mathcal{P}(\hat{\mathbf{u}}) u_i u_j u_k u_l dS_{\hat{\mathbf{u}}}. \quad (3.20)$$

These later tensors have the same information as the spherical harmonics of the reciprocal order and contain the corresponding ODF averages. Previous works have worked with the same tensors but named them differently,  $H_{(2)ij} = \langle n_i n_j \rangle$  and  $H_{(4)ijkl} = \langle n_i n_j n_k n_l \rangle$  [14]. Note that  $H_{(4)ijkk} = H_{(2)ij}$  and  $H_{(2)ii} = 1$ . For NODDIDA we get  $h_2(\xi, \kappa) = H_{(2)ij} n_i n_j$  and  $h_4(\xi, \kappa) = H_{(4)ijkl} n_i n_j n_k n_l$ , with  $\xi = \hat{\boldsymbol{\mu}} \cdot \hat{\mathbf{n}}$ . The cross-terms of  $\mathbf{C}$  present new information not accessible from SDE. This makes the DDE signal able to resolve the degeneracy. To make this explicit, we can write the components isolated in Eq. 3.18 in the adapted coordinate frame in terms of BP parameters:

$$\begin{aligned} \frac{3}{2}\zeta_2 &= C_{xxyy} - C_{xyxy} = (1-f) \left[ D_e^\perp \Delta_e \left( H_{(2)xx} + H_{(2)yy} \right) + D_e^{\perp 2} \right] - D_{xx}D_{yy} \\ &= (1-f) \left[ 2D_e^\perp \Delta_e h_2(0, \kappa) + D_e^{\perp 2} \right] - D_\perp^2 \\ \frac{3}{4}\zeta_1 &= C_{xxzz} - C_{zzxx} - C_{xxyy} + C_{xyxy} = (1-f)D_e^\perp \Delta_e \left( H_{(2)zz} - H_{(2)yy} \right) - D_{xx}(D_{zz} - D_{yy}) \\ &= (1-f)D_e^\perp \Delta_e (h_2(1, \kappa) - h_2(0, \kappa)) - D_\perp(D_\parallel - D_\perp) \end{aligned} \quad (3.21)$$

Those 2 equations are independent to the ones in Eq. 3.4, adding complementary information to the mapping between DK and BP spaces (see last column in Table 3.1). Using the same variables defined in Eq. 3.5 we get

$$2h_2(0, \kappa)\delta + \epsilon = \frac{3}{2}\zeta_2 + D_\perp^2 \quad \text{and} \quad (h_2(1, \kappa) - h_2(0, \kappa))\delta = \frac{3}{4}\zeta_1 + D_\perp(D_\parallel - D_\perp) \quad (3.22)$$

These two equations enlarge the system in Eq. 3.6. Following the derivation in Appendix A.2, we demonstrate that they determine a single solution for  $\kappa$ :

$$\frac{h_4(1, \kappa)}{h_4(0, \kappa)} = \frac{\frac{1}{3}W_\parallel \bar{D}^2 - \frac{3}{2}(\zeta_1 + \zeta_2) + (D_\parallel - D_\perp)^2}{\frac{1}{3}W_\perp \bar{D}^2 - \frac{3}{2}\zeta_2} \quad (3.23)$$

since the left-hand side is a strictly monotone increasing function on  $\kappa$ . This agrees with recent work by Cotaar *et al.* [92], who showed that combining different b-tensor shapes

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

---

can determine robustly fibre dispersion. Observe that the cases  $f = 0$  or  $f = 1$  reflect only an apparent degeneracy, as different parameter sets represent the same physical model. In contrast, the case of  $\kappa = 0$  presents a proper degeneracy of the model due to lack of information, where different model instances have identical  $\mathbf{D}$  and  $\mathbf{C}$  tensors.

#### 3.3.3 *In silico* experiments

We have shown theoretically that complementary information from DDE available in  $\mathcal{O}(b^2)$  acquisitions is sufficient for unique estimation of Watson SM parameters. Furthermore, we performed *in silico* experiments in a wide variety of model parameter combinations to show that in the presence of noise DDE estimates are more robust.

#### Signal generation

All synthetic measurements were generated from substrates composed of  $1\mu\text{m}$  diameter cylinders to simultaneously assess our stick approximation. We found this difference was below the noise level. We computed the signal attenuation in the cylinder's perpendicular plane with the Gaussian Phase Approximation (GPA) for both SDE [93] and DDE [80].

Since there is no closed analytical solution for the integral on the sphere in Eq. 3.9, we computed the spherical convolution using Lebedev's quadrature [94]:

$$\int_{\mathbb{S}^2} f(\hat{\mathbf{u}}) dS_{\hat{\mathbf{u}}} \approx \sum_i w_i f(\hat{\mathbf{u}}_i), \quad (3.24)$$

where  $w_i$  are the quadrature weights of each grid point  $\hat{\mathbf{u}}_i$  across the unit sphere. For all configurations of SDE and DDE we used 1,202 quadrature points, which guarantee an exact result up to a  $59^{\text{th}}$  order spherical harmonics decomposition of the ODF. No practical differences were found between the results from our SDE implementation and the analytic summation for SDE in [46].

Finally, Rician noise was added to the synthetic signals, normalising it to obtain a  $\text{SNR} = 50$  for the  $b_0$  measurements, like in [57].

#### Parameter estimation algorithm

Parameter estimation was based on a nonlinear least squares estimator. This was justified due to the relatively high SNR considered for the experiments, where Rician noise can be approximated as Gaussian [95]. We used the Trust Region Reflective algorithm implemented in the MATLAB (R2016a, MathWorks, Natick, MA, USA) optimisation toolbox.

### 3.3 Solving the degeneracy with orthogonal information

The objective cost function was

$$F(\boldsymbol{\theta}) = \sum_i^N (S(\mathbf{B}_i, \boldsymbol{\theta}) - S_{\text{NOIDDIDA}}(\mathbf{B}_i, \boldsymbol{\theta}))^2, \quad (3.25)$$

where  $N$  is the total number of measurements,  $\mathbf{B}_i$  indicates the b-tensor used in the  $i$ -th measurement and  $\boldsymbol{\theta} = \{f, D_a, D_e^{\parallel}, D_e^{\perp}, \kappa\}$  are the model parameters. The main direction of the fibres,  $\hat{\boldsymbol{\mu}}$ , and  $S_0$  were fitted independently in a first stage through a DTI fitting like in [57]. For all configurations, this optimisation procedure was repeated using 30 independent random initialisations for the model parameters to avoid local minima of the five-dimensional cost function. The local solution with the lowest residue was the global optimum.

#### SDE and DDE tested configurations

Five encoding configurations were considered:  $\text{DDE}_{60+0}$ ,  $\text{DDE}_{40+20}$ ,  $\text{DDE}_{30+30}$ ,  $\text{DDE}_{20+40}$ , and  $\text{DDE}_{0+60}$ , with progressively increasing proportions of perpendicular direction pairs,  $b$ , with respect to parallel direction pairs,  $a$ , denoted as  $\text{DDE}_{a+b}$ . Observe that  $\text{DDE}_{60+0}$  is equivalent to SDE if multiple Gaussian compartments are assumed.

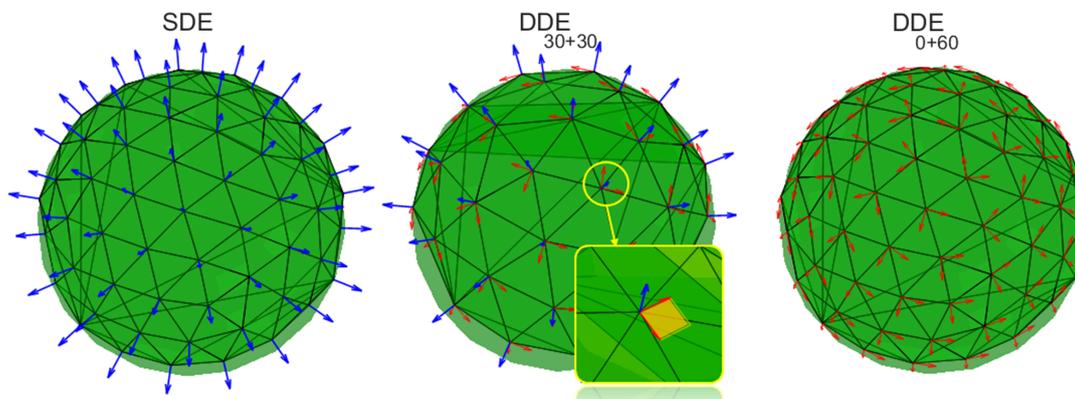


Figure 3.1: Diagram of different measurement protocols ( $\text{SDE}$ ,  $\text{DDE}_{30+30}$ , and  $\text{DDE}_{0+60}$ ). Only  $\text{SDE}$  and  $\text{DDE}_{30+30}$  were used in experiment 1, while they all were used in experiment 2. Blue colours denote the  $\text{SDE}$  directions or  $\text{DDE}$  parallel direction pairs.  $\text{DDE}$  perpendicular direction pairs are in red.

We compared the  $\text{SDE}$  protocol used in [57] against different  $\text{DDE}$  acquisitions with the same number of measurements that can be measured in a similar experimental time. The  $\text{SDE}$  measurement protocol had two shells with b-values of 1 and 2  $ms/\mu m^2$  with 30 directions each [57]. These directions were generated using the Sparse and Optimal

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

---

Acquisition (SOA) scheme [96]. DDE configurations were also divided in 2 shells with the same b-values as above and both directions in each pair had equal individual diffusion weightings,  $b_1 = b_2 = \frac{1}{2}b$ . Thus, perpendicular direction pairs define axially symmetric planar b-tensors, uniquely defined by their normal vector. We generated homogeneously distributed normal vectors using the same algorithm used for the SDE directions. The DDE<sub>30+30</sub> acquisition had 30 parallel direction pairs and 30 perpendicular direction pairs with normal vectors coinciding with the parallel pairs [81] (see middle diagram in Fig. 3.1). The DDE<sub>0+60</sub> protocol had only perpendicular directions pairs (right diagram in Fig. 3.1). Configuration DDE<sub>40+20</sub> had two parallel per each perpendicular directions pair, and DDE<sub>20+40</sub> two perpendicular per each parallel directions pair. All acquisitions had 5 non diffusion-weighted measurements (*i.e.*  $b_0$  measurements).

#### Experimental settings

We performed two *in silico* experiments to assess whether the addition of DDE measurements can enhance the parameter estimation in the presence of typical noise in the measurements.

In the first experiment, we considered two possible instances of NODDIDA parameter values for a voxel in the posterior limb of the internal capsule (PLIC) taken from [57] (see Table 3.2), for which SDE estimates showed a bimodal distribution. We explored in detail whether DDE solves the degeneracy between these particular cases. Only SDE and DDE<sub>30+30</sub> acquisition configurations were considered for this experiment. 2500 independent realisations of Rician noise were added to the synthetic SDE and DDE signals. The SNR was set to 50 in the  $b_0$  measurements to compare our results with [57].

Table 3.2: Ground truth NODDIDA parameters used in experiment 1.

Model parameter	SET A	SET B
f	0.38	0.77
$D_a [\mu m^2/ms]$	0.50	2.23
$D_e^{\parallel} [\mu m^2/ms]$	2.10	0.16
$D_e^{\perp} [\mu m^2/ms]$	0.74	1.48
$c_2 (\kappa)$	0.98 (64)	0.70 (4)

The second experiment aims to compare the accuracy and precision provided by SDE and the different DDE configurations extensively along the feasible region of the full five-dimensional (5D) space of parameters (diffusivities between 0 and  $3\mu m^2/ms$ , fraction between 0 and 1, and  $\kappa$  positive). This allows exploring whether there are subregions

presenting different behaviours. A 5D grid was generated by all the combinations of  $f = [0.1, 0.3, 0.5, 0.7, 0.9]$ ,  $D_a = [0.3, 0.8, 1.3, 1.8, 2.3]\mu m^2/ms$ ,  $D_e^{\parallel} = [0.8, 1.3, 1.8]\mu m^2/ms$ ,  $D_e^{\perp} = [0.5, 1, 1.5]\mu m^2/ms$ , and  $\kappa = [0.84, 2.58, 4.75, 9.27, 15.53, 33.70]$ . The fraction and the diffusivities were selected from a uniform discretisation of the expected range, and  $\kappa$  values were chosen such that the mean-squared-cosine corresponding angle,  $\langle \cos^2 \varphi \rangle = c_2 = \langle (\hat{\mathbf{u}} \cdot \hat{\boldsymbol{\mu}})^2 \rangle = (2\sqrt{\kappa}F(\sqrt{\kappa}))^{-1} - (2\kappa)^{-1}$ , was  $\varphi = [50^\circ, 40^\circ, 30^\circ, 20^\circ, 15^\circ, 10^\circ]$  ( $c_2 = [0.41, 0.59, 0.75, 0.88, 0.93, 0.97]$ ). We generated 50 independent Rician noise realisations (SNR=50 to enable a direct comparison with experiment 1) for the measurements of each combination of the parameters for the five configurations.

### 3.4 Results

Histograms of the estimated model parameters from the first experiment (Fig. 3.2) show an increase in the accuracy and precision of the estimates with the DDE scheme. The bimodal distribution of the estimated parameters is evident with the SDE acquisition, confirming that it is not possible to differentiate true and spurious minima. This effect is removed when using the DDE sequence.

We analysed the shapes of the SDE and DDE objective functions from the synthetic measurements of SET A (sum of squared differences:  $F_A(\boldsymbol{\theta})$ ). To facilitate the visualisation of these 5D functions, we performed a 1D cut through a straight line joining the true and spurious minima of SDE. This was parametrised with the scalar variable  $t$ :  $\boldsymbol{\theta} = t\boldsymbol{\theta}_{\text{spur}} + (1-t)\boldsymbol{\theta}_{\text{true}}$ ;  $t \in [0, 1]$ , where  $\boldsymbol{\theta}_{\text{true}} = [0.38, 0.5, 2.1, 0.74, 64]$  and  $\boldsymbol{\theta}_{\text{spur}} = [0.78, 2.67, 0.32, 0.85, 3.65]$ , with diffusivities expressed in  $\mu m^2/ms$ . Figure 3.3 shows the behaviour of  $F_A(\boldsymbol{\theta})$  along this cut as a function of  $t$ . It can be observed that although the DDE objective function is still bimodal, the spurious and true minima have significantly different absolute values (due to the contribution of the tensor  $\mathbf{C}$  to the DDE signal). This enables to distinguish both peaks in conditions where SDE cannot (*i.e.*  $b_{\text{max}} = 2ms/\mu m^2$ ). Adding more directions to the SDE acquisition would not help to differentiate the peaks. Only by increasing the SDE diffusion weighting the spurious minimum could be differentiated from the true one.

For each point in the 5D grid of parameters, the Root Mean Square Error (RMSE, for definition see for instance [97]) of each parameter has been computed from 50 independent noise realisations. The distributions of RMSE of the parameter estimates from this second experiment are displayed in Fig. 3.4 with violin plots (similar to box plots but showing also estimated probability density [98]). The summary statistics of the RMSE distributions are shown in Table 3.3. On average, DDE<sub>40+20</sub> and DDE<sub>30+30</sub> are the most accurate

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

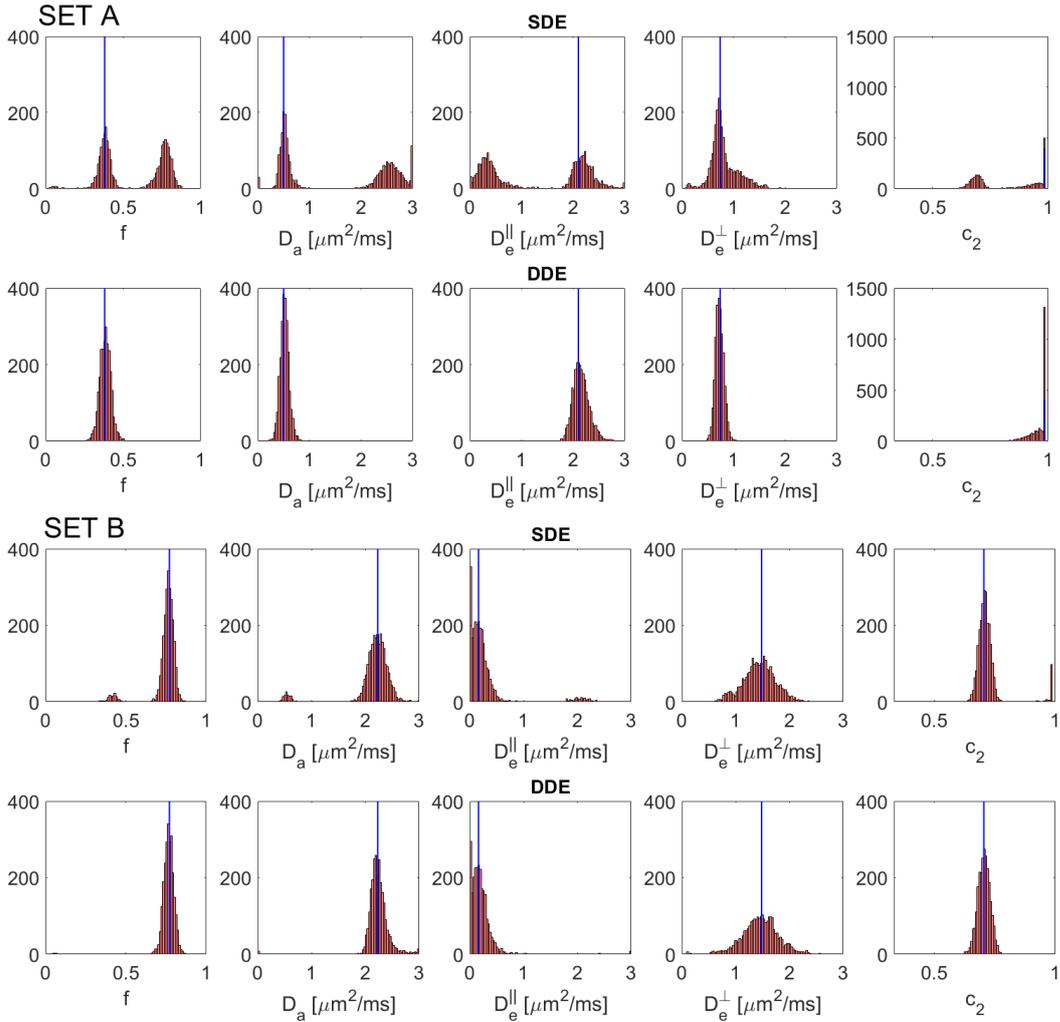


Figure 3.2: Histograms of the estimated model parameters for SDE (top row) and DDE<sub>30+30</sub> (bottom row) schemes in the first experiment for 2,500 independent noise realisations (SNR=50). The ground truth represents two possible solutions of the NODDIDA model applied to a voxel in the PLIC (Table 3.2). These values are shown in blue lines and correspond to set A (upper two rows), and set B (lower two rows).

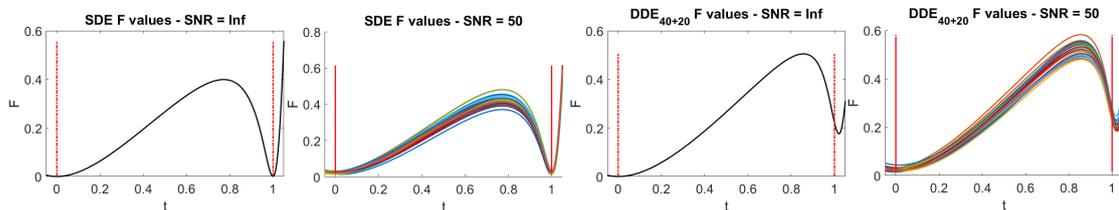


Figure 3.3: Plots of  $F_A(\boldsymbol{\theta}(t))$  for different values of  $t \in [-0.05, 1.05]$ , with  $\boldsymbol{\theta}(t) = t\boldsymbol{\theta}_{\text{spur}} + (1-t)\boldsymbol{\theta}_{\text{true}}$ . Black curves show  $F_A$  values for noise-free SDE and  $\text{DDE}_{30+30}$  acquisitions. The coloured curves show 30 independent realisations of  $F_A$  for  $\text{SNR}=50$ .

configurations for estimating all parameters. This suggests that the incorporation of even a small proportion of DDE measurements can remove the degeneracy, leading to an increase in accuracy and precision. To compare the performance of SDE and DDE in different

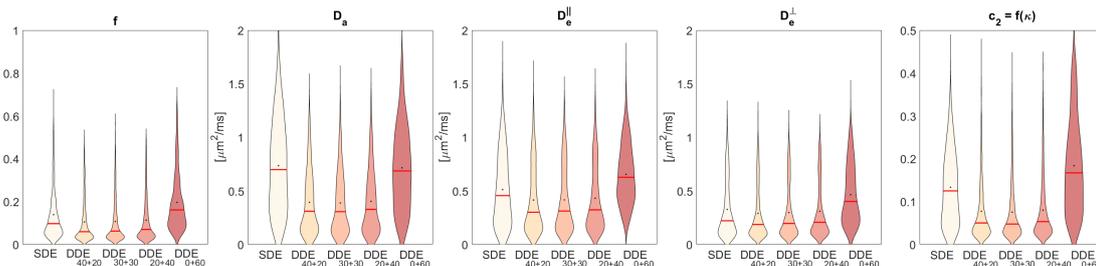


Figure 3.4: Violin plots of the RMSE for all model parameters for all voxels in the 5D grid (a total of  $5 \times 5 \times 3 \times 3 \times 6$ ). Black dots denote the mean and red lines the median.

regions of the parameter space, we projected the 5D RMSE map onto different 3D subspaces. Projections were made by computing the square root of the quadratic mean of the errors in the remaining 2 dimensions,  $E_{\text{proj},ijk} = \sqrt{\sum_{\ell} \sum_{m} E_{ijklm}^2 / (N_{\ell} N_m)}$ . Figure 3.5 shows two different 3D projections, over  $(D_e^{\parallel}, D_e^{\perp}, c_2(\kappa))$  and over  $(f, D_a, c_2(\kappa))$ , of the RMSE of  $f$  and  $D_a$ , respectively. The highest improvement of DDE with respect to SDE is associated with low  $c_2$  values, while for highly aligned voxels the performances of both schemes is similar.

### 3.5 Discussion

In this chapter, we show that modifying the diffusion MRI pulse sequence can mitigate the degeneracy on the Watson-SM (NODDIDA) parameter estimation. Our proposal circumvents the need of presetting diffusivities to *a priori* values as in NODDI. We showed

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

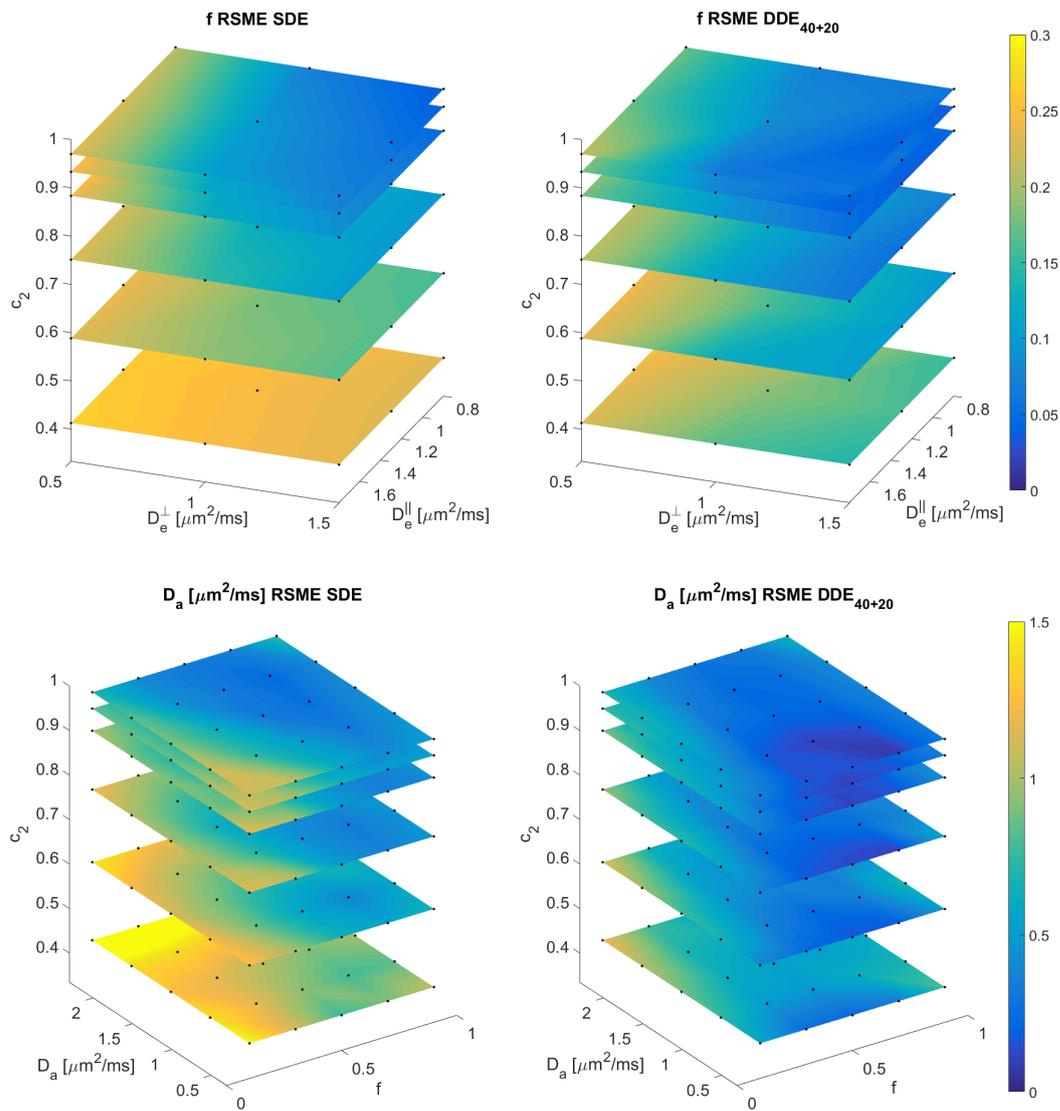


Figure 3.5: The RMSE of  $f$  and  $D_a$ , for SDE and DDE<sub>40+20</sub> acquisition protocols, is shown in top and bottom rows, respectively. These 3D plots show the projection over  $D_e^\parallel$ ,  $D_e^\perp$ , and  $c_2$  for  $f$  and over  $f$ ,  $D_a$ , and  $c_2$  for  $D_a$  of all the points in the 5D grid. Black dots denote actual grid points, linear interpolation was used to generate the colour figures.

