Universidad de Valladolid

Escuela de Doctorado Universidad de Valladolid

PROGRAMA DE DOCTORADO EN TECNOLOGÍAS DE LA INFORMACIÓN Y LAS TELECOMUNICACIONES

TESIS DOCTORAL:

DESIGN AND RELIABILITY VALIDATION OF DIFFUSION MRI BIOMARKERS FOR THE BRAIN'S WHITE MATTER

Presentada por Guillem París para optar al grado de
Doctor por la Universidad de Valladolid

Dirigida por:

Santiago Aja Fernández
y
Antonio Tristán Vega

Design and reliability validation of diffusion MRI biomarkers for the brain's white matter

Ph.D. Thesis
June 2025

Guillem París

Acknowledgements

First and foremost, I want to give a heartfelt thank you to Santiago Aja and Antonio Tristán, my supervisors at LPI, Universidad de Valladolid. It's hard to put into words how much I appreciate your constant support, generosity, and kindness throughout this journey. You've been more than just mentors—you've been genuinely caring, always willing to help, guide, and encourage me even when things got tough (or weirdly technical). I've learned so much from you both, not only about research but also about how to be a thoughtful and respectful scientist. And, perhaps most admirably, you've patiently endured my endless questions, half-baked ideas, and the occasional scientific nonsense without ever making me feel like I didn't belong. I truly wish the very best for both of you—you deserve it all.

To the whole LPI group: you've made this experience way brighter than I ever expected. Your patience, humor, and no-nonsense advice made a real difference. Special shoutout to Tomasz Pieciak—for your sharp insights, for sharing the emotional rollercoaster of reviewer responses, and for somehow surviving the chaos of working with me. Thanks for every discussion, every shared complaint, and every laugh in between. I always felt lucky to be part of this team. It's been a pleasure working alongside such a smart, supportive, and fun bunch of people.

I also want to sincerely thank Jelle Veraart, my supervisor at CBI, NYU Grossman School of Medicine. Professor, your guidance, patience, and openness made a huge difference. You've been incredibly generous with your time and insights, and I always felt encouraged and supported, even from across the Atlantic. Your kindness and clarity have made this process smoother, and I've genuinely enjoyed every chat, and every bit of feedback. It's been an honor to work with you.

To my friends and family—thank you for being there. For cheering me on, for listening to my rants, for pretending that you understood what I was talking about, and for being a safe space to just be myself. Whether you offered a hug, a meme, or a distraction when I needed it most, you've helped keep me sane. I owe you more than a paragraph. A very special shoutout to Juanpa—the best friend one could ever hope for: endlessly kind, fiercely fair, occasionally a full-blown drama queen, and somehow always exactly what I needed. Thank you for being there through all of it. Emma and Natalia—you're persistently in my heart, no matter the distance or the silence. And to all my current and past crew at Prince—you've shaped me in ways you probably didn't realize, and put up with me even at times when I probably wouldn't have. I'm deeply grateful for the laughter, the talks, and the quiet support. To all of you: thank you.

Finally, I want to remember two people whose presence left a mark far beyond this thesis. Tere Caja—an example of relentless courage and an unshakable voice for social justice. Your strength, integrity, and sense of purpose continue to inspire me. And Pau Verdaguer—a bright soul who radiated happiness and resilience, even in the face of an unfair world. Thank you both for showing me what it means to live with conviction, light, and heart. This work carries a little piece of you with it.

To conclude, a thought for the living: The everyday is not dull, it is simply overlooked. We move through routines, forgetting that each person carries invisible battles, quiet hopes, silent weight. The real work, perhaps, is not brilliance, but attention. To see the world with open eyes, and respond with decency and care. This is water.

Abstract

Diffusion Magnetic Resonance Imaging (dMRI) offers a unique, non-invasive approach to study the microstructural properties of the human brain. Yet, a fundamental challenge persists: while water diffusion occurs at the microscopic scale (μm), dMRI measurements are acquired at a much coarser spatial resolution (mm), introducing a considerable gap between what is measured and what is inferred. Modeling approaches attempt to bridge this gap by linking the observed signal to meaningful, often quantitative, descriptions of the underlying tissue structure. However, substantial disparities remain between what is technically achievable in state-of-the-art research environments and what is attainable in standard clinical or non-specialized research settings. This work aims at narrowing that gap by improving the practicality, reliability, and robustness of microstructural biomarkers derived from dMRI data.

Three main objectives guide the development of this thesis. First, we aim to define and implement novel and efficient biomarkers for white matter that provide relevant insights by describing the local diffusion properties. These biomarkers are designed to be meaningful—that is, to capture features known to have clinical or neuroscientific relevance—but also accessible, avoiding dependence on ultra-high-field MRI scanners, extensive acquisition protocols, or computationally intractable models. Second, we assess and mitigate the effect of thermal noise on the stability of dMRI-derived metrics. Despite being an unavoidable feature of MR data acquisition, thermal noise is rarely modeled explicitly and yet can significantly degrade the interpretability and reproducibility of estimates in modeling frameworks. Finally, we propose a framework to extend the state-of-the-art reliability assessment methodology in dMRI by explicitly distinguishing between two key components: repeatability (i.e., the consistency of a metric within the same condition) and separability (i.e., the ability of a metric to distinguish between different conditions or subjects). This distinction allows for a more nuanced evaluation of biomarkers, particularly in studies with limited sample size.

To address these objectives, we designed and analyzed a set of synthetic and real datasets with multiple repeated scans per subject, enabling the quantification of metric stability and discriminative power. Through this, we demonstrate that reliability is not a one-dimensional concept and cannot be fully captured by test-retest measures alone. We also show that thermal noise has a systematic and non-negligible impact on parameter estimation, especially for models with degenerate or ill-posed solution spaces—a common trait in biophysically inspired

approaches. Mitigation strategies, including robust estimators and informed regularization, are proposed and validated.

In summary, this thesis contributes a set of methodological tools and conceptual insights for advancing the development of microstructural biomarkers that are not only theoretically sound but also practical and reliable in real-world research and clinical settings. It promotes a shift from purely repeatability-driven evaluations to a broader understanding of reliability, while highlighting the crucial, often overlooked, role of thermal noise in dMRI modeling. The tools and results presented herein aim to facilitate the transition of advanced dMRI techniques from specialized labs to broader applications, bridging the gap between innovation and implementation.

Resumen

La imagen por resonancia magnética de difusión (dMRI, por sus siglas en inglés) ofrece un enfoque único y no invasivo para estudiar las propiedades microestructurales del cerebro humano. Sin embargo, persiste un desafío fundamental: aunque la difusión del agua ocurre a escala microscópica (μm), las mediciones de dMRI se adquieren a una resolución espacial mucho más gruesa (mm), lo que introduce una brecha considerable entre lo que se mide y lo que se infiere. Los enfoques de modelado intentan cerrar esta brecha al vincular la señal observada con descripciones significativas, a menudo cuantitativas, de la estructura tisular subyacente. No obstante, siguen existiendo disparidades sustanciales entre lo que es técnicamente posible en entornos de investigación de vanguardia y lo que se puede alcanzar en contextos clínicos estándar o en investigaciones no especializadas. Este trabajo busca reducir esa brecha mejorando la practicidad, fiabilidad y robustez de los biomarcadores microestructurales derivados de datos de dMRI.

El desarrollo de esta tesis se guía por tres objetivos principales. En primer lugar, se busca definir e implementar biomarcadores novedosos y eficientes para la sustancia blanca que proporcionen información relevante describiendo las propiedades locales de difusión. Estos biomarcadores están diseñados para ser significativos—es decir, captar características conocidas por su relevancia clínica o neurocientífica—pero también accesibles, evitando la dependencia de escáneres de RM de ultra-alto campo, protocolos de adquisición extensos o modelos computacionalmente intratables. En segundo lugar, se analiza y mitiga el efecto del ruido térmico sobre la estabilidad de las métricas derivadas de dMRI. A pesar de ser una característica inevitable en la adquisición de datos por RM, el ruido térmico rara vez se modela de forma explícita, y sin embargo puede degradar significativamente la interpretabilidad y reproducibilidad de las estimaciones en enfoques de modelado. Por último, se propone un marco para extender la metodología de evaluación de fiabilidad más avanzada en dMRI, distinguiendo explícitamente entre dos componentes clave: la repetibilidad (la consistencia de una métrica bajo una misma condición) y la separabilidad (la capacidad de una métrica para distinguir entre distintas condiciones o sujetos). Esta distinción permite una evaluación más matizada de los biomarcadores, especialmente en estudios con tamaños muestrales reducidos.

Para abordar estos objetivos, se diseñaron y analizaron conjuntos de datos sintéticos y reales con múltiples adquisiciones repetidas por sujeto, lo que permitió cuantificar la estabilidad de las métricas y su poder discriminativo. A través de estos análisis, se demuestra que la fiabilidad no

es un concepto unidimensional y que no puede capturarse plenamente solo mediante medidas de test-retest. También se muestra que el ruido térmico tiene un impacto sistemático y no despreciable en la estimación de parámetros, especialmente en modelos con espacios de solución degenerados o mal condicionados—una característica común en los enfoques de inspiración biofísica. Se proponen y validan estrategias de mitigación, que incluyen estimadores robustos y regularización informada.

En resumen, esta tesis aporta un conjunto de herramientas metodológicas y reflexiones conceptuales orientadas a avanzar en el desarrollo de biomarcadores microestructurales que no solo sean sólidos desde el punto de vista teórico, sino también prácticos y fiables en entornos de investigación y clínicos reales. Se promueve un cambio desde evaluaciones centradas exclusivamente en la repetibilidad hacia una comprensión más amplia de la fiabilidad, al tiempo que se subraya el papel crucial—y a menudo ignorado—del ruido térmico en el modelado de dMRI. Las herramientas y resultados presentados aquí tienen como objetivo facilitar la transición de técnicas avanzadas de dMRI desde laboratorios especializados hacia aplicaciones más amplias, cerrando la brecha entre la innovación y su implementación.

Contents

List of Abbreviations **xi**

1 Introduction **1**

1.1 Motivation . . . 1
1.2 Objectives . . . 5
1.3 Methodology . . . 6
1.4 Materials . . . 8
1.4.1 MRI Datasets . . . 9
1.4.2 Atlases . . . 11
1.4.3 Computational Resources . . . 11
1.5 List of Publications . . . 12
1.6 Thesis Overview . . . 15

2 Background **17**

2.1 Introduction to diffusion MRI . . . 17
2.1.1 Diffusion MRI principles . . . 18
2.1.2 Technical viewpoint . . . 19
2.2 Diffusion Tensor Imaging . . . 21
2.2.1 Tractography . . . 22
2.3 Beyond the tensor . . . 24
2.3.1 Limitations of DTI . . . 24
2.3.2 Q-space sampling . . . 24
2.3.3 Mathematical Tools . . . 27
2.4 Phenomenological vs. Biophysical Modeling . . . 29
2.4.1 Incompatibility . . . 31
2.4.2 Advanced Diffusion Metrics . . . 31
2.4.3 Phenomenological Models . . . 38
2.4.4 Biophysical Models . . . 52
2.5 Clinical Validation . . . 55
2.6 Preclinical Validation through Repeatability and Separability . . . 56
2.7 Contributions to the state of the art . . . 58

3 Contribution 1: Efficient estimation of PA and NG for MiSFIT **61**
3.1 Introduction and Purpose . . . 62
3.2 Theory . . . 63
3.2.1 MiSFIT's signal representation . . . 63
3.2.2 PA for MiSFIT's composite signal . . . 64
3.2.3 NG for MiSFIT's composite signal . . . 65
3.3 Materials . . . 66
3.3.1 In vivo evaluation . . . 66
3.3.2 Numerical simulations . . . 66
3.4 Methods . . . 67
3.4.1 Implementation details . . . 67
3.4.2 Ground-truth based evaluation . . . 67
3.4.3 Reliability Study . . . 68
3.4.4 Computational Time Evaluation . . . 68
3.5 Results . . . 69
3.5.1 Visual assessment . . . 69
3.5.2 Quantitative analysis based on ground-truth . . . 70
3.5.3 Reliability Study . . . 73
3.5.4 Computational Time Evaluation . . . 74
3.6 Discussion and Conclusions . . . 75

4 Contribution 2: Rician bias effects on biophysical models **79**
4.1 Introduction and Purpose . . . 80
4.2 Theory . . . 81
4.2.1 Rotationally-invariant spherical harmonic features . . . 81
4.2.2 Biophysical modeling . . . 82
4.2.3 Data distributions . . . 84
4.3 Materials . . . 86
4.3.1 In silico data . . . 86
4.3.2 In vivo dataset . . . 86
4.4 Method . . . 87
4.4.1 Parameter Estimation . . . 87
4.4.2 Statistical Analyses . . . 88
4.5 Results . . . 88
4.5.1 Simulations: Thermal noise effect on RISH features . . . 88
4.5.2 Simulations: Thermal noise effect on biophysical modeling parameters across estimators . . . 90
4.5.3 In vivo results . . . 91
4.5.4 Protocol dependency . . . 92
4.5.5 Reproducibility analysis . . . 94
4.6 Discussion and Conclusions . . . 94

5 Contribution 3: A reliability comparison of advanced descriptors **99**
5.1 Introduction and Purpose . . . 100
5.2 Theory . . . 102
5.2.1 Variability sources . . . 102
5.2.2 Effect size, reliability and sample size . . . 103
5.3 Materials . . . 104
5.3.1 In vivo datasets . . . 104
5.4 Methods . . . 104
5.4.1 Segmentation . . . 104
5.4.2 Computation of metrics . . . 105
5.5 Results . . . 105
5.5.1 Reliability of biomarkers . . . 106
5.5.2 Displacement between datasets . . . 107
5.5.3 Effect and sample sizes . . . 109
5.6 Discussion and Conclusions . . . 111
5.6.1 On biomarkers' reliability . . . 112
5.6.2 On dataset quality . . . 112
5.6.3 On effect and sample sizes . . . 113
5.6.4 Limitations . . . 114
5.6.5 Conclusion . . . 114

6 Discussion and Conclusions **115**
6.1 Discussion . . . 115
6.1.1 On efficient biomarkers development . . . 116
6.1.2 On thermal noise effects in biophysical modeling . . . 116
6.1.3 On biomarkers reliability . . . 117
6.1.4 Limitations . . . 118
6.2 Conclusions . . . 119
6.2.1 Key Findings . . . 120
6.2.2 Future Directions . . . 120
6.2.3 Final Thoughts . . . 121

Bibliography **123**

A Legendre Polynomials Integration **137**
A.0.1 Propagator Anisotropy for EAP composite signal . . . 138

B Non-Gaussianity for EAP composite signal **141**
B.0.1 Efficient Sampling of Spherical Function . . . 142

C Other results **145**

D Reliability computation from a standard two-way ANOVA **149**

E Reliability – repeatability Tables: MICRA **153**

F Reliability – repeatability Tables: ZJU **161**

List of Abbreviations

AD . . . . . Alzheimer's Disease

ADC . . . . . Apparent Diffusion Coefficient

AMURA . . . . . Apparent Measures Using Reduced Acquisitions

CC . . . . . Corpus Callosum

CLS . . . . . Conditional Least Squares

CUBRIC . . . . . Cardiff University Brain Research Imaging Centre

CV . . . . . Coefficient of Variation

DIPY . . . . . Diffusion Imaging in Python

DMRI . . . . . Diffusion Magnetic Resonance Imaging

DTI . . . . . Diffusion Tensor Imaging

DOT . . . . . Diffusion Orientation Transform

DSI . . . . . Diffusion Spectrum Imaging

EAP . . . . . Ensemble Average Propagator

FA . . . . . Fractional Anisotropy

FFT . . . . . Fast Fourier Transform

FMRIB . . . . . Oxford Centre for Functional Magnetic Resonance Imaging of the Brain

fODF . . . . . Fiber Orientation Distribution Function

FOM . . . . . Figure of Merit

FSL . . . . . FMRIB Software Library

GM . . . . . Grey Matter

HARDI . . . . . High Angular Resolution Diffusion Imaging

HCP . . . . . Human Connectome Project

HCP-MGH . . . Human Connectome Project — Massachusetts General Hospital

HYDI . . . . . Hybrid Diffusion Imaging

ICBM . . . . . International Consortium for Brain Mapping

ICC . . . . . Intra-class Correlation Coefficient

JHU . . . . . John Hopkins University

LS . . . . . Least Squares

MAPL . . . . . Laplacian-regularized Mean Apparent Propagator

MAP-MRI . . . . Mean Apparent Propagator MRI

MD . . . . . Mean Diffusivity

MiSFIT . . . . . Microstructure adaptive convolution kernels and dual Fourier Integral Transforms

MK . . . . . Mean Kurtosis

MLE . . . . . Maximum Likelihood Estimation

MNI . . . . . Montreal Neuroscience Institute

MPPCA . . . . . Marchenko-Pastur Principal Component Analysis

MR . . . . . Magnetic Resonance

MRI . . . . . Magnetic Resonance Imaging

NG . . . . . Non-Gaussianity

NLLS . . . . . Nonlinear Least Squares

ODF . . . . . Orientation Distribution Function

PA . . . . . Propagator Anisotropy

PD . . . . . Parkinson's Disease

RISH . . . . . Rotational Invariants from Spherical Harmonics

ROI . . . . . Region of Interest

RTOP . . . . . Return to Origin Probability

RTAP . . . . . Return to Axis Probability

RTPP . . . . . Return to Plane Probability

SH . . . . . Spherical Harmonics

SHORE . . . . . Simple Harmonic Oscillator Reconstruction and Estimation

SMI . . . . . Standard Model of diffusion Imaging

SNR . . . . . Signal-to-noise ratio

SOA . . . . . State-of-the-art

TBI . . . . . Traumatic Brain Injury

WM . . . . . White Matter

WM-TDI . . . . . White Matter Tract Density Imaging

ZJU . . . . . Zhejiang University

In the neoliberal regime, the soul has become a production unit.

— Han (2017)

1 Introduction

1.1 Motivation

The frontiers of human research expand relentlessly in all possible directions. While some try to uncover the misteries on distant galaxies and unveil the origins of our universe, others focus on the interactions of the smallest of particles and how, in them, rests the unknown. It is in this landscape where one recurrently comes across a complex system, the puzzle that has been accompanying us for the entirety of our history. Sculpted from cores of dying stars, emerges the most perfect tool, an entire universe within each of us: the human brain.

Consisting of approximately 86 billion neurons and about 150.000 km of myelinated axons (half of the distance to the moon), the human brain not only regulates essential bodily functions, but also is the seat for consciousness, behavior and decision-making. Its research entail a series of challenges, ranging from limited accessibility, to unpredictability of effects; yet, it is considered necessary for the study of brain-related conditions, human cognition and behavior or even the development of novel intelligent agents. In clinical endeavors, invasive brain procedures carry inherent risks: 14.3% of neurosurgical procedures result in complications (Fugate, 2015), and more than 19% of the inverventions exhibit adverse effects (Dao Trong et al., 2023)—such as memory loss or fatigue.

The emergence of noninvasive imaging techniques has lead to a paradigm shift in neuroscience. While not all noninvasive techniques excel on every application, Magnetic Resonance Imaging (MRI) stand out as a prominent tool for portraying brain anatomy, brain activity and brain connectivity (Bernstein-Eliav and Tavor, 2022), all topics around which this thesis revolves.

The brain consists of two distinct regions in terms of cytoarchitecture and function: Gray matter (GM)—consisting of neuron cell bodies, dendrites and synapses—and White matter (WM)—

primarily composed of myelinated axons. Given the difference in these two regions, relaxation properties by themselves are not enough to distinguish biological tissues, and are thus perceived as homogeneous in MRI images. Diffusion MRI (dMRI) was proposed as a solution to this: sensitize the MRI to the diffusion movement of water molecules within such tissues.

Since the appearance of dMRI, great advances have been made in the study of brain connectivity and microstructure, leading to the successful detection of neural processes and biophysical changes, as well as providing insights into brain cytoarchitecture, which has resulted in the profiling of neurological diseases and pathological processes, such as:

Multiple sclerosis or MS is an autoimmune disease that mistakenly attacks the myelin sheaths of axons in the central nervous system (CNS), a process called demyelination. It affects roughly 2.8 million people worldwide (Walton et al., 2020). The damages manifest as lesions in the WM, intimately related to cognitive decline.

Stroke entails a halt of blood supply to a brain part. It might be caused by a clot (ischemic stroke) or the burst of a blood vessel (hemorragyc stroke). It affects around 15 million people annually (World Health Organization, 2025b). The blockage might result in lack of oxygen and tissue death.

Traumatic Brain Injury or TBI is the result of an external force (trauma) onto the brain. It affects 69 million people worldwide (Dewan et al., 2019) and might entail disturbances in cognitive, motor or emotional healthy activities.

Neurodegenerative Diseases (Alzheimer's, Parkinson's, etc.) result from the progressive degeneration of neurons and axons. Alzheimer's disease (AD), affecting 55 million people worldwide (World Health Organization, 2025a), is linked to abnormal accumulation of amyloid plaques. Parkinson's disease (PD), on the other hand, affects approximately 10 million people worldwide (Parkinson's Foundation, 2025) and is associated with the loss of dopamine-producing neurons. Both AD and PD often lead to WM atrophy, that ultimately results in cognitive decline.

These pathologies are all characterized by distinctive diffusion profiles in the affected regions, which can now be effectively captured. In doing so, dMRI paves the way to a full understanding of their evolution, their early diagnosis (sometimes even pre-symptomatic) and monitoring, which ultimately holds potential for better health outcomes and disease prevention.

Similarly, dMRI has enabled crucial contributions to the neuroscience research, setting the stage for a clearer understanding of brain functions, connectivity, and development:

Brain Connectivity is of upmost importance given that the brain network enables the coordination of brain functions and the emergence of intelligence. Diffusion MRI plays a crucial role in understanding brain networks, and can ultimately be used to create maps such as tractograms (Jeurissen et al., 2017).

Brain Development refers to the growth and maturation of brain structures. The evolution of the brain during childhood is impacted by several changes including myelination, synpatic pruning, or strenthening of neural connections; which are essential in the development of appropriate cognitive functions. Diffusion MRI is able to capture the unique diffusion profile that each of these alterations produce, enabling a complete picture of the hidden brain biophysics (Feng et al., 2023).

Neuroplasticity or the brain's ability to reorganize itself by forming new connections (or dropping old ones) as a consequence of experiences, learning, injuries, or psychological traumas. Extending from Brain Connectivity, the dMRI capability in mapping the connections in the brain is important in the study of this brain reorganization (Thomas et al., 2014).

Cognitive Functions that allow individuals to acquire knowledge and understanding, from memory and attention, to problem-solving and language. Given that all these functions stem from the neurophysiological environments, their study using diffusion MRI provides valuable insights into the structural integrity of the brain's white matter and microstructure. Diffusion MRI enables the detailed analysis of the brain's white matter tracts, which are fundamental in supporting cognitive processes. In doing so, it can detect subtle changes in tissue structure that may correlate with cognitive abilities performance. This allows researchers to examine how variations in white matter integrity affect cognitive function, providing a window into the biological substrates of learning, memory, and executive functions, and helping to identify biomarkers for cognitive aging or early cognitive impairments (Nestor et al., 2014).

The proficiency of dMRI in characterizing diffusion behavior, with potential applications in various dimensions, has lead to a surge in its research. As with previous scientific paradigm shifts, the dMRI field still fails on fully finding its place, and has experienced an expansion in all possible directions: dozens of modeling frameworks, hundreds of processing pipelines, and a myriad of biomarkers and other metrics. Far from complaining on these natural processes—the author of this thesis understand this is the natural path of science—they have to ultimately be counteracted with convergence efforts. Recent calls for field convergence include—but are not limited to—standardization of methodologies, adoption of common processing pipelines (Joseph et al., 2021; Tax et al., 2022; Esteban, 2024) and protocol acquisitions (Campaña Perilla et al., 2023), and benchmarking research (Brainlife.io, 2025).

A byproduct of this field divergence is the increasing technological gap between high-end environments and regular ones. While well-funded research settings aim at increasing the accuracy, stability and reliability of their markers by employing powerful and expensive infrastructure, regular settings are condemned to see from afar the development of the field by agents with whom they cannot compete. In other words, standard environments are often unable to develop biomarkers that compete in accuracy or stability with those developed by high-end settings that employ extensive sampling schemes, powerful computational infrastructures and advanced scanners with higher diffusion gradients and lower echo times. This is also reflected in the scarce application of dMRI in clinical environments. After 30 years since its introduction, the healthcare sector still relies on the diffusion tensor (Jones and Leemans, 2010; Chung et al., 2010) to assess the status of pathological conditions and, only in certain surgical procedures,

tractography plays a role in preoperative planning to avoid complications and preserve neurological function (Soni et al., 2017). Appart from the expansion that diffusion MRI field is experiencing, there are other factors that affect this lack of clinical applicability: from the purely economical risk-benefit assessment of acquiring a multi-million dollar MR scanner in environments that are already economically constrained, to the increasing sociocultural interest in mind-related topics[1] which leads to better funded research settings and, in turn, increases the technological gap between regular settings and well-funded ones.

One of the goals of this thesis is to develop biomarkers that are more easily adoptable by regular environments, specially aimed at their adoption by medical settings. For that, we aim at implementing biomarkers that are compatible with regular MR scans and acquisition procedures, while also being able to be computed with the computational resources, and time available in this kind of environments.

The search for biomarkers adoptability by regular environments must take into account the decrease in SNR that reduced acquisitions with less powerful MR scans lead to. In a field like diffusion MRI, a decrease in SNR might be critical in fitting problems that are already prone to stability or accuracy issues. Novel biophysical models often employ projections onto lower dimensional spaces so that the complex non-linear fitting problems are more computationally tractable, less sensitive to noise, and better conditioned for optimization. A common approach for this is the use of rotational invariants from spherical harmonics (RISH) to decouple the microstructure from orientation information which, theoretically, also improves numerical stability, the robustness to noise, and accelerate fitting problems. However, in pursuit of these properties, researchers have overlooked the noise effect on said rotational invariants. One of these effects is that Rician noise-contaminated signals skew the distribution of higher-order RISH projections, which leads to inaccurate and weakly-reproducible results.

For that, and given that biophysical models are of great interest to both the research and clinical community, we aim at studying the noise impact on biophysical biomarkers derived from rotational invariants, while proposing new alternatives to fit the diffusion-weighted signal onto biophysical models that have been proven to be more accurate, stable, reproducible, and reliable than current implementations.

Another issue with the adoptability of current biomarkers by regular settings (specially clinical ones) is a lack of confidence on novel biomarkers given their troublesome reliability evaluation. Neurological biomarkers are difficult to assess given their symbolic nature—which may not hold any purely biophysical meaning to be evaluated against—or the clinical impossibility of accessing a living brain to evaluate their trustworthiness. In the interest of rapid expansion, diffusion MRI community overlooks the reliability assessment of novel biomarkers, which are evaluted in group studies that portray weak statistical inferences with datasets that are often poor in terms of quality. All this only adds another barrier to the full integration of biomarkers, as clinical environments are unable to adopt innovative implementations that might be more insightful than previous ones.

[1] During the last decade, there has been an explosion in the cultural interest in topics such as psychology and mental health (Mitsea et al., 2023); nootropics or cognitive enhancers (Dresler et al., 2018) linked to productivity and motivational sciences; mindfulness, meditation and other derivates from eastern medicine; or robotics, brain-machine interfaces or neuroprosthetics (Jangwan et al., 2022), with a special mention to artificial intelligence

In an effort to tackle this issue, we aim at developing and implementing a new methodology of reliability assessment, which is intended not to be obscured by statistical weaknesses. By focusing on separability and repeatability, rather than on repeatability alone, we aim at portraying properties that are closer to the medical concepts of especificity and sensibility. The proposed methodology offers a way of estimating inter-group differences, a property that has a direct effect on the sample size required for subsequent studies.

To sum up, there exists an imbalance of the adoptability of novel biomarkers. While a revision of currernt economical funding to research or clinical settings could have the biggest impact, it is a challenging topic this thesis does not aim at dealing with. It aims, nevertheless, at laying another stone in the path for field convergence, and at closing the ever-increasing technological void between well-funded environments and regular ones. This is done by impacting a series of domains that have an effect on said gap, such as developing novel, useful and efficient biomarkers steeming from advanced phenomenological or biophysical models, that portray brain function, connectivity and development in health or disease, and that are able to be computed in a timely manner, with regular MR scanners and computational infrastructure, while ensuring a level of reliability comparable to those of the state-of-the-art. Here, we also aim at implementing efficient fitting strategies to increase the applicability of biophysical models in standard environments by reducing the effect of noise in estimates and, thus, increase their reliability and stability. Finally, the thesis aims at closing the gap by implementing methodologies that assess the trustworthiness of biomarkers in a way that does not get obscured by the nature of the data acquired or the statistical weaknesses the study might have, while ensuring a link between the clinical's especificity and sensibility.

1.2 Objectives

This thesis, thus, revolves around the hindered applicability of current diffusion MRI-based white matter biomarkers on regular research and clinical settings. These limitations, that often derive from economical factors, have led to a gap in the adoption of state-of-the-art biomarkers between high technologically-advanced settings and regular ones. We aim at closing the gap by developing novel biomarkers that are relevant in both clinical and research environments, that show a level of reliability comparable to those from state-of-the-art, and that are able to be efficiently estimated with the economical limitations of typical standard environments.

For this we have proposed the following main objective: **To develop advanced and efficient diffusion MRI-based biomarkers for the brain's white matter that are both reliable and applicable in standard research and clinical settings**. The main objective is divided in the following subobjectives:

SO1. To define and implement novel and efficient biomarkers for the brain's WM that provide relevant insights by describing diffusion properties.

Currently, biomarkers are being developed at a fast pace by research communities. However, most of them require highly-advanced MR scanners, powerful computational resources or time-expensive acquisition sequences. We aim at describing diffusion phenomena via useful

biomarkers—i.e. that have shown to be relevant in clinical or research fields—and that are attainable under the technological, computational and time-related limitations of regular settings.

SO2. To enhance the reliability and stability of microstructural biomarkers by assessing and mitigating the thermal noise effect.

A number of biophysical models have been developed over recent years by the diffusion community. While they theoretically capture important aspects of the microstructural landscape in which diffusion takes place, they are also affected by thermal noise MR scanners capture. Such thermal noise further affects the reproducibility, reliability and stability of the metrics.

SO3. To provide a reliability comparison of EAP-based biomarkers via an extended state-of-the-art validity evaluation methodology, with a direct assessment of their effect sizes, and providing required sample sizes for statistically powerful results in subsequent group analyses.

The neuroimaging community is acquiring a big number of diffusion MRI data to assess the validity of the biomarkers currently being developed. However, such assessment is troublesome, as it often focuses on the test-retest capabilities of the metrics (i.e. repeatability) rather than on their reliability. At the same time, most of the datasets are limited in regards to their sample size, which might confound the statistical insights drawn from them.

1.3 Methodology

Exploring the depths of quantitative diffusion MRI, one might think this project is primarily developed under a positivistic perspective. From a preclinical point of view, either the signal processing approaches or the biophysical models are used in order to obtain more accurate and reliable descriptions of the diffusion processes. It is important to notice, however, that the biomarker's theorized, developed and implemented are not necessarily linked to any measurable phenomena. This ultimately results in a pragmatic view of the results, which are not intended to be the most accurate—there might be no accuracy in depicting phenomenological aspects of diffusion—but rather the most *useful* in terms of clinical and research applicability. Therefore, the project itself is developed under a pragmatic cosmovision, where diffusion processes characterizations, constrained by the quantity and quality of the data available, are evaluated in said dimensions while, at the same time, assessed in its applicability and usability.

Regarding the research design of the project, the engineering research method (Koen, 2003) will be followed. This involves the observation of already adopted solutions, the definition of areas of improvement, the proposal of approaches that tackle such areas, their development, and the iterative analysis for their refinement, finally leading to their evaluation with regards to the series of dimensions that have to be improved.

Therefore, the following phases are detailed:

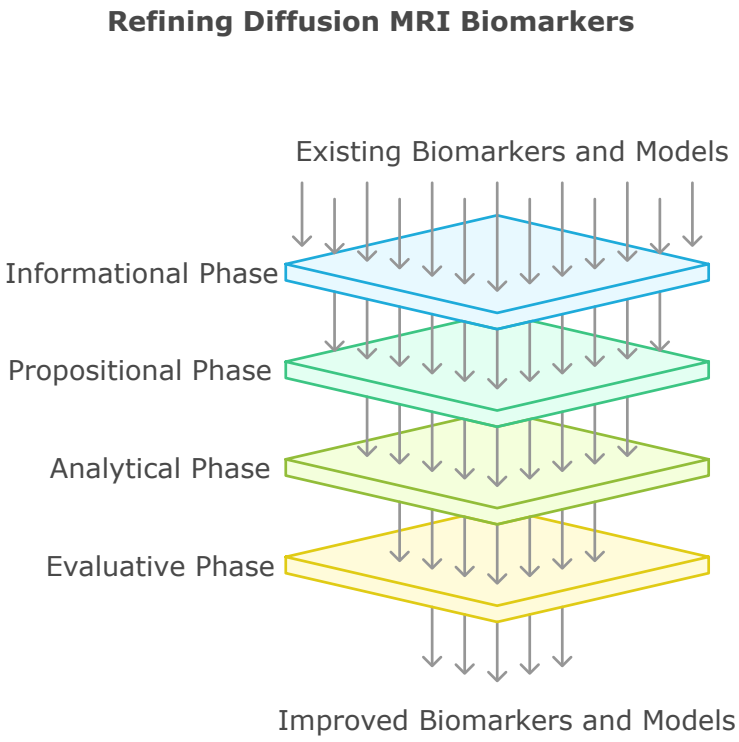


Figure 1.1: Methodology Schema.

The Informational Phase which involves the careful review of the existing solutions to a given problem, often through a literature review, and the examination of their possible drawbacks. In this thesis, it entails the assessment of existing biomarkers derived from either phenomenological or biophysical models, their purpose in regards to clinical or research environments, the theory upon which they are developed and their limitations or weaknesses. This latter includes the analysis of the formal limitations within the developed theory in regards to the diffusion phenomena or microstructural environment they aim to uncover, the biomarkers' reliability, and their applicability on regular clinical settings. This phase resulted in the review of some of the most important state-of-the-art biomarkers in diffusion MRI, and their analysis with respect to the aforementioned domains. Specially, we detected growing concerns in the evaluation of biomarker's reliability and their troublesome adoption by the clinical community (or regular research settings), who lack the technological infrastructure and time-expensive requirements of such biomarkers.

The Propositional Phase which entails the formulation of hypotheses, methods, markers, models, or theoretical frameworks that tackle one or more of the objectives. This phase culminated in the proposal of a series of hypotheses, each one adressing a particular aspect of the deficiencies depicted in the previous phase and resulting in a given contribution. The domains of deficiencies are tackled in three contributions: First, the development of biomarkers that are (1) easily adoptable by the standard environments without the need for advanced

scanners, time-expensive procedures or computationally-intensive resources; and (2) that show levels of reliability comparable to those in the state-of-the-art. Second, the improvement in reliability, reproducibility and applicability of state-of-the-art biophysical models and their derived biomarkers. Third, the implementation of a validity evaluation methodology that aims at uncovering *real* reliability of biomarkers, so that trustworthiness on these biomarkers is comparable to that of the ones currently being used by clinical and research communities.

The Analytical Phase in which the proposed solutions are analysed and probed. In the context of the thesis, this phase comprises the implementation of the proposed solutions and their evaluation via *in vivo* or *in silico* experiments that tackle the domains of interest, and culminates in the appraisal of the three proposals with regards to the deficiencies already pointed out.

For example, a series of *in vivo* experiments have been carried out to quantify the level of accuracy and robustness of the developed metrics. Also, by using the *multidimensional* datasets shown in Section 1.4, reliability will be evaluated by assessing the metrics performance on several sessions from the same subject. However, as mentioned earlier, accuracy validation in diffusion MRI is troublesome as it is often hard (or even impossible) to obtain real ground truth data. For this, *in silico* experiments are carried out by simulating biophysical environments from where to compare the estimated data.

The Evaluative Phase, whose goal is the assessment of the contributions in terms of accuracy, reliability, and robustness. This phase led to the careful evaluation of the proposed biomarkers with regards to their computational efficiency, and time-consumption nature, in order to derive assess their potential applicability. At the same time, it also resulted in the evaluation of the biomarkers reliability by using the proposed methodology. Finally, this phase also resulted in the evaluation of some of the state-of-the-art biomarkers via the proposed reliability assessment methodology.

The result of the previous phases are a series of contributions in the form of scientific publications in journals or conferences. These contributions are shown in Section 1.5.

1.4 Materials

This section outlines the materials employed throughout the project. Central to this work is the use of diffusion MRI datasets, which are essential for developing and validating the proposed contributions. However, given that this project focuses on the reliability of metrics, dMRI datasets should ideally include multiple subjects, each with repeated imaging sessions. This setup allows for evaluating a metric's repeatability—the consistency of its estimates within the same subject—and separability—the ability to distinguish between different subjects.

It is important to note that this project remains within the preclinical stage of biomarker development. Accordingly, all subjects in the datasets are healthy individuals, intended to

represent a general cross-section of the human brain across demographic factors such as age, biological sex, and ancestry.

1.4.1 MRI Datasets

MICRA or Micro-structural Image Compilation with Repeated Acquisitions database contains five repeated sets of multishell samplings DWI for each of six healthy volunteers. The dataset was acquired in the Cardiff University Brain Research Imaging Centre (CUBRIC) on 2021.

Technical details: The images were acquired within a 2-week period, approximately at the same time for each participant—avoiding potential diurnal effects—on an ultra-strong gradient 3T Connectome MRI scanner using a single-shot spin echo EPI with TR / TE = 3000 / 59 ms, six different shells at $b = [200, 400, 1200, 2400, 4000, 6000]$ s/mm^2, with [20, 20, 30, 61, 61, 61] gradient directions, respectively, in-plane resolution 2 $\times$ 2 mm^2, slice thickness 2 mm and $\Delta/\delta = 24/7$ ms.

Processing pipeline: The data were preprocessed by removing Gibbs ringing artifacts (Kellner et al., 2016), via MRtrix3 (Tournier et al., 2019) and correcting susceptibility-induced distortions (with FSL's `topup`; Andersson et al., 2003; Smith et al., 2004) and B1 field inhomogeneity (Zhang et al., 2001) with MRtrix3.

HCP-MGH or Human Connectome Project MGH-USC dataset comprises diffusion data from two repeated sets of multishell data for each of 45 healthy monozygotic twins (21 pairs plus 4 MZ twins without retest of cotwin). The dataset was acquired from 2012 to 2015 in the Massachusetts General Hospital.

Technical details: The high-quality DWI volumes acquired on a Siemens 3T Connectome Scanner (Siemens) with maximum gradient strength 300 mT/m. The data were acquired with a spin-echo echo planar imaging (EPI) with TR / TE = 8000 / 57 ms, four different shells at $b = [1000, 3000, 5000, 10000]$ s/mm^2 with [64, 64, 128, 256] diffusion gradients each, and 40 interleaved nonweighted baselines, in-plane resolution 1.5×1.5mm^2 and slice thickness 1.5 mm, and pulse separation time/diffusion gradients length $\Delta/\delta = 21.8/12.9$ ms.

Preprocessing: The outermost shell has been removed in order to validate our proposal with more standard acquisitions.

CUBRIC - Test/Retest consists of diffusion data from five healthy subjects, each scanned across two sessions. The dataset was originally intended for Henriques et al. (2021) and Veraart et al. (2021); and was acquired in the Cardiff University Brain Research Imaging Centre (CUBRIC) on 2021.

Technical details: All data were acquired on a Siemens Connectom 3T MRI scanner using a 32-channel receiver coil. Subjects underwent a test and retest scans during the same visit, but they were re-positioned between the scans. All dMRI data were acquired with a multi-band blipped-CAIPI accelerated (SMS = 2) EPI sequence. The acquisition was further accelerated using GRAPPA acceleration (R = 2), using adaptive combine reconstruction in order to obtain Rician distributed results (Walsh et al., 2000). All images were acquired with TR / TE = 3500 / 66 ms and an in-plane spatial resolution of 2.5 $\times$ 2.5mm^2 and slice thickness of 2.5

	Subjects and Sessions	b-shells ($\cdot 10^3$ s/mm^2)	Samples per shell	Contribution
MICRA Connectom 3T	6 × 5	[0.2, 0.4, 1.2, 2.4, 4, 6]	[20, 30, 61, 61, 61, 61]	① ③
HCP Connectom 3T	45 × 2	[1, 3, 5, 10]	[64, 64, 128, 256]	①
CUB - T/R Connectom 3T	5 × 2	[0.5, 1, 2.5, 6]	[30, 30, 30, 120]	②
ZJU Magnetom 3T	3 × 3	[1, 2, 3]	[30, 30, 30]	③

Table 1.1: Summary of the datasets used throughout this thesis.

mm. In addition to 23 non-DWI, DW images were recorded by applying diffusion gradients with $\Delta/\delta = 30/15$ ms and varying gradient amplitude. The acquisition distributed DWIs across the following shells $b = [500, 1000, 2500, 6000]$ s/mm^2, with [30, 30, 30, 120] diffusion gradients each. The last shell, however, was acquired in two sets of 60 uniformly-distributed gradients. Ten non-DW images with reversed phase encoding were acquired in support of susceptibility-induced geometrical distortion correction.

ZJU - Multicenter or Zhejiang University multicenter diffusion MRI dataset contains diffusion MRI data from three traveling subjects scanned at eight centers. Additionally, for intra-center research purposes, the subjects were scanned across three sessions in one center. Since inter-center reproducibility is out of the scope of this thesis, only the three sessions from the three subjects at one center will be used. The data was acquired for the work by Tong et al. (2019).

Technical details: The study was carried out using eight identical Siemens 3T MR Magnetom Prisma scanners, each running the Syngo MR VD13D software and equipped with the same gradient system, capable of a maximum gradient strength of 80 mT/m and a slew rate of 200 T/m/s. Three healthy participants (one male aged 23, and two females aged 26 and 23) with prior experience undergoing MRI scans and known for minimal head movement were scanned at all eight sites using a 64-channel head coil. All scans were performed within a one-month window and followed a uniform acquisition protocol. To assess intra-center variability, the experiment was repeated three times at Center H with these sessions labeled H-1, H-2, and H-3. DWIs were obtained using a simultaneous multi-slice (SMS) spin-echo echo planar imaging (EPI) prototype sequence. The imaging parameters were as follows: TR / TE = 5400/71 ms, a resolution of $1.5 \times 1.5 \times 1.5$ mm^3, bandwidth GRAPPA factor = 2, and SMS factor = 3, with reversed phase-encoding (PE) directions along anteroposterior (AP) and posteroanterior (PA) separately. The diffusion time and duration Δ/δ were 34.4/15.9 ms, respectively for the monopolar diffusion gradients. The diffusion scheme distributed DWIs across the following shells $b = [1000, 2000, 3000]$ s/mm^2, each consisting of 30 diffusion gradients, non-colinear between any two shells. Six non-diffusion frames were equally distributed in the scheme.

Table 1.1 provides a summary of the technical specifications for the diffusion MRI datasets employed throughout this work. For each dataset, the table lists the scanner used, the number of subjects and scanning sessions per subject, the number of diffusion shells, the number of gradient directions per shell, and the corresponding scientific publication in which the dataset has been utilized.

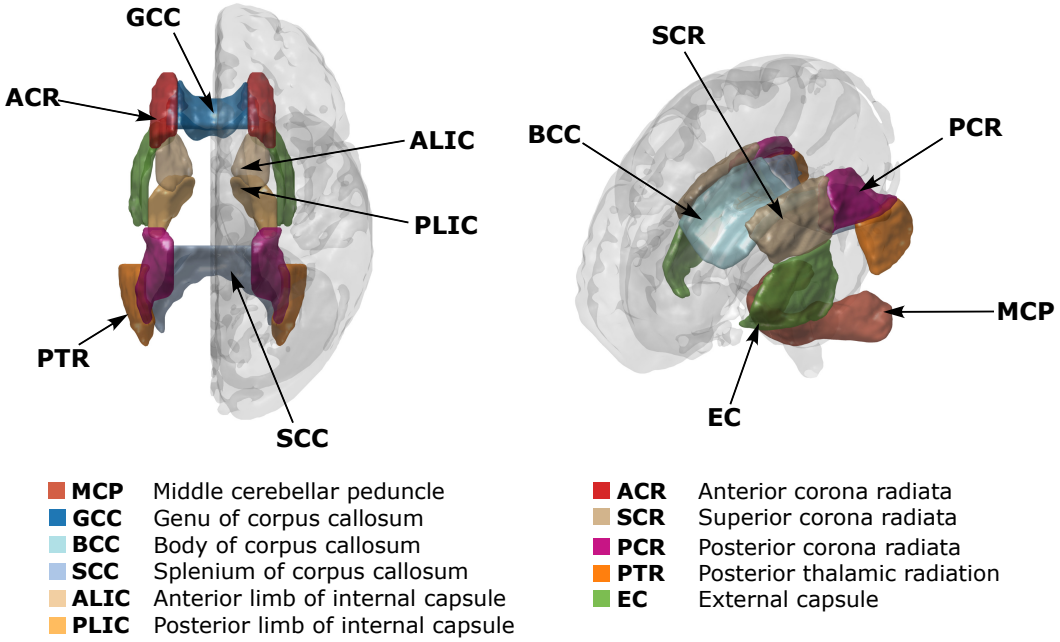


Figure 1.2: 3D visualization of 11 JHU DTI-based labels in standard space. Note that some of the regions are not shown in either subfigure, in order to ease the visualization.

1.4.2 Atlases

Now that the MRI datasets used in this study have been described, it's important to define the anatomical framework used for analysis. In diffusion MRI, atlases play a key role in identifying and segmenting white matter structures in a standardized way. They allow for region-based analyses, making it possible to compare diffusion metrics across subjects and groups while accounting for anatomical variability. By mapping individual data to a common space, atlases help ensure that results are both interpretable and reproducible.

One of the most widely used atlases for white matter analysis in diffusion MRI is the JHU ICBM-DTI-81 White Matter Atlas (Oishi et al., 2008). This atlas was created from a population-averaged DTI template and provides a standardized set of labels for major white matter tracts in the brain. It is defined in MNI space and includes both deterministic tract segmentations and probabilistic maps, making it useful for different types of analyses. The JHU atlas allows for consistent identification of white matter structures across subjects, which is particularly important when comparing diffusion-derived measures like FA or MD across groups.

Figure 1.2 shows the 3D representation of the 11 JHU labels used in the first contribution (París et al., 2023a).

1.4.3 Computational Resources

Diffusion MRI datasets often contain dozens of sessions, each session consisting of hundreds of thousands of voxels, and each voxel scanned over a set of dozens of gradient directions and

magnitudes. That results in overwhelmingly large data structures that often must be fit onto models by complex nonlinear and regularized fitting problems for each of hundreds of voxels. Needless to say, the processing of such data structures require intensive computational resources and infrastructure in order to keep processing times in realistic ranges.

The Laboratorio de Procesado de Imagen (LPI) in Universidad de Valladolid uses three powerful servers located at their offices, which include three high-performance computing servers, Tebas, Tanis and Isis, all running Debian GNU/Linux 12. Tebas is equipped with an Intel Xeon E5-2697 v4 CPU (34 physical cores, 2.30 GHz) and 500 GB of RAM, while Tanis features an Intel Xeon E5-2695 v3 CPU (26 physical cores, 2.30 GHz) and 115 GB of RAM. Each server has dual-socket processors, ranging from 56 to 128 threads each, and RAM capacities suitable for processing very large data volumes (between 128 and 1024 GB). Each server has several server-class GPUs (2 or 4 GPUs each) with between 8 and 24GB of RAM. In addition, LPI owns a network-attached, NVMe-based storage space of more than 40 TB with a high capacity for hardware fault tolerance, a sustained transfer rate of 10 Gbps (scalable) and very low latency access to databases, thanks to the aforementioned NVMe disks. The group has its own dedicated network infrastructure, uninterruptible power supplies for all critical equipment and a fully automated backup system with a capacity of more than 500 TB. In terms of software, the group has professional-level database servers currently in use, software for the management of its critical equipment and network infrastructure. The software environment included Python 3.11.2 and MATLAB® (R2022b, R2023b, and R2024b; The MathWorks, Inc., Natick, MA) , providing the necessary computational tools for data processing and analysis, including free state-of-the-art software, such asFreeSurfer, FSL, SPM and MRtrix3.

