UCLouvain
Institute for Information
and Communication Technologies,
Electronics and Applied Mathematics

Brain biomarkers from diffusion MRI: correction of motion artifacts, acceleration of microstructure estimation, and disease classification with transformers

Quentin Dessain

Thesis submitted in partial fulfillment
of the requirements for the degree of
Docteur en sciences de l'ingénieur et technologie

Dissertation committee:

Prof. Benoît Macq (UCLouvain, advisor)
Dr. Laurence Dricot (UCLouvain)
Prof. Christophe De Vleeschouwer (UCLouvain)
Prof. Christophe Craeye (UCLouvain)
Dr. Gaëtan Rensonnet (Boston Consulting Group)
Dr. Maxime Taquet (University of Oxford)

Abstract

Understanding the biological mechanisms underlying brain disorders remains a significant challenge in neuroscience. Traditional diagnostic approaches rely heavily on subjective assessments, leading to variability in diagnosis and treatment outcomes. Diffusion magnetic resonance imaging (dMRI) has emerged as a powerful tool for investigating brain microstructure, offering insights into white matter integrity and connectivity. However, several technical challenges hinder its clinical impact, including motion artifacts, the computational demands of complex tissue modeling and challenges in integrating deep learning with advanced microstructural models for disease diagnosis.

This thesis presents methodological innovations to address these challenges across the dMRI pipeline, from acquisition to analysis. The research is structured around three main contributions. First, a motion correction framework tailored for gradient cycling acquisitions in dMRI is introduced, significantly improving data quality in high b-value acquisitions. Second, two deep learning-based approaches are developed to accelerate the estimation of microstructural parameters in Microstructure Fingerprinting. Third, a Swin Transformer-based deep learning model is applied to multi-shell dMRI data for Alzheimer's disease classification, demonstrating the potential of deep learning in integrating diffusion-based biomarkers for early diagnosis.

Together, these contributions advance the practical utility of dMRI by improving data quality, computational efficiency, and clinical applicability.

Acknowledgments

The journey of completing a PhD is not one I could have done alone, and I am deeply grateful for the support, guidance, and encouragement of all the people who played an important role along the way.

First, I would like to express my gratitude to my supervisor, Prof. Benoît Macq, who played a decisive role in steering me toward a research career. His high-level insights, optimism, trust, and ability to distill complex ideas into valuable perspectives helped shape the trajectory of this thesis. Beyond his technical expertise, his mastery of soft skills and his extensive, world-spanning network have been invaluable assets, creating opportunities that enriched my academic journey. One such opportunity was my stay in Montreal, which would not have been possible without him.

I extend my deepest appreciation to Dr. Gaëtan Rensonnet whose mentorship during my first year of PhD proved foundational. His patient guidance helped me navigate the daunting early stages of research.

I owe a special thanks to Dr. Laurence Dricot, who was my reference at IoNS and a consistent source of support throughout my PhD. Her readiness to assist, her thoughtful advice, and her genuine interest in my progress made an immeasurable difference. Laurence's expertise and mentorship extended far beyond technical aspects, offering me a steady anchor and a source of encouragement during critical moments.

Another fascinating person I encountered during my PhD journey was Prof. Ron Kupers, with whom I had the pleasure of collaborating. His expertise and never-ending enthusiasm added great depth to my research, making our collaborations a highlight of this journey. Beyond our academic discussions, I also enjoyed our broader conversations and the occasional troubleshooting sessions, which created moments that blended work with my personal interest in technology.

To my colleagues and friends who shared the same workspace, Nicolas Delinte, Colin Vandenbulcke, Manon Dausort, and Clément Fuschs, thank you for making the daily work easier and more enjoyable. Your camaraderie and willingness to brainstorm over coffee breaks created a positive

and productive environment.

I also want to recognize both the PiLab and IoNS teams for their support and for providing an intellectually stimulating environment. The vibrant research culture and collaborative spirit of these groups provided me with inspiration and motivation throughout my PhD journey.

A special acknowledgment goes to both Prof. Maurice Ptito and his team, as well as Prof. Pierre Bellec and his team, who warmly welcomed me during my year abroad in Montreal. Splitting my time between their labs was an enriching experience that broadened my scientific horizons and allowed me to explore new perspectives with their guidance and collaboration. Their expertise, openness, and the stimulating discussions made my time in Montreal not only highly productive but also truly enjoyable. I am deeply grateful for their support and the opportunities they provided.

I will always be thankful to my family for their unwavering love, interest, and support throughout this journey. Their encouragement has been a source of strength, making this path more manageable and meaningful.

No acknowledgment would be complete without mentioning Delphine Scheirman, who has been a steady light through the highs and lows of this journey. Your support, patience, and understanding have made even the most challenging moments easier to navigate. I cannot thank you enough for always being there.

To my fellow friends who are also pursuing a PhD, Colin Timmers, Jean-Luc Timmermans, Mathieu Simon, and Quentin Delhaye, thank you for the countless discussions about the joys and struggles of academic life. Sharing this journey with you has been both inspiring and reassuring, making the challenges more manageable and the achievements more rewarding.

I also deeply appreciate my broader circle of friends, who have stood by me throughout this process. Your ability to pull me away from the intensity of work provided me with much-needed balance and perspective. Whether through conversations, shared meals, or unforgettable adventures, you reminded me of the importance of stepping back and embracing the bigger picture.

Finally, I want to express my gratitude to the individuals whose presence may not be captured in this acknowledgment but whose contributions have left a mark on my PhD experience. Every interaction, no matter how brief, played a role in shaping the researcher and individual I am today.

To all of you, thank you for being part of this incredible journey.

Author's publication list

Journal articles

Dessain, Q, Fuchs, C, Macq, B, and Rensonnet, G (2024). Fast multi-compartment microstructure fingerprinting in brain white matter. *Frontiers in Neuroscience,* 18:1400499.

Fuchs, C, Dessain, Q, Delinte, N, Dausort, M, and Macq, B (2024). Sparse blind spherical deconvolution of diffusion weighted MRI. *Frontiers in Neuroscience,* 18:1385975.

Conference proceedings

Dessain, Q, Dricot, L, Delinte, N, Macq, B, de Timary, P, and Kupers, R (2024). Brain functional connectivity predicts emotional awareness. In *OHBM 2024 Annual Meeting*.

Dessain, Q, Simon, M, Dricot, L, and Macq, B (2024). Advanced motion correction in diffusion MRI acquisition with gradient cycling. In *International Symposium on Biomedical Imaging (ISBI 2024)*.

Huyghe, L, Quenon, L, Salman, Y, Colmant, L, Gérard, T, Delhaye, E, Besson, G, Malotaux, V, Boyer, E, Dessain, Q, Dricot, L, Lhommel, R, Ivanoiu, A, Bastin, C and Hanseeuw, B (2024). Association between the ability of discriminate conceptually confusable objects and tau burden in the medial temporal lobe. In *European Meeting on Neurodegenerative Disease (EU-Mind)*.

Wang, HT, Clarke, N, Dessain, Q, Paugam, F, and Bellec, P (2024). Nilearn and big data facilitates transdiagnostic brain biomarkers. *Biological Psychiatry,* 95(10):S22.

Dessain, Q, Dricot, L, Delinte, N, Daussort, M, Macq, B, de Timary, P and Kupers, R (2023). Mood repair in long-term meditators is predicted by rsfmri functional connectivity of the insula and correlates with fa of the uncinate fasciculus. In *Proc. Intl. Soc. Mag. Reson. Med,* volume 31, page 0194.

Dausort, M, Delinte, N, Dessain, Q, Vanden Bulcke, C, and Macq, B (2023). A multi-compartment fingerprinting model for non-invasive tumor cell characterization via diffusion MRI. In *Proc. Intl. Soc. Mag. Reson. Med.,* volume 31.

Delinte, N, Dessain, Q, Dausort, M, Vanden Bulcke, C, and Macq, B (2023). Tractography of the subcortical u-fibers using a position-dependent maximum angle. In *Proc. Intl. Soc. Mag. Reson. Med.,* volume 31.

Dessain, Q, Delinte, N, Macq, B, and Rensonnet, G (2022). Fast multi-compartment microstructure fingerprinting using deep neural networks. In *Proc. Intl. Soc. Mag. Reson. Med,* volume 30, page 0560.

Delinte, N, Gosse, C, Dricot, L, Dessain, Q, Simon, M, Macq, B, Van Reybroeck, M and Rensonnet, G (2021). Microstructural alterations in the white matter of children with dyslexia assessed by multi-fascicle diffusion compartment imaging. In *Proc. Intl. Soc. Mag. Reson. Med.,* volume 29.

Publications in preparation and under review

Dessain, Q, Delinte, N, Hanseeuw, B, Dricot, L and Macq, B. Leveraging swin transformer for enhanced diagnosis of Alzheimer's disease using multi-shell diffusion MRI.

Derosiere, G, Vassiliadis, P, Dricot, L, Dessain, Q, Delinte, N, Zenon, A and Duque, J. Fronto-motor circuits linked to subclinical apathy.

Hao-Ting, W, Gau, R, Clarke, N, Dessain, Q and Bellec, P. Giga Connectome: a BIDS-app for time series and functional connectome extraction.

Dessain, Q, Dricot, L, Moniotte, S, Petit, G, Laureys, S, Darkner, S, de Timary, P and Kupers, R. Brain structural and functional connectivity differentiate long-term meditators from controls and predict emotional awareness and mood repair.

Dessain, Q, Dricot, L, Javadi, P, Chen, J, Zangenehpour, S, Ptito, M and Kupers, R. Does the smoker's brain mind the cigarette health warnings?

Contents

List of Figures . . . . . **xiii**

List of Tables . . . . . **xv**

Acronyms . . . . . **xvii**

1 Introduction **1**
1.1 Context and significance 1
1.2 Organization and contributions of this thesis 3

2 Theoretical background **9**
2.1 Diffusion MRI 11
2.1.1 Principles of diffusion MRI 11
2.1.2 Diffusion Tensor Imaging 17
2.1.3 Orientation Distribution Function 21
2.1.4 Neurite Orientation Dispersion and Density Imaging 23
2.1.5 Microstructure Fingerprinting 26
2.2 Deep learning 28
2.2.1 Multi-Layer Perceptron 28
2.2.2 Vision transformer and the Shifted window architecture 31

3 Motion correction in diffusion MRI acquisition with gradient cycling **35**
3.1 Introduction 36
3.2 Methods 40
3.2.1 Inter-volume motion correction (V2V) 40
3.2.2 Intra-volume motion correction (S2V) 43
3.2.3 Experiments 46
3.3 Results 50
3.3.1 Synthetic results 50
3.3.2 In-vivo results 52
3.4 Discussion and conclusion 54

4 Fast multi-compartment microstructure fingerprinting in brain white matter . . . 59
4.1 Introduction . . . 60
4.2 Materials and methods . . . 65
4.2.1 Microstructure Fingerprinting . . . 65
4.2.2 The Hybrid Method . . . 66
4.2.3 The Fully-Learned Method . . . 67
4.2.4 Diffusion protocol . . . 68
4.2.5 Experiment I: Performance assessment on synthetic test set . . . 68
4.2.6 Experiment II: Generalizability to unseen acquisition protocols 69
4.2.7 Experiment III: In-vivo population . . . 70
4.3 Results . . . 73
4.3.1 Experiment I: Performance assessment on synthetic test set . . . 73
4.3.2 Experiment II: Generalizability to unseen acquisition protocols 77
4.3.3 Experiment III: In-vivo population . . . 77
4.4 Discussion . . . 82
4.4.1 Performance . . . 82
4.4.2 Accuracy . . . 82
4.4.3 Interpretability . . . 82
4.4.4 Generalizability to diffusion protocol changes . . . 83
4.4.5 Generalizability to other microstructural models . . . 83
4.4.6 In-vivo application . . . 84
4.5 Conclusion . . . 84

5 Leveraging Swin Transformer for enhanced diagnosis of Alzheimer's disease using multi-shell diffusion MRI . . . 87
5.1 Introduction . . . 88
5.1.1 Contributions . . . 90
5.2 Materials and methods . . . 91
5.2.1 Dataset . . . 91
5.2.2 Processing the dMRI data . . . 92
5.2.3 Network architecture . . . 93
5.2.4 Experiments . . . 96
5.3 Results . . . 100
5.3.1 Classification performance . . . 100
5.3.2 Ablation study . . . 101
5.3.3 Amyloid classification results . . . 101
5.3.4 Explainability analysis . . . 102
5.4 Discussion . . . 104

5.4.1 Improved classification performance . . . 104
5.4.2 Insights from the ablation study . . . 105
5.4.3 Explainability and clinical relevance . . . 106
5.4.4 Framework flexibility and adaptability . . . 106
5.4.5 Comparison with state-of-the-art methods . . . 107
5.4.6 Limitations and future work . . . 107
5.5 Concluding remarks . . . 108

6 Discussion and conclusion . . . **111**
6.1 Key themes and challenges . . . 112
6.1.1 Reconciling precision and scalability in dMRI acquisition and processing . . . 112
6.1.2 Bridging the gap between engineering and biological interpretability . . . 113
6.2 Concluding remarks . . . 115

Appendices . . . **117**

A Supplementary materials for motion correction in diffusion MRI acquisition with gradient cycling . . . **117**

B Supplementary materials for fast multi-compartment microstructure fingerprinting . . . **119**

C Multi-modal study of brain changes in long-term meditators . . . **123**

Bibliography . . . **125**

List of Figures

1.1 Overview of the thesis structure 4

2.1 Venn diagram summarizing the key theoretical components of this thesis and their interconnections. 10
2.2 T_1 and T_2 relaxation after a 90° RF pulse 13
2.3 Schematic representation of a PGSE sequence 14
2.4 Illustration of single-shell and multi-shell schemes in q-space . . 17
2.5 Diffusion tensor representations 18
2.6 DTI maps visualization 20
2.7 Signal modeling with spherical deconvolution 22
2.8 Compartmental diffusion model in NODDI 24
2.9 Microstructure Fingerprinting workflow 27
2.10 MLP structure and computational process 29
2.11 Comparison of CNNs and ViTs 32

3.1 Effect of motion on intra-volume slice alignment in dMRI acquisitions 38
3.2 Comparison of diffusion MRI acquisition strategies for managing thermal load 39
3.3 Intra-volume motion artifacts and their correction in dMRI acquisitions 45
3.4 Simulated motion trajectory used in synthetic experiments . . . 47
3.5 Synthetic diffusion MRI volume with gradient cycling at $b = 700\,\mathrm{s/mm^2}$ 51
3.6 Synthetic diffusion MRI volume with gradient cycling at $b = 2000\,\mathrm{s/mm^2}$ 51
3.7 In-vivo diffusion MRI volume with gradient cycling at $b = 2000\,\mathrm{s/mm^2}$ 53
3.8 DTI FA maps with gradient cycling with V2V and S2V correction 53
3.9 FOD with gradient cycling with V2V and S2V correction 55

4.1 DW-MRI signal model of a white matter voxel 63

4.2 Overview of CSD & MF, Hybrid Method and Fully-Learned Method architectures 64
4.3 Mean absolute errors on synthetic structured test set (HCP protocol) 71
4.4 Correlation-accuracy plots for the four models on the test set with a two-fiber configuration (HCP protocol) 72
4.5 Angular error: CSD vs the Fully-Learned Method 75
4.6 t-SNE projection of the Hybrid Method network activations . . 75
4.7 Correlation-accuracy plots for the four models on the test set with a three-fiber configuration (HCP protocol) 76
4.8 Mean absolute errors on structured test set (clinical protocol) . 78
4.9 Histogram of the signed differences between CSD & MF and the Hybrid and Fully-Learned Methods for fvf_1 80
4.10 Detailed visualization on subject MGH 1001 of voxel-wise maps of fvf_1 81

5.1 Architecture overview 94
5.2 Shifted window approach in Swin Transformer 95
5.3 Grad-CAM activation map 104

A.1 In-vivo diffusion MRI volume with gradient cycling using at $b = 1000\,\mathrm{s/mm^2}$. 118
A.2 In-vivo diffusion MRI volume with gradient cycling using at $b = 3000\,\mathrm{s/mm^2}$. 118

B.1 Performance after coarser microstructural parameters during training 120

List of Tables

3.1 Mean voxel displacement error of S2V and V2V correction on synthetic data . . . 50
3.2 Computational efficiency of V2V and S2V correction across gradient cycling configurations . . . 54

4.1 Comparison of state-of-the-art methods . . . 63
4.2 Efficiency of the three methods when inference is performed on synthetic data using the MGH adult diffusion protocol . 73
4.3 MAEs comparison on the test set from Experiment 1 . . . 74
4.4 Efficiency of the three methods when inference is performed on 34 subjects from the MGH adult diffusion dataset . . . 79
4.5 Advantages and disadvantages of proposed methods . . . 85

5.1 Demographics of ADNI participants . . . 92
5.2 Classification results on ADNI data . . . 100
5.3 Ablation study for the AD vs. CN classification task . . . 101
5.4 Ablation study for the MCI vs. CN classification task . . . 102
5.5 Amyloid classification performance . . . 103
5.6 Top activated subregions identified by Grad-CAM . . . 103

B.1 Meta parameters and architectural characteristics of both accelerated methods . . . 121
B.2 Detailed computation times for the three methods when performing inference on synthetic data . . . 121

Acronyms

AD Alzheimer's Disease Dementia

ADC Apparent Diffusion Coefficient

ADNI Alzheimer's Disease Neuroimaging Initiative

AxD Axial Diffusivity

BIDS Brain Imaging Data Structure

BOLD Blood Oxygenation Level-Dependent

CN Cognitively Normal

CNN Convolutional Neural Network

CSD Constrained Spherical Deconvolution

CSF Cerebrospinal Fluid

DCT Discrete Cosine Transform

dMRI Diffusion Magnetic Resonance Imaging

DNN Deep Neural Network

DTI Diffusion Tensor Imaging

DW-MRI Diffusion-Weighted Magnetic Resonance Imaging

FA Fractional Anisotropy

FastMF Fast Microstructure Fingerprinting

fMRI Functional Magnetic Resonance Imaging

FOD Fiber Orientation Distributions

GM Gray Matter

GP Gaussian Process

Grad-CAM Gradient-Weighted Class Activation Mapping

HARDI High Angular Resolution Diffusion Imaging

HCP Human Connectome Project

MAE Mean Absolute Error

MCI Mild Cognitive Impairment

MD Mean Diffusivity

MF Microstructure Fingerprinting

MLP Multi-Layer Perceptron

MP-PCA Marchenko-Pastur Principal Component Analysis

MRF Magnetic Resonance Fingerprinting

MRI Magnetic Resonance Imaging

MSE Mean Squared Error

MSMT-CSD Multi-Shell Multi-Tissue Constrained Spherical Deconvolution

NNLS Non-Negative Linear Least-Squares

NODDI Neurite Orientation Dispersion and Density Imaging

ODF Orientation Distribution Function

ODI Orientation Dispersion Index

PET Positron Emission Tomography

PGSE Pulsed Gradient Spin Echo

qMRI Quantitative Magnetic Resonance Imaging

RD Radial Diffusivity

ReLU Rectified Linear Unit

RF Radiofrequency

ROI Regions of Interest

rs-fMRI Resting-State Functional Magnetic Resonance Imaging

RSNs Resting-State Networks

S2V Slice-to-Volume

SD Spherical Deconvolution

SH Spherical Harmonics

sMRI Structural Magnetic Resonance Imaging

SNR Signal-to-Noise Ratio

Swin Shifted Window

TE Echo Time

TI Inversion Time

TR Repetition Time

V2V Volume-to-Volume

ViT Vision Transformer

WM White Matter

xAI Explainable Artificial Intelligence

1 Introduction

1.1 Context and significance

Despite advances in neuroscience, understanding the biological basis of brain disorders remains an ongoing challenge, with implications for millions of patients worldwide. Diagnosing disorders for conditions like major depressive disorder, schizophrenia, bipolar disorder, and neurodegenerative diseases such as Alzheimer's disease often relies on subjective symptom assessment, which contributes to variability in both diagnosis and treatment outcomes (Allsopp, Read, Corcoran et al., 2019; Pearlson, Clementz, Sweeney et al., 2016). This diagnostic uncertainty significantly impacts patient care, as treatment decisions often depend on accurate classification of disease states and progression. The pressing need for more objective and precise diagnostic tools has fueled the search for reliable biomarkers, as these biological indicators hold the potential to improve how we diagnose, monitor, and treat these conditions (Thompson, Jahanshad, Ching et al., 2020).

Magnetic Resonance Imaging (MRI) has revolutionized our ability to non-invasively visualize the living brain. Conventional MRI techniques, such as T1- and T2-weighted imaging, excel at revealing macroscopic brain structure (brain morphometry) and detecting lesions. Functional MRI (fMRI) provides invaluable insights into brain activity and large-scale network connectivity by measuring blood-oxygen-level-dependent (BOLD) signals. These methods offer limited information about the brain's microscopic architecture, the arrangement and integrity of cells and nerve fibers at a scale far smaller than typical imaging resolution.

Diffusion MRI (dMRI) has emerged as a unique and powerful MRI modality specifically designed to probe this tissue microstructure in vivo.

Its fundamental principle lies in measuring the random motion of water molecules within brain tissue. In simple fluids, water diffuses freely and equally in all directions (isotropically). However, within the complex environment of the brain, water diffusion is significantly influenced by microscopic structures like cell membranes, organelles, and particularly the long axonal fibers that form white matter tracts. These structures act as barriers, hindering or restricting water movement more in some directions than others (anisotropic diffusion).

By sensitizing the MRI signal to the extent and directionality of this microscopic water displacement over milliseconds, dMRI provides indirect but highly valuable information about the underlying tissue architecture. For instance, in healthy white matter, water diffuses more easily along the tightly packed, myelinated axons than perpendicular to them. This directional preference allows dMRI to:

- infer the orientation of dominant nerve fiber bundles within each imaging voxel,
- reconstruct white matter pathways (tractography), providing a map of the brain's structural connections,
- quantify metrics reflecting tissue integrity, such as axonal density (how many axons are packed together) or other microstructural changes.

Because many neurological and psychiatric disorders involve subtle or plain alterations in white matter microstructure (e.g., demyelination in multiple sclerosis, axonal damage in traumatic brain injury, pathway degeneration in Alzheimer's disease), dMRI offers immense potential for developing sensitive biomarkers.

However, despite its potential, the widespread clinical application of advanced dMRI faces several significant technical challenges, which this thesis aims to address (Jones and Cercignani, 2010; Tax, Bastiani, Veraart et al., 2022).

1. **Sensitivity to motion:** dMRI sequences measure microscopic water displacement over time. Consequently, even minuscule head movements (sub-millimeter translations or small rotations) during the scan can severely corrupt the diffusion signal, obscuring the subtle microstructural information. This is particularly problematic because acquiring the rich data needed for advanced analysis often requires long scan times, increasing the likelihood of subject motion.

2. **Modeling complexity and computational demands:** Simple dMRI models, like Diffusion Tensor Imaging (DTI), assume a single, coherent fiber orientation per voxel. However, it's estimated that 60-90% of white matter voxels contain complex configurations, such as multiple crossing or fanning fiber populations (Jeurissen, Leemans, Tournier et al., 2013). Accurately characterizing such complex tissue requires sophisticated multi-compartment models that attempt to mathematically disentangle the signal contributions from different components (like intra-axonal water, extra-axonal water, and cerebrospinal fluid) and axonal populations. While offering richer microstructural detail, fitting these models is computationally intensive, often involving complex optimizations or large-scale dictionary matching, which can take hours or days per subject, hindering analysis in large clinical or research cohorts (Rensonnet, Adam, and Macq, 2021).
3. **Integrating advanced models and analysis techniques**: Translating the detailed outputs of complex dMRI models into clinically meaningful insights remains a challenge. Furthermore, effectively integrating these rich, high-dimensional diffusion datasets with powerful analytical tools like deep learning for tasks such as automated disease classification or prediction is an active area of research. Standard deep learning architectures may not be optimally suited for the unique structure of dMRI data, necessitating the development of methods capable of leveraging both spatial information and microstructural parameters effectively (Chen, Pawar, Ekanayake et al., 2023; Yan, Qu, Hu et al., 2022).

This thesis helps address these challenges by developing and applying novel methodologies that enhance the robustness, interpretability, and clinical applicability of microstructural imaging in the brain. Through innovations in motion correction, computational efficiency, and analysis frameworks, this research contributes to bridging the gap between dMRI's theoretical capabilities and its practical implementation in clinical settings.

1.2 Organization and contributions of this thesis

This thesis is structured around developing innovative solutions to the challenges outlined above, each addressing a distinct aspect of diffusion

MRI acquisition, processing, or application. Figure 1.1 provides a high-level overview of the thesis structure, summarizing the key contributions across multiple domains.

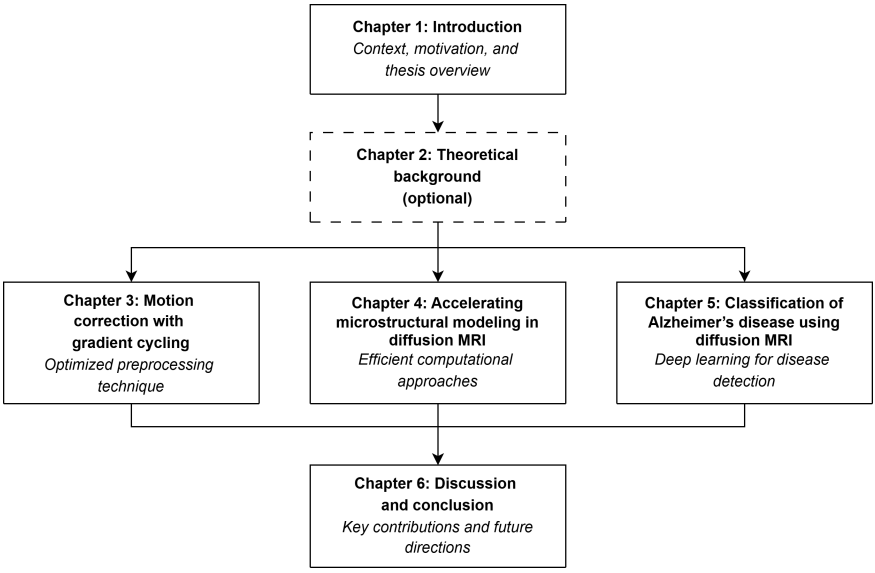


Fig. 1.1 Overview of the thesis structure.

Chapter 2: Theoretical background

This chapter provides a review of the fundamental principles and techniques relevant to this thesis. It covers the physics underlying dMRI measurements, explains how dMRI contrast is generated, and introduces key computational models used to interpret the dMRI signal. Starting with the basics of water diffusion in tissues, it progresses to established models like Diffusion Tensor Imaging and more advanced multi-compartment models such as Neurite Orientation Dispersion and Density Imaging (NODDI), and Microstructure Fingerprinting (MF). It also introduces essential deep learning concepts, including the Multi-Layer Perceptron (MLP), the Vision Transformer (ViT) and the Swin Transformer architectures, which are employed in later chapters.

Chapter 3: Motion correction in diffusion MRI acquisition with gradient cycling

Motivation: Advanced diffusion MRI protocols often require long scan times and strong magnetic gradients to measure microscopic movements in detail. This puts significant thermal stress on the scanner's hardware, especially the gradient coils, which can limit how much data can be safely acquired. To address this, a method called gradient cycling has been introduced (Hutter, Christiaens, Deprez et al., 2017; Hutter, Nilsson, Christiaens et al., 2018). Instead of using the same diffusion encoding measurement for every slice of the brain, gradient cycling alternates the diffusion encoding pattern between slices. This helps spread the workload, reducing heat buildup and making longer and more advanced scans possible.

However, while gradient cycling effectively manages thermal load, it comes at the cost of increased susceptibility to motion artifacts. Moreover, this technique introduces a new challenge: standard motion correction methods assume that all slices within a 3D image are acquired with the same settings. With gradient cycling, this is no longer true, making these traditional correction approaches ineffective. Even small head movements can then cause slices to misalign, leading to errors in the final images.

To fully benefit from gradient cycling without sacrificing image quality, it is essential to adapt motion correction methods to this specific acquisition scheme. Correcting these motion-related distortions is a key step toward reliable and efficient diffusion MRI for research and clinical use.

Contribution: This chapter presents a modified motion correction framework, specifically designed for gradient cycling acquisitions in diffusion MRI by extending the capabilities of FSL's eddy tool (Andersson, Graham, Drobnjak et al., 2017), a state-of-the-art motion correction algorithm. The new framework modifies the widely-used eddy algorithm to accurately model the slice-specific timing and diffusion encoding inherent in gradient cycling. The proposed approach effectively corrects both slice-to-slice misalignments within volumes and larger movements between volumes, reducing motion-induced artifacts and improving signal consistency across diffusion volumes. These contributions ensure that gradient cycling remains a viable strategy for advanced dMRI sequences, enhancing both acquisition efficiency and data quality.

Chapter 4: Fast multi-compartment microstructure fingerprinting in brain white matter

Motivation: Microstructure Fingerprinting is a powerful technique for estimating detailed microstructural properties of white matter using diffusion MRI. It infers these properties by matching the measured dMRI signal in each brain location (voxel) against a large, pre-computed dictionary of simulated dMRI signals, where each dictionary entry represents a unique, simulated tissue fingerprint. This matching process allows MF to estimate tissue parameters such as fiber volume fraction (how densely packed fibers are), fiber orientation, and extra-axonal diffusivity (diffusion outside fibers) from the raw dMRI data. However, its practical application is hindered by its computational complexity. The combinatorial nature of dictionary matching results in an exponential increase in processing time as the number of fascicles in a voxel grows. This bottleneck restricts the application of MF in large neuroimaging studies, where slow processing limits dataset scalability. Current dictionary-based approaches require an exhaustive search algorithm which is computationally expensive and do not scale well to whole-brain multi-fascicle modeling.

Contribution: This chapter introduces Fast Microstructure Fingerprinting (FastMF) (Dessain, Fuchs, Macq et al., 2024), a deep learning-based framework designed to dramatically accelerate the estimation of Microstructure Fingerprinting. The proposed methodology includes two distinct approaches to achieve this acceleration.

- **Hybrid Method:** Combines dictionary-based fingerprinting with a feed-forward neural network, enabling rapid estimation of microstructural parameters while preserving interpretability and protocol transferability.

- **Fully-Learned Method:** Uses a deep neural network with a spherical harmonics representation of the diffusion signal as input. It allows for an end-to-end estimation of tissue properties without the need for an external routine for estimating fiber orientations.

These methods achieve several orders of magnitude speedup compared to standard MF, while also improving estimation accuracy, unlocking the potential of applying MF to large-scale studies.

Chapter 5: Leveraging Swin Transformer for enhanced diagnosis of Alzheimer's disease using multi-shell diffusion MRI

Motivation: Alzheimer's disease (AD) is a complex neurodegenerative disorder characterized by progressive white matter degeneration, which can be investigated using multi-shell diffusion MRI. However, translating dMRI-derived biomarkers into clinically useful diagnostic tools remains a challenge due to the complexity and high dimensionality of diffusion data. Approaches relying on handcrafted features or statistical analysis of model-derived diffusion metrics may not fully capture the subtle microstructural changes associated with early-stage AD.

Deep learning offers a promising alternative by automatically learning meaningful patterns from dMRI data, potentially improving diagnostic accuracy and biomarker discovery. However, deep learning models have not been widely applied to AD classification using dMRI, limiting the integration of diffusion-based microstructural metrics into automated classification frameworks.

Contribution: This chapter presents a swin transformer-based deep learning model for AD classification using multi-shell dMRI. Key advancements include:

- leveraging diffusion models such as DTI and NODDI to capture AD-related microstructural changes in white matter,
- predicting amyloid beta presence from dMRI data, providing a potential biomarker for early AD diagnosis,
- applying Grad-CAM, an explainable AI (xAI) technique, to identify clinically relevant brain regions that contribute to model classification decisions.

Chapter 6: Discussion and conclusion

The thesis concludes with a broader discussion examining persistent challenges in diffusion MRI: balancing protocol richness with practical feasibility, and maintaining biological interpretability while advancing technical innovation. The chapter synthesizes the thesis contributions in addressing these challenges and presents future directions, including ongoing work that integrates together all developed methods.

Software contributions

Several computational tools have been developed as part of my PhD research to support MRI studies. In addition to the tool presented in the main chapters, the following contributions have been made:

- ElikoPy, a preprocessing pipeline for diffusion MRI (Dessain, Mathieu, and Delinte, 2024) which has been used for multiple ongoing studies at the IoNS institute.

  Among them, Appendix C presents a multi-modal neuroimaging study investigating brain structural and functional changes in long-term meditators. The study combines diffusion MRI, functional MRI, and anatomical imaging with extensive neuropsychological profiling to understand how meditation practice impacts the brain. By examining a cohort of 50 experienced meditators against matched controls, the study offers insights into how sustained cognitive practices can influence brain structure and function, while showcasing the value of a multi-modal analysis approach.

- Contributions to Giga-Connectome, a BIDS-App that streamlines functional connectivity analysis by generating connectomes directly from fMRIPrep (Esteban, Markiewicz, Blair et al., 2019) outputs. It integrates Nilearn and TemplateFlow (Ciric, Lorenz, Thompson et al., 2021) for time-series extraction, denoising, and parcellation, with customizable settings (bids-apps/giga_connectome).

- A workflow for processing neuroimaging data from the UK Biobank (Sudlow, Gallacher, Allen et al., 2015), a large-scale biomedical database containing brain scans from over 60 000 participants. The workflow handles both resting-state functional MRI and T1-weighted anatomical scans, adapting the fMRIPrep tool for efficient processing on High-Performance Computing clusters using dynamic job scheduling while maintaining standardized quality control aggregated in a centralized database (Hyedryn/ukbb-scripts).

- An open-source tool developed for GE scanners to mitigate gradient coil heating during gradient cycling dMRI sequences. The tool optimizes the temporal order of diffusion gradient within cycling blocks (e.g., 2TR or 3TR) to distributes the thermal load more effectively than standard vendor implementations, enabling more efficient acquisition protocols (Hyedryn/dMRI_Gradient_Cycling_Optimizer).

2

Theoretical background

Preliminary note:

This chapter provides a concise overview of the magnetic resonance imaging modalities and deep learning models relevant to this thesis.

Readers already familiar with the following concepts may skip this chapter.

- *Diffusion MRI*
  - *Principles of diffusion MRI (2.1.1)*
  - *Diffusion Tensor Imaging (DTI) (2.1.2)*
  - *Orientation Distribution Function (ODF) (2.1.3)*
  - *Neurite Orientation Dispersion and Density Imaging (NODDI) (2.1.4)*
  - *Microstructure Fingerprinting (MF) (2.1.5)*
- *Deep learning*
  - *Multi-Layer Perceptron (MLP) (2.2.1)*
  - *Vision Transformer (ViT) and the Shifted Window (Swin) architecture (2.2.2)*

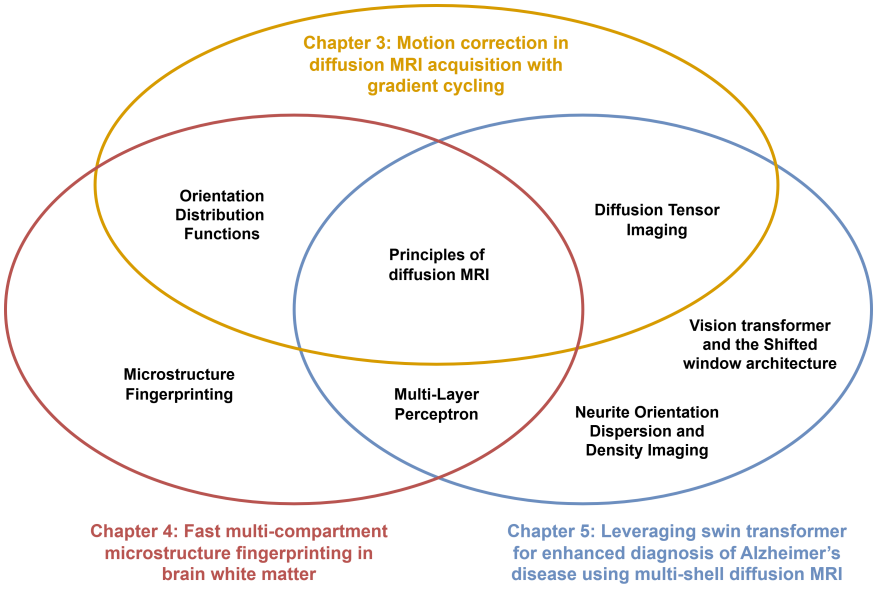


Fig. 2.1 Venn diagram summarizing the key theoretical components of this thesis and their interconnections.