Table 3.3: Mean and standard deviation of the RMSE over the whole grid for each acquisition protocol and each of the estimated parameters.

RMSE ( $\mu; \sigma$ )	$f$	$D_a[\mu m^2/ms]$	$D_e^{\parallel}[\mu m^2/ms]$	$D_e^{\perp}[\mu m^2/ms]$	$c_2 = f(\kappa)$
SDE	0.14 ; 0.12	0.74 ; 0.43	0.51 ; 0.33	0.33 ; 0.27	0.13 ; 0.08
DDE <sub>40+20</sub>	0.10 ; 0.10	0.39 ; 0.30	0.41 ; 0.31	0.29 ; 0.25	0.08 ; 0.07
DDE <sub>30+30</sub>	0.11 ; 0.10	0.39 ; 0.29	0.42 ; 0.30	0.30 ; 0.25	0.08 ; 0.07
DDE <sub>20+40</sub>	0.11 ; 0.10	0.40 ; 0.30	0.43 ; 0.31	0.31 ; 0.25	0.08 ; 0.07
DDE <sub>0+60</sub>	0.20 ; 0.13	0.72 ; 0.38	0.65 ; 0.28	0.46 ; 0.27	0.18 ; 0.11

that estimating the NODDIDA model through SDE is generally an ill-posed problem. Depending on the specific combination of model parameters, multiple parameter sets may produce the same signal profile. We show analytically that DDE makes parameter estimation well-posed, and illustrate for a particular model instance the better behaved cost function obtained with DDE. *In silico* experiments over a wide range of model parameter combinations confirmed that extending the acquisition to DDE makes the inverse problem well-posed and solves the degeneracy in the parameter estimation. Combining DDE parallel (*i.e.* LTE) and perpendicular (*i.e.* PTE) direction pairs not only provides more stable parameter estimates but also increases their accuracy and precision.

In Section 3.3.2, we showed that considering a noise-free scenario, in the case of fibres following a Watson ODF with known (nonzero) concentration parameter  $\kappa$  (including the case of parallel fibres), the inverse problem of recovering biophysical parameters from SDE measurements is well-posed. This is not the case for arbitrary unknown concentration  $\kappa$ , where Jelescu [57] first showed experimentally that the parameter estimation from SDE with intermediate b-values was degenerated. This was analysed in [86] showing that there were two nonlinear equations providing possible solutions. We demonstrated that in the absence of noise the number of BP parameter sets that describe the signal equally well up to  $\mathcal{O}(b^2)$  can be either 1, 2, or 4. In contrast, we showed analytically that the  $\mathbf{C}$  tensor includes non-symmetric independent components that are accessible by DDE, but not by SDE, allowing the complete inverse mapping between the cumulant signal representation and BP parameter space. Consistently, the first experiment showed that in both of the PLIC synthetic voxels, DDE leads to more accurate and precise parameter estimations. This is clearly seen when analysing the optimisation cost-function which shows that although DDE also presents multiple local minima, the global minimum is substantially deeper, unlike SDE, thus it can be picked in typical noise levels. However, two points in the 5D model parameter space are insufficient to draw more general conclusions. Therefore, the second

### 3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY

---

experiment swept the parameter space extensively using a regular grid. Mean results (see Table 3.3) showed the minimum RMSE for an acquisition consisting of both linear and planar b-tensors, suggesting that the optimal combination for the scenario considered is between  $DDE_{40+20}$  and  $DDE_{30+30}$  configurations.

Increasing the total number of measurements and SNR will have a larger impact in enhancing DDE parameter estimation than with SDE, since the bimodality present in SDE implies a non-zero lower bound for the achievable MSE even without noise. Results from [82] show that the addition of STE data also leads to an increase in the precision of  $D_a$  and  $f$  in *in vivo* experiments. In our synthetic experiments the addition of PTE data reduces the RMSE in all the parameter estimates (to a lesser extent in  $f$  and  $D_e^\perp$ ). Recently, Dhital *et al.* [83] showed through *in silico* experiments that incorporating PTE data to LTE data enabled us to discriminate spurious solutions in the cost-function. This latter result is explained by our theoretical analysis in Section 3.3.2 where we derive the independent equations provided by DDE that make the inverse problem well-posed. A recent paper by Reisert *et al.* [99] reached similar conclusions for the general SM.

Biophysical models are promising for extracting microstructure-specific information but care must be taken when applying them in dMRI. Some assumptions are more meaningful than others and hence their impact on parameter estimation must be assessed [6]. Invalid assumptions in the model will likely produce bias in the resulting microstructural information, which is epistemic and thus such biases cannot be removed simply by removing the degeneracy. Releasing the diffusivities in the typical two-compartment model eliminates an invalid assumption, reduces possible biases in the estimated parameters, and provides extra information amenable to be used as a biomarker of microstructural integrity and sensitive to specific disease processes [15, 58, 100]. In this chapter, we have focused on analysing the estimability of the model under different acquisition settings. The validation against complementary real data is an independent problem. Limitations in both fronts should be understood further to bring biophysical models to the clinic.

Jespersen *et al.* [49] showed the estimation of the SM was stable and without degeneracy when using extremely high b-values ( $15ms/\mu m^2$ ) on *ex vivo* data. Recent work by Novikov *et al.* [14] studied the unconstrained SM and concluded that if high b-values are unfeasible then orthogonal measurements might be an alternative to uniquely relate the kernel parameters to the signal. Veraart *et al.* extended the SM to acquisitions with varying echo time (TE) [101]. This latter work goes in a similar direction to our work here, *i.e.* adding extra dimensions to the experiment and changing the objective function to avoid ill-posedness. However, measuring multiple directions while varying the TE implies increasing the acquisition time and TE, affecting the SNR. This approach can be combined

with DDE leading to a DDE acquisition with multiple TEs. Recently, Lampinen *et al.* [75] showed that by acquiring data with linear and spherical tensor encodings the accuracy in estimating the microstructural anisotropy was increased compared to that derived from NODDI's parameters. Additionally, Dhital *et al.* [102] measured the intracellular diffusivity using isotropic encoding. These two works point in a similar direction than ours, *i.e.* extending the acquisition to combine different shapes of b-tensors to maximise accuracy and precision. The next chapter studies the generalisation of the model to a multidimensional diffusion acquisition, since the  $\mathbf{C}$  tensor can be fully sampled using different combinations of b-tensor shapes, not only by LTE + PTE.

This chapter's aim was to demonstrate that it is possible to solve the intrinsic degeneracy of the SM with a Watson ODF by using DDE. Although a cylindrically symmetric ODF might be insufficient to model crossing fibres, it may provide a reasonable approximation in the spinal cord and certain other white matter fibre bundles [103], or in highly dispersed tissues like gray matter. Work by Tariq *et al.* has extended the initial NODDI model to a Bingham ODF [104]. Additionally, Novikov *et al.* [14] proposed the unconstrained SM with ODF to be described by a series of spherical harmonics. The next chapter will extend this analysis to general ODFs and will explore biophysical models in general multidimensional dMRI acquisitions [38]. The comparisons made in this chapter between SDE and DDE protocols do not consider the optimisation of the diffusion directions in DDE, just taking four arbitrary chosen DDE protocols extrapolated from an optimised SDE. We expect that further optimisation of the DDE acquisition protocol may also lead to larger improvements. Finally, the largest errors in the parameter estimates occur for  $\kappa \rightarrow 0$ . This might mean that for highly dispersed tissue (*i.e.* grey matter) many measurements might be needed to accurately estimate model parameters.

### **3. RESOLVING THE WATSON STANDARD MODEL DEGENERACY**

---

# CHAPTER 4

---

## Optimal experiment design in Multidimensional dMRI

Biophysical tissue models have been proposed in dMRI research to elucidate the link between microstructural properties and the MR signal. For brain white matter, the research community has developed the SM which has been widely used. However, in clinically applicable acquisition protocols, the inverse problem of recovering SM parameters from a set of diffusion MRI measurements using pairs of short pulsed field gradients was shown to be ill-posed. As it was discussed in Chapter 3, combining linear and planar tensor encoding data allows us to resolve the degeneracy in SM parameter estimation. Given sufficient measurements, multiple combinations of b-tensor shapes may provide enough information to estimate all SM parameters. However, in the presence of noise, some b-tensor sets will provide better results. In this chapter, we focus on combining multidimensional dMRI and biophysical modelling. We develop a framework for optimal experimental design of multidimensional dMRI sequences and apply to the SM although it can be applied to any model. We analyse different metrics that assess the goodness of the protocols and compute their optimal protocols. Finally, we compare their performance against non-optimised acquisitions with noise propagation *in silico* experiments and analysing *in vivo* human data.

### 4.1 Optimal experiment design literature in dMRI

Optimal experimental designs are of key importance in dMRI. Due to the lengthy acquisitions and hardware requirements for most dMRI techniques, especially those going beyond DTI, it is crucial that we squeeze our resources. Limited availability of both acquisition time and advanced hardware, *e.g.* powerful diffusion gradients, hinder clinical translation of dMRI methods. Optimal acquisition strategies provide leverage in these regards by reducing scan times and/or improve the quality of computed diffusion parametric images.

Jones *et al.* [33] were among the first to study the impact of the diffusion directions and b-value for precise diffusion tensor imaging (DTI). With numerical simulations they concluded that the optimal diffusion weighting should be approximately  $1.1 \times 3/\text{Tr}(\mathbf{D})$ , and that diffusion directions should be isotropically distributed for maximum precision. Alexander *et al.* used Monte Carlo simulations to obtain optimal imaging parameters for fibre orientation estimation [105]. They considered one and two fibres per voxel, which they modelled with diffusion tensors. Brihuega-Moreno *et al.* [106] used the well-known Cramér-Rao bounds (CRB) [107, 108] to derive an optimal protocol for apparent diffusion coefficient estimation considering Gaussian noise. The use of such bounds for analysing experimental designs is widespread in estimation theory [97]. CRB were also used by Beltrachini *et al.* to analyse errors in DTI derived metrics when multiple coil acquisition systems are used [109]. Poot *et al.* [110] also used CRB but to propose an optimal acquisition for Diffusion Kurtosis Imaging (DKI) considering Rician distributed noise. Moving away from signal representations, Alexander [111] was one of the first to optimise the diffusion sequence parameters for a biophysical model. He considered the composite hindered and restricted model of diffusion (CHARMED) model [44], and selected the sequence parameters ( $\delta$ ,  $\Delta$ ,  $|g|$  and  $\hat{n}$ ) that minimised the CRB of the model parameters (axon radius, water fraction, parallel and perpendicular diffusivities) considering a Rician noise distribution. Furthermore, CRB have also been used in a DDE scheme to optimise the acquisition parameters in filter exchange imaging [112]. The use real valued data as proposed in [113] has the advantage of making the noise distribution become Gaussian. This simplifies CRB computation, although for high SNR one might approximate the Rician distribution with a Gaussian [95].

Most approaches targeted at optimising the data acquisition have been developed for SDE. Recently, multidimensional dMRI has gained attention in this regard as well. From the acquisition perspective, there are works focusing on the generation of efficient q-space trajectories, *i.e.* gradient waveforms. Dronbjak *et al.* [114, 115] developed a matrix formalism to design optimal gradient waveforms for maximising microstructural. Other

works have focused on minimising the echo time of q-space trajectories while hardware constraints are satisfied [116], and also on generating waveforms that reduce image artifacts [117]. Our goal comes from the modelling viewpoint, which is to optimally select the combination of b-tensor encodings that best inform the SM. Afzali *et al.* [118] used numerical simulations to compare the estimation errors of the SM for various combinations of LTE, PTE and STE encodings. This points in a similar direction to our work here but it was limited to a few discrete combinations of b-tensors and only to WM regions where the ODF can be modelled with a Watson distribution. Additionally, Bates *et al.* [119] proposed to use a quadrature design to recover the underlying diffusion tensor distribution. They explored the full space of axially symmetric b-tensors but only considered the DIAMOND signal representation [120]. Our work in [121] is the most aligned with this chapter. There, we explored the full space of isotropically distributed b-tensors and selected the combination that maximised the precision in the elements of the cumulant expansion, assuming this may be a good enough suboptimal for the full SM.

## 4.2 Rationale of our approach

### 4.2.1 Problem statement

B-tensors are rank-2 and symmetric. Thus, a set of b-tensors (B-set) with  $K$  measurements  $\{\mathbf{B}_k\}_{k=1}^K$ , has  $6K$  independent parameters. We consider that the total number of measurements is fixed. This means that our goal is to find in this  $6K$ -dimensional space the B-set that minimises the estimation error. Thus, our problem factorises into two sub-problems. Firstly, we must define a metric that assesses the estimation error that a certain B-set will have. We propose a loss-function based on the Cramér-Rao Bounds of the SM kernel parameters ( $f, D_a, D_e^{\parallel}$ , and  $D_e^{\perp}$ ). These bounds were computed using an approximation based on the cumulant expansion, which let us discard acquisitions leading to degenerate parameter estimates. Secondly, we have to find the B-set that minimises such metric in the high-dimensional acquisition space, for which we use stochastic optimisation.

Our goal is to maximise accuracy and precision of the SM kernel parameters. A standard loss function that considers both is the mean squared error ( $\text{MSE}(\hat{\theta}) = \text{bias}(\hat{\theta})^2 + \text{var}(\hat{\theta})$ ), where  $\hat{\theta}$  is our parameter estimator. We could use the MSE to choose the optimal B-set but, due to the nonlinearity of the problem, we would have to compute it with numerical simulations of noise propagation like it was done in Section 3.4. Our optimisation demands a large number of loss-function evaluations making MSE minimisation unfeasible. Since a faster metric is needed, we propose to construct our loss-function using CRB.

## 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

---

### 4.2.2 Cramér-Rao Bounds

A widespread approach for optimal experiment design is the minimisation of a loss function based on the CRB of the parameters of interest. If some mild regularity conditions are satisfied (see [97]), these bounds provide a theoretical limit to the variance that any unbiased estimator of the model parameters  $\boldsymbol{\theta} = [\theta_1 \dots \theta_N]$  can have and are defined as:

$$\text{cov}_\theta \geq \text{CRB}_\theta = [I(\boldsymbol{\theta})]^{-1}, \quad I(\boldsymbol{\theta})_{ij} = -\text{E} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(x|\boldsymbol{\theta}) \right], \quad i, j = 1 \dots N, \quad (4.1)$$

where  $f(x|\boldsymbol{\theta})$  is the likelihood,  $I(\boldsymbol{\theta})$  is the Fisher information (a matrix in the multivariate case), and  $\text{cov}_\theta$  is the covariance matrix. If we are not interested in  $\boldsymbol{\theta}$  but in a vector function  $\boldsymbol{\psi}(\boldsymbol{\theta})$ , we may linearly propagate these bounds such that

$$\text{cov}_\psi \geq \text{CRB}_\psi = \frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \text{CRB}_\theta \frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})^t}{\partial \boldsymbol{\theta}}, \quad (4.2)$$

where  $\frac{\partial \boldsymbol{\psi}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$  is the Jacobian matrix and  $\text{cov}_\psi \geq \text{CRB}_\psi$  means that  $\text{cov}_\psi - \text{CRB}_\psi$  is positive semidefinite [97]. CRB are fast to compute when analytical formulations are available. Thus, they are frequently employed to explore the ‘goodness’ of different experimental designs across the acquisition parameter space.

### 4.2.3 Metric definition

We have seen that CRB has several desirable properties that makes it a very good candidate to define our metric. However, Figure 3.3 shows that even acquisitions where the solution to the inverse problem is unique present a multi-modal likelihood. The problem with multi-modal likelihoods is that the existence of local minima far away from the true solution is not penalised by CRB. This is because Fisher information is defined as the expected local curvature, *cf.* Eq. 4.1. For our purpose, this insensitivity is undesired because if these minima have similar values to the global solution, maximum likelihood estimators become degenerate, *e.g.* when using only LTE. Multimodality does not affect the validity of CRB since CRB are still smaller than any achievable variance by unbiased estimators. The problem is that in such cases the actual variance of maximum likelihood estimators is far bigger than the CRB. Thus, is it possible that a B-set with a smaller CRB gives degenerate parameter estimates, which we want to avoid. As shown in Fig. 4.1, a higher local curvature, *i.e.* a smaller CRB value, may not be preferred in scenarios involving multi-modal likelihoods. Thus, computing the CRB directly will not provide sensitivity to the existence of degenerate SM parameter estimates in multidimensional dMRI acquisitions.

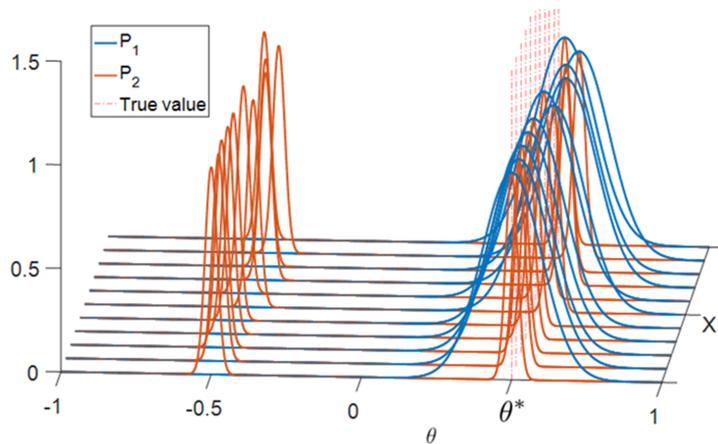


Figure 4.1: Different noise instances (along  $X$ ) of two 1D likelihoods are plotted. Although  $P_1$  has a lower curvature at  $\theta^*$  and thus, higher CRB, it is preferred to  $P_2$ .

To simultaneously avoid degeneracy and maximise precision, we propose a CRB-based metric that is not computed directly but in two steps. In the first step, we compute the CRB for the independent elements of  $\mathbf{D}$  and  $\mathbf{C}$  in the second order cumulant expansion (*cf.* Eq. 3.11). This presents interesting properties. First, it is a convex problem, thus unimodal, meaning that minimising CRB is a good criteria for optimal design. Second, the SM parameters can be fully determined from the cumulant tensors (see [122] and Appendix A.3). Thus, by maximising precision on all cumulants we guarantee that the SM estimation is not degenerate. This first step is similar to our previous work in [121], where we minimised the product of the CRB of all cumulants, *i.e.* the determinant of the Fisher Information matrix, instead of their sum (*cf.* C-optimality vs D-optimality [123]). Nonetheless, the accuracy of the SM parameters is not necessarily monotonic regarding the accuracy of the cumulant tensors, meaning that the cumulant-optimal acquisition may be distinct from the SM-optimal acquisition. Thus, in the second step, proposed in this chapter, kernel parameters are conceived as functions of the cumulants. Here, we propagate the cumulants' CRB to the kernel parameters (see Eq. 4.2), using the mapping between the cumulants and the kernel given by Eq. A.9. The benefit of this second step is that we weigh accordingly each cumulant since they are not equally important for the estimation of the kernel parameters. Therefore we can expect that, in general, our loss function would be a good approximation of the kernel's CRB for acquisitions determining the SM parameters without degeneration, while diverging for degenerate acquisitions. Note that accessing all cumulants is not necessary but a sufficient condition to avoid degeneracy in SM parameter estimation. It may be possible that there are acquisitions that achieve the

## 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

latter without the former although we have not been able to find them.

Like it was done in [111], to construct a scalar metric for the optimisation, we use the sum of the coefficients of variation of the kernel parameters with the CRB replacing the unknown kernel variances:

$$F = \sum_{a=1}^4 \text{CRB}_{\psi} \frac{aa}{x_a^2}. \quad (4.3)$$

Eq. 4.3 depends on the kernel and ODF specific values since these determine  $\text{CRB}_{\psi}$ . Thus, we define our scalar metric by integrating  $F$  over the kernel and ODF parameter space  $\mathcal{H}$  and dividing it by the integration volume:

$$\hat{F} = \frac{1}{\text{vol}(\mathcal{H})} \int_{\mathcal{H}} F(\mathbf{x}) d\mathbf{x}, \quad (4.4)$$

where  $\mathbf{x} = \{f, D_a, D_e^{\parallel}, D_e^{\perp}, p_{\ell m}\}$  and  $p_{\ell m}$  are the spherical harmonics of the ODF (*cf.* Eq. 2.13). This metric assesses how well on average a B-set recovers SM kernel parameters. Simulations showed that the ranking of B-sets by Eq. 4.3 does not vary much with the kernel parameters. Thus, to speed up the computation, we computed Eq. 4.4 averaging  $F$  over 30 different ODF that corresponded to isotropically oriented Watson distributions and a single set of kernel parameters. These values were fixed to  $f = 0.55$ ,  $D_a = 2.1 \mu\text{m}^2/\text{ms}$ ,  $D_e^{\parallel} = 1.8 \mu\text{m}^2/\text{ms}$ ,  $D_e^{\perp} = 0.7 \mu\text{m}^2/\text{ms}$ . Increasing the number of ODFs in the averaging did not vary the metric significantly.

Our interest lies in clinically feasible acquisitions, thus, we limited the maximum b-value to  $b_{\text{max}} = 2 \text{ms}/\mu\text{m}^2$ . We considered that the SNR was  $> 2$  and approximated MR Rician noise as Gaussian [95]. In this case, the Fisher information of the cumulants for a B-set can be written as:

$$I(\{\mathbf{B}_k\}_{k=1}^K; \boldsymbol{\theta})_{ij} = \frac{1}{\sigma^2} \sum_{k=1}^K \frac{\partial S(\mathbf{B}_k; \boldsymbol{\theta})}{\partial \theta_i} \frac{\partial S(\mathbf{B}_k; \boldsymbol{\theta})}{\partial \theta_j}, \quad i, j = 1 \dots 27, \quad (4.5)$$

where  $S(\mathbf{B}_k; \boldsymbol{\theta})$  is the second order cumulant expansion (*cf.* Eq. 3.11),  $\boldsymbol{\theta}$  is a vector containing all independent elements in  $\mathbf{D}$  and  $\mathbf{C}$  (27 in total),  $\sigma^2$  is the noise variance, and  $K$  is the number of total measurements. For more general scenarios, other noise models can be incorporated by modifying Eq. 4.5 but this comes at the cost of higher computational time.

### 4.2.4 Optimisation strategy and constraints

Due to the high dimensionality of the problem and multiple local minima in the space of b-tensors, a hybrid two-step optimisation strategy was adopted: First a stochastic optimization is applied, whose output initializes a subsequent gradient-descent local search.

## 4.2 Rationale of our approach

---

This combines the robustness against local minima of stochastic optimisation and the rapid convergence of greedy approaches. We tested the robustness of our hybrid strategy in high-dimensional toy examples such as Ackley’s function [124] for 50, 100, and 200 dimensions. The Self-Organising Migrating Algorithm (SOMA) [125] was chosen as the best stochastic optimisation from several tested state-of-the-art algorithms [126–128]. Our hybrid approach consistently found the global optimum (results not shown).

We performed the optimisation on three separate scenarios, defined by different sets of constraints to the B-sets, progressively increasing the measurement space. The b-tensors were parametrised with their eigenvalues  $\lambda_1, \lambda_2, \lambda_3$ , which define the tensor shape and size, and their eigenvectors  $\hat{v}_1, \hat{v}_2, \hat{v}_3$ , which define their orientation. Without symmetry constraints, tensor shape is characterised by two parameters, while its size, *i.e.* b-value, is defined by one. For axisymmetric shapes, only one parameter ( $\eta \in [-\frac{1}{2}, 1]$ ) describes the shape  $B_{ij} = b(\eta n_i n_j + \frac{1-\eta}{3} \delta_{ij})$ . LTE, STE and PTE correspond to  $\eta = 1$ ,  $\eta = 0$  and  $\eta = -\frac{1}{2}$ , respectively. Figure 2.8 shows all possible tensor shapes lie in a plane while axisymmetric shapes lie on a line (the two upper edges of the triangle). The first constraint, (C1), fixed the trace and eigenvectors of each b-tensor, leaving free their shapes but restricting these to have axial symmetry:

$$\lambda_{1,i} + \lambda_{2,i} + \lambda_{3,i} = b_i, \quad \text{with} \quad \lambda_{2,i} = \lambda_{3,i}. \quad (4.6)$$

Here, b-values were grouped into two shells of  $1 - 2 \text{ ms}/\mu\text{m}^2$  with half of the measurements each and eigenvectors were distributed uniformly in the hemisphere using a scheme based on [129]. The second constraint, (C2), fixed the eigenvectors like C1, leaving individual shapes (with axial symmetry) and b-values free to vary:

$$0 \leq \lambda_{1,i} + \lambda_{2,i} + \lambda_{3,i} \leq b_{\max}, \quad \text{with} \quad \lambda_{2,i} = \lambda_{3,i}. \quad (4.7)$$

The third constraint, (C3), only fixed the eigenvectors, leaving individual shapes (without imposing symmetries) and b-values free:

$$0 \leq \lambda_{1,i} + \lambda_{2,i} + \lambda_{3,i} \leq b_{\max}. \quad (4.8)$$

Since b-tensors are positive semidefinite and the maximum b-value is limited, the search space was restricted to  $0 \leq \lambda_1 + \lambda_2 + \lambda_3 = b \leq b_{\max} = 2 \text{ ms}/\mu\text{m}^2$ . Figure 4.2 shows a representation of how the three constraints progressively increase the acquisition space.