Brainlife

Brainlife (Brainlife.io, 2025) is a cloud-based platform intended for neuroimaging research. It is a useful tool that enables the processing, analysing, and sharing of neuroimaging data. Additionally, it provides researchers with a scalable and reproducible environment where to run complex pipelines without the need of intensive computational resources locally.

1.5 List of Publications

This section presents the main contributions resulting from this doctoral project, including contributions to scientific journals, talks and posters in conferences, or key methodological developments. The contributions reflect the impact of this project in the dMRI field, with an emphasis on reliability of the proposed methods. Figure 1.3 shows an overview of the thesis contributions to journals.

And, in what follows, a detailed list of the contributions is shown:

C1. Efficient estimation of propagator anisotropy and non-Gaussianity in multishell diffusion MRI with micro-structure adaptive convolution kernels and dual Fourier integral transforms.

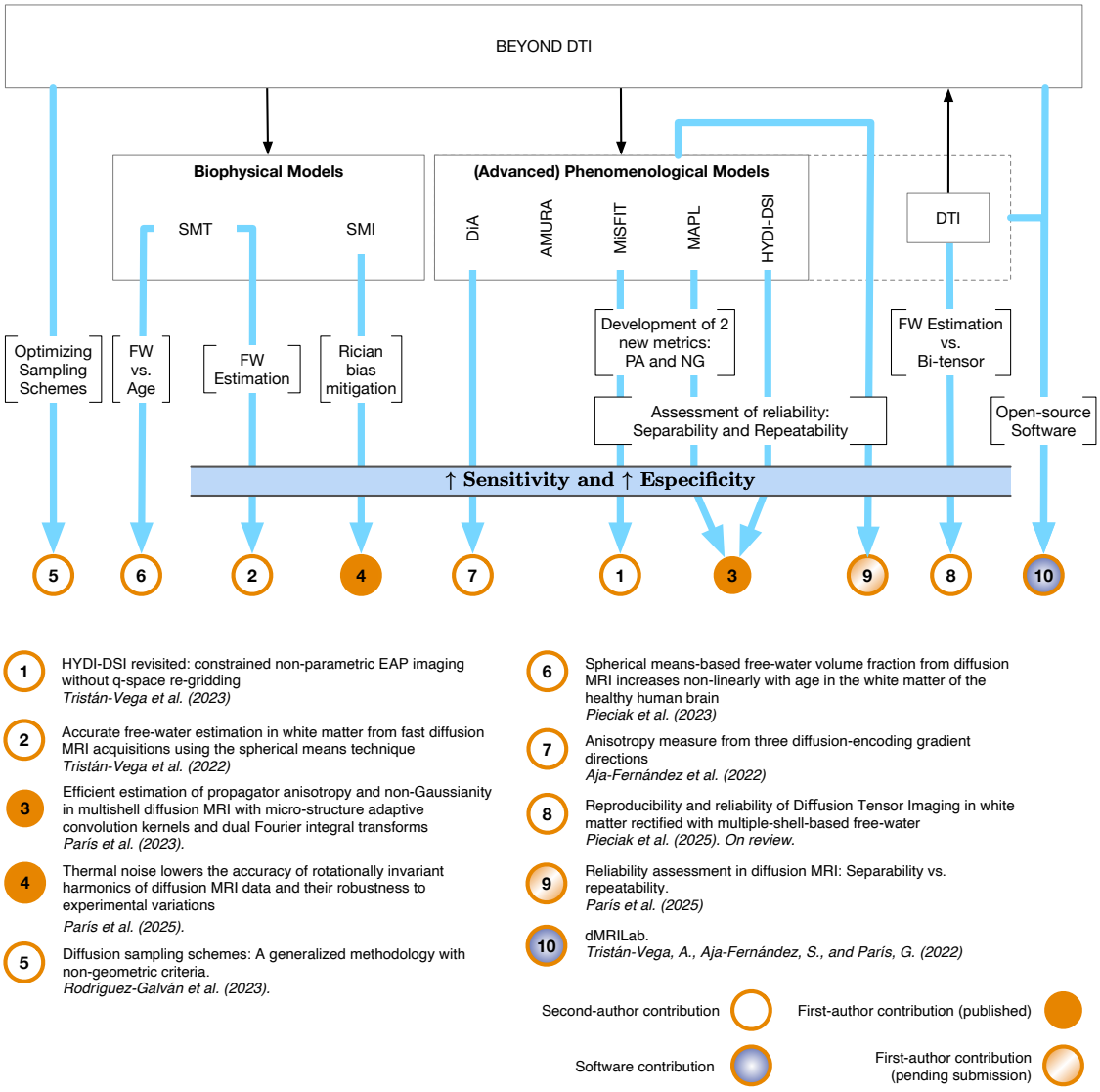


Figure 1.3: Thesis contributions overview.

París, Guillem; Pieciak, Tomasz; Aja-Fernández, Santiago; Tristán-Vega, Antonio. (2023). Magnetic Resonance in Medicine, Wiley.

C2. Thermal noise lowers the accuracy of rotationally invariant harmonics of diffusion MRI data and their robustness to experimental variations. **París, Guillem**; Pieciak, Tomasz; Jones, Derek K.; Aja-Fernández, Santiago; Tristán-Vega, Antonio; Veraart, Jelle. (2025). Magnetic Resonance in Medicine, Wiley.

C3. HYDI-DSI revisited: Constrained non-parametric EAP imaging without q-space re-gridding. Tristán-Vega, Antonio; Pieciak, Tomasz; **París, Guillem**; Rodríguez-Galván, Justino R; Aja-Fernández, Santiago. (2023). Medical Image Analysis, Elsevier.

C4. Accurate free-water estimation in white matter from fast diffusion MRI acquisitions using the spherical means technique. Tristán-Vega, Antonio; **París, Guillem**; de Luis-García, Rodrigo; Aja-Fernández, Santiago. (2022). Magnetic Resonance in Medicine, Wiley.

C5. Anisotropy measure from three diffusion-encoding gradient directions. Aja-Fernández, Santiago; **París, Guillem**; Martín-Martín, Carmen; Jones, Derek K; Tristán-Vega, Antonio. (2022). Magnetic Resonance Imaging, Elsevier.

C6. Diffusion sampling schemes: A generalized methodology with nongeometric criteria. Rodríguez-Galván, Justino R; **París, Guillem**; Tristán-Vega, Antonio; Alberola-López, Carlos. (2023). Magnetic Resonance in Medicine, Wiley.

C7. Spherical means-based free-water volume fraction from diffusion MRI increases non-linearly with age in the white matter of the healthy human brain. Pieciak, Tomasz; **París, Guillem**; Beck, Dani; Maximov, Ivan I; Tristán-Vega, Antonio; de Luis-García, Rodrigo; Westlye, Lars T; Aja-Fernández, Santiago. (2023). Neuroimage, Academic Press.

However, contributions come in many ways and forms. The second main category of scientific contributions come in the form of posters. Such posters are presented in national or international congresses and imply some degree of discussion of the poster with the attendants of such conferences.

P1. Impact of Rician Bias on diffusion MRI biophysical modeling. **París, Guillem**; Pieciak, Tomasz; Jones, Derek K.; Aja-Fernández, Santiago; Tristán-Vega, Antonio; Veraart, Jelle. (2024). ISMRM - Annual Meeting. Singapore, Singapore

P2. Efficient Estimation Of The Generalized Cross-Validation Regularization Weight: An Application For HYDI-DSI-QP. **París, Guillem**; Pieciak, Tomasz; Aja-Fernández, Santiago; Tristán-Vega, Antonio. (2023). ISMRM - Iberian Chapter. Valladolid, Spain

P3. Free-water correction of Diffusion Kurtosis Imaging in adult lifespan studies. Tristán-Vega, Antonio; Pieciak, Tomasz; **París, Guillem**; Rodríguez-Galván, Justino R; Aja-Fernández, Santiago. (2023). ISMRM - Iberian Chapter. Valladolid, Spain

P4. Free-water volume fraction increases non-linearly with age in the white matter of the healthy human brain. Tristán-Vega, Antonio; **París, Guillem**; de Luis-García, Rodrigo; Aja-Fernández, Santiago. (2023). ISMRM - Iberian Chapter. Valladolid, Spain

P5. HYDI-DSI: Efficient estimation of the Generalised Cross-Validation regularisation term from several diffusion parameters. **París, Guillem**; Aja-Fernández, Santiago; Tristán-Vega, Antonio. (2022). ISMRM - Workshop on diffusion MRI: From research to clinic. Amsterdam, Netherlands

P6. Efficient Estimation of Propagator Anisotropy and Non-Gaussianity with MiSFIT. **París, Guillem**; Pieciak, Tomasz; Aja-Fernández, Santiago; Tristán-Vega, Antonio. (2022). ISMRM - Iberian Chapter. Lisbon, Portugal

P7. Assessing reliability of white matter metrics in diffusion MRI based on ROI variability. **París, Guillem**; Pieciak, Tomasz; Aja-Fernández, Santiago; Tristán-Vega, Antonio. (2021). Brayn Conference. Pysa, Italy.

1.6 Thesis Overview

Chapter 2 introduces the necessary literature in order to fully understand the contributions of this project. After a quick introduction to the diffusion phenomena, the different analytical methods for diffusion signal characterization will be explained: from diffusion tensor imaging, to the state-of-the-art models and metrics. Finally, the chapter will cover one of the main topics of this project: the current validation of diffusion metrics via repeatability, separability, and reliability; and their relationship to the clinically-relevant sensitivity and specificity.

Chapters 3 to 5 consist of the three main contributions of this doctoral thesis. Namely, the development and implementation of the propagator anisotropy and non-Gaussianity biomarkers for MiSFIT framework (Chapter 3); the assessment of thermal noise and Rician bias' effect on rotationally invariant-based biophysical models, together with the proposal of fitting alternatives that reduce such effect (Chapter 4); and a reliability study of phenomenological biomarkers in neuroimage, which provides a comparison of state-of-the-art, EAP-based metrics via a variance decomposition methodology that aims at characterize the statistical performance of biomarkers while enhancing experimental planning (Chapter 5).

Finally, chapter 4 rounds up the thesis and summarizes its main findings. It also provides a quick view on the future work that might stem from these contributions as well as from the field.

Knowledge is not for knowing: knowledge is for cutting.

— Foucault (1972)

2

Background

2.1 Introduction to diffusion MRI

The dilution of a drop of ink in water may seem simple, but what are the underlying physical mechanisms at play? Under certain conditions—e.g., same-density liquids—the drop dilutes and ends up fully integrated into the water, achieving a perfect mix even without any water movement. This process is called diffusion and is crucial in the noninvasive study of tissues. Its study in the brain has become important, given its sensitivity at a molecular level, which enables the characterization of brain microstructure in both health and disease.

More technically, diffusion is the process by which molecules or particles move through a medium from a high-energy area to a lower-energy area (i.e., high concentration to low concentration) due to random thermal motion. The process is driven by the tendency of systems to reach an equilibrium state. Diffusion MRI, therefore, aims to sensitize MRI technologies to the diffusion of water molecules within tissues. In unrestricted environments, such as the glass of water, the movement follows Brownian motion—an erratic, random, and Gaussian-distributed movement of molecules immersed in a fluid affected by collisions with other molecules. In these cases, the diffusion is said to be isotropic (i.e., movement radiates to all directions equally). In restricted environments such as in biological tissues, however, molecules might have to interact with physical barriers. These interactions can be either hindered—molecules are slowed down by obstacles and can move through or around them—or restricted—molecules are confined within boundaries. When the obstacles hinder/restrict particle movements in a set of directions but not in the others, the diffusion is said to be anisotropic. This anisotropic diffusion profile remains consistent with Brownian motion and can be considered to follow a Gaussian distribution, provided that the

restrictions are not *extreme* (see Section 2.2 for a more detailed explanation).

Studying diffusion in specific environments reveals important aspects of the biophysical phenomena or microstructural landscape where such diffusion occurs. This understanding is ultimately the principle by which diffusion MRI works and where all conclusions are subsequently drawn. However, the validity of MRI extends to diffusion MRI, compromised by key challenges such as noise, artifacts, resolution limits, and the complexity of estimating reliable biomarkers. While critical advancements are being made in each of these dimensions, it is often helpful to remember that these traits are the very nature of the diffusion MRI.

Rather than obscuring the path for neuroscientific research, diffusion MRI has opened several ways into noninvasive brain exploration, becoming a crucial tool in neurological research settings. Nonetheless, its many limitations have hindered its capability to be fully adopted by the clinical community, which still relies on old (but reliable) modeling techniques. Roughly 40 years since the first application of diffusion MRI on an in vivo brain, this thesis aims to facilitate the incorporation of diffusion MRI in clinical settings while encouraging medical doctors to put its new models and biomarkers to the test.

2.1.1 Diffusion MRI principles

As the name indicates, the basic principle of diffusion MRI radicates in the Magnetic Resonance Imaging. Explaining how MRI works could already be a somewhat significant doctoral thesis and is thus out of this project's scope. Instead, we will explain how diffusion MRI systems differ from the classic MRI ones. That is, how they become sensitized to the movement of water molecules.

The only difference between a "classic" MRI scanner and a diffusion MRI scanner is the ability of the latter to use (emit and read from) strong diffusion-sensitizing gradients. By applying these so-called diffusion gradients, the spins that are in its reach and resonate at the Larmor frequency of the scanner dephase. This dephasing is relative to the molecule's position with respect to the location where the gradient is focused.

In a completely still (and theoretical) environment, if the opposite gradient is applied to the exact location, the molecules would be rephased to their original state (phase equal to 0), and thus producing no signal loss. However, since in biological tissues molecules experience motion, the second gradient does not completely rephase all molecules within its reach. This incomplete rephasing leads to a reduction—or attenuation—of the received signal, which is related to the number of spins that have moved between the application of both gradients (see Fig. 2.1). Therefore, in diffusion MRI, the amount of signal loss encodes information about the extent and directionality of water diffusion: greater attenuation indicates more substantial displacement along that gradient direction (see Fig. 2.1).

As previously introduced, biological tissues in the brain restrict the diffusion in specific directions, resulting in anisotropic diffusion. The diffusion phenomena can be reconstructed by acquiring attenuation signals over a set of diffusion-encoding gradient directions in a given voxel. This serves as the foundation for all diffusion MRI research, yet the field continuously strives for

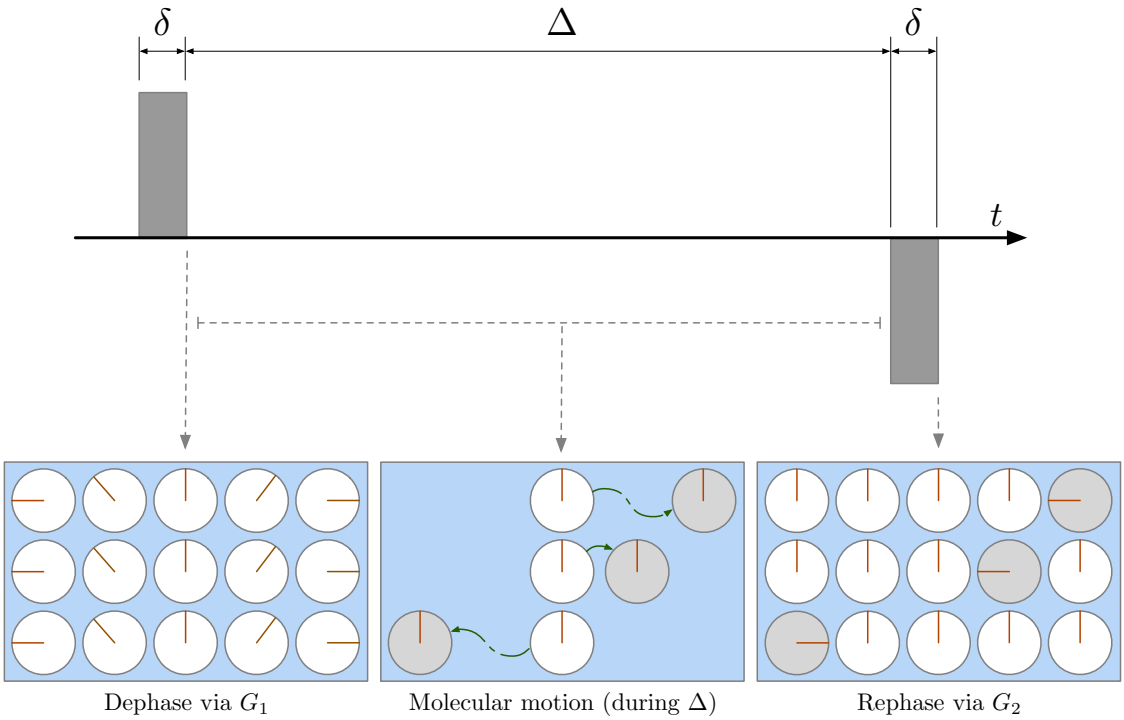


Figure 2.1: Gradients effect on the phase of diffusing spins. After the first diffusion gradient, spins become dephased relative to the center of the gradient application. Then, while the diffusion time Δ occurs, some spins diffuse onto other locations. Finally, after applying the last and opposite diffusion gradient, the spins become rephased, except for those that have moved from their previous location. There is, then, a fall in the net magnetization vector that is read as an attenuation of the diffusion signal.

progress in every direction: from acquisition protocols to allow a better or more efficient tissue characterization and the modeling of the diffusion phenomena or the biophysical environment in which they take place, to the efficient estimation of neural pathways across the brain.

2.1.2 Technical viewpoint

Bloch Equation

To understand the equations that describe the attenuation of signals due to the diffusion processes in the presence of diffusion-sensitizing gradients, we first have to describe the dynamics of nuclear magnetization in the presence of a magnetic field. For that, we must introduce the Bloch equations, which model the magnetization vector $\mathbf{M}(t) = (M_x(t), M_y(t), M_z(t))$ across time, under the influence of an external magnetic field $\mathbf{B}(t) = (B_x(t), B_y(t), B_z(t))$:

$$\frac{\partial \mathbf{M}(t)}{\partial t} = \gamma \mathbf{M}(t) \times \mathbf{B}(t) - \begin{pmatrix} \frac{M_x(t)}{T_2} \\ \frac{M_y(t)}{T_2} \\ \frac{M_z(t) - M_0}{T_1} \end{pmatrix} \tag{2.1}$$

where γ is the gyromagnetic ratio, T_1 is the longitudinal relaxation time, T_2 is the transverse relaxation time, and M_0 is the magnetization at thermal equilibrium in the presence of B_0.

Bloch-Torrey Equation

A modified version of the previous Bloch equations is needed in order to account for the effect of diffusion processes in the presence of diffusion gradients:

$$\frac{\partial \mathbf{M}(\mathbf{r},t)}{\partial t} = \mathbf{D}\nabla^2\mathbf{M}(\mathbf{r},t) + \gamma\left(\mathbf{M}(\mathbf{r},t)\times\mathbf{B}(\mathbf{r},t)\right) + \begin{pmatrix} \frac{M_x(\mathbf{r},t)}{T_2} \\ \frac{M_y(\mathbf{r},t)}{T_2} \\ \frac{M_z(\mathbf{r},t)-M_0}{T_1} \end{pmatrix} \tag{2.2}$$

where $\mathbf{M}(\mathbf{r},t) = (M_x(r_x,t), M_y(r_y,t), M_z(r_z,t))$ is the time-dependent magnetization vector of an excited sample located at $\mathbf{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ immersed in a magnetic field $\mathbf{B}$, $\mathbf{D}$ is the diffusion coefficient experienced by the particle, and ∇^2 is the Laplacian operator that accounts for spatial variation of the magnetization.

Stejskal-Tanner Equation

By solving the Bloch-Torrey equations under specific conditions, the Stejskal-Tanner equation—which relates the amount of signal attenuation due to diffusion in the presence of diffusion gradients—can be derived. Such conditions include, but are not limited to: (1) Constant background magnetic field, (2) Gaussian diffusion, (3) negligible relaxation effects during the diffusion time, (4) perfect gradient pulses with instantaneous switching, etc.

In practice, to isolate the effects of diffusion alone, the signal acquired with diffusion weighting is normalized by a reference signal without diffusion gradients, denoted $S(0)$. This yields the signal attenuation fraction:

$$E(\gamma,\delta,\Delta,G) = \frac{S(\gamma,\delta,\Delta,G)}{S(0)} = \exp\left(-\gamma^2\delta^2\left(\Delta - \frac{\delta}{3}\right)G^2\cdot D\right), \tag{2.3}$$

where G is the diffusion-sensitizing gradient magnitude $G = \|\mathbf{g}\|$, and both Δ and δ represent, respectively, the gradient separation and duration (see Fig. 2.1). The term D refers to the Apparent Diffusion Coefficient (ADC) and describes the mean diffusivity within the imaged voxel. For convenience, a variable that relates to the gradient strength and timing is defined as the b-value:

$$b = \gamma^2\delta^2\left(\Delta - \frac{\delta}{3}\right)G^2, \tag{2.4}$$

where the effective diffusion time is defined as $\tau = \Delta - \delta/3$.

Eq. (2.3), also referred to as Stejskal-Tanner equation, thus models the signal decay expected for free, isotropic diffusion. To capture more complex, anisotropic diffusion, the scalar coefficient

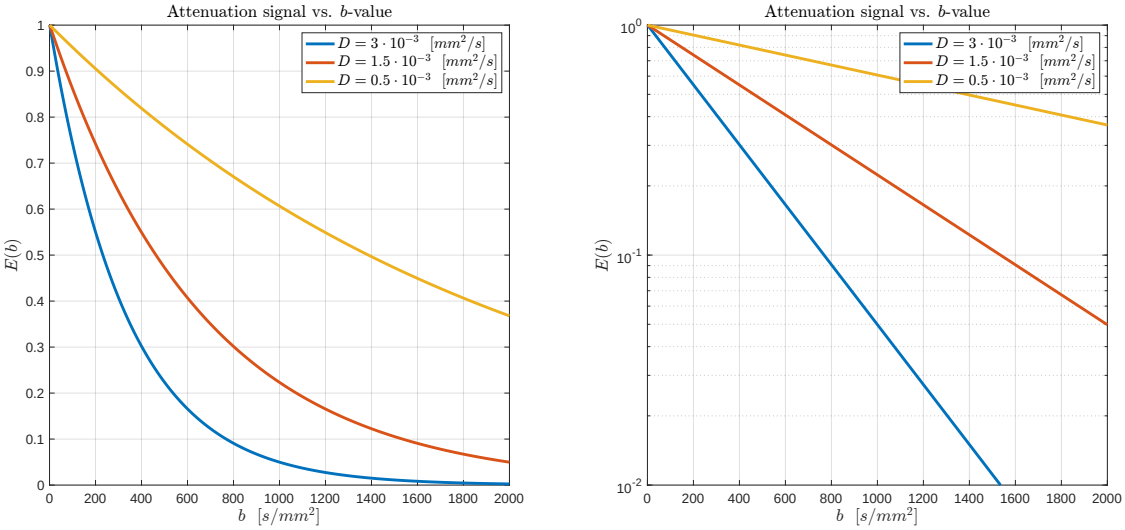


Figure 2.2: Attenuation signal for an increasing diffusion-weighted gradient magnitude. Up to a certain b-value, the attenuation signal is assumed to behave as a mono-exponential decay.

D is generalized to a second-order diffusion tensor $\mathbf{D}$. This tensor captures how the diffusion process varies with direction, and the signal attenuation for a given gradient direction depends on the projection of that direction onto the tensor. Mathematically, it entails the projection of the gradient direction onto the tensor, i.e. $\mathbf{g}^T\mathbf{D}\mathbf{g}$, thereby resulting in:

$$E(b, \widehat{\mathbf{g}}) = \frac{S(b, \widehat{\mathbf{g}})}{S(0)} = \exp\left(-b\ \widehat{\mathbf{g}}^T\mathbf{D}\widehat{\mathbf{g}}\right), \tag{2.5}$$

where $\widehat{\mathbf{g}}$ is a unitary vector pointing to the direction of the diffusion gradient applied.

Figure 2.2 (left) shows different realizations of the Stejskal-Tanner equation for a different range of b-values and diffusion coefficients. The values for said coefficients are not random: these are the values with which we will usually work. For example, the diffusion constant of free water at body temperature (37 °C) is $D = 3 \cdot 10^{-3}$ mm^2/s, often referred to as D_0. $D = 1.5 \cdot 10^{-3}$ mm^2/s is usually found in biological tissues; while $D = 0.5 \cdot 10^{-3}$ mm^2/s is found in tightly-packed structures. Note that since the behavior of the attenuation signal changes exponentially with respect to the b-values, these figures are usually displayed in logarithmic scales, such as the one in Fig. 2.2 (right).

2.2 Diffusion Tensor Imaging

One of the most prominent features of diffusion MRI is the ability to distinguish the direction in which anisotropic diffusion occurs. This is of particular interest in the brain, where anisotropic diffusion reveals properties of the underlying tissue. Diffusion Tensor Imaging (DTI, Basser et al., 1994) simplifies the diffusion within tissues so that it can be summarized in the tool that most efficiently packs 3D displacement: the tensor—a mathematical object that represents data varying across dimensions. Since, in this case, the data (i.e., diffusion) varies across the

three physical dimensions, a 3-by-3 rank-2 tensor is used:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix}. \tag{2.6}$$

Aligning with the biophysical properties of diffusion within tissues, the diffusion is presumed to have antipodal symmetry—identical in opposite directions—and to be positive definite—always positive. The tensor portrays these characteristics, resulting in a symmetric positive-definite matrix. Through a simple diagonalization process, the tensor can be parameterized by three orthogonal vectors describing the diffusion along them, together with three scalars representing the amount of diffusion in each direction. In this context, each eigenvalue (expressed in units of mm^2/s) represents the variance per unit time of the displacement distribution of water molecules along its corresponding eigenvector, which indicates the principal direction of diffusion. This squared unit reflects the fact that diffusion is modeled as a random walk: the mean squared displacement in three dimensions grows linearly with time, according to $\langle x^2 \rangle = 2Dt$, so the eigenvalues quantify how much molecules spread out (their variance) along each principal axis during the diffusion time.

Following from Eq. (2.6), the diagonalization results in:

$$\mathbf{D} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T, \tag{2.7}$$

where $\mathbf{\Lambda} = [\lambda_1, \lambda_2, \lambda_3] \cdot \mathbf{I}$ ($\mathbf{I}$ being the three-by-three identity matrix) contais the eigenvalues, and $\mathbf{Q}$ contains the (unitary) eigenvectors of $\mathbf{D}$, respectively:

$$\mathbf{Q} = \begin{bmatrix} v_{1x} & v_{2x} & v_{3x} \\ v_{1y} & v_{2y} & v_{3y} \\ v_{1z} & v_{2z} & v_{3z} \end{bmatrix} = \begin{bmatrix} \hat{v_x} \\ \hat{v_y} \\ \hat{v_z} \end{bmatrix}. \tag{2.8}$$

Suitably, the diagonalized tensor can be represented as an ellipsoid (see Fig. 2.3), whose axes are characterized by the eigenvalues in the directions of their respective eigenvectors. DTI rapidly became a breakthrough in diffusion MRI, enabling the characterization of microstructural changes associated with various neurological disorders and cognitive impairments (e.g., cognitive performance Maillard et al., 2019; AD Bergamino et al., 2021; Nakaya et al., 2022; aging Metzler-Baddeley et al., 2012; Chad et al., 2018; Kubicki et al., 2018; schizophernia Carreira Figueiredo et al., 2022; or first-episode psychosis Lyall et al., 2017, among others). The ability of DTI to distinguish anisotropic diffusion also provided a way of studying the white matter connectivity (Basser et al., 1994).

2.2.1 Tractography

One of the most widespread research applications of DTI is tractography (also known as fiber tracking). Using information derived from the diffusion tensor—in particular, the principal

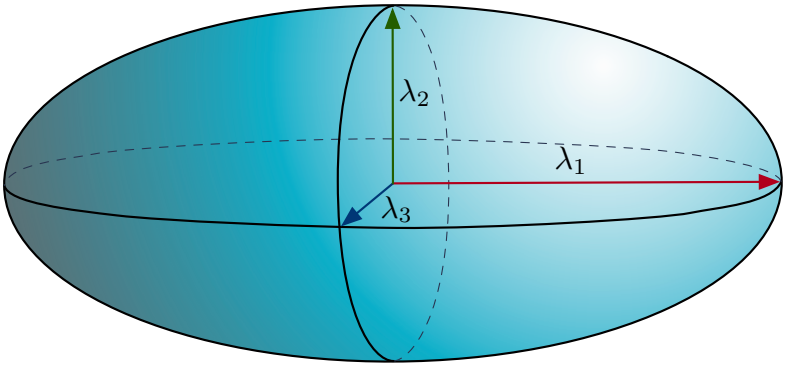


Figure 2.3: Diffusion Tensor. The diffusion tensor can advantageously be represented as an ellipsoid whose main axes are determinate by diffusivities (i.e. eigenvalues of the tensor, λ_1, λ_2, λ_3) along their respective eigenvectors ($\hat{v}_x$, $\hat{v}_y$, $\hat{v}_z$).

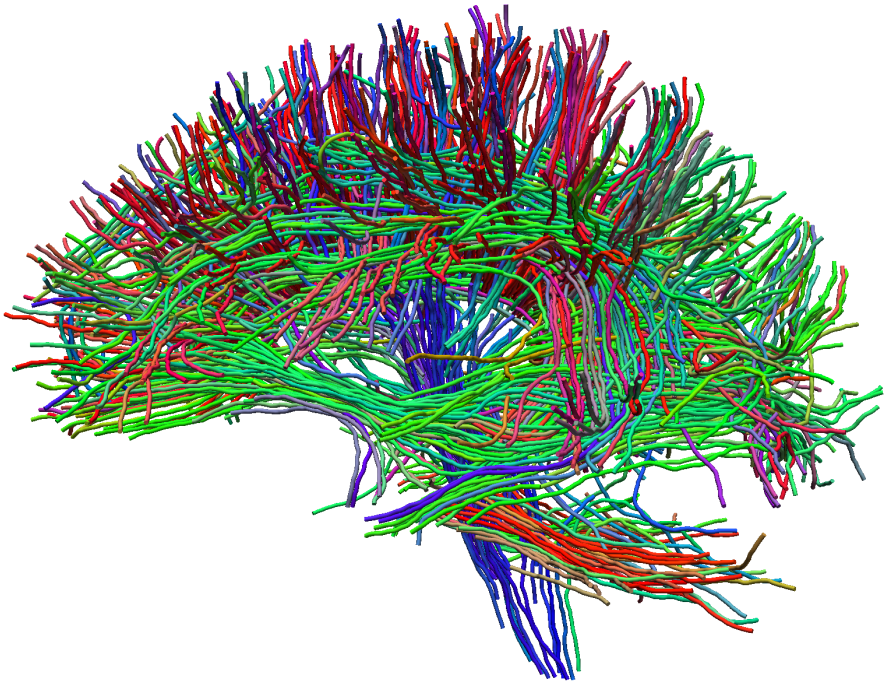


Figure 2.4: Tractography. Coloured by its main direction (vertical fibers blue, front-to-back fibers green, and side-to-side fibers red), the figure shows a tractography performed on a subject from CUBRIC test-retest dataset (see Section 1.4).

eigenvector that indicates the main orientation of the diffusion ellipsoid—tractography algorithms estimate the trajectories of white matter fiber bundles across the brain. Typically, trajectories are generated from selected voxels (seed points) that meet specific criteria, such as having a fractional anisotropy (FA) above a defined threshold. By stepping from voxel to voxel along the direction of maximum diffusivity (the eigenvector associated with the largest eigenvalue), the algorithm propagates new seed points until stopping criteria are reached. Due to the limited spatial resolution of diffusion MRI, tractography does not resolve individual axons but instead reconstructs pathways that represent coherent fiber bundles within the voxel grid.

Besides its clear contribution in elucidating brain connectivity, tractography is extremely

useful in aging studies (Davis et al., 2009) or brain plasticity (Blumenfeld-Katzir et al., 2011). Even in clinical environments, where classic dMRI still has trouble finding its applicability, tractography stands out for its usefulness in pre-surgical planning, stroke, TBI, MS, and other neurodevelopmental diseases (Yeh et al., 2021).

2.3 Beyond the tensor

2.3.1 Limitations of DTI

The need for a better characterization of diffusion in the brain aided the appearance of other frameworks that tackled the tensor's limitations. Although revolutionary, and despite still being used recurrently in most clinical environments, the tensor is defined upon some assumptions that do not always hold, especially when looking at diffusion in the complex WM structure.

First, the assumption of Gaussian distribution stemming from Brownian motion may not always be accurate. Specifically in WM, tightly packed structures, fiber arrangements, and tissue heterogeneity hinder water molecules' ability to diffuse following a Gaussian distribution. Also, applying high diffusion gradients can perturb the Brownian motion of the particles by introducing a bigger force on them, especially in areas with complex geometries.

Second, the tensor only holds macroscopic information. It is not intended to obtain microscopic information within the studied voxel. Thus, the parameters estimated with the tensor reflect the aggregated behavior of the diffusion processes within the voxel.

Third, by definition, the tensor can only distinguish one main diffusion direction. Picture the ellipsoid in Fig. 2.3; if imaging an environment with a second main diffusion direction (e.g., crossing fibers), the diffusivities λ_1 and λ_2 will become similar, resulting in a tensor that looks like an oblate. Thus, the tensor by itself is not enough to distinguish two fibers, but only an environment where diffusivity in one of the orthogonal directions to the main component is similar to that of the main component. This is known as the crossing fibers problem of DTI.

2.3.2 Q-space sampling

Following technological advances, the limitations of the tensor rapidly gave birth to refined models that could achieve a more accurate characterization of biomarkers in the WM. One of the ways to achieve a better diffusion profile estimation is to acquire more samples in the diffusion space (i.e., using more diffusion gradient directions and magnitudes), which was possible given the advancements in both hardware—higher diffusion gradients or shorter echo times, etc.—and software—efficient pulse sequence designs, or advanced reconstruction techniques, etc.

Said samples are obtained by applying diffusion gradients of different strengths and orientations. This concept prompted the appearance of the so-called **q**-space: a mathematical space that relates the diffusion gradients' strength and directions. Each data point $\mathbf{q}_i$ represents the application

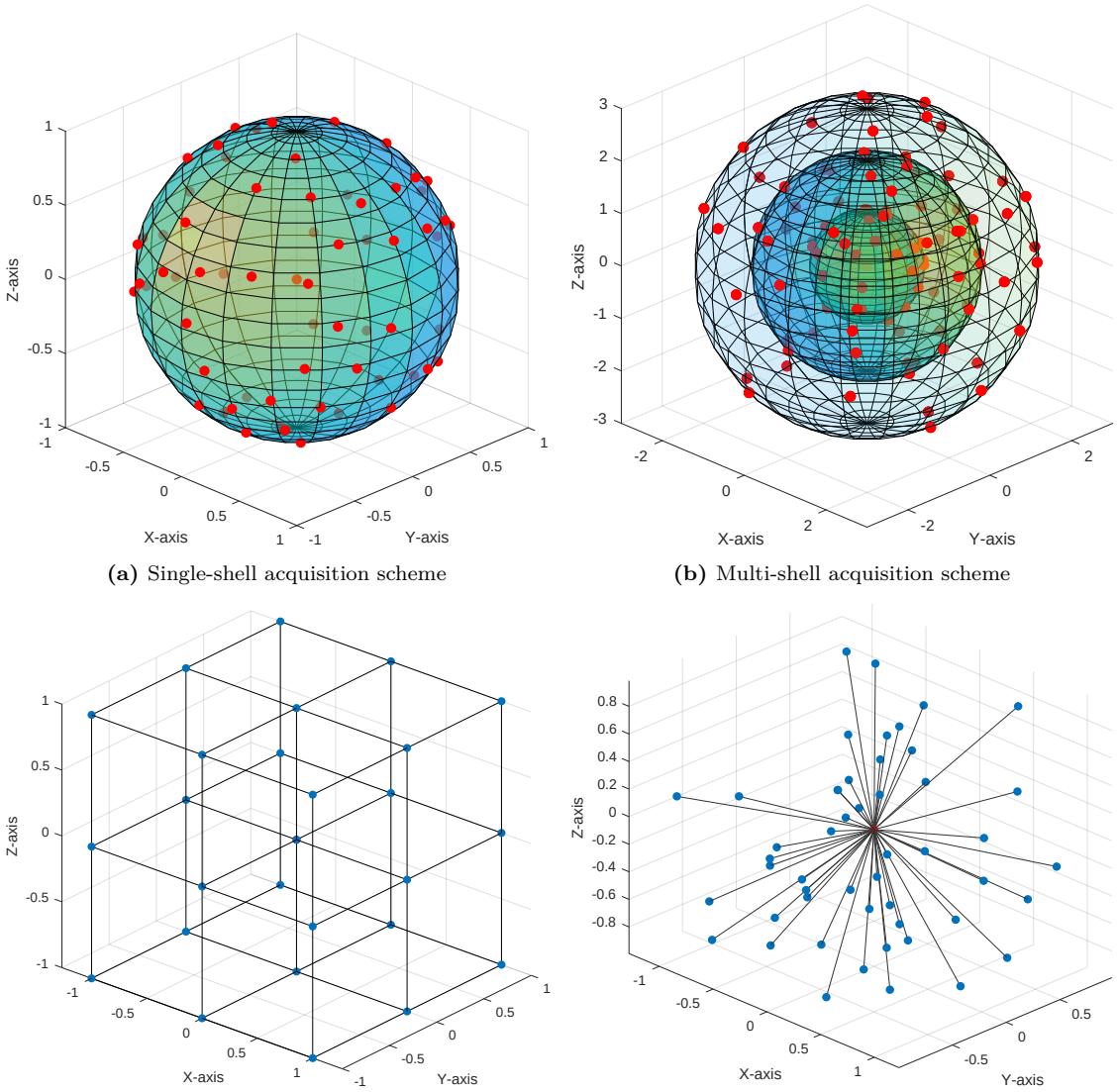
of a diffusion gradient with magnitude $\|\mathbf{q}_i\| = \gamma\delta G/2\pi \in \mathbb{R}^+$ and direction $\angle\mathbf{q}_i = \hat{\mathbf{g}}$. The more data points available, the more informed the model will be, but the more acquisition time is required. Figure 2.5 shows four representations of acquisition schemes in such space, corresponding to the most usual acquisition schemes in diffusion MRI.

While theoretically unrelated, the different **q**-space samplings appeared as a by-product of the models that introduced them. In other words, the introduction of more capable models was often associated with samplings that exploited the diffusion features said model intended to estimate. For example, the limitations of the DTI in characterizing crossing fibers were tackled by the development of High Angular Resolution Diffusion Imaging (HARDI) that leveraged spherical **q**-space samplings to increase the angular resolution. The growth of angular sampling with a fixed gradient strength gave rise to samplings that resembled a shell (samples along the surface of a sphere), which resulted in the designation of the name "single-shell" (see Fig. 2.5.a).

Single-shell Acquisitions (Fig. 2.5.a) aim at higher resolutions in the angular dimension of the diffusion profile. In turn, they seek to provide more information on the directionality of anisotropic diffusion, thus steering away from DTI problems such as crossing fibers. The position of the evenly spaced samples along the sphere's surface is usually found by either the use of regular polyhedra (Muthupallai et al., 1999) or by minimizing the electrostatic energy (EEM, Jones et al., 1999).

Multi-shell Acquisitions (Fig. 2.5.b) exploit the same improved angular resolution but also focus on characterizing the radial profile of the diffusion by sampling different diffusion strengths. The sampling of the angular domain remains denser, given that the radial profile of diffusion is assumed to be well captured by fewer parameters. In other words, it is often considered that the diffusion changes for different gradient strengths are well characterized with fewer data points than that along the angular domain, which encapsulates the structural complexity of the WM. Additionally, other limitations for denser radial samplings must be considered, such as the increase in signal attenuation for higher diffusion gradient strengths. An important aspect of multi-shell acquisitions is the required non-collinearity of samples across multiple shells (i.e., avoiding sampling the same directions in different shells). This not only entails a significant advantage in the angular sampling (it provides more angular information because more directions are sampled once accounted for all spheres) but also prevents further problems such as aliasing effects (not enough distinct angular directions), underfitting, or misinterpretation of some models, or directional bias (diffusion characterization is biased towards those points in space that are overrepresented). The implementation of these samplings is similar to that of single-shells but accounts for non-collinearity (Ye et al., 2012; Caruyer et al., 2013). However, other promising methods have been suggested, such as the weighted shells approach proposed in our recent work (Rodríguez-Galván et al., 2023).

Cartesian grid Acquisitions (Fig. 2.5.c). Rather than sampling **q**-space according to a biophysical rationale (angular vs. radial domain), this acquisition scheme seeks to fulfill a pragmatic approach for signal processing purposes. It aims at exploiting the relationship between the signal attenuation and the diffusion propagator, given by a Fourier transform (refer to Section 2.4.2 for a detailed explanation) while also preventing aliasing effects by



(a) Single-shell acquisition scheme

(b) Multi-shell acquisition scheme

(c) Cartesian grid acquisition scheme

(d) Compressed-sensing acquisition scheme

Figure 2.5: Q-space sampling schemes. Four sampling schemes are shown. Single-shell (a) and multi-shell (b) protocol acquisitions are the most common; some others rely on signal processing methods to estimate the diffusion processes by exploiting such acquisitions (c and d)

adhering to the Nyquist-Shannon sampling theorem. Its main limitation, however, is that a fully sampled cartesian grid includes many more samples than other sampling techniques, which in turn implies long acquisition protocols and, thus, longer scanning times.

Sparse Acquisitions (Fig. 2.5.d) refers to those acquisitions that randomly sample the **q**-space. Similarly to the previous case, the idea behind this acquisition seeks to be exploited by a signal processing technique called Compressed Sensing (CS, Donoho, 2006), by which a signal can be reconstructed from a reduced set of random acquisitions. In assuming signal sparsity, an optimizer tries to find the sparsest solution that best matches the measurements.

2.3.3 Mathematical Tools

Diffusion MRI is, in its very nature, a flawed system: from the noisy and distorted acquisition to the reconstruction of the attenuation signal, the ill-posed fitting of the models, etc. Also, diffusion-weighted signals are overwhelmingly large, so their computational analysis (in fitting, for example) always requires powerful computational resources so that applications do not become increasingly time-expensive in already time-constrained settings. Mathematics offers a way of dealing with this latter problem, providing tools to describe the most important signal features up to a desired degree of accuracy.

Spherical Harmonics

The pragmatic understanding of spherical diffusion acquisitions inside a voxel, captured by *shelled* samplings, favors using mathematical tools that exploit such behavior. Spherical Harmonics (SH) are mathematical objects that offer a way of representing functions over the surface of a sphere, allowing a decomposition of the spatial patterns on simpler harmonic components. More technically, SH serve as basis functions for square-integrable functions defined over the surface of a sphere; that is, functions that have finite energy when integrated over a sphere.

SH are solutions to the Laplace equation in spherical coordinates. They are defined as:

$$Y_l^m(\theta,\varphi) = N_l^m P_l^m(\cos\theta) e^{jm\varphi}, \tag{2.9}$$

where l is the degree ($l \in \mathbb{R}^+$), m the order ($-l \leq m \leq l$), $P_l^m(\cos\theta)$ is the associated Legendre polynomial, which is dependent on the zenithal domain θ. N_l^m is a normalization factor to ensure orthonormality in the basis functions:

$$N_l^m = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}. \tag{2.10}$$

Figure 2.6 shows the basis functions derived from SH. Each degree l is represented as a band of orders m from $-l \leq m \leq l$. Therefore, for a given degree L, the number of coefficients to estimate is $2L+1$. As other basis functions, the SH bases offer a set of useful properties:

Orthonormality among basis functions on the unit sphere, such that:

$$\int_0^{2\pi}\int_0^{\pi} Y_l^m(\theta,\varphi) Y_{l'}^{m'\star}(\theta,\varphi)\, d\theta\, d\varphi = \delta_{ll'}\delta_{mm'}, \tag{2.11}$$

where $Y_{l'}^{m'\star}(\theta,\varphi)$ is the complex conjugate of $Y_l^m(\theta,\varphi)$, and $\delta_{ll'}$ and $\delta_{mm'}$ are Kronecker's deltas that equal to 1 when $l = l'$ or $m = m'$ and 0 otherwise. This means that any given basis function is orthonormal to all the other basis functions that do not have the same degree and order.

Completness, which implies that any square-integrable function $g(\theta,\varphi)$ on the surface of a sphere

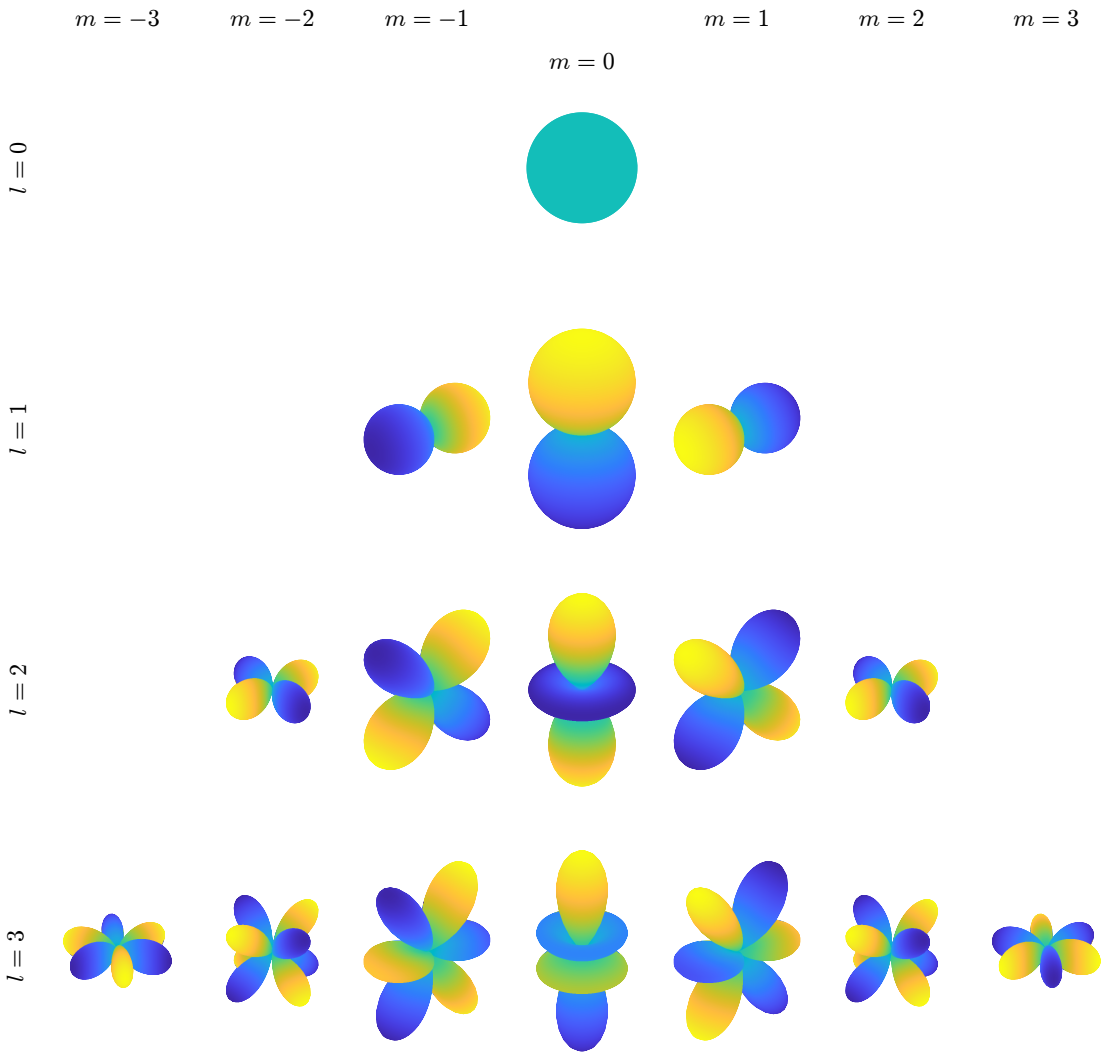


Figure 2.6: Spherical Harmonics. Visual representation of the different basis functions of spherical harmonics, for different degree $l = 1, 2, 3$ and order $-l \leq m \leq l$. The higher the degree of SH used, the more contributions of these basis functions will be used, leading to more accurate representations of the function.

can be expanded in terms of the basis functions:

$$g(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_l^m Y_l^m(\theta, \varphi), \tag{2.12}$$

where c_l^m is the coefficient corresponding to that basis function. The coefficients are defined as:

$$c_l^m = \int_{\mathcal{S}} g(\theta, \varphi) Y_l^{m*}(\theta, \varphi) dS. \tag{2.13}$$

While theoretically feasible, real-world settings do not always allow a full characterization of a function in terms of SH. That is, Eq. (2.12) is to be computed up to a certain degree $l = L_{\max}$,

which will thus result in a truncated (i.e., bandwidth-limited) version of the spherical signal:

$$g(\theta,\varphi) \simeq \sum_{l=0}^{L_{\max}} \sum_{m=-l}^{l} c_l^m Y_l^m(\theta,\varphi). \tag{2.14}$$

The choosing of $L_{\max}$ might be determined by the computational resources or time available to carry out the projection. However, it is also influenced by the diffusion sampling scheme used to acquire the signal itself: Using big enough $L_{\max}$ might try to capture high-frequency components that have been previously cropped by a poor **q**-space sampling.