Extracting meaningful information from diffusion MRI data requires a solid understanding of both the physical principles underlying diffusion measurements and the mathematical frameworks used to model tissue properties. This chapter provides the theoretical foundation necessary for the methodological advancements presented in this thesis.

The fundamental principles of the dMRI signal is first introduced, which serve as the basis for all subsequent chapters. Next, the chapter then explores diffusion models of increasing complexity, from classical diffusion tensor imaging to advanced multi-compartment approaches. This progression establishes the theoretical background for three key developments: the motion correction algorithm in Chapter 3, validated using DTI and ODFs; the accelerated microstructural models in Chapter 4, which combine ODF analysis with Microstructure Fingerprinting; and the classification framework in Chapter 5, which utilizes DTI and NODDI. The chapter concludes with deep learning architectures central to this thesis: the Multi-Layer Perceptron used in Chapter 4 for accelerated microstructure estimation and the Vision Transformer architecture employed in Chapter 5 for Alzheimer's disease classification.

2.1 Diffusion MRI

2.1.1 Principles of diffusion MRI

Quick overview

Magnetic resonance imaging (MRI) is a powerful non-invasive imaging technique that exploits the magnetic properties of atomic nuclei to generate detailed images of soft tissues. Diffusion MRI is a specialized form of MRI designed to measure the microscopic movement of water molecules in biological tissues. This movement, known as Brownian motion, is influenced by tissue microstructural properties, making dMRI a valuable tool for assessing cellular density, fiber orientation, and pathological changes. By leveraging magnetic field gradients, dMRI encodes diffusion information into the MRI signal, allowing for the reconstruction of white matter (WM) structures in the brain.

Fundamentals of MRI: Spin behavior and magnetization

MRI is based on the quantum mechanical property of spin, which is inherent to certain atomic nuclei, most notably hydrogen protons (^1H). Hydrogen is highly abundant in biological tissues, especially in water and fat, making it an ideal target for MRI imaging.

In the absence of an external magnetic field, these proton spins are randomly oriented. However, when a strong external magnetic field (B_0) is applied, typically using a superconducting magnet in the MRI scanner, the spins align in one of two possible states:

1. Parallel (low-energy state): Spins align with the magnetic field.
2. Antiparallel (high-energy state): Spins align against the magnetic field.

Due to a slight energy difference between these two states, slightly more spins align in the parallel direction, creating a net magnetization vector (M) along the direction of the main magnetic field B_0, typically along the longitudinal axis. This net magnetization is the basis of the MRI signal.

Additionally, in the presence of B_0, the spins begin to precess (wobble like a spinning top) around the magnetic field direction at a characteristic frequency known as the Larmor frequency:

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

where γ is the gyromagnetic ratio, a constant specific to hydrogen nuclei, and B_0 is the strength of the external magnetic field. The precession of spins generates an oscillating magnetic field, which can be manipulated to produce an MRI signal.

Excitation and signal formation: Role of RF pulses

To generate an MRI signal, a radiofrequency (RF) pulse is applied perpendicular to B_0 at the Larmor frequency. This pulse temporarily disturbs the equilibrium magnetization, producing two key effects:

1. Tips the magnetization vector away from the longitudinal axis (along B_0) and into the transverse plane. This occurs because the RF pulse, applied at the Larmor frequency, resonates with spins, promoting some to a higher-energy state and causing net magnetization to rotate. This reorientation into the transverse plane is essential for signal detection, as the receiver coils are designed to sense the magnetic changes produced by magnetization rotating within this plane.
2. Aligns the precessing spins in phase, meaning the individual spins, while precessing (wobbling) around B_0, are momentarily synchronized to point in the same direction in their cycle. This synchronization creates a strong oscillating magnetic field in the transverse plane that can be detected by the receiver coil.

Once the RF pulse is turned off, the magnetization gradually returns to its equilibrium state along B_0 through two relaxation processes, as illustrated in Figure 2.2:

- T1 relaxation (longitudinal relaxation): The magnetization realigns with B_0 as spins return to their lower-energy parallel state. The time constant T_1 describes how quickly this occurs.
- T2 relaxation (transverse relaxation): The synchronized spins gradually lose phase coherence (fall out of sync) due to small, unavoidable variations in the local magnetic field inhomogeneities which induce

slight differences in their precession rates. As they dephase, their individual signal contributions begin to cancel each other out, causing the net transverse magnetization to decay. The time constant T_2 describes how quickly this dephasing process occurs.

The oscillating transverse magnetization (while it lasts before T2 decay) induces an electromagnetic signal in the receiver coil, which detects the MRI signal perpendicular to B_0. However, standard MRI sequences primarily encode spatial information. To obtain diffusion-specific contrast, additional magnetic field gradients must be introduced.

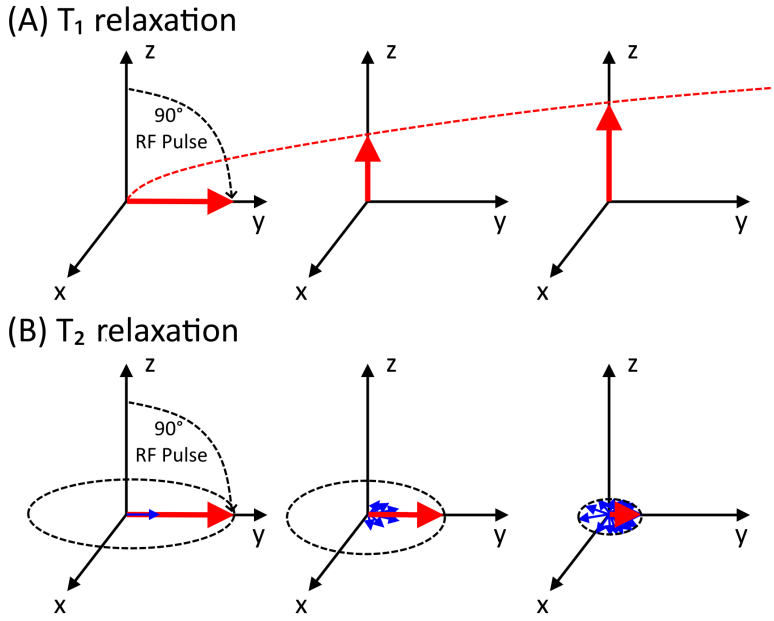


Fig. 2.2 (A) T_1 relaxation following a 90° RF pulse (B) T_2 relaxation following a 90° RF pulse. Magnetization vector in red and individual spins in blue.

Encoding diffusion contrast: The role of magnetic field gradients

Diffusion MRI takes advantage of the natural motion of water molecules to probe tissue microstructure. In free fluids, water molecules move randomly in all directions, a process known as Brownian motion. However, in biological tissues, cellular membranes, myelin, and extracellular structures

restrict this motion, leading to varying diffusion properties depending on tissue type and microstructural organization. Standard MRI primarily reveals tissue differences based on relaxation times, but isn't directly sensitive to this microscopic motion. To measure these microscopic movements, diffusion MRI employs magnetic field gradients which are temporary, controlled variations in the magnetic field strength across space.

The most widely used approach for encoding diffusion contrast is the Pulsed Gradient Spin Echo (PGSE) sequence, originally introduced by Stejskal and Tanner (1965). This sequence consists of a standard spin echo sequence (Hahn, 1950), enhanced by the application of two strong, matched diffusion-sensitizing gradient pulses on either side of the 180° refocusing RF pulse. These gradient pulses are the key to measuring diffusion. They introduce temporary magnetic field variations that cause the spins' precession phase (analogous to the precise alignment of multiple wobbling tops) to change depending on their spatial location. This phase encoding allows accurate tracking of their displacement, enabling the measurement of diffusion. A schematic representation of PGSE is illustrated in Figure 2.3(A).

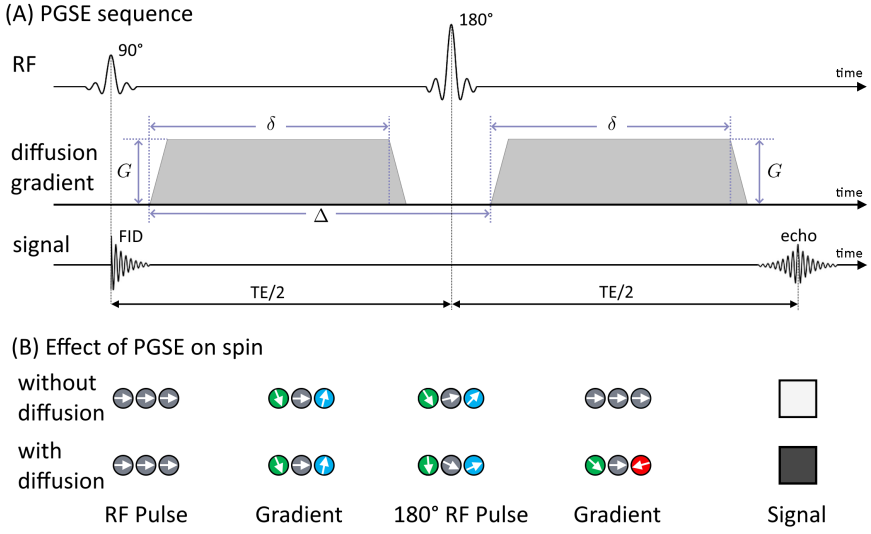


Fig. 2.3 (A) Schematic representation of a PGSE sequence with a 90° excitation pulse, 180° refocusing pulse, and spin echo readout. Diffusion is encoded by a pair of diffusion gradients, characterized by their length δ, amplitude G, and spacing Δ. (B) Effect of diffusion gradients on spins without and with water molecules diffusion.

The PGSE sequence consists of several key steps.

1. 90° **RF Pulse (excitation):** The process begins with a 90-degree RF pulse, which tips the net magnetization into the transverse plane, allowing the MRI signal to be detected. This is a standard step in many MRI sequences.
2. **First diffusion gradient:** Shortly after this, the first diffusion gradient is applied along a chosen direction. This gradient induces a position-dependent phase shift in the spins, meaning that spins in different spatial locations temporarily precess at slightly different rates and acquire different phases. If all spins remained stationary, this phase shift would be completely reversible. In effect, this gradient establishes a temporary mapping where each spin's initial location corresponds to a specific accumulated phase shift.
3. 180° **RF Pulse (refocusing spins):** To compensate for signal loss due to static magnetic field inhomogeneities, a 180-degree RF pulse is applied halfway through the diffusion period. This pulse inverts the spins' phase, meaning that those experiencing static field inhomogeneities will experience the opposite phase shift after inversion, allowing for a proper refocusing of the MRI signal, which is essential for forming the echo in a spin echo sequence. However, this pulse does not reverse the phase shifts introduced by the first diffusion gradient.
4. **Second diffusion gradient:** Following the 180-degree RF pulse, the second diffusion gradient is applied. This gradient is identical in strength, duration, and direction to the first. Its effect depends critically on whether the spin has moved:

    - *For stationary spins:* A spin that remains in the same position experiences a phase shift from the second gradient that, due to the preceding 180° pulse's inversion effect, perfectly cancels out the initial position-dependent phase shift acquired from the first gradient. This cancellation effectively reverses the phase mapping established by the first gradient, and the signal contribution from these spins is fully refocused.
    - *For moving spins:* If a spin has moved due to diffusion to a different position along the gradient direction, the phase shift imparted by the second gradient will be different from what is needed to cancel the (inverted) initial position-dependent phase

shift. Consequently, the phase mapping established by the first gradient is not fully reversed for these spins.

As a result, the phases of spins that have moved become dispersed or scrambled relative to each other. This phase dispersion across the many moving water molecules within a voxel leads to destructive interference of their signals, causing a signal attenuation.

5. **Signal detection:** Finally, the remaining transverse magnetization produces an echo signal. The intensity (strength) of this measured signal is lower if significant diffusion occurred along the gradient direction during the time between the gradient pulses, because moving spins led to phase dispersion and signal attenuation.

The amount of signal loss therefore depends directly on the extent of molecular displacement during the diffusion period, as illustrated conceptually in Figure 2.3(B). In free fluids like cerebrospinal fluid (CSF), water molecules move more freely over longer distances, leading to greater phase dispersion and thus greater signal attenuation. In structured tissues, such as white matter, diffusion is more restricted along certain directions (e.g., along axons), meaning that signal loss depends strongly on the fiber orientation relative to the applied gradient direction.

The degree to which the PGSE sequence sensitizes the MRI signal to molecular motion is determined by the strength and timing of the diffusion gradients, summarized by a parameter called the b-value. It is calculated as:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3}), \tag{2.2}$$

where γ is the gyromagnetic ratio of hydrogen, G is the gradient strength, δ is the duration of the gradient pulse, and Δ is the time between the two gradient pulses. Higher b-values correspond to stronger diffusion weighting, meaning the sequence becomes more sensitive to diffusion (even small movements cause significant signal loss), but this also generally reduces the overall signal-to-noise ratio (SNR). Choosing appropriate b-values is crucial. A typical clinical b-value ranges from $700\,\mathrm{s\,mm^{-2}}$ to $2000\,\mathrm{s\,mm^{-2}}$, while advanced diffusion imaging techniques may use values of $3000\,\mathrm{s\,mm^{-2}}$ or higher to probe specific microstructural features. dMRI acquisitions are often organized into shells, where diffusion measurements are taken in multiple directions at a constant b-value, forming a spherical sampling pattern in q-space as

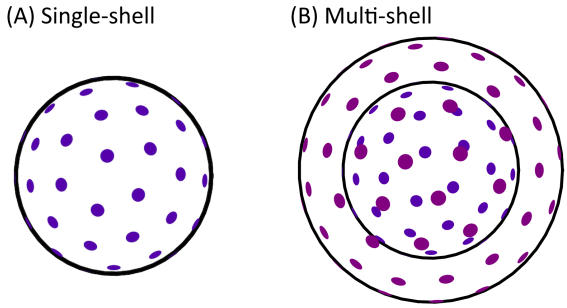


Fig. 2.4 Illustration of single-shell and multi-shell sampling schemes in diffusion MRI q-space. (A) Single-shell acquisition samples points on a single spherical shell in q-space. (B) Multi-shell acquisition samples multiple concentric shells, enabling advanced diffusion models.

illustrated in Figure 2.4. Q-space represents the space of diffusion-encoded signals, where each point corresponds to a specific gradient direction and b-value. Single-shell acquisitions collect data at a single b-value, while multi-shell acquisitions include multiple b-values, one by shell, providing richer microstructural information.

Through this mechanism, diffusion MRI encodes molecular motion into the MRI signal, providing indirect but powerful information about tissue microstructure. By analyzing how the signal attenuates in different gradient directions and b-values, diffusion properties can be estimated, and white matter pathways can be mapped.

2.1.2 Diffusion Tensor Imaging

Quick overview

Diffusion Tensor Imaging is a widely used model in dMRI that characterizes the anisotropic diffusion of water molecules within biological tissues (Basser, Mattiello, and LeBihan, 1994). By modeling diffusion as a 3×3 symmetric positive semi-definite tensor, DTI provides a mathematical framework for describing the three-dimensional diffusion process at the voxel level, enabling the study of white matter architecture and connectivity in the brain. Key limitations of DTI include its inability to resolve crossing fiber populations within a voxel, partial volume effects at tissue interfaces, and the assumption of Gaussian diffusion which breaks down at high b-values.

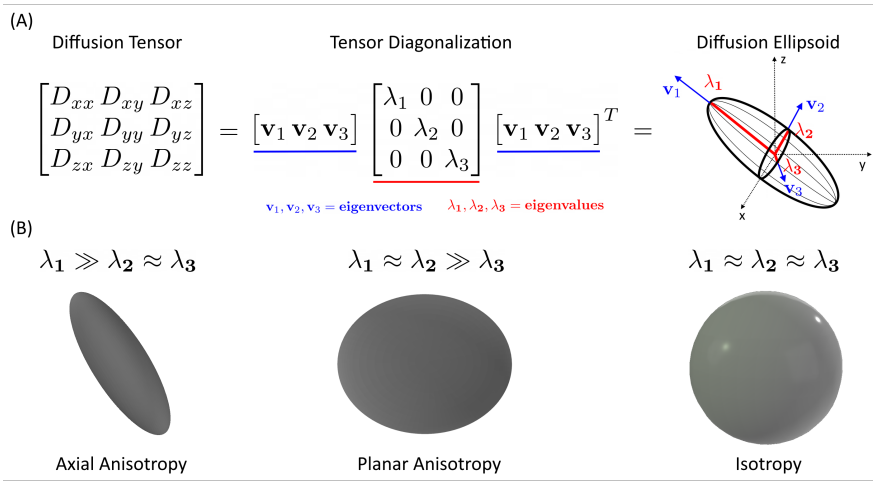


Fig. 2.5 Illustration of diffusion tensor representations. (A) The diffusion tensor is mathematically described as a 3×3 symmetric matrix, which can be diagonalized to obtain three eigenvalues ($\lambda_1, \lambda_2, \lambda_3$) and their corresponding eigenvectors (v_1, v_2, v_3). (B) Visualization of diffusion profiles using diffusion ellipsoids: an elongated ellipsoid indicates high anisotropy with dominant diffusion along one axis (axial anisotropy), a flattened ellipsoid represents planar anisotropy with significant diffusion within a plane, and a spherical shape corresponds to isotropic diffusion with equal diffusivity in all directions, such as in the CSF.

Understanding how DTI captures diffusion properties requires first defining the diffusion tensor, a mathematical representation of the directional dependence of water molecule movement. The diffusion tensor, denoted as $\boldsymbol{D}$, is represented as

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

The tensor can be decomposed into three eigenvalues ($\lambda_1, \lambda_2, \lambda_3$), which represent the principal diffusivities, and a set of eigenvectors (v_1, v_2, v_3), which define the corresponding diffusion orientations. This eigendecomposition enables the visualization of the diffusion tensor as an ellipsoid, with different diffusion profiles depending on the underlying microstructure, as illustrated in Figure 2.5.

The relationship between the diffusion tensor and the MRI signal was described by Stejskal and Tanner (1965) using an equation which models the attenuation of the diffusion-weighted signal as a function of the applied gradient direction $\hat{\boldsymbol{g}}$ and the b-value:

$$\frac{S(b,\hat{\boldsymbol{g}})}{S_0} = e^{-b\hat{\boldsymbol{g}}^T \boldsymbol{D} \hat{\boldsymbol{g}}}, \tag{2.4}$$

where $S(b,\hat{\boldsymbol{g}})$ is the diffusion-weighted signal, S_0 is the non-diffusion-weighted signal (b0 image), and b is the diffusion weighting factor.

To estimate the diffusion tensor, at least six diffusion-weighted images acquired along non-collinear gradient directions are required in addition to a b0 image. However, in practical applications, 20–30 gradient directions are typically acquired to improve the robustness of tensor estimation and reduce the effects of noise and acquisition artifacts (Jones, 2004; Skare, Hedehus, Moseley et al., 2000).

Several scalar metrics can be derived from the diffusion tensor to quantify different aspects of water diffusion in brain tissue, as illustrated in Figure 2.6. Mean Diffusivity represents the average diffusivity within a voxel and is useful for detecting abnormalities such as increased extracellular water:

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \tag{2.5}$$

Axial Diffusivity measures diffusion along the primary fiber direction:

$$AxD = \lambda_1. \tag{2.6}$$

Radial Diffusivity quantifies diffusion perpendicular to the primary fiber direction:

$$RD = \frac{\lambda_2 + \lambda_3}{2}. \tag{2.7}$$

The most widely used metric, Fractional Anisotropy, indicates how directional diffusion is within a voxel, ranging from 0 (isotropic diffusion) to 1 (highly anisotropic diffusion):

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}. \tag{2.8}$$

Color FA maps enhance visualization by encoding fiber orientations in an RGB color scheme (Pajevic and Pierpaoli, 1999).

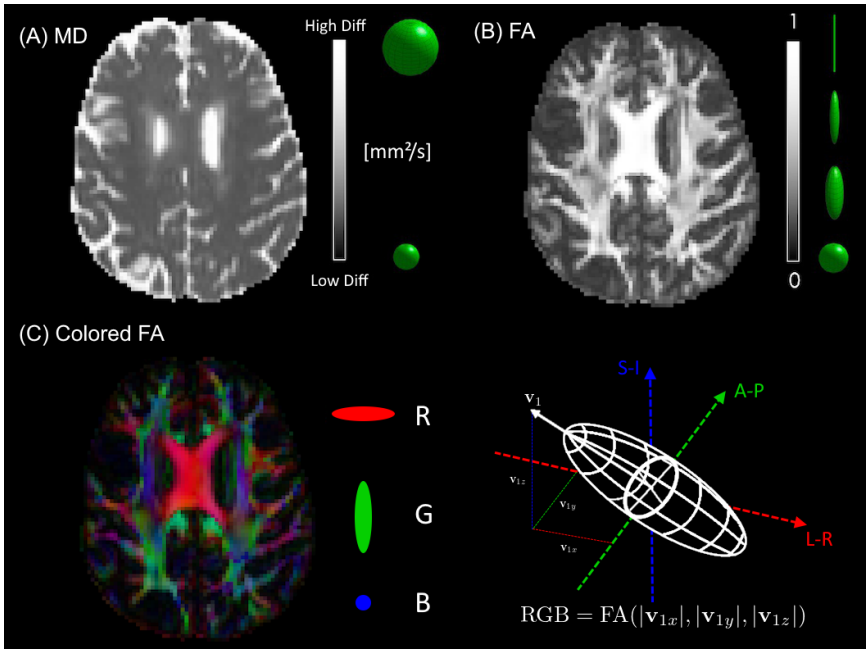


Fig. 2.6 Visualization of diffusion tensor imaging maps. (A) Mean Diffusivity. (B) Fractional Anisotropy. (C) Colored FA maps, where red, green, and blue represent left–right, anterior–posterior, and superior–inferior orientations.

While DTI has significantly advanced the study of white matter microstructure, it has several limitations. The model assumes a single dominant fiber direction per voxel, which limits its ability to resolve crossing fiber populations, a common occurrence in 60–90% of white matter voxels (Descoteaux, Deriche, Knosche et al., 2008; Jeurissen, Leemans, Tournier et al., 2013). In regions where multiple white matter pathways cross within a voxel, DTI can only provide an averaged representation of diffusion, leading to artificially reduced FA values and potential misinterpretations of white matter integrity. Another limitation arises from partial volume effects at tissue interfaces, which can lead to biased tensor estimates. Additionally, DTI models diffusion as a Gaussian process, which does not fully capture complex microstructural features such as fiber fanning, bending, or branching.

2.1.3 Orientation Distribution Function

Quick overview

The Orientation Distribution Function (ODF) is a continuous function defined on the unit sphere S^2, which describes the angular distribution of fiber orientations within a voxel in diffusion MRI. Unlike the DTI model, which assumes a single predominant fiber orientation per voxel and represents diffusion as an ellipsoid, the ODF provides a more detailed representation of fiber architecture. This is particularly beneficial in regions where multiple fiber populations cross, merge, or diverge, which are common in complex white matter structures.

Defining the ODF requires a function defined on the unit sphere S^2, mapping each orientation (θ, ϕ) to a real-valued scalar, representing the likelihood of water diffusion along that direction. Formally, the ODF is written as:

$$ODF(\theta, \phi, b) = \int_0^\infty P(r, \theta, \phi, b) r^2 \, dr, \tag{2.9}$$

where $P(r, \theta, \phi, b)$ is the diffusion propagator, describing the probability density of water displacement at distance r along direction (θ, ϕ) over the diffusion weighting factor b.

The integral projects the diffusion propagator onto the unit sphere, emphasizing its angular structure while marginalizing out the radial component. This formulation allows the ODF to capture the angular structure of diffusion, making it sensitive to multiple fiber orientations within a voxel.

Since direct computation of the diffusion propagator is impractical, as it would require probing a large extent of q-space, the ODF is typically estimated using spherical deconvolution (SD), a mathematical technique that separates contributions from different fiber populations. This is achieved by deconvolving the measured diffusion signal $S(y, \phi)$ using a predefined fiber response function R, which characterizes the expected signal from a single, well-aligned fiber population (see Figure 2.7):

$$S(y, \phi) = ODF(y, \phi) \circledast R, \tag{2.10}$$

where $\circledast$ denotes the convolution operation on the unit sphere. By solving this inverse problem, the ODF is reconstructed, providing an estimate of the fiber orientation distribution in each voxel. However, spherical de-

convolution lacks stability, meaning that small errors in the measured signal or response function can lead to large fluctuations in the estimated ODF. To address this issue, additional constraints are required to ensure robustness.

In the context of this thesis, Constrained Spherical Deconvolution (CSD) (Tournier, Calamante, and Connelly, 2007) and its extension, multi-shell multi-tissue CSD (MSMT-CSD) (Jeurissen, Tournier, Dhollander et al., 2014), are used to compute the ODF while ensuring numerical stability. CSD imposes a nonnegativity constraint, preventing spurious negative lobes in the estimated ODF. It assumes a single white matter response function and is widely used in single-shell diffusion MRI acquisitions, offering a reliable estimation of fiber orientations within each voxel. MSMT-CSD extends this framework by accounting for multiple tissue compartments, specifically white matter, gray matter (GM), and CSF. Unlike CSD, which models only a white matter signal, MSMT-CSD introduces separate response functions for each tissue type, leading to a more accurate decomposition of the diffusion signal. By incorporating tissue-specific response functions, MSMT-CSD mitigates partial volume effects, especially at WM-GM and WM-CSF interfaces, resulting in fewer spurious peaks.

Through these deconvolution techniques, the local maxima of the ODF can be extracted, which correspond to fiber orientations within a voxel. The integral or amplitude of each peak can provide a measure of relative fiber density. Thresholding peak amplitudes and setting angular separation criteria help mitigate noise artifacts and ensure robust fiber orientation extraction. These characteristics make ODF-based models particularly valuable for tractography, network-based connectivity analysis, and biomarker extraction models such as Microstructure Fingerprinting.

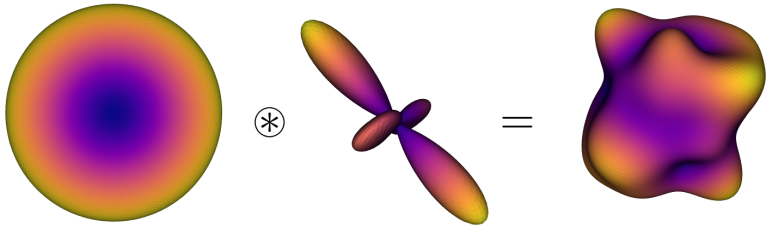


Fig. 2.7 The single-shell diffusion-weighted signal (right) is modeled as the spherical convolution of the response function (left) and the fiber orientation distribution function (middle).

2.1.4 Neurite Orientation Dispersion and Density Imaging

Quick overview

Neurite Orientation Dispersion and Density Imaging (NODDI) is a dMRI model that characterizes the structure of neuronal tissue by quantifying neurite density and orientation dispersion (Zhang, Schneider, Wheeler-Kingshott et al., 2012). Unlike DTI, which provides a general measure of anisotropic diffusion using a tensor, NODDI separates two key microstructural properties: the proportion of neurites (axons and dendrites) within a voxel and the degree to which these neurites are dispersed in different orientations.
At the core of NODDI lies a three-compartment model that describes water diffusion in different microstructural environments (see Figure 2.8). These compartments are the intra-cellular space which corresponds to neurites, the extra-cellular space which surrounds the neurites, and the CSF compartment which represents free water diffusion.

NODDI models the total diffusion-weighted signal as a weighted sum of the contributions from three compartments, following the principle of superposition:

$$A = (1 - \nu_{iso})(\nu_{ic} A_{ic} + (1 - \nu_{ic}) A_{ec}) + \nu_{iso} A_{iso}, \tag{2.11}$$

where A_{ic} and ν_{ic} represent the signal and volume fraction of the intra-cellular compartment, A_{ec} corresponds to the extra-cellular compartment, and A_{iso} and ν_{iso} describe the CSF compartment.

The intra-cellular compartment models the diffusion of water molecules inside neurites, where water movement is highly restricted perpendicular to the neurite's axis due to cellular membranes but relatively free along its length. This compartment is represented as an ensemble of zero-radius cylinders to capture the anisotropic nature of intra-axonal diffusion. Since neurites in brain tissue are not perfectly aligned, their orientation follows a statistical distribution. NODDI uses the Watson distribution to model this dispersion, which describes how tightly neurites cluster around a dominant orientation. This allows the model to quantify orientation dispersion, distinguishing between highly organized fiber tracts, such as the corpus callosum, and regions where fibers fan out or bend.

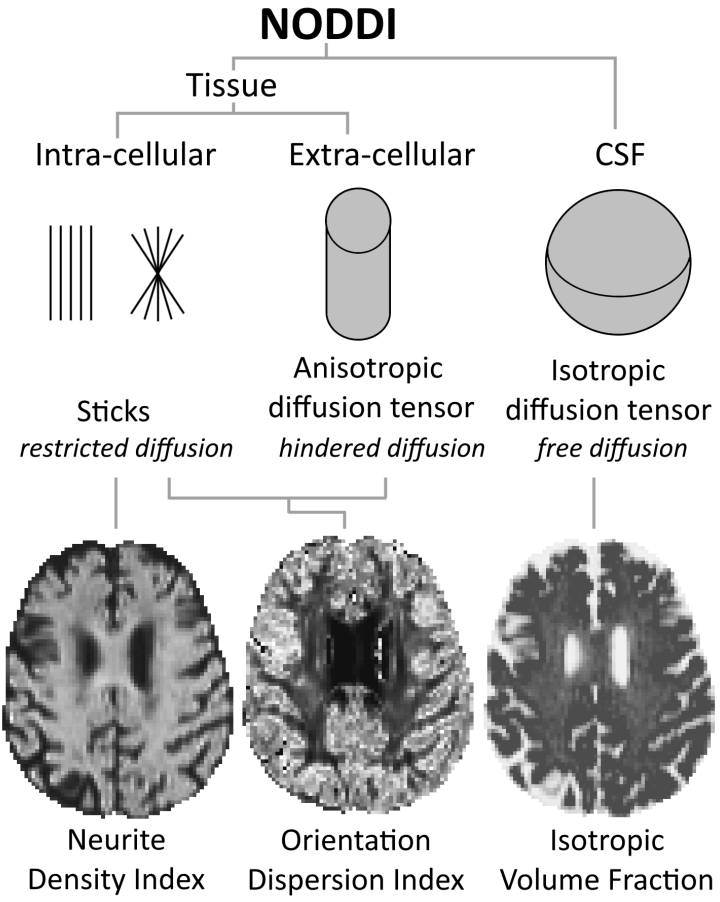


Fig. 2.8 Schematic representation of the NODDI model, which partitions water diffusion into three compartments: (left) restricted diffusion in the intra-cellular space (neurites), (middle) hindered diffusion in the extra-cellular space, where water interacts with neurite membranes, and (right) free, isotropic diffusion in the CSF.

The extra-cellular compartment represents the space surrounding neurites, where diffusion is hindered by cellular structures but not entirely restricted. This compartment is modeled using a Gaussian anisotropic diffusion process, meaning that diffusion is more hindered in some directions than others due to the presence of neurites. To improve model stability, NODDI incorporates a tortuosity constraint, which links the perpendicular diffusivity of the extra-cellular space ($d_\perp$) to the neurite density (ν_{ic}) and the parallel diffusivity ($d_\parallel$) through the relationship:

$$d_\perp = d_\parallel(1 - \nu_{ic}). \tag{2.12}$$

This ensures that as neurite density increases, diffusion perpendicular to neurites becomes more restricted, effectively capturing the interplay between neurite density and hindered diffusion.

The cerebrospinal fluid (CSF) compartment accounts for isotropic diffusion in fluid-filled spaces such as the ventricles. Since CSF diffusion is largely unrestricted, this compartment follows a simple Gaussian diffusion model with a fixed diffusivity.

Beyond estimating the volume fractions of these three compartments, NODDI provides an additional microstructural indice, the Orientation Dispersion Index (ODI), that measures how neurites spread in different directions, ranging from 0 (highly aligned neurites) to 1 (completely isotropic distribution). This metric overcomes one limitation of DTI-derived measures such as Fractional Anisotropy, which cannot separate the effects of fiber density and dispersion.

Despite its advantages, NODDI has some limitations. The Watson distribution assumes that neurites follow a unimodal distribution around a single mean direction, making it incapable of modeling crossing fibers, which are present in a large proportion of white matter voxels. This is a significant limitation in brain regions where multiple fiber populations intersect. Additionally, the model does not account for radial diffusivity within axons, assuming that all the intra-cellular diffusion is completely restricted perpendicular to neurites.

Nevertheless, NODDI has become a widely used tool. By providing biologically interpretable microstructural indices, it surpasses traditional DTI-based metrics in characterizing brain tissue. Its ability to disentangle neurite density and dispersion makes it particularly valuable for understanding white and gray matter microstructure, offering deeper insights into brain development, aging, and disease.

2.1.5 Microstructure Fingerprinting

Quick overview

Microstructure Fingerprinting (MF) is a multi-compartment model designed to estimate white matter microstructure in voxels containing one or multiple fascicles (Rensonnet, Scherrer, Girard et al., 2019). Unlike traditional analytical models that rely on simplified mathematical formulations, MF leverages Monte Carlo simulations to generate physically accurate diffusion signals. These precomputed signals form a dictionary of microstructural fingerprints, which serves as a reference to estimate the best-fitting tissue configuration in each voxel. This approach enables MF to capture complex features such as fiber volume fraction, and extra-axonal diffusivity while avoiding the restrictive assumption of a single fascicle per voxel made by other models such as DTI and NODDI.

MF replaces analytical models with a library of simulated diffusion signals, allowing for a precise estimation of microstructure configuration. Instead of relying on closed-form mathematical approximations, MF leverages Monte Carlo simulations to create a dictionary of microstructural fingerprints, which represent a wide range of tissue configurations. By matching the measured diffusion signal to rotated versions of the precomputed dictionary, where each version is aligned with the orientation of a fascicle found in the voxel, MF can identify the tissue configuration that best matches the observed signal. A visual representation of the MF workflow is provided in Figure 2.9.

A key advantage of MF is its flexibility. Monte Carlo simulations allow the dictionary to incorporate any microstructural configuration, making it possible to model additional parameters such as axonal undulation or permeability. This adaptability makes MF particularly useful for investigating complex white matter structures, especially in regions with crossing fibers, where simpler models like DTI or NODDI struggle to provide an accurate representation.

A more detailed presentation of MF, including its mathematical formulation, advantages, and limitations, is provided in Chapter 4, where we introduce two accelerated implementations of MF: the Hybrid Method and the Fully-Learned Method.