B-sets of  $K = 60$  b-tensors were considered for the optimisation. We assumed a constant noise level, independent of the tensor shape and diffusion weighting. For each of the 3 constraints, SOMA was run with a population of 60 times the number of free parameters and 500 migrations.

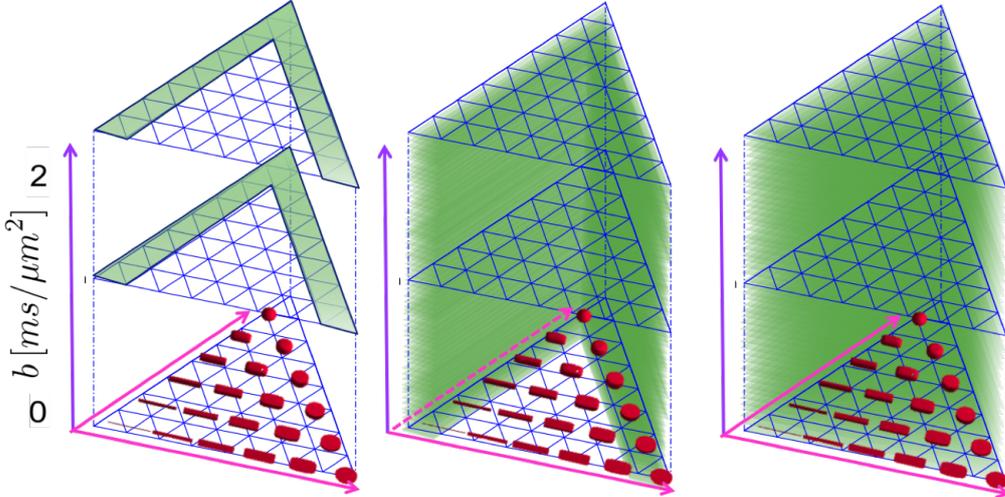


Figure 4.2: Representation of the three optimisation constraints. Since all eigenvectors are fixed, each b-tensor is represented by a point in this 3D space. Here, green indicates the subspace where the b-tensors in each constraint are allowed to be. From left to right: C1 (one free parameter per b-tensor), C2 (two free parameters per b-tensor), and C3 (three free parameters per b-tensor).

### 4.3 Experiments

Using CRB, we have defined a metric that assesses the SM estimation performance of a given combination of b-tensors. With stochastic optimisation, we search in the continuous acquisition space for the B-set that minimises such metric. Finally, we perform a set of synthetic and *in vivo* experiments comparing the estimation errors from the optimised protocol and combinations of LTE, PTE, and STE.

B-sets of  $K = 60$  b-tensors were considered for the optimisation. A constant noise level, independent of the tensor shape and diffusion weighting was assumed. SOMA was run with a population of 60 times the number of free parameters and 500 migrations for each of the constraints.

To assess the performance of the optimal B-set for recovering model parameters, we performed an *in silico* noise propagation analysis. We compared the optimal protocol from C3 against balanced combinations of LTE+PTE, LTE+STE and LTE<sub>only</sub> data (see Section 4.3.1 for acquisition details). Although LTE<sub>only</sub> scanned data had  $b_{\max} = 2\text{ms}/\mu\text{m}^2$ , for *in silico* experiments we considered  $b_{\max} = 5\text{ms}/\mu\text{m}^2$  to highlight the benefit of combining different diffusion encodings. We generated synthetic signals from 7000 random voxels corresponding to each acquisition, added noise (SNR=100, to ensure the existence of

non-degenerate parameter estimation on the optimal acquisition), estimated the kernel parameters and computed the root mean squared error (RMSE). Signals were generated with Eq. 3.9 considering random kernel sets and a crossing fibres ODF with  $\ell_{\max} = 4$ . Additionally, we acquired multidimensional dMRI data on two healthy volunteers to assess the variability in the kernel estimates of our optimised acquisition and compare it with suboptimal acquisitions.

### 4.3.1 MRI data acquisition

Three healthy volunteers were imaged on a Siemens Prisma 3T whole-body MRI scanner. The local Institutional Review Board approved the study and informed consent was obtained and documented from all participants. The MRI scanner had a  $80mT/m$  gradient system and a 64-channel receiver head coil. The body coil was used for transmission. Maxwell-compensated asymmetric waveforms were used for LTE, PTE, and STE acquisitions [117]. Four B-sets were acquired, each with 80 isotropically distributed b-tensors and 6  $b_0$  images. The first B-set LTE+PTE<sub>optimal</sub>, was based on the results from C3 optimisation. It had 30 LTE measurements at  $b = 1ms/\mu m^2$  and 16 LTE + 34 PTE at  $b = 2ms/\mu m^2$ . The second B-set LTE+PTE<sub>suboptimal</sub>, had the same number of LTE and PTE measurements, 10 LTE + 10 PTE at  $b = 1ms/\mu m^2$  and 30 LTE + 30 PTE at  $b = 2ms/\mu m^2$ . The third B-set LTE+STE, had 15 LTE at  $b = 1ms/\mu m^2$ , 41 LTE at  $b = 2ms/\mu m^2$  and 6 STE at  $b = 0.5, 1, 1.5, 2ms/\mu m^2$ . The fourth B-set LTE<sub>only</sub>, had 20 LTE at  $b = 1ms/\mu m^2$  and 60 LTE at  $b = 2ms/\mu m^2$ . The following imaging parameters were kept constant throughout the data acquisition sequences: repetition time: 7.2s, echo time: 94 ms, NEX: 1, resolution: 2.5mm isotropic, in plane FOV: 220mm, slices: 58, parallel imaging: GRAPPA with acceleration factor 2, reconstructed using the adaptive combine algorithm and no partial Fourier. The scan time for all four B-sets was around 45 minutes per participant and LTE measurements in the LTE+STE B-set were subsampled from the LTE<sub>only</sub> dataset.

### 4.3.2 Data processing

MRI data was first denoised using a Marchenko-Pastur principal component analysis method [130]. By preserving only the significant principal components in the signal, this method achieves a reduction of the noise without smoothing. Denoised images were corrected for Gibbs ringing artifacts based on re-sampling the image using local sub-voxel shifts [131]. For both methods we used implementations available in MRtrix3 [132]. These images were rigidly aligned and then corrected for eddy current distortions and subject

## 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

---

motion simultaneously [133]. A  $b_0$  image with reverse phase encoding was used for robust estimation of susceptibility maps [134].

Parameter estimation was based on a nonlinear least squares estimator

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|S(\mathbf{B}) - S(\mathbf{B}, \mathbf{x})\|^2, \quad (4.9)$$

where  $\hat{\mathbf{x}} = \{S_0, f, f_{\text{ISO}}, D_a, D_e^{\parallel}, D_e^{\perp}, p_{\ell m}\}$ . Note that an isotropic compartment with water fraction  $f_{\text{ISO}}$  and fixed diffusivity  $D_{\text{ISO}} = 3 \mu\text{m}^2/\text{ms}$  was added to the estimation procedure of the *in vivo* data. We considered  $\ell_{\text{max}} = 4$  for the spherical harmonic decomposition of the ODF because  $b_{\text{max}} = 2\text{ms}/\mu\text{m}^2$ . We used the Trust Region Reflective algorithm implemented in the MATLAB (R2018a, MathWorks, Natick, MA, USA) optimisation toolbox. For all configurations, this optimisation procedure was repeated using 40 independent random initialisations of the model parameters to avoid local minima of the cost function. The local solution with the lowest residue was considered the global optimum. To analyse SM parametric maps in WM, a region of interest was drawn using an FA map and a T1-weighted image.

## 4.4 Results

### 4.4.1 *In silico* experiments

Optimal configurations are very similar for all sets of constraints. Figure 4.3 shows a representation of the shapes of resulting B-sets. The shapes of the b-tensors are in all cases distributed between linear and planar encoding only. There are no spherical tensors. This is unsurprising since STE measurements are only sensitive to the traces of  $\mathbf{D}$  and  $\mathbf{C}$ , while LTE and PTE each *excite* a different subset of 21 of the 27 independent cumulants. However, there are no mixed or asymmetric shapes either. On the less restricted optimisation (C3), the optimal proportion of LTE data is around 55%. An interesting result is that in the C2 and C3 optimisations, where tensor traces are free parameters, measurements group into two shells with unequal number of measurements and only LTE measurements on the lower shell.

Numerical noise propagation experiments confirmed that B-sets from our optimal experiment design framework are more robust than non-informed combinations of LTE+PTE or LTE+STE data, or even high-b  $\text{LTE}_{\text{only}}$  data. Table 4.1 shows the resulting RMSE from each of the protocols tested. By using the B-set obtained from C3 optimisation, the RMSE was lowered by 10%, 18% and 29% on average, regarding suboptimal combinations of LTE and PTE, LTE and STE or using  $\text{LTE}_{\text{only}}$  respectively. The largest reduction in the RMSE was seen in  $D_e^{\parallel}$  and  $D_e^{\perp}$ .

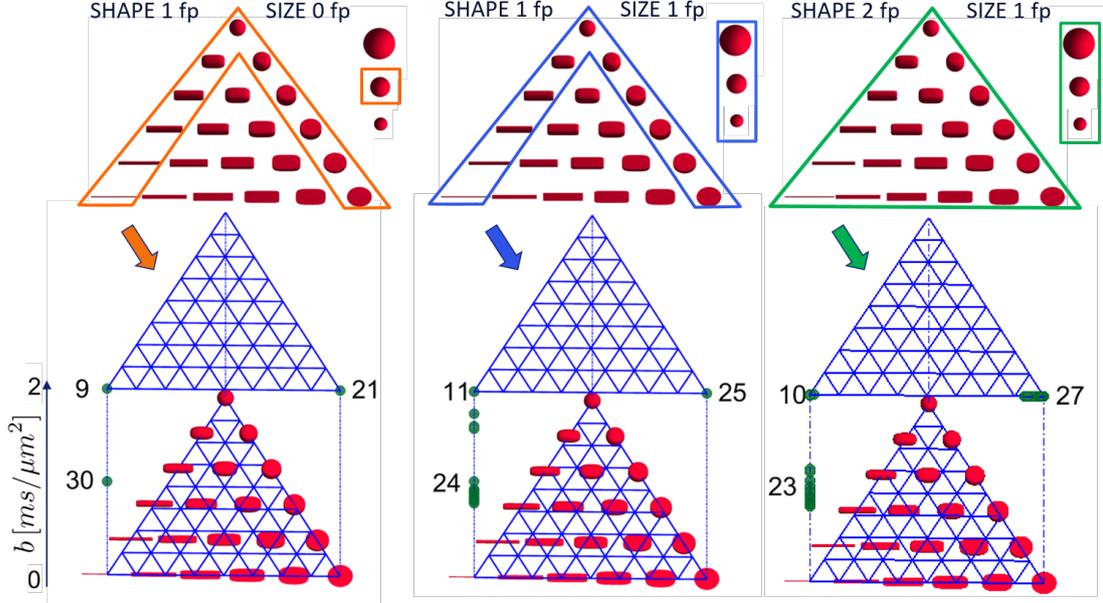


Figure 4.3: Results from the optimisation search for the three constraints, C1, C2 and C3 from left to right. Each green dot represents a b-tensor with its corresponding shape (in plane position) and b-value (vertical position). Although these solutions initially had 1, 2 and 3 free parameters (fp) per tensor, measurements clustered themselves into two shells ( $b \approx 1 - 2 \text{ms}/\mu\text{m}^2$ ) and two different shapes (LTE and PTE).

Table 4.1: RMSE for each acquisition protocol and each estimated parameter (SNR = 100).

RMSE	$f$	$D_a[\mu\text{m}^2/\text{ms}]$	$D_e^{\parallel}[\mu\text{m}^2/\text{ms}]$	$D_e^{\perp}[\mu\text{m}^2/\text{ms}]$
LTE + PTE <sub>optimal</sub>	0.170	0.456	0.509	0.313
LTE + PTE <sub>suboptimal</sub>	0.190	0.496	0.548	0.351
LTE + STE	0.181	0.520	0.689	0.367
LTE <sub>only</sub> ( $b_{\max} = 5 \text{ms}/\mu\text{m}^2$ )	0.229	0.756	0.577	0.327

## 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

### 4.4.2 Human brain experiments

Fig. 4.4 shows SM parametric maps across the white matter for one subject. These were computed independently from the 4 B-sets that were acquired. It can be seen that LTE+PTE parametric maps are less noisy and that values fall in the expected range. Across practically all the WM we find that  $D_a > D_e^{\parallel}$ . To compare the reproducibility of all acquisition protocols, we combined the independent measurements in these four B-sets and estimated the SM parameters from this extended set of 264 DWIs ( $80 \times 3 + 24$ ). Such estimates were used as gold standard and the estimation error was computed for each of the individual B-sets, shown in Fig. 4.5. Averaged across the WM of all subjects, the RMSE was reduced by using the optimised LTE+PTE combination, see Table 4.2. On average, the improvement in the RMSE was of 7%, 27% and 32% with respect to the LTE+PTE<sub>suboptimal</sub>, LTE+STE and LTE<sub>only</sub> acquisitions. Like with *in silico* experiments, the largest improvements are seen in  $D_e^{\parallel}, D_e^{\perp}$ .

Table 4.2: RMSE for each individual B-set and each of the estimated parameters averaged for a coronal WM slice for both subjects. Ground truth values were estimated from merging the four B-sets as done in Fig. 4.5.

RMSE in WM region	$f$	$D_a[\mu m^2/ms]$	$D_e^{\parallel}[\mu m^2/ms]$	$D_e^{\perp}[\mu m^2/ms]$
LTE + PTE <sub>optimal</sub>	0.253	1.251	0.993	0.298
LTE + PTE <sub>suboptimal</sub>	0.266	1.290	1.040	0.346
LTE + STE	0.265	1.281	1.353	0.495
LTE <sub>only</sub>	0.312	1.656	1.321	0.412

## 4.5 Discussion

This chapter proposed a novel framework to optimally select the combination of b-tensors that maximise the accuracy and precision in the estimation of the *Standard Model* microstructural parameters. For this, we defined a metric based on the CRB. Our results were consistent between settings with different constraints: for all of them the optimal measurement scheme was reached by combinations of linear and planar encodings with two diffusion weightings. The proposed proportions of LTE and PTE measurements reduced the RMSE of each parameter relative to non-informed combinations of LTE+PTE or LTE+STE for both *in silico* and *in vivo* experiments. This improvement is cost free.

This is the first work where multidimensional dMRI acquisitions are optimised for biophysical models by exploring the full b-tensor space. Previous work on signal repres-

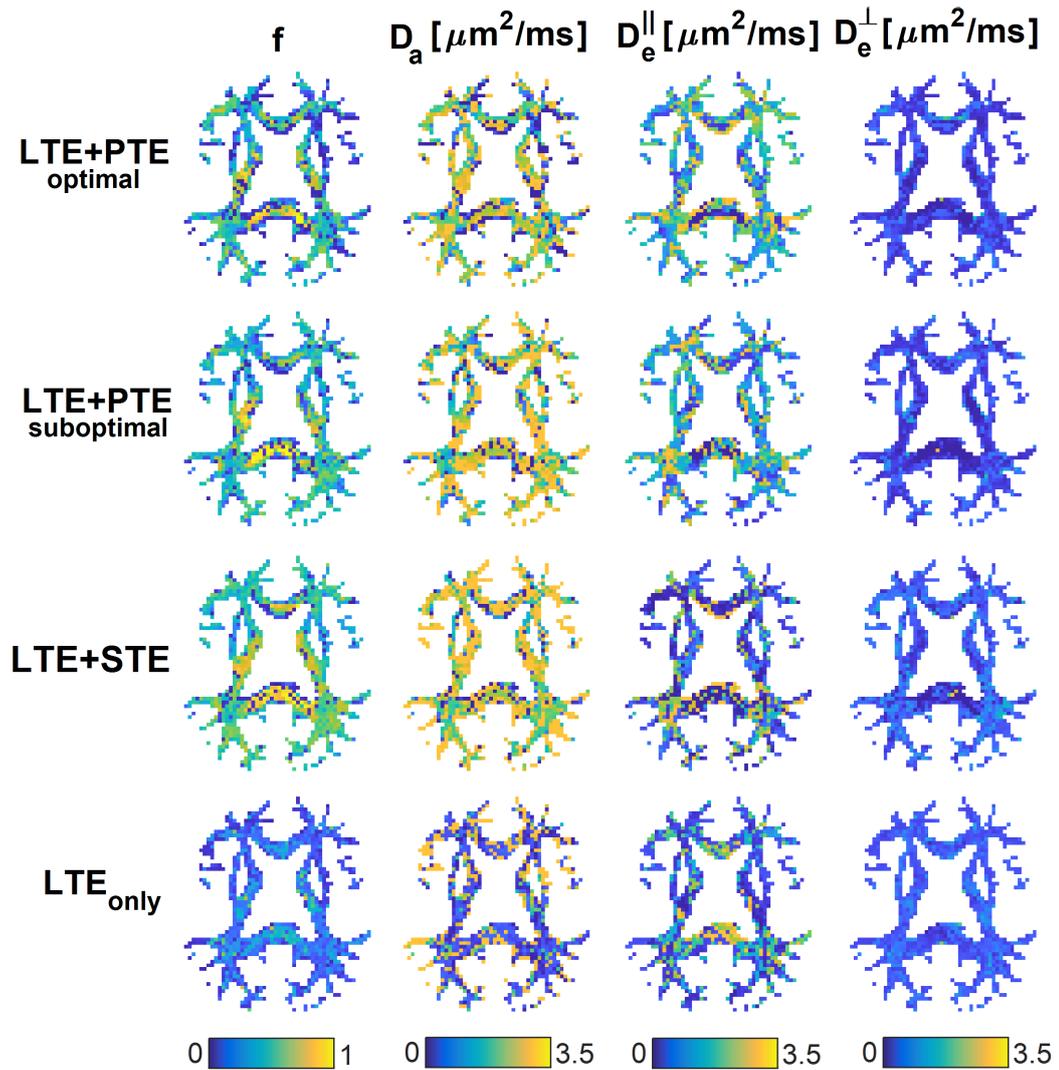


Figure 4.4: White matter maps of  $f$ ,  $D_a$ ,  $D_e^{\parallel}$  and  $D_e^{\perp}$  (columns) estimated from four different B-sets (rows) for a healthy volunteer. It can be seen qualitatively that LTE+PTE B-sets provide better looking maps. Note that the noisy LTE maps evidence the need of additional information. Each B-set took less than 10 minutes to acquire.

#### 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

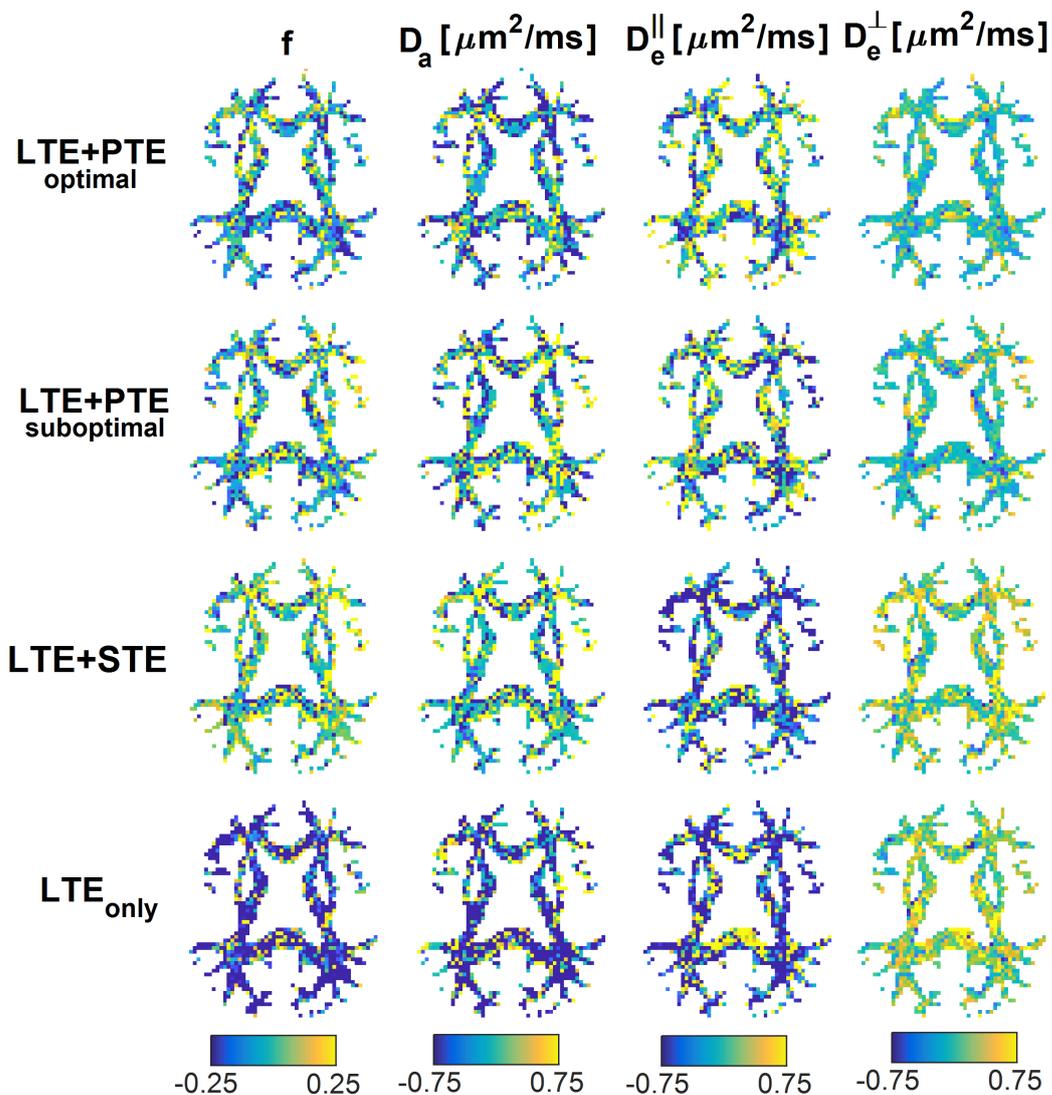


Figure 4.5: White matter error maps of  $f$ ,  $D_a$ ,  $D_e^{\parallel}$  and  $D_e^{\perp}$  (columns) for a healthy volunteer scanned with four different B-sets (rows). The ground truth for reference was estimated from merging all four B-sets. It can be seen qualitatively that LTE+PTE B-sets have smaller errors.

entations analysed the impact of different combinations of LTE, PTE, and STE, in the variation of microscopic anisotropy and isotropic kurtosis estimates [135] and in the estimation of constrained spherical deconvolution [136]. Afzali *et al.* [118] assessed the impact of different combinations of these b-tensors in parameter estimation of the SM with a Watson ODF. Furthermore, [119] considered the continuous space of axially symmetric b-tensors but aimed at the DIAMOND signal representation. This chapter is aligned with our previous paper [121], however, there the focus was not on the SM directly but on maximising the precision in the cumulant expansion.

Since almost no works have discussed optimal experimental designs for multidimensional dMRI [118, 119], we aimed to provide insights on how to best sample the full b-tensor space. The confirmation that no STE or non-axially symmetric b-tensors are involved in optimal acquisition schemes is interesting and somewhat unexpected. This might be explained by looking at Eq. 3.11, which shows that STE only couples to two degrees of freedom of the 2nd-order cumulant expansion (the traces of  $\mathbf{D}$  and  $\mathbf{C}$ ). This possibly indicates that LTE and PTE measurements are potentially more informative. Modifying slightly the proportions of LTE and PTE data does not significantly affect the loss function. However, the optimal B-set outperformed a non-informed combination of LTE and PTE data in both *in silico* and *in vivo* experiments (See Table 4.1 and Table 4.2). Although low-b signal features, *i.e.*  $\mathcal{O}(b)$ , are accessible by both LTE and PTE, it seems LTE is more efficient for extracting these since no PTE measurements are selected with low b-values. Replacing PTE with STE measurements has a high impact on the kernel estimation error as shown in Tables 4.1 and 4.2.

Even for the optimal acquisition,  $D_a$  and  $D_e^{\parallel}$  estimates were noisy due to the small  $b_{\max}$ .  $D_a$  estimated values were centred at  $2 \mu\text{m}^2/\text{ms}$  in the WM (see second column in Fig. 4.4), which follows measurements using high diffusion weighted PTE done by [102] and with LTE measurements analysing the  $1/\sqrt{b}$  scaling with very high b [53]. Our estimates in most of the WM show that  $D_a > D_e^{\parallel}$ , agreeing with gadolinium-based contrast experiments in the rat corpus callosum [137]. We would like to stress that each acquired B-set had only 80 diffusion-weighted measurements since we wanted to compare different acquisition settings. A clear improvement can be seen when comparing parametric maps from Fig. 4.4 against those in [14] with over 1300 LTE measurements including high-b, showing the gain achieved by combining LTE and PTE measurements.

Cumulants are defined as derivatives of  $\log(S)$  at  $b = 0$ , *i.e.* *true cumulants*. When fitting Eq. 3.11 to the measurements, if we consider appropriate b-ranges, we get a good approximation, *i.e.* *fitted cumulants*. We do not propose to compute the SM parameters from the fitted cumulants, similar to what is done by [103]. The nonlinearity of the

## 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

---

cumulants-SM mapping (Eq. A.9), amplifies the bias in the *fitted cumulants* to the kernel [14]. Instead, we propose to use the cumulants to help us discard B-sets leading to degeneracy in the parameter estimation. We assume *true* and *fitted* cumulants have similar variances, and thus, we can compute  $\text{CRB}_\theta$  by plugging in Eq. 3.11 into Eq. 4.5. Furthermore, the linear propagation given by Eq. 4.2 assumes small deviations so that higher order terms are negligible. Thus, our metric has two strengths. First, it assigns a large value of  $\text{CRB}_\theta$  and thus, of  $\text{CRB}_\psi$ , to acquisitions that do not access all cumulants, which makes the optimisation avoid degeneracy. Second, it weighs  $\text{CRB}_\theta$  so that  $\text{CRB}_\psi$  relates to precision in the kernel parameters, penalising large variances over smaller ones.