Rotational invariance. SH offer great flexibility for functions that are rotated. This is illustrated in Fig. 2.6. Suppose a 3D rotation is applied to a signal represented on this basis. In that case, the aggregated power in each order band l remains unchanged, albeit subject to different coefficients for each basis across $m = 2l + 1$ orders. In other words, while the coefficients themselves are not invariant under rotation, certain derived quantities from them are. A common example of a rotational invariant feature is the 0-th order of the SH projection (i.e., $l = 0$), proportional to the Spherical Means (SM)—the spherical average of the signal—which clearly does not change with data rotations.

Aiming at a more robust characterization of metrics and comparisons across subjects, rotational invariant features are often used in dMRI to decouple tissue microstructure from fiber orientation distributions (Novikov et al., 2018a; Zucchelli et al., 2020). This way, rotations of the diffusion encoding schemes or the subject's head become irrelevant since they pose no effect on the scalar quantities. Additionally, using rotational invariant features entails a projection onto lower-dimensional spaces, streamlining the fitting problems in models that use them. This will be seen, for example, in the Standard Model of Diffusion Imaging (see Section 2.4.4).

Particularly in dMRI, certain biophysical assumptions about the signal can be exploited to simplify the projections onto the SH. For example, if antipodal symmetry is assumed in the scanned voxel, i.e. $E(-\mathbf{q}) = E(\mathbf{q})$, only even degrees $l = 2, 4, 6, \ldots$ are needed. Also, if the reconstructed signal is real, the real spherical harmonics (rSH) are used:

$$Y_l^m(\theta,\varphi) = \begin{cases} \sqrt{2}\, N_l^m\, P_l^m(\cos\theta)\cos(m\varphi), & m > 0 \\ N_l^0\, P_l^0(\cos\theta), & m = 0 \\ \sqrt{2}\, N_l^{|m|}\, P_l^{|m|}(\cos\theta)\sin(|m|\varphi), & m < 0 \end{cases} \tag{2.15}$$

2.4 Phenomenological vs. Biophysical Modeling

The ability to describe the diffusion processes—or the microstructure landscape in which they take place—from the attenuation signal is not necessarily straightforward or, at least, as intuitive as in DTI. Given that diffusion-weighted images are complex data sets rich in hidden information, a usual approach involves projecting them onto lower-dimensional spaces so that inferences—

such as biomarkers—can be drawn.

In a nutshell, there are two main ways of lowering the dimensionality of diffusion-weighted signals onto a mathematical representation of diffusion. One relies on the modeling of the diffusion phenomena without any biophysical assumptions. The other relies on modeling the biophysical landscape in which diffusion occurs. Novikov et al. (2018a), possibly aiming to unify research perspectives, emphasized the importance of clarifying this distinction and proposed referring to them as signal representations and models, respectively. In this work, however, we have chosen to refer to these two philosophical perspectives as phenomenological[1] modeling and biophysical modeling. This reflects an effort to regard both frameworks as equally valid, particularly in clinical research.

In Novikov's work, phenomenological modeling is introduced as an alternative name but quickly set aside, favoring "signal representation". While the term "signal representations" is technically accurate, it reflects on the mathematical aspect of these models. Modeling, however, seems to bias the reader towards the thought that metrics derived from them are *more* linked to the actual diffusion process taking place. These assumptions are wrong: Biophysical and phenomenological modeling aim to describe diffusion processes by taking a look at different aspects of such diffusion processes.

Phenomenological modeling involves projecting the signal onto a basis usually chosen by balancing simplicity, assumptions, and available measurements. That is, for a model to fit adequately onto measured data that is scarce and poor in quality, the more biophysical assumptions will have to be made, or the more simplistic the model will have to be. Phenomenological modeling is undoubtedly linked to mathematics, but mathematics is merely a means to an end: It models some properties of the diffusion processes in a voxel without assuming a biophysical landscape. While this is technically true, some phenomenological models do rely on biophysical assumptions in order to simplify the signal processing techniques employed.

Clinically, phenomenological modeling relies on the acceptance of a baseline set of processes against which current observations are compared. As such, one cannot directly link phenomenological metrics to biophysical conclusions, since these metrics lack inherent biophysical meaning. However, given a specific set of phenomenological outcomes, it is possible to infer certain underlying biophysical mechanisms. For example, if diffusion appears isotropic, one might conclude that the underlying microstructure is unlikely to consist of coherently aligned, myelinated fibers, even though the model itself does not explicitly capture that level of structural detail.

Biophysical modeling is based on the projection of a signal onto a model that encapsulates some biophysical meaning. Given that the model gives absolute interpretations rather than perceptual, no baseline is needed to make inferences once the model has been validated. This is because a direct link exists between the metrics derived from biophysical modeling and the biophysical landscape where the diffusion takes place.

This naturally raises an important question: why model phenomena at all, when one could

[1] Drawing from the philosophy of phenomenology, which focuses on describing phenomena without explaining the underlying cause.

instead model biophysics and aim to reconstruct the underlying microstructural landscape from the diffusion signal? The answer lies in the inherent limitations of both the data and our modeling assumptions. Biophysical modeling aims at constructing a plausible environment that could result in the observed diffusion signal. Its strength lies in the interpretability of the estimated parameters. However, the weak link in this approach is the inversion itself: estimating complex microstructural parameters from a noisy and indirect signal is inherently ill-posed and model-dependent. Phenomenological modeling, by contrast, focuses on drawing conclusions from phenomena already captured by the diffusion signal. Yet its weakness lies in interpretation: while it models the phenomena effectively, it remains agnostic about their biological significance—leaving a gap between observed signal patterns and their relevance to research or clinical questions.

2.4.1 Incompatibility

It is important to clarify the incompatibilities of these two viewpoints: Modeling the signal regarding the diffusion phenomena necessarily implies the disregard of biophysical grounding. The flexibility of phenomenological models arises from the lack of assumptions about the underlying tissue but directly affects the link to microstructural landscape interpretability. On the other hand, biophysical models aim at interpretability but are constrained to the validity of their assumptions.

Since the goal is different, the parameters used to describe phenomena and biophysics are challenging to reconcile in a single framework: Phenomenological frameworks describe diffusion properties mathematically (e.g., using tensors, Kurtosis, spherical harmonics, etc.) or by diffusion descriptors (e.g. mean squared displacement, propagator anisotropy, return probabilities), whereas biophysical frameworks are described through tissue parameters (e.g., neurite density, axon diameter, etc.).

It should be noted, however, that beyond their conceptual differences, each modeling perspective also faces technical and mathematical constraints. For example, assumptions about the signal's decay behavior and the feasibility of its Fourier relationship to real-space propagators impose further practical limitations. These aspects are discussed in more detail in Section 2.4.2.

Moreover, regardless of their primary goal, some signal representation techniques are adopted by biophysical models to ease the fitting process. Similarly, some biophysical assumptions and priors are exploited and translated into the modeling of phenomena to allow more straightforward fitting problems.

2.4.2 Advanced Diffusion Metrics

Metrics are the essential layer that allows for interpretability of diffusion models. Mathematically, it entails a projection onto a lower-dimensional space, and they often take form of a 3D scalar map. Such projections are intended to provide a certain aspect of information about the diffusion within the imaged voxel.

Generalizing Phenomena: The Ensemble Average Propagator

Moving away from the simpler models such as the DT, it becomes important to define a concept that will be able to generalize the diffusion phenomena. Within phenomenological models, the Ensemble Average Propagator (EAP) is a mathematical object that conceptualizes the set of possible locations a water molecule can be displaced as diffusion time passes (Wedeen et al., 2005; Özarslan et al., 2006). In other words, the EAP is just a generalization of the diffusion profile so more accurate estimations might be made. It is important to note that the EAP aims at describing macroscopic diffusion processes, that is the reason behind the words "ensemble average", as it is assumed to portray the aggregate behavior of the individual diffusion processes. As a tool, it is built upon the set of assumptions the model is based on. That is, each phenomenological model defines their own propagator by extending the phenomena assumptions upon which the model is built.

The EAP equation represents the main relationship that phenomenological models exploit. The application of diffusion-sensitizing gradients yields the measured attenuation signal $E(\mathbf{q})$. Under the narrow pulse approximation (i.e., $\delta \to 0$, which practically translates to $\delta \ll \Delta$) this signal is linked to the EAP $P(\mathbf{R})$, the displacement probability density function in real ($\mathbf{R}$)-space. Then, for a diffusion wavevector $\mathbf{q}$, the EAP results in (Callaghan, 1993):

$$P(\mathbf{R}) = \int_{\mathcal{R}^3} E(\mathbf{q}) e^{-2\pi j \mathbf{q}\mathbf{R}} d\mathbf{q} = \mathcal{F}\{|E(\mathbf{q})|\}(\mathbf{R}). \tag{2.16}$$

For this relationship to hold, the signal $E(\mathbf{q})$ must be smooth enough and decay at infinity, which ensures the Fourier transform is well-behaved. In practice, smoothness is often imposed as a prior to regularize the inversion and avoid spurious features caused by noise or modeling errors. The assumption of rapid decay in $P(\mathbf{R})$ is also crticial, as it underlies the feasibility of reconstructing it from a limited $\mathbf{q}$-space samplings. However, such assumption does not always hold, particularly in cases of restricted and hindered diffusion. Consequently, phenomenological models relying on limited $\mathbf{q}$-space sampling may struggle to accurately reconstruct $P(\mathbf{R})$, especially when the underlying diffusion deviates significantly from free or Gaussian behavior.

In DTI, the EAP is characterized by a multi-variate normal distribution, whose second-order moment fully determines the shape of the propagator: the ellipsoid. Diffusion Kurtosis Imaging (DKI, Jensen et al., 2005) uses the fourth-order moment to assess deviations from such Gaussian distribution. By allowing more complex formulations of the EAP, more accurate descriptions of the diffusion profile can be made. However, more complex profiles might require more data (or more assumptions) in order for the fitting problem to be well defined. Mathematically, said complexity is captured in independent parameters that vary according to the system's behavior: the degrees of freedom (DOF).

The accuracy up to which diffusion phenomena is able to represent the diffusion profile from the EAP depends on the number of DOF the encompassing model offers. Higher DOF implies more flexibility of the model to adapt to certain signal characteristics. However, a higher number of DOF also requires more—and often higher-quality—data, as the model will be more difficult to fit accurately.

Given that clinical environments are still not ready for lengthy acquisition protocols, there is a need for developing methods on all ranges of degrees of freedom, exploiting all kinds of acquisition sequences: In time-restricted settings (such as clinical settings), shorter acquisitions might be used, which call for simpler EAP definitions (or more biophysical assumptions). In settings where time is not strictly restricted (e.g., some research settings), longer acquisition protocols might be used, which in turn permits the characterization of more complex EAP (or the avoidance of some assumptions).

The estimation of the EAP entails several issues that might compromise its accuracy. First, as a derivative of the phenomenological model, the EAP is represented as a superposition of functions from a basis in which the EAP is presumed to be sparse (i.e. can be approximated with a few non-zero coefficients on such basis). Second, its estimation often relies in non-linear problems whose stability might not be guaranteed, and whose convergence might be conditioned by the computational time restrictions. Third, some EAP definitions might require the acquisition of high diffusion-weighted gradients that negatively impact the SNR of the samples. Conversely, the limitation to a maximum b-value restricts the EAP interpretability to only the lower frequencies. Fourth, the Fourier relationship between EAP and diffusion signal might not always hold due to the non-linearity, diffraction or non-negligible diffusion effects during the δ (diffusion time duration).

Directionality descriptors: The ODF

The Orientation Distribution Function (ODF) is an orientation descriptor that reflects the probability of water molecules diffusing in different directions. In other words, it characterizes the shape of diffusion rather than directly describing the axons or microstructural features that give rise to such patterns. ODFs are typically derived from advanced diffusion models and form the foundation for tractography by indicating likely fiber orientations. To ease estimation and reduce computational complexity, ODFs are commonly represented over SH basis, which provides a compact and analytically tractable form. While not scalar metrics themselves, ODFs offer rich directional information that complements quantitative measures and enhances our understanding of brain microstructure, surpassing the limitations of simpler models like the diffusion tensor. Figure 2.7 shows a comparison between an ODF and the oblate shape resulting from its estimation using the tensor model.

One common definition of the ODF is as a radial projection of the EAP along a given direction $\mathbf{r}$, expressed as:

$$\mathbf{\Phi}(\mathbf{r}) = \int_0^\infty P(R\mathbf{r})R^2 dR \tag{2.17}$$

where $P(\mathbf{R})$ is the probability density function of net water molecule displacements and $R\mathbf{r}$ denotes points along the direction $\mathbf{r}$ scaled by the radial distance R.

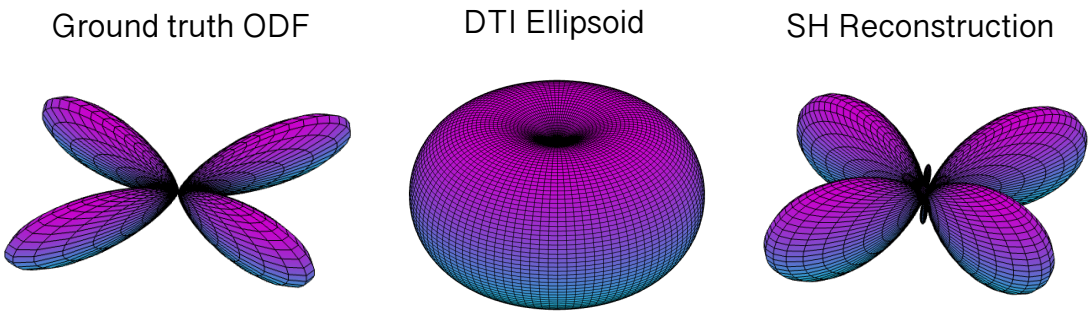


Figure 2.7: The fiber Orientation Distribution Function. Comparison between the ground truth ODF (left), the DTI ellipsoid (middle), and the SH reconstruction (right, with SH order $L = 6$). Note how DTI is unable to resolve crossing fibers, while SH reconstruction provides additional information for a more accurate ODF estimation.

Phenomenological metrics

In what follows, common metrics for diffusion phenomena, together with the state-of-the-art diffusion metrics, are presented. All of them are derived from specific EAP formulations, and as such, may require different amounts or qualities of data to accurately capture the aspects of diffusion they aim to represent.

Mean Diffusivity or MD is the mean over all the tensor's eigenvalues. It quantifies the amount of diffusion, regardless of its anisotropy.

$$\mathrm{MD} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{2.18}$$

MD not only is useful in differentiating WM from GM, but also has been associated with changes in edema and inflammation, neuron body swelling in the hyperacute/acute phase of stroke (Osa García et al., 2022), aging (Madden et al., 2009, 2012), AD's episodic memory deficits (Bosch et al., 2012; Sexton et al., 2010).

Fractional Anisotropy or FA is the most famous of diffusion metrics. It is derived from the tensor and indicates how anisotropic a diffusion process is by the relationship among the eigenvalues: 0 means isotropic, 1 means completely anisotropic. It is defined as:

$$\mathrm{FA} = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - \mathrm{MD})^2 + (\lambda_2 - \mathrm{MD})^2 + (\lambda_3 - \mathrm{MD})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{2.19}$$

Lower FA often reflects damage to the axons' myelin sheath, enlarged axonal diameter, reduced axonal packing density or increased membrane permeability (Farag and Kayser, 2017). Studies reveal that FA is sensitive to changes in WM integrity derived from schizophrenic disorders (Clark et al., 2011), vasogenic edema in stroke (Osa García et al., 2022), aging (Sullivan and Pfefferbaum, 2006; Madden et al., 2009, 2012), AD's episodic memory deficits (Bosch et al., 2012; Sexton et al., 2010) even in asymptomatic individuals with a family history of AD (Bendlin et al., 2010).

The tensor representation captures diffusion behavior through three eigenvalues (and corresponding eigenvectors), inherently limiting the richness of the information that can be extracted. Metrics like FA and MD, though not the only ones derivable from the tensor, remain among the most widely used. More advanced models (i.e., those that incorporate greater flexibility by introducing additional DOF) enable the estimation of a broader set of parameters, thereby uncovering aspects of the diffusion signal that remain hidden in the diffusion signal.

Return Probabilities, also called probabilities of zero displacement (Özarslan et al., 2013), are indices that quantify features from the diffusion processes captured by the EAP. More specifically, return to origin (RTOP), return to axis (RTAP) and return to plane (RTPP) probabilities enable the quantification of diffusion restrictions without the need of biophysical models. RTOP quantifies the probability of a water molecule to return to its original location, RTAP quantifies the probability of a water molecule to return to its original place along an axis. Finally, RTPP quantifies the probability of a water molecule to return to its original place within a specific plane. All of them are closely related to a diffusion feature within the environment: RTOP aims at quantifying the degree of restrictivity in the medium, RTAP the degree of diffusivity along a fiber, and RTPP the degree of diffusivity across the plane perpendicular to a fiber. They are defined over the propagator, as well as over the diffusion signal, as follows:

$$\text{RTOP} \triangleq P(0) = \int_{\mathbb{R}^3} E(\mathbf{q}) d\mathbf{q} \tag{2.20}$$

$$\text{RTAP} \triangleq \int_{\mathbb{R}} P(R\mathbf{r}_{\parallel} | \mathbf{r}_{\perp} = 0) dR = \int_{\mathbb{R}^2} E(\mathbf{q}_{\perp}) d\mathbf{q}_{\perp} \tag{2.21}$$

$$\text{RTPP} \triangleq \int_{\mathbb{R}} P(R\mathbf{r}_{\perp} | \mathbf{r}_{\parallel} = 0) dR = \int_{\mathbb{R}} E(q_{\parallel}) dq_{\parallel} \tag{2.22}$$

Assuming a Gaussian diffusion propagator, Aja-Fernández et al. (2020) derives the return probabilities from the tensor's three eigenvalues, yielding:

$$\text{RTOP} = \frac{1}{(4\pi\tau)^{3/2}} (\lambda_1 \cdot \lambda_2 \cdot \lambda_3)^{-1/2} = \text{RTAP} \cdot \text{RTPP} \tag{2.23}$$

$$\text{RTAP} = \frac{1}{\sqrt{4\pi\tau}} (\lambda_1)^{-1/2} \tag{2.24}$$

$$\text{RTPP} = \frac{1}{4\pi\tau} (\lambda_2 \cdot \lambda_3)^{-1/2} \tag{2.25}$$

These indices are primarily sensitive to the underlying microstructural environment. In particular, RTOP is largely independent of macroscopic fiber orientation, making it a more robust marker in complex fiber regions compared to measures like FA, which drop significantly in crossing-fiber voxels. However, RTAP and RTPP are computed along specific directions and can therefore reflect changes due to fiber dispersion or orientation.

Microstructure from Return Probabilities

Consider for example, RTOP, or the Return to origin probability. This quantity refers to the probability that a water molecule returns to its origin—i.e. net displacement equals 0. RTOP is conceptually linked to entropy, which in diffusion can be thought of as how constrained is diffusion. Higher RTOP values—meaning greater probability of return to origin—are associated with increased microstructural order, e.g., tightly packed fibers and small restricted compartments.

Under certain assumptions, such that diffusion gradients are infinitesimally short, Özarslan et al. (2009) prove the relationship of the metrics to the microstructural environment of the voxel analysed:

$$\text{RTOP} = \langle V \rangle^{-1} \tag{2.26}$$

$$\text{RTAP} = \langle A \rangle^{-1} \tag{2.27}$$

$$\text{RTPP} = \langle L \rangle^{-1} \tag{2.28}$$

where $\langle V \rangle$ is the mean pore volume, $\langle A \rangle$ the mean cross-sectional area, and $\langle L \rangle$ is the mean length of the cilynders. RTAP for example can thus linked to mean axon diameter $\langle D \rangle = 2\sqrt{1/(\text{RTAP}\pi)}$.

Mean Squared Displacement or MSD is a measure that quantifies the (mean squared) displacement of water molecules in biological tissues during the diffusion experiment. It is linked to the amount of diffusion that is facilitated or restricted by biological tissues, thus being closely related to MD (Wu et al., 2008; Ning et al., 2015), but more sensitive to restricted diffusion (Wang et al., 2021a; Ning et al., 2015).

$$\text{MSD} = \int_{\mathbb{R}^3} \text{P}(\mathbf{R})\text{R}^2 \text{d}\mathbf{R} = -\frac{1}{4\pi^2}\Delta \text{E}(\mathbf{q})|_{\mathbf{q}=0} \tag{2.29}$$

QMSD refers to the **q**-space analogy of the MSD metric. For Gaussian propagators, QMSD and MSD have very similar contrast in tissues. However, when deviating from Gaussianity in propagators, QMSD may offer a distinct contrast profile in the WM (McClymont et al., 2016; Hori et al., 2022). QMSD provides insights about the displacement of molecules across different spatial scales (Assaf et al., 2000; Caporale et al., 2022).

$$\text{QMSD} = \int_{\mathbb{R}^2} \text{E}(\mathbf{q})\text{q}^2 \text{d}\mathbf{q} = -\frac{1}{4\pi^2}\Delta \text{P}(\mathbf{R})|_{\mathbf{q}=0} \tag{2.30}$$

Q-space Inverse Variance or QIV is defined as the reciprocal of the QMSD. It allows identifying tissue regions with different diffusion profiles, such as tumors, demyelination or lesions (Fick et al., 2016b; Wu et al., 2008).

$$\text{QIV} = \text{QMSD}^{-1} \tag{2.31}$$

Propagator Anisotropy or PA measures how similar the propagator is to its isotropic version. Developed by Özarslan et al. (2013), it represents a generalization of the FA but, similarly to

the return probabilities, is independent from fiber architecture. It is defined as:

$$\mathrm{PA} = \gamma\Big(\sin\theta_{\mathrm{EO}}, \epsilon\Big), \tag{2.32}$$

which, by means of the Parseval theorem, can be expressed as the distance from $E(\mathbf{q})$ to its closest isotropic counterpart $O(\mathbf{q})$. The angular distance between the two vectors is defined as:

$$\cos\theta_{EO} = \frac{\langle E(\mathbf{q}), O(\mathbf{q})\rangle}{\|E(\mathbf{q})\|\|O(\mathbf{q})\|}, \tag{2.33}$$

and $\gamma(t, \epsilon)$ is a contrast enhancement function to project the PA onto a range of $[0, 1]$.

$$\gamma(t, \epsilon) = \frac{t^{3\epsilon}}{1 - 3t^{\epsilon} + 3t^{2\epsilon}}, \tag{2.34}$$

with $\epsilon = 0.4$.

PA has shown great potential in clinical studies, from the assessment of longitudinal changes within subjects (Avram et al., 2016), to characterization of impairment in cognition after TBI (Kraus et al., 2007) or social cognition in autism (Barnea-Goraly et al., 2004), and even on the age-dependent neuronal demise in transgenic Alzheimer's rats (Fick et al., 2017).

Non-Gaussianity or NG (Özarslan et al., 2013) measures the similarity between the propagator and its Gaussian version $G(\mathbf{q})$. Thus, NG stands out as a generalizability of the diffusion kurtosis. Similarly to PA, it is defined as:

$$\mathrm{NG} = \sin\theta_{\mathrm{EO}} = \sqrt{1 - \left(\frac{\langle \mathrm{E}(\mathbf{q}), \mathrm{G}(\mathbf{q})\rangle}{\|\mathrm{E}(\mathbf{q})\|\|\mathrm{G}(\mathbf{q})\|}\right)^2}, \tag{2.35}$$

where the trigonometric identity has been used in the second equality.

Even though its clinical applicability has not been as extensively assessed as PA, NG has been shown promising results in differentiating grade II from grade III and IV gliomas (Wang et al., 2021a) important in preosurgical evaluation of tumour pathological grading, in axonal loss and demyelination (Song et al., 2005), as well as head and neck cancer (Jansen et al., 2010; Yuan et al., 2014).

Biophysical metrics

Biophysical metrics describe the environment in which diffusion happens. Given the limitation of resolution in dMRI (i.e., around the mm), together with the size of microstructural features in the brain (i.e., around the μm), it is not possible to obtain direct measurements of individual microstructural elements.

For this reason, and similar to other metrics, biophysical metrics aim at depicting the overall contributions of all microstructural components to the diffusion phenomena, characterizing a voxel-size average estimation of the microstructure.

Intra-cellular signal fraction f represents the signal fraction that is assumed to come from

molecules diffusing within cellular bodies. Within the brain, abnormal values of f are related to more (or less) intra-cellular spaces which, for example, have been linked to axonal loss (Fieremans et al., 2012).

Longitudinal intra-cellular diffusivity D_a depicts the diffusivity that is permitted along the axons or other intra-cellular bodies. In the WM, it has been linked to the abnormal diffusivity along the axons, which has been proven to be caused by, for example, swellings or enlargements along the axon shaft (i.e., beading, Budde and Frank, 2010), as well as by inflammation and oedema (Unterberg et al., 2004).

Extra-cellular parallel diffusivity $D_e^{\parallel}$ denotes the diffusivity of molecules parallel to the cellular bodies (i.e., parallel to the axon shafts, in WM). It has been proven to be related to inflammation and edema (Unterberg et al., 2004).

Extra-cellular perpendicular diffusivity $D_e^{\perp}$ refers to the diffusivity of molecules perpendicular to the cellular membranes. In the WM, it portrays the transverse diffusivity across myelinated axons, which has been proven to be affected by processes such as demyelination (Fieremans et al., 2012; Jelescu et al., 2016b) and inflammation and edema (Unterberg et al., 2004).

2.4.3 Phenomenological Models

As already introduced, phenomenological models aim at representing diffusion processes by means of the simpler representation of diffusion phenomena. By projecting the signal onto a frame, and estimating mathematical objects like the EAP, features from the diffusion can be elucidated.

Contrary to popular belief, neither DTI nor DKI are biophysical models: They do not assume Gaussian, or a kurtosis-like deviation from Gaussian diffusion. Rather they just capture the lower orders of the Taylor expansion of the logarithmic diffusion signal (Novikov et al., 2018a). While DTI captures only the first cumulant (second order moment), DKI adds the deviation from Gaussianity—i.e., kurtosis—from the second cumulant (fourth order moment).

$$\ln(S(b)) = \ln(bS(0)) - bD + b^2C - \dots \tag{2.36}$$

In what follows, state-of-the-art phenomenological models are listed and sorted by their DOF number, as shown in Fig. 2.8. These models range from simpler to more complex representations. The simplest models require less data and computational resources and thus yield fast and efficient metrics that may lack generalizability or accuracy. In contrast, the most complex models aim to describe intricate EAP interpretations, which necessitate a larger dataset for reliable results, and hence the use of more computational resources.

Apart from the classification shown in Fig. 2.8, phenomenological models typically fall into one of the following categories: (1) HARDI-based models, (2) spherical deconvolution models, (3) Fourier-based models, and (4) basis function projection models. HARDI-based models focus on acquiring and analyzing diffusion signals sampled on a single shell, emphasizing angular resolution. Spherical deconvolution models leverage the convolutional relationship between the diffusion

signal and the ODF, assuming a single-fiber response function. Fourier-based models rely on the Fourier relationship between the diffusion signal and the EAP (see Eq. 2.16), either by directly acquiring, or interpolating into, a densely populated Cartesian sampling of **q**-space. Finally, basis function models project the signal onto a predefined set of analytical basis functions, which facilitate estimation, manipulation, and interpretation of the signal and its derived quantities. In the context of this work, the selected phenomenological models are representative of each of these categories: AMURA falls under HARDI-based models, MiSFIT represents spherical deconvolution models, MAP-MRI belongs to basis function projection models, and DSI exemplifies Fourier-based models.

The description of each model is organized according to the following domains:

1. **Diffusion model:** The main mathematical description of the model for the interpretation of the signal, together with the assumptions and requirements to be met.
2. **Numerical Implementation:** The mathematical and computational implementation of the theorized model from the acquired data.
3. **Derivation of common biomarkers:** The estimation of the most common metrics from the implemented model.
4. **Advantages, limitations, and applicability:** A brief explanation on the advantages and limitations of the model, together with the main applications of the model in the literature, and its relevance.

AMURA

Diffusion model. The accurate estimation of the EAP requires a full sampling of the **q**-space so that the Fourier transform from the attenuation signal to the propagator (Eq. (2.16)) is well defined on all its domain. While denser acquisitions are possible, they critically increase the scanning time in settings that might already be time-restricted, inducing patient discomfort, as well as prompting the appearance of artifacts due to subject movement. Considering this, the Apparent Measures Using Reduced Acquisitions (AMURA; Aja-Fernández et al., 2020) aims at computing scalar indices similar to those of the state-of-the-art EAP-based models, but taking advantage of modest acquisitions that are available in all said environments. Thus, AMURA does not seek to accurately characterize the full EAP, but rather to obtain a mathematical description of the propagator's moments that have been deemed to be sensitive to diffusion processes and pathologies.

By surrendering to single-shell acquisitions, the mono-exponential decay assumption has to be adopted. Similarly, by assuming that diffusivity on the radial domain is independent from the anisotropy, AMURA arrives at closed-form expressions of scalar maps that do not require extensive sampling of the **q**-space.

AMURA begins with a generalization of the DT: the Diffusion Orientation Transform (DOT), introduced earlier by Özarslan et al. (2005) and Özarslan et al. (2006). The DOT describes

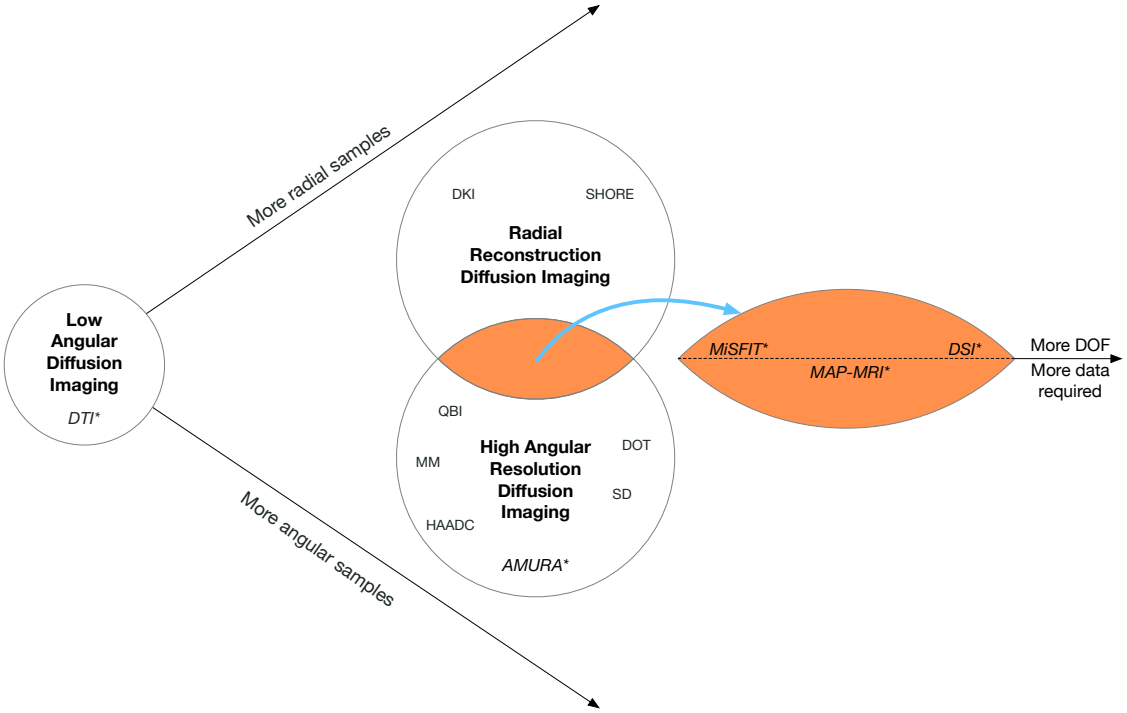


Figure 2.8: Classification of phenomenological models. Different implementations of the diffusion propagator formalism can be classified by the number of diffusion-weighted samples they require. Diffusion Tensor Imaging (DTI) is suitable for low angular resolution acquisitions and does not aim to reconstruct the radial structure of the diffusion propagator. With the advent of faster acquisition sequences, more angular samples became available, giving rise to High Angular Resolution Diffusion Imaging (HARDI; Tuch et al., 1999) and associated models such as High Angular ADC (HAADC; Frank, 2002; Alexander et al., 2002; Tuch, 2002), Q-Ball Imaging (QBI; Tuch, 2004), Spherical Deconvolution (SD; Tournier et al., 2004), and Diffusion Orientation Transform (DOT; Özarslan et al., 2005; Özarslan et al., 2006). Independently, the use of stronger diffusion gradients and the growing interest in the radial structure of the propagator led to models such as Diffusion Kurtosis Imaging (DKI; Jensen et al., 2005) and SHORE (Simple Harmonic Oscillator-based Reconstruction and Estimation; Özarslan et al., 2009). The intersection of these two approaches result in models that leverage both radial and angular resolution. Models marked with an asterisk are those explored throughout this thesis.

the attenuation signal of diffusion as:

$$E(\mathbf{q}) = \exp\left(-b \cdot D(\mathbf{q})\right), \tag{2.37}$$

where $D(\mathbf{q}) = D(q, \theta, \varphi)$ is the Apparent Diffusion Coefficient (ADC).

AMURA does not intend to further develop on the estimated EAP. Instead, its contribution entails the computation of the scalar maps of RTOP, RTAP and RTPP. The assumption of independency between radial diffusivity and anisotropy results in the ADC being independent from the radial domain. Which results in:

$$D(\mathbf{q}) = D(\theta, \varphi) \rightarrow E(\mathbf{q}) = \exp\left(-b \cdot D(\theta, \varphi)\right). \tag{2.38}$$

Biomarkers derivation. This pragmatic approach allows the return probabilities to be derived from Eqs. (2.20), (2.21) and (2.22) to an analytical solution independent from the $\mathbf{q}$-space sam-

plings:

$$\text{RTOP} = \frac{1}{4}\frac{\sqrt{\pi}}{(4\pi^2\tau)^{3/2}}\int_{\mathcal{S}}\frac{1}{D(\theta,\varphi)^{3/2}}dS, \tag{2.39}$$

$$\text{RTAP} = \frac{1}{2\cdot 4\pi^2\tau}\int_0^{2\pi}\frac{1}{D(\theta')d\theta'}, \tag{2.40}$$

$$\text{RTPP} = \sqrt{\frac{\pi}{4\pi^2\tau}}\sqrt{\frac{1}{D(r_0)}}. \tag{2.41}$$

Since the mono-exponential assumption in Eq. (2.37) only holds for a limited b-value range—i.e. around $b = 2000$ s/mm^2 up to $b = 3000$ s/mm^2 in the mammalian brain (Assaf and Basser, 2005; Jensen et al., 2005)—the diffusion features are said to be *apparent*. The theorized metrics, thus, will only reflect the *apparent* behavior of diffusion for a selected shell.

Numerical Implementation. Regarding the numerical implementation, Aja-Fernández et al. (2020) make use of the SH expansions and their useful property relating the integral of a signal $g(\theta,\varphi)$ over the surface of the unit sphere S to the first SH coefficient:

$$c_0^0\{g(\theta,\varphi)\} = \frac{1}{\sqrt{4\pi}}\int_{\mathcal{S}} g(\theta,\varphi)dS, \tag{2.42}$$

so that,

$$\text{RTOP} = \frac{1}{(4\pi)^2\tau^{3/2}}c_0^0\{H(\theta,\varphi)^{-3/2}\}. \tag{2.43}$$

Also, they make use of the Funk-Radon Transform $\mathcal{G}$ of $D(\theta')^{-1}$ evaluated at the direction of maximum diffusion $\mathbf{r}_0$, equivalent to computing the line integral in Eq. (2.40) along the equator perpendicular to that direction of maximum diffusion(Tuch, 2004).

$$\text{RTAP} = \frac{1}{2\cdot 4\pi^2\tau}\mathcal{G}\{\frac{1}{D(\theta')}\}(\mathbf{r}_0). \tag{2.44}$$

Finally, RTPP makes use of the regularized projection of $D(\theta,\varphi)$ onto the SH, $D_{\text{SH}}(\theta,\varphi)$:

$$\text{RTPP} = \frac{1}{\sqrt{4\pi\tau}}\frac{1}{\sqrt{D_{SH}(\mathbf{r}_0)}}. \tag{2.45}$$

In their work Generalized AMURA, Aja-Fernández et al. (2022) extended the framework so that general moments of the diffusion signal can be estimated. These moments describe properties of the diffusion process with higher accuracy and allow their contrast enhancement. Figures 2.10 and 2.11 show an example of EAP-based moments computed by MiSFIT model.

Advantages, limitations and applicability. AMURA exploits the difficulty for clinical environments for acquiring large data sets, and focuses on computing scalar indices from usual (and reduced) **q**-*samplings*, tackling the main limitations for this kind of long acquisitions.

Despite its recent introduction, the metrics derived from the AMURA framework have

already been proved to be sensitive to chronic vs. episodic migraine (Planchuelo-Gómez et al., 2020; Martín-Martín et al., 2023), and exhibit lower dependency on sample size than DTI (Martín-Martín et al., 2023).

MiSFIT

Diffusion model. The Micro-Structure adaptive convolution kernels and dual Fourier Integral Transforms (MiSFIT, Tristán-Vega and Aja-Fernández, 2021) takes advantage of *shelled* acquisitions to compute scalar indices that are proven to be more computationally efficient and reliable than *de facto* standards like MAPL, while also critically reducing the processing time. MiSFIT is defined by a semi-parametric approach, encompassing a fully non-parametric ODF $\mathbf{\Phi}$ convoluted to an adaptive and parametric microstructural kernel. Said kernel describes the radial behavior of diffusion, and is parameterized by up to three variables: longitudinal diffusivity $\lambda_\parallel$, transverse diffusivity $\lambda_\perp$, and non-free-water signal fraction f. The attenuation signal obtained is, therefore, defined as a spherical convolution of the adaptive Gaussian kernel to with the orientations histogram:

$$E(\mathbf{q}) = \iint_{\mathcal{S}} \mathbf{\Phi}(\mathbf{n}) \exp\left(-b(\mathbf{g}^T\mathbf{n})(\lambda_\parallel - \lambda_\perp) + \lambda_\parallel\right) d\mathbf{n}, \tag{2.46}$$

where $\mathcal{S} = \{\mathbf{n} \in \mathbb{R}^3 : \|\mathbf{n}\| = 1\}$.

Numerical Implementation. While MiSFIT is intended to work with multi-shell samplings, it might be used with single-shells provided that $\lambda_\perp$ is assumed to be 0. In this case, the longitudinal $\lambda_\parallel$ is inferred from the spherical average of the signal $\bar{E}$. Also, if more than two shells are acquired, the model allows to refine the representation by adding a third parameter that portrays the signal fraction of free-diffusing component, so that:

$${}_cE(\mathbf{q}) = (1-f)\exp(-bD_{\text{iso}}) + fE(\mathbf{q}). \tag{2.47}$$

The schematic shown in Fig. 2.9 describes the workflow in MiSFIT framework. In a nutshell, the N_s different shells acquired as the composite attenuation signal ${}_cE(\mathbf{q})$ are projected onto SH, and the spherical means (SM, i.e. the 0-th coefficient of the SH expansion, see Section 2.3.3) of each shell ${}_c\bar{E}(q_i)$ are computed. By solving a least squares (LS) problem, both the microstructural kernel parameters (f,$\lambda_\parallel$, $\lambda_\perp$) and the ODF $\Phi(\mathbf{r})$ are fit to the N_s shells' SM, and to the SH coefficients of the non-composite signal $E(\mathbf{q})$, respectively. As both the microstructural kernel and the ODF fully describe the EAP, moments can then be computed.

Biomarkers derivation. As in previous models, the EAP information has to be condensed into scalar indices for the interpretability of the captured diffusion processes. MiSFIT generalizes the return probabilities described in Section 2.4.2 as moments in the $\mathbf{R}$ or $\mathbf{q}$ domains. In doing so, it classifies moments in three categories:

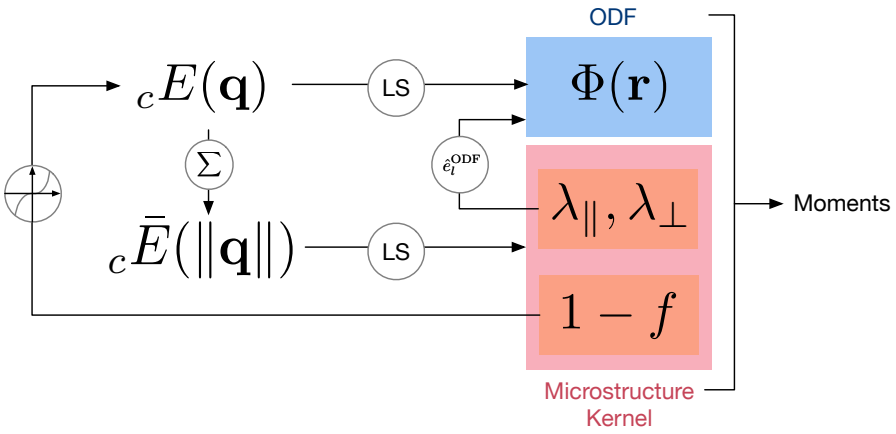


Figure 2.9: MiSFIT computational analysis workflow. The SM $_c\bar{E}(||\mathbf{q}||)$ at each shell is derived from the composite signal $_cE(\mathbf{q})$, and together they are fit to a LS problem so that the microstructural kernel and the ODF can be derived. Once the parameters are obtained, the computation of moments rely on numerical implementations.

Full moments: Those obtained by full integration of the whole 3D space (in either domain).

$$\Upsilon^\nu = \iiint_{\mathbb{R}^3} q^\nu E(\mathbf{q}) d\mathbf{q} \Longleftrightarrow \upsilon^\nu = \iiint_{\mathbb{R}^3} R^\nu P(\mathbf{R}) d\mathbf{R} \tag{2.48}$$

Axial moments: Those obtained by integrating along a line of direction $\mathbf{u}_0$.

$$\Upsilon_a^\nu(\mathbf{g}_0) = \int_{-\infty}^{\infty} q^\nu E(q\mathbf{g}_0) dq \Longleftrightarrow \upsilon_a^\nu(\mathbf{r}_0) = \int_{-\infty}^{\infty} R^\nu P(R\mathbf{r}_0) dR \tag{2.49}$$

Planar moments: Those obtained by integrating over a surface of the plane perpendicular to a certain direction $\mathbf{u}_0$. The origin must be contained within the plane.

$$\Upsilon_P^\nu(\mathbf{g}_0) = \iint_{q\mathbf{g}\perp\mathbf{g}_0} q^\nu E(q\mathbf{g}) dq\mathbf{u} \Longleftrightarrow \upsilon_p^\nu(\mathbf{r}_0) = \iint_{R\mathbf{r}\perp\mathbf{r}_0} R^\nu P(R\mathbf{r}) dR\mathbf{r} \tag{2.50}$$

From these representations, not only the most common metrics can be derived (i.e. MSD $= \upsilon^2$, qMSD $= \Upsilon^2$, RTOP $= \Upsilon^0$, RTAP $= \Upsilon_p^0(\mathbf{u}_\parallel) = \upsilon_a^0(\mathbf{u}_\parallel)$, and RTPP $= \Upsilon_a^0(\mathbf{u}_\parallel) = \upsilon_p^0(\mathbf{u}_\parallel)$) but also other moments of both the propagator and the signal can be computed. Similar to AMURA, the generalization of propagator and **q**-space signal moments allow the enhancement of contrast from different diffusion properties captured by multi-shell acquisitions (see Figs. 2.10 and 2.11).

Advantages, limitations and applicability. One of the main limitations of MiSFIT with respect to other phenomenological models is the need of acquiring *shelled* (spherical) **q**-space samplings. Also, while theoretically a single-shell would suffice, results would rely on the validity of assuming $\lambda_\perp = 0$. Other limitations include those derived from the projection onto SH, which by definition can crop higher frequencies of the EAP.