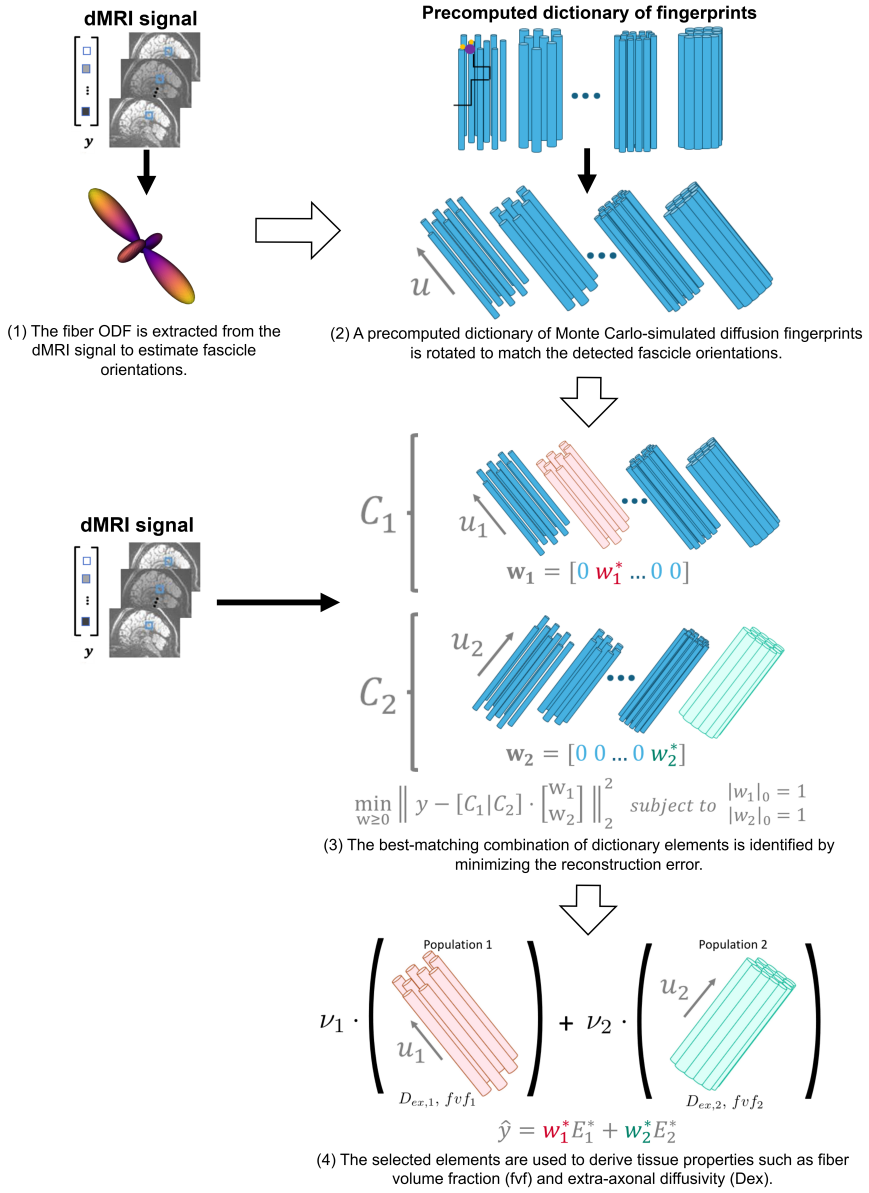


Fig. 2.9 Workflow of Microstructure Fingerprinting for estimating white matter microstructure in a voxel containing two fascicles. (1) ODF extraction. (2) Dictionary rotation. (3) Exhaustive search for optimal solution. (4) Microstructural parameter extraction.

2.2 Deep learning

2.2.1 Multi-Layer Perceptron

Quick overview

A Multi-Layer Perceptron (Haykin, 1994; Rumelhart, Hinton, and Williams, 1986) is a foundational type of artificial neural network designed to model complex relationships in data. It consists of an input layer that receives structured data, one or more hidden layers where computations take place, and an output layer that produces predictions. Each neuron in the hidden layers applies a weighted sum of inputs followed by a nonlinear activation function, allowing the model to capture patterns beyond simple linear relationships. MLPs are particularly effective for structured data but struggle with high-dimensional inputs like raw MRI scans, where alternative architectures such as Convolutional Neural Networks (LeCun, Bengio et al., 1995) or transformers (Vaswani, 2017) are more suitable.

An MLP follows a feedforward structure, meaning information flows sequentially from the input layer through hidden layers to the output layer, as illustrated in Figure 2.10. The input layer acts as the entry point for preprocessed data, such as numerical features or encoded information. In the context of MRI, the input features could include clinical biomarkers derived from imaging studies, voxel intensities or statistical measures derived from a connectivity analysis. These inputs are propagated through the hidden layers, where the core computational processes occur.

Within the hidden layers, each neuron performs two critical operations. First, it computes a weighted sum of inputs from the previous layer, where the weights determine the strength of connections between neurons. Second, it applies a nonlinear activation function, such as ReLU (Rectified Linear Unit) or sigmoid, to introduce nonlinearity. Without it, the MLP could only model linear relationships, severely limiting its predictive capability. The presence of multiple hidden layers enables hierarchical feature extraction, where initial layers may detect low-level patterns, such as variations in a microstructural feature, while deeper layers can learn more abstract representations, such as distinguishing normal from pathological brain structures.

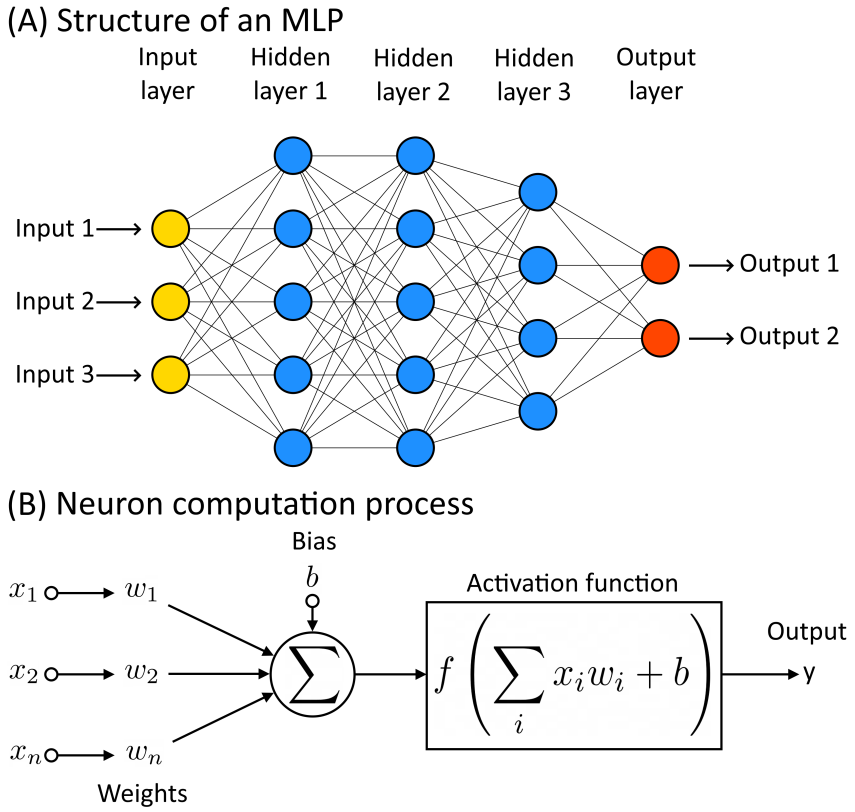


Fig. 2.10 Illustration of an MLP. (A) The overall architecture of an MLP consists of an input layer, multiple hidden layers, and an output layer. Each neuron is fully connected to neurons in adjacent layers. (B) The computational process of a single neuron involves computing a weighted sum of its inputs, adding a bias term, and applying a nonlinear activation function to produce the output.

The final stage, the output layer, produces the network's prediction. The choice of activation function at this stage depends on the specific task. For classification, a softmax function is commonly used to assign probabilities to different categories. In contrast, regression tasks often use a linear activation or no activation at all to produce direct numerical results.

Training an MLP revolves around minimizing prediction errors through an iterative optimization process. A loss function, such as cross-entropy for classification or mean squared error for regression, quantifies how far the network's predictions deviate from the true values. The network then refines its weights using backpropagation (Rumelhart, Hinton, and Williams, 1986), an algorithm that calculates the contribution of each weight to the error by propagating gradients backward through the layers. These gradients guide gradient descent, an optimization step where weights are adjusted incrementally to reduce the loss. The size of these adjustments is controlled by the learning rate, a hyperparameter balancing training speed and stability. Training continues until the model converges or is halted early to prevent overfitting, which is common in medical imaging when dealing with limited datasets.

A key theoretical strength of MLPs is their universal approximation property, which states that with sufficient neurons and layers, they can approximate any continuous function. This theoretical capability shows the power of MLPs as function approximators. Nevertheless, the practical realization of this property depends heavily on the quality and structure of the training data.

Indeed, they are particularly effective for structured, tabular data, such as patient demographics combined with extracted imaging features. However, when dealing with high-dimensional inputs like raw MRI scans or multidimensional microstructural maps, MLPs face significant challenges due to the explosion of parameters. For instance, treating a 3D MRI scan as a flat vector of voxel intensities would lead to enormous weight matrices, making the model computationally impractical and prone to overfitting. Moreover, MLPs lack spatial awareness, meaning that they treat each voxel independently rather than recognizing anatomical structures or spatial dependencies. These limitations have led to the development of specialized architectures like Convolutional Neural Networks or transformers, which handle high-dimensional inputs more efficiently. Despite these constraints, MLPs remain widely used for tasks involving low-to-medium dimensionality and non-sequential data.

2.2.2 Vision transformer and the Shifted window architecture

Quick overview

Vision Transformers (Dosovitskiy, Beyer, Kolesnikov et al., 2020) introduced a new approach to image analysis by applying the self-attention mechanism from natural language processing to visual data by splitting images into patches and capturing both local and global dependencies in a single framework. While ViT excels in modeling long-range interactions, it can be computationally expensive and requires large-scale datasets for optimal performance. The Shifted window (Swin) Transformer (Liu, Lin, Cao et al., 2021) addresses these challenges by applying self-attention within shifted windows, reducing computational overhead and enabling hierarchical feature extraction. This design allows the model to scale efficiently to high-resolution images and maintain a global receptive field, making it suitable for a wide range of vision tasks.

In the field of computer vision, deep learning has revolutionized the way machines interpret and analyze images. Traditionally, Convolutional Neural Networks have been the reference for image processing. CNNs work by applying small, learnable filters across an image, capturing local patterns such as edges, textures, and shapes. These local features are then gradually combined through multiple layers, allowing the network to recognize more complex structures, from simple edges in early layers to entire objects in deeper layers. This hierarchical approach makes CNNs highly effective for tasks like image classification and object detection. However, CNNs inherently focus on local relationships within an image and struggle to capture long-range dependencies without additional mechanisms like very deep architectures or specialized layers.

A major shift in computer vision occurred with the introduction of the Vision Transformer, which borrowed the transformer architecture from natural language processing and adapted it for visual data. Unlike CNNs, which process images through convolutions, ViT treats an image as a sequence of smaller fixed-size patches, such as 16×16 pixels. Each patch is transformed into a numerical representation and combined with positional information to preserve spatial relationships. The key innovation of ViT is its use of self-attention, a mechanism that allows each patch to directly consider all other patches in the image, capturing both local and global depen-

dencies in a single step. This global perspective enables ViT to understand the overall structure of an image more effectively than CNN. However, the downside of this approach is its computational cost: as the image size increases, the number of patches grows, making the self-attention mechanism increasingly expensive in terms of memory and processing power. Additionally, ViT requires very large datasets for training, as it lacks the built-in inductive biases of CNN that help in learning meaningful features from smaller datasets. Figure 2.11 illustrates this fundamental difference in processing, showing how CNNs extract features through localized receptive fields, while ViT models long-range dependencies using self-attention.

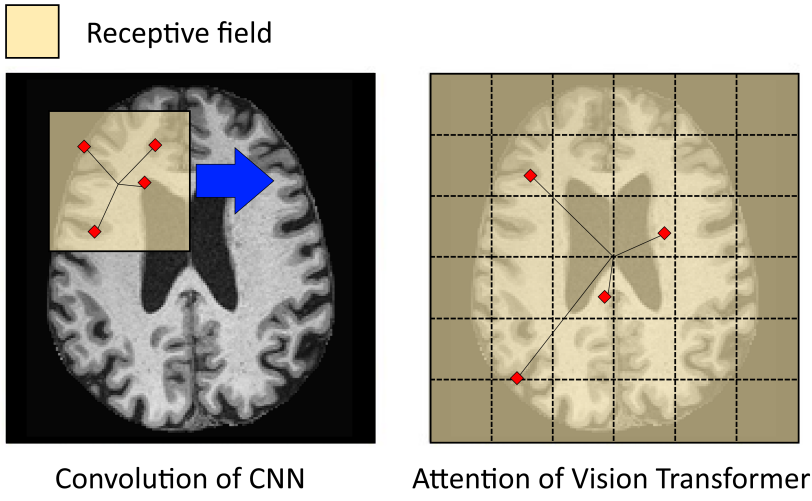


Fig. 2.11 A comparison of image processing in CNNs and Vision Transformers. CNNs (left) rely on localized receptive fields and gradually build hierarchical features, while ViTs (right) use self-attention mechanisms to model long-range dependencies directly.

To overcome these limitations, the Swin Transformer was introduced, offering another way to apply self-attention to images. Instead of computing attention over the entire image at once, Swin Transformer divides the image into smaller windows and applies self-attention within each window separately. This localized attention significantly reduces computational complexity, making the model more scalable for high-resolution images. A key innovation of Swin Transformer is its shifted windowing mechanism, which allows information to flow between different regions of

the image. By slightly shifting the window boundaries between layers, the model ensures that different parts of the image interact over time, improving its ability to capture complex structures. Another advantage of Swin Transformer is its hierarchical structure: as the network progresses, small image patches are gradually merged into larger ones, allowing the model to learn multi-scale representations, similar to how CNNs operate.

Together, ViTs and Swin Transformers represent a shift in computer vision, by using self-attention to analyze images with a global context that CNNs struggle to achieve.

Training and adapting these large transformer models poses significant practical challenges. Architectures like ViT and Swin Transformer typically contain tens to hundreds of millions of parameters distributed across numerous layers. For example, the Swin-Tiny model has approximately 29 million parameters, while the base version has around 88 million. Training these models from scratch requires substantial computational resources, often involving multiple high-end GPUs running for days or weeks, along with large labeled datasets such as ImageNet, which contains over a million images. Consequently, a common practice, especially in fields like medical imaging where large labeled datasets are scarce, is to use models pre-trained on large-scale natural image datasets (e.g., ImageNet) and then fine-tune them on the specific target task. This transfer learning approach significantly reduces training time and data requirements, leveraging the general visual features learned during pre-training.

However, even fine-tuning can be computationally expensive if all model parameters are updated (full fine-tuning). Moreover, fully fine-tuning a large pre-trained model on a smaller target dataset increases the risk of overfitting. To address these challenges, parameter-efficient fine-tuning methods have been developed. One prominent approach is Low-Rank Adaptation (LoRA) (Hu, Shen, Wallis et al., 2021), which targets the large weight matrices W used in key components of the transformer architecture, typically in the attention layers. Instead of updating the entire original weight matrix $W \in \mathbb{R}^{d \times k}$ (where d and k are typically large), LoRA freezes W and introduces a trainable low-rank update. Specifically, it learns two small matrices, $A \in \mathbb{R}^{d \times r}$ and $B \in \mathbb{R}^{r \times k}$, where $r \ll \min(d, k)$. The updated weight matrix is expressed as:

$$W' = W + \Delta W = W + AB$$

where $\Delta W = AB$ is a low-rank approximation to the full-rank update.

By training only A and B, the number of trainable parameters is reduced from $O(dk)$ to $O(r(d+k))$, often leading to reductions by orders of magnitude. This makes fine-tuning faster, more memory-efficient, and less prone to overfitting, while still allowing the model to adapt effectively to new tasks.

3

Motion correction in diffusion MRI acquisition with gradient cycling

Preliminary note:

On the GE Signa Premier MRI scanner jointly owned by UCLouvain and the Cliniques Universitaires Saint-Luc, gradient cycling is used in high b-value multi-shell dMRI sequences to reduce gradient coil heating. While effective, this increases susceptibility to motion artifacts, requiring the use of advanced preprocessing techniques. This chapter introduces a motion correction algorithm tailored for gradient cycling, providing researchers using dMRI sequences with an effective tool to reduce motion artifacts in their data.

Before starting the chapter, readers are encouraged to familiarize themselves with the following basic principles in Chapter 2:

- Principles of diffusion MRI (2.1.1)
- Diffusion Tensor Imaging (DTI) (2.1.2)
- Orientation Distribution Function (ODF) (2.1.3)

Extended version of an abstract published as:

Dessain, Q., Simon, M., Dricot, L. and Macq, B. (2024). Advanced motion correction in diffusion MRI acquisition with gradient cycling. In *International Symposium on Biomedical Imaging (ISBI 2024)*.

3.1 Introduction

Multi-shell diffusion MRI (dMRI) provides an unparalleled window into the microstructural organization of brain tissue, enabling in-vivo insights into properties such as axonal diameter (Assaf, Blumenfeld-Katzir, Yovel et al., 2008) and orientations (Jeurissen, Tournier, Dhollander et al., 2014; Tournier, Calamante, and Connelly, 2007), neurite morphology (Zhang, Schneider, Wheeler-Kingshott et al., 2012), cellular tissue composition (Le Bihan, Breton, Lallemand et al., 1988; Panagiotaki, Chan, Dikaios et al., 2015), and global connectivity patterns (Gong, He, Concha et al., 2009; Tournier, Calamante, and Connelly, 2012). Multi-shell dMRI captures diffusion signals at various b-values, with higher b-values offering finer sensitivity to restricted diffusion, thus enhancing resolution of white matter structures and fiber orientations (Fan, Eichner, Afzali et al., 2022). These acquisitions require a large amount of diffusion gradients directions to accurately model the underlying tissue properties.

However, the use of high b-values and a greater number of gradient directions introduces challenges, including increased thermal load on gradient coils, necessitating extended repetition times (TR) to allow for cooling period. The prolonged TR increases the overall scan duration, potentially affecting subject comfort, exacerbating motion artifacts, and limiting the total number of gradient directions that can be acquired within a session.

To address these limitations, several techniques have been developed to optimize diffusion MRI acquisitions under high thermal loads. The conventional shell-by-shell acquisition method captures entire shells and their associated volumes sequentially based on their b-values, with interleaved b=0 volumes to establish a baseline (Figure 3.2, panel a). While this method is straightforward and widely used, it does not alleviate gradient heating issues. A more effective strategy involves interleaving volumes with high and low b-values instead of doing acquisition one shell at a time (Figure 3.2, panel b). This approach helps distribute the gradient load more evenly and reduces peak heating. Additionally, it offers an advantage over the baseline method by improving the overall spherical coverage of partial scans across all acquired shells (Cook, Symms, Boulby et al., 2007). However, this method may still fall short in acquisition schemes involving high b-values. Slice-level gradient cycling, pioneered by Hutter, Christiaens, Deprez et al. (2017); Hutter, Nilsson, Christiaens et al. (2018), represents

an effective solution. This technique alternates slices with different gradient encodings (intensity and directions) within a single volume, efficiently managing thermal load (Figure 3.2, panel c). While more complex to implement, slice-level interleaving is a robust approach for controlling gradient heating and minimizing the impact of extended repetition times on acquisition efficiency.

However, gradient cycling comes at the cost of increased susceptibility to motion artifacts. In conventional dMRI acquisitions, each volume (acquired over one Repetition Time, TR) uses a single, consistent diffusion encoding for all its slices, allowing straightforward correction of motion occurring within a volume. In gradient cycling, however, each slice within a single volume is acquired with varying diffusion encodings (as shown conceptually in Figure 3.2(C)). This slice-wise encoding variation means that subject movement during acquisition may propagate motion artifacts across multiple volumes, complicating post-reconstruction corrections.

To mitigate this increased sensitivity to intra-volume motion, the pattern of gradient encodings applied slice-by-slice can be structured into repeating blocks that span a short duration, typically the time taken to acquire two or three full volumes (i.e., two or three TRs), instead of having a unique pattern that extends across the entire acquisition (Hutter, Christiaens, Deprez et al., 2017). For example, in a 3TR block configuration (illustrated in Figure 3.2(C)), a specific sequence of diffusion encodings is distributed across the slices of three consecutively acquired volumes. Grouping the cycling into these shorter temporal blocks limits the time window over which motion can introduce inconsistencies among slices intended for the same reconstructed volume, simplifying artifact identification and correction while still retaining significant benefits for thermal load reduction compared to non-cycled acquisitions.

Although this approach mitigates motion-induced distortions to some extent, it does not eliminate the fundamental challenge posed by the non-uniform slice timing introduced by interleaved slice encoding. Indeed, FSL's eddy (Andersson, Graham, Drobnjak et al., 2017), recognized as the state-of-the-art motion correction algorithm in dMRI preprocessing (Tax, Bastiani, Veraart et al., 2022), assume uniform diffusion encoding within each volume and cannot accommodate the variable slice timing inherent in gradient cycling. To address this, we propose a novel adaptation of the eddy motion correction algorithm specifically tailored for acquisitions with slice-level variation in diffusion encoding. This adaptation builds on eddy's framework and introduces modifications to provide an effective so-

lution for motion correction in gradient-cycled acquisitions, preserving the integrity of high-resolution, multi-shell dMRI data by addressing the temporal and spatial complexities introduced by this acquisition technique.

To evaluate the effectiveness of this adaptation, a comprehensive evaluation strategy was employed, combining both in-vivo and synthetic data experiments. First, experiments were conducted on an in-vivo dataset acquired under controlled motion conditions and with multiple different gradient cycling configurations to assess the framework's performance in a realistic setting. These in-vivo tests focused on the visual reduction of artifacts and the impact on downstream diffusion metrics across different gradient cycling configurations and levels of subject movement. Second, to provide a quantitative assessment against a known ground truth, a validation study using realistic synthetic diffusion MRI data with simulated motion was performed. This allowed for direct measurement of the accuracy in recovering motion parameters under various gradient cycling schemes.

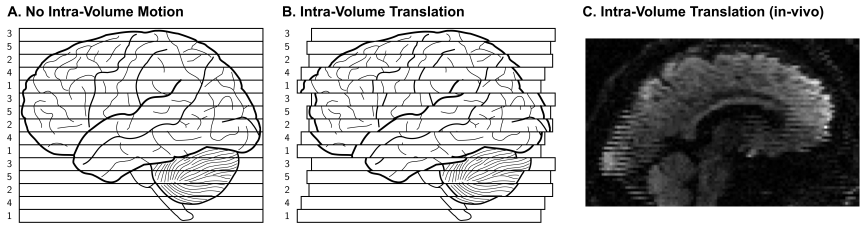


Fig. 3.1 Motion during acquisition causes jagged edges. (A) The outline of a brain with 15 slices acquired using a multiband factor of 3, with the acquisition order indicated by the numbers on the left. (B) The same brain undergoing continuous translation during slice acquisition, leading to misalignments between slices. (C) An actual dMRI scan illustrating the typical motion artifact pattern caused by intra-volume misalignments.

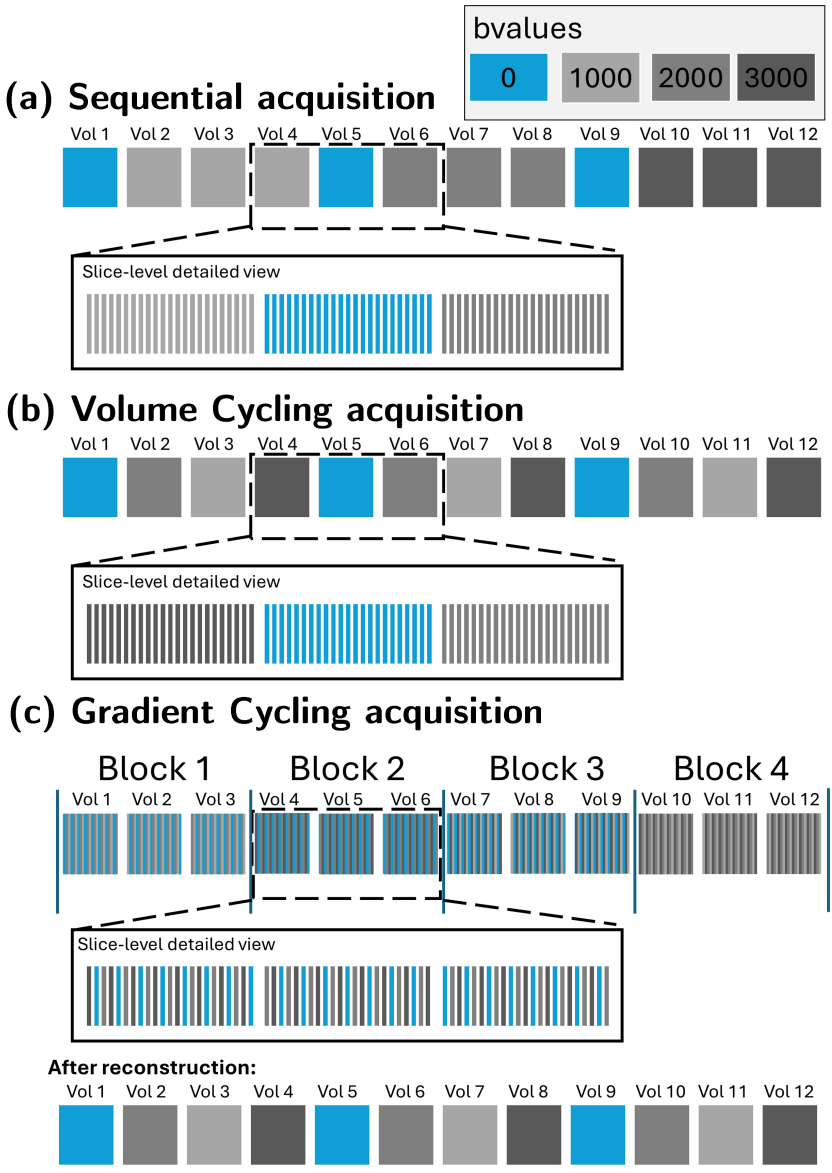


Fig. 3.2 Comparison of three diffusion MRI acquisition strategies to manage thermal load: (a) sequential shell-by-shell acquisition, (b) acquisition with volume interleaving of high and low b-values, and (c) slice-level gradient cycling acquisition with interleaved slices using a 3TR block configuration. Slice-level cycling offers the most efficient control of thermal demands while optimizing acquisition quality.

3.2 Methods

By leveraging a two-step approach, the motion correction framework proposed in this work ensures robust correction of both inter-volume and intra-volume motion artifacts. The first step involves volume-to-volume (V2V) correction, which aligns entire volumes and corrects gross motion and global distortions. The second step introduces slice-to-volume (S2V) correction, which resolves finer intra-volume motion effects that occur within individual reconstructed volumes. These artifacts arise when slices within a single volume are misaligned due to subject motion during acquisition, as illustrated in Figure 3.1. Gradient cycling further exacerbates this issue by introducing interleaved slice encodings and non-uniform acquisition timing.

To maintain consistency with standard preprocessing workflows, the motion correction framework starts with the established V2V algorithm implemented in FSL's eddy tool. This step provides initial estimates of motion and distortion parameters, laying the groundwork for finer intra-volume corrections. The subsequent S2V correction adapts and extends the original eddy's framework to incorporate volume-specific slice timing and interleaved encoding. By modeling motion as a continuous function of time and leveraging Gaussian process predictions, the adapted algorithm ensures precise alignment at the slice level, even in the presence of complex motion patterns and acquisition variability.

The combined V2V and S2V strategy is designed to handle the specific challenges introduced by gradient cycling, providing a robust pipeline for motion correction. The following sections detail each step of the algorithm, beginning with the inter-volume correction and progressing to the slice-to-volume adaptation.

3.2.1 Inter-volume motion correction (V2V)

The first step in the motion correction framework is inter-volume correction, which addresses motion and distortions that occur between the acquisitions of successive volumes. In the case of gradient cycling, this algorithm will correct the motion between reconstructed volume and not the motion in the acquired slice varying diffusion encoding volumes. After acquisition, the MRI system reconstructs the data such that each vol-

ume represents a single diffusion encoding. The method used for this step is the FSL's eddy inter-volume motion correction algorithm (Andersson and Sotiropoulos, 2016) which provides a robust foundation for subsequent intra-volume corrections by simultaneously handling susceptibility-induced distortions, eddy current-induced distortions, and rigid-body motion.

To achieve this, FSL's eddy aligns all volumes to a common space while estimating distortion fields in parallel. This process relies on a model-driven approach where the observed volumes are iteratively registered to a predicted reference space generated using Gaussian Process (GP) modeling (Andersson and Sotiropoulos, 2015).

Gaussian process prediction for distortion and motion estimation The Gaussian Process framework is central to the inter-volume correction process. It predicts how an undistorted version of each volume should appear, based on the smooth variation of the diffusion signal across b-values and gradient directions. The GP model interpolates the diffusion signal using observed data from all volumes, ensuring that the predictions are consistent with the acquired diffusion encoding. This prediction allows the algorithm to estimate how the volume should look in the absence of motion and distortions, providing a target for registration.

Simultaneous estimation of motion and distortions The motion and distortion parameters are optimized by minimizing the residual differences between the predicted and observed volumes. This involves:

- **Rigid-body motion estimation**: Each volume's rigid-body motion parameters (translations and rotations) are estimated to align it with the predicted reference space.
- **Distortion field modeling**: The susceptibility-induced distortion field $h(x, y, z)$ and the eddy current-induced distortion field $e(x, y, z)$ are estimated simultaneously. These fields are modeled as:

$$\phi(x, y, z) = h(x, y, z) + e(x, y, z), \tag{3.1}$$

  where $\phi(x, y, z)$ represents the total displacement of each voxel caused by the combined effects of susceptibility and eddy current. To prevent overfitting and ensure physical plausibility, the distortion fields are constrained. While the susceptibility field h is often

considered fixed (estimated separately), the eddy current field e is parameterized using a low-order basis, typically spatial polynomials (linear, quadratic, or cubic) (Andersson and Sotiropoulos, 2016). This means the optimization estimates a small set of coefficients for these smooth basis functions, rather than independent displacements for every voxel. This parameterization enforces spatial smoothness on the estimated eddy current field, preventing unrealistic distortions from being introduced to simply minimize the cost function.

Iterative optimization The correction process is iterative and consists of the following steps:

1. **Prediction**: Generate a predicted undistorted volume using the GP model.
2. **Alignment**: Align the observed volume to the predicted reference by estimating motion parameters r and distortion fields ϕ.
3. **Update**: Update the prediction based on the refined motion and distortion estimates.
4. **Convergence**: Repeat the prediction-alignment-update cycle until the residual differences between observed and predicted volumes converge.

At each iteration, the motion parameters r and distortion fields ϕ are updated using a Gauss-Newton optimization scheme:

$$\begin{bmatrix} r^{(k+1)} \\ \phi^{(k+1)} \end{bmatrix} = \begin{bmatrix} r^{(k)} \\ \phi^{(k)} \end{bmatrix} - \left(\mathbf{J}^T\mathbf{J}\right)^{-1}\mathbf{J}^T\mathbf{r}, \tag{3.2}$$

where $\mathbf{J}$ is the Jacobian matrix of partial derivatives of the predicted signal with respect to the motion and distortion parameters.

Benefits of V2V correction The inter-volume correction step ensures that all volumes are aligned in a common space and that global distortions are mitigated. This provides a consistent starting point for the more complex intra-volume corrections required for gradient cycling acquisitions. By leveraging the predictive power of Gaussian Processes and a unified framework for motion and distortion correction, this step achieves sub-millimeter alignment precision and ensures that the distortion fields are estimated with high fidelity.

In the subsequent section, we build on this foundation by introducing slice-to-volume correction, which addresses the intra-volume motion artifacts common in gradient cycling acquisitions.

3.2.2 Intra-volume motion correction (S2V)

While the inter-volume correction described previously aligns whole volumes, it assumes the subject is perfectly still during the acquisition of all slices within a single volume. Subject motion during this acquisition period leads to intra-volume artifacts, often visible as jagged edges when slices are stacked (Figure 3.1). The S2V correction, originally presented by Andersson, Graham, Drobnjak et al. (2017), addresses this by modeling motion as a continuous process occurring throughout the volume acquisition time. Our implementation adapts this framework, particularly considering the variable slice timing introduced by gradient cycling.

Modeling intra-volume motion Intra-volume motion is modeled as a smooth, continuous function of time using a discrete cosine transform (DCT) basis. This approach captures the temporal dynamics of motion during the acquisition of the volume's slices. The 6 rigid-body motion parameters at time t are represented by the vector $\boldsymbol{r}(t)$, expressed as:

$$\boldsymbol{r}(t) = \sum_{k=0}^{M-1} \boldsymbol{b}_k \cos\left(\frac{k\pi t}{T}\right) \tag{3.3}$$

where M is the number of DCT basis functions, $\boldsymbol{b_k}$ are the DCT coefficients vector estimated during the optimization process and T is the total acquisition duration for the volume. This representation ensures that motion parameters vary smoothly over time, with regularization typically applied to penalize high accelerations (large second derivatives of $\boldsymbol{r}(t)$), improving stability during optimization.

Accounting for gradient cycling slice timing To apply the continuous motion model, we need to know precisely when each slice (or slice group in multi-band) was acquired. For gradient cycling, the acquisition sequence interleaves slices with varying diffusion encodings. After reconstruction, volumes are organized such that each has a single diffusion encoding. However, the original interleaved slice timing persists. To address this, the proposed algorithm introduces a volume-specific slice timing function

$Z_i(t_n)$, which maps acquisition times t_n to their corresponding slice indices within each reconstructed volume i, where it is assumed that the volume was acquired at N discrete time-points (these would correspond to slices or multi-band groups). This function ensures that motion correction aligns with the true temporal order of slice acquisition and can be expressed as:

$$\text{slice index(es)} = Z_i(t_n) \tag{3.4}$$

This function $Z_i(t_n)$ is crucial as it links the generic time t in the DCT motion model (Eq. 3.3) to the specific slices being acquired at time t_n within the context of volume i's reconstruction, correctly handling the interleaved nature of gradient cycling.

Iterative correction process Similar to the V2V correction, S2V operates within an iterative prediction-registration framework. The key difference lies in how the alignment step estimates and uses the motion parameters. The core steps within each iteration for a given volume i are:

1. **Prediction:** As in V2V, a GP model predicts what an undistorted version of the current volume i should look like based on all other volumes, leveraging the expected smooth signal variation in q-space.
2. **Motion parameter estimation:** The goal is to find the optimal DCT coefficients $\{\boldsymbol{b}_k\}_{k=0}^{M-1}$ that best explain the observed (motion-corrupted) data. This is done by minimizing the difference between the observed data and the predicted data, accounting for the slice-wise motion. Conceptually, the optimization tries to find the motion parameters $\{\boldsymbol{b}_k\}$ such that if we were to uncorrect the GP prediction slice-by-slice using the motion model $\boldsymbol{r}(t)$, the result would best match the actually observed volume.

   During the optimization, the current estimate of the motion $\boldsymbol{r}_i(t_n)$ (derived from the current estimate of $\{\boldsymbol{b}_k\}$) is used to transform the relevant slice data. For each acquisition time point t_n:

   - The motion state $\boldsymbol{r}_i(t_n)$ is calculated using Eq. 3.3.
   - The corresponding rigid-body transformation matrix $\boldsymbol{P}(\boldsymbol{r}_i(t_n))$ is formed.
   - The slice(s) acquired at this time, identified by $Z_i(t_n)$, are considered. The transformation is used either to warp the predicted

data to match the observed space or to align the observed slices to the model space during the residual calculation.

3. **Parameter update:** The residuals calculated across all slices and time points are used to update the DCT motion coefficients $\{\boldsymbol{b}_k\}$ via a Gauss-Newton optimization step.
4. **Iteration and convergence:** The process (GP prediction -> Motion estimation -> Parameter update) repeats until convergence or for a fixed number of iterations.

This iterative process refines the estimate of the continuous motion trajectory $\boldsymbol{r}_i(t)$ for each volume.

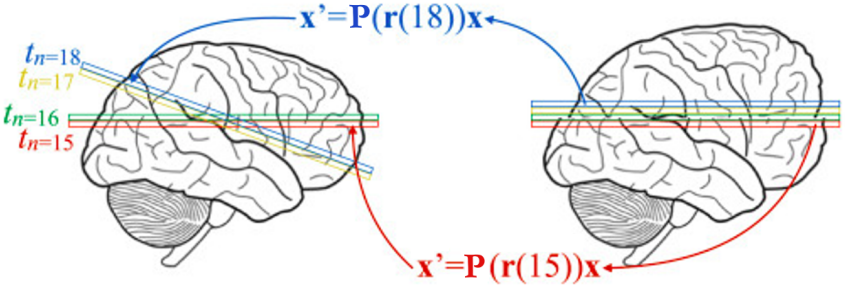


Fig. 3.3 Illustration of the intra-volume movement problem and its correction. The left panel shows the true slice positions before and after a subject movement occurs during the acquisition of slices. The right panel demonstrates the reconstructed volume without correction, where slices are misaligned due to the unaccounted motion. This misalignment creates a distorted representation of the brain volume. The intra-volume correction algorithm addresses these misalignments by modeling motion as a continuous function over time and realigning slices to their true spatial positions, preserving the anatomical integrity of the reconstructed volume. Adapted from Andersson, Graham, Drobnjak et al. (2017).