In this chapter we are limited to intermediate diffusion-weightings, *i.e.*  $b_{\max} = 2 \text{ ms}/\mu\text{m}^2$ . This enabled us to define our metric based on the cumulant expansion appropriate approximation only in this b-range [85]. To look for the optimal B-set in a larger b-range we would need to change the metric since the assumption of similarity in the variances of *true* and *fitted* cumulants may not hold.

Our framework searches for the optimal B-set in the acquisition subspace where there is no degeneracy, this might be too restrictive. Degeneracy is harmless when spurious solutions are close to the true solution, or if they can be discarded for falling outside the range of feasible values (see discussion in A.4). Since our loss function takes high values in degenerate cases, the optimisation does not rule out the possibility that there exists a b-tensor combination that leads to a degenerate parameter estimation while achieving a smaller MSE (see discussion in A.4). To test this we included an LTE+STE B-set in both *in silico* and *in vivo* experiments, which still had larger RMSE than the optimal protocol.

The high-dimensional optimisation has multiple local minima, thus, further improvements on the optimisation strategy may improve the results. However, our hybrid approach showed robust estimations of global optima in a toy function of similar dimensionality and complexity. We do not explore the  $6K$ -dimensional acquisition space but restrict ourselves to isotropically distributed b-tensor orientations ( $3K$ -dimensional). Releasing this constraint could improve the estimation if the optimal orientations of each group of LTE and PTE measurements involved some relationship between them. However, this would likely depend on the details of the ODF, potentially excluding whole brain optimization. Therefore, constraining them to be independent and isotropically oriented is reasonable and simplifies the problem considerably.

For *in vivo* experiments, the ground truth is unknown and this hinders the analysis of the RMSE. However, by merging all datasets (all encodings), the estimation error is significantly reduced and this provides a fair estimate for comparison.

Finally, we did not consider different SNR for each b-tensor shape, although the dif-

ferences in echo times for optimised waveforms are generally not negligible [116]. Furthermore, the assumption that a given q-space trajectory leads to planar or spherical encoding only holds if we can assume that voxels are comprised of multiple Gaussian compartments. Violations to this assumption, *i.e.* the presence of non-Gaussian diffusion, will likely cause orientational dependence in STE measurements and imperfect PTE [138]. This is more evident when shorter waveforms are used, however, a way to reduce such effect is by averaging rotated STE measurements [139].

Future work will explore modifications to the metric to analyse high-b regimes and noise level dependence on the tensor shape. This may be aligned with an acquisition with different echo times to simultaneously measure compartmental  $T_2$  values as it was proposed by [101]. Accounting for higher diffusion weightings and different TE for each shape will likely increase the proportion of LTE measurements. At high b LTE provides additional information and it can always be performed with a smaller TE than PTE or STE with the same b-value.

#### 4. OPTIMAL EXPERIMENT DESIGN IN MULTIDIMENSIONAL DMRI

---

# CHAPTER 5

---

## Tissue statistics from histological images

In this chapter, we statistically describe histologically stained white matter brain sections to subsequently inform and validate diffusion MRI techniques. For the first time, we characterise volume fraction distributions of three of the main structures in deep subcortical white matter (axons, astrocytes, and myelinated axons) in a representative cohort of an ageing population for which well-characterized neuropathology data is available. We analysed a set of samples from 90 subjects of the Cognitive Function and Ageing Study (CFAS), stratified into three groups of 30 subjects each, in relation to the presence of age-associated deep subcortical lesions. This provides volume fraction distributions in different scenarios relevant to brain diffusion MRI in dementia. We also assess statistically significant differences found between these groups. In addition, we introduced a framework to quantify volume fraction distributions from 2D immunohistochemistry images, which is validated against *in silico* simulations. Since a trade-off between precision and resolution emerged, we also performed an assessment of the optimal scale for computing such distributions.

### 5.1 Histology to inform diffusion MRI

Brain tissue microstructural damage can result from neurodegenerative diseases such as amyotrophic lateral sclerosis, Parkinson's disease, and Alzheimer's disease [2, 140, 141]. These conditions produce gradual deterioration or even death of neurons with concomitant alterations in brain structure and function. Devising imaging techniques capable

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

of characterising brain tissue microstructure *in vivo* is topical within neuroimaging. Key information about brain microstructure is provided by the volumetric densities of the different WM structures [142]. This knowledge might be valuable not only for research but also for its potential to help in developing early stage diagnosis of neurodegenerative diseases through dMRI techniques. The aim of this chapter is to characterise the local volume fraction distribution of axons, astrocytes, and myelinated axons in deep white matter for different populations. These are important stereological parameters, but their distribution has not been previously identified.

Age-associated cerebral white matter lesions can be sub-classified into those within deep white matter (DWM) of the *centrum semiovale* (deep subcortical lesion, DSCL) and those close to the angles of the lateral ventricles (periventricular lesion, PVL). Each has its own clinical relevance [143], but both are thought to be the consequence of small vessel-related vascular pathology such as vascular dementia. We analyse DSCLs, which are associated with loss of myelin components [144] and astrogliosis [145, 146]. To this purpose, various subjects belonging to groups that represent healthy and diseased conditions were imaged. We analyse immunohistochemically stained sections of three populations of DWM samples: Control (no DSCLs were present in the subject), Lesion (the sample presented DSCLs), and Normal Appearing White Matter (NAWM, the subject presented DSCLs but not in the sampled tissue).

Tens of thousands of structures such as axons, coexist in  $1\text{mm}^3$  of brain tissue [147]. Their arrangement varies between different subjects and also with the presence of disease. The information obtained from histological analysis has the potential to help in the description and understanding of healthy tissue, and also in a diverse range of conditions including multiple sclerosis [148, 149], schizophrenia [150], and Alzheimer’s disease [151]. Volume fraction maps of the main white matter structures can further inform and validate various MRI techniques. Prior distributions on the microstructural parameters of biophysical models can be generated from this kind of information.

MRI has become a clinical standard to diagnose brain diseases among other conditions in several body organs [152]. It has a spatial resolution considerably lower than histology. While MRI voxels are in the order of the millimetres, light microscopy can resolve structures smaller than a micron. While microscopy can discern individual structures, MRI can only detect the aggregate signal of the distribution of components within a voxel. However, MRI has the advantage of being a non-invasive imaging technique that can be used *in vivo*. As discussed in Chapter 3 an imaging modality that has gained popularity is dMRI. Within it, a number of biophysical tissue models [13, 43, 47, 54, 103] have been proposed, aiming to describe degeneration processes with high sensitivity and specificity.

As in any other physical problem involving a model, the accuracy of the results relies on how representative the model is for the phenomenon under study (see recent review [39]). The validation of dMRI biophysical models is generally hindered by the complexity and unavailability of the ground truth. Some of the prominent dMRI biophysical models make unrealistic assumptions and hence renders the results of these models dubious [75]. In addition, in absence of additional information, the precise estimation of the model parameters requires a huge amount of measurements. This is where the characterisation of volume fraction ( $V_V$ ) distributions, or more generally information derived from histology, can play a key role. This information has the potential to improve the performance of existing tissue models and help in the validation of new ones. For example, Clayden *et al.* [153] showed that by introducing structured prior information on model parameters, the accuracy in the estimation is improved. Similar work was done by Mozumder *et al.* [154] using MRI-derived prior distributions on NODDIDA parameters to solve the degeneracy in the estimation. Furthermore, Howard *et al.* [155] proposed to jointly analyse dMRI and histology data. The interpretation of parameters from several existing dMRI techniques such as DTI or biophysical models has been previously validated using histological sections (*cf.* [45, 49, 74, 156–158]). Additionally, combined analyses of histology and dMRI have been performed to further understand the development of certain diseases and the healthy brain [159–162]. Information from histology can also help developing realistic *in silico* biomimetic phantoms of brain tissue [163, 164]. Phantoms provide controlled ground truth that can test different dMRI acquisition schemes and post-processing methods.

Local volume fractions depend on the scale of the windows of observation. Previous works have only considered the global average  $V_V$  of white matter structures for the whole brain or over complete regions [157, 158, 165]. There is little information on which scale should be considered for computing local volumetric density maps. As in other imaging fields, there is a trade-off here between precision and resolution [166, 167]. The choice of a small scale can lead to imprecise estimates due to the comparable size of the structures and the averaging window. Larger scales yield stable density estimates, but at the price of losing microstructural detail and hence be uninformative. To define a convenient scale of analysis, we computed the standard error in volume fraction estimates for windows of observation of various scales, together with the significant differences found between adjacent windows. In order to characterise different populations, we required histology data from a large cohort of subjects. The best option for this was immunohistochemistry. However, this modality produces slices with non-negligible thickness in comparison with the structures of interest. Thus, to recover the volume fraction from area fraction measures, we had to adapt and develop new stereological methods. These methods are an interesting

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

additional contribution in themselves.

### 5.1.1 Methodological overview

This chapter addresses first the challenge of automatic segmentation of histological colour images, which is achieved with a colour deconvolution step followed by local thresholding. Then, the high-resolution binary masks are converted into local area fraction ( $A_A$ ) maps, after analysing the optimal scale for computing such values. Then, a method is developed for automatic computation of the  $V_V$  intra-subject distributions from  $A_A$  considering our thin histology sections. These steps are performed to analyse samples from different populations of deep white matter with varying degrees of age-related lesions.

## 5.2 Image acquisition

### 5.2.1 Tissue sample selection

The tissue samples for this work came from the Cognitive Function and Ageing Study (CFAS) neuropathology cohort [168, 169]. Brains were removed with the consent of the next of kin and with multicentre research ethics committee approval, according to standard CFAS protocols [170]. Brains were removed within 60 hours of death, one cerebral hemisphere was fixed in buffered formaldehyde and sliced into 10 mm thick coronal slices. These slices were: 1) immediately anterior to the temporal stem (anterior), 2) at the level of the pulvinar (middle), and 3) at the posterior most limit of the occipital horn of the lateral ventricle (posterior). For grouping the samples according to the presence of age-associated white matter lesions, these slices were scanned using  $T_1$  and  $T_2$  weighted MRI. The MR images were rated by three experienced observers (blind to clinical status) and given a score for DSCLs using a modified Scheltens' scale [171]. Following this scoring, the coronal slices were stored in formalin until required for this study (at least four weeks). From every subject one block of approximately  $20mm \times 20mm \times 10mm$  was sampled from one of the slices. Blocks were allocated in three groups: Control, NAWM, and Lesion. Control blocks were taken from cases where all three levels were scored as 0 on this scale or where only one slice had a score of a maximum of 1. Lesion blocks were taken from regions with a Scheltens' score of 4 or greater. NAWM blocks were taken from lesion free regions of deep white matter in which a DSCL of score 3 or greater was present elsewhere. For additional information about the sample classification and MR imaging details see [170]

To decide the total number of samples for the study, we performed a pilot study using

five samples for each group. We required that the standard error of the mean  $V_V$  for every group needed to be below 0.5% for all structures. This resulted in the need of at least 25 samples from each group. To guarantee our requirement, we decided to run the complete experiment with 30 samples per group. Table 5.1 presents the main information of the selected patient cohort. Additionally, a baseline demographic analysis was performed to assess significant differences in the position of the samples or the sex of the subjects between the groups. No statistically significant differences were found using a Pearson’s  $\chi^2$ -test.

Table 5.1: Patient cohort details: Number of samples per group (N), age of death (mean and standard deviation), sex (M=Male, F=Female), and level (A=Anterior, M=Middle, P=Posterior).

Tissue	N	Age [y-o]	Sex	Level
Control	30	85 $\pm$ 8	13M-17F	9A-17M-4P
NAWM	30	86 $\pm$ 6	14M-16F	10A-16M-4P
Lesion	30	87 $\pm$ 7	12M-18F	12A-14M-4P

### 5.2.2 Histology acquisition

The formalin-fixed blocks of tissue were processed to paraffin and embedded in paraffin wax using conventional protocols [170]. For each block, three sections of 5  $\mu\text{m}$  thickness were cut for immunohistochemistry (in plane dimensions of the samples were around  $20\text{mm} \times 20\text{mm}$ , see Fig. 5.1). The sections were collected onto charged slides and underwent Ag retrieval with Access Revelation RTU (A. Menarini Diagnostics Ltd, Winnersh, UK) in a pressure cooker. Sections were immunostained for phosphorylated neurofilament (SMI31, an axonal marker), glial fibrillary acidic protein (GFAP, an astrocyte marker), and proteolipid protein (PLP, a myelin marker) using an IntelliPATH FLX system (A. Menarini Diagnostics Ltd, Winnersh, UK). These immunostaining markers were chosen due to being the best option for analysing *ex vivo* samples of these structures [172–174]. Immunohistochemistry was performed using a standard ABC method, visualised with diaminobenzidine tetra-chloride (DAB), and the sections counterstained with haematoxylin. Prepared sections were scanned and digitised at 40X magnification using a Nanozoomer XR (Hamamatsu, Photonics Ltd., Hertfordshire, UK). The final resolution of the images was of 0.23  $\mu\text{m}/\text{pix}$ .

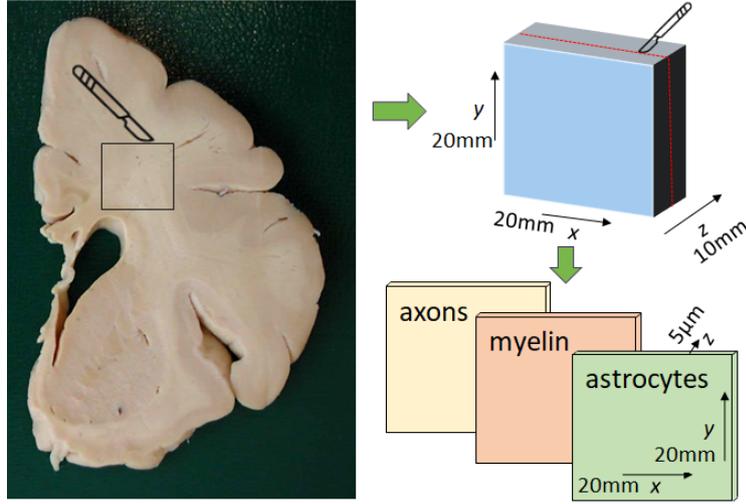


Figure 5.1: Diagram representing the cuts of the blocks and slices from the coronal sections. Each block belonged to a different subject.

### 5.3 Image analysis framework

#### 5.3.1 Histology segmentation

The area fraction ( $A_A$ ) of a certain structure in a region of an image is defined as the ratio of the area occupied by the structure of interest against the total area of the region. We computed the  $A_A$  maps using the area method [175]. To this effect, two segmentation methods were compared regarding their influence on the final results. The first step of the segmentations was to apply colour deconvolution [176] to change the representation of the RGB images to DAB, hematoxylin, and background channels. Only the DAB intensities were kept, leaving a single-channel image. The second step was local thresholding. The first segmentation was based on Otsu’s method [177], while the second one used k-means clustering [178] with two means. The DAB-channel images were divided into patches of 3001 by 3001 pixels to capture a sufficient amount of structures of interest but at the same time being small enough to have a constant illumination across the patch. The optimal thresholds were computed for each patch with the corresponding methods. These patch-wise thresholds were assigned to a cubic spline grid of control points centred at the middle of the corresponding patches. A smooth image with the original resolution was generated by this cubic spline interpolation, representing pixel-wise thresholds. Finally, they were applied to the corresponding pixels of the DAB-channel image resulting in the segmented binary image.

To assess the accuracy of the segmentation algorithms, we compared their performance against two experienced observers. For each preparation (axons, astrocytes and myelin), 200 random pixels belonging to two ROIs of two samples were assessed. These were segmented first by the automatic algorithms, and twice by the two experts on separate sessions at least 4 weeks apart. Intra-rater and inter-rater reliability were evaluated for both experts. These are defined as agreement between repeated segmentations performed by the same person and between segmentations performed by different individuals, respectively. The agreement between the manual and the automatic segmentations was also quantified. All agreement scores were computed using Cohen’s  $\kappa$ -statistic [179], the variation of information [180], and Rand’s index [181], between each pair of segmentations. Fig. 5.2 shows examples of the raw colour images for the three structures, together with the extracted DAB channel after the colour deconvolution step, and the final segmentation on top of the original image.

#### 5.3.2 Scale dependency of the area fraction

Brain tissue has a heterogeneous  $V_V$  spatial distribution. Our goal is to characterise  $V_V$  local variations (*i.e.* intra-subject) for a population. Hence, we need not only accurate estimates but also to capture the changes across a sample. To define an appropriate scale for analysis, we investigated the trade-off between resolution and precision. First, we computed the standard error of the  $A_A$  values. Second, we quantified our ability to find statistically significant differences between neighbouring regions. Both analyses were performed for five scales ranging from 118  $\mu\text{m}$  to 1884  $\mu\text{m}$  (512 pixels-8192 pixels). To this purpose, images were divided in squared windows of corresponding width  $s$  (Fig. 5.3). White matter regions of interest (ROIs) were drawn in each sample to avoid any bias in the results due to contributions from grey matter or artefacts in the images. Only windows that fitted entirely inside the ROI were considered for the analysis.

To compute the  $A_A$  local variance at scale  $s$ , we subdivided each corresponding window into four equal sub-windows of width  $s/2$  (see Fig. 5.3a). If the expected  $A_A$  is homogeneous inside the window, the standard error of the local  $A_A$  can be estimated by:

$$\sigma_i(s) = \sqrt{\frac{1}{12} \sum_{k=1}^4 (a_{i,k} - a_i(s))^2} \text{ for } i = 1 \dots N, \quad (5.1)$$

where  $i$  indexes the window (width  $s$ ) in the image,  $N$  is the total number of windows,  $k$  indicates the position of the sub-window,  $a_{i,k}$  indicates the area fraction measurement of a sub-window, and  $a_i(s)$  is the mean of the four measurements inside a window, *i.e.*

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

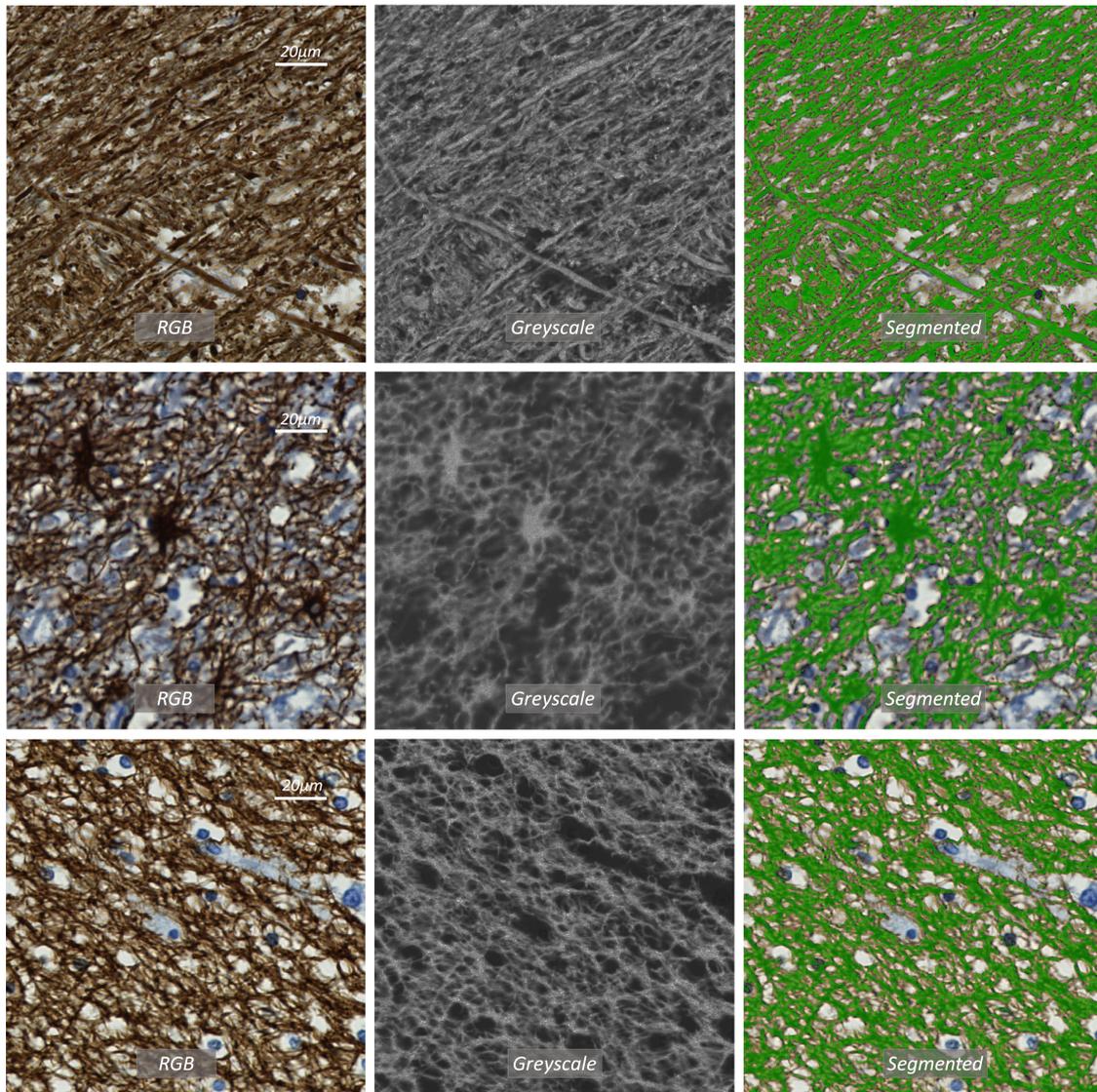


Figure 5.2: Histology raw RGB images of axons (top), astrocytes (middle), and myelinated axons (bottom) with their corresponding greyscale and segmented images for comparison.

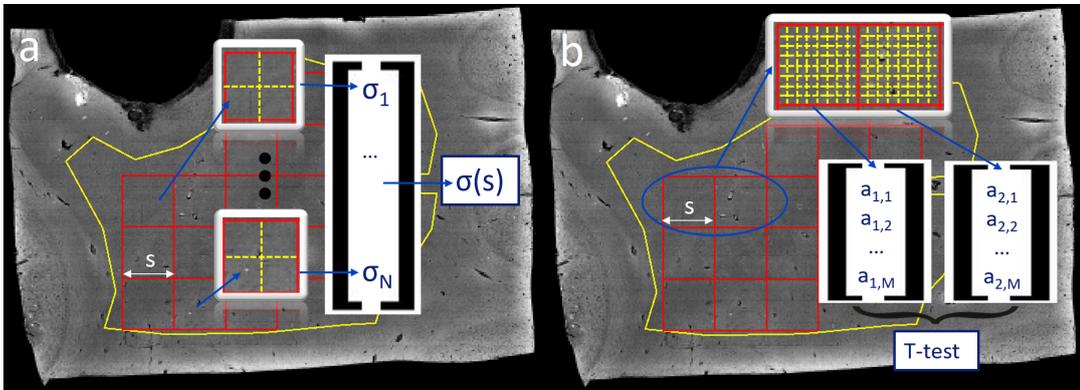


Figure 5.3: Schematic of the window arrangement for the scale dependence analysis for a scale  $s$ . In the computation of the standard error (a), the ROI is divided into square windows of width  $s$ . These are further subdivided into four sub-windows from which a  $\sigma_i(s)$  is computed. Then, a pooled analysis with all  $\sigma_i(s)$  is made and  $\sigma(s)$  is obtained. In the statistical comparisons (b), the ROI is again divided into square windows of width  $s$ . Here the sub-window size is kept fix (not the number of them) to the minimum scale that guarantees independence between  $A_A$  values. Then, statistical tests are made between all possible pairs of adjacent windows to detect significant differences between their mean  $A_A$  values.

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

the area fraction measurement of the window. The global standard error for each scale is estimated by the pooled analysis of the local standard error of all windows by:

$$\sigma(s) = \sqrt{\frac{1}{N} \sum_{i=1}^N \sigma_i^2(s)}. \quad (5.2)$$

Statistical hypothesis testing was used to assess the degree of differentiability between neighbouring regions. The null hypothesis ( $H_0$ ) was formulated as a pair of neighbour windows having the same  $A_A$ . We observed the percentage of rejections of  $H_0$  for each scale over the total number of comparisons. This was tested for every pair of possible neighbours. For all scales, each window was divided into sub-windows that were large enough to guarantee the independence of their  $A_A$  values (*i.e.* much larger than the typical size of the structures within it). Sub-windows had fixed sizes of  $29.44 \mu\text{m} \times 29.44 \mu\text{m}$  ( $128 \times 128$  pixels) in axons and myelinated axons, and  $58.88 \mu\text{m} \times 58.88 \mu\text{m}$  ( $256 \times 256$  pixels) in astrocytes. A two-sample Student's  $t$ -test was performed between all  $A_A$  values belonging to each pair of neighbouring windows to assess the validity of  $H_0$  (see Fig. 5.3b). As the size of the sub-windows was constant, their number inside a window increased with the scale. Thus, incrementing the size of the windows provided more independent samples, which helped to distinguish smaller differences between the mean  $A_A$  of neighbouring windows. An overall rule for the rejection of  $H_0$  was defined by applying a False Discovery Rate (FDR) correction [182] over all comparisons made for each given scale.

### 5.3.3 Volume fraction from histology images

When assuming sections of negligible thickness in comparison to the microstructural elements within (infinitesimally thin planar cuts), it is straightforward to estimate  $V_V$  from  $A_A$ . In this case, it can be shown that the average  $V_V$  is equal to the average  $A_A$  following properties of the expectation of random variables [175]. Given we aimed to analyse relatively large samples from an extensive population, we used digitised whole slide images of paraffin-based histology. While this allows larger tissue expanses to be assessed efficiently in larger cohort sizes, it does not yield 2D infinitesimal planar cuts. The samples used in this study have non-negligible thickness ( $5 \mu\text{m}$ ) compared to the size of axons or astrocytes ( $\sim 1 - 10 \mu\text{m}$ ). They are not infinitesimal planar cuts but thin sections and,  $A_A$  and  $V_V$  are unequal. The digitised images generated from these thin sections show projections of the structures within the slice (cf. Fig. 5.4). Complete 3D information from the projected image cannot be recovered. However, for basic characteristics, such as  $V_V$ , we can find theoretical formulas relating the projected image to the thin section. These depend on the section thickness  $t$  and the structures' size, and are valid under specific assumptions.