Despite being a novel contribution, the fast and efficient computation of moments derived from MiSFIT has been found to be useful in ischemic stroke patients: not only metrics correlate with microstructural changes resulting from acute ischemic stroke (Diamandi et al., 2025a), but also

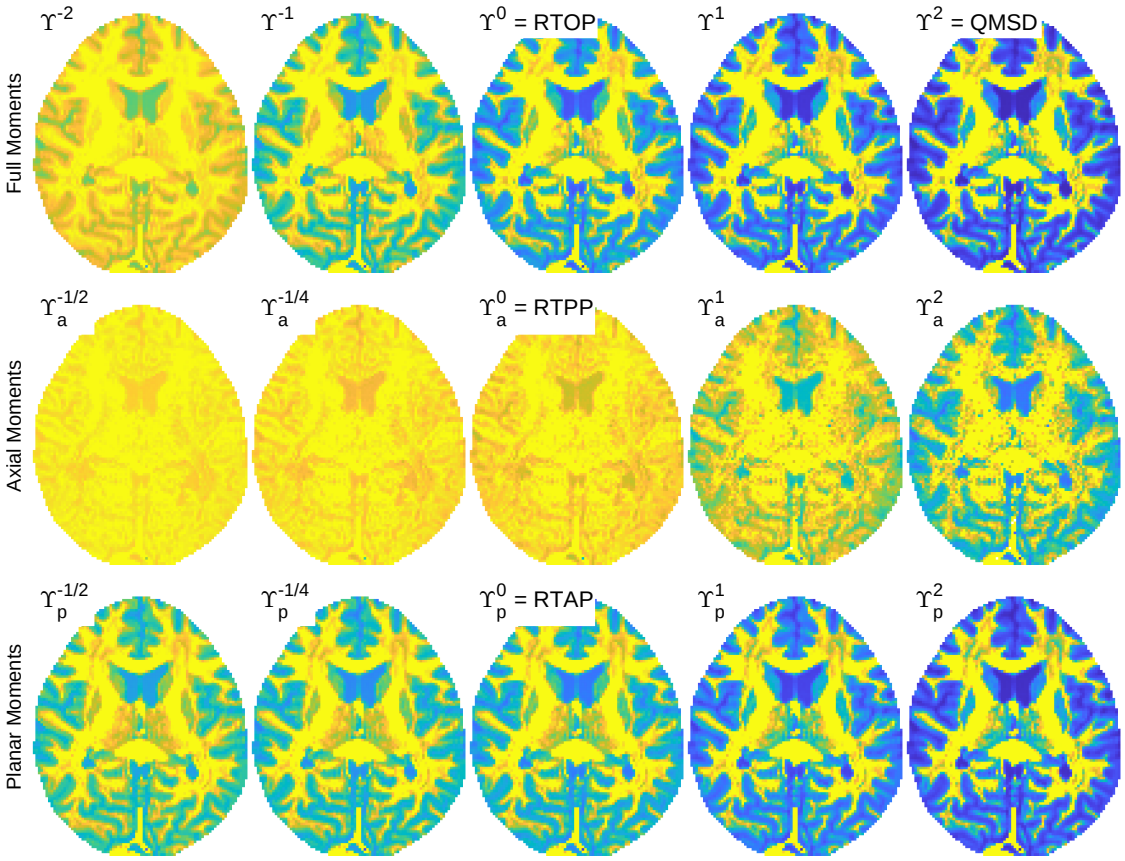


Figure 2.10: MiSFIT Q-space Signal Moments. Different order moments of the **q**-space signal derived from MiSFIT. Each moment is scaled according to its 25 and 75 quantiles. Noticeably, different orders result in distinct contrasts in cerebral tissues.

have been proven to be linked to functional outcomes from such patients (Diamandi et al., 2025b).

MAP-MRI

Diffusion model. Following the introduction of the one dimension Simple Harmonic Oscillator Reconstruction and Estimation (SHORE, Koay and Basser, 2008), Özarslan et al. introduced its 3D counterpart in 2009 (3D-SHORE, Özarslan et al., 2009). This first 3D generalization, however, modeled the diffusion with the same scalar parameter in all directions, thus labeling the diffusion isotropic. To overcome this limitation, Özarslan et al. (2013) implemented MAP-MRI, a framework that follows the same principles but generalizes the scaling to a tensorial form, allowing for a more accurate description of anisotropic diffusion processes.

MAP-MRI aims at projecting the diffusion signal onto a basis whose functions are eigenvectors of the Fourier transform, so that the propagator and the diffusion signal share the same coefficients:

$$E(\mathbf{q}) = \sum_i c_i \Phi_{n_i}(\mathbf{q}) \longrightarrow P(\mathbf{q}) = \sum_i c_i \Psi_{n_i}(\mathbf{q}), \tag{2.51}$$

where $\mathbf{\Psi} = \mathcal{F}^{-1}\{\mathbf{\Phi}\}$, and $\mathbf{c} = \{c_i\}$ are the basis coefficients. The basis functions stem from the

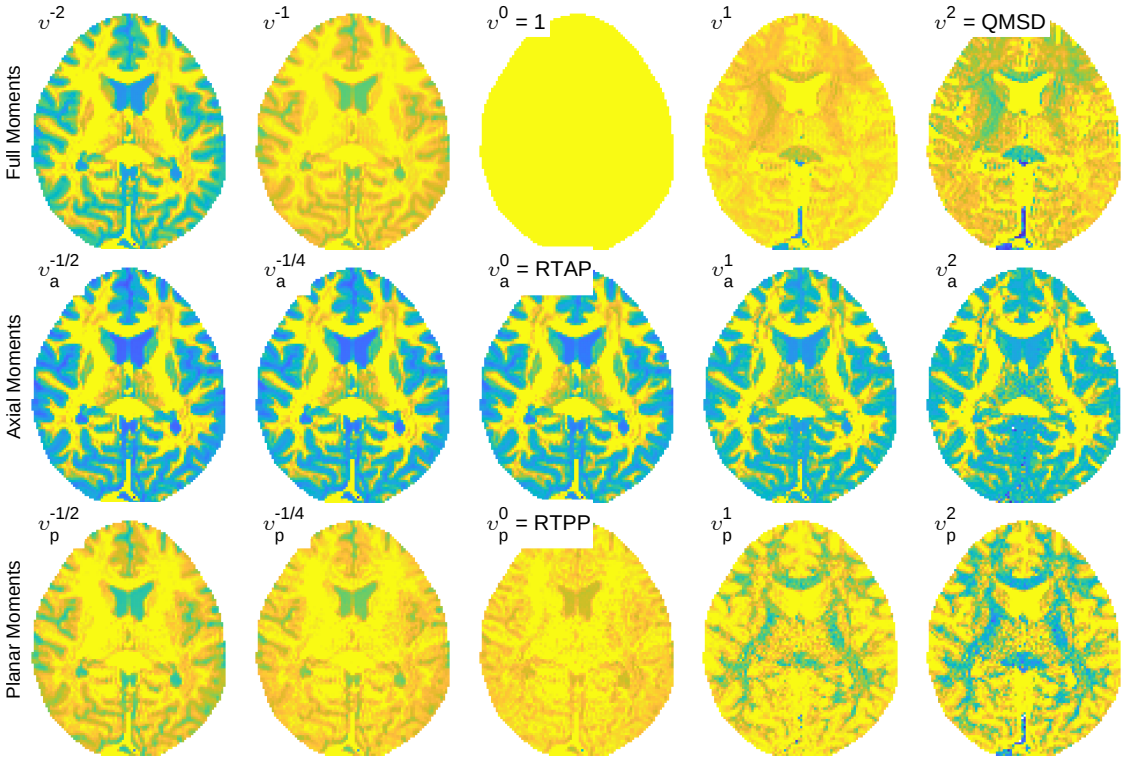


Figure 2.11: MiSFIT Propagator Moments. Visualization of various propagator moments derived from MiSFIT. Each moment is individually scaled to its 25th and 75th percentiles for contrast enhancement. Distinct moment orders reveal complementary contrasts across different cerebral tissue types.

eigenfunctions of the simple harmonic oscillator (SHO) Hamiltonian—also known as Hermite functions. The SHO Hamiltonian is a quantum mechanical operator that describes systems facing simple harmonic motion, such as the one molecules undergo in vibration (McQuarrie and Simon, 1997). By doing so, MAP-MRI mimics previous works in statistics (Schwartz, 1967), and expands **q**-space signal data onto Hermite functions, which has been shown to rapidly converge in both Fourier and real spaces. For the 1-dimensional case, the basis functions in **q**-space are thus defined as follows:

$$a \tag{2.52}$$

where $H_n(x)$ is the Hermite polynomial. The choice of u relies on the signal decay rate $u = \sqrt{2Dt_d}$, which ensures rapid convergence of the series. Note that, with this choice, the **q**-space signal in Eq. (2.51) is the Stejskal-Tanner formula followed by a series of cumulants describing higher order terms.

In 3D MAP-MRI, the signal is represented as a linear combination of basis functions that are separable along three orthogonal directions. These basis functions are constructed as the product of one-dimensional Hermite-like functions, denoted as $\Phi_{n_1}(x')\Phi_{n_2}(y')\Phi_{n_3}(z')$, where $\{n_1, n_2, n_3\}$ are integer indices defining the order along each axis. The axes x', y', z' are not fixed in space but are defined per voxel by the eigenvectors of the diffusion tensor estimated from the same signal. This voxel-wise diagonalization of the diffusion tensor defines a local

reference frame aligned with the principal directions of diffusion, which often coincide with underlying anatomical features such as fiber bundles. By expressing the signal in this "anatomically" informed frame, the basis functions become directionally adapted to the local diffusion structure. Generalizing Eq. (2.51) for the 3D case results in:

$$E(\mathbf{q}) = \sum_{N=0}^{N_{\max}} \sum_{\{n_1,n_2,n_3\}} a_{n_1n_2n_3} \mathbf{\Phi}_{n_1n_2n_3}(\mathbf{A}, \mathbf{q}) \tag{2.53}$$

where $a_{n_1n_2n_3} = \tilde{a}_{n_1n_2n_3}/S(0)$, $\mathbf{A} = 2\mathbf{D}t_d$ is a symmetric, positive-definite, rank-2 tensor that accounts for the spring's stiffness and coupling, as well as its anisotropy (Özarslan et al., 2013), revealing the relationship of MAP-MRI basis with the diffusion tensor $\mathbf{D}$; and

$$S(0) = \sum_{N=0}^{N_{\max}} \sum_{\{n_1,n_2,n_3\}} \tilde{a}_{n_1n_2n_3} B_{n_1n_2n_3}, \tag{2.54}$$

with

$$B_{n_1n_2n_3} = \Gamma_{n_1n_2n_3} \frac{\sqrt{n_1!\; n_2!\; n_3!}}{n_1!!\; n_2!!\; n_3!!}, \tag{2.55}$$

where $\Gamma_{n_1n_2n_3} = 1$ if n_1, n_2, n_3 are all even, and 0 otherwise.

As previously mentioned, when $n_1 = n_2 = n_3 = 0$, the basis function reduces to the zeroth-order term, since $H_0 = 1$. In this case, MAP-MRI matches the diffusion tensor approximation within a rotated anatomical frame where the axes x', y', z' are aligned with the principal diffusion directions. This term alone captures the Gaussian approximation of diffusion, while higher-order terms progressively encode non-Gaussian features of the displacement distribution.

Numerical Implementation: MAPL. Developed upon MAP-MRI, Fick et al. (2016a) proposed MAPL, a Laplacian-regularized MAP-MRI that imposes certain smoothness in the estimated EAP so that estimates are more robust. This allows MAPL to work with less samples and at lower b-values, and achieves a notably proficient EAP reconstruction even with subsampled data. In this formulation, the coefficients in Eq. (2.51) are found by minimizing

$$\mathbf{c} = \arg\min_{\mathbf{c}} \|\mathbf{y} - \mathbf{Q}\mathbf{c}\|^2 + \lambda_\Delta U(\mathbf{c}) \tag{2.56}$$

where λ_Δ weights the Laplacian regularization functional

$$U(\mathbf{c}) = \int_{\mathbb{R}^3} \|\Delta E_c(\mathbf{q})\|^2 d\mathbf{q} \tag{2.57}$$

with $\Delta E_c(\mathbf{q}) = \sum_i c_i \Delta \Phi_{N_i}(\mathbf{q})$ being the Laplacian of the reconstructed signal. This formulation has a unique minimum:

$$\mathbf{c} = \left(\mathbf{Q}^T\mathbf{Q} + \lambda_\Delta \mathbf{U}\right)^{-1} \mathbf{Q}^T\mathbf{y}. \tag{2.58}$$

Furthermore, to find the optimal regularization weights λ_Δ, the Generalized Cross Validation (GCV)

is used

$$\mathrm{GCV}(\lambda, \mathbf{y}) = \frac{\|\mathbf{y} - \hat{\mathbf{y}}_\lambda\|}{N_{\text{data}} - \mathrm{Tr}(\mathbf{S}_\lambda)}, \tag{2.59}$$

where Tr stands for trace, $\mathbf{S}_\lambda = \mathbf{Q}(\mathbf{Q}^T\mathbf{Q} + \lambda\mathbf{U})^{-1}\mathbf{Q}^T$ and $\hat{\mathbf{y}}_\lambda = \mathbf{S}_\lambda\mathbf{y}$.

From a numerical perspective, MAPL can also be formulated as a quadratic programming (QP) problem. Specifically, it is posed as a linear least-squares optimization with linear inequality constraints and a single equality constraint. The goal is to minimize the error between the measured diffusion signal and its reconstruction using MAP-MRI basis functions, while ensuring (constraining) that the estimated EAP integrates to one—a requirement for any valid probability density function. Additionally, non-negativity constraints are often imposed on the EAP at a discrete set of spatial locations to improve physical plausibility.

A more stringent variant of this approach, known as the strictly positive MAP-MRI, was proposed by Dela Haije et al. (2020). This formulation casts the problem as a semi-definite programming (SDP) task, where the EAP is constrained to remain strictly positive throughout its domain by enforcing the positive semi-definiteness of a matrix constructed from the MAP coefficients. Though computationally more demanding, this formulation provides enhanced robustness—particularly under noisy acquisition or reduced angular sampling—by guaranteeing global positivity of the propagator.

Given the theoretical framework in which the MAP models are derived, both MAPL and its predecessor MAP-MRI can compute EAP-derived microstructural metrics independently of the specific **q**-space acquisition scheme. However, high b-values remain essential to accurately capture fine features of brain tissue microstructure.

Biomarkers derivation. Once the EAP has been reconstructed, the moments can be derived with greater or lesser complexity. Given the basis upon which MAPL is built, in which the coefficients are estimated in a reference frame determined by the anatomy, the estimation of the return probabilities from MAPL coefficients is trivial. RTOP is the integral of the signal over all $\mathbb{R}^3$, RTAP is the integral of the $P(0, y, z)$ over the yz-plane, and RTPP is the integral of $P(x, 0, 0)$ over the x-axis (Özarslan et al., 2013).

$$\mathrm{RTOP} = \frac{1}{\sqrt{8\pi^3|A|}} \sum_{N=0}^{N_{\max}} \sum_{n_1,n_2,n_3} (-1)^{N/2} a_{n_1n_2n_3} B_{n_1n_2n_3}, \tag{2.60}$$

$$\mathrm{RTAP} = \frac{1}{2\pi u_y u_z} \sum_{N=0}^{N_{\max}} \sum_{n_1,n_2,n_3} (-1)^{(n_2+n_3)/2} a_{n_1n_2n_3} B_{n_1n_2n_3}, \tag{2.61}$$

$$\mathrm{RTPP} = \frac{1}{\sqrt{2\pi} u_x} \sum_{N=0}^{N_{\max}} \sum_{n_1,n_2,n_3} (-1)^{n_1/2} a_{n_1n_2n_3} B_{n_1n_2n_3}. \tag{2.62}$$

Similarly, due to the fact that the first coefficient of the MAPL series is built as the closest Gaussian profile to the EAP, the NG is easily derived as the angular distance between said

first coefficient to the whole signal:

$$\mathrm{NG} = \sin\theta_{EO} \text{ where } \cos\theta_{EO} = \frac{a_{000}}{(\sum_{N=0}^{N_{max}} \sum_{\{n_1,n_2,n_3\}} a_{n_1,n_2,n_3}^2)^{1/2}} \tag{2.63}$$

The PA, on the other hand, results in complex and time-consuming computations due to the fact that the isotropic equivalent's coefficients have to be fully recalculated from the anisotropic ones.

Advantages, limitations and applicability. The powerful properties of MAPL framework make of this model one of the most used EAP imaging tools in the neuroscientific community. For example, its capability to reconstruct the EAP with no assumptions over the diffusion in tissues can be beneficial in testing pathologies, where theoretical biophysical assumptions might not hold (Lampinen et al., 2019; Palmqvist et al., 2020). However, it is only fair to point out its limitations: its time-expensive nature. The fitting of the basis functions, together with the generalized-cross-validation (GCV)-derived Laplacian regularization and the positivity constraints result in overwhelming computational complexities, which in turn translates in processing times that are unfeasible in clinical practice (París et al., 2023a). Part of this is tackled in a recent Matlab implementation (Tristán-Vega, A. and Aja-Fernández, S. and París, G., 2022), similar to (París et al., 2023b).

MAP-MRI, with over 400 citations, together with MAPL, with around 120, have demonstrated to be powerful frameworks to disentangle the propagator from the diffusion signal, as well as to compute metrics that are linked to brain's microstructural changes. For this, together with their implementation into the DIPY package (Garyfallidis et al., 2014), they are currently considered to be the *de facto* standard in EAP imaging. Their derived metrics have been proved useful in both clinical and research settings: in glioma differential diagnostic (Wang et al., 2021a), and genotyping (Gao et al., 2022), sensitive on microstructural changes in AD (Spotorno et al., 2022) or more sensitive to age effect in neurodevelopment characterization (Pines et al., 2020), sensitive to sleep disorders (Örzsik et al., 2023), etc.

DSI

Diffusion model. Diffusion Spectrum Imaging (DSI; Wedeen et al., 2000, 2005) takes advantage of the mathematical definition of the propagator in Eq. (2.16). The framework aims at directly estimating the propagator by exploiting a Cartesian grid **q**-space acquisition, so that the propagator is efficiently computed by means of a Fast Fourier Transform (FFT). The use of Cartesian lattice minimizes aliasing effects, and makes it possible to avoid the use of non-uniform Fourier techniques (NUFFT) which are computationally expensive and less stable. The result is a fully non-parametric EAP characterization.

Numerical Implementation: HYDI-DSI-QP. The main limitation of the DSI framework is the extensive **q**-space acquisition required for an accurate propagator estimation, typically consisting of approximately 200 samples. Recent contributions addressed this limitation by developing on interpolation techniques to locally map the diffusion signal onto Cartesian grids from spherical or sparse acquisitions (Wu and Alexander, 2007; Özarslan et al., 2013; Setsompop

et al., 2013; Ning et al., 2015; Sjölund et al., 2016; Tobisch et al., 2018). Ultimately, this enables the non-parametric characterization of the EAP from shorter acquisitions, albeit with the inherent limitations of interpolation techniques: loss of high frequency information, introduction of biases and artifacts, etc.

One of the proposed alternatives to DSI is Hybrid Diffusion Imaging (HYDI; Alexander et al., 2006; Wu and Alexander, 2007), which builds upon a series of scientific contributions. First, HYDI is a dMRI framework designed to characterize complex diffusion patterns by leveraging a multi-shell acquisition strategy. It separates inner shells—where diffusion is assumed to follow a Gaussian distribution—from outer shells—where non-Gaussian diffusion is captured. HYDI employs the DTI model for inner shells and HARDI for outer ones, taking advantage of the higher angular resolution of the latter. Additionally, HYDI incorporates HYDI-DSI, which re-grids and interpolates the spherical sampling of the full $\mathbf{q}$-space onto a Cartesian lattice, so that DSI can be estimated from it. In the original work by Alexander et al. (2006), this regridding is performed using trilinear interpolation of the raw $\mathbf{q}$-space data, followed by the estimation of the displacement spectrum for each voxel via a 3D FFT.

Despite its proved clinical applicability in characterizing TBI (Wu et al., 2018), AD (Daianu et al., 2015), or gender and age-related changes (Ouyang et al., 2021), HYDI-DSI is not free from certain limitations. For example, the framework assumes that the EAP is compact supported (i.e., it is zero outside the acquired bandwidth), which directly depends on the maximum b-value acquired. Also, the framework suggests specific $\mathbf{q}$-space acquisitions (evenly spaced shells) so that interpolation does not use some samples more than others. Furthermore, positivity constraints, which have been proved to be meaningful in dMRI (Dela Haije et al., 2020; Herberthson et al., 2021) cannot be ensured.

HYDI-DSI-QP (Tristán-Vega et al., 2023) reformulates the original HYDI-DSI idea by means of a Fourier Transform encoding matrix that eliminates the need of $\mathbf{q}$-space regridding, while maintaining the non-parametric nature of the EAP estimation. With this approach, not only unit-mass, but also positivity constraints can be enforced in the solution. Also, the return probabilities (see Section 2.4.2) can be computed without the need for more regridding. Figure 2.12 shows the comparison between the original HYDI-DSI workflow and the proposed one.

First, the encoding matrix is built for each voxel according to a DTI estimation using only $b < 2000$ mm^2/s. This step involves the computation of the parameters in Eq. (2.7) (i.e., $\mathbf{\Lambda}$ and $\mathbf{Q}$). The vectors contained in $\mathbf{Q}$ are used to align the Cartesian lattice to the principal directions of the $\mathbf{D}$: With a change of variable $\mathbf{R}' = \mathbf{Q}^T\mathbf{R} \Rightarrow \mathbf{R} = \mathbf{Q}\mathbf{R}'$, the direction of maximum diffusion is aligned to the 'z' axis of the Cartesian grid.

$$E_{\mathbf{Q}}(\mathbf{q}') = E(\mathbf{Q}\mathbf{q}') \tag{2.64}$$

Second, by assuming that the attenuated signal is compact supported—i.e. it vanishes outside the range of study $\mathbf{\Omega}$—the shifted versions of $E_{\mathbf{Q}}(\mathbf{q}')$ can be arranged to form a periodic

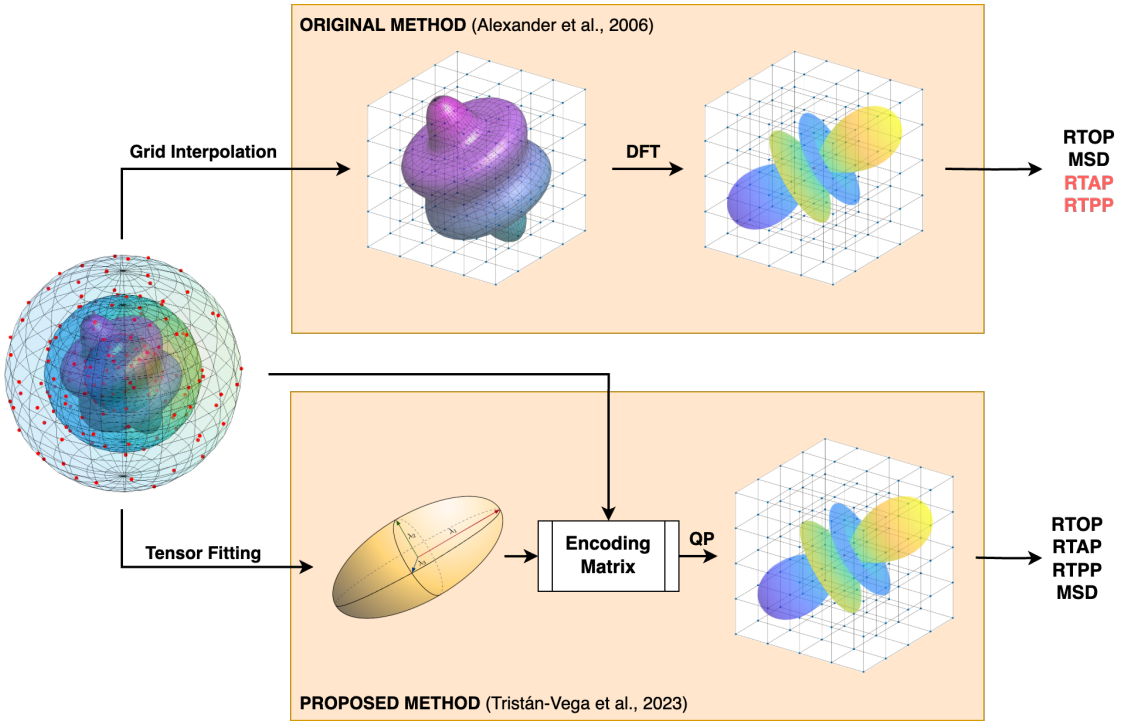


Figure 2.12: HYDI-DSI workflow comparison. In the original work, the attenuation signal in **q**-space is regridded and interpolated onto a Cartesian lattice, which then enables a direct characterization of the EAP via DFT. In the proposed work, a tensor is fit to the diffusion-weighted signal and then used to create the encoding matrix. By doing this, the EAP can be characterized by means of a QP problem. Moments in red represent those that must be estimated by means of numerical interpolation.

signal in the three coordinates $[q_x, q_y, q_z]^T$:

$$\tilde{E}_Q(\mathbf{q}') = E_Q(\tilde{\mathbf{q}}') \;\forall\; \tilde{\mathbf{q}}' \in \{\tilde{q_x}, \tilde{q_y}, \tilde{q_z}\}, \tag{2.65}$$

so that,

$$\tilde{q}_u = q_u - Q_u \left\lceil \frac{q_u - Q_u/2}{Q_u} \right\rceil \;\forall\; u \in \{x, y, z\}. \tag{2.66}$$

By further assuming $\tilde{E}_Q(\mathbf{q}')$ has finite power (i.e., it can be written in terms of Fourier series expansion; Oppenheim et al., 1997), and is both real and exhibits antipodal symmetry, a cosine series can be used:

$$\tilde{E}_Q(\mathbf{q}') = \frac{1}{\Omega} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} P_Q(\mathbf{R}'_{k,l,m}) \cos\left(2\pi \mathbf{q}'^T \mathbf{R}'_{k,l,m}\right). \tag{2.67}$$

Since the available **q**-space samples N_i are limited, in practice the Fourier series expansion will be cropped, so that:

$$\tilde{E}_Q(\mathbf{q}') \simeq \frac{1}{\Omega} \sum_{k=-N_x}^{N_x} \sum_{l=-N_y}^{N_y} \sum_{m=-N_z}^{N_z} P_Q(\mathbf{R}'_{k,l,m}) \cos\left(2\pi \mathbf{q}'^T \mathbf{R}'_{k,l,m}\right), \tag{2.68}$$

where N_x, N_y and N_z are chosen to fully cover the entire x, y, and z axes with evenly spaced samples at each $1/Q_x$, $1/Q_y$, and $1/Q_z$. The higher the N_x, N_y, and N_z, the fewer samples that will be discarded by the compact support assumption. However, it must balance $N_l = (2N_x+1)(2N_y+1)(2N_z+1)$, which should be similar to N_i for well-posedness sake. This previous restriction is relaxed once taken into account the antipodal symmetry of the propagator, which lowers the DOF number to estimate to $n = (N_l+1)/2$, or add a regularization term.

From Eq. (2.68), the encoding matrix is defined as the relationship between the $N_i \times n$ vector of measurements $\mathbf{E}$ with the $n \times 1$ vector of unknowns $\mathbf{P}$, so that:

$$\mathbf{E} \approx \mathcal{F}\mathbf{P} : [\mathcal{F}]_{i,j} = \frac{\kappa}{Q}\cos\left(2\pi \mathbf{q}_i'^T \mathbf{R}_j'\right) \tag{2.69}$$

All in all, resulting in a Quadratic Programming (QP) problem that is defined as:

$$\min_{\mathbf{P}} \frac{1}{2}\|\mathbf{E}-\mathcal{F}\mathbf{P}\|^2 + \frac{\lambda}{2}\|\mathcal{L}\mathbf{P}\|^2 \text{ s.t. } \mathbf{P} \geq 0 \text{ and } f_0^T\mathbf{P} = 0, \tag{2.70}$$

where $\Omega = Q_xQ_yQ_z$, $\mathbf{R}'_{k,l,m} = [k/Q_x, l/Q_y, m/Q_z]^T$, and the regularization term $\frac{\lambda}{2}||\mathcal{L}\mathbf{P}||^2$ (with $\lambda > 0$ a small positive constant and $\mathcal{L}$ some linear operator) is proposed resorting to a Laplacian penalty to promote smoothness over the solution (Descoteaux et al., 2007; Caruyer and Deriche, 2012; Fick et al., 2016a).

Biomarkers derivation. The derived metrics such as the return probabilities and MSD can be computed from Eq. (2.68):

$$\text{RTOP} = P(0) \tag{2.71}$$

$$\text{RTAP} \simeq \frac{1}{Q_z}\left(P_Q(0) + 2\sum_{m=1}^{N_z} P_Q(\mathbf{R'}_{0,0,m})\right) \tag{2.72}$$

$$\text{RTPP} \simeq \frac{1}{Q_xQ_y}\left(P_Q(0) + 2\sum_{k=-N_x}^{N_x}\sum_{l=1}^{N_y} P_Q(\mathbf{R'}_{k,l,0}) + 2\sum_{k=1}^{N_x} P_Q(\mathbf{R'}_{k,0,0})\right) \tag{2.73}$$

$$\begin{aligned}\text{MSD} \simeq \frac{2}{Q}\Bigg(&\sum_{k=1}^{N_x} P_Q(\mathbf{R'}_{k,0,0})\|\mathbf{R'}_{k,0,0}\|^2 + \sum_{k=-N_x}^{N_x}\sum_{l=1}^{N_y} P_Q(\mathbf{R'}_{k,l,0})\|\mathbf{R'}_{k,l,0}\|^2 + \\ &+ \sum_{k=-N_x}^{N_x}\sum_{l=-N_y}^{N_y}\sum_{m=1}^{N_z} P_Q(\mathbf{R'}_{k,l,m})\|\mathbf{R'}_{k,l,m}\|^2\Bigg).\end{aligned} \tag{2.74}$$

Advantages, limitations and applicability. HYDI-DSI-QP is able to estimate fully non-parametric and unit-mass constrained EAP from arbitrary **q**-space samplings. The framework offers a series of advantages when compared to the previous implementation, from more robust characterization of the EAP indices, to the time-efficiency derived from replacing the regridding and linear interpolation of HYDI-DSI to the Laplacian-regularized optimization problem. Also, it allows the possibility of enforcing positiveness, which is not feasible in interpolation-based methods,

and that comes forth as a non-negligible problem in common WM structures (Dela Haije et al., 2020). All in all, HYDI-DSI-QP is proven to achieve comparable, if not better, results than those of MAPL, even in terms of repeatability and reliability (see Section 5). It is only fair to notice that avoiding the interpolation in **q**-space yields to the introduction of two parameters: bandwidth selection μ, related to the maximum b-value that will be sampled, and Laplacian penalty weighting λ, which were not present in the original HYDI. These two parameters, nonetheless, are proven to be stable and can be fixed regardless of the **q**-space sampling used.

The recently introduced HYDI has not been extensively tested yet in clinical group studies. Nevertheless, the performed experiments carried out within this thesis show the potential of this framework (see Section 5). The high number of DOF through its fully-non-parametric nature results in a flexible method that can adapt to all kinds of propagators. Recent studies have employed HYDI-DSI-QP in the heart and obtained promising results Rodríguez-Galván et al. (2023).

2.4.4 Biophysical Models

As previously mentioned, biophysical modeling aims at bridging this gap by representing the diffusion signal as originated from multiple tissue compartments, each with different microstructural or biophysical properties.

The NODDI Model

Before diving into the Standard Model, it will be useful to briefly introduce the Neurite Orientation Dispersion and Density Imaging (NODDI, Zhang et al., 2012), as it is used in one of the contributions. NODDI is a biophysical model that represents the brain tissue as a mixture of intra-neurite, extra-neurite and isotropic compartments. From this framework, it allows the estimation microstructural indices such as the neurite density index (NDI), and the orientation dispersion index (ODI), together with the isotropic volume fraction.

While NODDI provides biophysically interpretable information, it ultimately relies on several assumptions that limit its ability to capture the full complexity of the diffusion signal. These assumptions consist on the non-exchanging nature of the compartments, the fixed intra-neurite diffusivity, or the attribution of isotropic diffusion to free water (CSF), among others.

The Standard Model

The Standard Model of Diffusion Imaging (SMI, Novikov et al., 2018b) has been recently proposed as the overarching picture drawn from previously developed biophysical models: Jespersen et al. (2007); Assaf et al. (2004); Zhang et al. (2012); Setsompop et al. (2013); Jelescu et al. (2016b). In turn, these models follow as special cases. As biophysical models, their main goal is to elucidate the microstructural origins of the diffusion behavior portrayed in the diffusion signal.

The diffusion-weighted signal $S(\mathbf{q})$, acquired for a diffusion gradient $\mathbf{q}$ with magnitude $\|\mathbf{q}\| = b$

and direction $\angle\mathbf{q} = \widehat{\mathbf{g}}$, can be modeled as:

$$S(\mathbf{q}) = \iint_{\mathcal{S}} \mathbf{\Phi}(\mathbf{n})\mathcal{K}(\mathbf{q})d\mathbf{n} = (\mathbf{\Phi} \circledast \mathcal{K})(\mathbf{q}), \tag{2.75}$$

where $\mathbf{\Phi}(\mathbf{n})$ is the normalized fODF and $\mathcal{K}$ the signal response microstructural kernel of a fascicle aligned to the direction $\mathbf{n}$ (Jespersen et al., 2007). To ease subsequent formulations, (Novikov et al., 2018b) normalize $d\mathbf{n} \equiv \frac{\sin\theta d\theta d\varphi}{4\pi}$ so that $\int d\mathbf{n} \equiv 1$. This general form, where the measured signal is modeled as the convolution of a microscopic kernel with an ODF, is also the basis of MiSFIT (see Eq. 2.46); however, while SMI imposes explicit biophysical constraints on the kernel parameters, MiSFIT models the kernel phenomenologically.

With this, SMI often parameterizes the kernel as:

$$\mathcal{K}(b,\xi) = S(0)\left(fe^{-bD_a\xi^2} + (1-f)e^{-bD_e^{\perp} - b(D_e^{\parallel} - D_e^{\perp})\xi^2}\right), \tag{2.76}$$

where $\xi = \mathbf{g}^T\mathbf{n}$ depicts the relative angle between gradient direction and the fascicle, i.e. $\cos\theta = \mathbf{g}^T\mathbf{n}$. The kernel parameters have direct link to the biophysical environment of the voxel. As shown in Fig. 2.13, they are defined as intracellular signal fraction f, longitudinal intracellular diffusivity D_a, extracellular parallel diffusivity $D_e^{\parallel}$ and extracellular perpendicular diffusivity $D_e^{\perp}$. Notice that intracellular perpendicular diffusivity is assumed to be 0 (Veraart et al., 2018).

Note the similarity between the previous equations and MiSFIT's formulation in Eq. (2.46). In fact, MiSFIT can be interpreted as a particular case of the more general SMI framework, where the microstructural kernel is not a mixture of multiple components, but rather a single, non-degenerate Gaussian—i.e., a 3D Gaussian with full covariance and nonzero diffusivity along all directions. However, while the mathematical structure of MiSFIT's formulation resembles that of SMI, it is important to stress that the two approaches differ in their modeling philosophy. MiSFIT adopts a phenomenological perspective, using a fixed Gaussian kernel purely to improve signal representation without directly interpreting it in biophysical terms. In contrast, SMI is a biophysical model explicitly designed to infer tissue microstructure by modeling the signal as arising from a mixture of restricted and hindered compartments. Therefore, despite the formal similarity, the two models serve different purposes and make different assumptions about the origin of the diffusion signal.

Following the parametrization of the fODF in the SH basis p_l^m, Eq. (2.75) can be rewritten as:

$$S(\mathbf{q}) = \sum_{\substack{l=0,\\ l\ \text{even}}}^{\infty} \sum_{m=-l}^{l} S_l^m(b) Y_l^m(\mathbf{g}) = \sum_{\substack{l=0,\\ l\ \text{even}}}^{\infty} \sum_{m=-l}^{l} p_l^m K_l(b,\boldsymbol{\theta}) Y_l^m(\mathbf{g}), \tag{2.77}$$

where K_l is the projection of the response kernel in Eq. (2.76) onto the Legendre Polynomials for a given set of microstructural parameters $\boldsymbol{\theta} = \{f, D_a, D_e^{\parallel}, D_e^{\perp}\}$, as per Funk-Hecke's Theorem (Hobson, 2012). Y_l^m are the SH basis functions (*cf.* Section 2.3.3). In order to further reduce the complexity of the fitting problem, Novikov et al. (2018b) proposes the use of rotational invariant features derived from the spherical harmonics representation (RISH) to lower the

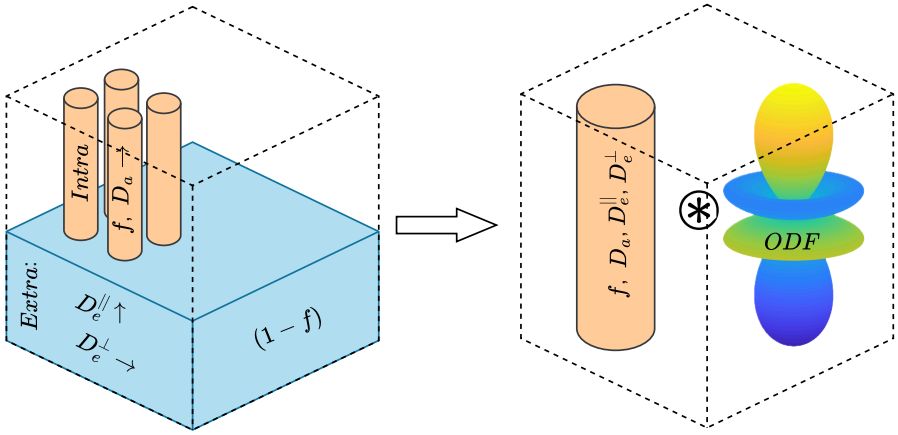


Figure 2.13: SMI diagram. On the left, the diffusion in the scanned voxel is assumed to be caused by microstructure consisting of, at least, 4 parameters: f, D_a, $D_e^{\parallel}$, and $D_e^{\perp}$. On the right, SMI models the signal from the voxel as contributions of fiber segments according to the ODF.

dimensionality of the signal:

$$S_l^m(b) = p_l^m K_l(b, \theta), \tag{2.78}$$

so that,

$$S_l(b) = p_l K_l(b, \theta), \tag{2.79}$$

where $p_l = \frac{1}{N_l} \sum_{m=-l}^{l} |p_l^m|^2$. Note that, for $l = 0$, the number of parameters to estimate is 4 (the microstructural parameters in the response kernel); whereas for $l \geq 2$ additional p_l have to be estimated.

The projection of the signal on its RISH features provides a low-dimensional, rotation-invariant representation of the signal, which eases the fitting procedure. The parameters θ are oftentimes estimated via a linear least squares (LLS) fitting problem, resulting in:

$$\widehat{\theta}^{(2)} = \arg\min_{\theta} \sum_{j=1}^{N_b} \sum_{\substack{l=0, \\ l \text{ even}}}^{L} \frac{1}{\sigma_{lj}^2} \left| S_l(b_j) - p_l K_l(b, \theta) \right|^2, \tag{2.80}$$

where b_j is the b-value for the j-th shell (among the N_b shells), L is the maximum SH degree onto which the signal is projected, and $\sigma_{lj} \sim (2l+1)/N_g$ weights to compensate for the heteroscedasticity, with N_g being the number of diffusion gradients in the j-th shell (Novikov et al., 2018b). The parameters maps derived from SMI can be seen in Fig. 2.14.

It is important to note, however, that SMI does not inherently require the use of RISH features or LLS-based fittings. Alternative estimation strategies, such as conditional least squares (CLS) or maximum likelihood estimation (MLE), can also be used within the SMI framework. Each approach, however, brings its own challenges, particularly related to the signal distribution and the specification of priors, as will be discussed in Chapter 4.

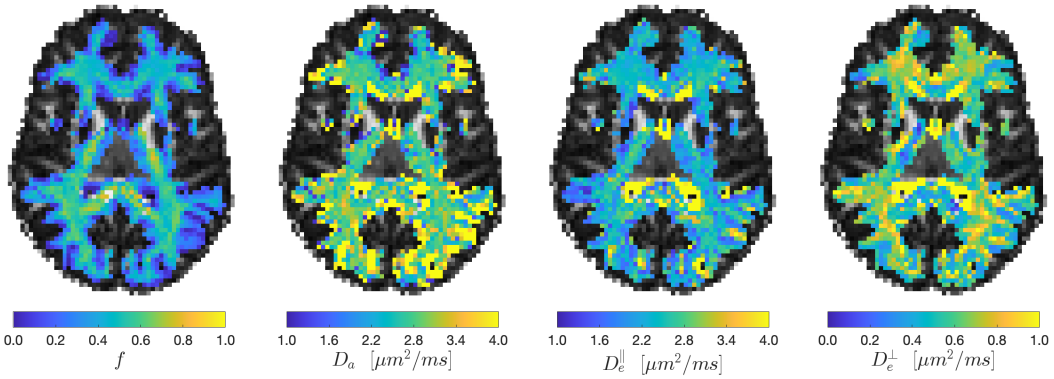


Figure 2.14: SMI microstructural estimates. From left to right, f, D_a, $D_e^{\parallel}$, and $D_e^{\perp}$ microstructural parameters. Only the WM is shown, overlaid to the gray-scaled FA. For more details, refer to Chapter 4.

Unveiling the microstructural environment catalyzing certain diffusion phenomena is a promising application of diffusion MRI, one that has been taking momentum recently. SMI offers great interpretability, and the possibility of discriminating between tissues, adding flexibility to current specificity and sensitivities models.

However, SMI still has a difficult time finding coherent fittings, even with oversampled acquisitions. The noise has a great impact on the model, which is degenerate in origin. Other limitations include the assumption of equal scalar parameters across all fibers in a given voxel, which might result in inaccuracies in fiber crossings. Later in this work, however, we focus on the effect of thermal noise onto the rotational invariant representations used by SMI, which preemptively further affects accuracy and stability of SMI estimates. We also propose an alternative estimation method that avoids the use of rotational invariant features and that has been proven to be more stable, reproducible and robust to protocol acquisition variability (see Chapter 4).

Regardless of its limitations, SMI has been proved to show both sensitivity and specificity in certain applications. Its overarching essence causes it to inherit previous applications. For example, the scalar parameters in the response kernel have been proven sensitive to demyelination (Fieremans et al., 2012; Jelescu et al., 2016b, depicted by $D_e^{\perp}$), axonal loss (Fieremans et al., 2012, f), beading (Budde and Frank, 2010, D_a), inflammation and edema (Unterberg et al., 2004, D_a, $D_e^{\parallel}$, and $D_e^{\perp}$), etc.

2.5 Clinical Validation

Having reviewed some of the most important state-of-the-art biophysical and phenomenological models, it becomes clear that numerous frameworks already exist to assess diffusion signals. The vast amount of models, and scalar metrics that are derived from them increase year by year, yet convergence seems to remain one step ahead. In line with this, if strides for convergence are to be made, it should be clarified how it should be produced in a consensual manner. Moreover, the lack of ground truth in diffusion MRI only hinders the evaluation of measures, and despite advances in the development of phantoms that replicate biological structures, there is no possibility of dealing

with the abundant variety of tissues and microscopic systems that coexist in the brain.

A common approach to deal with this issue is the evaluation of systems, models or metrics with regards to their sensitivity to clinical variations over group studies. That is, if a system, model or metric is able to find statistical differences in a group study, and is thus able to classify healthy (control) subjects from patients in a statistically significant manner, it is presumed to be sensitive to a given disease. However, such clinical validation is prone to misconceptions. As already introduced, diffusion MRI data is often scarce because of the time-expensive nature of diffusion acquisitions. Extrapolating to full test-retest datasets, these are often poor in terms of the number of subjects and sessions per subject. All in all, these datasets features result in findings that may lack statistical power (Button et al., 2013). Similarly, dMRI data is often poor in terms of quality, which may reduce the reliability of metrics, therefore leading to analogous weaknesses.

The previously introduced problems are common in neuroscience community. Statistical power is crucial in finding individual differences, as it determines the ability to detect effects of interest. Zuo et al. (2019) notes the interplay between reliability, effect size and sample size is usually overlooked. Reliability refers to the consistency, stability and dependability of a system, measure or process over time. It quantifies how well the system produces the same result under consistent conditions. Sample size refers to the number of observations or data points used in a study. It determines the study's power to detect an effect and the precision of its estimates. Effect size refers to the magnitude of the difference or relationship being measured. It reflects how strong or meaningful the observed effect is in practical or statistical terms. Figure 2.15 shows the interaction between these three domains assuming a statistical power of 0.8, and how sample size requirements depend on both effect size and reliability. If few data measurements are provided (low sample size), only big effect sizes and high reliability will lead to statistical powerful results. This previous point is important given that effect size is limited by the reliability of a measure. Thus, non-reliable metrics attenuate the effect sizes. Also, statistically significant differences between groups (e.g. patients vs. controls) that derive from low reliability measures might not be reproducible in other contexts or settings, specially given the quality and scarcity problems that dMRI data. In other words: low quality data is usually more noise-driven which, together with low sample sizes may lead to low reliability of metrics. Now, measures that exhibit low reliability attenuate effect sizes, up to a point where statistical significance might not even be reproducible across studies.

2.6 Preclinical Validation through Repeatability and Separability

Once acknowledged the interaction between data quality, sample size, effect size, and reliability, it is easy to see that neuroimaging researchers may not be able to directly address the lack of sample size or data quality in publicly available datasets. In light of this, it is important to recognize the complexity of inferring insights from current experiments and note that such insights might be driven by noise or artifacts rather than proper individual differences. One thing the neuroimaging community is able to address, nonetheless, is the evaluation of the reliability of the metrics being developed.

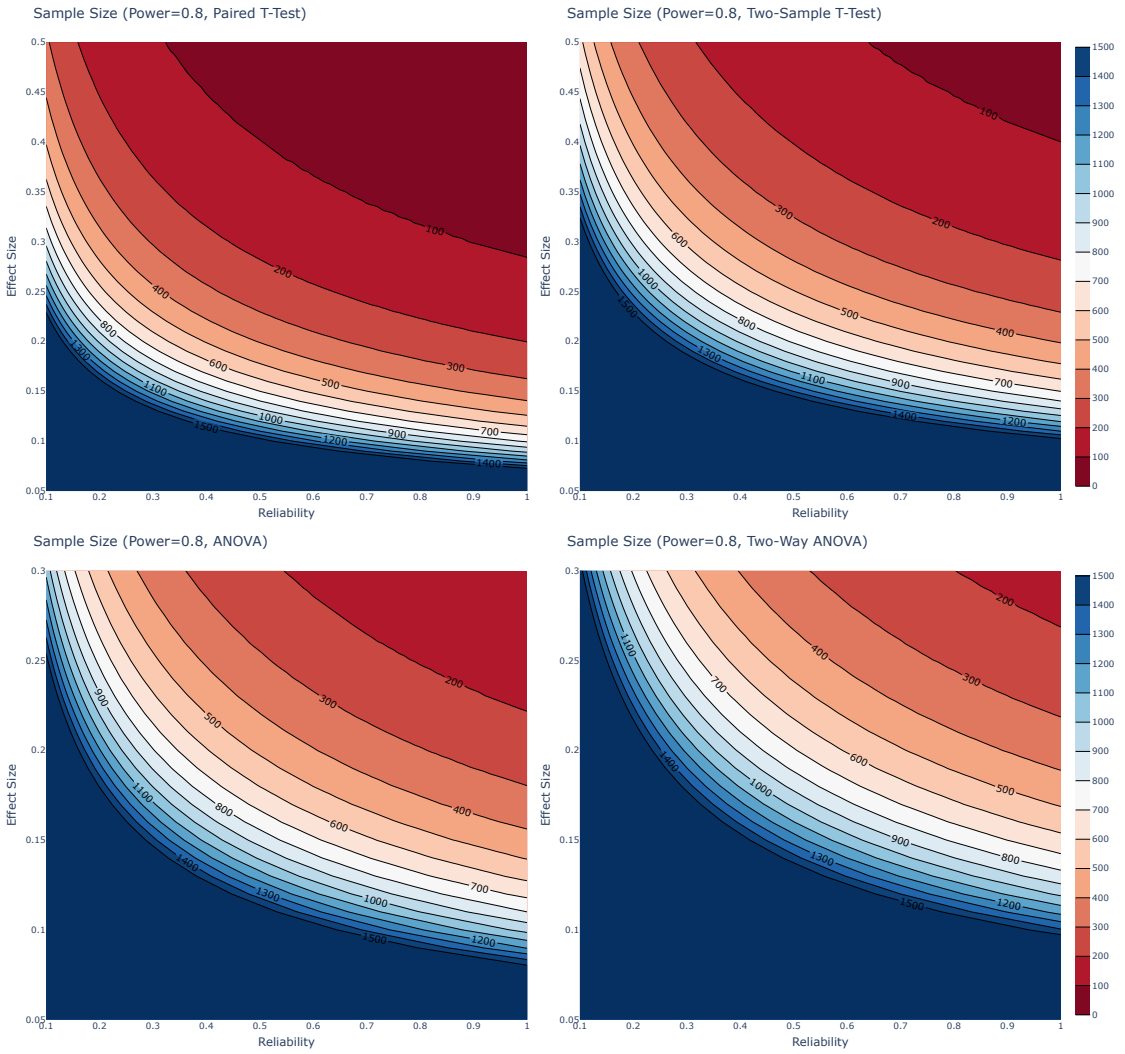


Figure 2.15: Interplay between reliability, effect size and sample size for a given statistical power of 0.8. Relationship between reliability and effect size for three different types of statistical tests: Pairet t-test, Two-sample t-test and ANOVA.