Final Resampling into Corrected Volume Finally, once the optimal motion parameters $\{\hat{\boldsymbol{b}}_k\}$ are estimated, the corrected slices are reconstructed into a motion-corrected volume using hybrid resampling. Each acquired slice is first transformed according to its specific estimated motion $\hat{\boldsymbol{r}}(t_n)$. Then, 2D in-plane resampling interpolates the transformed slice data onto a regular grid within its plane. Lastly, 1D out-of-plane resampling along

the slice direction uses the correctly positioned slice data (which may be irregularly spaced due to motion) to generate the final volume on a regular 3D grid. This hybrid 2D+1D approach minimizes interpolation artifacts while preserving spatial resolution.

By using volume-specific slice-timing modeling ($Z_i(t_n)$) and slice-specific motion modeling ($\boldsymbol{r}(t)$), the algorithm ensures precise realignment, enabling accurate downstream analyses such as advanced microstructural models and tractography.

3.2.3 Experiments

The proposed motion correction framework was evaluated using a realistic synthetic dataset with known motion parameters and an in-vivo dataset acquired under controlled motion scenarios. This dual approach allowed for both quantitative validation against a ground truth and qualitative assessment in a real-world setting. The evaluation focused on comparing the performance of the standard volume-to-volume correction algorithm with the enhanced slice-to-volume correction method, specifically adapted for gradient cycling.

Synthetic Experiments

To quantitatively validate the motion correction framework against a known ground truth, realistic synthetic diffusion MRI datasets were generated. Motion-free baseline data were created using FSL's DW-POSSUM (Graham, Drobnjak, and Zhang, 2016), an extension of POSSUM (Drobnjak, Gavaghan, Süli et al., 2006) capable of simulating diffusion MRI signals with realistic noise properties and artifacts. The diffusion sequence simulated included 12 reference b0 images, alongside 2 PGSE shells distributed isotropically on the sphere with 32 directions at $b = 700$, and 64 at $b = 2000\,\mu s\,\mu m^{-2}$. The acquisition used a matrix of $72 \times 86 \times 54$ with isotropic voxel size of 2.5 mm, a TR of 2700 ms, and a multiband factor of 3.

A pre-defined, realistic motion trajectory was simulated to mimic typical subject movement patterns. This trajectory was generated as a combination of two components superimposed over time: (1) a continuous slow drift representing gradual changes in head position, modeled using cosine functions, and (2) intermittent, sharp, transient movements (spikes)

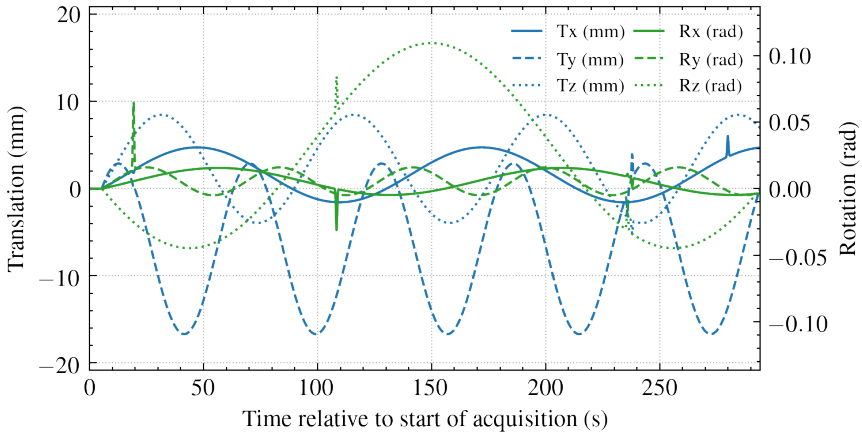


Fig. 3.4 Illustration of the simulated motion trajectory applied to the synthetic DW-POSSUM data. The plots show the six rigid-body parameters varying over the simulated acquisition time. This trajectory combines slow drift and sharp transient movements.

simulating sudden adjustments, modeled as Gaussian-shaped events occurring probabilistically. Both components affected translational and rotational axes. This known motion trajectory was applied post-simulation to the motion-free DW-POSSUM output volumes to create motion-corrupted data.

To investigate the interaction between motion and different acquisition strategies, this process was used to generate four distinct synthetic datasets:

1. **No Cycling**: Motion applied to a standard sequential acquisition.
2. **ALL TR blocks**: Motion applied considering gradient cycling across all TRs.
3. **Three TR blocks**: Motion applied considering a gradient cycling scheme with blocks of three TRs.
4. **Two TR blocks**: Motion applied considering a gradient cycling scheme with blocks of two TRs.

The effective ground truth motion for evaluating the correction algorithms was defined as the motion trajectory after the gradient cycling reconstruction process transformed the raw applied motion. For compatibility with the TR block grouping in the cycling schemes, the number of

b=2000 directions was slightly adjusted to 63 for the 2TR simulation and 62 for the 3TR simulation.

The performance of the V2V and S2V correction algorithms was then evaluated on these synthetic datasets by quantitatively comparing the estimated motion parameters against the known ground truth parameters (shown in Figure 3.4) and qualitatively assessing the reduction of motion artifacts in the corrected volumes compared to the motion-free reference.

In-Vivo Experiments

Complementing the quantitative synthetic validation, the in-vivo experiments aimed to qualitatively assess the framework's effectiveness in mitigating motion artifacts under realistic scanning conditions and its impact on downstream analyses. The evaluation comprised visual inspections of multiple diffusion MRI modalities. The first inspection focused on corrected volumes, examining reductions in common motion-related artifacts such as jagged edges, intensity modulations, and misaligned slices. The second inspection assessed fiber orientation distributions (FOD) generated using multi-shell multi-tissue constrained spherical deconvolution (MSMT-CSD) (Jeurissen, Tournier, Dhollander et al., 2014) from V2V- and S2V-corrected data, with comparisons of FOD continuity and artifact suppression providing insights into the impact of motion correction on downstream analyses. Lastly, colored fractional anisotropy (FA) maps derived from the diffusion tensor imaging (DTI) model (Basser, Mattiello, and LeBihan, 1994) were visually inspected to evaluate how motion correction influenced the representation of white matter structures, with attention to structural consistency and continuity across slices. For the FA analysis, only the $b = 0$, $b = 1000$, and $b = 2000\,\mathrm{s/mm^2}$ shells were used, as higher b-value shells, such as $b = 3000\,\mathrm{s/mm^2}$, violate the Gaussian diffusion assumptions underlying DTI (Jensen, Helpern, Ramani et al., 2005; Jones, Knösche, and Turner, 2013).

MRI acquisition MRI data were acquired on a 3T GE SIGNA Premier scanner using an identical diffusion scheme across all experiments. A single experienced volunteer participated in multiple scans within the same session. The acquisition was divided into an alternance of the two following phases for each scan:

1. **Motionless phase**: The subject was instructed to remain completely

still throughout the scan. This phase served as the baseline, representing ideal conditions with minimal motion artifacts.

2. **Controlled motion phase**: The subject performed deliberate movements, ranging from small translations and rotations to larger, more pronounced motions.

The MRI session included a T1-weighted anatomical scan and three diffusion MRI sequences, alongside a reverse-phase b=0 acquisition for distortion estimation. The T1-weighted anatomical image was obtained with the following parameters: echo time (TE) = 2.96 ms, TR = 2238.93 ms, inversion time (TI) = 900 ms, 170 slices, 1 mm isotropic, in-plane FOV: 256 × 256 mm^2. The diffusion-weighted sequences were acquired using a multi-shell PGSE protocol with 9 reference b0 images, and 4 PGSE shells of 32 directions at $b = 1000$ ($G = 69\,mT\,m^{-1}$), 64 at $b = 2000$ ($G = 120\,mT\,m^{-1}$) and 64 at $b = 3000\,\mu s\,\mu m^{-2}$ ($G = 219\,mT\,m^{-1}$) with gradient duration $\delta = 12.9\,ms$ and diffusion time $\Delta = 21.8\,ms$. The acquisition also used the following parameters: TE = 70.4ms, 2 mm isotropic voxels, in-plane FOV: 220 × 220 mm^2, 110x110x69 matrix and a multiband factor of 3.

To assess the influence of gradient cycling on motion correction, three distinct gradient cycling configurations were implemented:

1. **ALL TR blocks**: Gradient cycling applied across all volumes with a TR of 3584 ms. The total acquisition time was 10 minutes and 9 seconds.
2. **Three TR blocks**: Gradient cycling limited to blocks of three TRs with TR = 3079 ms. The total acquisition time was 8 minutes and 43 seconds.
3. **Two TR blocks**: Gradient cycling applied in two TR blocks with TR = 3957 ms. The total acquisition time was 11 minutes and 13 seconds.

No acquisitions were performed without gradient cycling due to the prohibitive scan time required.

MRI preprocessing Preprocessing of diffusion MRI data began with MP-PCA denoising (Veraart, Novikov, Christiaens et al., 2016), applied through MRtrix3's dwidenoise function (version 3.0.3) (Tournier, Smith, Raffelt et al., 2019). Following denoising, the FSL topup tool leveraged two b=0 images acquired with opposing phase encoding directions to estimate and correct susceptibility induced distortions. Next, depending

on the experiment, either the in-house adapted for gradient cycling version or the original version with only inter-volume correction of the FSL's eddy tool was employed for both head motion correction and eddy current correction. Finally, to ensure consistency between the applied diffusion weighting directions and the motion-corrected image data, the corresponding rotation component estimated by eddy was applied to the gradient vectors (b-vectors). These reoriented b-vectors were subsequently used for all diffusion model fitting.

3.3 Results

The efficacy of the proposed slice-to-volume motion correction framework was evaluated against the standard volume-to-volume correction approach using both synthetic data with known ground truth motion and in-vivo data with controlled motion.

3.3.1 Synthetic results

Table 3.1 S2V correction yields lower error in displacement field estimation compared to V2V correction. Error in displacement field estimates between the S2V and V2V motion correction methods and the ground truth displacement field on synthetic data. Errors are measured in millimeters (mm).

Configuration	S2V Error (mm) (Mean ± Std Dev)	V2V Error (mm) (Mean ± Std Dev)
No cycling	0.19 ± 0.16	0.46 ± 0.19
2 TR blocks	0.28 ± 0.19	0.47 ± 0.15
3 TR blocks	0.34 ± 0.24	0.42 ± 0.20
ALL TR blocks	2.81 ± 1.03	5.33 ± 3.32

The synthetic data experiments provided a quantitative means to assess the accuracy of both correction methods against a known ground truth motion trajectory (shown in Figure 3.4). Table 3.1 presents the primary quantitative results, comparing the mean voxel displacement error magnitude (in millimeters) for both S2V and V2V correction methods relative

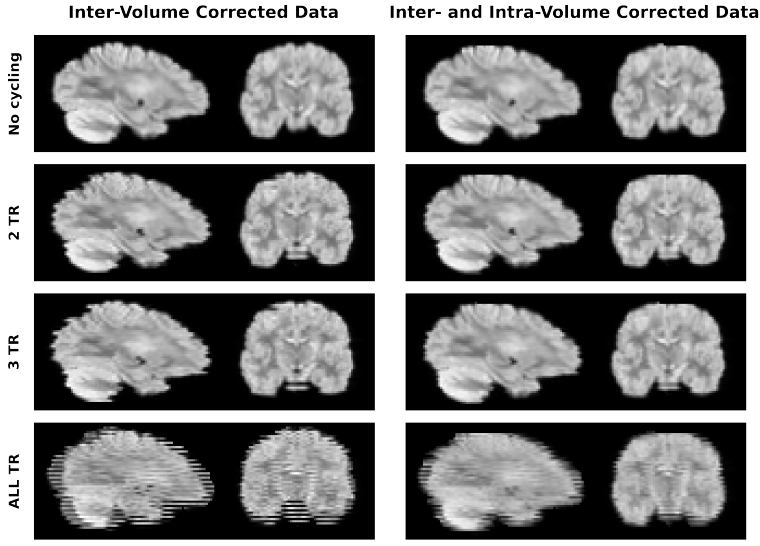


Fig. 3.5 S2V correction significantly reduces motion artifacts in synthetic data, approaching ground truth quality. Representative axial (or sagittal/coronal) slices from the synthetic dataset ($b = 700\,\mu s\,\mu m^{-2}$).

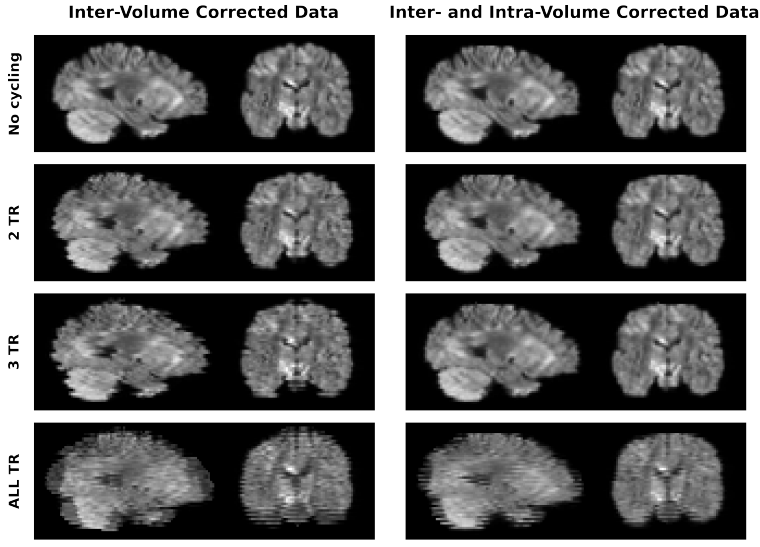


Fig. 3.6 S2V correction significantly reduces motion artifacts in synthetic data, approaching ground truth quality. Representative axial (or sagittal/coronal) slices from the synthetic dataset ($b = 2000\,\mu s\,\mu m^{-2}$).

to the ground truth across the different gradient cycling configurations. Across all simulated conditions, the S2V correction consistently yielded lower mean displacement errors compared to the standard V2V approach. This indicates that S2V is more effective at compensating for the complex motion patterns interacting with the gradient cycling reconstruction, leading to a final corrected volume that is spatially closer to the true, motion-free state.

Visual comparison of the corrected synthetic volumes further confirmed the quantitative results. Figure 3.6 and Figure 3.5 present representative slices from the synthetic dataset, comparing the V2V corrected volume and the S2V corrected volume. The S2V correction resulted in volumes with less visible artifacts. However, residuals artifacts persist in the ALL TR configuration.

3.3.2 In-vivo results

Visual inspection of the in-vivo diffusion MRI data revealed benefits of the S2V correction. Representative sagittal and coronal views of preprocessed diffusion MRI volumes acquired under gradient cycling are shown in Figure 3.7 for b-values of $2000\,\mathrm{s/mm^2}$. The comparison highlights the impact of the S2V correction in reducing motion artifacts compared to the standard V2V correction. Artifacts such as jagged edges and misaligned slices, more prominent in V2V-corrected volumes, are mitigated with the S2V correction. However, residual artifacts persist in the ALL TR configuration, even after S2V correction. Additional results for volumes with b-values of $1000\,\mathrm{s/mm^2}$ and $3000\,\mathrm{s/mm^2}$ are provided in Appendix A.

Beyond the raw diffusion-weighted volumes, downstream diffusion metrics offer further evidence of improved data quality with S2V. Figure 3.8 illustrates the impact of motion correction methods on colored FA maps derived from DTI model. The maps provide a visual representation of anisotropy and orientation within white matter pathways, highlighting differences between V2V and S2V corrections. In the V2V-corrected maps (leftmost panels), misaligned slices disrupt the continuity and clarity of white matter structures, particularly in regions with complex fiber configurations. In contrast, the S2V-corrected maps (rightmost panels) show improved structural consistency.

Similarly, fiber orientation distributions reconstructed with MSMT-CSD show that reducing slice misalignment directly improves angular

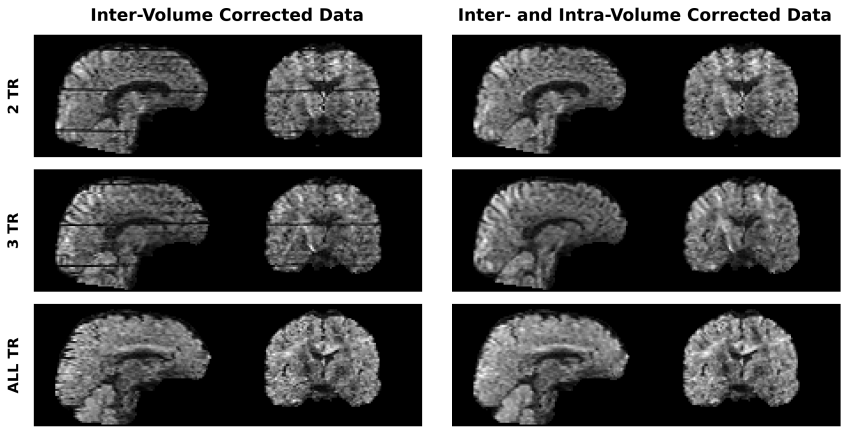


Fig. 3.7 S2V correction reduces motion artifacts and improves slice alignment compared to V2V correction. Sagittal and coronal views of preprocessed diffusion MRI volumes acquired under gradient cycling for volumes with b-values of 2000 s/mm^2. The leftmost panels display results with V2V correction, while the rightmost panels show results with S2V correction.

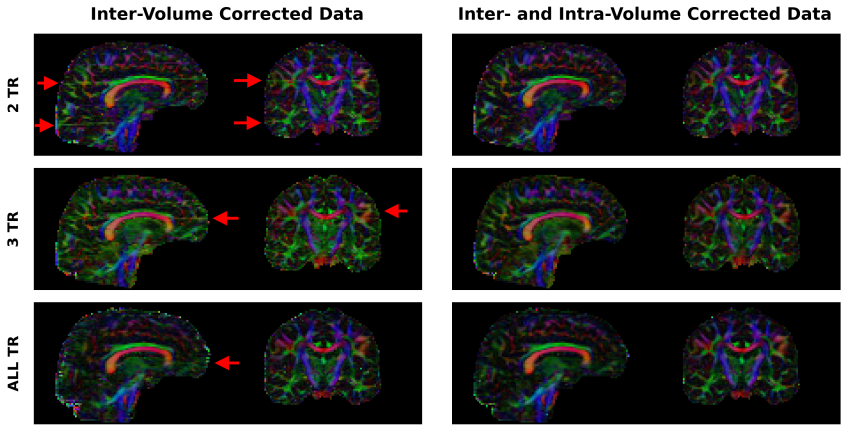


Fig. 3.8 S2V correction improve rgb FA maps compared to V2V correction. Sagittal and coronal views of colored FA maps. The leftmost panels display results with V2V correction, while the rightmost panels show results with S2V correction.

fidelity. Indeed, Figure 3.9 compares fiber orientation distribution maps generated using MSMT-CSD from diffusion MRI data acquired with gradient cycling on 2 TR block and corrected with V2V and S2V methods. The FOD maps in the upper panel, corrected using only the V2V approach, display visible slice artifacts and discontinuities. In contrast, the lower panel, corrected with the S2V method, shows improved FOD quality with smoother orientation continuity and reduced slice-level inconsistencies.

Table 3.2 Computational efficiency of V2V and S2V correction across gradient cycling configurations.

Configuration	V2V Time (min)	S2V Time (min)	Δ Time (min)
2 TR blocks	21	32	+11 (+52%)
3 TR blocks	23	36	+13 (+57%)
ALL TR blocks	43	82	+39 (+91%)

The computational overhead of the proposed S2V correction framework was evaluated across different gradient cycling configurations using a computer equipped with an Intel i7-11700K CPU and an NVIDIA RTX 3080 GPU. As shown in Table 3.2, S2V correction introduced a computational time increase ranging from 52% to 91% compared to the conventional V2V approach, with longer TR block durations resulting in greater processing times. For the 2 TR block configuration, S2V correction required 32 minutes, an increase of 52% over the 21 minutes needed for V2V correction. In the 3 TR block configuration, the processing time increased from 23 minutes (V2V) to 36 minutes (S2V), representing a 57% increase. The most significant overhead was observed in the ALL-TR configuration, where processing time nearly doubled from 43 minutes (V2V) to 82 minutes (S2V), reflecting a 91% increase.

3.4 Discussion and conclusion

The proposed motion correction framework for gradient cycling in diffusion MRI successfully combines inter-volume and intra-volume correction methods to address motion artifacts introduced during acquisition. The results demonstrate that the S2V correction method improves alignment and reduces artifacts compared to the standard V2V approach.

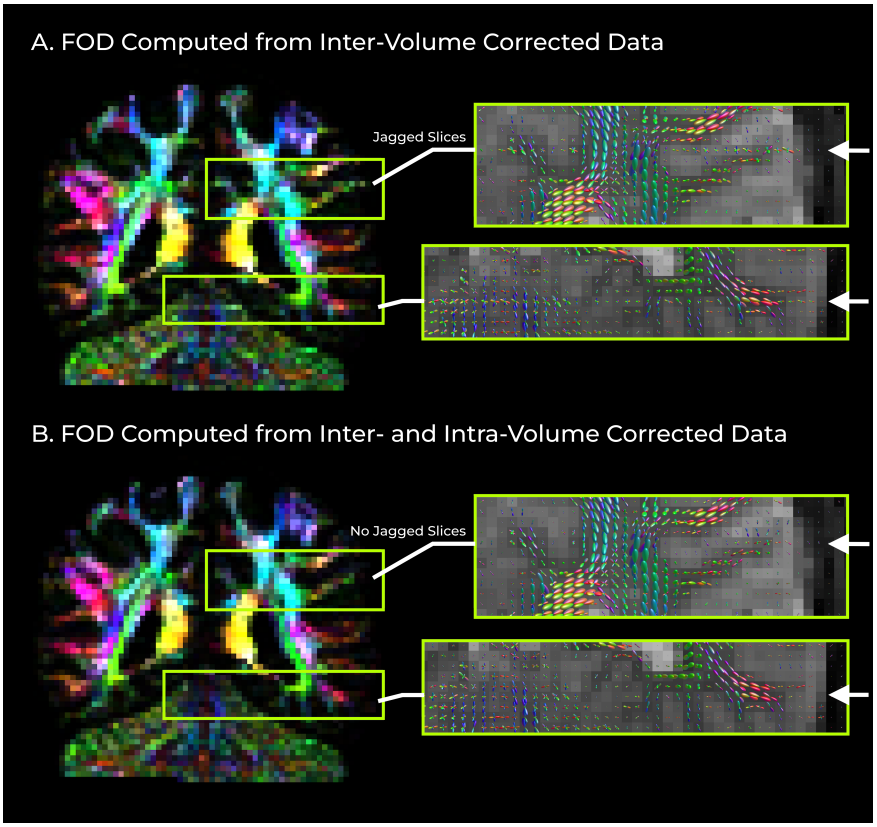


Fig. 3.9 S2V correction enhances FOD continuity and reduces slice artifacts compared to V2V correction. Coronal views of FOD maps generated using MSMT-CSD from diffusion MRI data acquired with gradient cycling using a 2 TR block. The upper panel shows FOD maps from V2V-corrected data, with visible slice artifacts and discontinuities. The lower panel presents FOD maps from S2V-corrected data, displaying smoother orientations and reduced inconsistencies.

In datasets with gradient cycling using shorter TR blocks (2 TR and 3 TR configurations), the S2V correction demonstrated clear improvements over the V2V approach in the in-vivo experiments. These improvements were characterized by reduced jagged edges and enhanced anatomical consistency in volumes affected by motion. This qualitative assessment is strongly supported by the quantitative synthetic data validation, which confirmed S2V's superior accuracy in estimating motion parameters (Table 3.1) compared to V2V when faced with known motion patterns. This improved performance stems from S2V's ability to model motion continuously at the slice level, explicitly accounting for the specific timing introduced by gradient cycling.

However, in configurations where gradient cycling was applied across all TR blocks, the benefits of S2V correction were less pronounced on synthetic data and in-vivo, with more residual artifacts persisting. This outcome can be attributed to two key factors. First, in the ALL TR configuration, motion artifacts are more omnipresent since any movement during acquisition introduces intra-volume artifacts across all volumes, rather than being confined to a single TR block. Second, the extended time interval between the acquisition of slice groups in the ALL TR configuration results in a lack of temporal correlation between slice positions, making it more challenging for the S2V correction to effectively model and compensate for motion. Additionally, these factors likely explain the significant increase in processing time observed in the ALL TR configuration, as the model struggles to converge due to the greater complexity of motion patterns compared to the more temporally constrained 2 TR and 3 TR configurations.

Beyond the primary contribution of the S2V motion correction framework detailed in this chapter, this research also encompassed a crucial upstream optimization: the optimization of the gradient cycling sequence implementation itself. Recognizing that the specific temporal order of diffusion gradients significantly impacts thermal efficiency, we developed and implemented an algorithm to reorder gradients within specified TR blocks. This optimization aims to enhance thermal management beyond standard vendor approaches, minimizing peak thermal load and thereby enabling more efficient sequence timing than default implementations. This optimized gradient ordering was applied to the sequences acquired for the experiments presented, contributing significantly to the feasibility of these demanding acquisitions. The code implementing this gradient cycling sequence optimization is open-source and available on

GitHub (Dessain, 2025).

In conclusion, this chapter advanced the use of gradient cycling in diffusion MRI through two key contributions. Firstly, the development of an open-source gradient reordering algorithm enhances the thermal efficiency and feasibility of the acquisition itself. Secondly, this work presented and validated an adapted motion correction framework (S2V) designed to mitigate motion artifacts in these sequences. Supported by both qualitative in-vivo results and quantitative validation against synthetic ground truth data, the framework demonstrates significant improvements over standard V2V correction, particularly for 2TR and 3TR cycling configurations. By combining V2V pre-alignment with slice-specific S2V refinement that accounts for gradient cycling timing, the framework enhances the quality and reliability of gradient-cycled data. While S2V introduces additional computational cost, the resulting improvement in data integrity makes it a valuable tool for leveraging the benefits of gradient cycling in advanced dMRI studies. Future investigations could explore integrating motion simulation directly within the MRI physics simulation, but the current validation confirms the efficacy of the S2V correction approach.

4

Fast multi-compartment microstructure fingerprinting in brain white matter

Preliminary note:

Before starting the chapter, readers are encouraged to familiarize themselves with the following basic principles in Chapter 2:

- Principles of diffusion MRI (2.1.1)
- Orientation Distribution Function (ODF) (2.1.3)
- Microstructure Fingerprinting (MF) (2.1.5)
- Multi-Layer Perceptron (MLP) (2.2.1)

Adapted version of an article published as:

Dessain, Q., Fuchs, C., Macq, B. and Rensonnet, G. (2024). Fast multi-compartment Microstructure Fingerprinting in brain white matter. *Frontiers in Neuroscience*, 18:1400499.

4.1 Introduction

Quantitative MRI (qMRI) aims to supply objective measurable metrics that specifically depict the morphology, microstructure, and/or chemical composition of tissues in order to provide a deeper knowledge of the physiology of the brain in-vivo (Cercignani, Dowell, and Tofts, 2018). For instance, Magnetic Resonance Fingerprinting (MRF) (Ma, Gulani, Seiberlich et al., 2013) pioneered an approach in which MRI data are acquired with time-varying acquisition parameters (e.g., flip angle, echo time, time of repetition) and then matched to a dictionary of pre-simulated fingerprints with identical parameters time course. Such a process determines the T1, T2, B0 and proton density parameters used to generate the fingerprint. This not only enables a very fast estimation of T1, T2, B0, and proton density maps, reduced to a simple dictionary matching, but also allows the use of the Bloch Simulator (Doneva, Börnert, Eggers et al., 2010), which is one of the most accurate ways to model the physics of MRI. The MRF model has then been extended to account for more complex tissue parameters such as B1+, T2*, perfusion (Su, Mao, Liu et al., 2017; Wright, Jiang, Ma et al., 2018), hemodynamic (Christen, Pannetier, Ni et al., 2014; Lemasson, Pannetier, Coquery et al., 2016) and diffusion (Cao, Liao, Zhou et al., 2024) related properties. However, this requires the use of ever growing dictionaries which renders the dictionary matching step of MRF very time consuming due to the considerable amount of fingerprints required.

Subsequent advancements demonstrated the feasibility of extracting morphological characteristics of brain tissues using diffusion-weighted magnetic resonance imaging (DW-MRI) by matching the observed DW-MRI signal to a simulated signal counterpart mirroring predefined morphological features (Palombo, Ligneul, Najac et al., 2016; Rensonnet, Jacobs, Macq et al., 2015). This approach inspired Microstructure Fingerprinting (Rensonnet, Scherrer, Girard et al., 2019), which uses a pre-calculated dictionary to estimate microstructural properties of the white matter from DW-MRI. Instead of the Bloch equations, a Monte Carlo simulator solves the Bloch-Torrey equation to simulate diffusion fingerprints in a variety of white matter tissue configurations, flexibly incorporating information about the axonal density, diameter, principal orientation, their geometries, and sizes. MF was designed to deal with multi-compartment models from the beginning because of the nature of the white matter,

composed of multiple populations of axons intersecting throughout the brain. When the fascicles overlap in a voxel, the model assumes that the total diffusion-weighted MRI signal is the sum of the signals originating from each fascicle independently (Rensonnet, Scherrer, Warfield et al., 2018). This allows the use of multiple small dictionaries instead of a much larger one, accounting for all the compartment orientations and fractions, considerably decreasing the time and memory requirements by reducing the problem size. However, selecting the right combination of fingerprints remains combinatorial. The runtime complexity of MF is therefore affected by the number of fascicles in a voxel as $\mathcal{O}\left(K \cdot N^K\right)$, where N represents the number of elements in the (single-fascicle) dictionary and K the number of fascicles in the voxel (Rensonnet, Adam, and Macq, 2021). This severely limits the size and thus the complexity of the precalculated dictionary compared to the theoretically limitless possibilities of Monte Carlo simulation tools.

Kiselev, Körzdörfer, and Gall (2021) highlighted the importance of physiology-based models and the potential synergy between microstructure MRI and magnetic resonance fingerprinting for achieving specific, target-oriented diagnostic tools. Aligning with these recommendations, other advanced models similar to MF have been developed. For instance, Filipiak, Shepherd, Lin et al. (2022) employed a dictionary of presimulated Orientation Distribution Functions based on a biophysically plausible multicompartment diffusion model to accurately estimate fiber crossings with shallow angles.

In both MRF (always dictionary-based) and DW-MRI estimation (either dictionary-based or not), deep neural networks (DNN) are increasingly used to accelerate estimations of complex tissue models. Indeed, in order to find hierarchical representations that can handle challenging tasks with large-scale datasets, DNNs have proven to be a very effective tool (Najafabadi, Villanustre, Khoshgoftaar et al., 2015). As illustrated in Table 4.1, a number of studies have proposed deep learning methods to bypass long non-linear models fitting used in dMRI and to circumvent lookup in dictionaries of ever-increasing sizes in MRF.

Simulated data have been used for training (de Almeida Martins, Nilsson, Lampinen et al., 2021; Cai, Wang, Zeng et al., 2018; Karimi, Vasung, Jaimes et al., 2021; Nedjati-Gilani, Schneider, Hall et al., 2017; Sabidussi, Klein, Caan et al., 2021) because they allow any combination of parameter values and facilitate the creation of substantial datasets. Moreover, simulated data offer a ground truth which is not available using real data.

When generating the synthetic dataset, Sabidussi, Klein, Caan et al. (2021), de Almeida Martins, Nilsson, Lampinen et al. (2021) and Karimi, Vasung, Jaimes et al. (2021) opted for a closed form model based on several assumptions, while Cai, Wang, Zeng et al. (2018) used the Bloch equations to numerically generate simulated data to train CNNs (LeCun, Bottou, Bengio et al., 1998) used for T1 and T2 mapping. By using a numerical simulator instead of a closed-form model, an enhanced biophysical accuracy is obtained at the cost of an increased use of computational resources.

Ye (2017) and Ye, Li, and Chen (2019) proposed a two-stages method for the acceleration of the estimation of the NODDI parameters (Zhang, Schneider, Wheeler-Kingshott et al., 2012). The first stage is similar to a solution to a dictionary-based sparse reconstruction problem while the second stage computes the final NODDI microstructure estimation. They jointly learn the weights of the two stages by minimizing the mean squared error (MSE) of microstructure estimation. The separation of the problem into multiple steps instead of using an end-to-end DNN allows them to reduce the complexity of the problem and accelerate training.

In this paper, we propose two methods aimed at accelerating DW-MRI microstructure estimation using DNN. The first method, the "Hybrid Method", exploits the power of presimulated Monte Carlo dictionaries by projecting the measured DW-MRI signal y into an interpretable latent space, representing the signal in terms of dictionary fingerprints, before providing it to a feed-forward neural network. This allows for generalization to different acquisition schemes with little to no additional training. The second method, the "Fully-Learned Method", is based on a feed-forward neural network that uses the spherical harmonics representation of the DW-MRI signal as input to jointly estimate microstructural features and fascicles orientations. This method is less interpretable than the Hybrid Method but is significantly faster since no minimization step is required.

We evaluated each method in terms of the inference speed and accuracy of their fittings on synthetic and in-vivo datasets. First, we estimated the accuracy and noise robustness via experiments on simulated data with varying noise levels. Second, we performed a second accuracy and noise robustness assessment on the Hybrid Method with data generated from an unseen acquisition protocol to evaluate its generalizability. Finally, we tested the quality of each method to retain small structures within the brain, using 34 in-vivo scans from the MGH-USC young adult cohort (Van Essen, Ugurbil, Auerbach et al., 2012).

Table 4.1 Fast estimation of tissue parameters from magnetic resonance measurements has relied on a variety of machine and deep learning approaches, usually assuming one main population of axons. Comparison of state-of-the-art methods.

	Trained on synthetic data	Trained on real data	Neural Network architecture	Type of data simulator	Type of acquisition protocol	# of axonal populations per voxel
Cai, Wang, Zeng et al. (2018)	✓	✗	CNN	Bloch equations	MRF	
Nedjati-Gilani, Schneider, Hall et al. (2017)	✓	✗	Random Forest regression	Monte Carlo simulator	dMRI	1
de Almeida Martins, Nilsson, Lampinen et al. (2021)	✓	✗	MLP	Standard Model of diffusion with relaxation	dMRI	1
Sabidussi, Klein, Caan et al. (2021)	✓	✗	Recurrent Inference Machines	Model based generation framework with gaussian noise	CINE	1
Karimi, Vasung, Jaimes et al. (2021)	✓	✓	MLP with ReLU	Based on a multi tensor model	dMRI	1
Hill, Palombo, Santin et al. (2018)	✓	✗	MLP	Monte Carlo simulator	dMRI	1
Ye, Li, and Chen (2019)	✗	✓	Two-stage architecture using a latent space	Not Applicable	dMRI	1
Proposed Hybrid Method	✓	✗	Two-stage architecture using a latent space	Monte Carlo simulator	dMRI	2
Proposed Fully-Learned Method	✓	✗	Spherical harmonics decomposition + MLP	Monte Carlo simulator	dMRI	2

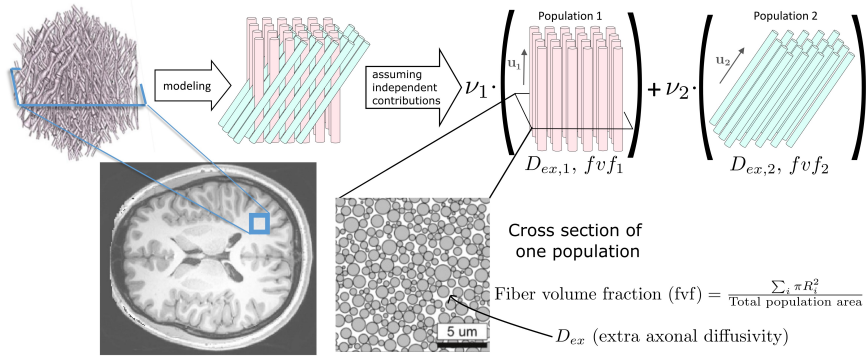


Fig. 4.1 Our forward signal model exploits Monte Carlo numerical simulations. A voxel of white matter (artistic view borrowed from Ginsburger, Poupon, Beaujoin et al. (2018)), located here on a T1 anatomical scan of a healthy young adult from the Human Connectome Project (Van Essen, Ugurbil, Auerbach et al., 2012), is represented by axon populations modeled as straight cylinders with signal contributions assumed independent. The DW-MRI signal is obtained by Monte Carlo simulations in each population (Hall and Alexander, 2009; Rensonnet, Jacobs, Macq et al., 2015).