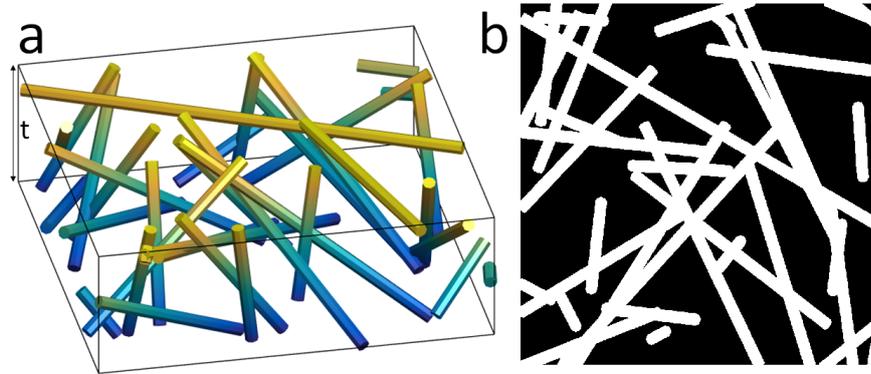


Figure 5.4: Substrate composed of randomly distributed cylinders. A 3D volume is shown on the left (a), and its corresponding 2D projection is on the right (b). In this example  $t/\text{diameter} = 10$ ,  $V_V = 0.043$  and  $A_A = 0.357$ .

The  $A_A$  was computed as the number of positively segmented pixels over the total pixels in the window. Myelin sheaths surround some axons in white matter. In the images stained for myelin, the measured  $A_A$  and hence the computed  $V_V$ , correspond to the volume of myelinated axons (*i.e.* the outer myelin volume plus the inner axon volume).

Based on the work of Miles [183], Weibel *et al.* [184] derived factors for correcting the section thickness effect on volume density estimates. These factors apply to substrates composed of spheres and randomly oriented cylinders (see Eqs. 5.3, 5.4). We modelled axons and myelinated axons as randomly distributed cylinders. For astrocytes, we considered spheres with cylinders protruding from the surface (see Fig. 5.5b), representing the body and processes of the astrocyte, respectively.

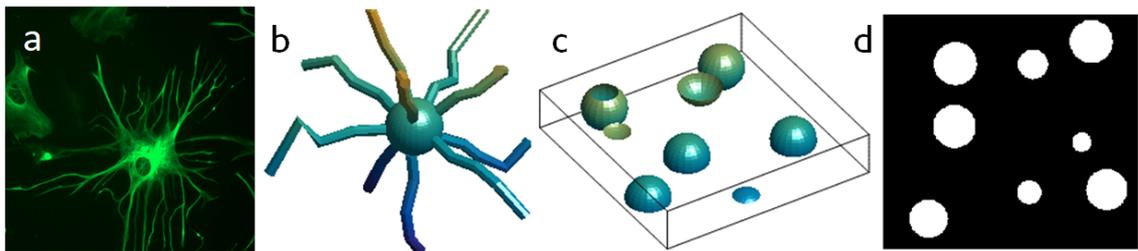


Figure 5.5: Confocal microscopy image of an astrocyte [185] (a) with the sphere+cylinders model (b). Randomly placed truncated spheres with constant radius  $r$ , and  $2r = t$  (c), and its corresponding projected image (d).

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

The correction factors for cylinders and spheres are [184]:

$$K_{V,\text{cyl, const}} = \frac{V_V}{A_A} = \frac{1 - \exp(-\xi\lambda)}{1 - \exp\left(-\xi\left(\lambda + g\left(\lambda + \frac{1}{2}\right)\right)\right)}, \quad (5.3)$$

$$K_{V,\text{sph}} = \frac{V_V}{A_A} = \frac{4E(R^3)}{4E(R^3) + 3tE(R^2)}, \quad (5.4)$$

where,  $\lambda = L/d$  is the ratio between the length  $L$  of the cylinders with the diameter  $d$ ,  $g = t/d$  is the relative section thickness, and  $\xi = V_V/\lambda$ .  $E(R^2)$  and  $E(R^3)$  are the second and third moments of the spheres radius distribution. We considered the cylinders to be long for our scales of analysis [186]. In that case,  $K_{V,\text{cyl, const}}$  was practically independent of  $\lambda$ , which we fixed to 100.  $\xi$  depends on the true  $V_V$ , which is what we were trying to estimate. Hence, we implemented an iterative estimation of  $K_V$ . We computed  $K_V^0$  as a function of  $\xi(A_A)$ , and repeated the process using  $K_V^n = f(K_V^{n-1} \times A_A)$  until  $K_V^n - K_V^{n-1} < \epsilon$ . Simulations showed that  $K_V^n$  converges fast to its true value considering  $\epsilon = 10^{-3}$ .

The main limitation in Eq. 5.3 is that the cylinders are assumed to have an equal radius. This is not realistic if we plan to use them to model axons in brain tissue. To remove this assumption, we derived a more general correction factor (see Appendix A.5) that accounts for a size distribution in the cylinders' diameter. It is proportional to the factor with constant radius:  $K_{V,\text{cyl}} = K_{\text{disp, cyl}} \times K_{V,\text{cyl, const}}$ . The additional factor is

$$K_{\text{disp, cyl}} = \frac{\mu^2 + \sigma^2}{\mu^2}, \quad (5.5)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the diameter distribution. This factor is unitary when there is no dispersion in the cylinders' size.

We computed the volume fraction of astrocytes as the volume fraction of processes and bodies. Each was estimated independently from their corresponding area fractions:

$$V_{V,\text{astro}} = V_{V,\text{proc}} + V_{V,\text{bodies}} = K_{V,\text{cyl}} A_{A,\text{proc}} + K_{V,\text{sph}} A_{A,\text{bodies}}, \quad (5.6)$$

where  $A_{A,\text{proc}}$  and  $A_{A,\text{bodies}}$  are the area fractions from the processes and bodies, respectively. These were computed by performing a segmentation of the astrocytes binary image that allowed us to separate the thin elongated structures (processes) from the larger and rounded ones (bodies).

These correction factors depend on the size of the modelling structures. However, when analysing histological samples, the statistics of axons, myelinated axons, and astrocytes sizes are unknown. To overcome this issue, 100 axons or processes were manually selected from each binary image, and the mean and standard deviation of the structures' radii were

estimated automatically based on structural tensor analysis [187]. For the astrocytes' bodies, their comparable size to the slice thickness introduced an extra challenge. We found a correspondence between the distributions of the projected radii and the actual radii, by assuming the latter one to be a Gaussian distribution (see Appendix A.6).

## 5.4 Results

Our hypothesis was that the previously presented methodology should enable the computation of volumetric statistics from thin histology sections. Below we present a set of experiments validating the main assumptions and we show the results from applying the pipeline to the histology images.

### 5.4.1 Validation of automated segmentation

The agreement scores computed with Cohen's  $\kappa$ -statistic, the variation of information, and Rand's index led to the same conclusions. Hence, we only present in Figure 5.6 a summary of Cohen's  $\kappa$ -statistic for each structure. For axons and astrocytes both methods

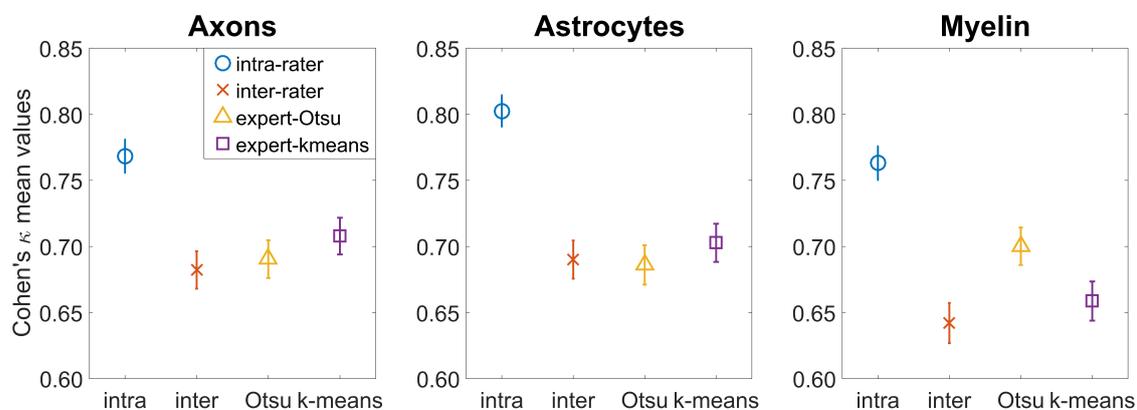


Figure 5.6: Agreement scores between different raters measured by Cohen's  $\kappa$  [179] with their confidence intervals. The first two symbols represent the mean score corresponding to intra-rater and inter-rater reliability. The third and fourth symbols show the mean agreement between the four manual segmentations and the Otsu and k-means automatic segmentations.

had similar agreement scores with the experts. We chose the Otsu-based segmentation as it was the fastest for the computation. For myelinated axons, we chose the Otsu-based segmentation as the score was significantly higher than using k-means. For the three

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

preparations the mean agreement score between the chosen algorithm and the experts was similar to the mean agreement between the two experts.

The segmentation accuracy was also qualitatively assessed in two randomly selected samples. For two ROIs from each sample, RGB images were overlaid with segmentation masks and visually examined by one expert. The expert confirmed that the segmentation produced acceptable results. See supplementary figures in [188] for segmentation examples.

### 5.4.2 Optimal scale selection

We computed the standard error of our  $A_A$  estimates and assessed the significance of the differences between neighbouring regions for different scales. These analyses provided information on the changes in the area fraction we can detect at each scale, and their relation with the resolution-precision trade-off. The statistical tests provided an answer

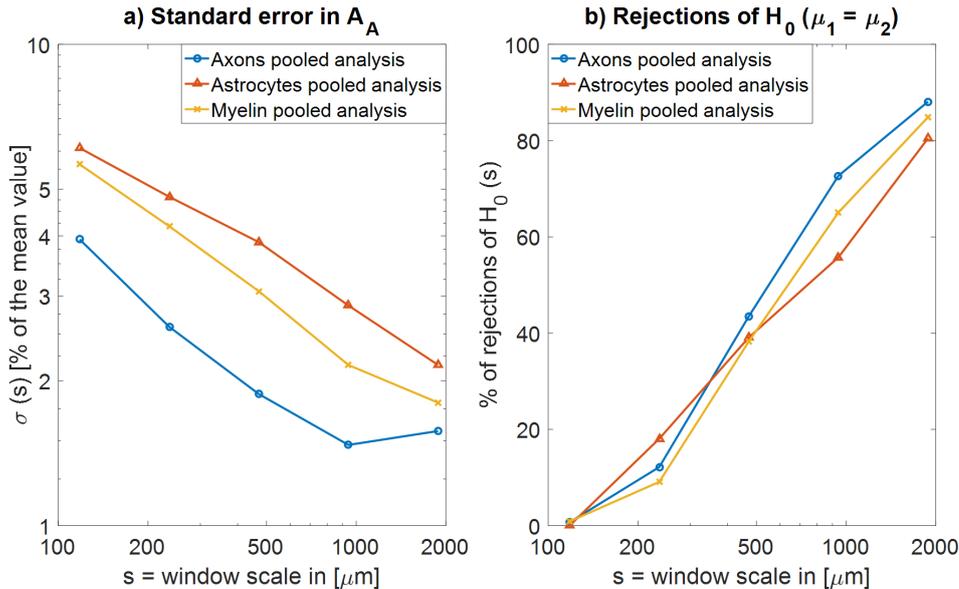


Figure 5.7: Analysis of the resolution-precision trade-off. (a) Standard error of the area fraction for varying scales. This is the pooled result from all the subjects normalised against the mean value across all subjects. (b) Percentage of adjacent windows with statistically significant differences in area fraction. The scales used were [118, 236, 471, 942, 1884]  $\mu\text{m}$ .

to the question: “What proportion of adjacent windows present statistically significant differences at each scale?”. Figure 5.7b shows that for windows whose width is smaller than 236  $\mu\text{m}$ , we cannot distinguish neighbours. This is either because we have insufficient

precision due to the small window size, or because the differences between windows was too small, or both. Figure 5.7a shows that as resolution increases (*i.e.* decreasing the width of the windows), the standard error of the  $A_A$  estimates is further increased. Our criteria was to select as an optimal scale the one which presented statistically significant differences between more than 5% of the possible neighbours. We found that neighbouring patches ( $\sim 10\%$  of them for all structures) presented statistically significant differences between them at a scale of  $236 \mu\text{m}$  and the accuracy of the  $A_A$  estimation was reasonable. Thus,  $236 \mu\text{m} \times 236 \mu\text{m}$  was the scale selected for the computation of  $A_A$  maps for the three structures.

### 5.4.3 Volume fraction distributions

Simulations of *in silico* substrates showed an excellent agreement between the true  $V_V$  and that estimated from the  $A_A$  of the projected image (see Appendix A.7). These comprised cylinders and spheres+cylinders with size distributions similar to those in the structures of interest. The simulated slice thickness was the same as that of the histology images. The robustness of the full orientation dispersion assumption was also tested. We simulated cylinders which were not fully dispersed but with dispersion values typical of DWM [189]. Results show that errors were low even in that scenario.

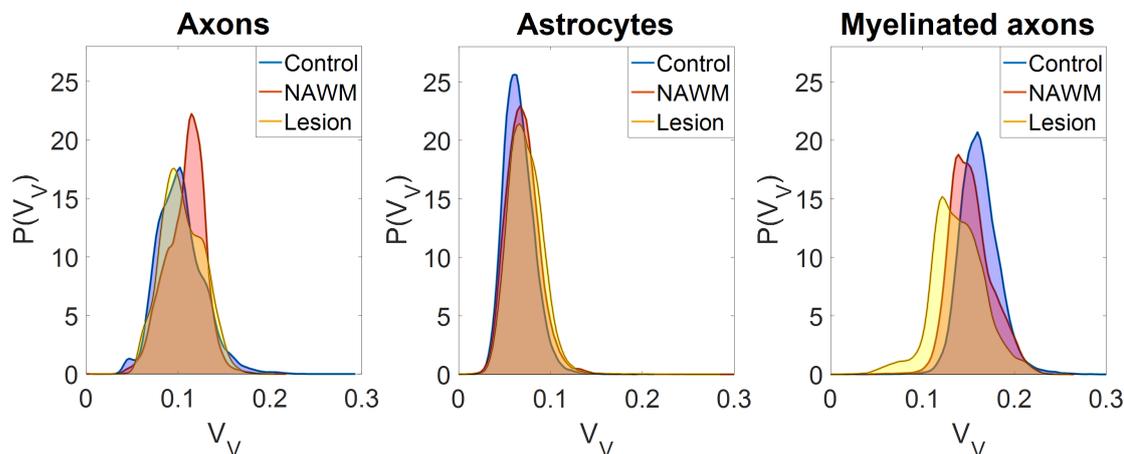


Figure 5.8: Estimated probability distribution functions of the local  $V_V$  of axons, astrocytes, and myelinated axons for the three groups.

The local volume fractions,  $V_V$ , were computed from the corresponding  $A_A$  following the method described in Section 5.3.3. Then, these values were gathered for each group of subjects (control, lesion, and NAWM) to estimate the group-wise distribution of  $V_V$ .

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

Table 5.2: Mean and standard deviation of the group-wise  $V_V$  distributions. For the complete distribution see Fig. 5.8.

<i>Global results (N=30)</i>										
Groups	Axons		Astrocytes		Astrocytes bodies		Astrocytes processes		Myelinated axons	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
Control	0.102	0.026	0.066	0.017	0.007	0.008	0.059	0.014	0.162	0.021
NAWM	0.108	0.021	0.070	0.019	0.007	0.009	0.063	0.017	0.153	0.023
Lesion	0.104	0.023	0.073	0.019	0.008	0.009	0.066	0.016	0.139	0.029

Figure 5.8 shows the estimated probability densities for the three structures. The group mean and the intra-group standard deviation are given in Table 5.2. The group means and variances for axons across different tissue groups are remarkably similar. For astrocytes, however, the mean of the controls is 10% smaller than that of lesions. Finally, myelinated axons display the largest difference in means, especially for the lesion group.

Figure 5.9 shows the group means and the corresponding 95% confidence intervals. The differences in the group means were statistically tested with the Student’s  $t$ -test. The obtained p-values were corrected for the multiple comparisons made using the Bonferroni correction [190]. In axons, no significant differences were observed. In astrocytes, we found a significant increase in the lesion mean regarding control mean ( $p < 0.05$ ). In myelinated axons, we found significant differences between lesion and control ( $p < 0.01$ ).

### 5.5 Discussion

Nervous tissue comprises countless different structures such as neurons and glial cells. An accurate characterisation of their complex arrangement in healthy and lesional tissue can help in the understanding of neurological diseases. The focus of this chapter was to characterise the volume fraction distribution of axons, astrocytes, and myelinated axons in DWM, which has not been previously addressed in the literature. Populations with and without DSCLs were analysed, since this conditions can alter the distribution of the studied WM structures. The volume densities represented in Figure 5.8 and their summary provided in Table 5.2 can be relevant prior information for a more efficient analysis of *in vivo* measurements of new patients.

When comparing the mean of the  $V_V$  distributions we see an agreement with previous works on the differences observed between groups. We observe that DSCLs lead to a sig-

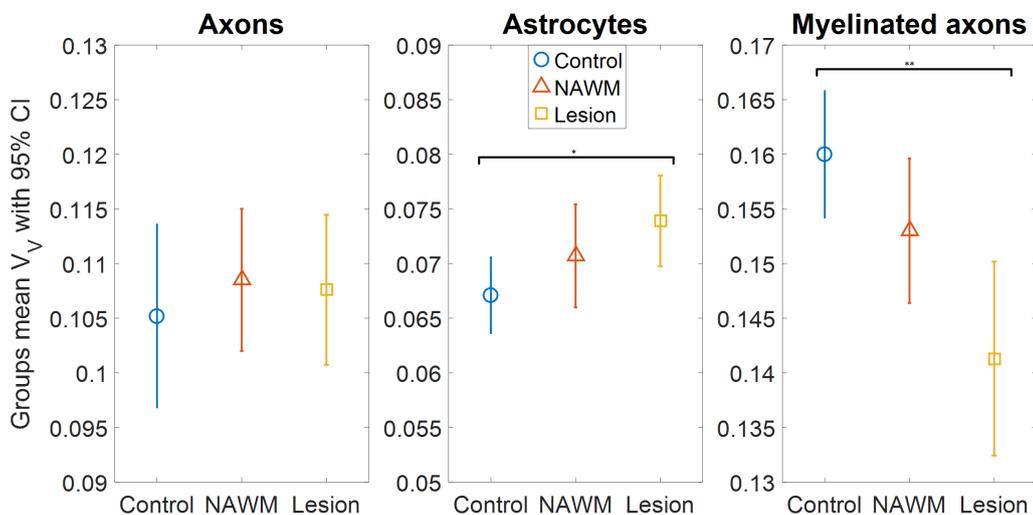


Figure 5.9: Mean  $V_V$  for each group and structure with its corresponding 95% confidence intervals. \* Indicates  $p < 0.05$ , \*\* indicates  $p < 0.01$ .

nificant decrease in the myelinated axons fraction [144]. We also found an increase in the astrocytic fraction as reported in [146]. Unlike them, we detected statistical significant differences between lesion and control groups due to analysing a larger cohort. We did not find significant differences between groups in the mean fraction of axons, which agrees with the work in [191]. These results demonstrate that DSCLs are associated with significant demyelination and astrogliosis, which may lead to axonal dysfunction and impact on central conducting pathways passing through the white matter. While control and normal appearing samples look similar on standard T2-weighted MRI scans, this analysis suggests potential (non-significant) astrogliosis and demyelination in the NAWM which may be interpreted as reflecting lesion progression.

Previous works have also focused on estimating the volume fraction of certain structures in the central nervous system. Histologically derived axonal mean volume fractions have been reported in the grey matter [192], *corpus callosum* [193], and spinal cord [157]. Our results are considerably lower than the ones in these studies. The axonal mean volume fraction we obtained in DWM was around 9 – 12%, while the reported volume fractions in grey matter and *corpus callosum* were 30% approximately. This discrepancy is not so surprising since the axonal density is expected to be quite different in each of these regions. The *corpus callosum* for example, is characterized by a very compact tract of axons while the DWM contains shorter and thinner fibres. Additionally, this study was performed on subjects belonging to an ageing population, where decreased values of ax-

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

onal [194] and myelin [195] content have been reported. This age-related decrease agrees with further tests we performed on a *corpus callosum* sample belonging to the same age cohort (see A.8), observed in both transversal and tangential cuts. The potential relative errors introduced by the mapping from  $A_A$  to  $V_V$  are quite smaller (relative error  $< 10\%$ , see A.7). Segmentation errors were around 7%, 6%, and 10% for axons, astrocytes, and myelin, respectively. We were not able to quantify errors due to incomplete staining, however, images were subject to qualitative quality control check for artifacts, repeating any unsatisfactory image with an adjacent slice from the same sample block. Finally, complex axonal structures (*i.e.* curved axons with varying radius) may have different relations between  $A_A$  and  $V_V$  than the ones modelled, but they are expected to be negligible compared to the errors previously mentioned. In contrast, no significant changes with ageing in the astrocytic fraction were reported in [196], which agrees with our results and previous measurements of astrocyte volume fractions in the *corpus callosum*, the *centrum semiovale* and the corticospinal tract [162]. Thus, with ageing, we seem to get closer proportions of axons and astrocytes.

We were interested in quantifying the local volume fraction across the tissue, as opposed to the global average volume fraction. Considering windows of observation of too large scales would not correctly capture the spatial variations. However, a too small scale would be dominated by statistical fluctuations reducing the precision of the  $V_V$  estimations. The volume fraction cannot be defined for scales comparables to the size of the structures of interest. The appearing precision-resolution trade-off was analysed in a sequence of scales by computing the standard error of the  $A_A$  estimations (precision) and the statistically significant differences observed between adjacent windows (resolution worth). The obtained insight led us to define an optimal scale ( $236 \mu\text{m} \times 236 \mu\text{m}$ ) for computing  $A_A$  distributions and,  $V_V$  distributions.

The chosen histology technique enabled the acquisition of an extensive population of large samples for the current study. This had a considerable impact on the accuracy of the characterisation of the  $V_V$  distribution of the population and in the detection of differences between groups. However, this adds the challenge of computing  $V_V$  from  $A_A$ . Many works that quantify volume fractions from histology images base their calculations on the stereological relationship  $V_V = A_A$ . However, this is only valid if the thickness of the imaged sample is negligible when compared to the size of the structures of interest [175]. Fortunately, there exists some theory [184] tackling this problem when certain conditions are satisfied regarding the structures' geometry. This methodology was extended to allow the computation of the volume fraction of structures that present an unknown size distribution.

A first application of the results obtained in this chapter could be as benchmark for diffusion MRI techniques aiming to estimate the axonal volume fraction, such as those recently reviewed in [39], and newer approaches such as [78, 101]. Our results could also contribute to the definition of new biomarkers [58, 161]. Knowledge of the typical volume fractions of the compartments present in white matter is a must for generating realistic *in silico* phantoms. Most dMRI biophysical models compute compartments' water volume fractions weighted by the  $T_2$  relaxation. Thus, their relationship needs to be appropriately modelled and care must be taken when comparing dMRI measures with the  $V_V$  distributions. Finally, the characterisation of the  $V_V$  distributions in the population could inform dMRI biophysical models as prior distributions to improve the accuracy of the estimated parameters [153].

The framework introduced in this chapter has the potential to aid diagnostic histopathologists and neuropathologists by providing a tool for automatic quantification of the volume fraction of specific WM structures. Digital pathology techniques can enhance pathologists' precision in biomarker assessment and accelerate diagnosis. Hence, there is an increasing interest from this community in automated image analysis technologies that support histopathological assessment of tissue structure (see for example [197–199]).

One of the limitations of the work presented in this chapter is that the immersion of the samples in fixative may produce shrinking in the tissue. However, this effect would affect the estimation of the relative fractions only if differential shrinkage occurs of the different structures and extracellular space. We found no available literature quantifying this latter effect. The chosen histology technique enabled the acquisition of samples from 90 subjects for each preparation. But, one disadvantage of this methodology is this came at the cost of measuring the volume fractions indirectly. Projected statistics were computed and related to their 3D counterpart. Although we tested the accuracy of the volume fraction estimation in synthetic phantoms that slightly violated the models assumptions, the method could still introduce bias in the results if the considered models are not applicable to the real structures. Having thinner sections or high resolution 3D data would be useful to further benchmark this approach. Our results depend highly on the accuracy of the segmentation. However, the agreement between the automatic algorithm and the pathologists equaled the agreement between the two expert pathologists, which is considered satisfactory. Finally, another limitation was that it was not possible to separate the myelin fraction from the axon fraction in myelinated axons. When we look at myelin images, we see practically the same projection as if the cylindrical sheaths were filled. Hence, as it is not possible to compute and subtract the axonal fraction we report the total fraction (*i.e.* myelin plus axon volume).

## 5. TISSUE STATISTICS FROM HISTOLOGICAL IMAGES

---

---

# CHAPTER 6

---

## Conclusions

The increase in sensitivity and specificity in detecting brain microstructural changes is a major motivation for developing biophysical models. The objective of this thesis was to explore ways to make dMRI biophysical modelling more robust while maintaining time and hardware requirements that are feasible in clinical conditions.

Although diffusion MRI is being applied in clinical routine, currently, this is not the case for biophysical tissue models. The impact of the application of some of the improvements proposed in this thesis might help their translation from research into the clinic. This has the potential to increase our specificity to microstructural alterations without compromising sensitivity.