In that regard, there is a crucial need to implement methodologies to assess the reliability of the scalar measures derived from diffusion frameworks. Even though researchers have come up with different ways of assessing the reproducibility of their findings, metrics, methods, and frameworks; there has been growing criticism of the current methodologies used for such assessment (Laumann et al., 2015; Xu et al., 2016; O'Connor et al., 2017), even in the use or meaning of the word "reliability" itself (Nichols et al., 2017). This thesis also makes echo of such criticism and elaborates on how to address these shortcomings.

In the scope of this document, we will call repeatability to the consistency that a metric exhibits across several measurements from the same subject. Separability, on the other hand, refers to the ability of a metric to consistently exhibit differences between classes (i.e., between subjects). Having said this, nowadays, a portion of contributions tend to assess the adequacy of scalar maps by evaluating their metrics through test-retest Coefficient of Variation (COV,

Grech-Sollars et al., 2015; Paudyal et al., 2019; McDonald et al., 2023; Ades-Aron et al., 2025), or simply by computing test-retest correlations. While within-subject robustness (i.e., repeatability) is a must-have for any biomarker, it does not automatically imply its suitability. Conversely, other researchers make use of the intraclass correlation coefficient (ICC), which does not only quantify repeatability but also considers inter-subject variability (i.e., separability), providing a more comprehensive picture of a metric's validity in differentiating between individuals.

It is clear that both repeatability and separability are desirable properties for any neuroimaging-derived marker. However, they alone do not ensure the biological validity of a metric, nor its relevance in a clinical or cognitive context. This is where notions such as specificity and sensitivity come into play. In essence, sensitivity refers to the ability of a metric to correctly detect the presence of an effect (e.g., a disease, condition, or cognitive trait), while specificity describes its ability to avoid false positives—detecting an effect where none exists. These two concepts are deeply embedded in the broader framework of biomarker validation and offer a useful lens through which to interpret the utility of a scalar map.

While repeatability and separability reflect the internal structure of a metric (its stability across time and its discriminative power across subjects), sensitivity and specificity relate more directly to the external validity of that metric—its ability to respond appropriately to meaningful biological variation while remaining robust against noise. Interestingly, there is a subtle but critical interaction between these dimensions. A metric that lacks repeatability will inevitably struggle with sensitivity, as the noise within a subject may obscure any true signal. Likewise, poor separability may impair specificity, blurring distinctions between groups and increasing the likelihood of type I errors.

In this sense, reliability acts as a foundational requirement for any attempt to interpret or deploy neuroimaging metrics in a meaningful way. Without sufficient reliability, no claim about specificity or sensitivity can be made with confidence, as the signal being measured may not reflect a stable biological feature but rather the artifact of methodological noise. Throughout this thesis, we explore methodologies aimed at reinforcing this foundation—ensuring that the proposed metrics not only perform consistently within and across subjects but also hold the potential to operate meaningfully in clinical or cognitive scenarios, where specificity and sensitivity are not optional features, but essential guarantees.

2.7 Contributions to the state of the art

The contributions made during this thesis revolve around most of the previous topics: Some of them serve as basis, whereas some others serve as beacons to check the correctness of newer developments. The next chapter describes the contributions made in this thesis to the state-of-the-art methods and techniques that allow interpretability of diffusion-weighted signals for both research and clinical environments.

The first contribution consists in the development of two EAP-based metrics, PA and NG, for the MiSFIT framework; the second assesses the influence of thermal noise in biophysical models that employ rotational invariants derived from spherical harmonics; and the third presents a

systematic analysis of the reliability of diffusion MRI techniques based on signal representations. All three contributions share a common thread: they examine the reliability of derived metrics —with the first two using it as a means to validate and refine modeling approaches, and the third focusing on reliability as its primary objective.

A concept is a brick. It can be used to build a courthouse of reason. Or it can be thrown through the window.

— Deleuze (1994)

3

Contribution: Efficient estimation of PA and NG for MiSFIT

Abstract

We seek to reformulate the so-called Propagator Anisotropy and Non-Gaussianity (PA and NG, see Section 2.4.2), originally conceived for the Mean Apparent Propagator diffusion MRI (MAP-MRI), to the Micro-Structure adaptive convolution kernels and dual Fourier Integral Transforms (MiSFIT). These measures describe relevant normalized features of the EAP

First, the indices, which are defined as the EAP's dissimilarity from an isotropic (PA) or a Gaussian (NG) one, are analytically reformulated within the MiSFIT framework. Then a comparison between the resulting maps is drawn by means of a visual analysis, a quantitative assessment via numerical simulations, a test-retest study across the MICRA dataset (6 subjects scanned five times) and, finally, a computational time evaluation.

Findings illustrate the visual similarity between the indices computed with either technique. Evaluation against synthetic ground truth data, however, demonstrates MiSFIT's improved accuracy. In addition, the test–retest study reveals MiSFIT's higher degree of reliability in most of white matter regions. Finally, the computational time evaluation shows MiSFIT's time reduction up to two orders of magnitude. Despite being a direct development of the MAP-MRI representation, the PA and the NG can be reliably and efficiently computed within MiSFIT's framework.

3.1 Introduction and Purpose

In line with this thesis objectives, we seek to develop and implement two metrics for the MiSFIT framework (see Section 2.4.3): the Propagator anisotropy and non-Gaussianity (see Section 2.4.2). The metrics were originally proposed in the MAP-MRI framework (see Section 2.4.3) and describe relevant normalized features of the EAP. For reader convenience, the main principles and derivation of PA and NG metrics are summarized here, despite having been discussed in Chapter 2.

Reconstructing the 3D EAP allows to compute several scalar indices that embed useful descriptors of WM. Among them RTOP, RTPP and RTAP, as well as PA or NG are the most often used (Fick et al., 2016a; Ning et al., 2015; Özarslan et al., 2013). In this chapter we focus on PA and NG, which come to generalize the popular Fractional Anisotropy (FA) and diffusion Kurtosis.

As previously mentioned, PA is defined in Özarslan et al. (2013) as the distance from the propagator $P(\mathbf{R})$ to its isotropic equivalent. By means of Parseval's theorem, it can also be defined as the distance from the attenuation signal $E(\mathbf{q})$ (the diffusion signal characterized by wave vector $\mathbf{q}$ over the unweighted T2 baseline) to its closest isotropic counterpart, $O(\mathbf{q})$:

$$\mathrm{PA} = \gamma\left(\sin(\angle(E(\mathbf{q}), O(\mathbf{q}))), \epsilon\right) \in [0, 1], \tag{3.1}$$

where $\angle(E(\mathbf{q}), O(\mathbf{q}))$ denotes the angular distance, defined implicitly through:

$$\cos(\angle(E(\mathbf{q}), O(\mathbf{q}))) = \frac{\langle E(\mathbf{q}), O(\mathbf{q})\rangle}{\parallel E(\mathbf{q}) \parallel \parallel O(\mathbf{q}) \parallel}, \tag{3.2}$$

and:

$$\gamma(t, \epsilon) = \frac{t^{3\epsilon}}{1 - 3t^{\epsilon} + 3t^{2\epsilon}}, \tag{3.3}$$

for $\epsilon = 0.4$, which stands for a contrast enhancement of PA within the normalized range $[0, 1]$. PA has shown the ability to characterize morphological and cytoarchitectural attributes, even in Gray Matter (GM) regions where the FA is non-informative (Galazzo et al., 2018). Yet, PA offers a more accurate assessment of the anisotropic behavior in crossing fibers regions. Within clinical setups, PA has shown a great potential in the analysis of longitudinal changes within subjects (Avram et al., 2016), the characterization of cognitive impairment after traumatic brain injury (TBI) (Kraus et al., 2007), impaired social cognition in autism (Barnea-Goraly et al., 2004), or age-dependent neuronal demise in transgenic Alzheimer rats (Fick et al., 2017).

In turn, NG is defined as the distance from $E(\mathbf{q})$ to its closest Gaussian representation, $G(\mathbf{q})$ (Özarslan et al., 2013):

$$\mathrm{NG} = \sin\left(\angle(E(\mathbf{q}), G(\mathbf{q}))\right). \tag{3.4}$$

Though the clinical applicability of NG has not been as thoroughly tested as that of PA, it has been lately proven useful at distinguishing grade II from grade III and IV gliomas (Wang et al., 2021b), relevant for the noninvasive preoperative evaluation of tumour pathological grading. Some additional studies on axonal loss and demyelination (Song et al., 2005), as well as head and neck cancer (Jansen et al., 2010; Yuan et al., 2014) are also available.

Though they can be computed resorting to other estimation techniques (Ning et al., 2015), PA and NG naturally arise from MAP-MRI. Therein, the NG can be easily computed from the energy of the non-DC components, since MAP-MRI develops the diffusion signal by successively refining a Gaussian model. Conversely, the PA is related to the non-DC components of the isotropic, non-voxel adaptive version of MAP-MRI, also known as 3D-SHORE.

In this chapter, we aim at formulating and evaluating both PA and NG for the newly developed Micro-Structure adaptive convolution kernels and dual Fourier Integral Transforms (MiSFIT, Tristán-Vega and Aja-Fernández, 2021), both of them from the very same signal representation. By taking advantage of the computational efficiency of MISFIT, we expect the already-demonstrated higher accuracy in the RTxP computation with respect to MAPL (the *de facto* standard, regularized version of MAP-MRI developed by Fick et al., 2016a) to also be translated for PA and NG. In addition, we also expect this accuracy to result in a higher reliability of the measures, which together with the critical reduction of the computational time, qualify MiSFIT's PA and NG to be incorporated not only in research studies, but also in clinical settings.

3.2 Theory

3.2.1 MiSFIT's signal representation

As already introduced in Section 2.4.3, MiSFIT's composite representation comprises the aggregate of a free-diffusing isotropic component and a semi-parametric component that accounts for the partial volume fraction $f \in [0,1]$ of constrained diffusion, as in Eq. (2.47). The constrained diffusion signal $E(\mathbf{q})$ is drawn as the spherical convolution of a parametric kernel, defined by the longitudinal $\lambda_\parallel$ and transverse $\lambda_\perp$ diffusivities, with a non-parametric fODF, Φ, as in Eq. (2.46). By representing the ODF in the basis of SH, MiSFIT lastly represents the attenuation signal as:

$$E(q\mathbf{u}) = \sum_{\substack{l=0 \\ l \text{ even}}}^{L} \sum_{m=-l}^{l} \widehat{e}_l^{\,\mathrm{ODF}}(q)\phi_l^m Y_l^m(\mathbf{u}), \tag{3.5}$$

where $Y_l^m(\mathbf{u})$ are the (even) SH functions, ϕ_l^m stand for the SH coefficients of the ODF, and $\widehat{e}_l^{\,\mathrm{ODF}}(q)$ are $\lambda_\parallel$ and $\lambda_\perp$-dependent convolution multipliers (for further details, see Tristán-Vega and Aja-Fernández, 2021). A central part to our developments, as can be easily deduced from Eqs. (3.1) and (3.4), is the computation of scalar products between SH-spanned functions like in Eq. (3.5). We prove in Appendix A that their calculation relies on the evaluation of the following integral, $\mathcal{I}_l(\rho_\lambda)$, involving the Legendre Polynomials $P_l(x_i)$:

$$\mathcal{I}_l(\rho_\lambda) = \iint_{-1}^{1} \frac{P_l(x_1)P_l(x_2)}{(2\rho_\lambda + x_1^2 + x_2^2)^{3/2}} dx_1 dx_2, \tag{3.6}$$

for $\rho_\lambda = \lambda_\perp/(\lambda_\parallel - \lambda_\perp)$. These integrals do not admit a closed form. However, since only the first few even orders of l are needed, they can be pre-computed for a wide range of ρ_λ and with an accuracy up to the numerical precision of the machine. Fig. 3.1 shows their values for the first few l.

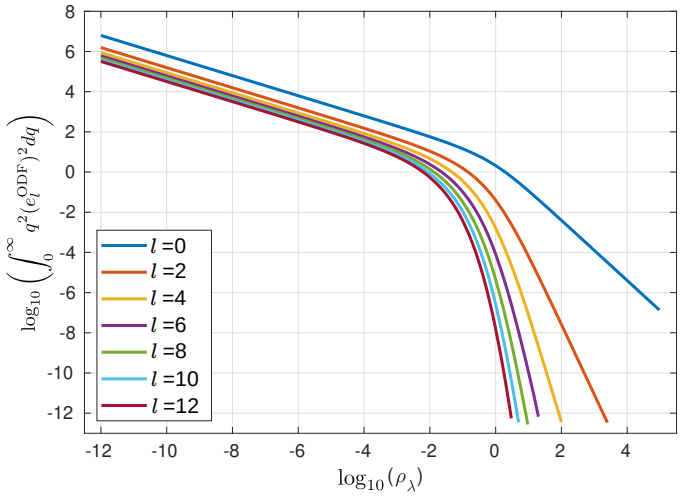


Figure 3.1: Numerical values of the integrals $\mathcal{I}_l(\rho_\lambda)$ as described in Eq. (3.6).

3.2.2 PA for MiSFIT's composite signal

From the definition in Eq. (3.1), we develop into:

$$\sin\left(\angle({}_cE(\mathbf{q}),\ {}_cO(\mathbf{q}))\right) = \sqrt{1-\cos^2\left(\angle({}_cE(\mathbf{q}),\ {}_cO(\mathbf{q}))\right)} \tag{3.7}$$

We demonstrate in Appendix A.0.1 that the squared cosine in the last equation leads to:

$$\cos^2\left(\angle({}_cE(\mathbf{q}),\ {}_cO(\mathbf{q}))\right) = \frac{\|{}_cO(q)\|^2}{\|{}_cE(q\mathbf{u})\|^2} = \frac{(1-f)^2O_{\mathrm{iso}} + f^2O_{\mathrm{ani}} + f(1-f)O_{\mathrm{mix}}}{(1-f)^2E_{\mathrm{iso}} + f^2E_{\mathrm{ani}} + f(1-f)E_{\mathrm{mix}}}, \tag{3.8}$$

where the isotropic terms (E_{iso}, O_{iso}), the anisotropic terms (E_{ani}, O_{ani}), and the mixed terms (E_{mix}, O_{mix}) come from the composite representation (i.e. isotropic plus anisotropic parts) in Eq. (2.47):

$$\begin{aligned}
O_{\mathrm{ani}} &= (\phi_0^0)^2\pi\delta_\lambda^{-3/2}\mathcal{I}_0(\rho_\lambda);\\
E_{\mathrm{ani}} &= \sum_{l,m}(\phi_l^m)^2\pi\delta_\lambda^{-3/2}\mathcal{I}_l(\rho_\lambda);\\
E_{\mathrm{iso}} &= O_{\mathrm{iso}} = (2D_0)^{-3/2};\\
E_{\mathrm{mix}} &= O_{\mathrm{mix}} = 4\sqrt{\pi}\phi_0^0\Big((D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}\Big)^{-1},
\end{aligned} \tag{3.9}$$

for $\delta_\lambda = \lambda_\parallel - \lambda_\perp$. Therefore, the PA finally reads:

$$\mathrm{PA} = \gamma\left(\sqrt{1 - \frac{\frac{(1-f)^2}{(2D_0)^{3/2}} + \frac{4\sqrt{\pi}f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}}\phi_0^0 + f^2(\phi_0^0)^2\pi\delta_\lambda^{-3/2}\mathcal{I}_0(\rho_\lambda)}{\frac{(1-f)^2}{(2D_0)^{3/2}} + \frac{4\sqrt{\pi}f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}}\phi_0^0 + f^2\sum_{l,m}(\phi_l^m)^2\pi\delta_\lambda^{-3/2}\mathcal{I}_l(\rho_\lambda)}}, \epsilon\right). \tag{3.10}$$

Noticeably, removing the isotropic compartment (i.e. taking $f = 1$) results in a much simpler expression:

$$\mathrm{PA} = \gamma\left(\sqrt{1 - \frac{O_{\mathrm{ani}}}{E_{\mathrm{ani}}}}, \epsilon\right) = \gamma\left(\sqrt{1 - \frac{(\phi_0^0)^2 \pi \delta_\lambda^{-3/2} \mathcal{I}_0(\rho_\lambda)}{\sum_{l,m} (\phi_l^m)^2 \pi \delta_\lambda^{-3/2} \mathcal{I}_l(\rho_\lambda)}}, \epsilon\right). \tag{3.11}$$

Conversely, fitting a voxel with $f = 0$ (i.e. a free-water voxel), results in both $\|{}_cO(q)\|^2$ and $\|{}_cE(q\mathbf{u})\|^2$ being equal. Hence, the PA as defined in Eq. (3.7), will be 0 as expected.

3.2.3 NG for MiSFIT's composite signal

Now, from the definition in Eq. (3.4):

$$\sin(\angle({}_cE(\mathbf{q}),\ {}_cG(\mathbf{q}, \widehat{\mathbf{D}}))) = \sqrt{1 - \left(\frac{\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q}, \widehat{\mathbf{D}})\rangle}{\|{}_cE(\mathbf{q})\| \|{}_cG(\mathbf{q}, \widehat{\mathbf{D}})\|}\right)^2}, \tag{3.12}$$

where ${}_cG(\mathbf{q}, \widehat{\mathbf{D}})$ is the closest Gaussian propagator. In the MiSFIT approach, the equivalent Gaussian propagator is estimated by fitting a tensor to the EAP-based reconstruction of the attenuation signal (for b-values < 2000 s/mm^2). The reason behind this is that computing the actual closest Gaussian propagator (i.e. that which results in the smallest possible Mean Squared Difference, MSD) is not trivial and would require solving a calculus of variations problem that could easily maim the computational efficiency nature upon which MiSFIT is build. Hence, the problem is solved in the logarithmic domain of $E(\mathbf{q})$, to make it convex and permit a closed-form solution, while keeping the computationally-efficient feature that makes MiSFIT desirable for specific settings. Thus, from now on, our implementation of the *closest* Gaussian propagator will be referred to as "DTI-like propagator". More information about the Gaussian propagator and the method chosen to compute it is included in the Appendix C. We demonstrate in Appendix B that the quotient in Eq. (3.12) equals:

$$\left(\frac{\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q}, \widehat{\mathbf{D}})\rangle}{\|{}_cE(\mathbf{q})\| \|{}_cG(\mathbf{q}, \widehat{\mathbf{D}})\|}\right)^2 = \frac{\left(\frac{1-f}{\sqrt{|\widehat{\mathbf{D}}+\mathbf{D}_0|}} + f \sum_{l,m} \phi_l^m \xi_l^m\right)^2 \sqrt{8|\widehat{\mathbf{D}}|}}{(1-f)^2 E_{\mathrm{iso}} + f^2 E_{\mathrm{ani}} + f(1-f) E_{\mathrm{mix}}}, \tag{3.13}$$

where $\widehat{\mathbf{D}}$ is the DTI-like propagator; $\mathbf{D}_0 = D_0 I_3$ is D_0 times the identity matrix; $|\cdot|$ stands for the determinant; ξ_l^m are the SH coefficients of a spherical function defined in Appendix B that depends on $\widehat{\mathbf{D}}$, $\lambda_\parallel$ and $\lambda_\perp$. Finally, the NG reads:

$$\mathrm{NG} = \sqrt{1 - \frac{\left(\frac{1-f}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + f \sum_{l,m} \phi_l^m \xi_l^m\right)^2 \sqrt{8|\widehat{\mathbf{D}}|}}{\frac{(1-f)^2}{(2D_0)^{3/2}} + f^2 \sum_{l,m} (\phi_l^m)^2 \pi \delta_\lambda^{-3/2} \mathcal{I}_l(\rho_\lambda) + \frac{4\sqrt{\pi} f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}} \phi_0^0}}. \tag{3.14}$$

Again, evaluating the NG for $f = 1$ yields to a simpler expression:

$$\mathrm{NG} = \sqrt{1 - \left(8\delta_\lambda^3 |\widehat{\mathbf{D}}|\right)^{1/2} \frac{\left(\sum_{l,m} \phi_l^m \xi_l^m\right)^2}{\pi \sum_{l,m} (\phi_l^m)^2 \mathcal{I}_l(\rho_\lambda)}}. \tag{3.15}$$

For values of f other than 0, finding the Gaussian counterpart to the MiSFIT-estimated EAP is not equally easy. This is not an issue within the (anisotropic) MAP-MRI framework, since it represents the diffusion signal as a series of orthogonal cumulants, being the DTI estimation the first one, i.e.:

$$E_{\mathrm{MAP}}(\mathbf{q}) = E_{\mathrm{DTI}}(\mathbf{q}) + E_{\mathrm{MAP}\setminus\mathrm{DTI}}(\mathbf{q}), \tag{3.16}$$

so that the closest Gaussian is just the first addend. As already explained in the beginning of the section, MiSFIT can only compute a "DTI-like" propagator in order not to maim its computational efficiency.

3.3 Materials

3.3.1 In vivo evaluation

In-vivo validation has been carried out by using some of the publicly available datasets introduced earlier in Section 1.4. Specifically, we have used one subject from the MGH-HCP dataset (MGH 1007) and the whole MICRA dataset to evaluate the metrics' reliability.

3.3.2 Numerical simulations

Numeric comparisons over ground-truth data are based on the methodology originally proposed in (Tristán-Vega and Aja-Fernández, 2021). A microstructure model is estimated at representative regions of the WM using NODDI (previously introduced in Section 2.4.4). Afterwards, a statistical model is built upon the estimated parameters, and further used to draw random samples that are fed to the forward NODDI model to generate synthetic samples simulating 1, 2, or 3 crossing fibers at will with a known Peak Signal to Noise Ratio (PSNR):

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{I_{\max}^2}{\mathrm{MSE}} \right) [\mathrm{dB}] \tag{3.17}$$

where $I_{\max}^2$ is the maximum possible intensity value of the image, and MSE is the mean squared error between the reference image and the reconstructed one.

As long as the generative model can be sampled for any gradient direction and b-value, ground-truth values are easily obtained for any dMRI measure with arbitrary precision by numerical integration. See Tristán-Vega and Aja-Fernández (2021) for further details on this methodology. While the ground truth for the PA is somehow trivial via the SH representation, determining the ground truth of NG implies solving a problem of calculus of variations. So this is transformed

into a Least Squares optimization one by minimizing the squared residuals of a preset number of $\mathbf{q}$-space samples and then solved in the logarithmic domain of $E(\mathbf{q})$, as explained in Section 2.3.

3.4 Methods

3.4.1 Implementation details

The fitting procedure for parameters f, $\lambda_{\parallel}$ and $\lambda_{\perp}$ in Eqs. (2.47) and (2.46), and the computation of the ODF's SH coefficients, is detailed in Tristán-Vega and Aja-Fernández (2021). We have used $L = 8$ as the maximum order for SH expansions, and empirically set $\mu = 1.2 \cdot 10^{-4}$ as the regularization parameter described therein to avoid singular convolution kernels. To compute $\mathcal{I}_l(\rho_\lambda)$, we have resorted to linear interpolation in the logarithmic domain from the values depicted in Fig. 3.1. The whole MiSFIT framework, including the newly introduced PA and NG, was coded in Matlab R2019b and is available for download[1] (Tristán-Vega, A. and Aja-Fernández, S. and París, G., 2022).

The computation of PA and NG with MiSFIT is validated by comparing them with the *de facto* standard in the related literature, i.e. the MAPL as described in Fick et al. (2016a). We have used the implementation in the DIPY package[2] under Python 3, though the actual code allowing the computation of the PA within the anisotropic MAP-MRI reconstruction was kindly provided by the authors on demand. We use MAPL with positivity constraints and cross-validation for setting the Laplacian penalty term. The maximum order for the basis functions was set to 6.

In both cases, and unless otherwise noticed, both PA and NG have been set up using the entirety of the acquisition's shells.

3.4.2 Ground-truth based evaluation

Numeric comparisons over ground-truth data are based on the methodology originally proposed in (Tristán-Vega and Aja-Fernández, 2021). A micro-structure model is estimated at representative regions of the WM using NODDI (Zhang et al., 2012). Afterwards, a statistical model is built upon the estimated parameters, and further used to draw random samples that are fed to the forward NODDI model to generate synthetic samples simulating 1, 2, or 3 crossing fibers at will with a known Peak Signal to Noise Ratio (PSNR). As long as the generative model can be sampled for any gradient direction and b-value, ground-truth values are easily obtained for any dMRI measure with arbitrary precision by numerical integration. See Tristán-Vega and Aja-Fernández (2021) for further details on this methodology. While the ground truth for the PA is somehow trivial via the SH representation, determining the ground truth of NG implies solving a problem of calculus of variations. So this is transformed into a Least Squares (LS) optimization one by minimizing the squared residuals of a preset number of $\mathbf{q}$-space samples and then solved in the

[1] http://www.lpi.tel.uva.es/dmrilab
[2] https://dipy.org/

logarithmic domain of $E(\mathbf{q})$, as explained in Section 3.2.3.

3.4.3 Reliability Study

The repeated acquisitions within MICRA dataset have been used to assess the repeatability of the computed measures (i.e., inter-session variability) and the separability they provide (i.e. inter-subject differences). These complementary properties together characterize the reliability of each method (MiSFIT/MAPL) and measure (PA/NG) (Zuo et al., 2019). Note that DTI's FA and DKI's mean Kurtosis have been included in the table for a wider comparison.

The design of the corresponding experiment is as follows: First, ROIs were back-projected by registering the JHU atlas' labels in MNI152 space (Mori et al., 2008) into each subjects' space by means of linear plus nonlinear registration (FSL's Smith et al., 2004 FLIRT and FNIRT Jenkinson and Smith, 2001; Jenkinson et al., 2002, respectively) of the subject's FA to the JHU atlas, followed by the application of the inverse warping to the JHU labels. An eroded mask of the regions with a kernel of size $2 \times 2 \times 2$ was then computed to palliate the effects of a possible misalignment in the registration, removing tissue regions potentially affected by a partial volume effect, followed by the removal of those values for which the FA in the region yielded outliers (defined as those values falling 1.5 times outside the interquartile range). More information about the outlier rejection procedure is included in the Appendix C. Finally, for each of the 12 ROIs in the subject space (see Fig. 1.2), one single-valued representative of each measure was computed as the median value. Note that ROIs in both right and left hemisphere have been merged together into a single region. Repeatability was then computed as the mean across subjects of the variances across sessions, while separability was computed as the variance across subjects of the means across sessions. A Figure of Merit (FOM) was lastly defined as the separability over repeatability ratio. Owing to the limited size of the database, a 200-runs bootstrap analysis over population's subsets was carried out to assess the confidence on such FOM, expressed as its coefficient of variation (CV):

$$\mathrm{CV}_{\mathrm{FOM}} = \left(1 + \frac{1}{4\mathrm{n}}\right)\frac{\mathrm{s}}{\bar{\mathrm{x}}}, \tag{3.18}$$

with $\bar{x}/s$ the mean/standard deviation over bootstrapped samples, and n the number of bootstrap samples, used in the first term to attain unbiased estimates of the CV.

3.4.4 Computational Time Evaluation

One main advantage of MiSFIT, as compared to MAP-MRI/MAPL, is its computational efficiency, which allows to estimate the EAP two orders of magnitude faster. Nonetheless, the computation of scalar measures from the reconstructed EAP can result in non-negligible computational overloads: while the computation of NG is straightforward after the non-isotropic MAP-MRI signal representation, this is not the case for PA, which requires non-trivial extra calculations.

For this reason, it makes sense to compare how long MiSFIT and MAPL take to compute (1) the signal representation, (2) the PA and (3) the NG. The experiment is carried out with

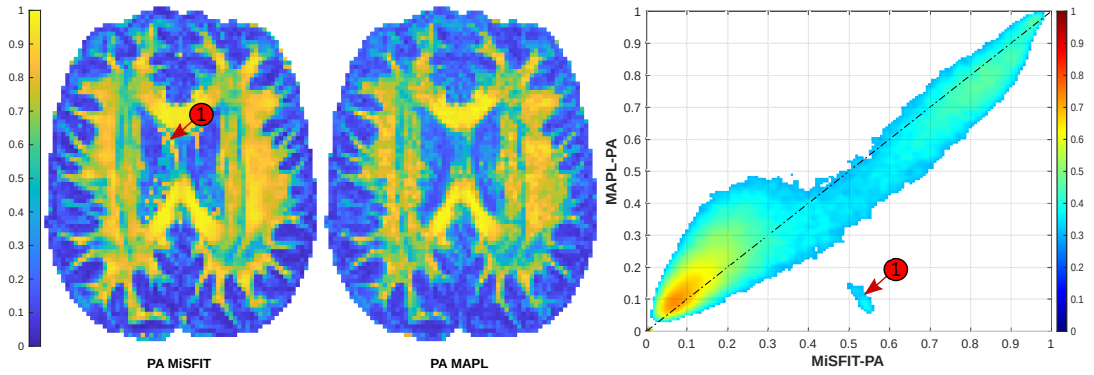


Figure 3.2: Comparison of the PA maps obtained with either MiSFIT and MAPL for the third session of the fourth MICRA's subject. Left: visual assessment within the whole range of variation $[0, 1]$. Right: joint 2-D histograms (logarithmic) of MAPL's values vs. MiSFIT's. The high-valued islands within MiSFIT (tag numbered '1') correspond to ambiguous configurations of the convolutional representation of MiSFIT.

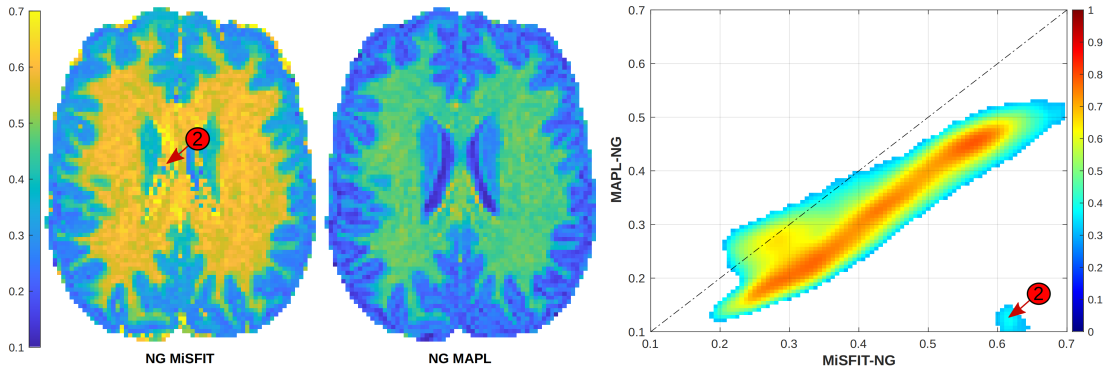


Figure 3.3: Comparison of the NG maps obtained with either MiSFIT and MAPL for the third session of the fourth MICRA's subject. Left: visual assessment within the cropped range $[0.1, 0.7]$. Right: joint 2-D histograms (logarithmic) of MAPL's values vs. MiSFIT's. The outliers tagged with '2' correspond to ambiguous configurations of the convolutional representation of MiSFIT.

the aforementioned Matlab's (for MiSFIT) and Python DIPY's (for MAPL) implementations running in a quad-core Intel© CoreTM i7-6500U processor with 8GB RAM.

3.5 Results

3.5.1 Visual assessment

Figs. 3.2 and 3.3 compare the computations of PA and NG by MiSFIT and MAPL for the randomly chosen third session of the fourth MICRA's subject. With regard to PA, MAPL presents a noisier behavior (Fig. 3.2, left), specially in those areas with lower anisotropy values—except the ventricles, which MiSFIT clearly fails to delineate.

This can also be observed in areas with higher anisotropy, resulting, for example, in some visible discontinuities in the *external capsule* (EC). In comparison, MiSFIT produces a more saturated map, clearly defining the major WM fiber tracts and their limits, making even more distinguishable the anisotropy of diffusion processes in some brain regions. Nonetheless, this contrast saturation does not imply, *per se*, a more accurate result. The comparison can also be analysed in terms of the 2D joint histogram (Fig. 3.2, right), which shows a linear correlation between both metrics, only disturbed for lower PA values, where the histogram broadens. As explained in Tristán-Vega and Aja-Fernández (2021), MiSFIT's estimation produces artifacts (tagged with number 1) due to ambiguities in highly-isotropic zones (such as the ventricles or the GM), where the optimizer has to decide, in Eqs. (2.47) and (2.46), whether $f = 1$ or $\lambda_\perp \simeq \lambda_\parallel$.

For NG, both MiSFIT and MAPL provide outcomes similar to those found with PA (Fig. 3.3, left), with the exception of the overall negative bias MAPL introduces with respect to MiSFIT. This shift is more clearly visible in Fig. 3.3, right: MAPL's values are strongly linearly correlated with MiSFIT's, but placed along a line with slope less than 1 and negative bias. Once again, MiSFIT fails to delineate the ventricles due to the ambiguity in the representation, yielding to the outlier tagged as 2 in the histogram.

3.5.2 Quantitative analysis based on ground-truth

The numerical assessment of the accuracy of each method is based on the ground-truth data described in Section 3.4.2: Figs. 3.4 and 3.5 show the similarity of PA and NG, as computed with each of MAPL and MiSFIT, with true values of PA and NG, respectively. Joint 2D histograms have been computed for different PSNR values (16, 32 and 64).

Fig. 3.6 shows the rMSE for PA and NG estimates from both MAPL and MiSFIT as a function of the ground-truth RTOP value, as in Tristán-Vega and Aja-Fernández (2021). This choice is based on the definition of RTOP as the integral of the attenuation signal in the whole $\mathbf{q}$-space (see Eq. (2.20)) which directly relates RTOP to the bandwidth of the EAP.

For PSNR=64, MiSFIT results in more accurate PA maps than MAPL from every possible point of view: Not only depicting smaller variability across the dynamic range, but also resulting in mean values closer to the ground truth. When conditions worsen, e.g. PSNR=16, MiSFIT obtains poorer results, overestimating the anisotropy in highly-isotropic regions. These PA estimates, however, get more accurate the more anisotropic the region is.

Regarding the relative MSE (top of Fig. 3.6), MiSFIT yields to smaller errors, consistently outperforming MAPL values for any given bandwidth, PSNR or fiber bundle configuration. This latter variable is responsible of the poorer behavior of MiSFIT in Fig. 3.4 with PSNR=16. Arguably, both PA estimates worsen as the number of crossing fibers increases, which may be caused by the representation's difficulties when computing the isotropic EAP counterpart from bundles that may have different diffusion properties. Notice that, despite the results being worse for the third bundle configuration, they still outperform MAPL's.

The discussion for the NG is not equally good to MiSFIT. First of all, the two left-most

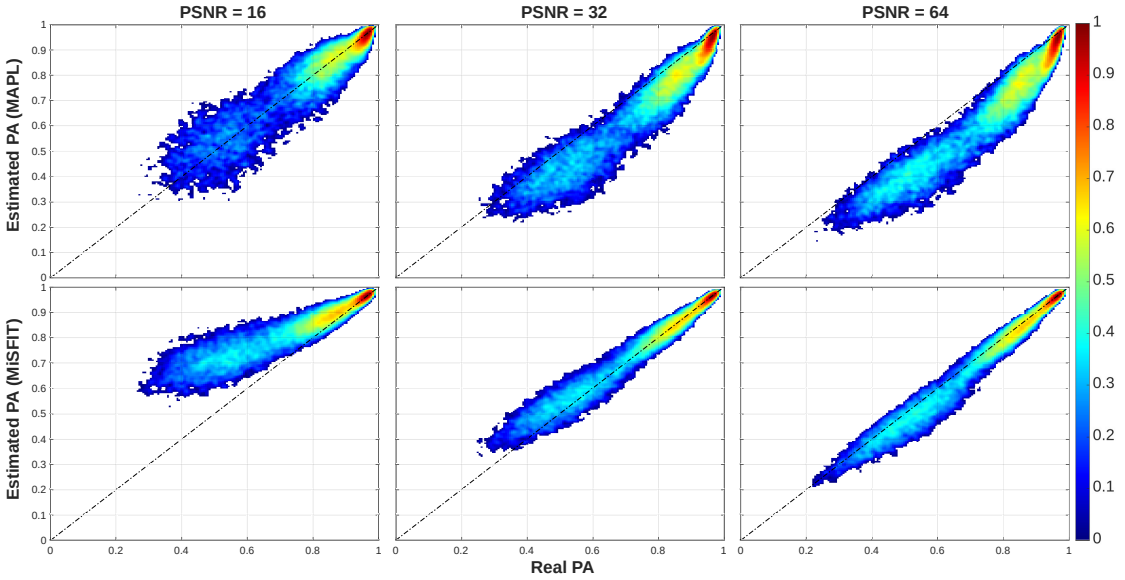


Figure 3.4: 2D Joint histogram of the estimated PA (y-axis) against synthetic ground truth values (x-axis) for different PSNR (16 left, 32 middle, 64 right). There exists a significant linear correlation between the estimated and real values, with MiSFIT being clearly more accurate and less variable than MAPL for PSNR > 32.

columns of Fig. 3.5 (PSNR 16 and 32, respectively) show bigger variability across the measure's dynamic range, even for those values with means closer to the ground truth than those reported by MAPL. With respect to MiSFIT, there are multiple sources of error that end up stacking one to each other, the main one is driven by the noise—which causes the metric to be underestimated, specially in regions with high-Gaussianity behavior—but also by the estimated EAP—which affects the non-Gaussian regions estimates—and the DTI-estimated equivalent Gaussian EAP—which is greatly palliated as the fiber crossings increases.

On the other hand, the rMSE depicted from MiSFIT's NG (bottom of Fig. 3.6) seems to be dominated by the EAP reconstruction error, which results in MiSFIT underestimating the NG. This behavior is consistent with the one explained in Tristán-Vega and Aja-Fernández (2021), Fig. 8. Interestingly, the rMSE increases when dealing with single-fiber configurations. This may be caused by MiSFIT's construction: The convolution of the Gaussian kernel with a very prolate but not-completely-sampled ODF (i.e. possibly non-Gaussian) results in Gaussian-like distributions, which yields to underestimated NG results. Alternatively, when adding fibers onto the configuration, the reconstructed EAP function gets smoother over the surface of the sphere, thus getting more accurately defined by the sampling scheme, which in turn leads to more precise estimations of the original EAP. This can also be appreciated in the bottom of Fig. 3.6.

In order to relate the numerically-obtained results with in-vivo images, the PSNR has been computed for a subject of the MICRA dataset by dividing the filtered (denoised) image by the estimated noise level. Both filtered image and noise level have been computed by using dwidenoise (MRtrix3, Tournier et al., 2019) on the raw images. Finally, a white matter mask has been crafted by merging all the (previously-eroded) JHU WM labels. The resulting PSNR is 28.12, a usual result for good-quality acquisitions such as MICRA. This result indicates that the in-vivo images

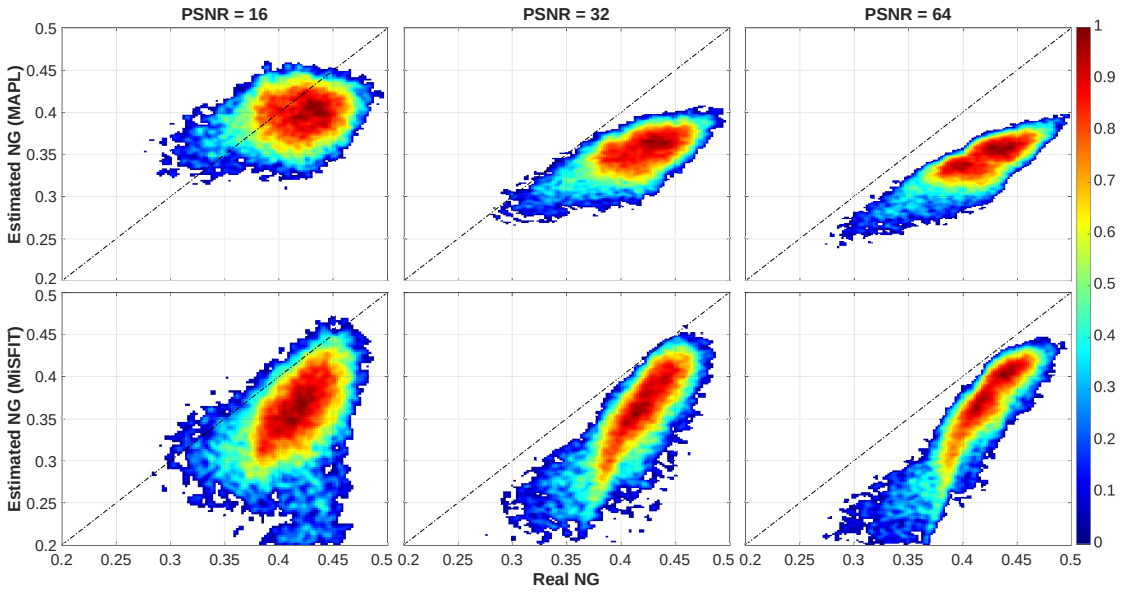


Figure 3.5: 2D Joint histogram of the estimated NG (y-axis) against synthetic ground truth values (x-axis) for different PSNR (16 left, 32 middle, 64 right). A linear correlation can be appreciated in both PSNR = 64 maps. MiSFIT tends to be more accurate in the high values, while MAPL estimation gets unbiased for lower ones, as well as for PSNR = 16.

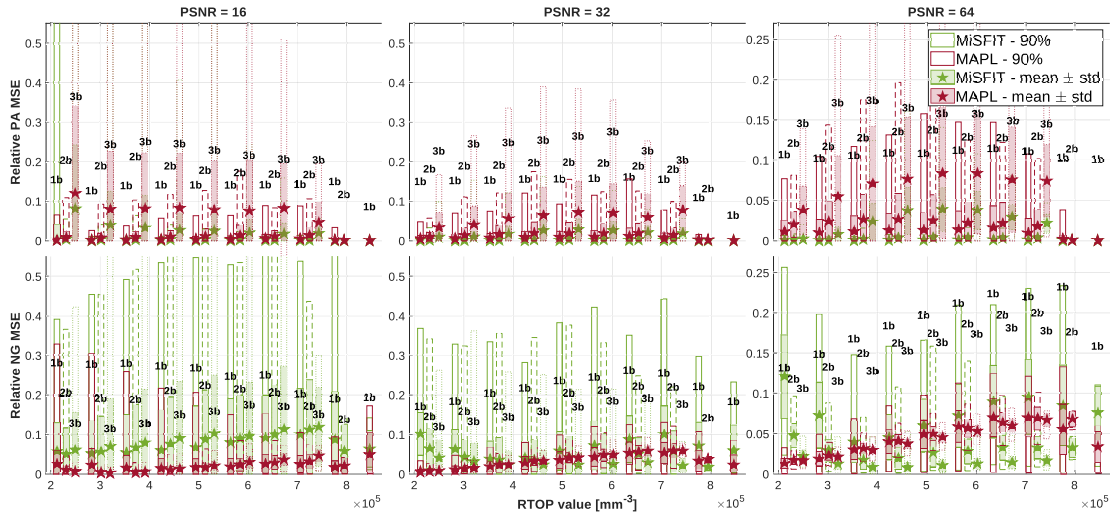


Figure 3.6: Estimated relative mean squared error (rMSE) of PA (top) and NG (bottom) as a function of signal's bandwidth (i.e. RTOP), for different PSNR (16 left, 32 middle and 64 right) and 1, 2 or 3 fiber crossing configuration. Missing bars correspond to those fiber bundle configurations for which PA, NG or RTOP values could not be computed. Statistical properties displayed: mean (star-shaped symbol), standard deviation (full-colored box), 90% confidence interval (empty box).

in Fig. 3.2 and 3.3 (from MICRA dataset) can be related to the middle column in Fig. 3.4, 3.5 and 3.6. In this PSNR range, the MiSFIT's PA portrays a much less variable estimation than MAPL's. MiSFIT tends to overestimate the mid-low PA values, while MAPL tends to subestimate them. MiSFIT's rMSE depicts a better accuracy in either bundle configuration and throughout the entire bandwidth range. Regarding the NG, the variance of the estimation increases significantly, yielding to subestimated results for both MiSFIT and MAPL. Concerning

Region	PA		FA	NG		MK
	MAPL	MiSFIT	DTI	MAPL	MiSFIT	DKI
MCP	0.113 / 3.54	0.064 / 2.09	0.184 / 2.64	0.035 / 2.66	0.136 / 3.82	0.039 / 1.44
GCC	0.067 / 13.49	0.151 / 4.59	0.192 / 13.20	0.120 / 3.59	0.070 / 19.74	0.072 / 8.58
BCC	0.057 / 4.66	0.061 / 4.31	0.077 / 7.42	0.087 / 1.98	0.016 / 4.37	0.010 / 11.83
SCC	0.006 / 14.7	0.026 / 10.86	0.018 / 33.42	0.049 / 5.77	0.003 / 40.94	0.025 / 17.46
ALIC	0.078 / 3.04	0.148 / 5.25	0.149 / 2.57	0.056 / 4.00	0.045 / 1.40	0.035 / 13.01
PLIC	0.045 / 6.55	0.049 / 3.72	0.059 / 7.07	0.030 / 3.31	0.015 / 8.43	0.021 / 9.43
ACR	0.152 / 3.42	0.024 / 3.39	0.033 / 11.33	0.026 / 4.55	0.036 / 14.30	0.024 / 14.37
SCR	0.232 / 3.38	0.019 / 3.21	0.067 / 24.30	0.023 / 2.26	0.040 / 35.15	0.017 / 5.32
PCR	0.157 / 3.14	0.034 / 2.00	0.349 / 1.20	0.021 / 2.34	0.152 / 2.12	0.037 / 6.70
PTR	0.192 / 5.05	0.023 / 5.98	0.252 / 2.06	0.036 / 3.42	0.083 / 2.03	0.016 / 18.82
EC	0.729 / 3.39	0.035 / 17.30	0.352 / 5.35	0.075 / 7.1	0.491 / 5.47	0.022 / 18.84

Table 3.1: Repeatability coefficient of variation (‰) over reliability (FOM) values for each of the regions, measures (PA, NG) and methods (MiSFIT, MAPL). DTI's FA and DKI's MK have been added for a wider comparison. Results indicate a good ratio between inter-subject and the intra-subject variability for both measures. On average, MiSFIT's NG obtain slightly better results than MAPL's, e.g. in the *corpus callosum*, while resulting in less reliable PA maps than those obtained with MAPL or DTI's FA.

Region	PA		FA	NG		MK
	MAPL	MiSFIT	DTI	MAPL	MiSFIT	DKI
MCP	0.596	0.658	0.613	0.494	0.225	0.325
GCC	0.385	0.230	0.243	0.318	0.303	0.199
BCC	0.524	0.557	0.588	0.310	0.408	0.437
SCC	0.488	0.431	0.516	0.324	0.293	0.486
ALIC	0.331	0.368	0.332	0.351	0.322	0.322
PLIC	0.719	0.399	0.553	0.409	0.342	0.203
ACR	0.359	0.275	0.416	0.488	0.649	0.323
SCR	0.679	0.522	0.491	0.502	0.488	0.430
PCR	0.256	0.559	0.334	0.378	0.382	0.326
PTR	0.170	0.304	0.333	0.388	0.513	0.586
EC	0.659	0.328	0.276	0.305	0.326	0.201

Table 3.2: Bootstrap-derived CV of the FOM values for each of the regions, measures (PA, NG) and methods (MiSFIT, MAPL), together with the DTI's FA and the DKI's MK, for a wider comparison. The bootstrap was computed by doing 200 runs.

the rMSE, for the lower bandwidth ranges the MiSFIT's NG results in slightly less accurate estimations, which are improved in the mid-to-high bandwidth ranges.