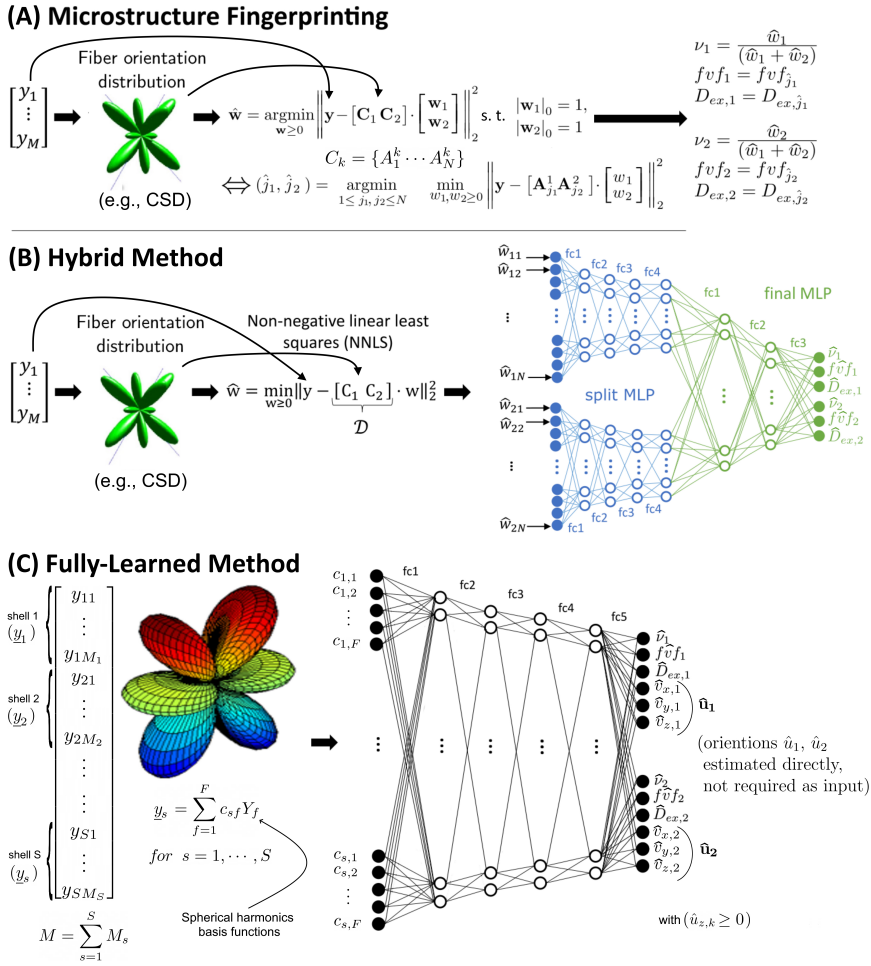


Fig. 4.2 CSD & MF, Hybrid Method and Fully-Learned Method (A) Microstructure Fingerprinting involves solving a large number of small NNLS problems, leading to long computation times as the dictionary size N increases. (B) The vector of measurements **y** is decomposed by NNLS into a sparse representation in the space of physics-based fingerprints. The weights are given to a multilayer perceptron with a split architecture to predict the tissue parameters (relative volume of axon population ν_k, fiber volume fraction fvf_k and extra-axonal diffusivity $D_{ex,k}$). (C) The Fully-Learned Method calculates shellwise spherical harmonics coefficients of order 12, which are fed into a DNN to predict the tissue parameters and axon populations orientation.

4.2 Materials and methods

This section begins with a detailed presentation of the reference methodology, known as Microstructure Fingerprinting. We then introduce and elaborate on the two proposed acceleration techniques: the Hybrid Method and the Fully-Learned Method. Subsequently, the section delineates the three experiments carried out to evaluate and compare these three methods, focusing on assessing their inference speed and mean absolute error (MAE). This comparative analysis aims to provide a comprehensive understanding of the performance and potential advantages of each method within the context of DW-MRI microstructure estimation.

4.2.1 Microstructure Fingerprinting

The Microstructure Fingerprinting framework first precomputes a dictionary of DW-MRI fingerprints (Rensonnet, Jacobs, Macq et al., 2015; Rensonnet, Scherrer, Girard et al., 2019). Each fingerprint in the dictionary corresponds to a unique microstructural configuration, obtained by leveraging the physical accuracy of Monte Carlo simulations (Hall and Alexander, 2009; Rensonnet, Jacobs, Macq et al., 2015). As depicted in Figure 4.1, in this work, every axonal configuration was a random 2D packing of straight cylinders with radii drawn from a gamma distribution with mean radius of 0.5 µm and a standard deviation of 0.3 µm. The intra-axonal diffusivity was set to $2.2\,\mu m^2\,ms^{-1}$ (Dhital, Reisert, Kellner et al., 2019). The different configurations present in the dictionary were obtained by choosing the fiber volume fraction (fvf) from 38 equally spaced values within the range $[0.06, 0.8]$ and the extra-axonal diffusivity (D_{ex}) from 10 equally spaced values in the $[0.6, 2.4]\,\mu m^2\,ms^{-1}$ range, totaling $N = 380$ precomputed fingerprints.

As shown in Figure 4.2 (A), at every runtime, MF requires an estimate of the orientations $\mathbf{u}_k$ of each fascicle k in every voxel. For this purpose, we here employed a constrained spherical deconvolution (CSD) (Tournier, Calamante, and Connelly, 2007). Assuming independent signal contributions from each fascicle (Rensonnet, Scherrer, Warfield et al., 2018), the precomputed dictionary was then rotated to create fascicle-specific subdictionaries $\mathbf{C}^k$. The method then searched for the optimal combination of single-fascicle configurations $(fvf_k, D_{ex,k})$ and volume fractions v_k, for

each fascicle $k = 1, \ldots, K$, given a vector $\mathbf{y} \in \mathbb{R}^M$ of M noisy DW-MRIs by selecting the best-fit solution out of N^K independent non-negative linear least-squares (NNLS) sub-problems of K variables.

4.2.2 The Hybrid Method

The Hybrid Method operates in two distinct stages. In the first stage (Figure 4.2 (B)), a relaxed version of the problem present in MF was used to project the DW-MRI signal into a latent space. Specifically, the same matrix $[C_1 \ldots C_K]$ was first assembled. However, instead of tackling the problem with its full multi-compartment complexity, one single NNLS problem with $K \cdot N$ variables was solved without any knowledge about its multi-compartment structure and without any sparsity constraints on the weight vector $\mathbf{w}$. This effectively projected the measurements $\mathbf{y}$ in the space of the dictionary fingerprints. NNLS is known to naturally enforce sparsity (Canales-Rodríguez, Legarreta, Pizzolato et al., 2019; Lawson and Hanson, 1995). Additionally, its empirical runtime is proportional to the number of matrix columns, which in this case is $K \cdot N$. This characteristic of NNLS is a key factor in the efficiency of the Hybrid Method.

As illustrated in Figure 4.2 (B), the second stage consisted in passing the NNLS weight vector $\mathbf{w}$ through a feed-forward neural network. The network initially had split arms, independently processing the weights related to each compartment by NNLS before merging into a common multi-layer perceptron (MLP) for final parameter inference. The advantage of this architecture is to reduce the number of free parameters and facilitate training. The output size of the two initial branches of the MLP, the learning rate and the number of layers were determined through a systematic hyperparameter optimization process using grid search. Our network used a rectified linear unit (ReLU) (Nair and Hinton, 2010) as an activation function for all layers, except for the last layer, where a sigmoid function was used. Between all layers, dropout regularization (Hinton, Srivastava, Krizhevsky et al., 2012) was used with a retention probability of 0.9. The network included a batch normalization layer (Ioffe and Szegedy, 2015) between the the split arms and common multi-layer perceptron. The Adam optimizer (Kingma and Ba, 2014) was utilized during training. The target output of the MLP for each fascicle k was the fascicle fraction ν, the fiber volume fraction fvf and the extra-axonal diffusivity D_{ex}. For more in-depth details on models architecture and parameters used during training,

see Table B.1 in appendix B.

The network was trained on a dataset comprising 2 000 000 synthetic voxels. These voxels were generated using Monte Carlo simulations based on the microstructural configuration presented in Figure 4.1, considering 2 fascicles per voxel. Microstructural parameters were drawn from uniform distributions: the crossing angle between the two axon populations $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ from $\mathcal{U}([15°, 90°])$ and the volume fractions ν from $\mathcal{U}([0.05, 0.95])$. For each voxel, the pairs $(fvf_k, D_{ex,k})$ were randomly selected from the $N = 380$ precomputed fingerprints described above. The resulting DW-MRI signal s was corrupted with Rician noise using an i.i.d. Gaussian realization ϵ with standard deviation $\sigma = \frac{s_{\text{noiseless}}(\text{b-value} = 0)}{\text{SNR}}$ by computing $s = \sqrt{s^2_{\text{noiseless}} + \epsilon^2}$. The signal-to-noise ratio (SNR) was drawn from $\mathcal{U}([10, 100])$.

4.2.3 The Fully-Learned Method

As depicted in Figure 4.2 (C), the Fully-Learned Method relies on a traditional MLP architecture and directly outputs estimates of microstructural parameters without resorting to any intermediate biophysical representation of the signal. The Fully-Learned Method requires the DW-MRI signal to be acquired in shells (Tuch, Reese, Wiegell et al., 2002) of fixed b-value as it first computes a spherical harmonics representation of each shell separately to provide the coefficients as inputs to the network. Compared to the use of raw DW-MRI signal directly, this approach is more flexible with respect to changes in the acquisition protocol (e.g., in the number of gradients in some shells), and robust to missing measurements.

An important difference with the Hybrid Method is that we do not need to give the orientations $\hat{\mathbf{u}}_k$ because they are estimated from the data jointly with the other fascicle-specific properties. Consequently, the approach does not require external estimates such as those of CSD. Note that, given the spherical symmetry of the DW-MRI signal, the orientations were restricted to the unit half-sphere $\left\{ \left[\hat{u}_x, \hat{u}_y, \hat{u}_z\right]^T | \hat{u}_z \geq 0 \right\}$ during training.

As in the Hybrid Method, a grid search approach was employed for the optimization of the neural network hyperparameters. ReLU and dropout were incorporated into each layer, except for the final layer, which utilized a sigmoid as the activation function and did not include dropout. The Adam optimizer was used during training. As we used real-valued

of even-degree spherical harmonics functions up to degree 12, the input contained 91 coefficients per shell.

The training of the Fully-Learned Method was conducted using the same set of simulated data as for the Hybrid Method, ensuring consistency in the evaluation of both methods.

4.2.4 Diffusion protocol

Unless otherwise specified in the experiments below, the acquisition protocol used to simulate the DW-MRI samples was the one from the MGH human connectome project (HCP) (Setsompop, Kimmlingen, Eberlein et al., 2013), consisting of 4 PGSE shells of 64 directions at $b = 1000$ ($G = 69\,\mathrm{mT\,m^{-1}}$), 64 at $b = 3000$ ($G = 120\,\mathrm{mT\,m^{-1}}$), 128 at $b = 5000$ ($G = 155\,\mathrm{mT\,m^{-1}}$) and 256 at $b = 10\,000\,\mathrm{\mu s\,\mu m^{-2}}$ ($G = 219\,\mathrm{mT\,m^{-1}}$) with gradient duration $\delta = 12.9\,\mathrm{ms}$, diffusion time $\Delta = 21.8\,\mathrm{ms}$, TE/TR=57/8800 ms and no MultiBand. A non-diffusion weighted image (b = 0) was collected every 14 volumes, yielding a total of 552 volumes. The total acquisition time was 89 min.

4.2.5 Experiment I: Performance assessment on synthetic test set

Two-Fascicle configuration

Based on the diffusion protocol used during training, a test set was created with 2-fascicle voxels having crossing angles drawn from $\mathcal{U}([15°, 90°])$, ν_1 values selected in the set [0.5, 0.6, 0.7, 0.8, 0.9], every $(fvf_k, D_{ex,k})$ pair from the $N = 380$ precomputed fingerprints described above and SNR values in the set [20, 30, 50], generating a total of 150 000 samples. All testing samples were never seen by either neural networks during training.

The experiment focused on assessing four distinct models. The first two are the proposed accelerated methods, known as the Hybrid Method and the Fully-Learned Method. Alongside these, two versions of the Microstructure Fingerprinting model were assessed. The True orientations & MF, which utilizes actual, known fascicle orientations, serving as a benchmark for the ideal scenario in Microstructure Fingerprinting. In contrast, the CSD & MF model operates without the advantage of known fascicle orientations and instead relies on orientations estimated through constrained spherical deconvolution, depicting a scenario more representative of typical clinical settings.

For each of these models, the absolute error and mean absolute error were calculated on the estimated volume fractions v_1 and v_2, extra-axonal diffusivities $D_{ex,1}$ and $D_{ex,2}$, fiber volume fractions fvf_1 and fvf_2. We also reported the mean angular error on $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, computed either by CSD or by the Fully-Learned Method directly.

Three-Fascicle configuration

To evaluate the performance of our models in more complex scenarios involving three-way fascicle crossings, the architectures of both the Hybrid and Fully-Learned neural networks were modified to process three-fascicle voxels. These models were retrained on synthetic data employing similar configurations to those used in the two-fascicle experiments, with an alteration in the method of volume fraction generation. Given the difficulty of using an uniform distribution for v_k in scenarios involving more than two fascicles while ensuring $\sum_k v_k = 1$, volume fractions v_k were generated using a Dirichlet distribution with $\alpha = 1$.

For the test set, the methodology from the two-fascicle experiment was maintained, but the uniform distribution for volume fractions was replaced with a Dirichlet distribution. Following the procedures established in the two-fascicle tests, all four methods were evaluated.

4.2.6 Experiment II: Generalizability to unseen acquisition protocols

The objective of this experiment was to evaluate the Hybrid Method's ability to generalize to new, previously unseen DW-MRI acquisition protocols. This was driven by the fact that, unlike the Fully-Learned Method, the Hybrid Method represents the DW-MRI signal by an intermediate weight vector $\mathbf{w}$ that does not depend on the number of measurements M but only on the underlying microstructural model via the dictionary of fingerprints.

For this experiment, a new simulated protocol was designed to mimic current clinical capabilities in DW-MRI with gradient duration $\delta = 22.9\,\text{ms}$, diffusion time $\Delta = 35.7\,\text{ms}$, TE/TR=77.4/4842 ms, 64 gradient directions at $b = 1000$ ($G = 31\,\text{mT}\,\text{m}^{-1}$) and $b = 2000$ ($G = 44\,\text{mT}\,\text{m}^{-1}$), and 128 directions at $b = 5000\,\text{µs}\,\text{µm}^{-2}$ ($G = 69\,\text{mT}\,\text{m}^{-1}$), along with 4 $b = 0$ acquisitions, for a total of 260 measurements. The simulated protocols excluded gradient directions at $b = 10\,000\,\text{µs}\,\text{µm}^{-2}$ since high b-values are typically impractical in clinical settings. Furthermore, to create a clear distinction

from the reference protocol used during training, gradient directions at $b = 3000\,\mu s\,\mu m^{-2}$ were also omitted.

The test set for this experiment was generated following the same methodology as in Experiment I, but adapted to the new protocol. In this experiment, four distinct methods were put to the test: The MF model with true fascicle orientations (True orientations & MF) and without true fascicle orientations (CSD & MF) as in Experiment I, and the Hybrid Method under two different conditions. The first condition involved applying the Hybrid Method without any retraining, directly to the new protocol. This gauges the method's inherent ability to adapt to different acquisition parameters based on its original training. The second condition involved fully retraining the Hybrid Method specifically for the new protocol, aiming to determine the advantages of tailoring the model to the unique characteristics of a new protocol.

4.2.7 Experiment III: In-vivo population

Experiment III focuses on estimating the microstructural properties of white matter on in-vivo data for the CSD & MF, Hybrid Method, and Fully-Learned Method. This assessment was performed on all 34 subjects from the MGH-USC Young Adult cohort (Van Essen, Ugurbil, Auerbach et al., 2012), whose ages ranged from 20 to 59 years old and scanned using a customized Siemens 3T Connectome scanner, providing a consistent and high-quality dataset for analysis. The computation time required by each model to process the MRI data from all 34 subjects was recorded to assess inference speed. The white matter maps generated by the two accelerated methods were compared with those produced by the reference MF method across all subjects and a visual inspection was performed. Given the lack of ground truth on in-vivo data, the experiment also aimed to discern the effects of transitioning from synthetic to real-world data. To achieve this, the differences in the estimated fiber volume fraction (fvf_1) between the reference MF and the two proposed methods were calculated for each voxel. These calculations were performed on both the simulated test set from Experiment I and the in-vivo data from the 34 subjects in the cohort. The results were presented as histograms of the signed differences, providing a statistical perspective on the variance between each method's estimations when applied to synthetic and real-world data.

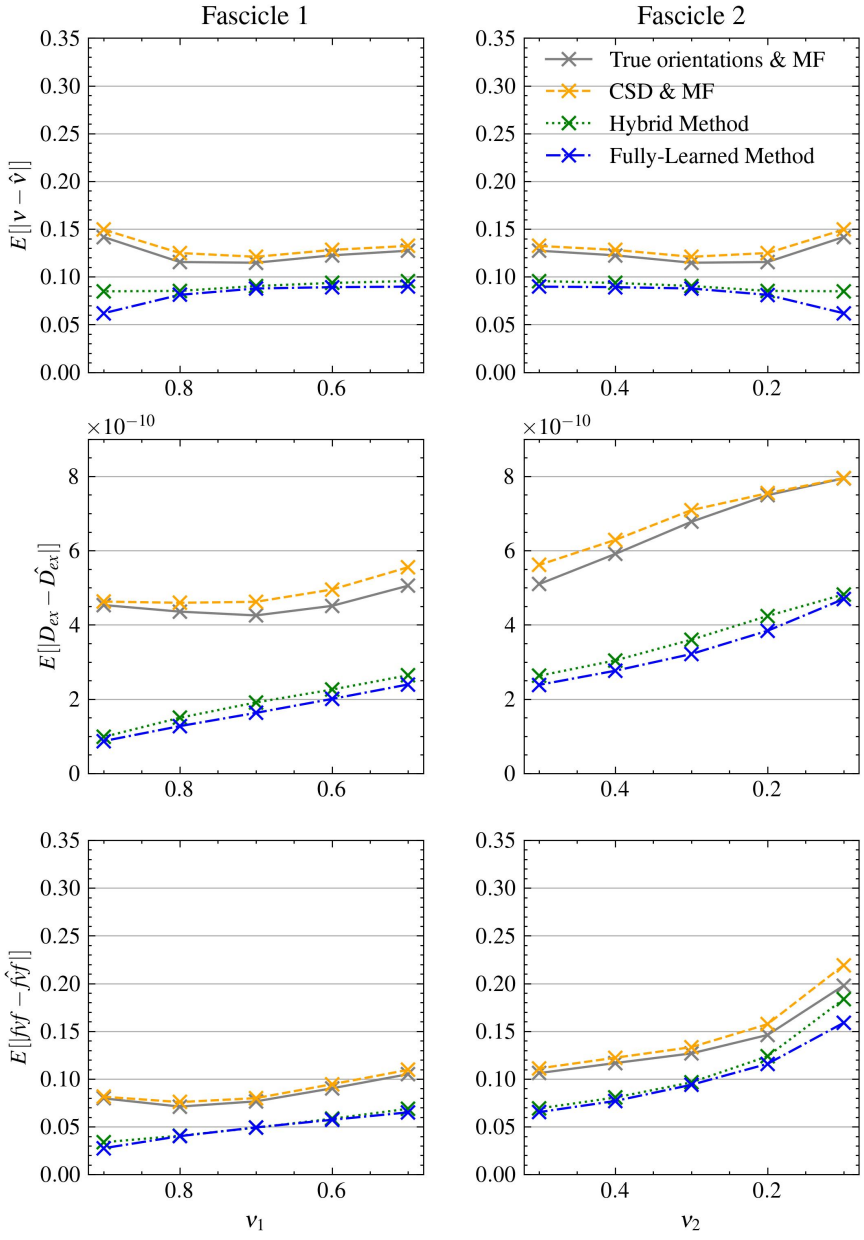


Fig. 4.3 The Fully-Learned Method has better estimation accuracy than the Hybrid Method, the CSD & MF and the MF algorithm with true orientations. MAE on structured test.

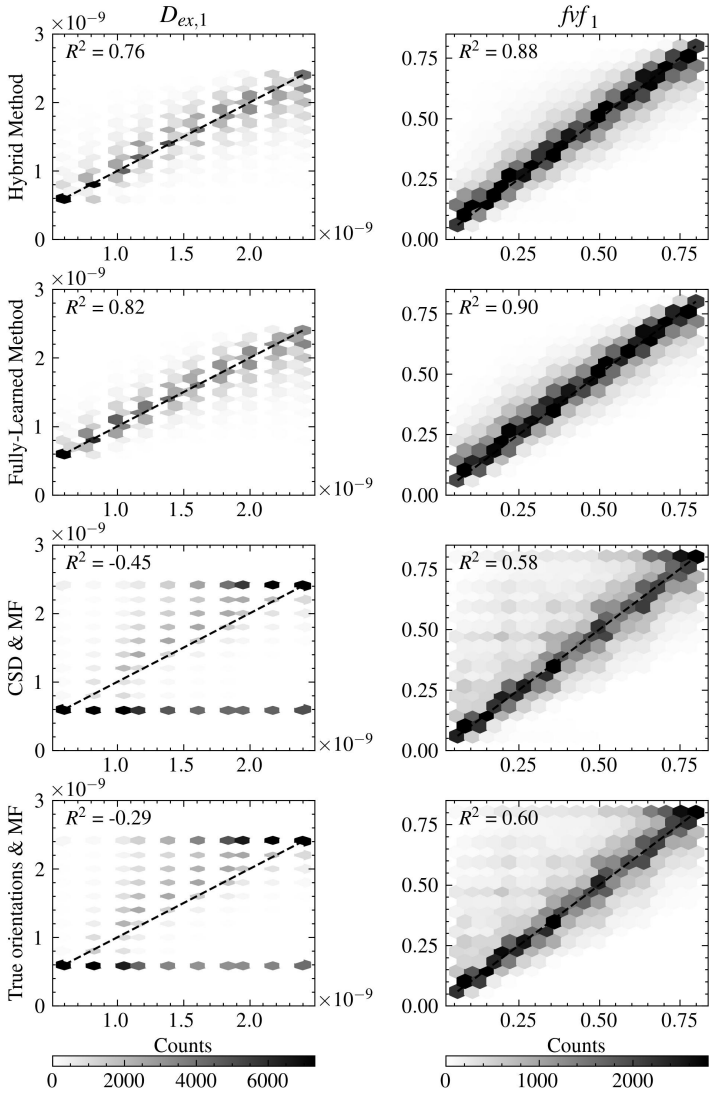


Fig. 4.4 Both accelerated method obtain estimates closer to the groundtruth for the *fvf* and D_{ex} metrics than the reference MF methods. Correlation-accuracy plots comparing the predicted values and the ground truth parameters for the four models on the test set. R^2 denotes the coefficient of determination, defined as $R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$, where SS_{res} is the residual sum of squares and SS_{tot} is the total sum of squares. Negative values indicate that the model performs worse than a baseline model that always predicts the mean of the target variable.

4.3 Results

4.3.1 Experiment I: Performance assessment on synthetic test set

Two-Fascicle configuration

Experiment I evaluated the accuracy and efficiency of our proposed methods, the Hybrid Method and the Fully-Learned Method, against conventional Microstructure Fingerprinting approaches, using a synthetic test set designed to simulate diverse white matter microstructures.

Table 4.2 Large speed-up factors were observed for our proposed methods, with theoretical projections suggesting even larger gains with increasing dictionary size (N) and number of compartments (K). Efficiency of the three methods when inference is performed on synthetic data using the MGH adult diffusion protocol. The test set was processed on a Skylake Xeon 4116 CPU cluster, without GPU acceleration. During the training phase, a Nvidia RTX 3090 GPU was used to train both proposed methods.

	Microstructure Fingerprinting	Hybrid Method	Fully-Learned Method
Precomputation time	$\approx 2\,\mathrm{d}$	$\approx 2\,\mathrm{d}$	Not applicable
CASE WITH K=2 FASCICLES:			
Model Training	Not applicable	7 min	21 min
Total inference time/voxel	$9.18 \times 10^{-1}\,\mathrm{s}$	$2.36 \times 10^{-2}\,\mathrm{s}$	$8.61 \times 10^{-4}\,\mathrm{s}$
Acceleration factor	**1**	**38.9**	**2090**
CASE WITH K=3 FASCICLES:			
Model Training	Not applicable	10 min	22 min
Total inference time/voxel	248 s	$2.97 \times 10^{-2}\,\mathrm{s}$	$8.82 \times 10^{-4}\,\mathrm{s}$
Acceleration factor	**1**	**8350**	**281179**
Inference complexity	$\mathcal{O}\left(N^K K\right)$	$\mathcal{O}\left(NK\right)$	$\mathcal{O}\left(1\right)$
Theoretical acceleration factor	$\mathcal{O}\left(1\right)$	$\mathcal{O}\left(N^{K-1}\right)$	$\mathcal{O}\left(N^K K\right)$

As visible in Table 4.2, both methods yielded the expected gain in computation time with the Hybrid Method reaching an acceleration factor of 38.9 and the Fully-Learned Method reaching an even higher acceleration factor of 2090. A more detailed view of computation time is available in Table B.2 of appendix B.

Table 4.3 Proposed acceleration methods outperform reference approaches. MAEs of all methods on the test set from Experiment 1. For each metric and SNR, the method with the lowest MAE is highlighted in bold.

	SNR=20			SNR=30			SNR=50		
	ν	D_{ex}	fvf	ν	D_{ex}	fvf	ν	D_{ex}	fvf
True orientations & MF	0.149	6.77e−10	0.123	0.125	5.65e−10	0.114	0.100	4.37e−10	0.098
CSD & MF	0.152	7.02e−10	0.129	0.131	5.82e−10	0.119	0.111	4.82e−10	0.109
Hybrid Method	0.107	3.21e−10	0.094	0.092	2.84e−10	0.083	0.071	2.24e−10	0.064
Fully-Learned Method	**0.095**	**2.92e−10**	**0.088**	**0.084**	**2.55e−10**	**0.076**	**0.068**	**2.06e−10**	**0.061**

Figure 4.3 presents the mean absolute errors (MAEs) by fascicle fraction ν for each microstructural feature estimated by the four evaluated models. Furthermore, Table 4.3 complements this by providing a detailed breakdown of the MAEs for each of the three signal-to-noise ratio levels within the synthetic test set, thus offering insights into each method's accuracy under different noise conditions. The Hybrid Method yielded MAEs of 0.0899 for ν, 0.0805 for fvf, and 2.76e−10 for D_{ex}. The Fully-Learned Method demonstrated marginally better performance, with MAEs of 0.0819 for ν, 0.0751 for fvf, and 2.51e−10 for D_{ex}. In contrast, the True Orientations & MF method, despite utilizing accurate fascicle orientations, reported higher MAEs of 0.124 for ν, 0.112 for fvf, and 5.59e−10 for D_{ex}. The CSD & MF method marked the highest MAEs, with 0.131 for ν, 0.118 for fvf, and 5.88e−10 for D_{ex}. These results underscore the superior accuracy of both proposed methods over the traditional MF approach in estimating key microstructural features.

Further insight into the models' performance is provided by Figure 4.4, which includes correlation-accuracy plots for the fvf_1 and $D_{ex,1}$ properties. This figure also reveals the coefficient of determination (R^2), showcasing superior predictive accuracy for the Hybrid (0.88 for fvf_1 and 0.76 for $D_{ex,1}$) and Fully-Learned Methods (0.90 for fvf_1 and 0.82 for $D_{ex,1}$), in stark contrast to the lower R^2 values observed for the CSD & MF (0.58 for fvf_1 and -0.45 for $D_{ex,1}$) and True Orientations & MF methods (0.60 for fvf_1 and -0.29 for $D_{ex,1}$).

As shown in Figure 4.5, the angular errors on the estimated orientations obtained by the Fully-Learned Method were smaller (mean $= 5.06°$) than those of CSD (mean $= 7.41°$), highlighting the capacity of the method to successfully estimate orientations.

Figure 4.6 shows the t-SNE projection of the neural network activations for the Hybrid Method. The activations form distinct clusters correspond-

ing to different ground truth microstructural configurations. The clustering indicates that the Hybrid Method organizes the data into a structured representation based on the input DW-MRI signal.

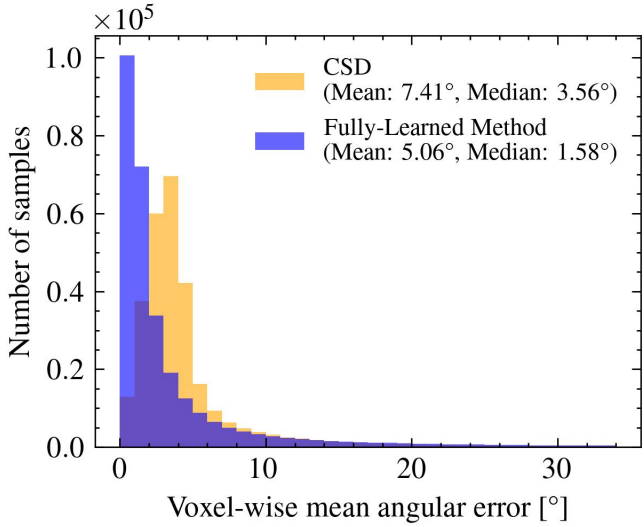


Fig. 4.5 The orientations obtained with the Fully-Learned Method outperform results from constrained spherical deconvolution in the synthetic test set. Angular error: CSD vs the Fully-Learned Method.

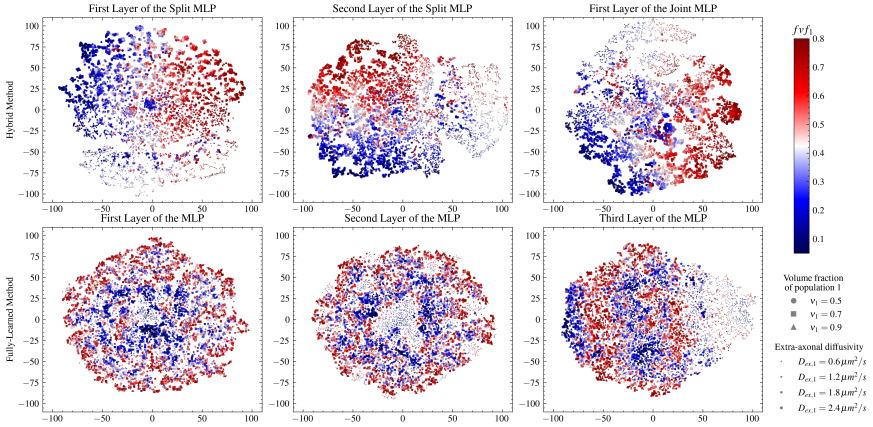


Fig. 4.6 The neural network of the Hybrid Method learns faster useful signal representations. Projection in 2-dimensional plane of network activations using t-SNE embedding (Van der Maaten and Hinton, 2008).

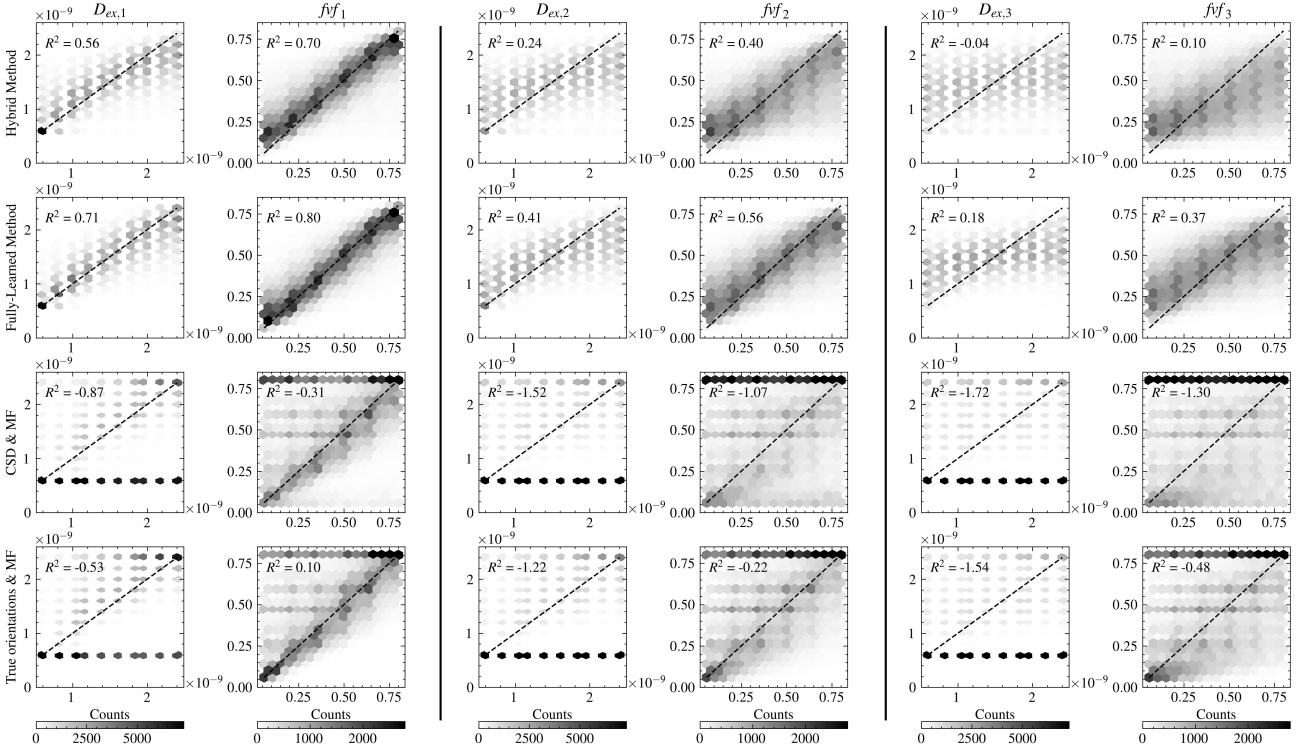


Fig. 4.7 The accelerations methods generalize to three-way crossings, outperforming the reference MF methods in estimation accuracy for the metrics *fvf* and D_{ex}. Correlation-accuracy plots comparing the predicted values against the ground truth parameters for the four models on the test set with a three-fiber configuration. R^2 denotes the coefficient of determination, defined as $R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$, where SS_{res} is the residual sum of squares and SS_{tot} is the total sum of squares.

Three-Fascicle configuration

To further assess the performance and scalability of the proposed methods, the second part of Experiment I evaluated all methods on voxels with a three-fascicle configuration. This extension aimed to test the robustness and efficiency of the methods under more complex scenarios.

As illustrated in Table 4.2, the acceleration factors achieved with the three-fascicle configuration are significantly higher than those observed with the two-fascicle configuration. Specifically, the Hybrid Method achieved an acceleration factor of 8350, while the Fully-Learned Method reached 281179.