### 6.1 Achievements

In Chapter 3 we have studied measurements *orthogonal* to SDE, *i.e.* functionally independent, and the extra information these provide about tissue. Nonlinear parameter estimation of the SM is not well-posed and can lead to unreliable parameter values. Through the cumulant expansion of the dMRI signal, we showed theoretically that information from DDE is sufficient to avoid degeneracies in the estimation of SM parameters. This is possible without the need of high diffusion weightings, something crucial if our ultimate goal is clinical translation. Furthermore, extending the SM to DDE leads to an increase in the accuracy and precision in the model parameter estimates in the presence of noise.

## 6. CONCLUSIONS

---

In Chapter 4, we have proposed a framework to compute the set of b-tensors that maximises the accuracy and precision in the SM estimation considering a multidimensional dMRI acquisition. This can help to reduce the acquisition time aimed at estimating the SM or any given biophysical model, or reduce the error in the estimated parameters. The framework was based on minimising a metric based on the Cramér-Rao Bounds of the SM kernel parameters. These bounds were computed by an approximation based on the cumulant expansion, which let us discard acquisitions leading to degenerate parameter estimates.

Interestingly, the optimal B-set resulting from the optimisation was reached by combinations of LTE and PTE data with only two diffusion weightings, *i.e.* shells. Furthermore, no PTE data was selected in the lower shell and the confirmation that STE data was not preferred for an optimal acquisition was not a trivial result. *In silico* noise propagation experiments confirmed that the optimal acquisition outperformed other protocols such as LTE+STE or LTE+PTE with equal number of measurements, by an average of 10% in each kernel parameter. Experiments with human data showed agreed, displaying a reduced variability in the kernel estimates with respect to parametric maps obtained from an extended acquisition.

In Chapter 5, we have successfully computed the distribution of the local volume fraction of axons, astrocytes, and myelinated axons in DWM for the first time. These distributions were estimated specifically for an ageing population (CFAS), and are not generalisable to younger populations. Since these distributions depend potentially on DWM pathologies, control, normal appearing, and lesion groups were analysed. The optimal scale for performing this analysis was also assessed.

The obtained distributions of local volume fractions have multiple implications in developing accurate dMRI biophysical tissue models. They can inform models, benchmark their performance, or help to construct realistic *in silico* phantoms. Volume fraction distributions can also be useful for the neuropathological assessment of quantitative changes in healthy and diseased DWM.

We also devised a framework that allows the computation of volume fraction maps from digitised whole slide images of paraffin-based histology. The optimal scale for applying such framework was also analysed as a balance between resolution and precision. Automatic image analyses that can support a histopathological assessment of tissue structure are a valuable tool for diagnostic histopathologists and neuropathologists.

## 6.2 Future work

Future work will focus on the clinical translation of the SM. To achieve it, further validation of model assumptions is crucial. A central assumption in the SM is that there is one set of microstructural parameters inside a voxel (kernel). This sounds reasonable but adding kernels may obscure the inverse problem. For large voxel sizes modelling one kernel may not be correct, meaning that optimal scales where this can be done may be derived [122].

Additionally, I will focus on two alternatives to improve tissue characterisation. The first one, by exploring the benefits of combining dMRI measurements with various Echo Times, in the line of [101]. This has not been explored for multidimensional dMRI, and it has the potential to outperform results shown in Chapter 4 by enabling the separation of intra and extra-axonal  $T_2$ . Simultaneously probing diffusion and relaxometry properties not only improves the estimation of both features but also has a big diagnostic potential since  $T_2$  has been shown to be sensitive to multiple disease processes. However, its clinical translation relies on an optimised acquisition that makes it feasible for clinical acquisition times. The framework we have derived in Chapter 4 will be extended in this direction by accounting a measurement space with b-tensors and multiple TE. Such framework must consider that different b-tensor shapes and sizes have different TE values, and thus, SNR, which will be embedded in the optimisation metric.

The second interesting way of obtaining robust information from the DWIs is not to keep adding orthogonal information, but finding adequate constraints that further simplify the physical system we want to measure. This may imply limiting our measurements to a more specific regime, where, for example only one compartment dominates the signal, like it is proposed to probe the intra-axonal diffusivity in [102] or to validate the sticks assumption in [53].

## 6. CONCLUSIONS

---

## Co-authored published work not included in the thesis

- Rachel Waller, Lynne Baxter, Daniel J. Fillingham, Santiago Coelho, Jose M. Pozo, Meghdoot Mozumder, Alejandro F. Frangi, Paul G. Ince, Julie E. Simpson, and J. Robin Highley. Iba-1-/CD68+ microglia are a prominent feature of age-associated deep subcortical white matter lesions. *PLOS ONE* 14(1): e0210888, 2019.
- Meghdoot Mozumder, Jose M. Pozo, Santiago Coelho, Marina Costantini, Julie Simpson, J. Robin Highley, Paul G. Ince, and Alejandro F. Frangi. Quantitative histomorphometry of capillary microstructure in deep white matter. *NeuroImage: Clinical*, 23:101839, 2019.
- Meghdoot Mozumder, Jose M. Pozo, Santiago Coelho, and Alejandro F. Frangi. Population-based Bayesian regularization for microstructural diffusion MRI with NODDIDA. *Magnetic Resonance in Medicine*, 82:1553-1565, 2019.
- Jose M. Pozo, Santiago Coelho and Alejandro F. Frangi. Tensorial formulation allowing to verify or falsify the microstructural standard model from multidimensional diffusion MRI. *Proceedings of the 27th Annual Meeting of ISMRM*, 2019.



---

# APPENDIX A

---

## Appendix

### A.1 Invariant decomposition of an axially symmetric 4th order tensor

In [88], the invariant decomposition of a fully symmetric 4th-order tensor with axial symmetry is presented. Here we follow an analogous procedure for the decomposition of a 4th order tensor,  $\mathbf{C}$ , with only major and minor symmetries.

Any axially symmetric tensor can be expressed in terms of two ingredients: the symmetry axis vector,  $\boldsymbol{\mu}$ , and the isotropic tensor (Kronecker delta),  $\delta_{ij}$ , which can be grouped by the number of copies of  $\boldsymbol{\mu}$  included. For a 4th order tensor, this gives  $\mathbf{C} = \mathbf{E} + \mathbf{F} + \mathbf{G}$ , where:

$$\begin{aligned} E_{ijkl} &= a \mu_i \mu_j \mu_k \mu_\ell, \\ F_{ijkl} &= b \mu_i \mu_j \delta_{kl} + c \mu_k \mu_\ell \delta_{ij} + d \mu_i \mu_k \delta_{j\ell} + e \mu_j \mu_\ell \delta_{ik} + f \mu_i \mu_\ell \delta_{jk} + g \mu_j \mu_k \delta_{il}, \\ G_{ijkl} &= h \delta_{ij} \delta_{kl} + m \delta_{ik} \delta_{j\ell} + n \delta_{il} \delta_{jk}, \end{aligned} \tag{A.1}$$

are linear combinations of 1 term with 4 times  $\boldsymbol{\mu}$ , 6 terms including it twice, and 3 terms not including it, respectively. Imposing the major and minor symmetries of the tensor  $\mathbf{C}$ , we find the constraints  $a, b = c, d = e = f = g, h$ , and  $m = n$ , resulting in only 5

## A. APPENDIX

---

independent components:

$$\begin{aligned}
 E_{ijkl} &= a \mu_i \mu_j \mu_k \mu_l, \\
 F_{ijkl} &= b(\mu_i \mu_j \delta_{kl} + \mu_k \mu_l \delta_{ij}) + d(\mu_i \mu_k \delta_{jl} + \mu_j \mu_k \delta_{il} + \mu_i \mu_l \delta_{jk} + \mu_j \mu_l \delta_{ik}), \\
 G_{ijkl} &= h \delta_{ij} \delta_{kl} + m(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}),
 \end{aligned} \tag{A.2}$$

Finally, splitting each of the 3 tensors above into its symmetric and complementary parts we get:

$$\mathbf{C} = a\mathbf{P} + (2b + 4d)\mathbf{Q} + (h + 2m)\mathbf{I} + \frac{4}{3}(b - d)\mathbf{R} + \frac{3}{2}(h - m)\mathbf{J}, \tag{A.3}$$

where

$$\begin{aligned}
 P_{ijkl} &= \frac{1}{a} E_{(ijkl)}, & Q_{ijkl} &= \frac{1}{2b + 4d} F_{(ijkl)}, & I_{ijkl} &= \frac{1}{h + 2m} G_{(ijkl)}, \\
 R_{ijkl} &= \frac{3}{4(b - d)} (F_{ijkl} - F_{(ijkl)}), & J_{ijkl} &= \frac{3}{2(h - m)} (G_{ijkl} - G_{(ijkl)}),
 \end{aligned} \tag{A.4}$$

result in the tensors defined in Eq. 3.16.

## A.2 Single solution for NODDIDA with DDE

From Eq. 3.6 and Eq. 3.22, we can select 4 equations generating the joint system:

$$\begin{bmatrix} h_4(1, \kappa) & 2h_2(1, \kappa) & 1 \\ h_4(0, \kappa) & 2h_2(0, \kappa) & 1 \\ 0 & h_2(1, \kappa) - h_2(0, \kappa) & 0 \\ 0 & 2h_2(0, \kappa) & 1 \end{bmatrix} \begin{bmatrix} \gamma \\ \delta \\ \epsilon \end{bmatrix} = \begin{bmatrix} \frac{1}{3}W_{\parallel}\bar{D}^2 + D_{\parallel}^2 \\ \frac{1}{3}W_{\perp}\bar{D}^2 + D_{\perp}^2 \\ \frac{3}{4}\zeta_1 + D_{\perp}D_{\parallel} - D_{\perp}^2 \\ \frac{3}{2}\zeta_2 + D_{\perp}^2 \end{bmatrix}. \quad (\text{A.5})$$

By simple linear combinations we reach

$$\begin{bmatrix} h_4(1, \kappa) & 2h_2(1, \kappa) & 1 \\ h_4(0, \kappa) & 2h_2(0, \kappa) & 1 \\ 0 & 2h_2(1, \kappa) & 1 \\ 0 & 2h_2(0, \kappa) & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3}W_{\parallel}\bar{D}^2 + D_{\parallel}^2 \\ \frac{1}{3}W_{\perp}\bar{D}^2 + D_{\perp}^2 \\ \frac{3}{2}(\zeta_1 + \zeta_2) - D_{\perp}^2 + 2D_{\perp}D_{\parallel} \\ \frac{3}{2}\zeta_2 + D_{\perp}^2 \end{bmatrix} \quad (\text{A.6})$$

$$\begin{bmatrix} h_4(1, \kappa) & 0 & 0 \\ h_4(0, \kappa) & 0 & 0 \\ 0 & 2h_2(1, \kappa) & 1 \\ 0 & 2h_2(0, \kappa) & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3}W_{\parallel}\bar{D}^2 - \frac{3}{2}(\zeta_1 + \zeta_2) + (D_{\parallel} - D_{\perp})^2 \\ \frac{1}{3}W_{\perp}\bar{D}^2 - \frac{3}{2}\zeta_2 \\ \frac{3}{2}(\zeta_1 + \zeta_2) - D_{\perp}^2 + 2D_{\perp}D_{\parallel} \\ \frac{3}{2}\zeta_2 + D_{\perp}^2 \end{bmatrix}. \quad (\text{A.7})$$

Finally, dividing the first and second equations in A.7, the dependency on  $\gamma$  cancels out, resulting in Eq. 3.23, providing a single solution for  $\kappa$ . This is possible since  $\gamma$  is strictly positive, unless there are no sticks ( $f = 0$ ) and the extracellular diffusion is isotropic ( $\Delta_e = 0$ ), or if there is no extra-neurite fraction ( $f = 1$ ) and no axial diffusion inside sticks ( $D_a = 0$ ), and  $h_4(0, \kappa) > 0$  for all finite  $\kappa$ .

### A.3 Tensorial formulation of the SM for a general ODF

Based on [84], Pozo *et al.* [122] recently derived a tensorial formulation of the SM for a general ODF. This provides a direct mapping between the cumulants  $\mathbf{D}$  and  $\mathbf{C}$  and the kernel parameters. We write this mapping below since it is a key component of the work presented in Chapter 4.

Following the discussion in Chapter 3, we know that the set of kernel parameters has a unique mapping to the parametrisation in Eq. 3.5. Thus, to obtain the kernel, we only need to express  $\alpha, \beta, \gamma, \delta, \epsilon$  as a function of the cumulants. According to [122] we can decompose  $D_{ij}$  and  $\hat{C}_{ijkl} = C_{ijkl} + D_{ij}D_{kl}$  into their irreducible independent tensors  $D^{(0)}, \mathbf{D}^{(2)}, \mathbf{S}^{(0)}, \mathbf{S}^{(2)}, \mathbf{S}^{(4)}, A^{(0)}$  and  $\mathbf{A}^{(2)}$

$$\begin{aligned} D_{ij}^{(2)} &= \alpha H_{ij}^{(2)} & D^{(0)} &= \alpha + 3\beta \\ S_{ijkl}^{(4)} = \gamma H_{ijkl}^{(4)} & S_{ij}^{(2)} = \left(\gamma + \frac{7}{3}\delta\right) H_{ij}^{(2)} & S^{(0)} &= \gamma + 5\left(\frac{2}{3}\delta + \epsilon\right) \\ A_{ij}^{(2)} &= \frac{2}{3}\delta H_{ij}^{(2)} & A^{(0)} &= 4\left(\frac{2}{3}\delta + \epsilon\right). \end{aligned} \quad (\text{A.8})$$

where  $H_{ij}^{(2)}$  and  $H_{ijkl}^{(4)}$  are the traceless tensors associated to the ODF averages  $H_{(2)ij}$  and  $H_{(4)ijkl}$ , defined in Eq. 3.20. From the above system of equations, we can see that as long as  $H_{ij}^{(2)}$  is not null (an isotropic ODF is one possibility), it is possible to uniquely recover the kernel and the ODF spherical harmonics up to fourth order from the cumulants:

$$\begin{aligned} \alpha &= \frac{D_{ij}^{(2)}}{H_{ij}^{(2)}}, & \beta &= \frac{1}{3} \left( D^{(0)} - \alpha \right), & \gamma &= S^{(0)} - \frac{5}{4} A^{(0)}, & \delta &= \frac{3}{2} \frac{A_{ij}^{(2)}}{H_{ij}^{(2)}}, & \epsilon &= \frac{1}{4} A^{(0)} - \frac{2}{3} \delta, \\ & & H_{ij}^{(2)} &= \frac{1}{\gamma} \left( S_{ij}^{(2)} - \frac{7}{2} A_{ij}^{(2)} \right), & \text{and} & & H_{ijkl}^{(4)} &= \frac{1}{\gamma} S_{ijkl}^{(4)}. \end{aligned} \quad (\text{A.9})$$

## A.4 Degeneracy in the SM with LTE and LTE+STE

Equation A.8 relates all independent cumulants and the kernel parameters (*cf.* Eq. 3.5). Accessing more cumulants means increasing the number of equations which may enable the inversion of the system. If all cumulants are probed (*e.g.* combining LTE and PTE measurements with two different diffusion weightings), the system is invertible and the solution is unique as long as  $H_{ij}^{(2)} \neq 0$  (see Eq. A.9). However, it is also interesting to analyse what happens when we have an incomplete access to the cumulants, which is the case of LTE or LTE+STE acquisitions.

As it was discussed in Chapter 3, LTE alone is not enough to uniquely recover SM parameters without extremely high diffusion weightings [14]. In an acquisition sensitive up to  $\mathcal{O}(b^2)$ , estimating the kernel parameters is equivalent to solving the resulting system:

$$\begin{aligned} D_{ij}^{(2)} &= \alpha H_{ij}^{(2)} & D^{(0)} &= \alpha + 3\beta \\ S_{ijkl}^{(4)} &= \gamma H_{ijkl}^{(4)} & S_{ij}^{(2)} &= \left(\gamma + \frac{7}{3}\delta\right) H_{ij}^{(2)} & S^{(0)} &= \gamma + 5\left(\frac{2}{3}\delta + \epsilon\right). \end{aligned} \quad (\text{A.10})$$

We may simplify the above system for the kernel parameters and get two independent nonlinear equations relating  $f$ ,  $\Delta_e$  and  $D_e^\perp$ . This means there is a 1-dimensional continuum of solutions, which coincides with results reported in [14]. Although some of these solutions fall outside the range of feasible parameters, there is an infinite amount of physically plausible kernel sets that explain the measurements equally well up to  $\mathcal{O}(b^2)$ .

In the same b-range, combining two shells of LTE and STE data adds an extra equation. STE probes the trace of  $\mathbf{C}$ , which let us compute  $A^{(0)}$ . Thus, the resulting system to solve is:

$$\begin{aligned} D_{ij}^{(2)} &= \alpha H_{ij}^{(2)} & D^{(0)} &= \alpha + 3\beta \\ S_{ijkl}^{(4)} &= \gamma H_{ijkl}^{(4)} & S_{ij}^{(2)} &= \left(\gamma + \frac{7}{3}\delta\right) H_{ij}^{(2)} & S^{(0)} &= \gamma + 5\left(\frac{2}{3}\delta + \epsilon\right) \\ & & & & A^{(0)} &= 4\left(\frac{2}{3}\delta + \epsilon\right). \end{aligned} \quad (\text{A.11})$$

It is possible to rearrange these equations to get a system with 4 nonlinear equations and 4 unknowns. However, the solution is not unique, we may write it solution in terms of  $\delta = (1 - f)\Delta_e D_e^\perp$  and get a quadratic equation. This means that there are *always* two kernel sets that solve the system. Although this may sound tragic at first, interestingly, when we evaluate the quadratic equation in a large number of random feasible combinations of the kernel parameters, we observe that in the majority of the cases, only one solution to the quadratic equation is physically plausible, *i.e.*  $\delta \geq 0$ . This means that although the inverse problem of inverting LTE+STE data, is mathematically degenerate, knowing that  $\delta$  cannot be negative is enough to guarantee a unique solution in most of cases.

## A. APPENDIX

---

Due to the *theoretical* degeneracy that an LTE+STE acquisition presents, our CRB-based metric will penalise such protocols. However, since the spurious solution generally falls outside the range of feasible parameters, any solver with box-constraints can pick the correct solution. Due to this, we added an acquisition with LTE and STE data to test how good these ‘almost not degenerate’ acquisitions behaved, both for *in silico* and *in vivo* experiments.

## A.5 Correction factor for cylinders' size dispersion

The relation between the projected area fraction and the volume fraction for randomly placed cylinders derived in [184] is given by:

$$V_V = K_V(r, t) A_A, \quad (\text{A.12})$$

where the correction factor  $K_V$  depends on the cylinders' radius  $r$  and the section thickness  $t$ . We will omit the dependence on  $t$ , as it will be constant for the remaining of our calculations. The main limitation of the equations provided in [184] is that cylinders are considered to have an equal radius. We are interested in substrates where the radius is given by a probability distribution. Let  $r'$  be the random variable that determines the cylinders' radii,  $A'_A$  and  $V'_V$  the corresponding area and volume fractions of such substrates. We would be interested in:

$$K'_V = \frac{E(V'_V)}{E(A'_A)}, \quad (\text{A.13})$$

where  $K'_V$  is the quotient between the expected fractions, which are computed regarding the radius. Our goal was to derive a correction factor that considers not only the section thickness effect but also the contribution from the dispersion in the radii. The latter one is given by:

$$K_{disp, cyl} = \frac{K'_V}{K_V} = \frac{E(V'_V) E(A_A)}{E(V_V) E(A'_A)}. \quad (\text{A.14})$$

We assumed that the projected area of a cylinder and its volume are approximately proportional to  $r$  and to  $r^2$ , respectively. Hence, the expectation of the area and volume fraction regarding the radius will be approximately proportional to  $E(r)$  and  $E(r^2)$ :

$$K_{disp, cyl} \approx \frac{E(r'^2) E(r)}{E(r^2) E(r')}. \quad (\text{A.15})$$

We consider that  $r'$  has mean  $\mu$  and variance  $\sigma^2$ , while  $r$  is constant and equal to  $\mu$ . Hence,  $E(r) = \mu$  and  $E(r^2) = \mu^2$ . Following this, the desired correction factor is given by:

$$K_{disp, cyl} \approx \frac{\mu^2 + \sigma^2}{\mu^2} = 1 + \frac{\sigma^2}{\mu^2}. \quad (\text{A.16})$$

In the absence of dispersion in the cylinders' radius the size dispersion factor is equal to the unit. As  $\frac{\sigma^2}{\mu^2}$  increases the errors in  $V_V$  estimations due to considering cylinders of equal radius become larger too. We tested the validity of this correction factor with simulations considering feasible ground truth values and results were satisfactory (see Appendix A.6 for *in silico* validation).

## A.6 Match between projected and true size statistics for spheres

We modelled astrocytes' bodies as spheres, where their radius  $R$  was a random variable. Each was considered randomly placed in an infinitely large thin section with thickness  $t$  (see Fig. 5.5 c). However, in the projected 2D images of this section (see Fig. 5.5 d), we observe circumferences with radii not drawn from  $R$  but from  $R_{proj}$ :

$$R_{proj} = f_{proj}(R, Z) = \begin{cases} \sqrt{R^2 - Z^2} & \text{if } -r < z \leq 0, \\ R & \text{if } 0 < z < t, \\ \sqrt{R^2 - (Z - t)^2} & \text{if } t + r \leq z < t, \end{cases} \quad (\text{A.17})$$

where  $Z$  is the coordinate of the centres of the spheres in the dimension of the slice thickness. We assumed that the radius distribution of the astrocytes bodies could be approximated by  $R \sim \mathcal{N}(\mu, \sigma^2)$ , and that  $Z \sim \mathcal{U}[-r, t + r]$ . Then, we can compute the mean and variance of  $R_{proj}$ :

$$E(R_{proj}) = \mu_{proj} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{proj}(r, z) \rho_R(r) \rho_Z(z) dr dz, \quad (\text{A.18})$$

$$Var(R_{proj}) = \sigma_{proj}^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f_{proj}(r, z) - \mu_{proj})^2 \rho_R(r) \rho_Z(z) dr dz, \quad (\text{A.19})$$

where  $\rho_R(r)$  and  $\rho_Z(z)$  are the probability density functions of  $R$  and  $Z$ , respectively. These two integrals provide the means to go from  $\mu_{true}, \sigma_{true} \rightarrow \mu_{proj}, \sigma_{proj}$ . In the real scenario, we can only measure the projected values from segmented astrocyte images. Based on these assumptions, we computed numerically the inverse relation (*i.e.*  $\mu_{proj}, \sigma_{proj} \rightarrow \mu_{true}, \sigma_{true}$ ) for obtaining the true distribution parameters from the projected ones. The Gaussian approximation was feasible because  $\mu_{true}/\sigma_{true} \geq 5$ .

## A.7 *In silico* validation of volume fraction computation

We performed *in silico* experiments to test the validity of the methodology in [184] and the extensions in the methods section. Substrates composed of randomly placed spheres and cylinders (see Fig. A.1) were generated for various feasible  $A_A$  values. From these, we computed their projected 2D binary images and then estimated the volume fractions. The dimensions of the synthetic 3D slices were  $200\ \mu\text{m} \times 200\ \mu\text{m} \times 5\ \mu\text{m}$ .

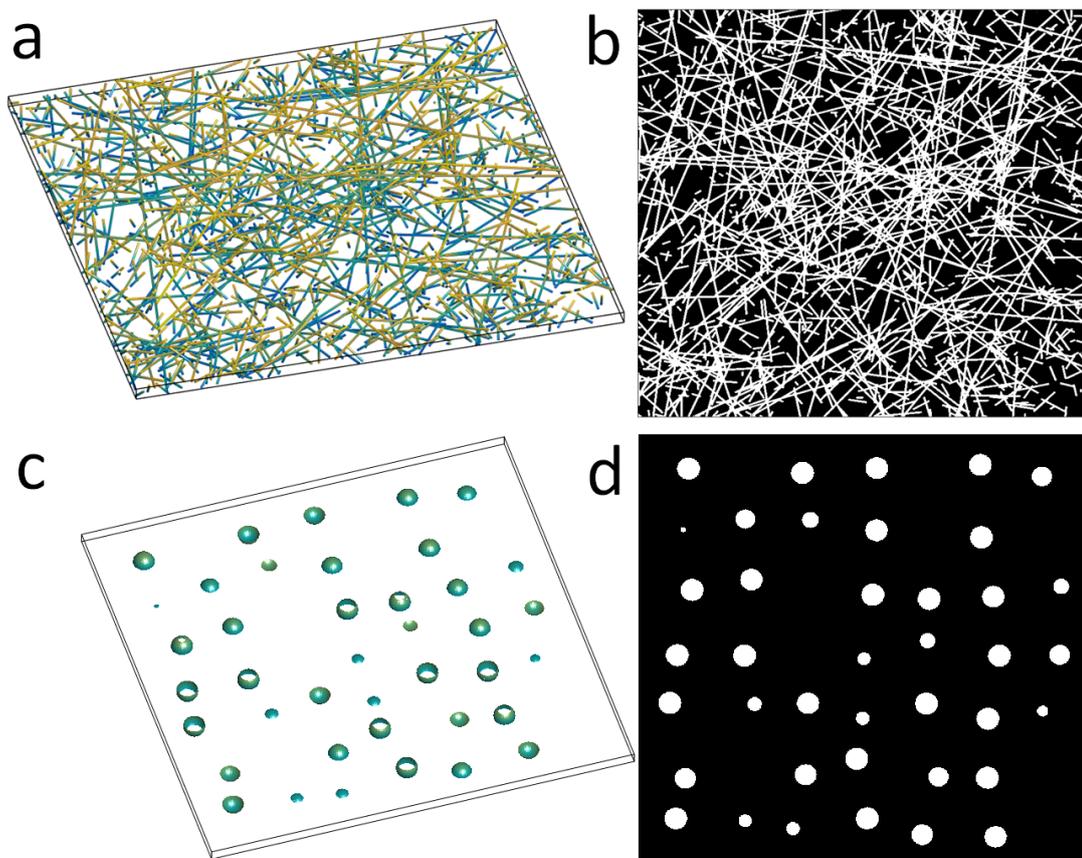


Figure A.1: Simulated 3D substrates of cylinders (a) and spheres (c), and their corresponding 2D projections (b, d).