3.5.3 Reliability Study

The results corresponding to the experiment described in Section 3.4.3 are summarized in Tables 3.1 and 3.2. Table 3.1 reports the FOM—the reliability measurement—together with the CV of the repeatability values. Table 3.2 shows the FOM's bootstrapped CV.

The first thing to notice from Table 3.1 is that all FOM values are greater than one, meaning that, for any given ROI, the inter-subject variability of the metrics is greater than the intra-subject variability. Regarding the PA, both MiSFIT and MAPL exhibit the same tendency, yielding to reliability values in the same order of magnitude, with the exception of the GCC where MAPL depicts much greater results than MiSFIT, and the EC where MiSFIT depicts greater results than MAPL. The reliability of the MiSFIT's NG, on the other hand, outperforms MAPL in most regions, for example in the 3 subregions of the *corpus callosum* (CC)—*genu*, *body* and

Phases	MAPL	MiSFIT
Fit Model	8 h 36 m	**17** s
PA Computation	2 h 19 m	**1** s
NG Computation	**3** s	6 s
Total time	10 h 56 m	**23** s

Table 3.3: Time consumption on the different phases of the metric's computation. Fitting the MAPL model is three-orders of magnitude more time-expensive than MiSFIT. MAPL requires almost 11 hours for a single DWI volume, in comparison with MiSFIT's 23 seconds, which makes the former unfeasible for clinical applicability.

splenium—which together form one of the biggest connective pathways in the brain.

In terms of repeatability, both measures result in stable outputs throughout the various sessions of a given subject, yielding low CV values. MAPL's PA results in more repeatable values in 6 out of 11 regions while giving virtually equal values for 2 of the rest (BCC and PLIC, with less than 10% difference between both frameworks); whereas MAPL's NG depicts higher repeatability in 5 out of the 11 regions. To widen the comparison with existing and previously-validated measures, the table also includes the reliability and repeatability values for FA (Fractional Anisotropy) and MK (mean Kurtosis) from DTI and DKI frameworks, respectively. As it can be seen, the FA shows the same tendency to that of MAPL and MiSFIT, with the only exceptions being ACR and SCR. Concerning the MK, which outperforms both MAPL and MiSFIT's reliability values in 6 of the regions, it is worth noticing that this measure is not as similar to the NG as the FA is to the PA. That is simply because, by definition, the mean Kurtosis is computed as the mean of individual kurtoses calculated along all gradient directions, whereas NG requires the 3D EAP and its 3D Gaussian counterpart to compute the dissimilarity of the projection onto a multidimensional vector; thus being affected by moments with order higher than kurtosis.

The results shown in the previous table are supplemented by the bootstrap-derived coefficients of variation of the reliability values in Table 3.2, which gives us insights about the reliability of the FOM values. For example, we can conclude that repeated acquisitions of the FOM value for the MAPL's PA in the *middle cerebellar peduncle* (MCP, 2.09) yield to variations 0.658 times its mean, thereby making us aware of the variation of such value, and able to question its trustworthiness.

3.5.4 Computational Time Evaluation

Arguably, one notorious factor that conditions the actual clinical applicability of EAP imaging resides in the computational time requirements. Table 3.3 shows the time needed by MiSFIT and MAPL to (1) fit the signal, (2) compute the PA, and (3) compute the NG. As it can be seen, fitting the signal to MAPL requires almost 9 hours for a single DWI MICRA volume with size (110, 110, 66, 266). MiSFIT does so in less than 20 seconds.

As expected, the PA computation is time-consuming for the MAPL approach, which has to re-calculate the MAP-MRI isotropic coefficients from those of the anisotropic representation. MiSFIT, on the contrary, only has to take the first SH coefficient to estimate the isotropic counterpart. This process ends up taking MAPL 2 hours 20 minutes in comparison with MiSFIT's 0.345 seconds.

Finally, the NG is easily computed by MAPL, which only has to take the first coefficient (Gaussian) from its series expansion and compute the angular divergence. In this case, MAPL computes the measure in 3.6 seconds while MiSFIT takes 5.8 seconds.

To sum up, the total time required by either framework to actually produce a meaningful map is no less than 8 hours 30 minutes for MAPL (fit the model and compute the NG) compared with MiSFIT's maximum of 22 seconds (fit the model and compute the NG). Needless to say, this time-consuming difference is of main importance not only in clinical settings, but also in the processing of large data sets that are becoming more common in the dMRI community.

3.6 Discussion and Conclusions

In this chapter we have shown how two descriptors of the white matter anatomy, PA and NG, can be computed within MiSFIT with very little computational effort. This way, we have generalized the analysis described in the original work Tristán-Vega and Aja-Fernández (2021), based on raw moments, to normalized indices that can be easier to interpret.

Regarding the PA, we have shown that MiSFIT yields to maps comparable to those obtained with MAPL. Yet, according to the numerical simulations, and within the context of NODDI-based simulations performed, MiSFIT results are more accurate. This remains true in regions with diverse anisotropic behavior, and becomes compromised only when dealing with attenuation signals highly contaminated by noise. The reliability evaluation depicts comparable outcomes for both frameworks; meaning that, in general, none is far superior than the other and result, on average, in estimates almost identical (5.85 for MAPL, 5.70 for MiSFIT). In some particular regions, however, substantial differences can be found. The repeatability assessment, on the other hand, yields to MiSFIT estimates more repeatable than MAPL's (on average, 0.06 and 0.17‰, respectively). Thus, we conclude not only that MiSFIT's PA leads to a desirable stability of measurements throughout the sessions (in its maximum, the average subject's measures of the GCC varies 0.151‰ its mean), but also that its reliability is comparable to that of MAPL.

Concerning the NG, we have shown the high correlation between MiSFIT's and MAPL's estimates, disturbed only by MiSFIT's noisier behavior in Gaussian regions. Overall, MiSFIT exhibits a positive bias compared to MAPL for all the range of NG. According to the numerical simulations in Fig. 3.5, however, both frameworks clearly underestimate the *actual* value of the NG. Paradoxically, the negative bias in MAPL increases with the PSNR. Yet, the (negative) bias introduced by MiSFIT is less severe than MAPL's, which is consistent with the experiment in Fig. 3.3 with real data. Therefore, the large relative errors reported in Fig. 3.6 for the NG (far larger than those for the PA) are likely explained by the bias in the estimation. Similar to PA, both frameworks depict highly repeatable results across intra-subject sessions (0.05‰ for MAPL, 0.10‰ for MiSFIT, averaging across regions). In terms of reliability, MiSFIT's estimates are, on average, far more reliable (12.52 against MAPL's 3.73).

Note the evaluation of the NG based on ground-truth is far more difficult than it is for the PA, since the definition of the NG is tightly related to the MAP-MRI representation: with MAP-MRI,

the closest Gaussian is trivially computed as the first addend of the expansion, so that the non-Gaussian part of the EAP is orthogonal to its Gaussian part. This property is unique to MAP-MRI, and does not hold neither for MiSFIT nor for the designed ground-truth signal. Consequently, it remains unclear if our ground truth is actually the desirable target. Given this situation, the reliability analysis grants a valuable quantitative information without the need of a ground-truth. In this sense, MiSFIT outperforms MAPL's estimation of the NG in 7 out of the 11 regions studied, including the CC. Taking also into account that MiSFIT offers a higher dynamic range for the NG, see Fig. 3.3 (left), we postulate that MiSFIT can be an attractive alternative to compute this index.

The focus on reliability rather than just repeatability was one of the main goals of this thesis. In this first contribution, we have taken initial steps toward developing this perspective and have begun to show that repeatability alone is not sufficient. In doing so, we concluded that assessing the reliability is mandatory when comparing different dMRI techniques, specially those that are based on the EAP, given that such imaging approaches crop the otherwise infinite bandwidth of the diffusion signal in different ways, depending on the representation used. The main limitation of the reliability assessment is the need for a robust inter-subject registration, so that the registration error does not become a critical confounding factor. In this sense, the repeatability values added to the analysis together with the computation of bootstrapped FoMs in Table 3.2 gives an idea of the confidence we can put on the results in Table 3.1. As long as the coefficients of variation are relatively large in all cases, we may conclude that more work is needed to pose the assessment of reliability as a state of the art procedure. Later in the document (see Chapter 5) in line with subobjective **SO3**, we dive into the reliability assessment and present a comparative study of EAP-based models and methods, together with a framework that is robust to intra-region registration.

A critical issue with previous implementations of both the PA and the NG is their time-consuming nature, unsuited for clinical practice. According to Table 3.3, the time required for processing an entire volume of the MICRA database with MAPL is well over 8 hours when Generalized Cross-Validation (GCV) and Positivity Constraints are used. For a subject in the HCP database, this time can grow up to 52 hours. Provided that for a clinical study a whole database needs to be processed, researchers will be compelled to use sub-optimal configurations, without positivity constraints, with fixed Laplacian penalty terms, or cropping the maximum order of the basis functions, which will compromise the accuracy of the measures and, in turn, the quality of the study. MiSFIT, on the other hand, is two orders of magnitude faster for any configuration.

It must be noted that we used MAPL's DIPY implementation because, at the time, it was the only available option and, for the sake of fairness, we wanted to use the original authors' implementation. Currently, newer implementations of MAPL exist that address its time-consuming nature (see Tristán-Vega, A. and Aja-Fernández, S. and París, G., 2022). While this implementation is still slower than MiSFIT, it is up to an order of magnitude faster than DIPY's for similar configurations; and is what has enabled the systematic study presented in Chapter 5. Although MAPL and MiSFIT use different optimization strategies (MAPL includes non-negativity constraints and GCV, while MiSFIT relies on regularization without positivity constraints) the computational time evaluation was done in the context of achieving an accurate representation of PA and NG markers.

To sum up, in this work the PA and NG measures have been introduced, their equations derived for MiSFIT's full composite attenuation signal and their performance within such framework validated. Both measures, with several proven clinical applications, result in consistent and reliable maps. All of this, together with MiSFIT's proven efficiency make the MiSFIT framework qualified for the new standardized dMRI protocol within clinical settings.

That which does not kill us makes us stronger.

— Nietzsche (1997)

4

Contribution: Rician bias effects on biophysical models

Abstract

Rotational invariants (RI) are at the root of many dMRI applications. Among others, they are presented as a sensible way of reducing the dimensionality of biophysical models. While thermal noise impact on diffusion metrics has been well studied, little is known on its effect on spherical harmonics-based RI (RISH) features and derived models. In this chapter, we evaluate the effect of noise on RISH features and downstream Standard Model Imaging (SMI) estimates.

Using simulated and test/retest multishell MRI data, we assess the accuracy and precision of RISH features and SMI parameters in the presence of thermal noise, as well as its robustness to variations in protocol design. We further evaluate the SMI estimator that elevates the need of RISH features.

RISH (and SMI) features are impacted by SNR-dependent Rician biases. However, higher-order RISH features are susceptible to a secondary noise-related source of bias, which not only depends on SNR, but also protocol and underlying microstructure. Rician bias-correcting techniques are insufficient to maximize the accuracy of RISH and SMI features, or to ensure consistency across protocols. SMI estimators that avoid RISH features by fitting the model to the directional diffusion MRI data outperform RISH-based approaches in accuracy, repeatability, and reproducibility across acquisition protocols.

RISH features are increasingly used in dMRI analysis, but prone to various sources of noise that lower their accuracy and reproducibility. Understanding the impact of noise, and mitigating such biases is critical to maximize the validity, repeatability, and reproducibility of dMRI studies.

4.1 Introduction and Purpose

Consistent with the aims of this thesis, specifically that in goal **SO2**, we seek to enhance the reliability and stability of microstructural biomarkers derived from SMI (see Section 2.4.4) by studying and mitigating the thermal noise and its effect on downstream estimates.

As mentioned in Chapter 2, biophysical models in anisotropic brain white matter tissue have a high dimensionality given the necessity to parametrize the fODF, possibly resulting in poor precision and robustness (Jelescu et al., 2016a; Novikov et al., 2018a). An initial strategy to reduce the number of model parameters of biophysical models is the adoption of simplified models of the fODFs, e.g. Watson or Bingham distribution, with the risk of reduced validity (Zhang et al., 2012; Fieremans et al., 2013; Tariq et al., 2016). Alternatively, the dimensionality of biophysical models can be reduced by extracting rotationally invariant signal features that encode the microstructural information of fiber fascicles, but not the macroscopic orientation distribution function of such fascicles, disentangling microstructural properties from their ODF (Mirzaalian et al., 2015; Reisert et al., 2017; Novikov et al., 2018a). In other words, rotationally invariant features, typically derived from the SH representation of the DW signal (see chapter 2.3.3), are agnostic to any kind of rotation that data might endure, thus removing the directional dependency of the signal.

Today, rotationally invariant signal features form the basis of numerous biophysical models of the white matter (e.g. Reisert et al., 2017; Pizzolato et al., 2022; Veraart et al., 2023, etc.), or the gray matter (Palombo et al., 2020; Jelescu et al., 2022). Common examples include the Spherical Mean Technique (Kaden et al., 2016), Standard Model Imaging (SMI, Novikov et al., 2018b), Axon Diameter Mapping (Veraart et al., 2021), Free Water Imaging (Pasternak et al., 2009; Tristán-Vega et al., 2021), Soma and Neurite Density Imaging (SANDI, Palombo et al., 2020), and Neurite Exchange Imaging model (NEXI, Jelescu et al., 2022). Within the context of modeling, the most commonly used rotationally invariant signal feature is the "spherical mean" (Mirzaalian et al., 2015; Kaden et al., 2016), which is proportional to the zeroth order spherical harmonic coefficient, but the use of higher orders have recently been promoted as it provides complimentary information (Reisert et al., 2017; Novikov et al., 2018b; Zucchelli et al., 2020; Pizzolato et al., 2022; Pieciak et al., 2024). Therefore, for anisotropic tissue, including higher order rotational invariants, e.g. $l = 2$ or above, provides additional equations to resolve the model while lowering the necessity for numerous b-shells, thereby making biophysical models compatible with many publicly available data sets (Coelho et al., 2024).

Alongside their use in biophysical modeling, rotationally invariant features are also used as sensitive biomarkers in the analysis of diffusion MRI data (Chamberland et al., 2021; Ward et al., 2023; Raven et al., 2023), or as features to drive data harmonization across scanners and protocols (Mirzaalian et al., 2015, 2016; Karayumak et al., 2019) and to train deep learning tools (Chamberland et al., 2021). As we witness an increased use of rotationally-invariant signal

features for a variety of analysis approaches, it becomes increasingly important to understand the confounds in its quantification, including its robustness to experimental factors such signal-to-noise ratio (SNR) or protocol design.

Despite increasing attempts to achieve higher SNR through hardware and software improvements, e.g. increasing field strengths (Balchandani and Naidich, 2015), stronger imaging gradients (Setsompop et al., 2013), better receiver coil arrays (Wiggins et al., 2009), denoising (Veraart et al., 2016a), etc., dMRI is an imaging technique that is intrinsically poor in SNR. Thermal noise impacts the quantification of diffusion metrics in various ways, with eigenvalue repulsion (Pierpaoli and Basser, 1996) and Rician biases (Jones and Basser, 2004) being notorious examples. With decreasing SNR, a noise-related signal bias, often referred to as Rician biases, result in an underestimated apparent diffusivities and fractional anisotropy and overestimated kurtosis values if parameters estimators do not explicitly account for the actual probability distribution function of magnitude MRI data (Jones and Basser, 2004; Veraart et al., 2011). In the context of biophysical modeling, Rician biases have been associated with a still water fraction (Uhl et al., 2024). Yet, little is known on the impact of noise on rotational invariants and on downstream derived biophysical model parameters.

In this chapter, we assess the effect of thermal noise on rotational invariants, via both simulated and in vivo data. We also evaluate the effect of noise in downstream biophysical models that are derived from such rotational invariants by taking a look at the SMI microstructural parameters. Finally, we propose and evaluate correction strategies.

4.2 Theory

4.2.1 Rotationally-invariant spherical harmonic features

Although previously discussed in Chapter 2, the full derivation of the SMI framework is included here to ensure this section is self-contained. The attenuation signal $S(b, \mathbf{g})$, acquired for a diffusion-weighting strength b and gradient direction $\mathbf{g}$ can be represented as:

$$S(\mathbf{q}) = \iint_{\mathcal{S}} \mathbf{\Phi}(\mathbf{n}) \mathcal{K}(\mathbf{q}) d\mathbf{n} = (\mathbf{\Phi} \circledast \mathcal{K})(\mathbf{q}), \tag{4.1}$$

where $\mathbf{\Phi}(\mathbf{n})$ is the fODF, normalized so that $\int d\mathbf{n} \mathbf{\Phi}(\mathbf{n}) = 1$, and $\mathcal{K}$ the signal response kernel of a fascicle aligned to the direction $\mathbf{n}$ (Jespersen et al., 2007; Reisert et al., 2017). Similar to Novikov et al. (2018b), we normalize $d\mathbf{n} = \frac{\sin\theta d\theta d\varphi}{4\pi}$ so that $\int d\mathbf{n} = 1$. The fODF $\mathbf{\Phi}(\mathbf{n})$ is typically parameterized by means of its projection onto the SH, such that:

$$\mathbf{\Phi}(\mathbf{n}) = 1 + \sum_{\substack{l=2, \\ l \text{ even}}}^{\infty} \sum_{m=-l}^{l} p_l^m Y_l^m(\mathbf{n}), \tag{4.2}$$

where the SH of degree l and order m is given by Eq. 2.9. Notice that only real and even SH (even orders of l) are being introduced under the assumption of antipodally symmetric

diffusion (Tournier et al., 2004).

Similarly, for a given b-value, $S(\mathbf{q}) = S(b, \mathbf{g})$ samples the signal Orientation Distribution Function, which can also be parameterized in such SH basis using coefficients $S_l^m(b)$:

$$S(b, \mathbf{g}) = \sum_{\substack{l=0, \\ l \text{ even}}}^{\infty} \sum_{m=-l}^{l} S_l^m(b) Y_l^m(\mathbf{g}). \tag{4.3}$$

While $S_l^m(b)$ are dependent on the orientation of the underlying microstructure, we can derive RISH features of the l-th order S_l:

$$S_l(b) = \frac{1}{N_l} \sqrt{\sum_{m=-l}^{l} \left| S_l^m(b) \right|^2}, \tag{4.4}$$

with $N_l = \sqrt{4\pi(2l+1)}$ being a normalization parameter (Mirzaalian et al., 2015; Reisert et al., 2017; Novikov et al., 2018b). We limit ourselves to the 0th and 2nd order RISH features ($l = 0$ and $l = 2$) assuming sufficient signal smoothness to ensure exponential decay of higher-order contributions, therefore resulting in:

$$S_0(b) = \sqrt{\frac{1}{4\pi}} S_0^0(b), \tag{4.5}$$

and

$$S_2(b) = \sqrt{\frac{1}{20\pi} \sum_{m=-2}^{2} \left[S_2^m \right]^2}. \tag{4.6}$$

Note that S_0 is often referred to as the spherical mean. S_2, on the other hand, is closely related to the recently introduced spherical variance (Pizzolato et al., 2022).

4.2.2 Biophysical modeling

Model

The SMI draws an overarching picture of biophysical models that describe DW signal as a sum of non-exchanging Gaussian compartments (Novikov et al., 2018b). The signal kernel $\mathcal{K}$ of Eq. (4.1) is defined as in Eq. (2.76).

The kernel parameters, corresponding to the microstructural parameters to be estimated within the white matter, are defined in Section 2.4.4 as the intracellular signal fraction f, and longitudinal intra-cellular diffusivity D_a; and the extracellular parallel $D_e^{\parallel}$, and perpendicular $D_e^{\perp}$ diffusivities. Note that intracellular perpendicular diffusivity is assumed to be zero (Veraart et al., 2018).

Following the parameterization of the fODF in the SH basis, Eq. (4.1) can be rewritten as:

$$S(b,\mathbf{g}) = \sum_{\substack{l=0,\\ l \text{ even}}}^{\infty} \sum_{m=-l}^{l} p_l^m K_l(b;\boldsymbol{\theta}) Y_l^m(\mathbf{g}), \tag{4.7}$$

where K_l is the projection of the response kernel onto the Legendre polynomials for a given set of parameters $\boldsymbol{\theta} = [f, D_a, D_e^{\parallel}, D_e^{\perp}]$, as per Funk-Hecke's theorem (Hobson, 2012).

We can further simplify this equation, by representing $S(b,\mathbf{g})$ in the SH basis,

$$S_l^m(b) = p_l^m K_l(b;\boldsymbol{\theta}), \tag{4.8}$$

or the derived rotational-invariant features:

$$S_l(b) = p_l K_l(b;\boldsymbol{\theta}), \tag{4.9}$$

with $p_l = \frac{1}{N_l}\sqrt{\sum_{m=-l}^{l} \left|p_l^m\right|^2}$.

The fODF rotational invariants p_l are either defined as $p_0 = 1$, or estimated as part of the microstructural parameters that measure the coherence of the fODF p_2.

Note that, for $l = 0$, the number of parameters is 4 $(f, D_a, D_e^{\parallel}, D_e^{\perp})$, whereas for $l \geq 2$, additional modeling parameters p_l have to be estimated. The coherence index p_2 varies between 0 and 1. For a fully isotropic fODF, $p_2 = 0$, while a delta distribution on a sphere has $p_2 = 1$.

Fitting

Given the multitude of SMI representations, e.g. Eqs. (4.7) and (4.9), it is clear there is no unique way of model fitting to extract the parameters of interest from dMRI data.

The biophysical model parameter vector $\boldsymbol{\theta}$ is most commonly estimated by minimizing the equation (2.80). In its essence, this presents a two-step fitting process SH-SMI, referencing the intermediate step of estimated SH coefficients $S_l^m(b)$ and derived rotation-invariants $S_0(b)$ and $S_2(b)$. $S_l^m(b)$ is commonly estimated using linear least squares (LLS) estimators as it is computationally efficient and various software libraries provide open source implementations of such estimators (e.g., MRtrix, by Tournier et al., 2019):

$$\widehat{S}_l^m(b) = \arg\min_c \sum_{j=1}^{N_g} \left(S(b,\mathbf{g}_j) - cY_l^m(\mathbf{g}_k)\right)^2, \tag{4.10}$$

which has a trivial closed-form solution. Note that N_g is the number of DW signals with b-value b. This two-step fitting process is commonly used in many of the above listed models (SMT from Kaden et al., 2016; SMI from Novikov et al., 2018b; SANDI from Palombo et al., 2020; NEXI from Jelescu et al., 2022, etc.).

Alternatively, one can bypass the intermediate step of computing rotational invariants.

Following (4.7), we can estimate the model parameters, alongside the fODF, directly from the DW data $S(b, \mathbf{g})$ using the following nonlinear least squares estimator:

$$\widehat{\boldsymbol{\theta}}^{(1)} = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{N} \left(S(b_j, \mathbf{g}_j) - \sum_{\substack{l=0, \\ l \text{ even}}}^{L} \sum_{m=-l}^{l} p_l^m K_l(b; \boldsymbol{\theta}) \right)^2 . \tag{4.11}$$

Note that the superscript denotes the number of estimation steps of the approach. Equation (4.9) first extracts RISH features and then fits the signal via Eq. (2.80) , whereas Eq. (4.11) fits the signal directly. This single-step fitting (or direct fit, DF) strategy was initially proposed by Jespersen et al. (2007). While SH-based modeling approaches gained interest, preliminary results suggested the impact of direct fitting in settings where DW data is not acquired in conventional b-shells, by design or due to gradient nonlinearities (Henriques et al., 2018).

The estimators of $S_l^m(b)$ and $\widehat{\boldsymbol{\theta}}^{(1)}$ are (asymptotically) unbiased if the DW data $S(b, \mathbf{g})$ are Gaussian distributed.

4.2.3 Data distributions

Rician distribution of dMRI data

The reconstruction of DW images typically includes the computation of the magnitude of the otherwise normally-distributed, complex-valued MR image. Their phase—sensitive to macroscopic tissue motion due to cardiac pulsation, perfusion, and others—varies from image to image, and impedes further modeling if not corrected or omitted (Eichner et al., 2015). In computing the magnitude s, its distribution is no longer Gaussian, but Rician under some often-met conditions (Aja-Fernández and Vegas-Sánchez-Ferrero, 2016). Therefore, the probability density function $p(s|v, \sigma)$ of magnitude MRI data is then:

$$p(s|v, \sigma) = \frac{s}{\sigma^2} \exp\left(-\frac{s^2 + v^2}{2\sigma^2} \right) I_0\left(\frac{sv}{\sigma^2} \right), \quad s \geq 0, \tag{4.12}$$

where v is the noiseless signal amplitude, σ the Gaussian noise standard deviation and I_0 the 0th-order modified Bessel function of the first kind (Gudbjartsson and Patz, 1995; Aja-Fernández and Vegas-Sánchez-Ferrero, 2016). Its conditional expectation is given by the formula:

$$E\left[s|v, \sigma\right] = \mu = \sigma \cdot \sqrt{\frac{\pi}{2}} \cdot L_{1/2}\left(\frac{-v^2}{2\sigma^2} \right) \tag{4.13}$$

with $L_{1/2}$ being Laguerre's polynomial. The positive offset between μ and v is SNR-dependent, exacerbated on the low-intensity data points such as those acquired in higher b-values, and often referred to as the Rician signal bias. This signal bias impacts the accuracy of the estimated model parameters (Pierpaoli and Basser, 1996; Jones and Basser, 2004; Veraart et al., 2011; Tristán-Vega et al., 2012; Kerkelä et al., 2020).

Correction approaches

Common estimators, such as *least squares estimators* are biased if the data is Rician distributed. The magnitude of the bias is SNR-dependent, but cannot be ignored if any of the DW signals has an SNR below 2 (Gudbjartsson and Patz, 1995). To address this issue, various strategies have been presented to mitigate the bias either (1) by lowering the initial noise level in the complex (Cordero-Grande et al., 2019; Moeller et al., 2021) or spatial domain (Varadarajan and Haldar, 2015; Pieciak et al., 2024), or (2) by mitigating its propagation to the dMRI metrics by means of maximum likelihood estimators (Sijbers et al., 1998; Aja-Fernández and Vegas-Sánchez-Ferrero, 2016), conditional least squares (CLS, Veraart et al., 2013, or signal transformations (Koay and Basser, 2006).

In this chapter, we make use of a CLS implementation embedding the Rician expectation (Eq. (4.13)) to offset model predictions (Veraart et al., 2013; Uhl et al., 2024):

$$v_{\text{CLS}} = \arg\min_{v} \sum_{i}^{N} \left(s_i - E[s|v,\sigma]\right)^2 \tag{4.14}$$

where s_i are the observed signals, and the model prediction term is replaced by their corresponding Rician expectation with v being the model prediction itself and σ the Gaussian noise standard deviation. We will here refer to Rician Bias Correction (RBC) when using this approach.

Such estimator is a generic approach that can be adopted to estimate $S_l^m(b)$ or $\widehat{\boldsymbol{\theta}}^{(1)}$, by modifying Eqs. (4.10) and (4.11), accordingly. Unlike maximum likelihood estimators, we model the signal offset due to the Rician signal bias using its expectation value only, thereby ignoring higher order terms of the data distribution. As demonstrated by Veraart et al. (2013), such an approach is favorable if the data distribution has been altered during image preprocessing.

Non-central Chi distribution of RISH features

Note that the $S_l^m(b)$ coefficients are not noise-free. Indeed, they are asymptotically normal (under the central limit theorem) with their standard deviation dependent on σ and scan protocol, e.g. number of gradient directions. By definition, the rotational invariant S_0 will thus be (asymptotically) Gaussian distributed. However, assuming the same variance on each of these S_2^m coefficients (with m ranging from -2 to 2), the S_2 follows a noncentral Chi distribution (Aja-Fernández and Vegas-Sánchez-Ferrero, 2016). Similar to the Rician distribution, the expectation value of the noncentral Chi distribution exceeds its underlying *noise free* value. Therefore, $S_2(b)$ is prone to a systematic numeric bias that depends on the noise level of $S_l^m(b)$, which is impacted by various factors, including SNR of data, data preprocessing steps such as denoising, and scan protocol. Note that this source of bias is complementary to any errors that arise from Rician signal biases.

4.3 Materials

4.3.1 In silico data

We evaluated the properties of the various estimates using simulated data for varying b-values, gradient directions, biophysical model parameters, fiber configurations, and SNR values. Eq. (4.7) was used as a forward model to generate simulated DW signals for given parameters, setting $L_{\max} = 6$. For crossing fiber configurations, we generated DWI signals for two fascicles that are rotated with respect to each other by a rotation angle α, but identical otherwise. These signals were added with relative weight w prior to adding noise. We added complex Gaussian noise with $\sigma = 1/\text{SNR}$, for different SNR values, prior to computing the signal's magnitude, cf. Veraart et al. (2013). The resulting DW data is Rician distributed (Gudbjartsson and Patz, 1995). All parameter settings are specified when relevant in the following sections. However, briefly, SNR varied between 5 and 45 in our experiments. In agreement with our *in vivo* MRI data, b-values were generally set to $[500, 1000, 2500, 6000]$ s/mm^2. Unless specified otherwise, the biophysical model parameters were uniformly sampled within the following ranges: $f \in [0.25, 0.75]$, $D_a \in [0.5, 2.5]$ μm^2/ms, $D_e^{\parallel} \in [0.5, 2.5]$ μm^2/ms, $D_e^{\perp} \in [0.1, 1]$ μm^2/ms and $p_2 \in [0, 1]$. Moreover, the elevation angle between two simulated fibers was randomly set to $[0, \pi/2]$. Finally, the number of gradient directions was set to 32.

4.3.2 In vivo dataset

This study presents the secondary analysis of test/retest dMRI data introduced in Section 1.4, specifically we have used the CUBRIC - Test/Retest dataset to evaluate the performance of the proposed strategies on real-life data (Henriques et al., 2021; Veraart et al., 2021).

Data Preprocessing: Preprocessing of the data included detection and removal of signal outliers (Andersson et al., 2016) and the corrections for Gibbs ringings (Kellner et al., 2015), subject motion, and susceptibility and eddy current distortions (Andersson et al., 2003). Gradient nonlinearities were corrected by scaling the b-value voxel-by-voxel (Rudrapatna et al., 2020). The noise level was estimated from the data acquired with b-values up to 1000 s/mm^2 using MPPCA (Marchenko-Pastur Principal Component Analysis) with a window size of [5,5,5] to avoid Rician biases in the noise map (Veraart et al., 2016a).

ROI Analysis: We performed image registration using the ANTsRegistrationSyN tool (ANTs toolbox) between the subject's FA map and the JHU FA template (JHU-ICBM-FA-1mm). The registration was only computed for the subject's test data and includes both an affine transformation and a diffeomorphic warp field. The JHU labels (JHU-ICBM-labels-1mm; Analysis Group, FMRIB, Oxford, UK.; Smith et al., 2004) were then transformed into the subject's space by applying the inverse transformation using nearest neighbor interpolation. We accounted for any misalignment between the test and retest data using an additional rigid registration transformation. Each of the white matter regions of interest (ROIs) was then masked by means of threshold masks of the white matter tract density imaging (WM-TDI; Calamante et al., 2010), removing voxels that were

not intersected by tracts and thus, possibly contaminated by partial volume effects. Finally, ROIs containing less than 50 voxels were discarded. A representative of each region (a total of 27 out of the 48 original labels) was computed as the median for each SMI estimate.

4.4 Method

4.4.1 Parameter Estimation

The microstructural parameters within the SMI framework are estimated by means of two estimation strategies, which are derived from Eqs. (2.80) and (4.11), respectively. For convenience, equations are included here.

1. The first approach $\widehat{\boldsymbol{\theta}}^{(2)}$, first estimates $\widehat{S}_l^m(b)$ (Eq. (4.10)) and derives the corresponding $\widehat{S}_l(b)$ (Eq. (4.9)). Thereafter, it estimates the standard model parameters $\widehat{\boldsymbol{\theta}}^{(2)}$ using Eq. (2.80):

$$\widehat{S}_l^m(b) = \arg\min_c \sum_{j=1}^{N_g} \left(S(b,\mathbf{g}_j) - cY_l^m(\mathbf{g}_k)\right)^2 \tag{4.15}$$

$$\downarrow$$

$$S_l(b) = \frac{1}{N_l}\sqrt{\sum_{m=-l}^{l}\left|S_l^m(b)\right|^2} \tag{4.16}$$

$$\downarrow$$

$$\widehat{\boldsymbol{\theta}}^{(2)} = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{N_b} \sum_{\substack{l=0,\\ l \text{ even}}}^{L} \frac{1}{\sigma_{lj}^2}\left|S_l(b_j) - p_l K_l(b,\boldsymbol{\theta})\right|^2. \tag{4.17}$$

2. The second approach, $\widehat{\boldsymbol{\theta}}^{(1)}$, fits the standard model directly to the direction DW signals as in Eq. (4.11):

$$\widehat{\boldsymbol{\theta}}^{(1)} = \arg\min_{\boldsymbol{\theta}} \sum_{j=1}^{N} \left(S(b_j,\mathbf{g}_j) - \sum_{\substack{l=0,\\ l \text{ even}}}^{L} \sum_{m=-l}^{l} p_l^m K_l(b;\boldsymbol{\theta})\right)^2. \tag{4.18}$$

The estimation of $\widehat{S}_l^m(b)$ and $\widehat{\boldsymbol{\theta}}^{(1)}$ can be performed using the least squares estimators of Eqs. (4.10) and (4.11), respectively. However, to account for Rician signal biases, we will compare such estimators to the CLS estimators, $\widehat{\boldsymbol{\theta}}^{(1)}_{\text{RBC}}$ and $\widehat{\boldsymbol{\theta}}^{(2)}_{\text{RBC}}$), that model the Rician expectation value operator at the relevant steps. Following Eq. (4.14), the Rician bias correction requires prior knowledge of the noise level. For simulated data, we know this value by design, while for the experimental data, we use the noise levels as estimated in Section 4.3.2.

All code was implemented using standard library functions from the optimization toolbox of Mat-

lab.

4.4.2 Statistical Analyses

In this study, we measure the test/retest repeatability of estimators and we quantify the agreement between two sets of measurements to quantify the robustness of the estimators to variations in scan protocols. The agreement between scan protocols is done by using half of the last b-shell ($b = 6000$ s/mm^2) gradient directions. We will use the coefficient of variation between two repeated measurements to measure test/retest reproducibility, whereas Lin's Concordance Correlation Coefficient (CCC; Lawrence and Lin, 1989) is used to measure agreement between two paired measurements.

Test/retest repeatability: To assess the repeatability and reliability of the metrics, the Coefficient of Variation (COV) were computed as follows:

$$\overline{\text{COV}} = \frac{\sqrt{\pi}}{N} \sum_{i=1}^{S} \frac{|x_{it} - x_{ir}|}{x_{it} + x_{ir}} \times 100[\%] \tag{4.19}$$

with x_{it} and x_{ir} being test and retest measurement of a diffusion metric-of-interest of the i-th subject, respectively, and S being the number of subjects.

Concordance: The agreement between two subject-matched sets of measurements of a diffusion metric of interest was computed as follows:

$$\rho_C = \frac{2\rho\sigma_X\sigma_Y}{(\mu_X - \mu_Y)^2 + \sigma_X^2 + \sigma_Y^2} \tag{4.20}$$

where ρ_C is the Pearson's correlation coefficient between X and Y, and μ_X and μ_Y their averaged value. X and Y are both $[N, 1] \in \mathbb{R}^N$ vectors, representing a diffusion metric for each subject and measurement measurement.

4.5 Results

4.5.1 Simulations: Thermal noise effect on RISH features

Figure 4.1 shows the impact of signal noise on the accuracy and precision of $\widehat{S}_0(b)$ and $\widehat{S}_2(b)$ for a wide range of parameter values ($n = 100.000$, sampled from distributions defined in Section 5.3) in relation to the distribution of the dMRI data and the used estimator (LLS versus CLS). We observe an SNR-dependent overestimation of $\widehat{S}_0(b)$ if the dMRI data is Rician distributed, but an ordinary LLS estimator has been used to estimate $\widehat{S}_l^m(b)$. The error is significantly reduced when using the CLS estimator instead. In contrast, $\widehat{S}_2(b)$ is prone to error, even in case of Gaussian distributed dMRI data. We observe SNR-dependent overestimation of $\widehat{S}_2(b)$, with an magnitude of error that is inversely proportional to its underlying value $\widehat{S}_2(b)$. Rician signal biases result in an underestimation of $\widehat{S}_2(b)$, with a magnitude that is proportional to $\widehat{S}_2(b)$. The use of a Rician

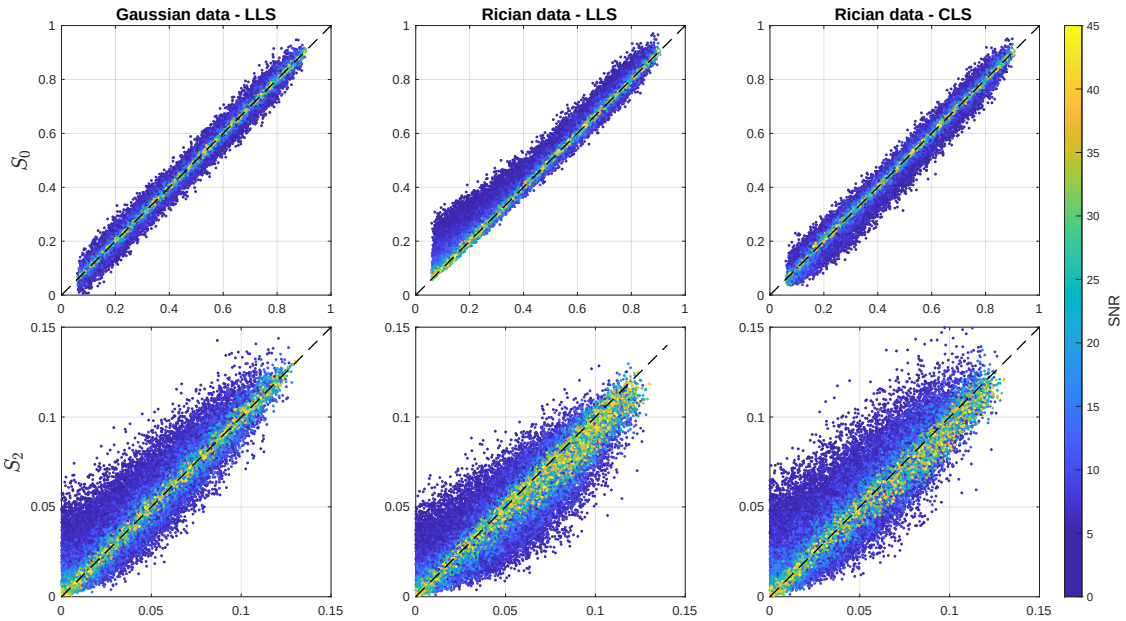


Figure 4.1: Rotational invariants $\widehat{S}_0$ and $\widehat{S}_2$ estimates against their ground truth for varying SNR levels: Gaussian noise with LLS estimation (left), Rician noise with LLS estimation (middle), and Rician noise with CLS estimation (right). Top row: Behavior of $\widehat{S}_0$. Under Gaussian noise, estimates are unbiased. Under Rician noise, however, low $\widehat{S}_0$ values are systematically overestimated. This bias is corrected when using CLS estimators, restoring an unbiased behavior. Bottom row: Behavior of $\widehat{S}_2$. Even under Gaussian noise, $\widehat{S}_2$ tends to be overestimated at low values. Under Rician noise, two sources of bias emerge: overestimation at low $\widehat{S}_2$ values and underestimation at high values. Applying CLS to Rician data mitigates the second source of bias, restoring the estimation pattern observed in the Gaussian case.

bias correcting estimator reduces the second error, while the first source of error remains.

Figure 4.2 outlines the impact of SNR, number of gradient directions, and coherence index p_2 (also called microscopic anisotropy index, i.e. the ODF anisotropy RI) on $\widehat{S}_0(b)$ and $\widehat{S}_2(b)$ for Gaussian and Rician distributed dMRI data. In this single fiber simulations ($w = 1$), the following parameters were fixed to $f = 0.75$, $D_a = 2.5$ μm^2/ms, $D_e^{\parallel} = 2$ μm^2/ms, and $D_e^{\perp} = 0.5$ μm^2/ms. When not varying, the number of gradient directions was set to 32 and SNR to 30. Moreover, p_2 was set to 0.2 or 0.6 when varying the number of gradient directions or SNR, respectively. First we observe that the error in the estimation of $\widehat{S}_0(b)$ depends on the SNR and p_2 in case of Rician distributed data only. The Rician signal bias results in an overestimation of $\widehat{S}_0(b)$, particularly at low SNR and high p_2 with relative errors up to 20%. The error in the estimation of $\widehat{S}_2(b)$ depends on SNR, number of gradient direction, and p_2. However, unlike $\widehat{S}_0(b)$, $\widehat{S}_2(b)$ is subject to such errors even in case of Gaussian distributed dMRI data. Indeed, for Gaussian distributed dMRI, we observe an overestimation $\widehat{S}_2(b)$ that increases with decreasing SNR, number of gradient direction, and p_2. In case of Rician distributed dMRI data, this overestimation is counterbalanced, oftentimes resulting in underestimation of $\widehat{S}_2(b)$. This effect is most pronounced at low SNR and high p_2.

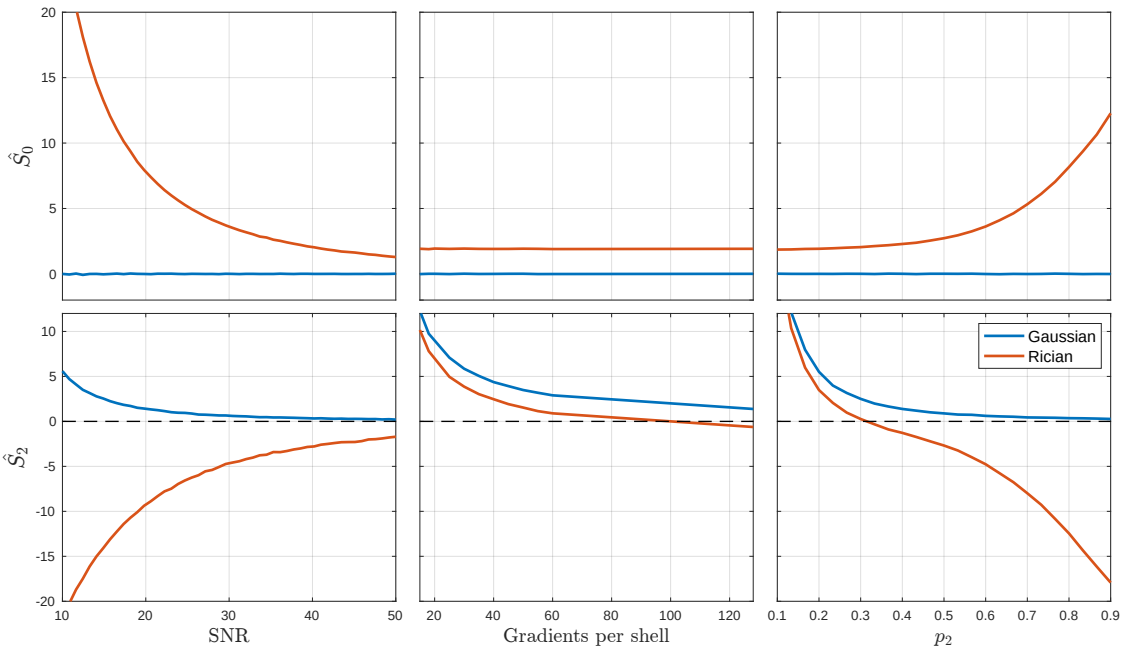


Figure 4.2: Relative error of $\widehat{S}_0$ and $\widehat{S}_2$ estimates in different scenarios: For varying SNR (left), varying number of gradients per shell (middle) and varying p_2 (right). When not varying, the number of gradient directions was set to 32 and SNR to 30. Also, p_2 was set to 0.2 or 0.6 when varying the number of gradient directions or SNR, respectively. Top row: Behavior of $\widehat{S}_0$. Gaussian noise is unbiased in all scenarios. Under Rician noise, however, $\widehat{S}_0$ is overestimated in all three cases. The overestimation diminishes for increasing SNR and decreasing axon dispersion p_2. The number of gradients per shell does not impact the overestimation. Bottom row: Behavior of $\widehat{S}_2$. Under Gaussian noise, $\widehat{S}_2$ is systematically overestimated and such overestimation diminishes with increasing SNR, number of gradients per shell and axon dispersion p_2. Under Rician noise, however, more complex interactions come to play.

4.5.2 Simulations: Thermal noise effect on biophysical modeling parameters across estimators

Figure 4.3 illustrates how the errors in the estimation of $\widehat{S}_0(b)$ and $\widehat{S}_2(b)$ propagate and impact the accuracy of $\widehat{\boldsymbol{\theta}}^{(2)}$ (i.e., SH-based) estimators. Using the same simulated data as used for Figure 4.2, we demonstrate the impact of SNR, number of gradient directions, and p_2 on $\widehat{\boldsymbol{\theta}}^{(2)}$ for Gaussian and Rician distributed dMRI data. As $\widehat{\boldsymbol{\theta}}^{(2)}$ is derived from $\widehat{S}_0(b)$ and $\widehat{S}_2(b)$ simultaneously, we observe errors in case of Gaussian and Rician distributed dMRI data, but the direction and magnitude of the error depends on the data distribution. However, even for Gaussian-distributed data, estimation errors increase across all parameters as SNR, number of gradients or p_2 decrease, with errors ranging from a few percent to over 50%. For Rician-distributed dMRI data the magnitude of the errors typically increases. However, instead of an overestimation, we observe an underestimation of $\widehat{D}_e^{\perp}$ and $\widehat{p}_2$. While $\widehat{\boldsymbol{\theta}}^{(2)}$ had relatively low errors for high p_2 values under Gaussian distributed data, we observe a strong error increase for Rician distributed data.