The correlation-accuracy plots in Figure 4.7 reveal the same trends as in the two-fascicle configuration, with acceleration methods obtaining estimates closer to the groundtruth than the reference MF methods. The errors are slightly higher overall in the three-fascicle configuration for all tested approaches.

4.3.2 Experiment II: Generalizability to unseen acquisition protocols

The results from Experiment II, as depicted in Figure 4.8, demonstrate the ability of the Hybrid Method to transfer across different protocols. Without retraining, the Hybrid Method achieved MAEs of 0.102 for ν, 0.0923 for fvf and 3.79e$-$10 for D_{ex}, consistently outperforming the CSD & MF method in all scenarios which had MAEs of 0.122 for ν, 0.114 for fvf and 5.68e$-$10 for D_{ex} along with the True orientations & MF approach which had MAEs of 0.112 for ν, 0.102 for fvf and 5.42e$-$10 for D_{ex}.

Upon retraining with the new protocol, the Hybrid Method's accuracy improved further, yielding even lower MAEs of 0.0942 for ν, 0.0843 for fvf and 2.86e$-$10 for D_{ex}, exhibiting an improvement in performance compared to own version trained on the source protocol. This enhanced performance was consistent across all metrics, further outstripping both reference methods.

4.3.3 Experiment III: In-vivo population

Table 4.4 presents the acceleration factors achieved by these methods. The results were in line with the theoretical predictions regarding inference complexity.

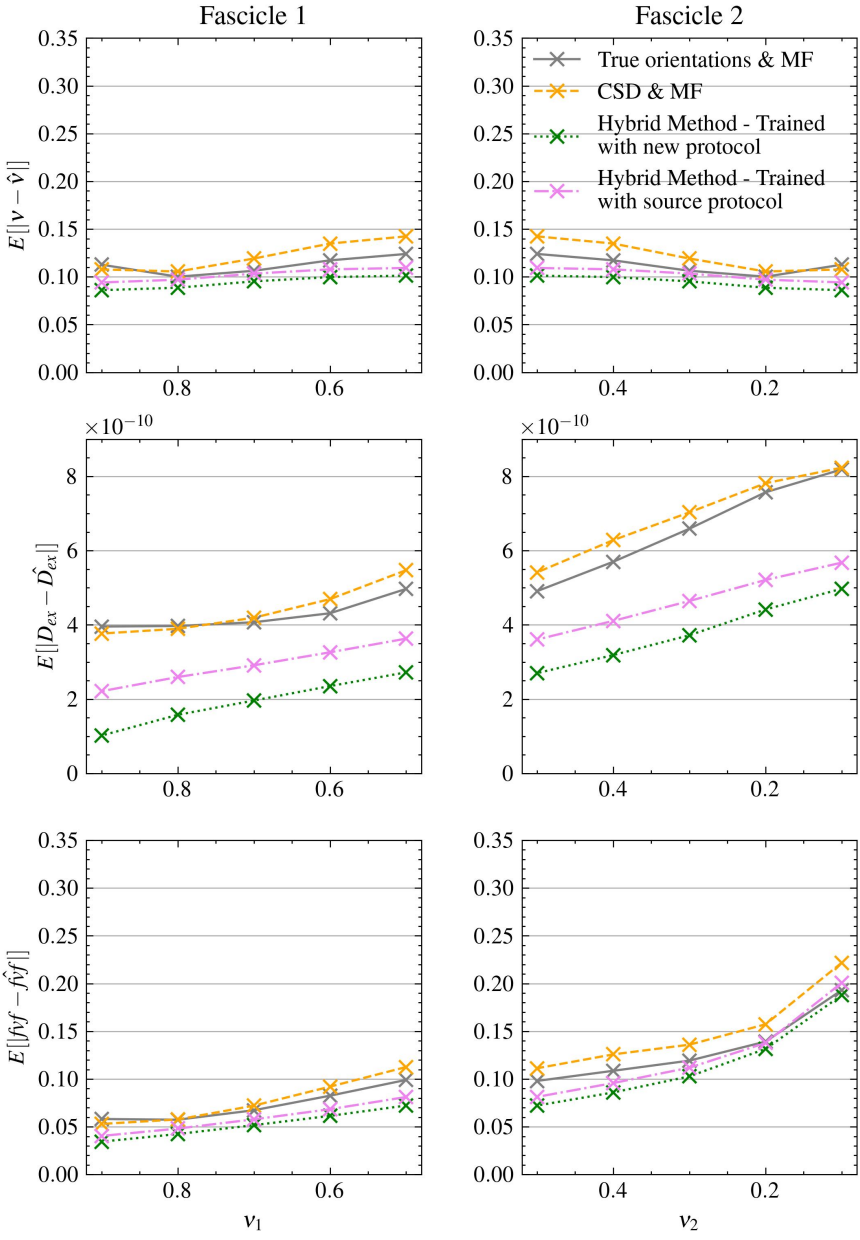


Fig. 4.8 The Hybrid Method is able to generalize to data acquired with a different experimental protocol with minimal retraining. MAE on structured test with the clinical protocol.

Table 4.4 The acceleration factor obtained on in-vivo data is of the same order of magnitude as the one obtained on synthetic data. Efficiency of the three methods when inference is performed on 34 subjects from the MGH adult diffusion dataset, processed on a Skylake Xeon 4116 CPU cluster with single-thread allocation per subject, without GPU acceleration.

	CSD & MF	Hybrid Method	Fully-Learned Method
Computation of SH coefficients/cohort	Not applicable	Not applicable	47 min
Constrained Spherical Deconvolution	9.3 h	9.3 h	Not applicable
Dictionary rotation time/cohort	14.2 h	14.2 h	Not applicable
Exhaustive fingerprinting time/cohort	1690.5 h	Not applicable	Not applicable
Solving NNLS time/cohort	Not applicable	26.5 h	Not applicable
NN forward pass time/cohort	Not applicable	26 min	9 min
Total inference time/cohort	1714 h	50.4 h	56 min
Total inference time/subject	50.4 h	1.5 h	96 s
Acceleration factor	**1**	**34**	**1836**

Figure 4.9 presents a comparative analysis between the CSD & MF method and each accelerated method, utilizing histograms to depict differences in fvf_1 estimates across synthetic and in-vivo datasets. Although the distribution of these differences changes slightly, the mean difference for both proposed methods maintains consistency as we transition from synthetic to in-vivo data. Notably, the in-vivo dataset shows a more tightly clustered signed difference between CSD & MF and the Hybrid Method around zero, indicating a closer alignment in real-world settings as opposed to synthetic environments. Meanwhile, the Fully-Learned Method consistently estimates higher fiber volume fraction values compared to the CSD & MF method.

In Figure 4.10(A), the voxel-wise values of the fiber volume fraction for the larger compartment are displayed. Here, the Fully-Learned Method generally estimated higher *fvf* values compared to the other models, indicating a tendency to estimate higher values than the CSD & MF method.

Figure 4.10(B) highlights the differences between the fully analytical CSD & MF model and the two Deep Learning models. The Hybrid Method showed results that were aligned with the CSD & MF model, suggesting its effectiveness in mirroring more traditional analytical approaches. On the other hand, the Fully-Learned Method exhibited a nearly constant bias, diverging from the CSD & MF model to a certain extent.

Overall, these findings from in-vivo data analysis underscore the practical applicability of the proposed methods, particularly highlighting the nuances in their performance when applied to real-world clinical data as opposed to a controlled synthetic environments.

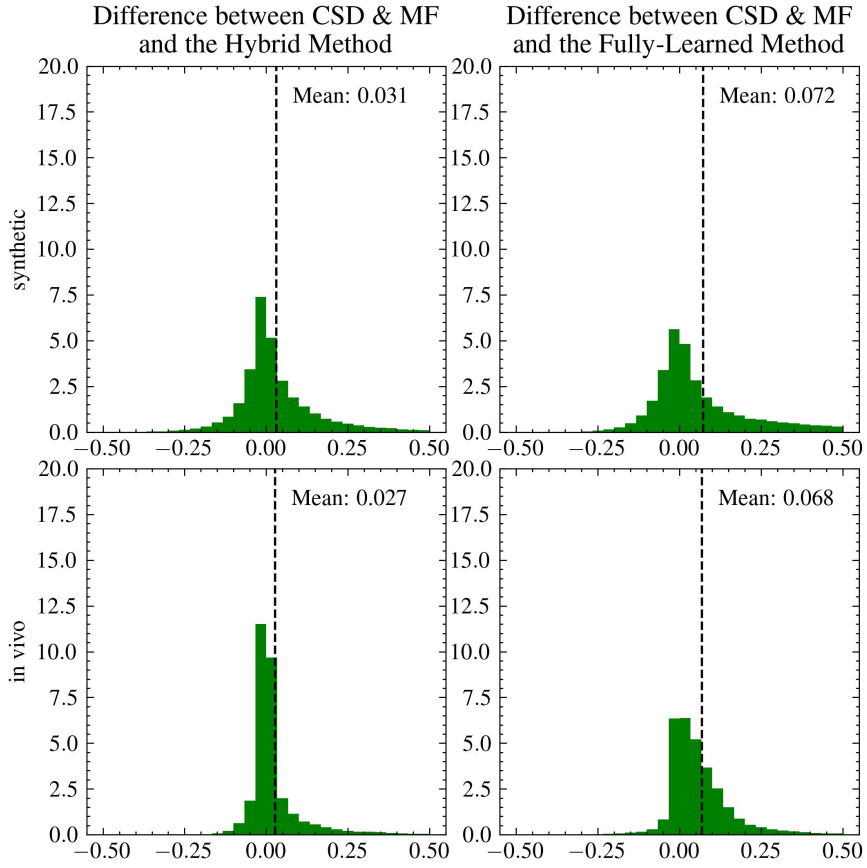


Fig. 4.9 The mean differences between the reference and the accelerated Microstructure Fingerprinting follow similar trends on simulated and on in-vivo data. Histogram (normalized to integrate to 1) of the signed differences between CSD & MF and the Hybrid (left) and Fully-Learned (right) Methods for fvf_1 on synthetic (top) and in-vivo (bottom) data. The vertical dotted lines represent the mean of the signed difference.

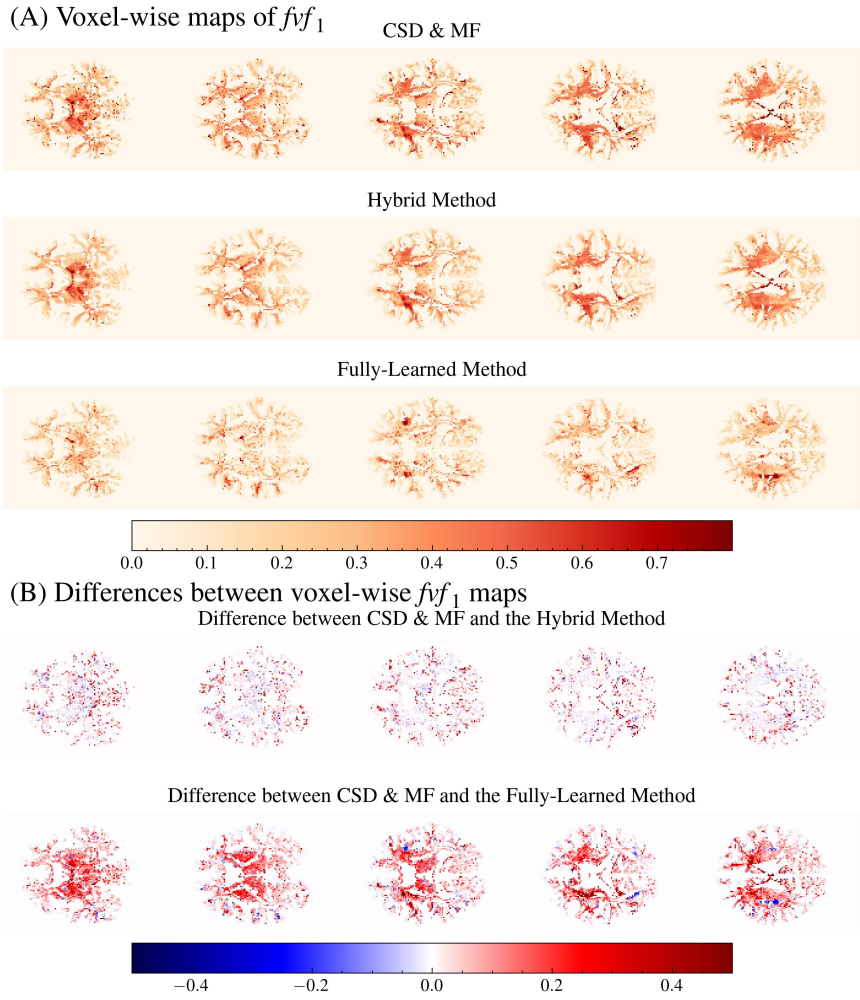


Fig. 4.10 Structural integrity between all methods seems to be maintained. Several voxels are estimated to a higher fiber volume fraction by the Fully-Learned Method. Detailed visualization on subject MGH 1001 of voxel-wise maps of fvf_1 and voxel-wise maps of the differences in fiber volume fraction for the first compartment (fvf_1) between the reference method and the Hybrid Method and the Fully-Learned Method. Constrained spherical deconvolution was used to estimate the number of fascicles and their orientations. The analysis specifically targets voxels identified by CSD as containing two or more fascicles.

4.4 Discussion

4.4.1 Performance

Both accelerated methods successfully improved the inference speed of microstructural parameters on both the test set and in in-vivo data. This achievement is further highlighted when dealing with voxels containing three fascicles. The acceleration factors become even more impressive due to the inference complexity of the reference method, which increases from $\mathcal{O}\left(2N^2\right)$ for two-fiber configurations to $\mathcal{O}\left(3N^3\right)$ for three-fiber configurations. This highlights the potential of the two accelerated methods to handle multiple-fascicles configurations and large dictionary sizes efficiently.

4.4.2 Accuracy

The results from our experiments demonstrate that both the Hybrid Method and the Fully-Learned Method significantly outperform the reference Microstructure Fingerprinting estimation across a variety of simulated scenarios. This superior performance can be attributed to the neural networks' ability to learn and model the complex relationship between the DW-MRI signal and the underlying tissue properties, even in noisy environments. A key factor in this success is the original MF's vulnerability to errors in fascicle orientation estimation, as evidenced in Experiments I and II. The Hybrid Method, which relies on these estimates only in its initial stage, appears to correct for any residual biases in the subsequent stage. In contrast, the Fully-Learned Method, which independently estimates orientations, outperformed CSD in orientation accuracy (as shown in Experiment I, Figure 4.5) suggesting the potential benefits of benchmarking neural network approaches against established methods in the literature (Canales-Rodríguez, Legarreta, Pizzolato et al., 2019; Karimi, Vasung, Jaimes et al., 2021; Ye, Li, and Chen, 2019).

4.4.3 Interpretability

One notable advantage of the Hybrid Method is its interpretability. The method's design allows it to learn useful signal representations more rapidly and with fewer layers. The t-SNE visualization of intermediate

outputs through the networks' layers, shown in Experiment I, Figure 4.6, demonstrates that the Hybrid Method effectively captures the fvf and D_{ex} properties after just the first layer, unlike the Fully-Learned Method, which requires more layers to achieve similar structure in its data representations. This rapid learning is likely due to the reliance on the dictionary of fingerprints, which grounds the method in physics-based properties.

4.4.4 Generalizability to diffusion protocol changes

Experiment II focused on the generalizability of the Hybrid Method, revealing that the model trained with the source protocol performed almost as well as the fully retrained model. This finding indicates that the Hybrid Method possesses robust protocol transfer capabilities, suggesting the feasibility of using a single DNN for multiple acquisition protocols. These findings illustrate the Hybrid Method's robust generalization capabilities. The method not only maintains strong performance across different protocols but also shows that a small, yet significant, marginal gain can be achieved by retraining the model on the new protocol. This adaptability is crucial for practical applications, suggesting that the Hybrid Method can be effectively applied in various clinical and research settings with different DW-MRI protocols, and can benefit from fine-tuning to specific acquisition parameters. For the Fully-Learned Method, theoretically, changes in the number or values of shells would necessitate retraining the entire network. However, the use of spherical harmonics implies that changes in gradient directions within a shell should not significantly impact performance.

4.4.5 Generalizability to other microstructural models

The trained models and estimated microstructural parameters in this study were based on a dictionary of presimulated fingerprints obtained with Monte Carlo simulations. This flexible methodology can inherently accommodate other fiber configurations by updating the single-fascicle dictionary, e.g., sampling an undulation, an intra-axonal diffusivity or a dispersion parameter to reflect different brain regions, age, or diseased tissue.

Incorporating more complex fiber configurations at the single-fascicle level such as fanning fibers (Sotiropoulos, Behrens, and Jbabdi, 2012; Tariq, Schneider, Alexander et al., 2016), undulation (Brabec, Lasic, and Nils-

son, 2020; Nilsson, Lätt, Ståhlberg et al., 2012; Rafael-Patino, Romascano, Ramirez-Manzanares et al., 2020) or glial cells (Taquet, Jankovski, Rensonnet et al., 2019) does however require slight modifications of the neural networks. The last layer must match the number of microstructural parameters to estimate and training must be specific to each single-fascicle model.

4.4.6 In-vivo application

The proposed methods led to massive acceleration of the dMRI data processing, with the Fully-Learned Method performing the estimation of the entire HCP-MGH cohort in 56 minutes, compared to 1714 hours (of single-core equivalent) for the original MF (Table 4.4). The in-vivo experiments showed differences between the original MF maps and our proposed methods, with CSD & MF and the Hybrid Method aligning and the Fully-Learned Method predicting higher *fvf* values in specific white matter areas (Figure 4.9, Figure 4.10). The stronger similarity between CSD & MF and the Hybrid Method can be explained by the use, in the Hybrid Method, of a signal decomposition in fingerprint space close to the analytical approach of MF. This contributes to the Hybrid Method interpretability, since it is closely aligned with the dictionary approach and its architecture reflects the biophysical organization of axons in white matter (Rensonnet, Adam, and Macq, 2021). Additionally, both methods also share a reliance on the same orientation estimation technique. As for the Fully-Learned Method, the consistency of the differences with CSD & MF observed between synthetic and in-vivo data (Figure 4.9), combined with its smaller MAEs over CSD & MF in most simulated experiments (Experiments I, II), gives ground to argue for superior accuracy. This suggestion, while promising, deserves further investigation and validation in a broader range of scenarios to fully substantiate its potential advantages over established techniques.

4.5 Conclusion

In this work, we proposed two novel deep learning-based approaches for estimating microstructural features of crossing fascicles in white matter from DW-MRI measurements. The Hybrid Method relies on two stages

Table 4.5 The Hybrid Method is capable of protocol transfer. However, the Hybrid Method is less accurate and depends on the accuracy of an external routine used for orientation estimation. Advantages and disadvantages of both fast approaches.

	Hybrid Method	Fully-Learned Method
Learning Speed	Fast learning: Requires a small volume of training data.	Slow learning: Requires a large volume of training data.
Inference Speed	Slow inference: The NNLS and the external estimation of fiber orientation reduce the inference time.	Fast inference: The computations of SH coefficients and the inference of the MLP are extremely efficient.
Robustness to missing data	Robust: NNLS can be performed with missing data.	Robust: Missing measurements are interpolated using spherical harmonics.
Protocol transfer capabilities	Capable: No fine-tuning is required for protocol transfer.	Not capable.
Sensitive to external routine used for orientation estimation	Very sensitive.	Not sensitive since no external routine is used.

and uses all diffusion measurements in a voxel along with external estimates of fascicle orientations to determine the properties of each fascicle. In contrast, the Fully-Learned Method leverages a spherical harmonics representation of shells derived from raw diffusion measurements within a voxel. A key distinction between these methods lies in the Hybrid Method's use of a latent space representation through non-negative least squares, which allows a high interpretability and protocol transfer. On the other hand, the Fully-Learned Method's employment of a single deep neural network contributes to its enhanced accuracy and speedup factor.

Our results demonstrate that both the Hybrid and Fully-Learned Methods yield massive acceleration factors compared to the reference Microstructure Fingerprinting method while outperforming the reference method in estimation accuracy, in various simulated settings. Each method presents its unique advantages and limitations (as summarized in Table 4.5), highlighting the potential of deep learning in accelerating multi-dictionary matching problems in DW-MRI analysis. Notably, our findings reveal that both proposed methods surpass the reference method in most volume fraction estimations, even when the reference method utilizes ground truth values for fascicle orientations. This underscores the efficacy of deep learning approaches in providing more accurate and efficient solutions for complex interpretation of DW-MRI data, paving the way for their increased application in the field.

Funding

Quentin Dessain is a research fellow of the Fonds de la Recherche Scientifique - FNRS of Belgium. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region.

Acknowledgments

Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. The authors thank Dr. Ron Kupers and Dr. Laurence Dricot for their constructive comments on the manuscript.

Data availability statement

The data and code supporting the findings of this study are available in the GitHub repository of the article at https://github.com/Hyedryn/FastMF_paper/. The repository includes a precalculated dictionary of microstructural features, weights of models trained with PyTorch, and the Python code that encompasses scripts used for voxels simulations, models training, and data analysis.

5

Leveraging Swin Transformer for enhanced diagnosis of Alzheimer's disease using multi-shell diffusion MRI

Preliminary note:

Before starting the chapter, readers are encouraged to familiarize themselves with the following basic principles in Chapter 2:

- Diffusion Tensor Imaging (DTI) (2.1.2)
- Neurite Orientation Dispersion and Density Imaging (NODDI) (2.1.4)
- Vision transformer and the Shifted window (Swin) architecture (2.2.2)

Adapted version of an article in preparation for submission as:

Dessain, Q., Delinte, N., Hanseeuw, B., Dricot, L., and Macq, B. Leveraging Swin Transformer for enhanced diagnosis of Alzheimer's disease Using multi-shell diffusion MRI.

5.1 Introduction

Diffusion Magnetic Resonance Imaging (dMRI) has emerged as a powerful modality to investigate the microstructural properties of brain tissue. By analyzing the diffusion of water molecules in the brain, dMRI provides insights into the integrity of white matter tracts and the microstructural organization of brain tissue. In particular, the combination of high angular resolution diffusion imaging (HARDI) with multi-shell dMRI, offers a detailed characterization of tissue microstructure, positioning it as a promising tool for diagnosing and understanding the underlying biological changes associated with neurodegenerative diseases (Frank, 2001).

However, the high dimensionality and complexity of multi-shell dMRI data present significant challenges for traditional analysis methods. Deep learning has shown promise in automating feature extraction and classification from high-dimensional neuroimaging data (Litjens, Kooi, Bejnordi et al., 2017). Convolutional Neural Networks (CNNs), in particular, have demonstrated substantial success in medical image analysis, including brain MRI tasks, due to their ability to capture local spatial features (LeCun, Bottou, Bengio et al., 1998). These approaches have been successfully applied to structural MRI (sMRI) and functional MRI (fMRI) for Alzheimer's disease classification (Wen, Thibeau-Sutre, Diaz-Melo et al., 2020), with models such as hierarchical fully CNNs extracting multi-scale features from sMRI (Lian, Liu, Zhang et al., 2020) and spatiotemporal 3D CNNs capturing dynamics in fMRI data (Parmar, Nutter, Long et al., 2020). However, CNNs primarily focus on local spatial features, which may limit their ability to model global context and long-range dependencies.

To address these limitations, transformer models, originally developed for natural language processing tasks, have recently gained traction in medical imaging due to their ability to capture both local and global context through self-attention mechanisms. The Vision Transformer, which applies these principles to image data, has shown competitive performance across various computer vision tasks by leveraging its ability to model long-range dependencies in images (Dosovitskiy, Beyer, Kolesnikov et al., 2020). In the neuroimaging field, pre-trained ViTs have demonstrated promising results in tasks such as brain tumor classification and Alzheimer's disease diagnosis using MRI (Alp, Akan, Bhuiyan

et al., 2024; He, Gan, Li et al., 2023; Shamshad, Khan, Zamir et al., 2023). However, ViTs typically require large-scale labeled datasets for effective training, which can be a limiting factor in medical imaging where labeled data are often scarce (Touvron, Cord, Douze et al., 2021). While using pre-trained ViT can alleviate the need for extensive labeled data, ViTs still face challenges in medical imaging due to their fixed input size. This constraint limits the usability of pre-trained models when the input dimensions of medical images differ from the size expected by the pre-trained ViT (Liu, Lin, Cao et al., 2021).

The Swin Transformer, a variant of the ViT, addresses the former challenges by introducing a hierarchical structure and a shifted window approach, allowing it to capture both fine-grained local features and broader global dependencies while supporting variable input sizes (Liu, Lin, Cao et al., 2021). This flexibility is particularly beneficial for medical imaging, where the scale and resolution of images can vary significantly, and the ability to leverage pre-trained weights is critical for maximizing performance in the context of limited data availability (Kim, Cosa-Linan, Santhanam et al., 2022; Morid, Borjali, and Del Fiol, 2021).

The Swin Transformer's ability to maintain high-resolution representations while efficiently handling large images makes it an ideal choice for analyzing complex neuroimaging data, such as those derived from multi-shell dMRI. This ability to work with high-resolution representations, combined with the use of pre-trained weights, allows Swin Transformers to not only improve the adaptability of models to different image sizes and resolutions, but also enhances the potential for accurate diagnosis and prognosis of neuropathologies such as Alzheimer's disease by leveraging diffusion MRI data.

The recent inclusion of multi-shell diffusion MRI in the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset presents a valuable opportunity for developing new diagnostic tools (Zavaliangos-Petropulu, Nir, Thomopoulos et al., 2019). Historically, most large diffusion MRI datasets were limited to single-shell acquisitions (Mueller, Weiner, Thal et al., 2005), suitable primarily for Apparent Diffusion Coefficient (ADC) and DTI (Basser, Mattiello, and LeBihan, 1994). While foundational, these models are constrained by simplified assumptions about water diffusion, limiting their ability to characterize complex microstructural features such as crossing fibers and tissue heterogeneity. Multi-shell dMRI overcomes these limitations by enabling more sophisticated models that provide deeper insights into tissue integrity and pathology, making it particularly

valuable for studying Alzheimer's disease, since early diagnosis and monitoring of Alzheimer's disease are crucial for effective disease management, with the potential to delay progression and improve the quality of life for affected individuals. However, despite significant advances in neuroimaging and clinical diagnostics using DTI (Acosta-Cabronero, Alley, Williams et al., 2012; Ahmed, Benois-Pineau, Allard et al., 2017; Demirhan, Nir, Zavaliangos-Petropulu et al., 2015; Lella, Pazienza, Lofù et al., 2021; Marzban, Eldeib, Yassine et al., 2020), identifying reliable biomarkers for the early stages of Alzheimer remains a significant challenge.

In this study, we propose an architecture that leverages a multi-shell dMRI dataset and the recent developments of vision-oriented transformer models. The imaging data are preprocessed, and microstructural models, such as DTI (Basser, Mattiello, and LeBihan, 1994) and NODDI (Zhang, Schneider, Wheeler-Kingshott et al., 2012), are derived. The 3D microstructural maps estimated with these models were projected onto 2D planes, enabling the Swin Transformer to effectively operate on these 2D projections while retaining the advantages of pre-trained weights. By employing this advanced architecture and utilizing dMRI's insights into microstructural changes, we aim to enhance the understanding and diagnosis of Alzheimer's disease.

5.1.1 Contributions

- We highlight the use of multi-shell diffusion MRI for Alzheimer's disease classification, leveraging the microstructural information available from diffusion models such as DTI and NODDI.
- We demonstrate the possibility of using multi-shell diffusion MRI for predicting amyloid beta status, a key biomarker for early detection of Alzheimer's disease.
- To efficiently process diffusion MRI data, we project 3D microstructural maps into 2D representations, allowing the use of pre-trained Swin Transformer while maintaining essential spatial information.
- We incorporate LoRA, a parameter-efficient fine-tuning approach, into our deep learning framework, enhancing adaptability to neuroimaging tasks with limited labeled data while leveraging pre-trained ImageNet weights.

- Through Grad-CAM-based explainability, we highlight the clinically relevant brain regions that contribute to model classification decisions, offering valuable interpretability for medical applications.
- Our framework is designed to be adaptable and extendable to other neuroimaging modalities and medical imaging tasks, paving the way for future applications beyond Alzheimer's disease.

5.2 Materials and methods

5.2.1 Dataset

The neuroimaging data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu), specifically from the ADNI3 and ADNI4 phases. The ADNI was launched in 2003 as a public–private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, Positron Emission Tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease.

For this study, we used both multi-shell dMRI data and T1-weighted sMRI data from the ADNI repository. The dataset includes high-resolution scans from participants categorized into three diagnostic groups: Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's disease Dementia (AD).

To leverage metrics from multi-shell models, single-shell dMRI acquisitions present in the ADNI dataset were excluded. All dMRI scans were acquired using a 3.0 Tesla Siemens MAGNETOM Prisma/Prisma^{fit} scanner with a 2D echo-planar sequence. The voxel size was 2.0 × 2.0 × 2.0 mm. The acquisition parameters included an echo time (TE) of 71.0 ms, a repetition time (TR) of 3400.0 ms, and a flip angle of 90 degrees. The diffusion-weighted imaging protocol had 6 gradient directions at $b = 500$, 48 directions at $b = 1000$ and 60 directions at $b = 2000\,\mu s\,\mu m^{-2}$, along with 13 $b = 0$ acquisitions, for a total of 127 measurements. The matrix size for the scans was 116 × 116 × 81 pixels.

T1-weighted sMRI scans using the MPRAGE sequence, acquired during the same session as the diffusion data, were included to provide

anatomical context for the neuroimaging analysis. These MPRAGE scans were acquired in the sagittal plane using a 3D gradient-recalled inversion recovery sequence, with a voxel size of 1.0 × 1.0 × 1.0 mm. The acquisition parameters included a TE of 3.0 ms, a TR of 2300.0 ms, and an inversion time of 900.0 ms. The flip angle was set to 9 degrees, and the matrix size was 240 × 256 × 208 pixels.

Table 5.1 Demographic characteristics of selected participants from the ADNI dataset. When available, subjects were further subclassified by amyloid beta (Aβ) status into Aβ^- and Aβ^+.

	AD	MCI	CN
Number of subjects	37	121	213
Amyloid status (Aβ^-/Aβ^+)	0/17	42/40	107/54
Sex (Male/Female)	17/20	65/56	72/141
Age (Mean ± SD)	77.2 ± 6.9	73.3 ± 7.2	72.9 ± 8.7

A total of 371 participants were included after preprocessing, comprising 37 individuals diagnosed with AD, 121 with MCI, and 213 CN subjects. The demographic characteristics are summarized in Table 5.1.

5.2.2 Processing the dMRI data

The preprocessing of diffusion MRI data was performed using ElikoPy (Dessain, Mathieu, and Delinte, 2024), incorporating a series of specialized algorithms to ensure high-quality data suitable for subsequent analysis. The preprocessing began with MP-PCA denoising (Veraart, Novikov, Christiaens et al., 2016), applied through MRtrix3's dwidenoise function (version 3.0.3) (Tournier, Smith, Raffelt et al., 2019). Following denoising, FSL's eddy tool (version 6.0.7) (Andersson and Sotiropoulos, 2016) was employed for both head motion correction and eddy current correction. Additionally, b=0 reference images were processed alongside T1-weighted images using Synb0disco (Schilling, Blaber, Huo et al., 2019) to estimate the susceptibility-induced off-resonance field. These field maps were then integrated into the eddy correction process, ensuring that both eddy current and head motion were accurately corrected during the interpolation phase. Finally, post-eddy alignment of shells was performed to further refine the data, ensuring precise spatial consistency across all diffusion-weighted images.

After preprocessing, microstructural models were derived from the multi-shell dMRI data, and three key metrics were selected for each model as inputs to the Swin Transformer. For DTI, the metrics included Fractional Anisotropy (FA), which measures the degree of anisotropy in water diffusion; Axial Diffusivity (AxD), indicating water diffusion along the principal axis of the diffusion tensor; and Radial Diffusivity (RD), representing diffusion perpendicular to the principal axis. For NODDI, the selected metrics were the Orientation Dispersion Index (ODI), which quantifies the spread of neurite orientations; the Intra-neurite Volume Fraction (f_{intra}), reflecting the proportion of the diffusion MRI signal from neurites; and the Extra-neurite Volume Fraction (f_{extra}), representing the signal from the extracellular space. These metrics were chosen for their relevance in characterizing microstructural properties and providing rich input features for the transformer model. The 3D microstructural maps derived from the dMRI data were subsequently used as input for the deep learning models developed in this study.

5.2.3 Network architecture

For each model, the three selected 3D microstructural maps are projected into 2D space through orthogonal slicing, creating three-channel images where each channel corresponds to a specific microstructural map, as illustrated in Figure 5.1.

$$\mathbf{I}(x, y, c) = \mathbf{M}_c(x, y) \quad \text{for } c \in \{1, 2, 3\}. \tag{5.1}$$

where $\mathbf{I}(x, y, c)$ represents the resulting 2D image at coordinates (x, y), with channel $c \in \{1, 2, 3\}$ corresponding to one of the microstructural maps: $\mathbf{M}_1(x, y)$, $\mathbf{M}_2(x, y)$, or $\mathbf{M}_3(x, y)$. These three-channel images are then input into a fine-tuned Swin Transformer architecture, pre-trained on the ImageNet dataset (Deng, Dong, Socher et al., 2009).

To enhance the model's generalization across different native spaces and reduce overfitting to specific subject characteristics, we applied a data augmentation strategy to the 3D microstructural volumes before slicing them into 2D images. The augmentation technique consisted in synchronized translations by up to 5 voxels applied uniformly across the 3D volumes for all three channels, corresponding to the different microstructural maps.

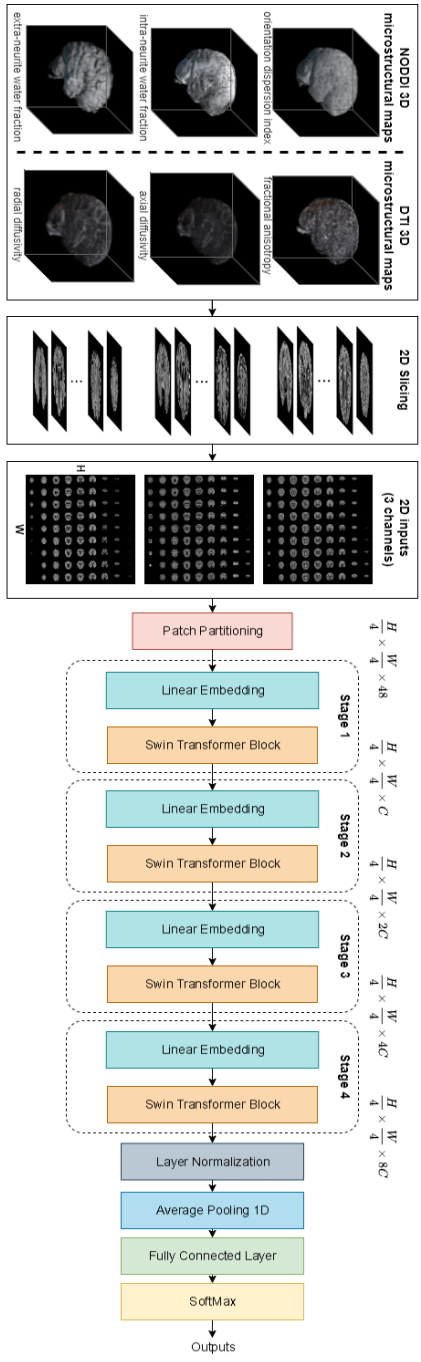


Fig. 5.1 Overview of the proposed architecture. The 3D microstructural maps derived from diffusion MRI data are projected into 2D slices, which serve as input to the Swin Transformer pre-trained on ImageNet. The Swin Transformer blocks are fine-tuned with LoRA.