Figures A.2 and A.3 show the relative error in the estimation of spheres and cylinders volume fractions. The variance of the  $V_{V,sph}$  estimates is larger than  $V_{V,cyl}$  ones because the number of spheres simulated per section was much smaller than the number of cylinders. The feasible number of axons or astrocyte processes (cylinders) inside a window of observation are much larger than the typical number of astrocyte bodies (spheres). Hence, as the number of elements increases, the relative dispersion between the mean projected

## A. APPENDIX

area and the mean volume gets smaller (see scatter plots in Fig. A.4 and A.5). The total volume fraction of astrocytes is mostly due to contribution from processes, hence, the variance in the error of the total volume is similar to the one for cylinders.

To test the robustness of the methodology synthetic experiments were performed. Astrocytes with tortuous processes and branching lengths equal to the bodies' radius were considered. Branching angles were drawn randomly from a uniform distribution in a cone with 20 degrees of aperture. Relative errors were also very small in this case.

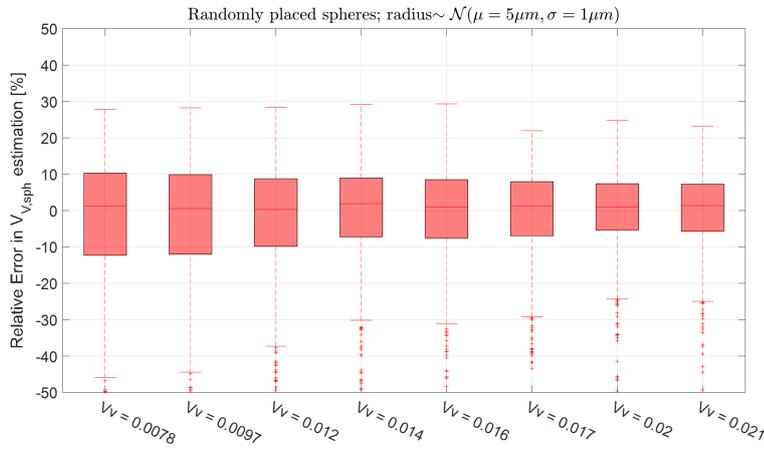


Figure A.2: Relative error in the  $V_V$  estimates for randomly placed spheres. The simulated radius distribution was Gaussian with  $\mu_{sph} = 5 \mu m$  and  $\sigma_{sph} = 1 \mu m$ .

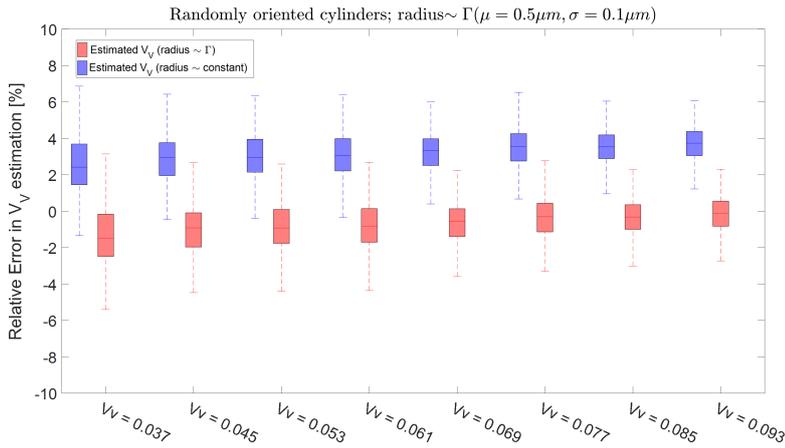


Figure A.3: Relative error in the  $V_V$  estimates for randomly placed cylinders, assuming constant radius (blue boxes) and a radius distribution (red boxes). The simulated radius distribution was Gamma with  $\mu_{cyl} = 0.5 \mu m$  and  $\sigma_{cyl} = 0.1 \mu m$ .

## A.7 *In silico* validation of volume fraction computation

---

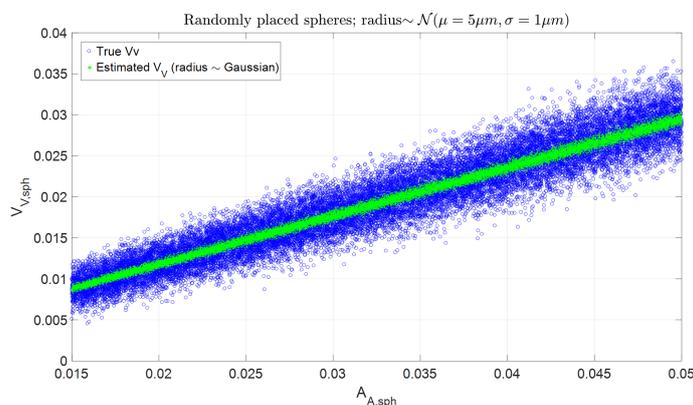


Figure A.4: Scatter plot of  $A_{A,sph}$  vs  $V_{V,sph}$ . Ground truth values are shown in blue circles while estimated values are shown in green dots. The simulated radius distribution was Gaussian with  $\mu_{sph} = 5 \mu\text{m}$  and  $\sigma_{sph} = 1 \mu\text{m}$ .

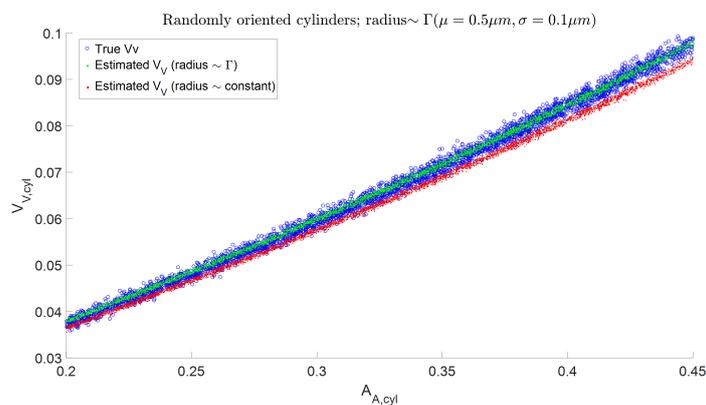


Figure A.5: Scatter plot of  $A_{A,cyl}$  vs  $V_{V,cyl}$ . Ground truth values are shown in blue circles while the estimated ones assuming a radius distribution or constant radius are shown in green and red dots, respectively. The simulated radius distribution was Gamma with  $\mu_{cyl} = 0.5 \mu\text{m}$  and  $\sigma_{cyl} = 0.1 \mu\text{m}$ .

## A.8 Experiments on *corpus callosum* data

There are no previous works that measure volume fractions in human DWM. While this was one of the motivations for our work, it made validating our results very challenging. The lack of literature might be due to the almost isotropic nature of brain tissue orientation in DWM, which makes approaches like those in [157, 193] infeasible as they heavily rely on fibre orientation being perpendicular to the imaging plane. As an assessment of the consistency of our experiments, we performed some tests on a *corpus callosum* sample from the same ageing cohort. We cut two slices from the *corpus callosum* of one control subject. In the first one, axons were perpendicular to the slice plane, while on the second one they were parallel to it (see Fig. A.6). We computed the mean volume fraction ( $V_V$ ) from the segmented image as explained in Section 5.3.3. However, as fibres in the *corpus callosum* cannot be assumed to be isotropically distributed, some adjustments were done to the  $V_V$  estimation.

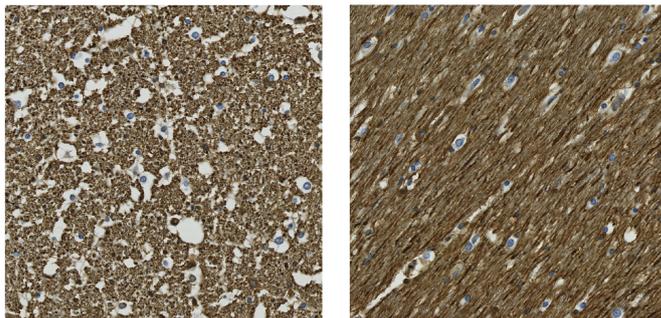


Figure A.6: Axon histology images of *corpus callosum* samples. The samples were cut such that axons were perpendicular (left) and parallel (right) to the imaging plane.

Both slices were imaged following the procedure explained in the section ‘Histology acquisition’. A semi-automatic segmentation was performed on both of them. Regions of interest (ROIs) were drawn in the middle of the *corpus callosum*, and the mean area fraction ( $A_A$ ) was computed on both. For the slice where axons were perpendicular to the imaging plane, we used the stereological relation  $E(V_V) = E(A_A)$ . To obtain the relation between  $V_V$  and  $A_A$  in the slice where axons were parallel to the imaging plane we performed an *in silico* experiment (like those shown in Appendix A.7). Parallel cylinders were randomly placed in a synthetic histological sample of  $5\mu\text{m}$  thickness, and the paired values of  $(V_V, A_A)$  were computed for multiple realisations of varying densities (see Fig. A.7).

The computed  $V_V$  value in the perpendicular slice was 0.226. For the parallel one, we

## A.8 Experiments on *corpus callosum* data

---

obtained  $A_A = 0.699$ , which was mapped to  $V_V = 0.220$  using a fourth order polynomial fit to the synthetic cases (see Fig. A.7). These results are in agreement between themselves and also show, as expected, a significant decrease with respect to the values reported in [193]. We attribute it to ageing as it has been previously reported in [194].

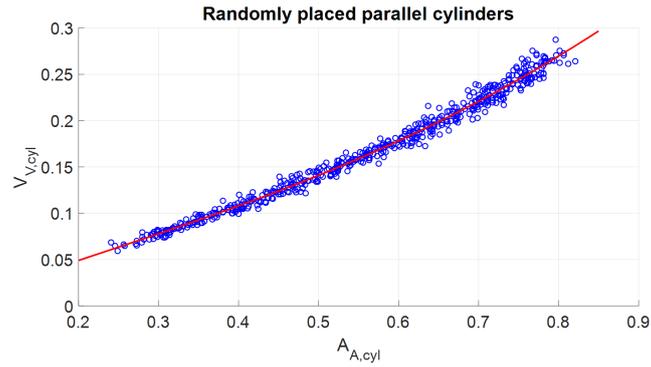


Figure A.7: Scatter plot of  $A_{A,cyl}$  vs  $V_{V,cyl}$  for parallel cylinders that are randomly placed in the slice volume. The best polynomial fit of fourth order is shown in red.



# BIBLIOGRAPHY

- [1] R. W. Williams and K. Herrup. The control of neuron number. *Annual Review of Neuroscience*, 11:423–453, 1988.
- [2] A. J. Stoessl. Neuroimaging in the early diagnosis of neurodegenerative disease. *Translational Neurodegeneration*, 1(1):1–6, 2012.
- [3] C. M. Filley. Neurobiology of white matter disorders. In D. K. Jones, editor, *Diffusion MRI: Theory, Methods and Applications*, chapter 2, pages 19–30. Oxford University Press, Oxford, 2010.
- [4] P. Mansfield and P. K. Grannel. NMR ‘diffraction’ in solids? *Journal of Physics C: Solid State Physics*, 6:422–426, 1973.
- [5] P. C. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242:1973–1974, 1973.
- [6] D. S. Novikov, V. G. Kiselev, and S. N. Jespersen. On modeling. *Magnetic Resonance in Medicine*, 79:3172 – 3193, 2018.
- [7] M. E. Moseley, Y. Cohen, J. Kucharczyk, J. Mintorovitch, H. S. Asgari, M. F. Wendland, J. Tsuruda, and D. Norman. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. *Radiology*, 176:439–445, 1990.
- [8] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology*, 161:401–407, 1986.
- [9] L. C. Maas and P. Mukherjee. Diffusion MRI: Overview and clinical applications in neuroradiology. *Applied Radiology*, 34:44, 2005.

## BIBLIOGRAPHY

---

- [10] S. Mori, B. J. Crain, V. Chacko, and P. C. M. van Zijl. Three-dimensional tracking of axonal projections in brain by magnetic resonance imaging. *Annals of Neurology*, 45:265–269, 1999.
- [11] W. I. Essayed, F. Zhang, P. Unadkat, G. R. Cosgrove, A. J. Golby, and L. J. O’Donnell. White matter tractography for neurosurgical planning: A topography-based review of the current state of the art. *NeuroImage: Clinical*, 15:659 – 672, 2017.
- [12] M. Nilsson, D. van Westen, F. Ståhlberg, P. C. Sundgren, and J. Lätt. The role of tissue microstructure and water exchange in biophysical modelling of diffusion in white matter. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 26:345–370, 2007.
- [13] S. N. Jespersen, C. D. Kroenke, L. Østergaard, J. J. H. Ackerman, and D. A. Yablonskiy. Modeling dendrite density from magnetic resonance diffusion measurements. *NeuroImage*, 34:1473–1486, 2007.
- [14] D. S. Novikov, J. Veraart, I. O. Jelescu, and E. Fieremans. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *NeuroImage*, 174:518 – 538, 2018.
- [15] E. Fieremans, J. H. Jensen, J. A. H. ans Sunghoon Kim, R. I. Grossman, M. Inglese, and D. S. Novikov. Diffusion distinguishes between axonal loss and demyelination in brain white matter. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2012.
- [16] M. D. Budde and J. A. Frank. Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke. *Proceedings of the National Academy of Sciences*, 107(32):14472–14477, 2010.
- [17] A. W. Unterberg, J. Stover, B. Kress, and K. L. Kiening. Edema and brain trauma. *Neuroscience*, 129(4):1019 – 1027, 2004. Brain Water Homeostasis.
- [18] NHS England. Diagnostic imaging dataset annual statistical release 2017/18, 2018. [Online; accessed 29-August-2019].
- [19] World Health Organization. *Neurological Disorders: Public Health Challenges*. Non-serial Publication. World Health Organization, 2006.
- [20] D. K. Jones. *Diffusion MRI*. Oxford University Press, 2010.

- [21] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, 1946.
- [22] H. C. Torrey. Bloch equations with diffusion terms. *Phys. Rev.*, 104:563–565, 1956.
- [23] E. L. Hahn. Spin echoes. *Phys. Rev.*, 80:580–594, 1950.
- [24] V. G. Kiselev. Fundamentals of diffusion MRI physics. *NMR in Biomedicine*, 30:1–18, 2017.
- [25] P. T. Callaghan. Physics of diffusion. In D. K. Jones, editor, *Diffusion MRI: Theory, Methods and Applications*, chapter 4, pages 45–56. Oxford University Press, Oxford, 2010.
- [26] E. O. Stejskal and T. E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42:288–292, 1965.
- [27] D. S. Grebenkov. From the microstructure to diffusion MRI, and back. *RSC Publishing*, 2015.
- [28] N. Shemesh, S. N. Jespersen, D. C. Alexander, Y. Cohen, I. Drobnjac, T. B. Dyrby, J. Finterbusch, M. A. Koch, T. Kuder, F. Laun, M. Lawrenz, H. Lundell, P. P. Mitra, M. Nilsson, E. Özarslan, D. Topgaard, and C.-F. Westin. Conventions and nomenclature for Double Diffusion Encoding NMR and MRI. *Magnetic Resonance in Medicine*, 75:82–87, 2015.
- [29] P. T. Callaghan. *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford Science Publications. Clarendon Press, 1991.
- [30] P. J. Basser, J. Mattiello, and D. LeBihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance*, 103:247–254, 1994.
- [31] D. K. Jones. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study. *Magnetic Resonance in Medicine*, 51:807–815, 2004.
- [32] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.
- [33] D. Jones, M. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(3):515–525, 1999.

## BIBLIOGRAPHY

---

- [34] V. G. Kiselev. The cumulant expansion: an overarching mathematical framework for understanding diffusion NMR. In D. K. Jones, editor, *Diffusion MRI: Theory, Methods and Applications*, chapter 10, pages 152–168. Oxford University Press, Oxford, 2010.
- [35] J. H. Jensen, J. A. Helpert, A. Ramani, H. Lu, and K. Kaczynski. Diffusional Kurtosis Imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53:1432–1440, 2005.
- [36] B. Hansen, N. Shemesh, and S. N. Jespersen. Fast imaging of mean, axial and radial diffusion kurtosis. *NeuroImage*, 142:381–393, 2016.
- [37] D. A. Yablonskiy, G. L. Bretthorst, and J. J. Ackerman. Statistical model for diffusion attenuated mr signal. *Magnetic Resonance in Medicine*, 50(4):664–669, 2003.
- [38] C.-F. Westin, H. Knutsson, O. Pasternak, F. Szczepankiewicz, E. Özarslan, D. van Westen, C. Mattisson, M. Bogren, L. J. O’Donnell, M. Kubicki, D. Topgaard, and M. Nilsson. q-space trajectory imaging for multidimensional diffusion MRI of the human brain. *NeuroImage*, 135:345–362, 2016.
- [39] D. S. Novikov, E. Fieremans, S. N. Jespersen, and V. G. Kiselev. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR in Biomedicine*, 32(4):e3998, 2019.
- [40] P. V. Gelderen, D. Despres, P. C. M. V. Zijl, and C. T. W. Moonen. Evaluation of restricted diffusion in cylinders phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance, Series B*, 103:255–260, 1994.
- [41] G. J. Stanisz, A. Szafer, G. A. Wright, and R. M. Henkelman. An analytical model of restricted diffusion in bovine optic nerve. *Magnetic Resonance in Medicine*, 37:103–111, 1997.
- [42] T. E. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted mr imaging. *Magnetic Resonance in Medicine*, 50(5):1077–1088, 2003.
- [43] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52:965–978, 2004.

- [44] Y. Assaf and P. J. Basser. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *NeuroImage*, 27:48–58, 2005.
- [45] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P. J. Basser. AxCaliber: A method for measuring axon diameter distribution from Diffusion MRI. *Magnetic Resonance in Medicine*, 59:1347–1354, 2008.
- [46] H. Zhang, P. L. Hubbard, G. J. Parker, and D. C. Alexander. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage*, 56:1301–1315, 2011.
- [47] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61:1000–1016, 2012.
- [48] E. Panagiotaki, T. Schneider, B. Siow, M. G. Hall, M. F. Lythgoe, and D. C. Alexander. Compartment models of the diffusion mr signal in brain white matter: a taxonomy and comparison. *Neuroimage*, 59(3):2241–2254, 2012.
- [49] S. N. Jespersen, C. R. Bjarkam, J. R. Nyengaard, M. M. Chakravarty, B. Hansen, T. Vosegaard, L. Østergaard, D. Yablonskiy, N. C. Nielsen, and P. Vestergaard-Poulsen. Neurite density from magnetic resonance diffusion measurements at ultrahigh field: Comparison with light microscopy and electron microscopy. *NeuroImage*, 49:205–216, 2010.
- [50] B. Lampinen, F. Szczepankiewicz, M. Novén, D. van Westen, O. Hansson, E. Englund, J. Mårtensson, C.-F. Westin, and M. Nilsson. Searching for the neurite density with diffusion mri: Challenges for biophysical modeling. *Human Brain Mapping*, 40(8):2529–2545, 2019.
- [51] C. D. Kroenke, J. J. Ackerman, and D. A. Yablonskiy. On the nature of the naa diffusion attenuated mr signal in the central nervous system. *Magnetic Resonance in Medicine*, 52(5):1052–1059, 2004.
- [52] M. Nilsson, S. Lasič, I. Drobnjak, D. Topgaard, and C.-F. Westin. Resolution limit of cylinder diameter estimation by diffusion mri: The impact of gradient waveform and orientation dispersion. *NMR in Biomedicine*, 30(7):e3711, 2017.
- [53] J. Veraart, E. Fieremans, and D. S. Novikov. On the scaling behavior of water diffusion in human brain white matter. *NeuroImage*, 185:379–387, 2019.

## BIBLIOGRAPHY

---

- [54] D. C. Alexander, P. L. Hubbard, M. G. Hall, E. A. Moore, M. Ptito, G. J. Parker, and T. B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage*, 52:1374–1389, 2010.
- [55] B. Dhital, E. Kellner, V. G. Kiselev, and M. Reisert. The absence of restricted water pool in brain white matter. *NeuroImage*, 182:398–406, 2018.
- [56] C. Tax, F. Szczepankiewicz, M. Nilsson, and D. Jones. The dot... wherefore art thou? search for the isotropic restricted diffusion compartment in the brain with spherical tensor encoding and strong gradients. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2018.
- [57] I. O. Jelescu, J. Veraart, E. Fieremans, and D. S. Novikov. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR in Biomedicine*, 29:33–47, 2016.
- [58] I. O. Jelescu, M. Zurek, K. V. Winters, J. Veraart, A. Rajaratnam, N. S. Kim, J. S. Babb, T. M. Shepherd, D. S. Novikov, S. G. Kim, and E. Fieremans. *In vivo* quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. *NeuroImage*, 132:104–114, 2016.
- [59] M. Palombo, A. Ianus, D. Nunes, M. Guerreri, D. C. Alexander, N. Shemesh, and H. Zhang. Sandi: a compartment-based model for non-invasive apparent soma and neurite imaging by diffusion mri, 2019.
- [60] N. G. Gyori, K. K. Seunarine, M. G. Hall, C. A. Clark, D. C. Alexander, and E. Kaden. In-vivo neural soma imaging using b-tensor encoding and deep learning. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2019.
- [61] D. G. Cory, A. N. Garroway, and J. B. Miller. Applications of spin transport as a probe of local geometry. *Polymer Preprints*, 31:149, 1990.
- [62] P. P. Mitra. Multiple wave-vector extensions of the NMR pulsed-field-gradient spin-echo diffusion measurement. *Physical Review B*, 51:15074–15078, 1995.
- [63] M. Lawrenz and J. Finsterbusch. Double-wave-vector diffusion-weighted imaging reveals microscopic diffusion anisotropy in the living human brain. *Magnetic resonance in medicine*, 69(4):1072–1082, 2013.

- 
- [64] S. N. Jespersen, H. Lundell, C. K. Sønderby, and T. B. Dyrby. Orientationally invariant metrics of apparent compartment eccentricity from double pulsed field gradient diffusion experiments. *NMR in Biomedicine*, 26:1647–1662, 2013.
- [65] E. Özarslan, N. Shemesh, and P. J. Basser. A general framework to quantify the effect of restricted diffusion on the NMR signal with applications to double pulsed field gradient NMR experiments. *The Journal of Chemical Physics*, 130:104702/1–104702/9, 2009.
- [66] D. S. Grebenkov. Multiple correlation function approach to study restricted diffusion under arbitrary magnetic field. *Magnetic Resonance Imaging*, 25:559, 2007.
- [67] E. Özarslan. Compartment shape anisotropy (CSA) revealed by double pulsed field gradient mr. *Journal of Magnetic Resonance*, 199:56–67, 2009.
- [68] M. Komlosh, E. Özarslan, M. Lizak, I. Horkayne-Szakaly, R. Freidlin, F. Horkay, and P. Basser. Mapping average axon diameters in porcine spinal cord white matter and rat corpus callosum using d-PFG MRI. *NeuroImage*, 78:210–216, 2013.
- [69] D. Benjamini, M. E. Komlosh, L. A. Holtzclaw, U. Nevo, and P. J. Basser. White matter microstructure from nonparametric axon diameter distribution mapping. *NeuroImage*, 135:333–344, 2016.
- [70] S. N. Jespersen. Equivalence of double and single wave vector diffusion contrast at low diffusion weighting. *NMR in Biomedicine*, 25:813–818, 2011.
- [71] S. Eriksson, S. Lasic, and D. Topgaard. Isotropic diffusion weighting in PGSE NMR by magic-angle spinning of the q-vector. *Journal of Magnetic Resonance*, 226:13 – 18, 2013.
- [72] S. Eriksson, S. Lasič, M. Nilsson, C.-F. Westin, and D. Topgaard. NMR diffusion-encoding with axial symmetry and variable anisotropy: Distinguishing between prolate and oblate microscopic diffusion tensors with unknown orientation distribution. *The Journal of Chemical Physics*, 142(10):104201, 2015.
- [73] D. Topgaard. Multidimensional diffusion mri. *Journal of Magnetic Resonance*, 275:98–113, 2017.
- [74] F. Szczepankiewicz, D. van Westen, E. Englund, C.-F. Westin, F. Ståhlberg, J. Lätt, P. C. Sundgren, and M. Nilsson. The link between diffusion MRI and tumor heterogeneity: Mapping cell eccentricity and density by diffusional variance decomposition (DIVIDE). *NeuroImage*, 142:522–532, 2016.

## BIBLIOGRAPHY

---

- [75] B. Lampinen, F. Szczepankiewicz, J. Mårtensson, D. van Westen, P. C. Sundgren, and M. Nilsson. Neurite density imaging versus imaging of microscopic anisotropy in diffusion mri: A model comparison using spherical tensor encoding. *NeuroImage*, 147:517–531, 2017.
- [76] E. B. Hutchinson, A. V. Avram, M. O. Irfanoglu, C. G. Koay, A. S. Barnett, M. E. Komlosh, E. Özarslan, S. C. Schwerin, S. L. Juliano, and C. Pierpaoli. Analysis of the effects of noise, dwi sampling, and value of assumed parameters in diffusion MRI models. *Magnetic Resonance in Medicine*, 78(5):1767–1780, 2017.
- [77] E. Kaden, N. D. Kelm, R. P. Carson, M. D. Does, and D. C. Alexander. Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139:346–359, 2016.
- [78] M. Reisert, E. Kellner, B. Dhital, J. Hennig, and V. G. Kiselev. Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. *NeuroImage*, 147:964–975, 2017.
- [79] D. Benjamini, M. E. Komlosh, P. J. Basser, and U. Nevo. Nonparametric pore size distribution using d-PFG: Comparison to s-PFG and migration to MRI. *Journal of Magnetic Resonance*, 246:36–45, 2014.
- [80] A. Ianus, I. Drobnjak, and D. C. Alexander. Model-based estimation of microscopic anisotropy using diffusion MRI: a simulation study. *NMR in Biomedicine*, 29:672–685, 2016.
- [81] S. Coelho, L. Beltrachini, J. M. Pozo, and A. F. Frangi. Double Diffusion Encoding vs Single Diffusion Encoding in parameter estimation of biophysical models in Diffusion-Weighted MRI. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2017.
- [82] E. Fieremans, J. Veraart, B. Ades-Aron, F. Szczepankiewicz, M. Nilsson, and D. S. Novikov. Effects of combining linear with spherical tensor encoding on estimating brain microstructural parameters. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2018.
- [83] B. Dhital, M. Reisert, E. Kellner, and V. G. Kiselev. Diffusion weighting with linear and planar encoding solves degeneracy in parameter estimation. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2018.