Figure 4.4 presents violin plots of the error in the estimation of $\widehat{\boldsymbol{\theta}}^{(2)}_{\mathrm{RBC}}$, $\widehat{\boldsymbol{\theta}}^{(2)}$, and $\widehat{\boldsymbol{\theta}}^{(1)}_{\mathrm{RBC}}$ for a wide range of parameter values ($n = 1500$, sampled from distributions defined in Section 4.3.1). Note that $\hat{\theta}^{(1)}$ is omitted here and in the following figures, as previous results have already demonstrates that the RBC correction significantly improves the estimation accuracy compared

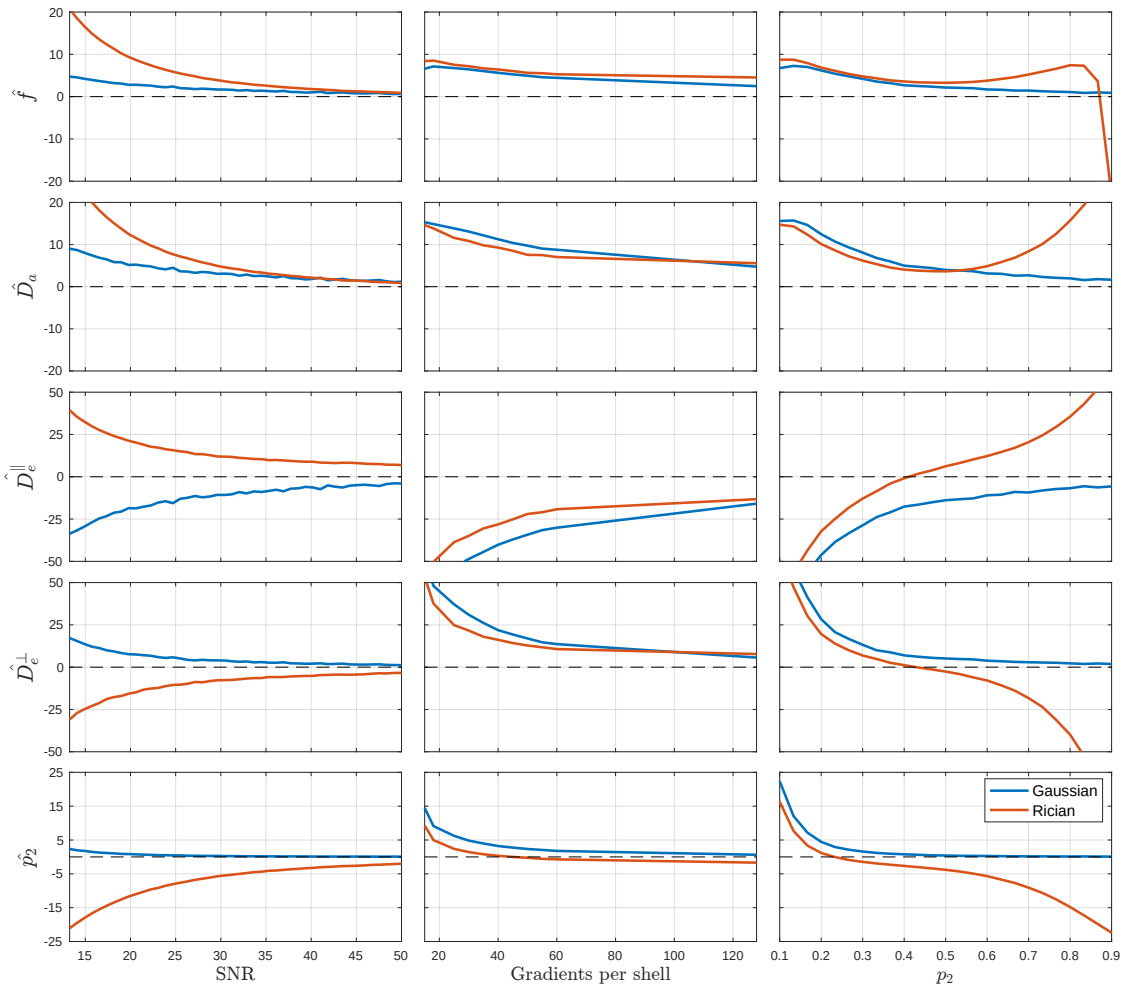


Figure 4.3: Percentage relative error in $\widehat{\theta}^{(2)}$ microstructural estimates for Gaussian- and Rician-distributed data for varying levels of SNR (left); number of gradients per shell (middle); and axon dispersion p_2 (right). When not varying, the number of gradient directions was set to 32 and SNR to 30; p_2 was set to 0.2 or 0.6 when varying the number of gradient directions or SNR, respectively. Errors arise from inaccuracies in both $\widehat{S}_0(b)$ and $\widehat{S}_2(b)$, with Gaussian data showing increasing overestimation as SNR, number of directions, or p_2 decrease. Rician noise leads to generally larger errors, typically manifesting as underestimation, particularly at low SNR and high p_2.

to the original implementation. Therefore, subsequent comparisons focus on the original $\hat{\theta}^{(2)}$ and the RBC-corrected estimators, $\hat{\theta}^{(1)}_{\text{RBC}}$ and $\hat{\theta}^{(2)}_{\text{RBC}}$. All simulated dMRI data was Rician-distributed with randomly varying SNR. In addition, we computed the mean squared error (MSE), see Table 4.1. For all metrics, the MSE—a measure that combines accuracy and precision—was lowest for $\widehat{\theta}^{(1)}_{\text{RBC}}$, indicating superior overall performance.

4.5.3 In vivo results

Figure 4.5 shows the in vivo maps of the different estimates for a single subject's white matter, overlapped with its grayscale FA. Overall, we observe a substantial noisy behavior on the original $\hat{\theta}^{(2)}$, probably intensified by the degeneracy of the problem and the inability of the

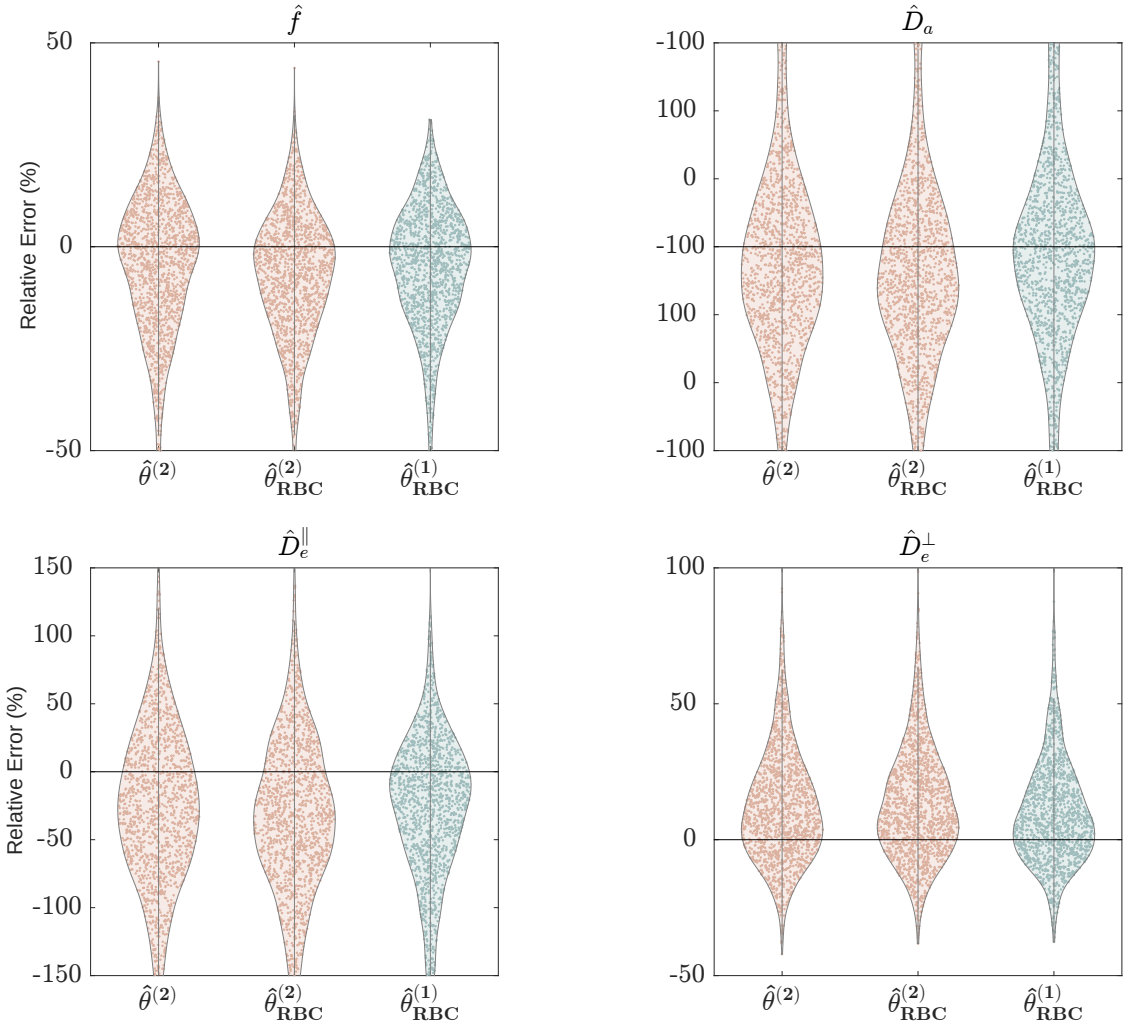


Figure 4.4: Relative error distributions of SMI microstructural estimates: Solid black line represents the ground truth (relative error equal to 0). Violin plots show errors in $\widehat{\theta}^{(2)}$, $\widehat{\theta}^{(2)}_{\mathrm{RBC}}$, and $\widehat{\theta}^{(1)}_{\mathrm{RBC}}$ across 1500 Rician-distributed dMRI simulations with varying parameters. The estimates' modes derived from $\widehat{\theta}^{(2)}$ and $\widehat{\theta}^{(2)}_{\mathrm{RBC}}$ show consistent bias, in contrast to the unbiased estimates obtained from $\widehat{\theta}^{(1)}_{\mathrm{RBC}}$.

branch selection algorithm to correctly decide the solution. In comparison, $\hat{\theta}^{(2)}_{\mathrm{RBC}}$ and $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ depict more robust and stable maps throughout the white matter. While the difference on stability/robustness between $\hat{\theta}^{(2)}$ and the other strategies can be observed in any of the maps, it is more prominent in the f and D_e parameters.

4.5.4 Protocol dependency

Table 4.2 shows Lin's CCC score obtained for each of the methods (the higher, the better). From the three methods under study, $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ outperforms the other methods nominally for all SMI metrics. Following the guidelines for interpretation (see Akoglu, 2018), the performance of $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ ranges from substantial (0.95 - 0.99) to almost perfect. The performance of $\hat{\theta}^{(2)}_{\mathrm{RBC}}$ is more variable, ranging from poor to substantial, while $\hat{\theta}^{(2)}$ is overall poor.

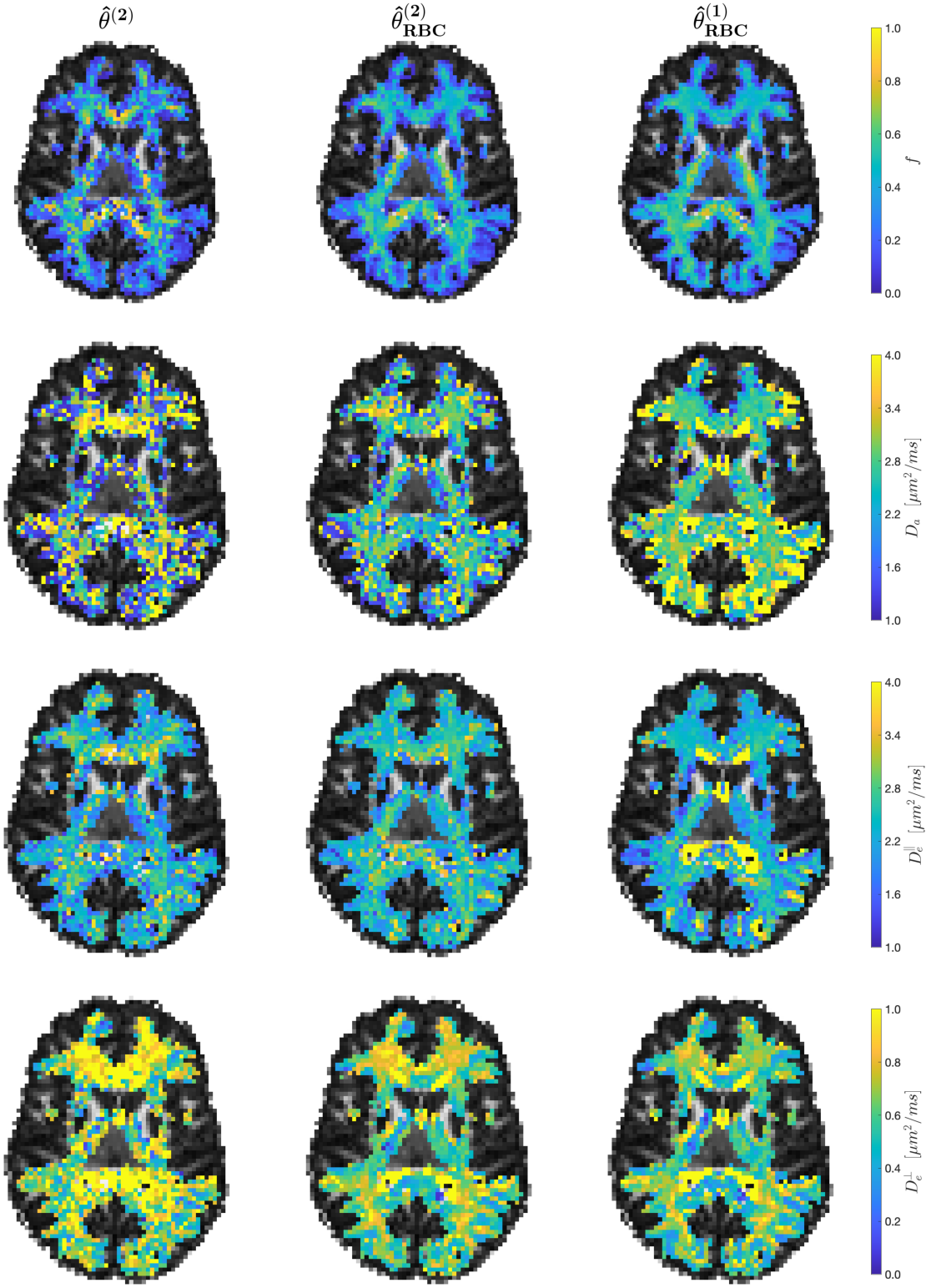


Figure 4.5: Real data comparison of the estimation strategies: Comparison between the three approaches under study of microstructural estimates from *in vivo* data, masked via WM-TDI (see more on Section 4.4). Noticeably, $\hat{\boldsymbol{\theta}}^{(2)}$ outputs noisier maps, probably due to the degenerate landscape of the problem. This is mitigated in the proposed strategies ($\hat{\boldsymbol{\theta}}^{(2)}_{\mathrm{RBC}}$ and $\hat{\boldsymbol{\theta}}^{(1)}_{\mathrm{RBC}}$), which result in more robust maps.

MSE	f	$\mathbf{D_a}$	$\mathbf{D_e^{\parallel}}$	$\mathbf{D_e^{\perp}}$
$\hat{\theta}^{(2)}$	0.86	0.84	0.85	0.71
$\hat{\theta}^{(2)}_{\mathrm{RBC}}$	0.96	0.84	0.95	**0.97**
$\hat{\theta}^{(1)}_{\mathrm{RBC}}$	**0.99**	**0.93**	**0.99**	**0.97**

Table 4.1: MSE of the estimated biophysical model parameters using simulated data with a wide range of underlying parameters: $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ outperforms the other strategies by resulting in the lower MSE. Bold values depict the lowest MSE values for a given parameter.

CCC	f	$\mathbf{D_a}$	$\mathbf{D_e^{\parallel}}$	$\mathbf{D_e^{\perp}}$
$\hat{\theta}^{(2)}$	0.86	0.84	0.85	0.71
$\hat{\theta}^{(2)}_{\mathrm{RBC}}$	0.96	0.84	0.95	**0.97**
$\hat{\theta}^{(1)}_{\mathrm{RBC}}$	**0.99**	**0.93**	**0.99**	**0.97**

Table 4.2: Lin's CCC of the estimates from the full acquisition protocol against the halved last shell protocol. $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ outperforms the other strategies scoring from substantial to almost perfect results in all parameters. Bold values depict the highest CCC scores for a given parameter.

4.5.5 Reproducibility analysis

Figure 4.6 shows the COV between ROIs medians within test-retest sessions for $\hat{\theta}^{(2)}_{\mathrm{RBC}}$ and $\hat{\theta}^{(1)}_{\mathrm{RBC}}$. To show the general trend, we perform a weighted linear fit, where the weights are defined by the number of voxels enclosed within each region.

We observe an improved test/retest repeatability when using $\hat{\theta}^{(1)}_{\mathrm{RBC}}$, in particular for the intra-cellular model parameters. The fit line slope remains below 0.5 in all cases except $D_e^{\parallel}$ (the slope is 1.12), implying an improved reproducibility of $\hat{\theta}^{(1)}_{\mathrm{RBC}}$ with respect to $\hat{\theta}^{(2)}_{\mathrm{RBC}}$. The average COV is 25.5% for the $\hat{\theta}^{(2)}$, 5.8% for $\hat{\theta}^{(2)}_{\mathrm{RBC}}$ and 3.12% for $\hat{\theta}^{(1)}_{\mathrm{RBC}}$.

4.6 Discussion and Conclusions

To promote the use of rotational invariants in downstream analyses, support more specific inferences, and maximize their inter- and intra-site reproducibility, we must understand the confounds in their quantification, including their robustness to experimental factors such as SNR or protocol design; as well as develop parameter estimators that have a high accuracy, precision, and robustness.

The so-called Rician signal biases are known to impact the accuracy of commonly used least-squares estimators of diffusion metrics. Various works have previously reported that ignoring the Rician data distributions in parameter estimation results in underestimated mean diffusivities and fractional anisotropy (FA), overestimated kurtosis values, and loss of angular resolution (Jones and Basser, 2004; Veraart et al., 2011). Here we complement previous works by evaluating the impact of Rician signal biases on RISH features and SMI parameters.

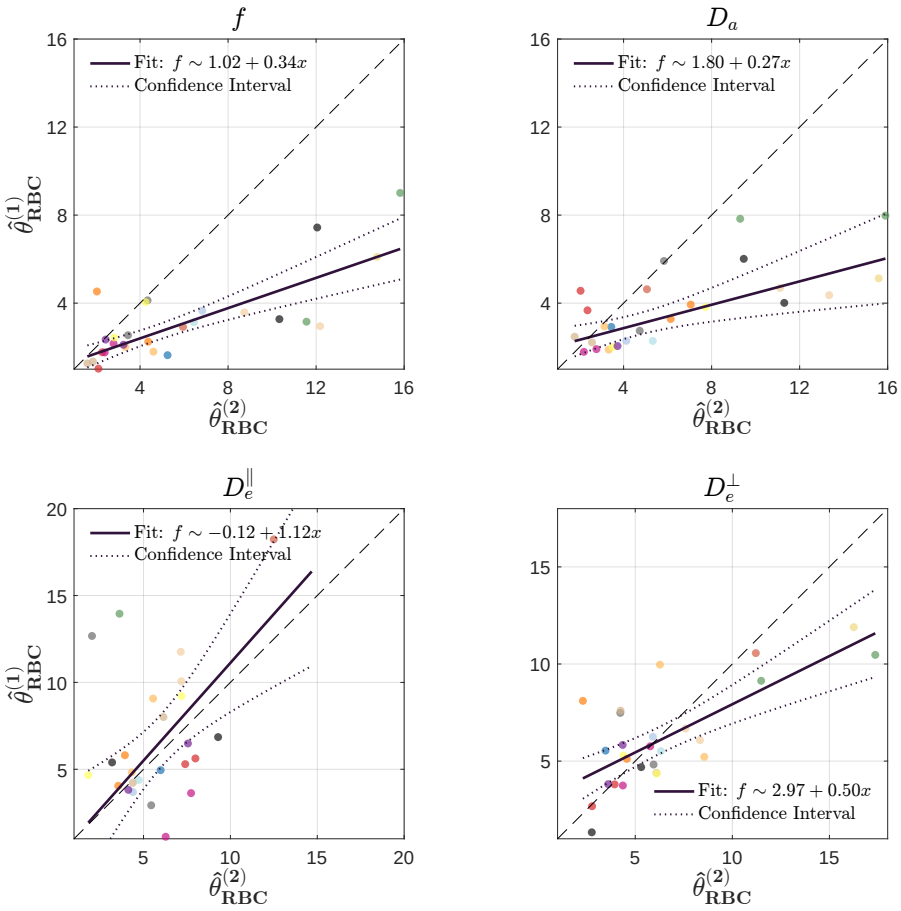


Figure 4.6: Test-Retest Reproducibility: Comparison of the COV between the two RBC-derived approaches for the four microstructural estimates from in vivo data. Colored dots depict COVs calculated from different ROIs. Dashed black line represents $x = y$ line. Black solid line represents the linear fit of the dots, weighted by the amount of voxels enclosed within each region. Dotted black lines depict the 95% confidence intervals of the fit. As shown, $\hat{\boldsymbol{\theta}}^{(1)}_{\mathrm{RBC}}$ consistently results in similar or smaller COV values, resulting in fitting lines that range from the identity line to slopes smaller than one. Thus, test-retest repeatability does not decrease with the proposed $\hat{\boldsymbol{\theta}}^{(1)}_{\mathrm{RBC}}$ strategy.

There are various strategies to mitigate the impact of the Rician signal bias. Real-valued MRI provides a promising avenue if the raw or complex MRI data is accessible (Eichner et al., 2015); however, most often, there is a need for estimators that model the data distributions explicitly. While MLE is a popular example due its favorable properties in terms of accuracy and precision, it has been shown that it is not compatible with the series of image preprocessing steps that alter the data distribution prior to fitting (Veraart et al., 2013). In contrast, RBC—a CLS estimator that offsets the model prediction by the Rician expectation value operator—has been shown to be a more accurate (Veraart et al., 2012) and efficient alternative in mitigating biases in various diffusion metrics (Veraart et al., 2012; Uhl et al., 2024), which also applies to the estimation of SHs. However, the Rician bias correction of SH is insufficient to maximize the accuracy of $\widehat{S}_2$.

We observe that $\widehat{S}_2$ is overestimated when not accounting for the Rician distribution of the DW data. The magnitude of the error depends on the SNR, underlying diffusion properties, and other factors that impact the nominal DW signal, including b-value and echo time. The impact of noise on S_2 is less trivial as multiple sources of error contribute. Overall, the Rician signal bias mainly

results in an underestimation of $\widehat{S}_2$, primarily in voxels with high anisotropy, when compared to simulations with Gaussian distributed data. Given the high correlation between S_2 and FA, it is not surprising that similar trends have been observed in previous DTI studies. However, when S_2 is low, e.g. in crossing fibers or isotropic structures, there is a significant overestimation of $\widehat{S}_2$, which is further amplified at low SNR or low number of gradient directions. Indeed, S_2 is vulnerable to a secondary source of noise bias that impacts its accuracy. All SH coefficient estimates are intrinsically noisy and their noise level depends on the gradient encoding schemes. The noise on each individual SH coefficient is asymptotically Gaussian, but the computation of the rotational invariant features skews the noise distribution if the order is larger than 0. Hence, the expectation value of S_2 exceeds its underlying value (Tristán-Vega et al., 2012; Pieciak et al., 2024). The offset depends on the underlying value and noise level of the SH coefficients, which in turn depends on the SNR of the dMRI data and the number of gradient directions. The dependency of the errors in the estimation of S_0 and S_2 on SNR, scan protocol settings, and/or microstructure is a barrier to data harmonization and reproducibility of diffusion metrics in a wide range of applications.

The RISH features are commonly used as a basis for biophysical modeling. Therefore, any of the biases above might impact the accuracy, the reproducibility and inter-site comparison of metrics derived from the models. We here focus our evaluation of such effects on SMI parameters. Given that SMI builds upon S_0 and S_2, it is vulnerable to both sources of biases. Indeed, when simulating Gaussian-distributed noise, p_2 and D_a are underestimated, while f and $D_e^{\parallel}$ are overestimated. Similar biases are observed when using a Rician-bias correcting estimator in case of Rician distributed data. The magnitude of the effects are dependent on the SNR, protocol, and underlying microstructure. The Rician bias itself might amplify or mask any of such effects, as it independently biases the parameters as follows: increased f, D_a and $D_e^{\parallel}$, and decreased $D_e^{\perp}$ and p_2. To maximize the accuracy and robustness of SMI parameters across protocols and studies, it is critical to minimize both noise biases.

Leysen et al. (2023) presented a voxel-wise rank-1 decomposition of SH coefficients of the DW signal to lower their noise level, hence improving the accuracy of higher order RISH features. Denoising the raw data or optimizing the experimental design favoring a large number of DW gradients per b-shell will also improve the accuracy of biophysical models that are derived from RISH features. While favorable, none of these approaches will resolve the effect in their entirety, and might not be feasible in limited scan times. Here, we evaluate an alternative fitting strategy for SMI that avoids the use of SH projections as an intermediate step, thereby evading the second source of noise. Despite its increased dimensionality, this alternative strategy shows an improved robustness to noise, reduced prevalence of degenerate solutions, enhanced consistency across estimates from different protocol acquisitions, and an improved test/retest repeatability scores when compared to SH-based SMI (see Figs. 4.4 to 4.6, and Tables 4.1 and 4.2). However, the accuracy of the estimator will depend on the adequacy of the selected SH order to capture the complexity of the underlying WM configuration. Therefore we recommend its use in DW data with sufficiently dense sampling of gradient directions (Tournier et al., 2013).

The intermediate use of RISH features was initially motivated by the decreased dimensionality of the model. Indeed, by disentangling the microstructural kernel from the fiber ODF, the number of parameters that are to be estimated drops significantly. However, we here show that the

simultaneous estimation of both the microstructural kernel and the fiber ODF does not negatively impact the precision of the typical microstructural parameters: f, D_a, $D_e^{\parallel}$, and $D_e^{\perp}$. Instead, following our experimental data, we observe that maps are overall cleaner, less contaminated by the significant degeneracies that are commonly observed in SMI. Moreover, the proposed method offers the opportunity for an integration of microstructure and tractography, by means of an fODF estimation without the need of a globally-estimated kernel. It must be noted, however, that the proposed model might be less tuned to the use of alternative estimators, including machine learning approaches (Nedjati-Gilani et al., 2017; Reisert et al., 2017; Palombo et al., 2020; Almeida Martins et al., 2021; Coelho et al., 2022).

A key limitation of this study is that the use of this estimator requires prior information on the data that is not always available. Prior to parameter estimation, (1) it must be verified that the data of the unprocessed magnitude MRI is Rician distributed, and (2) the noise level is assumed to be known. In our study, we adopted the MPPCA technique for noise map estimation as it has been shown to be accurate if applied to uncorrelated Gaussian distributed data. The error in the noise level has shown to be as low as 1% (Veraart et al., 2016a). In that case, the downstream impact on SMI parameters using RBC is below 0.5%. To meet that requirement, we estimated the noise level from the lower b-values only. Note, however, that other estimators might also be used if proven accurate and can be used interchangeably (e.g. Pieciak et al., 2017, under the assumption of Rician distributed data; or Henriques et al., 2023 in case of correlated noise).

Moreover, we evaluated the downstream impact of biases in RISH features on SMI parameters using realistic scan protocols. However, it is important to note that $\widehat{\boldsymbol{\theta}}^{(1)}$ only exhibits asymptotic normality and consistency (Veraart et al., 2012). As our accuracy measures are based on normality assumptions (e.g. mean of distribution), minimal scan protocols may give the appearance of inaccurate estimates. These errors, however, tend to diminish as the number of gradient directions and/or b-values increases, with estimates converging to the right solution. We further hypothesize that our findings can be translated to other models that currently employ a two-step fitting process centered around RISH features, including axon diameter mapping (Veraart et al., 2021), SANDI (Palombo et al., 2020), NEXI (Jelescu et al., 2022), and others.

In summary, we evaluated the impact of thermal noise on widely-used RISH features, and its effect on derived biophysical models such as the SMI. We propose an alternative fitting strategy bypassing RISH features and correcting for Rician bias via CLS estimators that has proven to be not only more accurate and robust, but also equal or more repeatable across sessions and more reproducible across protocol acquisitions.

The Truth Arises from Misrecognition.

— Zizek (1991)

5

Contribution: A systematic comparative study of the reliability of advanced descriptors extracted from usual signal representations

Abstract

This study aims to advance the assessment of reliability, repeatability and separability in region-based diffusion MRI by evaluating state-of-the-art EAP-derived biomarkers using a variance decomposition framework. We examine not only the consistency of measurements across multiple sessions per subject, but also the separability of individuals. Additionally, we investigate how scanner quality impacts reliability estimates. Finally, we analyze the statistical power associated with each biomarker and provide practical sample size recommendations for subsequent group comparisons.

We define repeatability, separability and reliability from a decomposition of a two-way random effects ANOVA. Two repeated-measures multishell datasets are used to assess the reliability, repeatability and separability of relevant diffusion metrics: the return probabilities (RTOP, RTAP, RTPP), non-Gaussianity (NG), Mean Squared Displacement (MSD), and **q**-space Inverse Variance (QIV).

EAP-based biomarkers show distinct behavior for the two different-quality datasets. The methods under study achieve moderate-to-high reliabilities consistently across datasets. Repeatability appears to be less affected by data quality, suggesting that improvements in data quality may primarily enhance the ability to detect inter-individual differences, rather than reducing within-subject

variability. While varying across statistical tests, datasets and biomarkers, results confirm that effect sizes are generally low among healthy population, leading to minimum sample sizes of 150-200.

We provide a reliability comparison of EAP-based biomarkers using an extended variance decomposition framework that accounts for intra-region variability and potential registration errors. Our results reveal that while most biomarkers exhibit high reliability, return probabilities and QIV emerge as the most robust measurements for group-level studies. Regarding the methods under comparison, HYDI-DSI is the most consistent across scanner quality, yielding the lowest sample size requirements. MiSFIT performs best with high-quality data, offering more regions with significant reliability and reduced sample size demands. These findings highlight the importance of both methodological robustness and data quality in biomarker selection for group-level diffusion MRI studies.

5.1 Introduction and Purpose

In pursuit of this thesis's objectives, specifically objective **SO3**, we aim to assess and compare the reliability of state-of-the-art phenomenological biomarkers by extending the variance decomposition framework used in Zuo et al. (2019). By doing so, we analyze the statistical significance that common biomarkers can attain, and determine the effect sizes each of them can achieve. Finally, we compute and evaluate the biomarkers with regards to the sample size required to ensure statistical powerful results in downstream group analyses.

As mentioned in Chapter 2, biophysical modeling, while theoretically more sensitive and specific, often results in ill-conditioned problems caused by parameter degeneracies. On the other hand, phenomenological modeling offers better conditioned problems, but might lack interpretability or specificity. Even so, phenomenological modeling has been proved useful in the assessment of disease or treatments evolution (Barnea-Goraly et al., 2004; Kraus et al., 2007; Wu et al., 2008; Madden et al., 2009, 2012; Bosch et al., 2012; Sexton et al., 2010; Clark et al., 2011; Ning et al., 2015; Fick et al., 2016b, 2017; Wang et al., 2021a; Osa García et al., 2022, etc.).

Still, the adoption of diffusion MRI by clinical settings is troublesome for several reasons, mainly derived from the economical constraints of such environments which include, but are not limited to, the extensive acquisition times and/or computational resources required. With this in mind, many researchers have aimed at developing efficient biomarkers that loosen these constraints by relying on progressively fewer assumptions. Yet, the challenge of ensuring biomarker reliability endures in clinical practice.

Reliability assessment in diffusion MRI is somewhat of an obscure science. The lack of ground truth, specifically in phenomenological modeling, pressures researchers to come up with ways of evaluating their biomarkers. Some studies end up relying on test-retest consistency estimates to assess the overall adequacy of the metrics under study (Grech-Sollars et al., 2015; Paudyal et al., 2019; McDonald et al., 2023; Ades-Aron et al., 2025). Inter-session repeatability is paramount in preclinical stages of biomarkers development, as it allows to assess treatments or disease progressions over longitudinal studies by ensuring that the changes do not come from artifacts or

noise. If a biomarker is not consistent across scanners, motion or preprocessing pipelines, it will not be adopted by clinical settings as results may reflect noise behavior rather than actual biological environment. However, repeatability is just not enough: Indeed, a marker that outputs the same value for all voxels in the brain, no matter the circumstances, will have a perfect repeatability. Separability, or the ability to distinguish between classes (or subjects), tends to be thrown out to a less relevant position. While it is true that separability alone may be considered a limited metric, it is important to recognize that diagnostics without separability is unattainable. Acknowledging this, other studies (Vollmar et al., 2010; Grech-Sollars et al., 2015; Duan et al., 2015; Venkatraman et al., 2015; Albi et al., 2017; Merisaari et al., 2019; Laguna et al., 2020; Rosberg et al., 2022) make use of intra-class correlation coefficient (ICC), which does take into account the separability amongst subjects by including a between-subject variance term (McGraw and Wong, 1996).

Most studies in the field focus on evaluating the performance of the metrics or models they introduce, often using their own criteria and datasets, rather than conducting direct comparisons with state-of-the-art alternatives. As a result, there is, to the best of our knowledge, no comprehensive systematic review that contrasts different signal representations under a common framework. An exception to this trend can be found in the case of free water elimination models, which have been more frequently benchmarked across studies (Pieciak et al., 2025). The work in this chapter aims to fill that gap by presenting a systematic and comparative evaluation of phenomenological diffusion models, establishing a foundation for consistent benchmarking and interpretation. In this context, we evaluate the performance of 5 models that differ on the number of degrees of freedom, in terms of the reliability they can attain: from the most assumption-driven and traditional DTI (Basser et al., 1994), to the single-shell AMURA (Aja-Fernández et al., 2020), the efficient MiSFIT (Tristán-Vega and Aja-Fernández, 2021), the *de facto* standard MAPL (Özarslan et al., 2013; Fick et al., 2016a), and the least-restrictive HYDI-DSI (Alexander et al., 2006; Wu and Alexander, 2007).

Our variance-based formulation of reliability offers a practical and interpretable alternative to the traditional ICC, particularly in neuroimaging settings where spatial variability and segmentation uncertainty may impact measurement quality. While ICC is designed to quantify consistency across repeated measurements using a two-way random effects model, it does not explicitly account for intra-region variability, which can arise from anatomical mismatches or partial volume effects. Our framework extends standard variance decomposition by incorporating this additional variability term, allowing for a more realistic characterization of measurement reliability in voxel- or region-based analyses. In other words, traditional ICC formulations assume that all variability not explained by differences between subjects or sessions is random noise. However, in neuroimaging applications, especially when working with location-specific measures such as voxels or anatomical regions, additional sources of variability may emerge. These include subtle misalignments between sessions, imperfections in image registration, or contamination from neighboring tissue types (e.g., partial volume effects with cerebrospinal fluid or gray matter). Given that inter-subject image registration is more complex than intra-subject registration, such errors often affect inter-subject comparisons more severely. By explicitly modeling this intra-region variability, our framework avoids overestimating reliability and provides a more accurate depiction of how stable and distinguishable a biomarker truly is across individuals and acquisitions. This makes the method especially valuable when comparing estimation models or designing experiments

where anatomical precision and inter-subject differentiation are critical.

Moreover, our framework provides natural extensions for assessing statistical significance and estimating effect sizes, which are critical for understanding measurement quality in practical terms. By expressing reliability as a ratio of variance components, we can apply statistical tests (e.g., F-tests) to determine its statistical significance—a feature not directly supported by standard ICC formulations. Additionally, by linking this variance structure to observed effect sizes, we can derive required sample sizes to detect group-level differences, enabling reliability to serve not just as a descriptive metric, but as a tool for experimental planning. This makes our approach particularly powerful in research settings where understanding both reproducibility and group discriminability is key.

5.2 Theory

As previously introduced, a thorough reliability analysis in diffusion MRI requires distinguishing between multiple sources of variability, including subject, session, and anatomical inconsistencies. In this section, we formalize a variance decomposition framework that captures these elements, adapting and extending the formulation introduced by Zuo et al. (2019). From these variance elements we define repeatability, separability and reliability metrics; which later will be used to evaluate and compare the aforementioned EAP-based biomarkers.

5.2.1 Variability sources

Our reliability assessment is based on a standard two-way ANOVA variance decomposition, from which the total sum of squares is split into four uncorrelated terms. Our methodology, depicted in Appendix D, results in three terms MS_{sep}, MS_{inc}, MS_{E} that portray, respectively, the deviation of within-subject means with respect to the overall mean (i.e., separability across subjects), the deviation of each session with respect to the corresponding within-subject mean (i.e., inconsistency across sessions) and, finally, the unexplained variability. By defining repeatability as the complementary of inconsistency (i.e. repeatability $= 1 -$ inconsistency), we obtain:

$$\text{repeatability} = 1 - \frac{\mathrm{MS}_{inc}}{\mathrm{MS}_{inc} + \mathrm{MS}_{E}}; \tag{5.1}$$

$$\text{separability} = \frac{\mathrm{MS}_{sep}}{\mathrm{MS}_{sep} + \mathrm{MS}_{E}}. \tag{5.2}$$

Note it makes sense to compare each magnitude to the unexplained variability MS_{E}, which in this case stands primarily for the inherent anatomical variation of the subject inside each tract, among other effects. For example, if the separability (i.e. inter-subject variation) is small compared to the intra-subject anatomical variability, a region-based study will not be robust, since the variability introduced by registration or segmentation will be on the same order as the true inter-subject differences. Additionally, this formalism allows to compute the statistical significance of each term by introducing the F statistic of each one in the same way we do with

ANOVA tests, by normalizing by their DOF number:

$$F_{rep} = \frac{\mathrm{MS}_{inc}}{\mathrm{MS}_E}; \tag{5.3}$$

$$F_{sep} = \frac{\mathrm{MS}_{sep}}{\mathrm{MS}_E}, \tag{5.4}$$

which will be tested for statistical significance against F distributions with $B(S-1)$ and $R-BS$ DOF and $B-1$ and $R-BS$ DOF, respectively. Where B is the number of subjects, S the number of sessions per subject and R the number of measurements (voxels) inside the region under study for a given session. It remains to define reliability itself. After the original work by Zuo, we define:

$$\text{reliability} = \frac{\mathrm{MS}_{sep}}{\mathrm{MS}_{sep} + \mathrm{MS}_{inc} + \mathrm{MS}_E}, \tag{5.5}$$

together with the F statistic:

$$F_{rel} = \frac{\mathrm{MS}_{sep}}{\mathrm{MS}_{inc} + \mathrm{MS}_E}, \tag{5.6}$$

to be tested against a F distribution with $B-1$ and $B(S-1)+R-B\,S = R-B$ DOF. Note MS_{sep} comprises both the actual anatomical differences between subjects and systematic, unwanted contaminants. MS_{inc} is an estimate of what Zuo et al. call V_r, i.e. the random noise added to our measurements as a consequence of the acquisition/estimation procedure. Besides, we add a third term (MS_E) that stands for the intra-region variability. The placement of this term in the denominator, instead of the numerator, contributes to the state-of-the-art methodology on reliability assessment, and avoids positive biases in reliability metrics.

5.2.2 Effect size, reliability and sample size

To understand how reliability impacts statistical inference, we rely on a theoretical framework that relates reliability, effect size, and required sample size under various statistical tests (Kanyongo et al., 2007). Similar to Fig. 2.15, this relationship is visualized in contour plots, where the x-axis represents measurement reliability (estimated using the methodology described in the previous section), the y-axis represents effect size, and the color-coded background denotes the required sample size needed to achieve a fixed statistical power (typically 80%) at a given significance level ($\alpha = 0.05$). For each statistical test (i.e., paired t-test, two-sample t-test, one-way ANOVA, and two-way ANOVA) a different power analysis model is used, as each test imposes different assumptions on data structure, variance decomposition, and DOF.

Importantly, when reliability is strictly less than 1, the observed effect size must be corrected to reflect the attenuating influence of measurement noise. Specifically, the observed Cohen's d is adjusted using the factor $\sqrt{R}$, where R is the estimated reliability, yielding a corrected effect size: $d_{\text{corrected}} = d_{\text{observed}} \cdot \sqrt{R}$. This adjustment accounts for the inflation of the variance due to unreliability and ensures that the sample size calculations accurately reflect the true discriminative power of the measure. The resulting plots thus serve as a visual tool to evaluate the interplay between a method's reliability and its sensitivity to group differences, allowing one to estimate

the minimum sample size required to detect statistically significant effects in real-world scenarios.

5.3 Materials

5.3.1 In vivo datasets

This work presents a reliability study based on ZJU and MICRA multiple-session datasets, both introduced in Section 1.4.

Data Preprocessing: MICRA was preprocessed by removing Gibbs ringing artifacts (Kellner et al., 2016), via MRtrix3; (Tournier et al., 2019) and correcting susceptibility-induced distortions (with FSL's `topup`; Andersson et al., 2003; Smith et al., 2004) and B1 field inhomogeneity (Zhang et al., 2001) with MRtrix3. ZJU database was publicly shared in a preprocessed variant. The preprocessing pipeline first covered an MP-PCA-based noise removal (Veraart et al., 2016b), a Gibbs ringing artifacts correction (Kellner et al., 2015), an estimation of susceptibility-induced distortions via FSL's `topup` and, finally, correction of both head movements and eddy current distortions with FSL's eddy. No further preprocessing was performed.

Additionally, to isolate and assess the impact of scanner quality on reliability estimates, a downsampled version of the MICRA dataset was created to closely mimic the acquisition protocol of the ZJU dataset. This involved reducing both the number of diffusion shells and the number of diffusion gradient directions per shell in MICRA. Specifically, the reduced MICRA subset retained only the three inner shells ($b = [1200, 2400, 4000]$ s/mm^2), with each shell subsampled to 30 diffusion directions. By matching the sampling scheme of ZJU while preserving MICRA's inherently higher-quality scanner data, this design allowed us to evaluate whether differences in reliability stem from acquisition protocols or underlying scanner quality (e.g. gradients strength).

5.4 Methods

5.4.1 Segmentation

The FA was computed for all subjects and sessions of the database from an in-house nonlinear, non-negative estimation pipeline of the tensor (Tristán-Vega, A. and Aja-Fernández, S. and París, G., 2022) where only those samples with $b \leq 2,000$ s/mm^2 were used. The FA maps were iteratively registered to the ENIGMA template (Thompson et al., 2014) following the procedure described in Jahanshad et al. (2013), where, at each iteration, all warped FA maps are averaged to construct a refined FA atlas. Following the method outlined by Tristan-Vega et al. (2008), we used an atlas that had been aligned with the ENIGMA template but adapted to the specific features of our dataset. This customized atlas then served as the new registration target. This process was repeated five times, and all subjects and sessions were then registered to the final atlas to estimate definitive warping fields. The warping fields were taken into account to map back the 48 labels of

the JHU atlas segmentation (Oishi et al., 2008) to each subject/session's native space. Note that, since the native resolution of the MICRA and ZJU datasets differ from the ENIGMA template, the voxel populations in the labels are a fraction of what they were in the original datasets.

5.4.2 Computation of metrics

Metrics were all computed using the in-house implementation (Tristán-Vega, A. and Aja-Fernández, S. and París, G., 2022). In all cases, only those samples with $b \leq 2,000$ s/mm^2 were used to estimate the DT and derived parameters (such as anisotropic scaling in MAPL and HYDI-DSI). Other model-specific design parameters are listed below:

AMURA: Given that the model works on a single-shell basis, the reliability results were assessed via MANOVA across b-shells, using subject and session as predictors. Roy's largest root was used as test statistic.

MiSFIT: The maximum order for the SH coefficients was fixed to $L = 6$, while the Laplacian energy-based penalty weighted by λ was fixed to 0.001. MiSFIT enforces the EAP to be strictly positive by using a similar approach as in Cheng et al. (2014).

MAPL: The maximum order for the basis functions was fixed to $N = 6$, with anisotropic scaling and a Laplacian energy-based penalty weighted by λ. This parameter was estimated using GCV with greedy search over the logarithmic domain, as in París et al. (2023b). After the coefficients were found, the scalar measurements were computed, together with the SH coefficients of the ODF up to order 8 using a mixed analytical-numerical technique (the radial integrals were analytically computed for a set of 128 evenly spaced orientations, then the SH coefficients were fitted using the approach described in Descoteaux et al., 2007).

HYDI-DSI: The in-house quadratic-programming (QP) implementation by Tristán-Vega et al. (2023) was used (i.e., HYDI-DSI-QP). The grid lattice was set to $[4, 4, 4]$. The Laplacian energy-based penalty was weighted by a term λ, which was estimated using GCV with greedy search over the logarithmic domain (París et al., 2023b).

5.5 Results

The tables E.1–F.6 in the Appendices E and F show the detailed results for MICRA and ZJU datasets, respectively. Each table represents one measurement, and each row one region. The performance of each EAP-based biomarker is directly represented in the rows, where the best result has been highlighted in boldface, and the worst in italics. Moreover, we used color-coding to easily see the statistical significance of all methods: orange numbers represent those results with p-values below 0.05 but over 0.01 (statistically significant results); red numbers represent those results whose p-values are over 0.05 (not statistically significant); finally, uncoloured numbers (black) represent those results whose p-values are less than 0.01 (i.e. highly statistically significant).

Note that, taking into account the logic behind the repeatability and separability computation (see Section 5.2), orange and red numbers are good results in the repeatability column, whereas bad results in separability and reliability columns.

To ease the interpretation, we include a graphical representation of the tables (see Figs. 5.1a and 5.1b), where the behavior of the metrics across regions is averaged in a 2D boxplot showing 25%–75% percentiles. Results are weighted by the number of voxels within each region, and do not include significance level at the regional level.

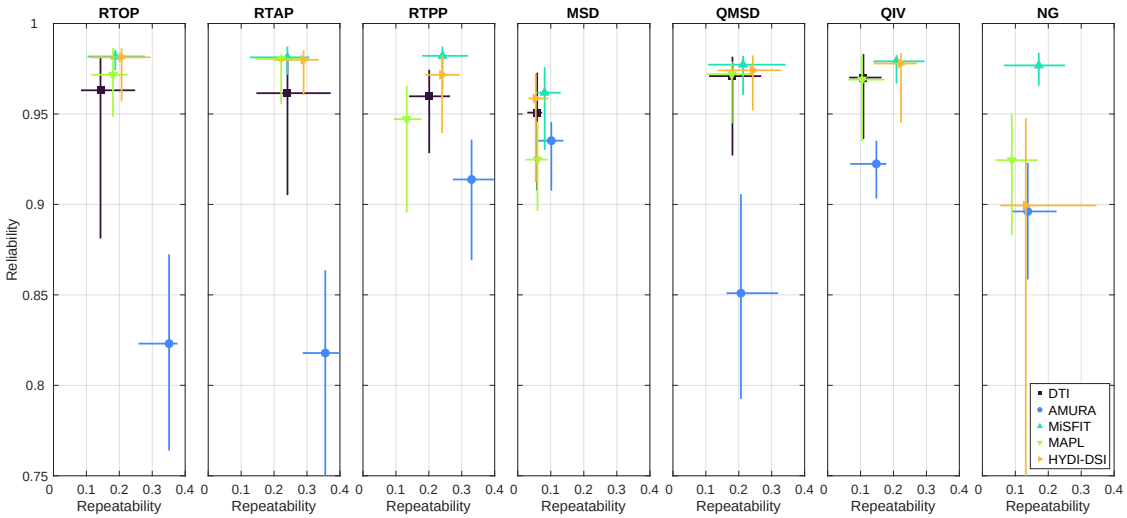
5.5.1 Reliability of biomarkers

Figures 5.1a and 5.1b show the reliability and repeatability of the methods across measures for the MICRA and ZJU datasets, respectively. The crosses width and height depict the 25th–75th percentiles of the distributions, intersecting in their median.

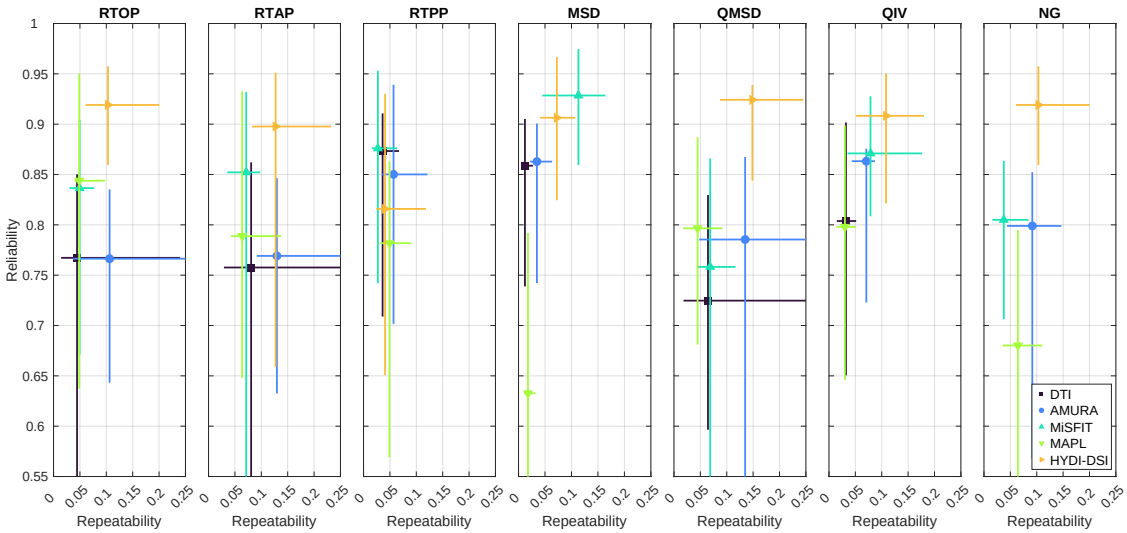
The first thing to notice in Fig. 5.1a is that higher repeatability (right-bound) does not necessarily imply higher reliability (top-bound). Repeatability might be a byproduct of a general lack of adaptability to the between-subject differences. An example of this can be seen in MICRA's AMURA: it scores high in repeatability (the highest in RTOP, RTAP, RTPP, and MSD) yet its low separability makes it the less reliable marker across all but one metric (MSD). This difference is mitigated in the ZJU dataset (Fig. 5.1b), possibly due to the smaller samples available per subject, which hinders the adaptability of the methods to subject differences. Note, however, that AMURA processes each shell independently, as it operates on a single-shell basis. This may limit its comparability with fully multi-shell models, as it integrates results from all shells separately.