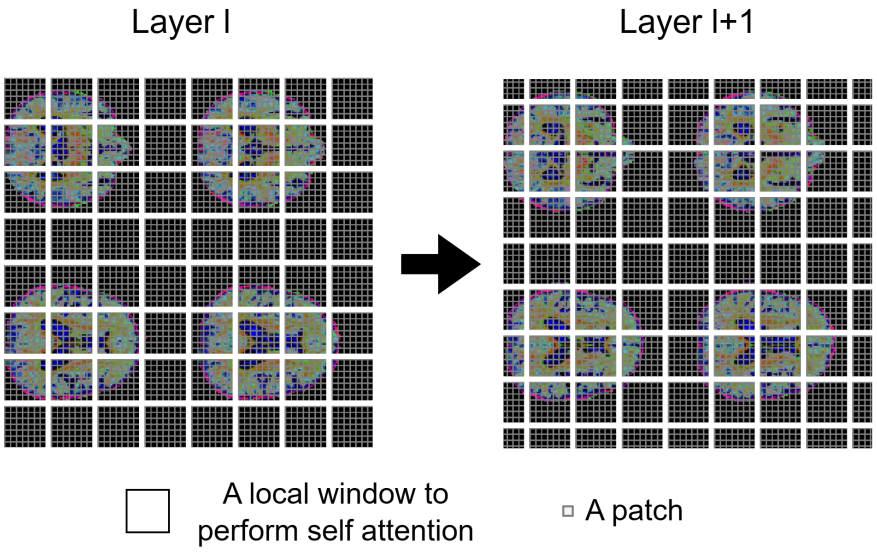


Fig. 5.2 An illustration of the shifted window approach in the Swin Transformer architecture. In the first layer (left), non-overlapping windows are used for self-attention computation within each window. In the subsequent layer (right), the window partitions are shifted, allowing for connections across the boundaries of the previous windows. This shift enables the model to capture dependencies between adjacent regions, improving its ability to model global context while maintaining computational efficiency.

The Swin Transformer employs a hierarchical attention mechanism that uses a shifted window approach, as depicted in Figure 5.2. This approach allows the model to capture both local and global dependencies within the input images. In the shifted window strategy, each self-attention layer operates on a partitioned image, with attention computed locally within non-overlapping windows. In the subsequent layer, the windows are shifted, allowing the model to compute attention across different image regions. This shifting mechanism ensures that the Swin Transformer can model long-range dependencies, which is crucial for processing the spatial relationships in neuroimaging data.

After adapting the Swin Transformer's input to handle 2D projections of 3D microstructural maps, we fine-tuned its weights using dMRI metrics. The fine-tuning involved the incorporation of LoRA (Hu, Shen, Wal-

lis et al., 2021) to refine the attention mechanisms efficiently, and the customization of the output layers for the classification of cognitive impairment categories such as CN, MCI and AD. By integrating LoRA into each attention layer of the Swin Transformer, we adapted the model for early Alzheimer's disease diagnosis while minimizing the number of additional parameters.

The overall architecture, from the generation of microstructural maps to the final classification, is illustrated in Figure 5.1.

5.2.4 Experiments

To evaluate the effectiveness of the proposed architecture, we conducted a series of experiments using diffusion MRI data from the ADNI dataset. A rigorous data splitting strategy was employed to ensure robust evaluation and prevent data leakage. First, 15% of unique subjects were randomly assigned to a fixed hold-out test set, reserved exclusively for final model assessment. The remaining 85% of subjects were used for 5-fold cross-validation. The cross-validation folds were generated using a grouped and stratified approach. This procedure guaranteed that all data from any single subject belonged entirely to either the training set or the validation set within any given fold, preventing optimistic bias due to subject-specific correlations. Concurrently, it aimed to preserve the original proportion of diagnostic classes within each fold, ensuring that each set was representative of the overall dataset's class balance. All splits were generated using a fixed random seed for reproducibility. Furthermore, to address class imbalance within the training data of each fold, class weights were utilized within the Cross-Entropy loss function during model training.

All subsequent experiments, including model training and inference, were performed on a NVIDIA Tesla A100 GPU to ensure efficient handling of the high-dimensional dMRI data and computational demands of the Swin Transformer. The model was trained using standard backpropagation with the Adam optimizer (Kingma and Ba, 2014) and a cosine annealing learning rate schedule (Loshchilov and Hutter, 2016). The experiments focused on comparing the impact of different microstructural models, analyzing the effect of architectural components via an ablation study, investigating the framework's diagnostic potential for classifying amyloid beta status, and interpreting model decisions through explainable AI (xAI).

dMRI model study

To assess the impact of different microstructural dMRI models on the performance of the proposed Swin Transformer-based architecture, we conducted experiments utilizing two common dMRI models: DTI and NODDI. These models were chosen to highlight the differences between a single-shell model (DTI) and a multi-shell model (NODDI), reflecting their distinct capabilities in capturing microstructural characteristics relevant to Alzheimer's disease classification.

The Swin Transformer was trained and evaluated separately on the datasets corresponding to each dMRI model. This allowed us to directly compare the classification performance of the architecture when utilizing DTI-derived metrics versus NODDI-derived metrics. By comparing these models, we aimed to determine which set of diffusion metrics offers better sensitivity and specificity for classifying AD, MCI, and CN subjects.

We evaluated the model's performance on three classification tasks involving different diagnostic groups: (1) distinguishing CN subjects from individuals with AD, (2) differentiating CN from subjects with MCI, and (3) separating MCI from AD cases. The performance metrics (accuracy, precision, recall, and F1 score) are presented in the results section, providing insight into the effectiveness of each dMRI model.

Ablation study

To further understand the contributions of each component in our proposed architecture, we conducted an ablation study that systematically evaluates the impact of different architectural choices, including the use of LoRA, and the Swin Transformer's hierarchical attention mechanism.

We designed four main experiments to isolate the effects of these components:

1. **Baseline Resnet**: A 2D Resnet CNN architecture pre-trained on ImageNet was used with the same 2D projections of the dMRI data but without the Swin Transformer or LoRA. This model served as a baseline to compare how the hierarchical attention mechanism and pre-trained weights affect performance.
2. **Swin Transformer without LoRA (full fine-tuning)**: In this configuration, we employed the Swin Transformer pre-trained on ImageNet without the incorporation of LoRA. This experiment assesses how

the Swin Transformer performs with full-weight fine-tuning on dMRI projections.

3. **Swin Transformer without LoRA (head fine-tuning)**: As before, in this configuration, we employed the Swin Transformer pre-trained on ImageNet without the incorporation of LoRA, but only the classification head is finetuned. This experiment assesses how the Swin Transformer performs with classification head weight fine-tuning on dMRI projections.

4. **Swin Transformer with LoRA**: In this experiment, we applied our proposed architecture with LoRA integrated into the Swin Transformer, evaluating the benefits of LoRA's fine-tuning on the Alzheimer's disease classification task.

Each configuration was trained and evaluated using the same dataset splits to ensure a fair comparison. The performance metrics, including balanced accuracy, precision, recall, and F1 score, were computed for each configuration and compared to understand the individual and combined effects of these components on classification performance.

Amyloid classification experiments

To further investigate the diagnostic capabilities of diffusion MRI combined with the Swin Transformer framework, we conducted two experiments focused on classifying amyloid beta status. Amyloid beta is a key biomarker of Alzheimer's disease, often accumulating years before cognitive symptoms appear, making it a valuable target for early detection and intervention in preclinical stages (Jack Jr, Bennett, Blennow et al., 2018; Sperling, Aisen, Beckett et al., 2011; Villemagne, Burnham, Bourgeat et al., 2013). Both DTI and NODDI-derived microstructural maps were used to evaluate the discriminatory power of these metrics for amyloid classification.

Amyloid positivity was determined using the UC Berkeley - Amyloid PET 6mm Res analysis derivatives available on the ADNI website. As described in their method, the labels were derived from cortical summary standardized uptake value ratios calculated from co-registered amyloid PET and MRI data. Subjects with missing amyloid status information were excluded from the analysis.

Experiment 1: Control vs. case classification This experiment aimed to distinguish amyloid-negative cognitively normal ($A\beta^-$ CN) subjects from amyloid-positive individuals diagnosed with either MCI or AD ($A\beta^+$ MCI/AD). The amyloid-positive group included 40 MCI and 17 AD subjects and the amyloid-negative group was composed of 107 control subjects. This approach ensures a clearer distinction between healthy control and amyloid-positive cases.

Experiment 2: Amyloid classification in cognitively normal subjects The second experiment focused on detecting amyloid beta positivity among CN subjects to assess whether diffusion MRI metrics could distinguish between preclinical amyloid-positive individuals and those without amyloid accumulation. Out of the original ADNI dMRI cohort of 213 CN subjects, 52 individuals were excluded due to missing amyloid status information. The remaining dataset consisted of 161 subjects, categorized into 54 amyloid-positive ($A\beta^+$) and 107 amyloid-negative ($A\beta^-$) cases.

Explainability

To provide insights into the decision-making process of the proposed Swin Transformer-based architecture, we employed Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju, Cogswell, Das et al., 2017) to visualize which regions of the input contributed most to the model's predictions. Grad-CAM highlights areas in the input data that are most influential for the classification, which is crucial for enhancing the interpretability of deep learning models in medical imaging.

We applied Grad-CAM to the test set, focusing on correctly classified subjects from each diagnostic group, namely AD, MCI and CN. To obtain a more comprehensive and consistent understanding of the model's decision-making, we averaged the Grad-CAM output across all correctly classified subjects within each group. This averaging allowed us to identify common patterns and regions of interest (ROI) in the brain that were consistently emphasized by the model during classification. The importance of each region was quantified by calculating the mean value assigned by Grad-CAM to each brain parcellation, based on the Brainnetome atlas (246 ROIs) (Fan, Li, Zhuo et al., 2016) and the SUIT cerebellum atlas (34 ROIs) (Diedrichsen, 2006).

5.3 Results

5.3.1 Classification performance

We evaluated the proposed framework's performance in classifying participants into diagnostic categories using the fine-tuned pre-trained Swin Transformer. We conducted experiments using microstructural maps derived from both DTI and NODDI models. The classification tasks included distinguishing between CN and AD, CN and MCI, and MCI and AD groups.

The classification metrics, including balanced accuracy, precision, recall, and F1 score, are presented in Table 5.2.

Table 5.2 Classification results using Swin Transformer on ADNI data for DTI and NODDI models.

	Balanced Accuracy	**Precision**	**Recall**	**F1 Score**
DTI				
CN vs AD	92.2%	94.6%	85.7%	88.4%
CN vs MCI	68.1%	73.2%	**66.7%**	**68.2%**
MCI vs AD	67.5%	73.3%	72.0%	72.5%
NODDI				
CN vs AD	**95.3%**	**95.7%**	**91.4%**	**92.7%**
CN vs MCI	**74.4%**	**80.8%**	**66.7%**	67.9%
MCI vs AD	**83.3%**	**87.1%**	**76.0%**	**77.2%**

The results indicate that the model achieved higher classification performance when using NODDI-derived microstructural maps compared to DTI. In particular, the CN vs AD classification task yielded a balanced accuracy of 95.3% with NODDI, compared to 92.2% with DTI. Similarly, the MCI vs AD classification showed significant improvement with NODDI, achieving a balanced accuracy of 83.3%.

5.3.2 Ablation study

The ablation study results, presented in Tables 5.3 and 5.4 highlight the impact of different model configurations on AD vs CN and MCI vs CN classification tasks.

For the AD vs CN classification, the baseline ResNet achieved a balanced accuracy of 90.6% with an F1-score of 86.3%. Full fine-tuning of the Swin Transformer underperformed with 50% balanced accuracy, likely due to overfitting. Head fine-tuning improved performance to 92.2% balanced accuracy and an F1-score of 88.4%. Adding LoRA further enhanced the model's performance, achieving the highest balanced accuracy of 95.3% and an F1-score of 92.6%.

For the MCI vs CN task, the baseline ResNet showed moderate performance with a balanced accuracy of 71.4%, and whereas precision was high (79.5%), recall was limited (62.5%), reflecting difficulty in identifying all MCI cases. Full fine-tuning resulted again in poor performance with a balanced accuracy of 50%. Head fine-tuning improved results, reaching 70.6% balanced accuracy and a precision of 82.8%, though recall remained low. The integration of LoRA achieved the best results with a balanced accuracy of 74.4% and improved recall and F1-scores.

Table 5.3 Ablation study: Performance of the model under different settings with the NODDI microstructural model for the AD vs CN classification task.

Model Configuration	Balanced Accuracy	Precision	Recall	F1 Score
Baseline Resnet (2D)	90.6%	94.3%	82.9%	86.3%
Swin Transformer (full fine-tuning)	50.0%	50.2%	70.8%	58.7%
Swin Transformer (head fine-tuning)	92.2%	94.6%	85.7%	88.4%
Swin Transformer (LoRA & head fine-tuning)	**95.3%**	**95.7%**	**91.4%**	**92.6%**

5.3.3 Amyloid classification results

We evaluated the performance of the Swin Transformer with LoRA fine-tuning in two amyloid classification tasks using DTI- and NODDI-derived

Table 5.4 Ablation study: Performance of the model under different settings with the NODDI microstructural model for the MCI vs CN classification task.

Model Configuration	Balanced Accuracy	Precision	Recall	F1 Score
Baseline Resnet (2D)	71.4%	79.5%	62.5%	63.5%
Swin Transformer (full fine-tuning)	50.0%	8.5%	29.2%	13.2%
Swin Transformer (head fine-tuning)	70.6%	**82.8%**	58.3%	58.3%
Swin Transformer (LoRA & head fine-tuning)	**74.4%**	80.8%	**66.7%**	**67.9%**

microstructural metrics. The model achieved balanced accuracy scores above chance level in both experiments, with NODDI metrics consistently outperforming DTI metrics.

Experiment 1: Control vs. case classification In this binary classification task, the model distinguished amyloid-negative cognitively normal ($A\beta^-$ CN) subjects from amyloid-positive cases diagnosed with MCI or AD ($A\beta^+$ MCI/AD). Table 5.5 summarizes the performance metrics, indicating that NODDI-derived metrics yielded higher accuracy, precision, and recall compared to DTI-derived metrics.

Experiment 2: Amyloid classification in cognitively normal subjects In this experiment, we evaluated the model's ability to classify amyloid status in cognitively normal subjects, identifying individuals with preclinical amyloid accumulation. As shown in Table 5.5, NODDI metrics again outperformed DTI metrics.

5.3.4 Explainability analysis

To interpret the model's decisions, we employed Grad-CAM (Selvaraju, Cogswell, Das et al., 2017) to visualize the regions contributing to the classification.

The activation map resulting from the classification of the CN vs AD groups with the Swin Transformer, enhanced with LoRA and NODDI metrics, is shown in Figure 5.3. Table 5.6 summarizes the most important brain regions identified by Grad-CAM. Specifically, the table displays the top

Table 5.5 Performance metrics for amyloid classification with Swin Transformer on ADNI data for DTI and NODDI models.

	Balanced Accuracy	Precision	Recall	F1 Score
DTI				
$A\beta^-$ CN vs $A\beta^+$ MCI/AD	71.4%	77.1%	66.7%	66.2%
$A\beta^-$ CN vs $A\beta^+$ CN	58.1%	69.5%	57.1%	60.3%
NODDI				
$A\beta^-$ CN vs $A\beta^+$ MCI/AD	**82.1%**	**88.7%**	**86.1%**	**85.3%**
$A\beta^-$ CN vs $A\beta^+$ CN	**77.5%**	**82.2%**	**76.2%**	**77.7%**

five most significant *gyri* based on their mean activation values. Within each gyrus, subregions with mean activation values exceeding a threshold of 0.37 were selected, emphasizing the areas that contributed most to the model's classification decisions.

The analysis revealed that the model attributed more weight to specific brain areas, such as the parahippocampal gyrus (including the entorhinal cortex), cerebellum, hippocampus, superior frontal gyrus, and paracentral lobule.

In contrast, the Grad-CAMs from other classification tasks, with lower balanced accuracy scores, did not reveal any distinct regions of interest.

Table 5.6 Top subregions within the five most important *gyri* identified by Grad-CAM, ranked by their mean activation values. Only subregions with mean values greater than 0.37 are displayed.

Lobe	Gyrus	Subregion	Mean Value
Temporal Lobe	Parahippocampal Gyrus (PhG)	TL (lateral posterior PhG)	0.445
		TH (medial posterior PhG)	0.413
		A28/34 (entorhinal cortex)	0.373
Frontal Lobe	Superior Frontal Gyrus	dorsolateral area	0.415
	Paracentral Lobule	area 4, (lower limb region)	0.396
Cerebellum	/	/	0.373
Subcortical Nuclei	Hippocampus	rHipp (rostral hippocampus)	0.379

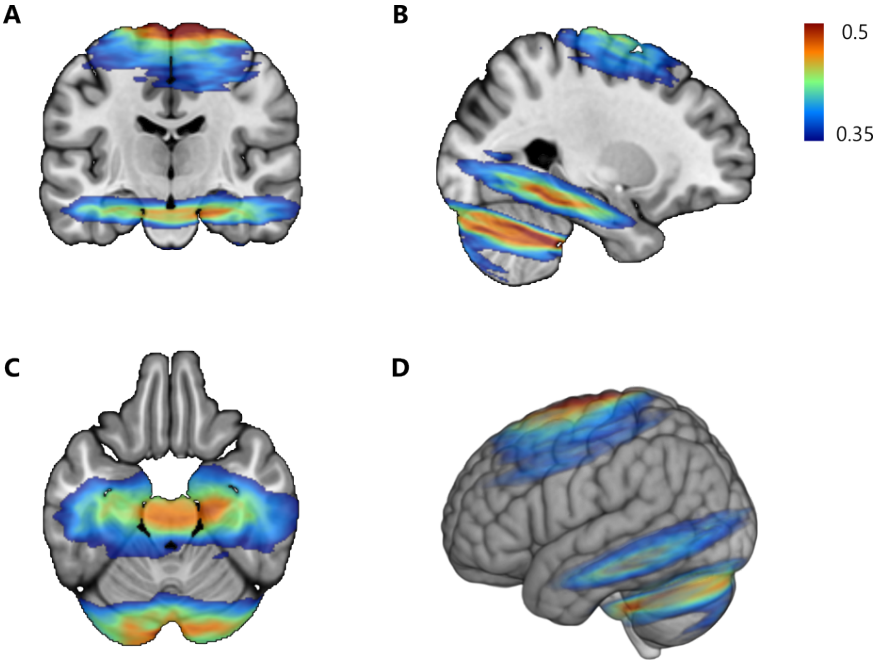


Fig. 5.3 Average activation map generated by Grad-CAM for the subjects correctly classified as AD, highlighting important regions inside the brain contributing to the classification. The regions are represented in 2D (**A** coronal, **B** sagittal and **C** axial) and 3D (**D**) views in the MNI space.

5.4 Discussion

5.4.1 Improved classification performance

Our study demonstrates that leveraging advanced microstructural imaging with the Swin Transformer architecture improves the classification performance for early diagnosis of Alzheimer's disease. The use of NODDI-derived microstructural maps resulted in higher accuracy compared to DTI, suggesting that NODDI provides more sensitive measures of microstructural changes associated with Alzheimer's disease.

Additionally, our framework was evaluated on amyloid classification tasks, where distinguishing between amyloid-positive and amyloid-negative subjects is critical for early detection of Alzheimer's disease

related pathology. In particular, we examined the classification of cognitively normal subjects based on amyloid positivity. This task targets preclinical stages of Alzheimer's disease where amyloid deposition is detectable long before the onset of cognitive symptoms. The classification performance for $A\beta^-$ CN vs $A\beta^+$ CN using NODDI-derived metrics yielded promising results, with high balanced accuracy and precision scores. This suggests that dMRI metrics, particularly those derived from NODDI, are sensitive enough to detect microstructural changes associated with early amyloid buildup.

5.4.2 Insights from the ablation study

The ablation study offers valuable insights into the contributions of various components in the proposed architecture. First, the baseline 2D Resnet model had inferior performances compared to the Swin Transformer with LoRa, suggesting that the hierarchical attention mechanism and the ability to model global dependencies are useful for capturing diffusion MRI patterns.

When comparing the different Swin Transformer configurations, the results indicate that fine-tuning the entire model (full fine-tuning) prevents the model from learning, with classification accuracy equivalent to random guessing. This suggests that full fine-tuning is ineffective for smaller datasets, such as the one used in this study, as it may cause the model to overfit or fail to learn meaningful representations.

The best performance was achieved when combining LoRA with head fine-tuning, underscoring the importance of parameter-efficient adaptation in medical imaging tasks, where the availability of labeled data is often limited. LoRA allows for a more targeted fine-tuning approach, updating only a small subset of parameters while maintaining strong generalization. The results also highlight LoRA's critical role in boosting performance. Without LoRA, head fine-tuning alone resulted in lower balanced accuracy, precision, recall, and F1 scores. By enabling more efficient training with few parameters, LoRA helps to preserve or enhance classification performance, making it a key component in this architecture. This is especially important in neuroimaging applications, where data scarcity can lead to overfitting, and parameter-efficient models are essential to ensure robust learning.

5.4.3 Explainability and clinical relevance

The explainability analysis using the Grad-CAM adapted for our architecture revealed that the model's attention was focused on regions known to be affected in Alzheimer's disease. This focus was observed in the classification task that achieved the highest balanced accuracy, namely the classification of CN versus AD cases using the Swin Transformer model with LoRA and NODDI metrics. The most prominent regions included the parahippocampal gyrus (including the entorhinal cortex), cerebellum, hippocampus, superior frontal gyrus and paracentral lobule.

These findings align with established neuropathological patterns in Alzheimer's disease. For example, a review by Rathore, Habes, Iftikhar et al. (2017) showed that DTI metrics were significantly impacted by Alzheimer's disease pathology in key regions such as the hippocampus and posterior cingulate cortex, consistent with our model's focus areas. This correspondence supports the biological plausibility of our model's learned features, enhancing its potential clinical relevance.

Furthermore, the regions highlighted by Grad-CAM are consistent with findings from previous studies on the ADNI dataset using CNNs and T1-weighted data, such as Jang and Hwang (2022). Additionally, studies on the ADNI dataset have reported variations in DTI metrics in the temporal and occipital lobes, as well as in the cingulate gyrus and hippocampus, when comparing AD patients to healthy controls (Mueller, Weiner, Thal et al., 2005). These regions are aligned with our observations, further supporting the model's potential to focus on clinically relevant areas. This alignment highlights the advantages of incorporating explainability features into advanced deep learning architectures.

5.4.4 Framework flexibility and adaptability

The proposed framework leverages pre-trained models and fine-tuning strategies, making it adaptable to other medical imaging applications. The use of 2D projections allows for the utilization of existing pre-trained transformer models, reducing the need for large training datasets. Moreover, the incorporation of LoRA enables efficient fine-tuning with minimal additional parameters, facilitating rapid adaptation to new tasks.

To encourage further research and promote reproducibility, the implementation of our framework is open-source and available on GitHub.

5.4.5 Comparison with state-of-the-art methods

A notable aspect of this study is the absence of a direct quantitative comparison between our proposed framework and other state-of-the-art methods for Alzheimer's disease or amyloid classification. This decision stems primarily from the distinct modality employed in our work. While numerous deep learning approaches have achieved high performance using structural MRI or functional MRI data (Lian, Liu, Zhang et al., 2020; Wen, Thibeau-Sutre, Diaz-Melo et al., 2020), the application of advanced deep learning architectures to multi-shell diffusion MRI data for these specific classification tasks remains relatively unexplored. A direct comparison across such disparate modalities could be misleading, given the fundamentally different types of biological information captured by dMRI versus sMRI or fMRI.

Instead of attempting a potentially unfair cross-modal comparison, we focused on rigorous internal validation through ablation studies and detailed performance reporting specific to dMRI metrics (DTI and NODDI). Furthermore, to facilitate future research and enable direct benchmarking within this specific domain, we have made our implementation open-source and provided the exact subject identifiers used for defining the training, validation, and hold-out test sets. This commitment to transparency and reproducibility aims to establish a reliable baseline against which future methods utilizing multi-shell dMRI for AD and amyloid classification can be directly and fairly compared.

5.4.6 Limitations and future work

Despite the promising results, our study has limitations. The sample size, particularly for the AD group, is relatively small, which may affect the generalizability of the findings. Future work should include larger and more diverse cohorts to validate the model's robustness. Additionally, while the use of 2D projections allows the framework to leverage pre-training on large-scale datasets like ImageNet, it may not fully capture the intricate 3D spatial relationships inherent in dMRI data. Extending the framework to process 3D volumes directly could potentially improve performance by preserving spatial context, but this would come at the cost of losing the advantages of pre-training and may require significantly larger datasets to train effectively from scratch. Balancing these trade-offs will be an important direction for future research.

Another limitation in our amyloid analysis relates to the sample size and dataset composition. While we compared Aβ^- CN versus Aβ^+ CN and Aβ^- CN versus Aβ^+ MCI/AD, the relatively small number of individuals with both multi-shell dMRI and amyloid status in the ADNI database limited the statistical power of these comparisons. This constraint may affect the robustness of the findings and restrict the exploration of amyloid's role in more granular diagnostic stratifications. However, as the ADNI database continues to grow, the availability of subjects with comprehensive imaging and biomarker data is expected to increase, enabling larger and more representative analyses.

Furthermore, tau PET imaging data are becoming increasingly available in the ADNI database. As tau is a critical biomarker of Alzheimer's disease alongside amyloid accumulation, incorporating both biomarkers could refine classification tasks and improve the biological interpretability of models. Future studies could integrate these complementary biomarkers to stratify subjects into more meaningful categories, enhancing the granularity and precision of diagnostic tools. These advances, coupled with larger and more diverse datasets, will also facilitate the training of more sophisticated multi-modal deep learning frameworks.

5.5 Concluding remarks

Our study demonstrates the potential of combining a pre-trained Swin Transformer architecture with advanced microstructural models for accurate classification using diffusion MRI. By leveraging diffusion models such as NODDI and DTI, along with parameter-efficient techniques like LoRA, we were able to capture subtle microstructural changes for neuroimaging-based classification tasks. The results underline LoRA's contribution to enhancing classification accuracy, particularly in data-constrained scenarios, by optimizing model adaptation without extensive retraining. As diffusion MRI continues to evolve as a critical tool for probing tissue microstructure, the presented methods hold significant promise for advancing research applications in neuroimaging.

Funding

Quentin Dessain is a research fellow of the Fonds de la Recherche Scientifique - FNRS of Belgium. Nicolas Delinte was funded by the MedReSyst-AI4Alzheimer project, which has been supported by the European Union and Wallonia as part of the "Wallonia 2021-2027" program. Bernard Hanseeuw is a clinical researcher of FNRS, supported by the Wel Research Institute (WelBIO research program) and the StopAlzheimer Foundation. Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated

by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data availability statement

The code supporting the findings of this study along with the specific subject identifiers defining the test set and each cross-validation fold are provided in the GitHub repository of the article at https://github.com/Hyedryn/Swin_dMRI_public. The entire framework was implemented using Python and PyTorch.

6

Discussion and conclusion

Throughout my PhD journey, I have developed novel methodologies to address challenges in diffusion MRI, advancing the field through innovations that span acquisition, preprocessing, and analysis. While each chapter presents distinct technical contributions, my daily experiences as a researcher, from troubleshooting codes to collaborating with clinicians, revealed recurring challenges that extend beyond individual technical solutions.

Two key themes emerged from these experiences: the persistent tension between acquiring rich diffusion MRI data and maintaining practical scalable protocols, and the challenge of maintaining biological interpretability while pursuing engineering innovation. These themes resonated not only in my own research but also in discussions with colleagues, collaborators, and clinicians, highlighting questions that the field must address as it continues to mature.

The following sections explore each theme, drawing from both my technical work and practical experiences to examine how our methodological contributions address current limitations while pointing toward future opportunities for advancement. Through this analysis, I hope to contribute to the ongoing dialogue about developing diffusion MRI as a powerful and practical tool for understanding brain structure and function.

6.1 Key themes and challenges

6.1.1 Reconciling precision and scalability in dMRI acquisition and processing

A fundamental challenge in diffusion MRI research lies in balancing the competing demands of acquisition protocol richness and practical feasibility. This tension manifests throughout the diffusion MRI pipeline, from data collection to analysis. Multi-shell acquisitions with high b-values provide rich microstructural information but face significant practical barriers: longer scan times amplify sensitivity to motion, increase patient discomfort, and raise study costs, while higher b-values and complex protocols impose a substantial increase on computational demands for preprocessing and analyses. Conversely, single-shell protocols offer greater accessibility and efficiency but sacrifice the detailed tissue characterization that more sophisticated acquisitions enable.

Chapter 5 with its work on the ADNI dataset illustrates these trade-offs. NODDI analysis demonstrated superior diagnostic performance compared to DTI in detecting Alzheimer's disease-related changes, yet its requirement for multi-shell data significantly restricted the available sample size. The disparity in dataset size is striking: ADNI contains 1320 subjects with single-shell data suitable for DTI but only 371 subjects with the multi-shell acquisitions necessary for NODDI. Moreover, even the multi-shell dataset, comprising 127 directions across two shells, fall short of the requirements for more advanced techniques like Microstructure Fingerprinting, which typically demands additional shells and directions for robust parameter estimation. The reduced availability of multi-shell data can affect statistical power, limits the generalizability of findings, and complicates multi-site studies.

This thesis has made several contributions toward addressing these challenges. The motion correction framework developed for gradient cycling acquisitions represents an advance in making complex protocols more practically feasible. By enabling efficient correction of motion artifacts while maintaining thermal efficiency, this method helps preserve data quality without extending scan duration. The framework particularly benefits high b-value acquisitions, where motion sensitivity has traditionally been a limiting factor.

Similarly, the Fast Microstructure Fingerprinting methods address the computational challenges associated with sophisticated tissue modeling. By drastically accelerating parameter estimation, these methods make complex microstructural analysis more accessible for large-scale studies. The Hybrid Method's ability to generalize across protocols is particularly valuable, as it enables the application of advanced models even when acquisition parameters vary across sites or time points.

These advances demonstrate that technical innovation can help bridge the gap between precision and scalability. However, they also highlight the need for continued development of methods that can extract maximum information from limited data. The field's progress will likely require a multidimensional approach: continuing to develop methods that make sophisticated acquisitions more practical while simultaneously improving our ability to extract meaningful information from simpler protocols.

6.1.2 Bridging the gap between engineering and biological interpretability

Another challenge in diffusion MRI research is maintaining biological relevance while pursuing technical innovation. Modern analysis methods, particularly those leveraging deep learning, introduce sophisticated computational approaches that must balance performance with interpretability. This thesis explored various ways to address this challenge, demonstrating how engineering solutions can be developed while maintaining meaningful connections to tissue properties.

The FastMF methods developed in Chapter 4 illustrate this balance through two contrasting approaches. The Hybrid Method maintains stronger biological interpretability by incorporating a physics-based dictionary, allowing clear connections between model parameters and tissue properties, while the Fully-Learned Method prioritizes performance with less direct biological interpretation. However, both approaches share a common limitation: the underlying tissue models, though based on sophisticated Monte Carlo simulations, cannot uniquely map changes in metrics to specific biological alterations. For instance, variations in fiber volume fraction could reflect multiple different pathological processes, making it challenging to infer exact biological mechanisms of change.

Our Alzheimer's disease classification framework further demonstrates this tension between engineering advancement and biological interpretability. The framework incorporates well-established microstruc-

tural metrics from DTI and NODDI as inputs, and employs Grad-CAM analysis to reveal anatomically relevant activation patterns. However, these approaches face similar challenges to FastMF. While Grad-CAM highlights relevant brain regions, it provides only a high-level view of the model's decision-making process. Similarly, changes in DTI and NODDI metrics can result from various biological alterations, making it difficult to attribute specific pathological mechanisms to observed changes in these measures.

These limitations in biological specificity, while significant, should be viewed within the broader context of neuroimaging research. The field has consistently progressed by developing methods that, despite lacking perfect biological specificity, provide valuable insights when interpreted carefully within their theoretical frameworks. Although many dMRI methods do not perfectly represent the reality, it still brings biological insights that can then be interpreted through accumulated knowledge and hypothesis.

The integration of multiple imaging modalities offers another pathway to enhance biological interpretability. For instance, in my ongoing work studying long-term meditators presented in Appendix C, the combination of diffusion MRI and functional connectivity analyses reveals complementary insights into meditation-induced brain alterations. This multi-modal approach helps disambiguate the biological basis of meditation-related traits by providing convergent evidence across different neural measures.

Finally, I think that a crucial step in bridging the gap between engineering innovation and clinical application lies in promoting better dialogue between technical developers and clinical users. While engineers need to better communicate their models' assumptions and limitations through accessible documentation and training materials, clinicians would benefit from deeper engagement with the theoretical foundations of the tools they use. This two-way exchange, engineers making their work more accessible and clinicians developing a stronger grasp of model fundamentals, could significantly improve how diffusion MRI tools are applied in practice.

6.2 Concluding remarks

This thesis has advanced the field of diffusion MRI through three main technical contributions: a specialized motion correction framework that enables accurate correction of motion artifacts in gradient cycling acquisitions while maintaining thermal efficiency, accelerated methods for microstructure fingerprinting that achieve orders of magnitude faster computation while improving accuracy, and a deep learning approach using Swin Transformers that leverages multi-shell diffusion data for both Alzheimer's disease classification and prediction of amyloid beta presence. Each contribution addresses specific challenges across the diffusion MRI pipeline, spanning from preprocessing improvements to efficient microstructural analysis and ultimately diagnostic tools development.

Looking ahead, an exciting development is the ongoing collection of a comprehensive dataset among the Institute of Neuroscience of UCLouvain that will enable the integration of all methods presented in this thesis. This large dataset, which includes cognitively normal subjects and individuals across the Alzheimer's disease spectrum, employs an advanced multi-shell protocol with gradient cycling. The acquisition specifically enables the application of our complete methodological pipeline: the specialized motion correction algorithm will address artifacts introduced by gradient cycling, the FastMF algorithm will extract detailed microstructural information, and the Swin Transformer framework will facilitate diagnostic classification. This integrated approach will not only help validate the transferability of our ADNI findings but also demonstrate how our various technical contributions can work together to enhance diffusion MRI analysis.

A

Supplementary materials for motion correction in diffusion MRI acquisition with gradient cycling

A.1 Additional results

This section provides additional sagittal and coronal views of preprocessed diffusion MRI volumes acquired under gradient cycling for b-values of $1000\,\mathrm{s/mm^2}$ and $3000\,\mathrm{s/mm^2}$. These figures complement the results presented in the main text (Section 3.7) and further illustrate the performance of the slice-to-volume (S2V) correction compared to the volume-to-volume (V2V) correction across a range of diffusion encoding.

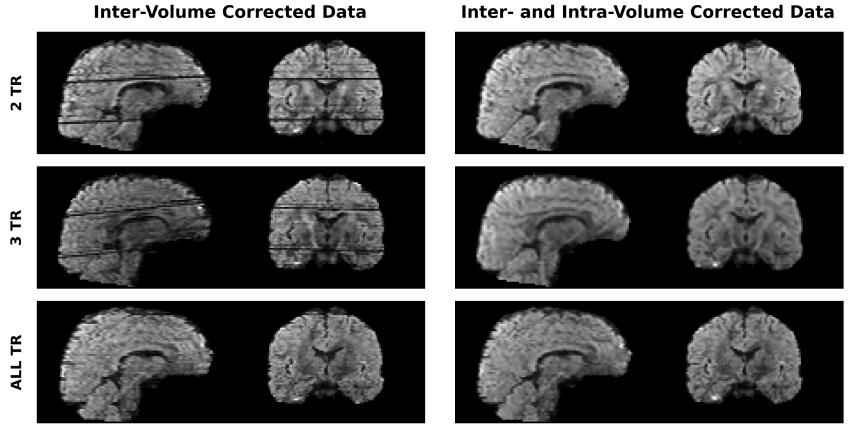


Fig. A.1 S2V correction reduces motion artifacts and improves slice alignment compared to V2V correction. Sagittal and coronal views of preprocessed diffusion MRI volumes acquired under gradient cycling for volumes with b-values of 1000 s/mm^2. The leftmost panels display results with volume-to-volume (V2V) correction, while the rightmost panels show results with slice-to-volume (S2V) correction.