- [84] S. Coelho, J. M. Pozo, S. N. Jespersen, D. K. Jones, and A. F. Frangi. Resolving degeneracy in diffusion MRI biophysical model parameter estimation using double diffusion encoding. *Magnetic Resonance in Medicine*, 82:395–410, 2019.
- [85] V. G. Kiselev and K. A. Il'yasov. Is the “biexponential diffusion” biexponential? *Magnetic Resonance in Medicine*, 57(3):464–469, 2007.
- [86] S. N. Jespersen, J. L. Olesen, B. Hansen, and N. Shemesh. Diffusion time dependence of microstructural parameters in fixed spinal cord. *NeuroImage*, 182:329–342, 2017.
- [87] M. Abramowitz and I. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1965.
- [88] B. Hansen, A. R. Khan, N. Shemesh, T. E. Lund, R. Sangill, S. F. Eskildsen, L. Østergaard, and S. N. Jespersen. White matter biomarkers from fast protocols using axially symmetric diffusion kurtosis imaging. *NMR in Biomedicine*, 30:1–17, 2017.
- [89] D. Jones, D. Alexander, R. Bowtell, M. Cercignani, F. Dell’Acqua, D. McHugh, K. Miller, M. Palombo, G. Parker, U. Rudrapatna, and C. Tax. Microstructural imaging of the human brain with a “super-scanner”™: 10 key advantages of ultra-strong gradients for diffusion mri. *NeuroImage*, 182:8–38, 2018. Microstructural Imaging.
- [90] C. Westin, F. Szczepankiewicz, O. Pasternak, E. Özarslan, D. Topgaard, H. Knutsson, and M. Nilsson. Measurement tensors in diffusion MRI: Generalizing the concept of diffusion encoding. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 8675, pages 209–216. Springer, 2014.
- [91] Y. Itin and F. W. Hehl. The constitutive tensor of linear elasticity: Its decompositions, cauchy relations, null lagrangians, and wave propagation. *Journal of Mathematical Physics*, 54(6042903), 2013.
- [92] M. Cottaar, F. Szczepankiewicz, M. Bastiani, S. N. Sotiropoulos, M. Nilsson, and S. Jbabdi. Accurate fibre dispersions from varying the b-tensor shape. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2018.
- [93] C. H. Neuman. Spin echo of spins diffusing in a bounded medium. *The Journal of Chemical Physics*, 60:4508–4511, 1974.
- [94] V. Lebedev and D. Laikov. A quadrature formula for the sphere of the 131st algebraic order of accuracy. *Doklady Mathematics*, 59:477–481, 1999.

## BIBLIOGRAPHY

---

- [95] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995.
- [96] C. G. Koay, E. Özarıslan, K. M. Johnson, and M. E. Meyerand. Sparse and optimal acquisition design for diffusion MRI and beyond. *Medical Physics*, 39:2499–2511, 2012.
- [97] S. M. Kay. *Fundamentals of statistical signal processing: estimation theory*. Englewood Cliffs, N.J : PTR Prentice-Hall, 1993.
- [98] J. L. Hintze and R. D. Nelson. Violin plots: A box plot-density trace synergism. *The American Statistician*, 52(2):181 – 184, 1998.
- [99] M. Reiserı, V. G. Kiselev, and B. Dhital. A unique analytical solution of the white matter standard model using linear and planar encodings. *Magnetic Resonance in Medicine*, 81:3819–3825, 2019.
- [100] I. O. Jelescu, J. Veraart, V. Adisetiyo, S. S. Milla, D. S. Novikov, and E. Fieremans. One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI. *NeuroImage*, 107:242–256, 2015.
- [101] J. Veraart, D. S. Novikov, and E. Fieremans. TE dependent Diffusion Imaging (TEdDI) distinguishes between compartmental T2 relaxation times. *NeuroImage*, 182:360–369, 2017.
- [102] B. Dhital, M. Reiserı, E. Kellner, and V. G. Kiselev. Intra-axonal diffusivity in brain white matter. *NeuroImage*, 189:543 – 550, 2019.
- [103] E. Fieremans, J. H. Jensen, and J. A. Helpen. White matter characterization with diffusional kurtosis imaging. *NeuroImage*, 58:177–188, 2011.
- [104] M. Tariq, T. Schneider, D. C. Alexander, C. A. G. Wheeler-Kingshott, and H. Zhang. Bingham-NODDI: Mapping anisotropic orientation dispersion of neurites using diffusion MRI. *NeuroImage*, 133:207 – 223, 2016.
- [105] D. C. Alexander and G. J. Barker. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. *NeuroImage*, 27:357–367, 2005.
- [106] O. Brihuega-Moreno, F. P. Heese, and L. D. Hall. Optimization of diffusion measurements using Cramer-Rao lower bound theory and its application to articular cartilage. *Magnetic Resonance in Medicine*, 50(5):1069–1076, 2003.

- [107] C. R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945.
- [108] H. Cramér. *Methods of Mathematical Statistics*. Princeton, NJ: Princeton University Press, 1946.
- [109] L. Beltrachini, N. von Ellenrieder, and C. H. Muravchik. Error bounds in diffusion tensor estimation using multiple-coil acquisition systems. *Magnetic Resonance Imaging*, 31(8):1372 – 1383, 2013.
- [110] D. H. J. Poot, A. J. den Dekker, E. Achten, M. Verhoye, and J. Sijbers. Optimal experimental design for diffusion kurtosis imaging. *IEEE Transactions on Medical Imaging*, 29(3):819–829, 2010.
- [111] D. C. Alexander. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine*, 60:439–448, 2008.
- [112] B. Lampinen, F. Szczepankiewicz, D. van Westen, E. Englund, P. Sundgren, J. Lätt, F. Ståhlberg, and M. Nilsson. Optimal experimental design for filter exchange imaging: Apparent exchange rate measurements in the healthy brain and in intracranial tumors. *Magnetic Resonance in Medicine*, 77:1104–1114, 2017.
- [113] C. Eichner, S. F. Cauley, J. Cohen-Adad, H. E. Müller, R. Turner, K. Setsompop, and L. L. Wald. Real diffusion-weighted mri enabling true signal averaging and increased diffusion contrast. *NeuroImage*, 122:373 – 384, 2015.
- [114] I. Drobnjak, B. Siow, and D. C. Alexander. Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. *Journal of Magnetic Resonance*, 206:41–51, 2010.
- [115] I. Drobnjak, H. Zhang, M. G. Hall, and D. C. Alexander. The matrix formalism for generalised gradients with time-varying orientation in diffusion NMR. *Journal of Magnetic Resonance*, 210:151–157, 2011.
- [116] J. Sjölund, F. Szczepankiewicz, M. Nilsson, D. Topgaard, C.-F. Westin, and H. Knutsson. Constrained optimization of gradient waveforms for generalized diffusion encoding. *Journal of Magnetic Resonance*, 261:157–168, 2015.
- [117] F. Szczepankiewicz, C.-F. Westin, and M. Nilsson. Maxwell-compensated design of asymmetric gradient waveforms for tensor-valued diffusion encoding. *Magnetic Resonance in Medicine*, 0:1–14, 2019.

## BIBLIOGRAPHY

---

- [118] M. Afzali, C. Chatziantoniou, and D. K. Jones. Comparison of different tensor encoding combinations in microstructural parameter estimation. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2019.
- [119] A. Bates, A. Daducci, and E. Caruyer. Multi-dimensional diffusion MRI sampling scheme: B-tensor design and accurate signal reconstruction. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2019.
- [120] B. Scherrer, A. Schwartzman, M. Taquet, M. Sahin, S. P. Prabhu, and S. K. Warfield. Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND). *Magnetic Resonance in Medicine*, 76(3):963–977, 2016.
- [121] S. Coelho, J. M. Pozo, S. N. Jespersen, and A. F. Frangi. Optimal experimental design for biophysical modelling in multidimensional diffusion MRI. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 3542. Springer, 2019.
- [122] J. M. Pozo, S. Coelho, and A. F. Frangi. Tensorial formulation allowing to verify or falsify the microstructural standard model from multidimensional diffusion MRI. In *Proceedings of the International Society of Magnetic Resonance in Medicine*, volume 3560. Wiley, 2019.
- [123] S. Silvey. *Optimal design: an introduction to the theory for parameter estimation*, volume 1. Chapman and Hall, 1980.
- [124] D. Ackley. *A connectionist machine for genetic hillclimbing*. Kluwer Academic Publishers, 1987.
- [125] I. Zelinka. SOMA — self-organizing migrating algorithm. In *New Optimization Techniques in Engineering*, chapter 7, pages 167–217. Springer Berlin Heidelberg, 2004.
- [126] L. Ingber. Adaptive simulated annealing (ASA): lessons learned. *Control and Cybernetics*, 25:33–54, 1996.
- [127] A. R. Conn, N. I. M. Gould, and P. L. Toint. A globally convergent augmented lagrangian barrier algorithm for optimization with general inequality constraints and simple bounds. *Mathematics of Computation*, 66:261–288, 1997.
- [128] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks*, pages 1942–1945. IEEE, 1995.

- [129] C. G. Koay. A simple scheme for generating nearly uniform distribution of antipodally symmetric points on the unit sphere. *Journal of Computational Science*, 2(4):377 – 381, 2011.
- [130] J. Veraart, E. Fieremans, and D. S. Novikov. Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, 76(5):1582–1593, 2016.
- [131] E. Kellner, B. Dhital, and M. Reisert. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*, 76, 01 2015.
- [132] J.-D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh, and A. Connelly. Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 202:116137, 2019.
- [133] S. M. Smith, M. Jenkinson, M. W Woolrich, C. F Beckmann, T. E J Behrens, H. Johansen-Berg, P. Bannister, M. Luca, I. Drobnjak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, M. Brady, and P. Matthews. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23 Suppl 1:S208–19, 02 2004.
- [134] J. L. Andersson, S. Skare, and J. Ashburner. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2):870–888, 2003.
- [135] P. Brynolfsson, M. Nilsson, L. E. Olsson, C.-F. Westin, and F. Szczepankiewicz. Technical validation of b-tensor encoding shows the need for standardized acquisition protocols. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2019.
- [136] B. Jeurissen and F. Szczepankiewicz. Spherical deconvolution of diffusion mri data with tensor-valued encodings. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2018.
- [137] N. Kunz, A. R. da Silva, and I. O. Jelescu. Intra- and extra-axonal axial diffusivities in the white matter: Which one is faster? *NeuroImage*, 181:314 – 322, 2018.
- [138] S. N. Jespersen, J. L. Olesen, an, and N. Shemesh. Effects of nongaussian diffusion on ”isotropic diffusion” measurements: An ex-vivo microimaging and simulation study. *Journal of Magnetic Resonance*, 300:84–94, 2019.

## BIBLIOGRAPHY

---

- [139] N. G. Gyori, K. K. Seunarine, M. G. Hall, C. A. Clark, D. C. Alexander, and E. Kaden. Time dependence in b-tensor encoding. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2019.
- [140] C. Tur, O. Goodkin, D. R. Altmann, T. M. Jenkins, K. Miszkiet, A. Mirigliani, C. Fini, C. A. M. G. Wheeler-Kingshott, A. J. Thompson, O. Ciccarelli, and A. T. Toosy. Longitudinal evidence for anterograde trans-synaptic degeneration after optic neuritis. *Brain*, 139(3):816–828, 2016.
- [141] R. Pearson, M. Esiri, R. Hiorns, G. Wilcoc, and T. Powell. Anatomical correlates of the distribution of the pathological changes in the neocortex in alzheimer disease. *Proceedings of the National Academy of Sciences of the United States of America*, 82(13):4531–4534, 1985.
- [142] M. Horsfield and D. Jones. Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases - a review. *NMR in Biomedicine*, 15(7-8):570–577, 2002.
- [143] K. H. Park, J.-Y. Lee, D. L. Na, S. Y. Kim, H.-K. Cheong, S. Y. Moon, Y. S. Shim, K. W. Park, B. D. Ku, S. H. Choi, H. Joo, J. S. Lee, S. M. Go, S. H. Kim, S. Kim, K. R. Cha, J. Lee, and S. W. Seo. Different associations of periventricular and deep white matter lesions with cognition, neuropsychiatric symptoms, and daily activities in dementia. *Journal of Geriatric Psychiatry and Neurology*, 24(2):84–90, 2011.
- [144] S. B. Wharton, J. E. Simpson, C. Brayne, and P. G. Ince. Age-associated white matter lesions: The MRC cognitive function and ageing study. *Brain Pathology*, 25(1):35–43, 2015.
- [145] J. E. Simpson, M. S. Fernando, L. Clark, P. G. Ince, F. M. G., Forster, J. T. O. R., Barber, R. N. Kalaria, C. Brayne, P. J. Shaw, C. E. Lewis, S. B. Wharton, and MRC Cognitive Function and Ageing Neuropathology Study Group. White matter lesions in an unselected cohort of the elderly: astrocytic, microglial and oligodendrocyte precursor cell responses. *Neuropathology and Applied Neurobiology*, 33(4):410–419, 2007.
- [146] J. E. Simpson, S. B. Wharton, J. Cooper, C. Gelsthorpe, L. Baxter, G. Forster, P. J. Shaw, G. Savva, F. E. Matthews, C. Brayne, and P. G. Ince. Alterations of the blood-brain barrier in cerebral white matter lesions in the ageing brain. *Neuroscience Letters*, 486(3):246–251, 2010.

- [147] F. A. C. Azevedo, L. R. B. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. L. Ferretti, R. E. P. Leite, J. F. Wilson, R. Lent, and S. Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology*, 513(5):532–541, 2009.
- [148] J. W. Peterson, L. Bö, S. Mörk, A. Chang, and B. D. Trapp. Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Annals of Neurology*, 50(3):389–400, 2001.
- [149] B. D. Trapp, J. Peterson, R. M. Ransohoff, R. Rudick, S. Mörk, and L. Bö. Axonal transection in the lesions of multiple sclerosis. *New England Journal of Medicine*, 338(5):278–285, 1998.
- [150] E. J. Colon. Quantitative cytoarchitectonics of the human cerebral cortex in schizophrenic dementia. *Acta Neuropathologica*, 20:1–10, 1972.
- [151] A. K. Stark, D. P. Pelvig, A.-M. Jorgensen, B. B. Anderser, and B. Pakkenberg. Measuring morphological and cellular changes in Alzheimer’s dementia: a review emphasizing stereology. *Current Alzheimer Research*, 2:449–481, 2005.
- [152] W. Hollingworth, C. J. Todd, M. I. Bell, Q. Arafat, S. Girling, K. R. Karia, and A. K. Dixon. The diagnostic and therapeutic impact of MRI: an observational multi-centre study. *Clinical Radiology*, 55(11):825–831, 2000.
- [153] J. D. Clayden, Z. Nagy, N. Weiskopf, D. C. Alexander, and C. A. Clark. Microstructural parameter estimation in vivo using diffusion MRI and structured prior information. *Magnetic Resonance in Medicine*, 75(4):1787–1796, 2016.
- [154] M. Mozumder, J. M. Pozo, S. Coelho, and A. F. Frangi. Population-based Bayesian regularization for microstructural diffusion MRI with NODDIDA. *Magnetic Resonance in Medicine*, 84:1553–1565, 2019.
- [155] A. F. Howard, J. Mollink, M. Kleinnijenhuis, M. Pallegage-Gamarallage, M. Bastiani, M. Cottaar, K. L. Miller, and S. Jbabdi. Joint modelling of diffusion mri and microscopy. *NeuroImage*, 201:116014, 2019.
- [156] T. L. Chenevert, L. D. Stegman, J. M. G. Taylor, P. L. Robertson, H. S. Greenberg, A. Rehemtulla, and B. D. Ross. Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors. *JNCI: Journal of the National Cancer Institute*, 92(24):2029–2036, 2000.

## BIBLIOGRAPHY

---

- [157] J. Xu, H. Li, K. D. Harkins, X. Jiang, J. Xie, H. Kang, M. D. Does, and J. C. Gore. Mapping mean axon diameter and axonal volume fraction by MRI using temporal diffusion spectroscopy. *NeuroImage*, 103:10–19, 2014.
- [158] F. Sepehrband, K. A. Clark, J. F. Ullmann, N. D. Kurniawan, G. Leanage, D. C. Reutens, and Z. Yang. Brain tissue compartment density estimated using diffusion-weighted MRI yields tissue parameters consistent with histology. *Human Brain Mapping*, 36(9):3687–3702, 2015.
- [159] M. D. Budde and J. A. Frank. Examining brain microstructure using structure tensor analysis of histological sections. *NeuroImage*, 63:1–10, 2012.
- [160] J. Kolasinski, C. J. Stagg, S. A. Chance, G. C. DeLuca, M. M. Esiri, E.-H. Chang, J. A. Palace, J. A. McNab, M. Jenkinson, K. L. Miller, and H. Johansen-Berg. A combined post-mortem magnetic resonance imaging and quantitative histological study of multiple sclerosis pathology. *Brain*, 135:2938–2951, 2012.
- [161] A. R. Khan, A. Chuhutin, O. Wiborg, C. D. Kroenke, J. R. Nyengaard, B. Hansen, and S. N. Jespersen. Biophysical modeling of high field diffusion MRI demonstrates micro-structural aberration in chronic mild stress rat brain. *NeuroImage*, 142:421–430, 2016.
- [162] J. Mollink, M. Kleinnijenhuis, A.-M. van Cappellen van Walsum, S. N. Sotiropoulos, M. Cottaar, C. Mirfin, M. P. Heinrich, M. Jenkinson, M. Pallegage-Gamarallage, O. Ansorge, S. Jbabdi, and K. L. Miller. Evaluating fibre orientation dispersion in white matter: Comparison of diffusion MRI, histology and polarized light imaging. *NeuroImage*, 157:561–574, 2017.
- [163] P. A. Cook, Y. Bai, S. Nedjati-Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker, and D. C. Alexander. Camino: Open-Source Diffusion-MRI Reconstruction and Processing. In *Proceedings of the International Society of Magnetic Resonance in Medicine*. Wiley, 2006.
- [164] L. Beltrachini, Z. A. Taylor, and A. F. Frangi. A parametric finite element solution of the generalised Bloch-Torrey equation for arbitrary domains. *Journal of Magnetic Resonance*, 259:126 – 134, 2015.
- [165] Y. Tang, J. R. Nyengaard, B. Pallenberg, and H. J. G. Gundersen. Age-induced white matter changes in the human brain: A stereological investigation. *Neurobiology of Aging*, 18(6):609–615, 1997.

- [166] T. Chen, P. B. Catrysse, A. E. Gamal, and B. A. Wandell. How small should pixel size be? In *Proceedings of SPIE*, volume 3965, pages 451–459, 2000.
- [167] S. C. Kale, X. J. Chen, and R. M. Henkelman. Trading off SNR and resolution in MR images. *NMR in Biomedicine*, 22(5):488–494, 2009.
- [168] C. Brayne, C. McCracken, and F. E. Matthews. Cohort profile: The medical research council cognitive function and ageing study (CFAS). *International Journal of Epidemiology*, 35:1140–1145, 2006.
- [169] Cognitive Function and Ageing Studies (CFAS) Collaboration. Cognitive Function & Ageing Study. "<http://www.cfas.ac.uk/>", 2017. Accessed: 2017-10-10.
- [170] M. S. Fernando, J. T. O'Brien, R. H. Perry, P. English, G. Forster, W. McMeekin, J. Y. Slade, A. Golkar, F. E. Matthews, R. Barber, R. N. Kalaria, P. G. Ince, and N. G. of MRC CFAS. Comparison of the pathology of cerebral white matter with post-mortem magnetic resonance imaging (MRI) in the elderly brain. *Neuropathology and Applied Neurobiology*, 30(4):385–395, 2004.
- [171] P. Scheltens, F. Barkhof, D. Leys, J. P. Pruvo, J. J. Nauta, P. Vermersch, M. Steinling, and J. Valk. A semiquantitative rating scale for the assessment of signal hyperintensities on magnetic resonance imaging. *Journal of the Neurological Sciences*, 114:7–12, 1993.
- [172] R. Barker, D. Wellington, M. M. Esiri, and S. Love. Assessing white matter ischemic damage in dementia patients by measurement of myelin proteins. *Journal of Cerebral Blood Flow & Metabolism*, 33(7):1050–1057, 2013.
- [173] C. Garwood, L. Ratcliffe, J. Simpson, P. Heath, P. Ince, and S. Wharton. Review: Astrocytes in alzheimer's disease and other age-associated dementias: a supporting player with a central role. *Neuropathology and Applied Neurobiology*, 43(4):281–298, 2016.
- [174] L. A. Sternberger and N. H. Sternberger. Monoclonal antibodies distinguish phosphorylated and nonphosphorylated forms of neurofilaments in situ. *Proceedings of the National Academy of Sciences of the United States of America*, 80(19):6126–6130, 1983.
- [175] J. Ohser and F. Mücklich. *Statistical analysis of microstructures in materials science*. Chichester, UK : Wiley, 2000.

## BIBLIOGRAPHY

---

- [176] A. C. Ruifrok and D. A. Johnston. Quantification of histochemical staining by color deconvolution. *Analytical and Quantitative Cytology and Histology*, 23:291–299, 2001.
- [177] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9:62–66, 1979.
- [178] D. Arthur and S. Vassilvitskii. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.
- [179] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.
- [180] M. Meila. Comparing clusterings. *COLT*, 2002.
- [181] W. M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.
- [182] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.
- [183] R. E. Miles. Estimating aggregate and overall characteristics from thick sections by transmission microscopy. *Journal of Microscopy*, 107:227–233, 1976.
- [184] E. R. Weibel and D. Paumgartner. Integrated stereological and biochemical studies on hepatocytic membranes. ii. Correction of section thickness effect on volume and surface density estimates. *Journal of Cell Biology*, 77:584–597, 1978.
- [185] B. Pascal. *23 week human culture astrocyte stained for GFAP*. Wikimedia Commons, 2012.
- [186] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell*. Garland Science, 4th edition edition, 2003.
- [187] J. Bigun and G. H. Granlund. Optimal orientation detection of linear symmetry. Technical Report Report LiTH-ISY-I-0828, Computer Vision Laboratory, Linköping University, Sweden, 1986.

- [188] S. Coelho, J. M. Pozo, M. Costantini, J. R. Highley, M. Mozumder, J. E. Simpson, P. G. Ince, and A. F. Frangi. Local volume fraction distributions of axons, astrocytes, and myelin in deep subcortical white matter. *NeuroImage*, 179:275–287, 2018.
- [189] Y. S. Chang, J. P. Owen, N. J. Pojman, T. Thieu, P. Bukshpun, M. L. J. Wakahiro, J. I. Berman, T. P. L. Roberts, S. S. Nagarajan, E. H. Sherr, and P. Mukherjee. White matter changes of neurite density and fiber orientation dispersion during human brain maturation. *PLOS ONE*, 10(6):1–23, 2015.
- [190] J. P. Shaffer. Multiple hypothesis testing. *Annual review of psychology*, 46(1):561–584, 1995.
- [191] J. Highley, O. Gebril, J. Simpson, S. Wharton, J. Kirby, J. O’Brien, R. Barber, R. Kalaria, C. Brayne, C. Lewis, P. Shaw, and P. Ince. Axonal preservation in deep subcortical white matter lesions in the ageing brain. *Journal of Aging Science*, 2(1), 2014.
- [192] D. B. Chklovskii, T. Schikorski, and C. F. Stevens. Wiring optimization in cortical circuits. *Neuron*, 34(3):341 – 347, 2002.
- [193] N. Stikov, J. S. W. Campbell, T. Stroh, M. Lavelée, S. Frey, J. Novek, S. Nuara, M. Ho, B. J. Bedell, R. F. Dougherty, I. R. Leppert, M. Boudreau, S. Narayanan, T. Duval, J. Cohen-Adad, P. Picard, A. Gasecka, D. Côté, and G. B. Pike. In vivo histology of the myelin g-ratio with magnetic resonance imaging. *NeuroImage*, 118:397 – 405, 2015.
- [194] D. Calkins. Age-related changes in the visual pathways: Blame it on the axon. *Investigative ophthalmology & visual science*, 54(14):ORSF37–ORSF41, 12 2013.
- [195] A. Peters. The effects of normal aging on myelin and nerve fibers: A review. *Journal of Neurocytology*, 31(8):581–593, 2002.
- [196] J. A. Sloane, W. Hollander, D. L. Rosene, M. B. Moss, T. Kemper, and C. R. Abraham. Astrocytic hypertrophy and altered GFAP degradation with age in subcortical white matter of the rhesus monkey. *Brain Research*, 862(1):1 – 10, 2000.
- [197] R. Paulik, T. Micsik, G. Kiszler, P. Kaszál, J. Székely, N. Paulik, E. Várhalmi, V. Prémusz, T. Krenács, and B. Molnár. An optimized image analysis algorithm for detecting nuclear signals in digital whole slides for histopathology. *Cytometry Part A*, 91(6):595–608, 2017.

## BIBLIOGRAPHY

---

- [198] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang, M. Lai, and E. I.-C. Chang. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*, 18(1):281, 2017.
- [199] J. Barker, A. Hoogi, A. Depeursinge, and D. L. Rubin. Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. *Medical Image Analysis*, 30:60–71, 2016.