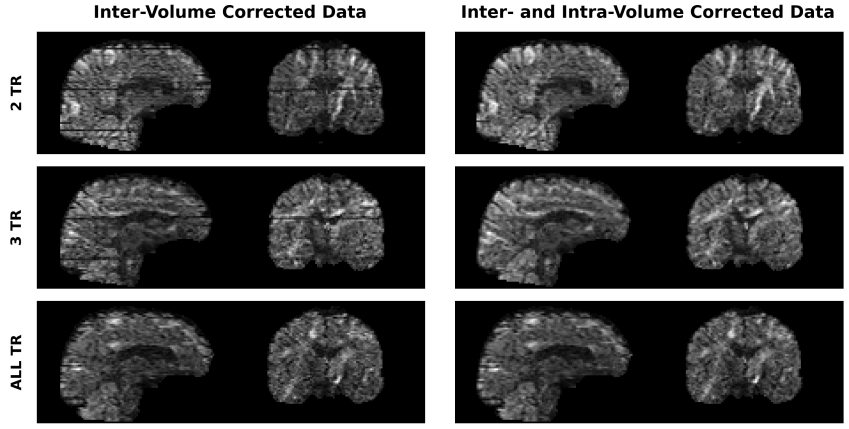


Fig. A.2 S2V correction reduces motion artifacts and improves slice alignment compared to V2V correction. Sagittal and coronal views of preprocessed diffusion MRI volumes acquired under gradient cycling for volumes with b-values of 3000 s/mm^2. The leftmost panels display results with volume-to-volume (V2V) correction, while the rightmost panels show results with slice-to-volume (S2V) correction.

B

Supplementary materials for fast multi-compartment microstructure fingerprinting

B.1 Effect of coarser microstructural parameters during training

To evaluate the necessity of close-range microstructural parameters during training, the performance of both accelerated methods were investigated when trained with a reduced number of fiber volume fraction (*fvf*) configurations. Specifically, a network was trained with a dictionary containing 19 equally spaced values within the range $[0.06, 0.78]$ for the fiber volume fraction instead of the original 38 equally spaced values within the range $[0.06, 0.8]$. The extra-axonal diffusivity (D_{ex}) was left unchanged with 10 equally spaced values in the $[0.6, 2.4]\,\mu m^2\,ms^{-1}$ range. Consequently, the total number of the precomputed fingerprints was reduced to $N = 190$ from the original $N = 380$.

The training process was identical to that used for the original two-fascicle problem, but with the adapted dictionary. The resulting models were evaluated on the test set from Experiment I.

As shown in Figure B.1, the performance of both accelerated methods remained consistent despite the reduced granularity in the training set. The mean absolute errors (MAEs) for the fiber volume fractions showed minimal deviation when compared to the models trained on the full training set. The Hybrid Method yielded MAEs of 0.05 for fvf_1 and 0.111 for fvf_2 with both the original model and the model trained with a reduced granularity. The Fully-Learned Method exhibited MAEs of 0.048 for fvf_1 and

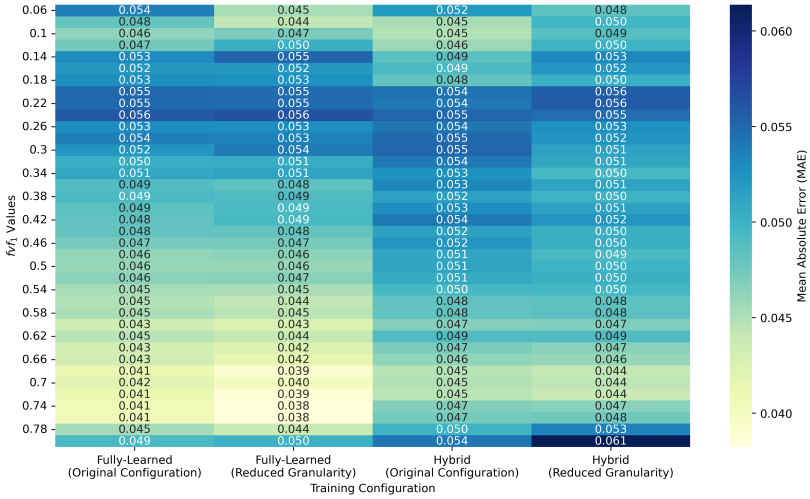


Fig. B.1 Consistent performance with minimal deviation were observed, highlighting both methods robustness to reduced granularity in the training set. Performance comparison of the Hybrid Method and Fully-Learned Method on the test set, trained with a reduced granularity of fvf_1 configurations and trained with the original configuration.

0.101 for fvf_2 with the original model and MAEs of 0.047 for fvf_1 and 0.10 for fvf_2 with the model trained with a reduced granularity.

These findings suggest that the models can effectively interpolate within the trained range and maintain reliable performance, reducing the necessity for closely spaced training data points.

B.2 Details on training parameters and inference performances

Table B.1 Our neural networks used meta parameters determined by systematic grid-search. Meta parameters and architectural characteristics of both accelerated methods. Note that the number of overall parameters in the network used in the Hybrid Method was lower than that of the Fully-Learned Method, due to its specific split-arm architecture.

	Hybrid Method	Fully-Learned Method
Minibatch size	5000	12 000
Dropout rate	0.1	0.1
Learning rate	3.5×10^{-4}	2×10^{-4}
Training samples	2×10^{6}	2×10^{6}
CASE WITH K=2 FASCICLES:		
Hidden units	2116	2108
Free parameters	5.30×10^{5}	2.11×10^{6}
CASE WITH K=3 FASCICLES:		
Hidden units	2999	2114
Free parameters	8.33×10^{5}	2.12×10^{6}

Table B.2 Large speed-up factors were observed for our proposed methods. Detailed computation times for the three methods when performing inference on synthetic data using the MGH Adult Diffusion protocol. The test set was processed on a Skylake Xeon 4116 CPU cluster, without GPU acceleration.

	Microstructure Fingerprinting	Hybrid Method	Fully-Learned Method
CASE WITH K=2 FASCICLES:			
Computation of SH coefficients/voxel	Not applicable	Not applicable	7.8×10^{-4} s
Dictionary rotation time/voxel	8.10×10^{-3} s	8.10×10^{-3} s	Not applicable
Exhaustive fingerprinting time/voxel	9.09×10^{-1} s	Not applicable	Not applicable
Solving NNLS time/voxel	Not applicable	1.47×10^{-2} s	Not applicable
NN forward pass time/voxel	Not applicable	4.24×10^{-5} s	8.12×10^{-5} s
Total inference time/voxel	9.18×10^{-1} s	2.36×10^{-2} s	8.61×10^{-4} s
CASE WITH K=3 FASCICLES:			
Computation of SH coefficients/voxel	Not applicable	Not applicable	7.8×10^{-4} s
Dictionary rotation time/voxel	1.03×10^{-2} s	1.03×10^{-2} s	Not applicable
Exhaustive fingerprinting time/voxel	248 s	Not applicable	Not applicable
Solving NNLS time/voxel	Not applicable	1.94×10^{-2} s	Not applicable
NN forward pass time/voxel	Not applicable	6.17×10^{-5} s	1.02×10^{-4} s
Total inference time/voxel	248 s	2.97×10^{-2} s	8.82×10^{-4} s

C

Multi-modal study of brain changes in long-term meditators

Summary of an article in preparation for submission as:

Dessain, Q, Dricot, L, Moniotte, S, Petit, G, Laureys, S, Darkner, S, de Timary, P and Kupers, R. Brain structural and functional connectivity differentiate long-term meditators from controls and predict emotional awareness and mood repair.

Meditation is a cognitive tool used for centuries to improve cognitive health and emotional coping. There is ample evidence that long-term meditation can have profound effects on brain structural and functional organization (Bowen, Witkiewitz, Clifasefi et al., 2014; Hilton, Maher, Colaiaco et al., 2017; Jamerson, Mehta, Michos et al., 2017; Kuyken, Hayes, Barrett et al., 2015). Although many brain imaging studies have been published on brain structural and functional changes linked to meditation, most of these studies are hampered by several methodological shortcomings including small sample sizes, the use of participants with insufficiently long meditation practice, the lack of neuropsychological profiling of the participants and the use of a single imaging modality, most often functional MRI and in other cases anatomical MRI or diffusion MRI. The aim of this study was therefore to test for brain structural and functional changes in a large homogeneous group of meditators

and to correlate these findings with a neuropsychological assessment of participants. We included a group of 50 meditators with an average daily meditation practice of 17.2±6.2 years. We also included a control group of 37 non-meditators who were matched to the meditators on various lifestyle characteristics. We used a multimodal brain imaging approach including anatomical T1-weigthted images (MPRAGE), resting state functional MRI (rsfMRI) and diffusion MRI (dMRI). The meditators were scanned at rest and during deep meditative state. A comprehensive neuropsychological assessment was also carried out, focusing on interoceptive awareness (MAIA) (Mehling, Price, Daubenmier et al., 2012), the cognitive modulation of emotions (CERQ) (Garnefski, Kraaij, and Spinhoven, 2001) and mood repair (Mayer, Salovey, and Caruso, 2004).

The behavioral data showed that meditators had a significantly higher interoceptive awareness compared to controls and also scored better on the cognitive modulation of emotions. This is in line with previous results. The anatomical data revealed significant increases in cortical surface in the meditators in superior frontal and middle frontal cortices, superior temporal cortex, supramarginal gyrus and parieto-occipital junction, postcentral and superior parietal cortex, and the occipital cortex. The fMRI data revealed widespread increases in functional connectivity during meditation, including occipital, insular, prefrontal, anterior temporal and cerebellar regions. Connectome-based predictive modeling (CPM) (Shen, Finn, Scheinost et al., 2017) revealed that mood repair behavioral scores could be predicted from rsfMRI and dMRI data. The highest predictive CPM node involved the right ventral agranular insula with the parieto-occipital sulcus, hippocampus, amygdala, precuneus and posterior cingulate cortex.

Bibliography

Acosta-Cabronero, J, Alley, S, Williams, GB, Pengas, G, and Nestor, PJ (2012). Diffusion Tensor Metrics as Biomarkers in Alzheimer's Disease. *PLoS ONE*, 7(11):e49072. doi: 10.1371/journal.pone.0049072.

Ahmed, OB, Benois-Pineau, J, Allard, M, Catheline, G, and Amar, CB (2017). Recognition of Alzheimer's disease and Mild Cognitive Impairment with multimodal image-derived biomarkers and Multiple Kernel Learning. *Neurocomputing*, 220:98–110. doi: 10.1016/j.neucom.2016.08.041.

Allsopp, K, Read, J, Corcoran, R, and Kinderman, P (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry research*, 279:15–22.

de Almeida Martins, JP, Nilsson, M, Lampinen, B, Palombo, M, While, PT, et al. (2021). Neural networks for parameter estimation in microstructural MRI: Application to a diffusion-relaxation model of white matter. *NeuroImage*, 244:118601.

Alp, S, Akan, T, Bhuiyan, MS, Disbrow, EA, Conrad, SA, et al. (2024). Joint transformer architecture in brain 3D MRI classification: its application in Alzheimer's disease classification. *Scientific Reports*, 14(1):8996. doi: 10.1038/s41598-024-59578-3.

Andersson, JL, Graham, MS, Drobnjak, I, Zhang, H, Filippini, N, et al. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion mr images: Within volume movement. *Neuroimage*, 152:450–466.

Andersson, JL and Sotiropoulos, SN (2015). Non-parametric representation and prediction of single-and multi-shell diffusion-weighted mri data using gaussian processes. *Neuroimage*, 122:166–176.

Andersson, JL and Sotiropoulos, SN (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging. *Neuroimage*, 125:1063–1078.

Assaf, Y, Blumenfeld-Katzir, T, Yovel, Y, and Basser, PJ (2008). Axcaliber: a method for measuring axon diameter distribution from diffusion mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 59(6):1347–1354.

Basser, P, Mattiello, J, and LeBihan, D (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267. doi: 10.1016/S0006-3495(94)80775-1.

Bowen, S, Witkiewitz, K, Clifasefi, SL, Grow, J, Chawla, N, et al. (2014). Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: a randomized clinical trial. *JAMA psychiatry*, 71(5):547–556.

Brabec, J, Lasič, S, and Nilsson, M (2020). Time-dependent diffusion in undulating thin fibers: Impact on axon diameter estimation. *NMR in Biomedicine*, 33(3):e4187.

Cai, C, Wang, C, Zeng, Y, Cai, S, Liang, D, et al. (2018). Single-shot T2 mapping using overlapping-echo detachment planar imaging and a deep convolutional neural network. *Magnetic resonance in medicine*, 80(5):2202–2214.

Canales-Rodríguez, EJ, Legarreta, JH, Pizzolato, M, Rensonnet, G, Girard, G, et al. (2019). Sparse wars: A survey and comparative study of spherical deconvolution algorithms for diffusion mri. *NeuroImage*, 184:140–160.

Cao, X, Liao, C, Zhou, Z, Zhong, Z, Li, Z, et al. (2024). Dti-mr fingerprinting for rapid high-resolution whole-brain t1, t2, proton density, adc, and fractional anisotropy mapping. *Magnetic resonance in medicine*, 91(3):987–1001.

Cercignani, M, Dowell, NG, and Tofts, PS (2018). *Quantitative MRI of the brain: principles of physical measurement*. CRC Press.

Chen, Z, Pawar, K, Ekanayake, M, Pain, C, Zhong, S, et al. (2023). Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. *Journal of Digital Imaging*, 36(1):204–230.

Christen, T, Pannetier, N, Ni, WW, Qiu, D, Moseley, ME, et al. (2014). Mr vascular fingerprinting: A new approach to compute cerebral blood vol-

ume, mean vessel radius, and oxygenation maps in the human brain. *Neuroimage*, 89:262–270.

Ciric, R, Lorenz, R, Thompson, W, Goncalves, M, MacNicol, E, et al. (2021). Templateflow: a community archive of imaging templates and atlases for improved consistency in neuroimaging. *bioRxiv*, pages 2021–02.

Cook, PA, Symms, M, Boulby, PA, and Alexander, DC (2007). Optimal acquisition orders of diffusion-weighted mri measurements. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 25(5):1051–1058.

Dausort, M, Delinte, N, Dessain, Q, Vanden Bulcke, C, and Macq, B (2023). A multi-compartment fingerprinting model for non-invasive tumor cell characterization via diffusion mri. In *Proc. Intl. Soc. Mag. Reson. Med.*, volume 31.

Delinte, N, Dessain, Q, Dausort, M, Vanden Bulcke, C, and Macq, B (2023). Tractography of the subcortical u-fibers using a position-dependent maximum angle. In *Proc. Intl. Soc. Mag. Reson. Med.*, volume 31.

Delinte, N, Gosse, C, Dricot, L, Dessain, Q, Simon, M, et al. (2021). Microstructural alterations in the white matter of children with dyslexia assessed by multi-fascicle diffusion compartment imaging. In *Proc. Intl. Soc. Mag. Reson. Med.*, volume 29.

Demirhan, A, Nir, TM, Zavaliangos-Petropulu, A, Jack, CR, Weiner, MW, et al. (2015). Feature selection improves the accuracy of classifying Alzheimer disease using diffusion tensor images. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 126–130. IEEE, Brooklyn, NY, USA. ISBN 978-1-4799-2374-8. doi: 10.1109/ISBI.2015.7163832.

Deng, J, Dong, W, Socher, R, Li, LJ, Li, K, et al. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee.

Derosiere, G, Vassiliadis, P, Dricot, L, Dessain, Q, Delinte, N, et al. (2024). Fronto-motor circuits linked to subclinical apathy. *bioRxiv*, pages 2024–10.

Descoteaux, M, Deriche, R, Knosche, TR, and Anwander, A (2008). Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE transactions on medical imaging*, 28(2):269–286.

Dessain, Q (2025). dMRI Gradient Cycling Optimizer: Software for optimizing GE dMRI gradient cycling sequences for thermal load management. https://github.com/Hyedryn/dMRI_Gradient_Cycling_Optimizer. GitHub repository.

Dessain, Q, Delinte, N, Macq, B, and Rensonnet, G (2022). Fast multi-compartment microstructure fingerprinting using deep neural networks. In *Proc. Intl. Soc. Mag. Reson. Med*, volume 30, page 0560.

Dessain, Q, Dricot, L, Delinte, N, Daussort, M, Macq, B, et al. (2023). Mood repair in long-term meditators is predicted by rsfmri functional connectivity of the insula and correlates with fa of the uncinate fasciculus. In *Proc. Intl. Soc. Mag. Reson. Med*, volume 31, page 0194.

Dessain, Q, Dricot, L, Delinte, N, Macq, B, de Timary, P, et al. (2024). Brain functional connectivity predicts emotional awareness. In *OHBM 2024 Annual Meeting*.

Dessain, Q, Fuchs, C, Macq, B, and Rensonnet, G (2024). Fast multi-compartment microstructure fingerprinting in brain white matter. *Frontiers in Neuroscience*, 18:1400499.

Dessain, Q, Mathieu, S, and Delinte, N (2024). Hyedryn/elikopy. doi: 10.5281/ZENODO.10514465.

Dessain, Q, Simon, M, Dricot, L, and Macq, B (2024). Advanced motion correction in diffusion mri acquisition with gradient cycling. In *International Symposium on Biomedical Imaging (ISBI 2024)*.

Dhital, B, Reisert, M, Kellner, E, and Kiselev, VG (2019). Intra-axonal diffusivity in brain white matter. *NeuroImage*, 189:543–550.

Diedrichsen, J (2006). A spatially unbiased atlas template of the human cerebellum. *Neuroimage*, 33(1):127–138.

Doneva, M, Börnert, P, Eggers, H, Stehning, C, Sénégas, J, et al. (2010). Compressed sensing reconstruction for magnetic resonance parameter mapping. *Magnetic Resonance in Medicine*, 64(4):1114–1120.

Dosovitskiy, A, Beyer, L, Kolesnikov, A, Weissenborn, D, Zhai, X, et al. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Drobnjak, I, Gavaghan, D, Süli, E, Pitt-Francis, J, and Jenkinson, M (2006). Development of a functional magnetic resonance imaging simulator for modeling realistic rigid-body motion artifacts. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 56(2):364–380.

Esteban, O, Markiewicz, CJ, Blair, RW, Moodie, CA, Isik, AI, et al. (2019). fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, 16(1):111–116.

Fan, L, Li, H, Zhuo, J, Zhang, Y, Wang, J, et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral cortex*, 26(8):3508–3526.

Fan, Q, Eichner, C, Afzali, M, Mueller, L, Tax, CM, et al. (2022). Mapping the human connectome using diffusion mri at 300 mt/m gradient strength: Methodological advances and scientific impact. *Neuroimage*, 254:118958.

Filipiak, P, Shepherd, T, Lin, YC, Placantonakis, DG, Boada, FE, et al. (2022). Performance of orientation distribution function-fingerprinting with a biophysical multicompartment diffusion model. *Magnetic resonance in medicine*, 88(1):418–435.

Frank, LR (2001). Anisotropy in high angular resolution diffusion-weighted mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 45(6):935–939.

Fuchs, C, Dessain, Q, Delinte, N, Dausort, M, and Macq, B (2024). Sparse blind spherical deconvolution of diffusion weighted mri. *Frontiers in Neuroscience*, 18:1385975.

Garnefski, N, Kraaij, V, and Spinhoven, P (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual differences*, 30(8):1311–1327.

Ginsburger, K, Poupon, F, Beaujoin, J, Estournet, D, Matuschke, F, et al. (2018). Improving the realism of white matter numerical phantoms: a

step toward a better understanding of the influence of structural disorders in diffusion MRI. *Frontiers in Physics*, 6:12.

Gong, G, He, Y, Concha, L, Lebel, C, Gross, DW, et al. (2009). Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral cortex*, 19(3):524–536.

Graham, MS, Drobnjak, I, and Zhang, H (2016). Realistic simulation of artefacts in diffusion mri for validating post-processing correction techniques. *NeuroImage*, 125:1079–1094.

Hahn, EL (1950). Spin echoes. *Physical review*, 80(4):580.

Hall, MG and Alexander, DC (2009). Convergence and parameter choice for Monte-Carlo simulations of diffusion MRI. *IEEE transactions on medical imaging*, 28(9):1354–1364.

Haykin, S (1994). *Neural networks: a comprehensive foundation*. Prentice hall PTR.

He, K, Gan, C, Li, Z, Rekik, I, Yin, Z, et al. (2023). Transformers in medical image analysis. *Intelligent Medicine*, 3(1):59–78.

Hill, ID, Palombo, M, Santin, MD, Branzoli, F, Philippe, AC, et al. (2018). Deep neural network based framework for in-vivo axonal permeability estimation. In *Joint Annual Meeting ISMRM-ESMRMB 2018*, volume 9681, page 968107. ISMRM.

Hilton, L, Maher, AR, Colaiaco, B, Apaydin, E, Sorbero, ME, et al. (2017). Meditation for posttraumatic stress: Systematic review and meta-analysis. *Psychological Trauma: Theory, Research, Practice, and Policy*, 9(4):453.

Hinton, GE, Srivastava, N, Krizhevsky, A, Sutskever, I, and Salakhutdinov, RR (2012). Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.

Hu, EJ, Shen, Y, Wallis, P, Allen-Zhu, Z, Li, Y, et al. (2021). Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

Hutter, J, Christiaens, D, Deprez, M, Cordero-Grande, L, Slator, P, et al. (2017). Dynamic field mapping and motion correction using interleaved double spin-echo diffusion mri. In *Medical Image Computing and Computer*

Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20, pages 523–531. Springer.

Hutter, J, Nilsson, M, Christiaens, D, Schneider, T, Price, AN, et al. (2018). Highly efficient diffusion mri by slice-interleaved free-waveform imaging (sifi). In *26th Annual Meeting of ISMRM, Montreal, Canada*.

Ioffe, S and Szegedy, C (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR.

Jack Jr, CR, Bennett, DA, Blennow, K, Carrillo, MC, Dunn, B, et al. (2018). Nia-aa research framework: toward a biological definition of alzheimer's disease. *Alzheimer's & dementia*, 14(4):535–562.

Jamerson, K, Mehta, PK, Michos, ED, Norris, K, Ray, IB, et al. (2017). Meditation and cardiovascular risk reduction. *J Am Heart Assoc*, 6:e002218.

Jang, J and Hwang, D (2022). M3T: three-dimensional Medical image classifier using Multi-plane and Multi-slice Transformer. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20686–20697. IEEE, New Orleans, LA, USA. ISBN 978-1-66546-946-3. doi: 10.1109/CVPR52688.2022.02006.

Jensen, JH, Helpern, JA, Ramani, A, Lu, H, and Kaczynski, K (2005). Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 53(6):1432–1440.

Jeurissen, B, Leemans, A, Tournier, JD, Jones, DK, and Sijbers, J (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human brain mapping*, 34(11):2747–2766.

Jeurissen, B, Tournier, JD, Dhollander, T, Connelly, A, and Sijbers, J (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. *NeuroImage*, 103:411–426.

Jones, DK (2004). The effect of gradient sampling schemes on measures derived from diffusion tensor mri: a monte carlo study. *Magnetic Resonance*

in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 51(4):807–815.

Jones, DK and Cercignani, M (2010). Twenty-five pitfalls in the analysis of diffusion mri data. *NMR in Biomedicine*, 23(7):803–820.

Jones, DK, Knösche, TR, and Turner, R (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion mri. *Neuroimage*, 73:239–254.

Karimi, D, Vasung, L, Jaimes, C, Machado-Rivas, F, Khan, S, et al. (2021). A machine learning-based method for estimating the number and orientations of major fascicles in diffusion-weighted magnetic resonance imaging. *Medical Image Analysis*, 72:102129.

Kim, HE, Cosa-Linan, A, Santhanam, N, Jannesari, M, Maros, ME, et al. (2022). Transfer learning for medical image classification: a literature review. *BMC medical imaging*, 22(1):69.

Kingma, DP and Ba, J (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kiselev, VG, Körzdörfer, G, and Gall, P (2021). Toward quantification: Microstructure and magnetic resonance fingerprinting. *Investigative Radiology*, 56(1):1–9.

Kuyken, W, Hayes, R, Barrett, B, Byng, R, Dalgleish, T, et al. (2015). Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (prevent): a randomised controlled trial. *The Lancet*, 386(9988):63–73.

Lawson, CL and Hanson, RJ (1995). *Solving least squares problems*. SIAM.

Le Bihan, D, Breton, E, Lallemand, D, Aubin, ML, Vignaud, J, et al. (1988). Separation of diffusion and perfusion in intravoxel incoherent motion mr imaging. *Radiology*, 168(2):497–505.

LeCun, Y, Bengio, Y, et al. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995.

LeCun, Y, Bottou, L, Bengio, Y, and Haffner, P (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Lella, E, Pazienza, A, Lofù, D, Anglani, R, and Vitulano, F (2021). An Ensemble Learning Approach Based on Diffusion Tensor Imaging Measures for Alzheimer's Disease Classification. *Electronics*, 10(3):249. doi: 10.3390/electronics10030249.

Lemasson, B, Pannetier, N, Coquery, N, Boisserand, LS, Collomb, N, et al. (2016). Mr vascular fingerprinting in stroke and brain tumors models. *Scientific reports*, 6(1):1–11.

Lian, C, Liu, M, Zhang, J, and Shen, D (2020). Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4):880–893. doi: 10.1109/TPAMI.2018.2889096.

Litjens, G, Kooi, T, Bejnordi, BE, Setio, AAA, Ciompi, F, et al. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88.

Liu, Z, Lin, Y, Cao, Y, Hu, H, Wei, Y, et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022.

Loshchilov, I and Hutter, F (2016). Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*.

Ma, D, Gulani, V, Seiberlich, N, Liu, K, Sunshine, JL, et al. (2013). Magnetic resonance fingerprinting. *Nature*, 495(7440):187–192.

Van der Maaten, L and Hinton, G (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

Marzban, EN, Eldeib, AM, Yassine, IA, Kadah, YM, and for the Alzheimer's Disease Neurodegenerative Initiative (2020). Alzheimer's disease diagnosis from diffusion tensor images using convolutional neural networks. *PLOS ONE*, 15(3):e0230409. doi: 10.1371/journal.pone.0230409.

Mayer, JD, Salovey, P, and Caruso, DR (2004). Target articles:" emotional intelligence: Theory, findings, and implications". *Psychological inquiry*, 15(3):197–215.

Mehling, WE, Price, C, Daubenmier, JJ, Acree, M, Bartmess, E, et al. (2012). The multidimensional assessment of interoceptive awareness (maia). *PloS one*, 7(11):e48230.

Morid, MA, Borjali, A, and Del Fiol, G (2021). A scoping review of transfer learning research on medical image analysis using imagenet. *Computers in biology and medicine*, 128:104115.

Mueller, SG, Weiner, MW, Thal, LJ, Petersen, RC, Jack, CR, et al. (2005). Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's & Dementia*, 1(1):55–66. doi: 10.1016/j.jalz.2005.06.003.

Nair, V and Hinton, GE (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814.

Najafabadi, MM, Villanustre, F, Khoshgoftaar, TM, Seliya, N, Wald, R, et al. (2015). Deep learning applications and challenges in big data analytics. *Journal of big data*, 2(1):1–21.

Nedjati-Gilani, GL, Schneider, T, Hall, MG, Cawley, N, Hill, I, et al. (2017). Machine learning based compartment models with permeability for white matter microstructure imaging. *NeuroImage*, 150:119–135. doi: https://doi.org/10.1016/j.neuroimage.2017.02.013.

Nilsson, M, Lätt, J, Ståhlberg, F, van Westen, D, and Hagslätt, H (2012). The importance of axonal undulation in diffusion mr measurements: a monte carlo simulation study. *NMR in Biomedicine*, 25(5):795–805.

Pajevic, S and Pierpaoli, C (1999). Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(3):526–540.

Palombo, M, Ligneul, C, Najac, C, and Valette, J (2016). New paradigm to assess brain cell morphology by diffusion-weighted mr spectroscopy in vivo. *PNAS USA*, 495(7440):187–192.

Panagiotaki, E, Chan, RW, Dikaios, N, Ahmed, HU, O'Callaghan, J, et al. (2015). Microstructural characterization of normal and malignant human prostate tissue with vascular, extracellular, and restricted diffusion for cytometry in tumours magnetic resonance imaging. *Investigative radiology*, 50(4):218–227.

Parmar, H, Nutter, B, Long, R, Antani, S, and Mitra, S (2020). Spatiotemporal feature extraction and classification of alzheimer's disease using deep learning 3d-cnn for fmri data. *Journal of Medical Imaging*, 7(5):056001–056001.

Pearlson, GD, Clementz, BA, Sweeney, JA, Keshavan, MS, and Tamminga, CA (2016). Does biology transcend the symptom-based boundaries of psychosis? *The Psychiatric clinics of North America*, 39(2):165.

Rafael-Patino, J, Romascano, D, Ramirez-Manzanares, A, Canales-Rodríguez, EJ, Girard, G, et al. (2020). Robust monte-carlo simulations in diffusion-mri: Effect of the substrate complexity and parameter choice on the reproducibility of results. *Frontiers in neuroinformatics*, 14:8.

Rathore, S, Habes, M, Iftikhar, MA, Shacklett, A, and Davatzikos, C (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage*, 155:530–548. doi: 10.1016/j.neuroimage.2017.03.057.

Rensonnet, G, Adam, L, and Macq, B (2021). Solving inverse problems with deep neural networks driven by sparse signal decomposition in a physics-based dictionary. *arXiv preprint arXiv:2107.10657*.

Rensonnet, G, Jacobs, D, Macq, B, and Taquet, M (2015). A hybrid method for efficient and accurate simulations of diffusion compartment imaging signals. In *11th International Symposium on Medical Information Processing and Analysis*, volume 9681, page 968107. International Society for Optics and Photonics.

Rensonnet, G, Scherrer, B, Girard, G, Jankovski, A, Warfield, SK, et al. (2019). Towards microstructure fingerprinting: estimation of tissue properties from a dictionary of Monte Carlo diffusion MRI simulations. *NeuroImage*, 184:964–980.

Rensonnet, G, Scherrer, B, Warfield, SK, Macq, B, and Taquet, M (2018). Assessing the validity of the approximation of diffusion-weighted-MRI signals from crossing fascicles by sums of signals from single fascicles. *Magnetic resonance in medicine*, 79(4):2332–2345.

Rumelhart, DE, Hinton, GE, and Williams, RJ (1986). Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, ed. de rumelhart and j. mcclelland. vol. 1. 1986. *Biometrika*, 71(599-607):6.

Sabidussi, ER, Klein, S, Caan, MW, Bazrafkan, S, den Dekker, AJ, et al. (2021). Recurrent inference machines as inverse problem solvers for MR relaxometry. *Medical Image Analysis*, 74:102220.

Schilling, KG, Blaber, J, Huo, Y, Newton, A, Hansen, C, et al. (2019). Synthesized b0 for diffusion distortion correction (synb0-disco). *Magnetic resonance imaging*, 64:62–70.

Selvaraju, RR, Cogswell, M, Das, A, Vedantam, R, Parikh, D, et al. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626.

Setsompop, K, Kimmlingen, R, Eberlein, E, Witzel, T, Cohen-Adad, J, et al. (2013). Pushing the limits of in vivo diffusion mri for the human connectome project. *Neuroimage*, 80:220–233.

Shamshad, F, Khan, S, Zamir, SW, Khan, MH, Hayat, M, et al. (2023). Transformers in medical imaging: A survey. *Medical Image Analysis*, 88:102802.

Shen, X, Finn, ES, Scheinost, D, Rosenberg, MD, Chun, MM, et al. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *nature protocols*, 12(3):506–518.

Skare, S, Hedehus, M, Moseley, ME, and Li, TQ (2000). Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with mri. *Journal of magnetic resonance*, 147(2):340–352.

Sotiropoulos, SN, Behrens, TE, and Jbabdi, S (2012). Ball and rackets: inferring fiber fanning from diffusion-weighted mri. *Neuroimage*, 60(2):1412–1425.

Sperling, RA, Aisen, PS, Beckett, LA, Bennett, DA, Craft, S, et al. (2011). Toward defining the preclinical stages of alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia*, 7(3):280–292.

Stejskal, EO and Tanner, JE (1965). Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292.

Su, P, Mao, D, Liu, P, Li, Y, Pinho, MC, et al. (2017). Multiparametric estimation of brain hemodynamics with mr fingerprinting asl. *Magnetic resonance in medicine*, 78(5):1812–1823.

Sudlow, C, Gallacher, J, Allen, N, Beral, V, Burton, P, et al. (2015). Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779.

Taquet, M, Jankovski, A, Rensonnet, G, Jacobs, D, Des Rieux, A, et al. (2019). Extra-axonal restricted diffusion as an in-vivo marker of reactive microglia. *Scientific reports*, 9(1):13874.

Tariq, M, Schneider, T, Alexander, DC, Wheeler-Kingshott, CAG, and Zhang, H (2016). Bingham–noddi: Mapping anisotropic orientation dispersion of neurites using diffusion mri. *Neuroimage*, 133:207–223.

Tax, CM, Bastiani, M, Veraart, J, Garyfallidis, E, and Irfanoglu, MO (2022). What's new and what's next in diffusion mri preprocessing. *NeuroImage*, 249:118830.

Thompson, PM, Jahanshad, N, Ching, CR, Salminen, LE, Thomopoulos, SI, et al. (2020). Enigma and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Translational psychiatry*, 10(1):100.

Tournier, JD, Calamante, F, and Connelly, A (2007). Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage*, 35(4):1459–1472.

Tournier, JD, Calamante, F, and Connelly, A (2012). Mrtrix: diffusion tractography in crossing fiber regions. *International journal of imaging systems and technology*, 22(1):53–66.

Tournier, JD, Smith, R, Raffelt, D, Tabbara, R, Dhollander, T, et al. (2019). Mrtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage*, 202:116137.

Touvron, H, Cord, M, Douze, M, Massa, F, Sablayrolles, A, et al. (2021). Training data-efficient image transformers & distillation through attention. In *International conference on machine learning*, pages 10347–10357. PMLR.

Tuch, DS, Reese, TG, Wiegell, MR, Makris, N, Belliveau, JW, et al. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 48(4):577–582.

Van Essen, DC, Ugurbil, K, Auerbach, E, Barch, D, Behrens, TE, et al. (2012). The Human Connectome Project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231.

Vaswani, A (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.

Veraart, J, Novikov, DS, Christiaens, D, Ades-Aron, B, Sijbers, J, et al. (2016). Denoising of diffusion mri using random matrix theory. *Neuroimage*, 142:394–406.

Villemagne, VL, Burnham, S, Bourgeat, P, Brown, B, Ellis, KA, et al. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic alzheimer's disease: a prospective cohort study. *The Lancet Neurology*, 12(4):357–367.

Wang, HT, Clarke, N, Dessain, Q, Paugam, F, and Bellec, P (2024). Nilearn and big data facilitates transdiagnostic brain biomarkers. *Biological Psychiatry*, 95(10):S22.

Wen, J, Thibeau-Sutre, E, Diaz-Melo, M, Samper-González, J, Routier, A, et al. (2020). Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation. *Medical image analysis*, 63:101694.

Wright, KL, Jiang, Y, Ma, D, Noll, DC, Griswold, MA, et al. (2018). Estimation of perfusion properties with mr fingerprinting arterial spin labeling. *Magnetic resonance imaging*, 50:68–77.

Yan, W, Qu, G, Hu, W, Abrol, A, Cai, B, et al. (2022). Deep learning in neuroimaging: Promises and challenges. *IEEE Signal Processing Magazine*, 39(2):87–98.

Ye, C (2017). Estimation of tissue microstructure using a deep network inspired by a sparse reconstruction framework. In *International Conference on Information Processing in Medical Imaging*, pages 466–477. Springer.

Ye, C, Li, X, and Chen, J (2019). A deep network for tissue microstructure estimation using modified lstm units. *Medical image analysis*, 55:49–64.

Zavaliangos-Petropulu, A, Nir, TM, Thomopoulos, SI, Reid, RI, Bernstein, MA, et al. (2019). Diffusion MRI Indices and Their Relation to Cognitive Impairment in Brain Aging: The Updated Multi-protocol Approach in ADNI3. *Frontiers in Neuroinformatics*, 13:2. doi: 10.3389/fninf.2019.00002.

Zhang, H, Schneider, T, Wheeler-Kingshott, CA, and Alexander, DC (2012). NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61(4):1000–1016. doi: 10.1016/j.neuroimage.2012.03.072.