The most reliable methods are MiSFIT and HYDI-DSI, scoring the highest in MICRA and ZJU datasets, respectively. Across all measures, MiSFIT scores an average of 0.98 in MICRA and 0.84 in ZJU. In other words, 98% or 84% of the variability within MiSFIT's results can be attributed to inter-subject differences. HYDI-DSI, on the other hand, scores an average of 0.96 in MICRA and 0.90 in ZJU. The lowest reliabilities are obtained by AMURA (with an average across measures of 0.88 in MICRA) and MAPL (with an average of 0.76). In fact, MAPL's reliability experiences the deepest impact when changing the datasets, falling 20.4% of its value in MICRA. Even though this fall in reliability is generalized across all methods, it is less relevant: DTI falls 17%, MiSFIT 13.4%, AMURA 7.5% and HYDI-DSI 6.7%.

Table 5.1 shows the amount of ROIs that result in highly statistically significant reliability estimates ($p < 0.01$), for all methods–markers pairs. As can be observed, all ROIs under study yielded highly statistically significant results for both MiSFIT and MAPL when using the MICRA dataset. However, when using the ZJU dataset, MAPL drops to 93% of the regions, while MiSFIT still obtains significant results in 95% of them.



(a) MICRA's reliability comparison across models: Reliability vs. repeatability across models for each biomarker under study. Crosses depict 25%-75% percentiles, and intersect in the median.



(b) ZJU's reliability comparison across models: Reliability vs. repeatability across models for each biomarker under study. Crosses depict 25%-75% percentiles, and intersect in the median.

Figure 5.1: Reliability comparison between models.

5.5.2 Displacement between datasets

Figure 5.2 shows the displacement map that reliability and repeatability suffer when using different quality datasets. In this figure, MICRA dataset has been reduced to match ZJU using the strategy described in Section 5.3. For clarity, each method has been assigned a numerical label as follows: 1 – DTI, 2 – AMURA, 3 – MiSFIT, 4 – MAPL, and 5 – HYDI-DSI. A blue line depicts the impact on repeatability and reliability for a given method.

There exists a clear reduction of both reliability and repeatability (samples move towards bottom-left corner) when reducing the quality of the datasets, though the fall in reliability is less pronounced (an average of 13.8% across all methods and measures) than repeatability

Method	MICRA (% Significant Regions)								ZJU (% Significant Regions)							
	RTOP	RTAP	RTPP	MSD	QMSD	QIV	NG	**Total**	RTOP	RTAP	RTPP	MSD	QMSD	QIV	NG	**Total**
DTI	90	90	100	100	90	100	–	96	83	81	96	98	81	98	–	91
AMURA	94	94	100	100	96	100	100	98	73	73	100	100	83	100	98	90
MiSFIT	100	100	100	100	100	100	100	**100**	94	96	96	92	96	96	96	**95**
MAPL	100	100	100	100	100	100	100	**100**	98	83	92	98	92	98	90	93
HYDI-DSI	100	100	98	100	98	98	83	97	92	90	96	94	88	94	92	92

Table 5.1: Percentage of white matter regions with highly statistically significant reliability ($p < 0.01$) for each method and measure in the MICRA and ZJU datasets.

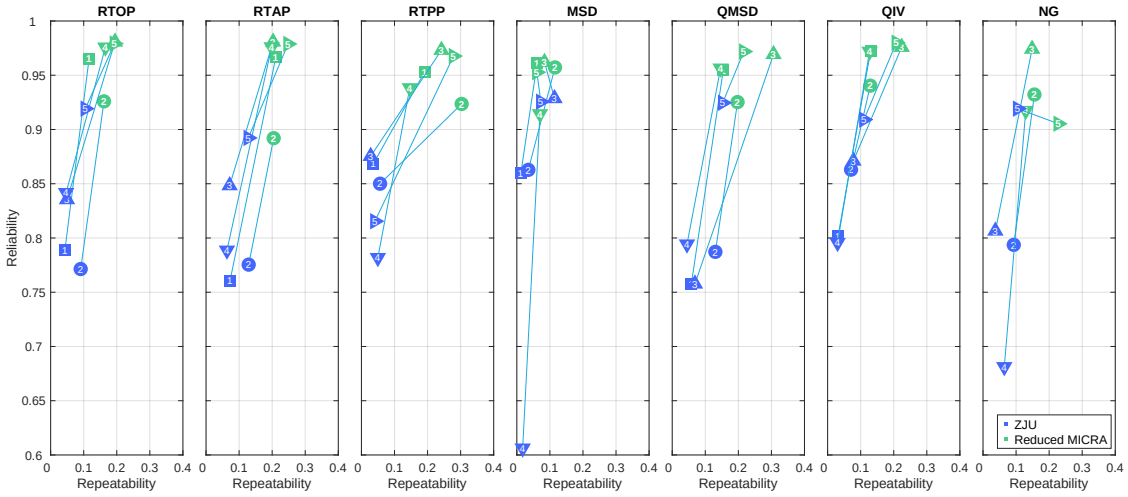


Figure 5.2: Reliability vs. repeatability displacement across datasets: Displacement of reliability and repeatability values for reduced MICRA (green) and ZJU (blue) datasets. Numbers represent each model: DTI (1), AMURA (2), MiSFIT (3), MAPL (4), HYDI-DSI (5). The blue line depicts the impact on repeatability and reliability for a given method.

(averaging 61.2%). In terms of reliability, MAPL still suffers the most when dealing with lower quality datasets, falling a 19.9% from the subsampled MICRA to ZJU. When studying the general behavior of the measures across the datasets, we observe a consistent reduction of reliability of 0.137 ± 0.018 (mean ± std).

On the other hand, Fig. 5.3 illustrates the changes in reliability and repeatability measures when comparing models fitted on the full MICRA dataset versus its subsampled counterpart.

The behavior is markedly more chaotic, with fluctuations that lack a consistent direction across both diffusion models and microstructural measures. Specifically, reliability estimates do not follow a monotonic trend: for example, AMURA exhibits a substantial average decrease of -4.81% in reliability, whereas HYDI-DSI increases modestly by +0.10%. Similarly, MiSFIT and MAPL show average gains (+0.34% and +0.56%, respectively), while DTI remains nearly unchanged (+0.04%). Across models, the variation is subtler, yet still inconsistent: RTOP and RTAP show slight decreases (-2.12% and -1.47%), while RTPP exhibits a marginal increase (+0.37%). These shifts suggest that denser sampling's impact on measurement stability is not uniform across different modeling frameworks. This lack of systematicity is visually evident in the figure, where

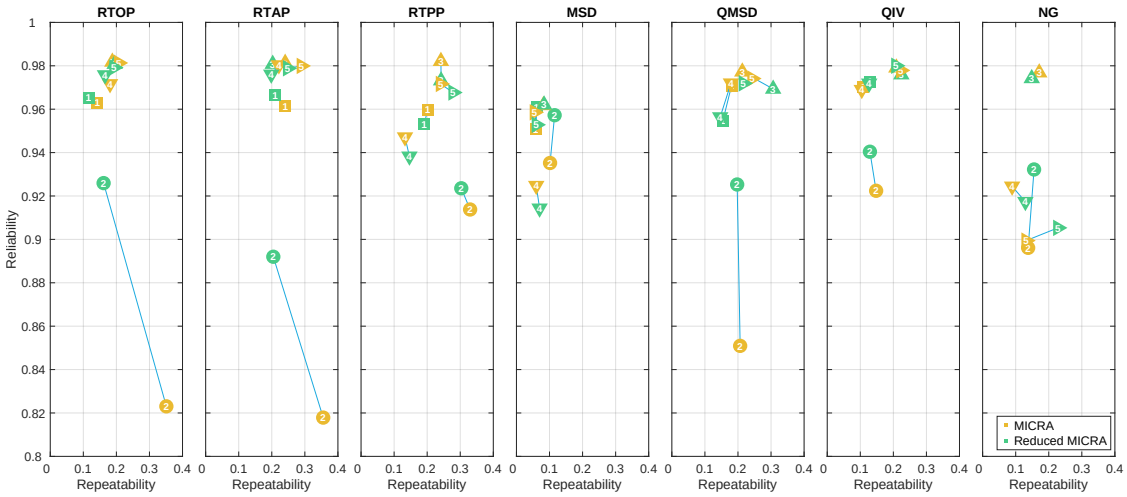


Figure 5.3: Reliability vs. repeatability displacement across datasets: Displacement of reliability and repeatability values for MICRA (yellow) and reduced MICRA (green) datasets. Numbers represent each model: DTI (1), AMURA (2), MiSFIT (3), MAPL (4), HYDI-DSI (5)

trajectories of reliability and repeatability crisscross rather than aligning in parallel or opposing trends (as in Fig. 5.2). The results also indicate a more subtle change in the reliability effect across sampling schemes than across datasets of varying quality.

5.5.3 Effect and sample sizes

Figure 5.4a shows the relationship between reliability, effect size, and sample size, with the behavior of the five models overlaid according to their ZJU reliability. The relationship between reliability, effect size and sample size was estimated using the analytical solution of a two-sample t-test power function (=0.05, power =80%), as in (Zuo et al., 2019). The lines effectively convey the minimum detectable effect size for a given number of subjects. The x-axis represents the sample size—or subjects per group—and the y-axis depicts the effect size in standard deviation units (SDUs). The SDUs are used here so that the comparison of aggregate behavior across measures is fair, as they account for the normalized variances. ZJU results are used here instead of MICRA's as they provide a wide range of reliabilities and thus, yield bigger differences in terms of effect and sample sizes.

As expected, higher reliability models (i.e., HYDI-DSI and MiSFIT) require less subjects for the same detectable effect size. This effect is barely noticeable at low subjects per group, but its impact grows with sample size. For example, HYDI-DSI can detect group differences of 1.2 SDUs with 14 subjects, whereas MAPL or DTI require 16 subjects. Lower effect sizes, however, yields higher sample size differences models with distinct reliability: to detect a group difference of 0.8 SDUs, MAPL requires 34 subjects, while HYDI-DSI only needs 29.

This difference is stressed in Figures 5.4b and 5.4c—inspired in the work by Zuo et al. (2019), as previously mentioned. These two figures are shown to visualize the relationship from another perspective, but in two different points in the curve of Fig. 5.4a. That is, Fig. 5.4b shows the neighborhood of the effect size of 1.2 SDUs, in terms of reliability (x-axis) and sample size

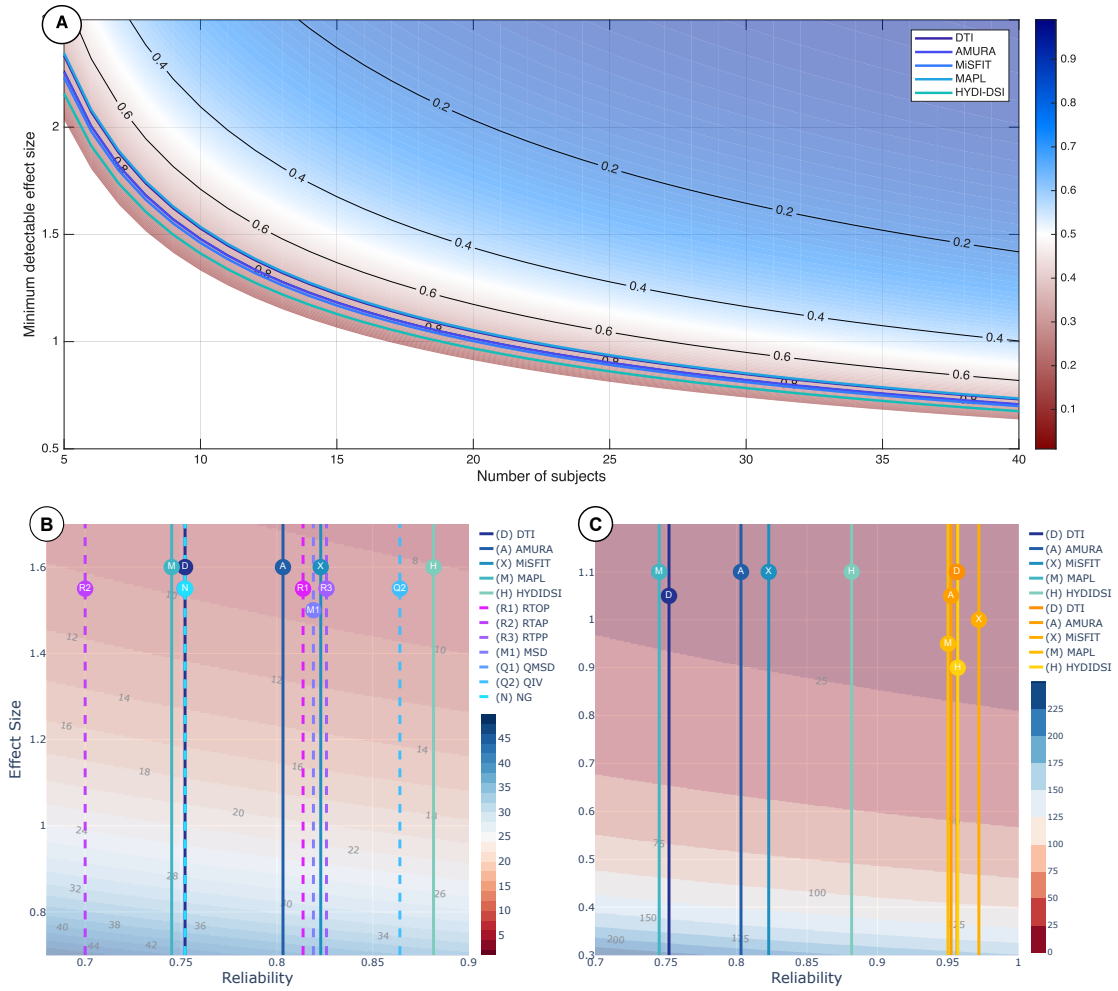


Figure 5.4: Effect size and reliability across models and measures. Following the reliabilities estimated per model, the minimum detectable effect size is estimated in terms of the sample size, or number of subjects (Fig. 6a). Solid lines depict the models reliabilities, and the color-coded background represent, together with the isocontours, the attainable effects for a wide range of reliabilities. Figures 6b and 6c describe the same curves when changing the axis (i.e., representing effect size against reliability and color-encoding the background and isocontours in terms of sample size.

(coloured background). Figure 5.4c on the other hand, shows the neighborhood around 0.8 SDUs.

Figure 5.4b shows the general behavior of measures (aggregated across models and regions) resulting from the ZJU dataset. To ease the readability, measures are represented by dashed lines and labeled as RTOP (R1), RTAP (R2), RTPP (R3), MSD (M1), QMSD (Q1), QIV (Q2), and NG (N). Models are represented by solid lines (following the color-coding in Fig. 5.4a) and labeled as in previous sections.

According to the Fig. 5.4b, QIV yields the highest (0.86) reliability across models, whereas RTAP results in the lowest (0.70). As expected, this difference leads to a range of minimum effect sizes that can be detected for a given number of subjects. Similar to the models previously explained, when the effect size to be detected is high, sample size requirements barely change (e.g.,

to detect 1.2 SDUs, RTAP requires 17 subjects, whereas QIV only 15). In the same way, detecting lower effect sizes (e.g., 0.8) requires a minimum of 36 subjects for RTAP and 30 subjects for QIV.

Finally, Fig. 5.4c shows the original ZJU lines (in the same color-coding as in previous Figs. 5.4a and 5.4b) for the effect size neighborhood around 0.8 SDUs. As mentioned, sample size requirements vary abruptly when trying to detect low effect sizes in models with different reliabilities. Here, we also show the model's behavior when using MICRA dataset (solid lines in yellow colors) to see the effect of data quality in effect and sample size estimations. Consistent with previous Sections, MICRA data results in excellent reliabilities across models (they all score higher than 0.95). Consequently, models that are more affected by data like MAPL or DTI, experience a pronounced change in terms of number of subjects for a given effect size when data quality varies.

5.6 Discussion and Conclusions

The widespread adoption of diffusion MRI biomarkers in clinical practice has been hindered by challenges in accurately estimating their reliability. Reliability estimation plays a crucial role in validating biomarkers, ensuring that they provide consistent and reproducible measurements across different individuals, sessions, and scanners. Despite significant advancements in diffusion MRI, the lack of standardized and robust methods for assessing reliability has limited the ability to confidently translate these biomarkers into clinical use.

In this chapter, we assess the reliability of diffusion MRI biomarkers in MICRA and ZJU datasets, by extending the methodology proposed by Zuo et al. (2019), whose reliability estimation method enables the decomposition into repeatability and separability. Repeatability quantifies consistency within the same condition, while separability evaluates the ability to differentiate between different groups, subjects or conditions. Importantly, the proposed approach eases the estimation of inter-class differences and the required sample sizes for group-level studies, which is crucial in the design of statistically powered experiments. While closely related to the ICC, our method emphasizes spatially localized reliability estimates across regions, helping to account for intra-region variabilities such as those introduced by misregistration. In addition, it offers a path to derive statistical quantities such as effect sizes and p-values, supporting more comprehensive group comparisons.

An important finding in this work is that repeatability alone is not a sufficient indicator of a method's reliability. While some measures, such as the return probabilities, exhibit high repeatability—with repeatability variance often not statistically significant—this does not necessarily imply high reliability. A clear example of this is AMURA, which shows the highest repeatability for RTxP but, paradoxically, the lowest reliability by a large margin. This result underscores the risk of drawing misleading conclusions when relying solely on repeatability metrics, such as test-retest variability.

5.6.1 On biomarkers' reliability

The MICRA and ZJU datasets differ across several acquisition parameters, including the number of subjects, sessions per subject, diffusion gradient directions and strengths, as well as the magnitude of diffusion weighting (see Section 1.4). In all these aspects, MICRA represents a more comprehensive and higher-quality dataset, so higher reliabilities are expected.

MICRA results confirm that reliability is generally high, often exceeding 90%, given adequate and high-quality data. A more detailed analysis reveals significant differences across methods, highlighting key insights into their performance and limitations. Among the evaluated methods, MiSFIT stands out as the most robust, consistently demonstrating the highest reliability.

ZJU results, on the other hand, confirm that reliability is strongly impacted by dataset adequacy. Reliability falls consistently an average of 13.8% across methods, often scoring in a range from 75% to 90%. In this case, the highest reliability is achieved by HYDI-DSI, consistently overperforming other models in all but 2 metrics: RTPP and MSD, in which MiSFIT results in more reliable estimates. MAPL's reliability suffers the most when dealing with poorer quality datasets, falling down from 95% with MICRA to 74% with ZJU.

Certain methods exhibit clear limitations across both datasets. AMURA generally performs worse than the other techniques, possibly due to its design being tailored for single-shell acquisitions. This is mitigated when employing datasets with fewer shells from which other techniques can benefit. A similar argument applies to DTI, which operates on a limited data subset, as b-values above 2000 s/mm^2 must be excluded. When comparing the three higher-order methods—MiSFIT, MAPL, and HYDI-DSI—across datasets, MiSFIT and HYDI-DSI consistently outperform MAPL. MiSFIT places a higher upper bound on reliability, exploiting the dataset adequacy. HYDI-DSI, on the other hand, is more consistent across datasets, outperforming other methods in ZJU and coming close to MiSFIT in MICRA.

Regarding the different measurements, MICRA's RTxP and QIV (or qMSD, since QIV = qMSD^{-1}) emerge as the most robust, achieving a balance between high repeatability and reliability. ZJU results in RTPP and MSD being the most robust metrics. MSD exhibits the poorest repeatability among all the analyzed metrics, yet its reliability is often over 85%, further reinforcing the importance of reliability as a key evaluation criterion.

5.6.2 On dataset quality

The effect of data quality is primarily examined in Figs. 5.2 and 5.3. However, this effect is also perceived in downstream analyses (Fig. 5.4).

While ZJU and the subsampled version of MICRA datasets were matched in terms of the number of subjects, sessions, and diffusion sampling scheme (i.e., number of shells and gradient directions per shell), they differ substantially in acquisition hardware. Specifically, MICRA was acquired with a Connectom scanner equipped with ultra-strong gradients (300mT/m), whereas ZJU was acquired on a scanner with conventional gradient strength (80mT/m). This key difference enables MICRA

to achieve the similar diffusion weighting with shorter gradient durations, leading to shorter echo times and, consequently, higher SNR. Additionally, MICRA images are acquired at a lower spatial resolution (2 mm^3 vs. ZJU's 1.5 mm^3), a factor shown to influence reproducibility assessment (Zhong et al., 2023), and which further amplifies the SNR advantage relative to the ZJU dataset.

The observed drop in reliability for the ZJU's lower-quality data appears to be primarily driven by a marked decrease in repeatability, while separability remains relatively stable. This suggests that data quality has a stronger impact on within-subject consistency than on between-subject variability. In other words, noisier acquisitions degrade the precision of repeated measurements from the same subject, reducing repeatability, whereas the distinctiveness between subjects is less affected, as it reflects underlying biological differences that are more robust to random noise.

These findings are aligned to the behavior of reliability and repeatability measures across MICRA dataset and its subsampled version (Fig. 5.3), where decreasing (or increasing) both the shells and the number of diffusion gradients per shell does not result in consistent reliability, repeatability or separability changes. On the contrary, the results follow a more unpredictable variability.

5.6.3 On effect and sample sizes

Effect sizes quantify the magnitude of differences between populations, providing a direct measure of how distinct two groups are in terms of a parameter. This concept is particularly related to sample size—or the number of observations included in the analysis—since the ability to detect a given effect size with statistical significance depends on how large the data sample is. Thus, estimating plausible effect sizes comes hand in hand with determining the sample size required to achieve sufficient statistical power. Effect sizes of events of interest (e.g., diseases that affect specific brain regions, and possibly alter physiological processes, and the ability up to which a measure can accurately identify such effects) can vary in different ways. These variations are ultimately tied to clinical aspects, and are out of the scope of this work.

From a preclinical point of view we can, instead, focus on reliability. Reliability mediates the relationship between effect and sample size. An unreliable model will yield noisier outputs—results vary randomly for the increased effect of unwanted variabilites—so the true difference between groups will be masked. In practice, a moderately reliable model will require bigger sample sizes for the same effect size. Thus, assessing reliability in these terms becomes paramount, as a model or measure that requires larger amounts of samples to detect the same effect sizes will unlikely succeed in its translation into clinical settings.

Inspired by the work of Zuo et al. (2019), we have assessed the effect of reliability on the propagator-based models' and measures' sample and effect sizes. By using an analytical solution of a two-sample t-test power function ($\alpha = 0.05$, statistical power of 80%), we effectively portray the relationship between the three factors—reliability, effect size and sample size—resulting in Fig. 5.4. Given that most the models and measures under study yield moderate-to-excellent reliabilities, the attainable effect sizes for a given sample size (or, in other words, the required sample sizes for a given observed effect size) are similar. However, this is only true when the magnitude of the effect is big (e.g., effect size $\sim$ 1.2 SDUs): detecting weaker events (e.g., effect

sizes ~ 0.8 SDUs) result in big differences in sample size requirements. In fact, the growth of sample size n as the effect size d decreases follows a quadratic relationship ($n \propto 1/d^2$).

These results provide a useful framework for optimizing study design and planning, offering quantitative guidance on how models and measures that yield a wide range of reliabilities might affect crucial aspects such as attainable effects or required sample sizes.

5.6.4 Limitations

Despite offering a comprehensive comparison across diffusion MRI biomarkers and modeling strategies, this study presents several limitations. First, all analyses were conducted exclusively on healthy participants, which, while allowing controlled estimation of variability and test-retest behavior, may not fully capture the heterogeneity or signal characteristics present in clinical populations. Consequently, generalizing sample size requirements or effect size expectations to patient groups must be done with caution. Second, the scope of reliability assessments relies on specific acquisition protocols, hardware configurations, and pre-processing pipelines, which may not be representative of general practices across sites or vendors. Finally, while we emphasize statistical power and reproducibility, the study does not account for biological validity or ground-truth correspondence, leaving open questions about the biomarkers' true specificity to underlying microstructural processes.

5.6.5 Conclusion

This work emphasizes the critical need for a comprehensive assessment of biomarker reliability in diffusion MRI and provides a comparison of EAP-based biomarkers derived from state-of-the-art models. In extending the variance decomposition framework by Zuo et al. (2019), we aim to account for intra-region variability and potential registration errors. We identify clear differences in biomarkers' reliability profiles, highlighting the need for an assessment beyond repeatability. Our results show that diffusion biomarkers tend to have high reliability. With regards to the models, HYDI-DSI provides the most consistent reliability across different quality datasets, yielding the lowest sample size requirements on average. In contrast, MiSFIT excels when data quality is high, producing the highest number of regions with statistically significant reliability and reduced sample size requirements. These findings underscore the relevance of considering both methodological robustness and acquisition quality when selecting biomarkers for group-level studies. The framework presented here offers a practical approach to quantify effect sizes and guide study design in future diffusion MRI research.

Words are never 'only words'; they matter because they define the contours of what we can do.

— Žižek (2006)

6

Discussion and Conclusions

6.1 Discussion

Diffusion MRI (dMRI) has seen significant advancements in the past decades, but its translation into standard research and clinical settings remains hindered by technological, economic, and practical constraints. While high-end laboratories and specialized centers push the boundaries of microstructural imaging, the widespread adoption of state-of-the-art biomarkers is still limited by factors such as scanner availability, acquisition time, and computational resources.

This thesis focuses on addressing these limitations by:

- Developing novel biomarkers that maintain clinical and research relevance while being feasible within standard imaging environments.
- Mitigating the effects of thermal noise to enhance the stability and reproducibility of microstructural biomarkers, particularly in settings where acquisition conditions are suboptimal.
- Extending reliability assessment methodologies to ensure that neuroimaging biomarkers are not only repeatable but also reliable and applicable in real-world conditions.

By tackling these three key aspects, this work contributes to closing the gap between cutting-edge research and practical clinical applications.

6.1.1 On efficient biomarkers development

The expansion of dMRI has led to a growing number of diffusion-based biomarkers, each attempting to capture different aspects of brain microstructure or phenomena. However, many of these metrics are inaccessible in standard clinical and research environments due to their reliance on high-field scanners, time-intensive acquisitions, or complex computational models. This has created a technological divide, where certain biomarkers remain confined to specialized centers, limiting their broader adoption.

A central aim of this thesis was to develop advanced yet efficient biomarkers—ones that retain the diagnostic and scientific value of complex models but can be implemented within the constraints of standard imaging protocols. This led to the proposal of two new metrics within the MiSFIT phenomenological model: PA and NG, which quantify anisotropic and non-Gaussian diffusion behaviors, respectively. Both metrics are defined over the EAP, allowing them to capture essential diffusion properties while remaining computationally efficient.

In relation to **SO1** (i.e., developing advanced and efficient biomarkers for the brain's white matter), our findings demonstrate that phenomenological modeling remains a viable and promising approach for clinical neuroimaging. By refining MiSFIT and introducing PA and NG, we show that accurate and reliable diffusion characterizations can be achieved without the need for high-end hardware or computationally expensive processing. These results challenge the assumption that only biophysical models can provide valuable information, reinforcing the role of phenomenology in practical neuroimaging.

However, despite these advancements, clinical acquisition constraints remain a limiting factor. Given the strict regulations on scan duration and patient comfort, as well as economic factors influencing hospital scanner capabilities, simple tensor models like DTI may continue to dominate. Overcoming this will require further efforts to streamline acquisition protocols while ensuring that richer diffusion metrics can still be extracted.

6.1.2 On thermal noise effects in biophysical modeling

The push toward biophysical modeling has been largely driven by the promise of more accurate and biologically interpretable biomarkers. However, these models are highly susceptible to acquisition-related biases, particularly those stemming from noise.

This thesis specifically investigated the impact of thermal noise on rotationally invariant fitting procedures, which are commonly used in dMRI biophysical models. These models often assume Gaussian noise distributions, ignoring the intrinsic Rician nature of magnitude-reconstructed MR signals. As a result, their parameter estimates become both biased and dependent on acquisition conditions, introducing inconsistencies in biomarker interpretation.

Our second contribution demonstrates that RISH-based modeling suffers from systematic biases due to noise, potentially affecting the accuracy of microstructural estimates. Particularly, the biases that stem from the computation of higher-order rotational invariant features not only

are dependent on SNR, but also on protocol acquisition and microstructural arrangement of WM fibers. These interactions render the RISH-based biophysical models, aside from inaccurate, weakly reproducible. To address this, we propose two potential solutions: (1) Incorporating noise-informed statistical priors into the estimation process to counteract the effects of thermal noise; or (2) Avoiding RISH-based fitting procedures, which are particularly susceptible to noise biases. These improvements directly align with **SO2**: enhancing the reliability and stability of microstructural biomarkers by mitigating thermal noise effects. By reducing bias, as well as their interaction to experimental settings, our findings contribute to more robust and generalizable biophysical estimates, which is essential for their broader adoption.

Beyond methodological refinements, these results have significant clinical implications. In standard hospital environments, scanner quality and SNR are often suboptimal, meaning that noise effects are even more pronounced than in research settings. Thus, ensuring that biophysical models are resilient to thermal noise effects and acquisition-related variability is a necessary step toward their real-world applicability. Additionally, these findings highlight the importance of protocol standardization, as variations in acquisition settings can introduce artifacts that may not be immediately apparent in controlled research environments. In this regard, further work is required to quantify the practical impact of these corrections across different scanner types and acquisition protocols.

Nonetheless, while our proposed solutions mitigate the problem, they do not entirely eliminate it. Magnitude-reconstructed dMRI signals will always be affected by the presence of Rician bias, and any correction remains subject to the accuracy of the noise estimation procedures. Furtermore, the solutions are based on the assumption of Rician distributed signals and, as such, prior signal processing steps that alter the underlying noise characteristics may compromise their effectiveness.

6.1.3 On biomarkers reliability

The rapid expansion of dMRI research has led to an influx of new biomarkers, yet rigorous reliability assessments remain a challenge. Most validation efforts focus on test-retest repeatability, but repeatability alone is insufficient—biomarkers must also be reliable across different acquisition conditions, scanners, and patient populations. This thesis proposes an alternative reliability assessment framework, grounded in specificity and sensitivity—two core concepts in clinical diagnostics. By translating these principles into the evaluation of dMRI biomarkers, we define a methodology based on:

1. Repeatability (consistency across multiple scans),
2. Separability (ability to differentiate meaningful biological differences),
3. Reliability (overall robustness in various conditions).

Chapters 3 and 4 offer a partial reliability assessment that focus not only on repeatability but also on separability, showing that repeatability alone is not sufficient to assess a biomarker's reliability. Chapter 5 systematizes and generalizes the reliability evaluation and applies it

to state-of-the-art EAP-based phenomenological biomarkers. This distinction is critical, as some biomarkers may demonstrate high repeatability yet low separability, leading to misleading reliability assessments. We also identify key factors affecting reliability, including data sampling strategies (e.g., DTI's limitation in high-b acquisitions) and model constraints (e.g., positivity constraints improving output stability).

Importantly, our results highlight the central role of dataset quality in shaping reliability estimates. High-quality acquisitions, as in MICRA, enable more stable within-subject measurements and thus higher repeatability and overall reliability. Conversely, lower-quality data—such as those acquired with conventional gradient strengths—tend to compromise repeatability while leaving separability relatively unaffected. This imbalance leads to a reduction in overall reliability and, consequently, an increase in the sample size required to detect statistically significant effects. In this context, reliability acts as a direct mediator between data quality and statistical power.

We also explored how reliability and effect sizes influence sample size considerations. Even with moderate-to-highly reliable models and measures, detecting smaller effects requires substantially larger sample sizes, whereas larger effects can be detected with similar sample sizes across most models. Importantly, as reliability decreases, the required sample sizes increase substantially, highlighting the need for robust acquisition protocols and careful model selection in study design. While we do not provide strict numerical boundaries, our results offer a quantitative framework linking reliability, effect size, and sample size, which can guide planning and optimization of future studies. In clinical populations, where effect sizes may be larger or more variable, these relationships still hold, but pilot data remain essential for accurate power analyses.

This work directly addresses **SO3** by extending and applying a robust framework for reliability quantification, rooted in a two-way ANOVA decomposition of variability. It provides a rigorous statistical foundation for comparing diffusion MRI biomarkers in a clinically meaningful way. These results are crucial to ensure that biomarkers developed in research settings can be confidently translated into clinical practice. Future studies should examine how reliability patterns shift in pathological populations and further refine power estimation strategies for clinical applications.

6.1.4 Limitations

Limitations in diffusion MRI are many and difficult to overcome. The primary challenge in this field remains the lack of ground truth, making validation inherently complex. High-quality acquisitions are also scarce, often limited by scanner capabilities, acquisition time, and the inherent constraints of in vivo imaging.

The limitations of this thesis can be broadly categorized into four main areas:

Data — All findings in this thesis, as in any diffusion MRI study, rely on the quality and availability of data. The lack of ground truth data makes it difficult to validate models with absolute certainty. While *ex vivo* studies and synthetic simulations provide partial solutions, they do not fully capture the complexity of *in vivo* diffusion. Additionally, clinical datasets

are often heterogeneous and subject to variability in acquisition protocols, which can influence the robustness of the proposed methods.

Noise — Noise is an intrinsic limitation of diffusion MRI, affecting all stages of data acquisition and modeling. Although this thesis proposes strategies to mitigate noise-induced bias in rotational invariant features for biophysical models, these methods cannot fully remove noise effects. In particular, their robustness has been validated under typical clinical SNR conditions, but their performance in extreme low-SNR scenarios—where scan time is severely constrained—remains to be fully characterized. As such, noise continues to be a relevant source of uncertainty for the proposed biomarkers.

Models — Every model, by definition, is a simplification of reality. Both phenomenological and biophysical approaches rely on assumptions that may not hold across all conditions. While the introduced biomarkers are designed to balance accuracy and feasibility, trade-offs remain—some diffusion properties are still challenging to capture without increased model complexity. Furthermore, while this work evaluates models under different acquisition conditions, their performance in pathological tissues or highly anisotropic environments could introduce additional uncertainties.

Group Studies — Reliability assessment and clinical translation require large, diverse datasets. However, diffusion MRI datasets are often limited in sample size, particularly when investigating specific patient populations. Small sample sizes can lead to statistical biases and overfitting of conclusions, particularly when analyzing biomarker reliability. While the methodology proposed in this thesis accounts for these factors, further validation in larger, multi-center studies is necessary to fully establish generalizability.

Despite its limitations, the work carried out in this thesis contributes to the ongoing effort to ensure that advances in neuroimaging research have a tangible impact on clinical workflows. Future work should focus on expanding dataset diversity, refining noise mitigation techniques, and optimizing models for broader applicability in both clinical and research settings.

6.2 Conclusions

This thesis tackles a fundamental challenge in diffusion MRI: bridging the gap between advanced research and practical clinical applications. We approach this by developing efficient biomarkers, improving reliability assessment methodologies, and addressing noise-induced biases in biophysical models.

Despite these contributions, economic and regulatory constraints remain a significant bottleneck in dMRI adoption. Clinical settings operate under strict time and cost restrictions, which dictate scan durations, acquisition protocols, and computational feasibility. This places a hard limit on what can realistically be implemented in hospitals and standard research facilities, regardless of technological advancements. Thus, the challenge is not just to develop better biomarkers

but to design solutions that work within these constraints.

Our findings reinforce the importance of balancing scientific ambition with clinical practicality. Future research should focus on streamlining acquisition strategies, improving the efficiency of advanced models, and ensuring that new biomarkers are not just accurate, but feasible. Ultimately, integrating dMRI into clinical practice will require a multidisciplinary effort—one that combines technical innovation, economic feasibility, and regulatory adaptability.

This thesis provides incremental but necessary steps toward that goal, contributing to the broader mission of making diffusion MRI a truly impactful tool in both neuroscience and medicine.

6.2.1 Key Findings

This thesis tackled the challenge of making diffusion MRI-based biomarkers more reliable and accessible in both research and clinical settings, despite technological, economic, and methodological constraints. The main contributions of this work can be summarized as follows:

- Development of novel phenomenological biomarkers — We introduced and implemented two new diffusion biomarkers (PA and NG) based on the state-of-the-art MiSFIT model. These metrics capture key diffusion properties while maintaining computational efficiency, making them suitable for standard research and clinical environments.
- Assessment of Rician noise effects on RISH-based features — We demonstrated how Rician bias significantly impacts rotationally invariant-based biophysical models, leading to systematic errors in parameter estimation. To mitigate these effects, we proposed strategies that improve the robustness and reliability of RISH-based biomarkers, ensuring their applicability under varying acquisition conditions.
- Systematic comparison of EAP-based models and markers via a novel and clinically-relevant framework for biomarker reliability assessment — We extended existing validation methodologies by incorporating separability and repeatability into biomarker evaluation. This approach provides a more comprehensive reliability assessment, ensuring that biomarkers are not only stable but also meaningful for clinical translation. By using this framework, we were able to provide a robust comparison of reliability among EAP-based models.

Together, these findings offer a meaningful step toward transferring advanced dMRI developments from the research environment to practical, real-world scenarios, helping diffusion MRI move toward broader accessibility in neuroscience, medicine, and beyond.

6.2.2 Future Directions

While this thesis addresses key barriers in dMRI adoption, open questions remain. Future research should focus on:

Refining and expanding biomarker validation — The proposed reliability framework provides an important step forward, but further validation in larger, multi-center studies is needed to ensure generalizability across different scanners, populations, and pathological conditions. Specifically for within-subject errors and variability, it would be insightful to evaluate the behavior of biomarkers at different longitudinal scales, comparing the across a series of sessions, each consisting of several repeated acquisitions. Also, a generalization into a three-way ANOVA would enable the simultaneous characterization of inter-subject, inter-session, and intra-session variability, along with their potential interactions, thus offering a more complete framework for assessing the robustness of the proposed diffusion MRI biomarkers. Finally, further studies should include biophysical models (SMI, NODDI, etc.) to verify their suitability for answering scientific questions in either clinical or research settings.

Optimizing acquisition strategies — Standard clinical settings operate under strict time and economic constraints, limiting the feasibility of advanced dMRI protocols. Future work should explore adaptive acquisition techniques that balance scan efficiency with biomarker accuracy.

Integrating noise-informed modeling — The findings on Rician bias open the door for further exploration of statistically-informed biophysical modeling, where noise characteristics are directly incorporated into the estimation process to improve reliability. Additionally, applying such informed statistics would be benefitial in phenomenological models like the ones visited throughout this thesis, as they often use least squares fitting algorithms with no regard for the underlying noise characteristics.

Ultimately, diffusion MRI still has much to offer, and its future is undoubtedly promising. The challenge is not just to develop more advanced methods, but to ensure that these advancements can be adopted in real-world settings, where their impact will be felt the most.

6.2.3 Final Thoughts

Diffusion MRI is a constantly evolving field, where methodological advancements push the boundaries of what we can extract from the brain's microstructural landscape. However, technological advancements alone are not enough—for these innovations to reach everyday research labs and hospitals, they must align with the economic, practical, and regulatory constraints of standard environments.

The work presented in this thesis stands as another step toward this goal. By developing biomarkers that are both reliable and efficient, addressing the impact of noise on biophysical models, and proposing a clinically relevant framework for reliability assessment, we contribute to making dMRI more robust, accessible, and ultimately, more useful.

Still, the road ahead is long. Just as a strong river carves its way through the landscape, diffusion MRI researchers face two fundamental strategies: some strive to control the flow, reducing its randomness through ever-more sophisticated models and acquisition techniques. Others aim to understand and harness its nature, adapting methodologies to work with the inherent complexities

of diffusion. For dMRI to be widely adopted, both approaches must coexist—technological innovation and practical adaptation must go hand in hand.

This work is just a small piece of that effort, contributing to a larger collective mission: ensuring that diffusion MRI reaches its full potential, whether in technological development, neuroscientific discovery, or clinical application.

Bibliography

Ades-Aron, B., Coelho, S., Lemberskiy, G., Veraart, J., Baete, S. H., Shepherd, T. M., Novikov, D. S., and Fieremans, E. (2025). Denoising Improves Cross-Scanner and Cross-Protocol Test–Retest Reproducibility of Diffusion Tensor and Kurtosis Imaging. *Human Brain Mapping*, 46(4).

Aja-Fernández, S., de Luis-García, R., Afzali, M., Molendowska, M., Pieciak, T., and Tristán-Vega, A. (2020). Micro-structure diffusion scalar measures from reduced MRI acquisitions. *PLoS One*, 15(3):e0229526.

Aja-Fernández, S., Pieciak, T., Martín-Martín, C., Planchuelo-Gomez, A., de Luis-García, R., and Tristán-Vega, A. (2022). Moment-based representation of the diffusion inside the brain from reduced DMRI acquisitions: Generalized AMURA. *Medical Image Analysis*, 77:102356.

Aja-Fernández, S. and Vegas-Sánchez-Ferrero, G. (2016). *Statistical Analysis of Noise in MRI*. Springer.

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3):91–93.

Albi, A., Pasternak, O., Minati, L., Marizzoni, M., Bartrés-Faz, D., Bargallo, N., Bosch, B., Rossini, P. M., Marra, C., Müller, B., et al. (2017). Free water elimination improves test–retest reproducibility of diffusion tensor imaging indices in the brain: A longitudinal multisite study of healthy elderly subjects. Technical report, Wiley Online Library.

Alexander, A. L., Wu, Y.-C., and Venkat, P. C. (2006). Hybrid diffusion imaging (HYDI). In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2245–2248. IEEE.

Alexander, D., Barker, G., and Arridge, S. (2002). Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340.

Almeida Martins, J. P. d., Nilsson, M., Lampinen, B., Palombo, M., While, P. T., Westin, C.-F., and Szczepankiewicz, F. (2021). Neural networks for parameter estimation in microstructural MRI: Application to a diffusion-relaxation model of white matter. *NeuroImage*, 244:118601.

Andersson, J. L. R. et al. (2016). Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *NeuroImage*, 141:556–572.

Andersson, J. L. R., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, 20(2):870–888.

Assaf, Y. and Basser, P. J. (2005). Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. *Neuroimage*, 27(1):48–58.

Assaf, Y., Freidlin, R. Z., Rohde, G. K., and Basser, P. J. (2004). New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5):965–978.

Assaf, Y., Mayk, A., and Cohen, Y. (2000). Displacement imaging of spinal cord using q-space diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 44(5):713–722.

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

Balchandani, P. and Naidich, T. P. (2015). Ultra-High-Field MR Neuroimaging. *AJNR. American journal of neuroradiology*, 36(7):1204–1215.

Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., and Reiss, A. L. (2004). White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biological Psychiatry*, 55(3):323–326.

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267.

Bendlin, B. B., Ries, M. L., Canu, E., Sodhi, A., Lazar, M., Alexander, A. L., Carlsson, C. M., Sager, M. A., Asthana, S., and Johnson, S. C. (2010). White matter is altered with parental family history of Alzheimer's disease. *Alzheimer's & Dementia*, 6(5):394–403.

Bergamino, M., Walsh, R. R., and Stokes, A. M. (2021). Free-water diffusion tensor imaging improves the accuracy and sensitivity of white matter analysis in Alzheimer's disease. *Scientific Reports*, 11(1).

Bernstein-Eliav, M. and Tavor, I. (2022). The Prediction of Brain Activity from Connectivity: Advances and Applications. *The Neuroscientist*, 30(3):367–377.

Blumenfeld-Katzir, T., Pasternak, O., Dagan, M., and Assaf, Y. (2011). Diffusion MRI of Structural Brain Plasticity Induced by a Learning and Memory Task. *PLoS ONE*, 6(6):e20678.

Bosch, B., Arenaza-Urquijo, E. M., Rami, L., Sala-Llonch, R., Junqué, C., Solé-Padullés, C., Peña-Gómez, C., Bargalló, N., Molinuevo, J. L., and Bartrés-Faz, D. (2012). Multiple DTI index analysis in normal aging, amnestic MCI and AD. Relationship with neuropsychological performance. *Neurobiology of Aging*, 33(1):61–74.

Brainlife.io (2025). Brainlife.io: A platform for computational neuroimaging research, benchmarking, and data sharing. Accessed: 2025-04-04.

Budde, M. D. and Frank, J. A. (2010). Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke. *Proceedings of the National Academy of Sciences*, 107(32):14472–14477.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., and Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5):365–376.

Calamante, F., Tournier, J.-D., Jackson, G. D., and Connelly, A. (2010). Track-density imaging (TDI): Super-resolution white matter imaging using whole-brain track-density mapping. *NeuroImage*, 53(4):1233–1243.

Callaghan, P. T. (1993). *Principles of nuclear magnetic resonance microscopy*. Clarendon Press, Oxford, England.

Campaña Perilla, L. A., Mahecha Carvajal, M. E., Cardona Ortegón, J. D., Barragán Corrales, C., and Barrera Patiño, A. M. (2023). Diffusion Tensor Imaging Protocol: The Need for Standardization of Measures. *Radiology*, 308(2).

Caporale, A., Bonomo, G. B., Tani Raffaelli, G., Tata, A. M., Avallone, B., Wehrli, F. W., and Capuani, S. (2022). Transient Anomalous Diffusion MRI in Excised Mouse Spinal Cord: Comparison Among Different Diffusion Metrics and Validation With Histology. *Frontiers in Neuroscience*, 15.

Carreira Figueiredo, I., Borgan, F., Pasternak, O., Turkheimer, F. E., and Howes, O. D. (2022). White-matter free-water diffusion MRI in schizophrenia: a systematic review and meta-analysis. *Neuropsychopharmacology*, 47(7):1413–1420.

Caruyer, E. and Deriche, R. (2012). Diffusion MRI signal reconstruction with continuity constraint and optimal regularization. *Medical Image Analysis*, 16(6):1113–1120.

Caruyer, E., Lenglet, C., Sapiro, G., and Deriche, R. (2013). Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine*, 69(6):1534–1540.

Chad, J. A., Pasternak, O., Salat, D. H., and Chen, J. J. (2018). Re-examining age-related differences in white matter microstructure with free-water corrected diffusion tensor imaging. *Neurobiology of Aging*, 71:161–170.

Chamberland, M., Genc, S., Tax, C. M. W., Shastin, D., Koller, K., Raven, E. P., Cunningham, A., Doherty, J., van den Bree, M. B. M., Parker, G. D., Hamandi, K., Gray, W. P., and Jones, D. K. (2021). Detecting microstructural deviations in individuals with deep diffusion MRI tractometry. *Nature Computational Science*, 1(9):598–606.

Cheng, J., Deriche, R., Jiang, T., Shen, D., and Yap, P.-T. (2014). Non-Negative Spherical Deconvolution (NNSD) for estimation of fiber Orientation Distribution Function in single-/multi-shell diffusion MRI. *NeuroImage*, 101:750–764.

Chung, H.-W., Chou, M.-C., and Chen, C.-Y. (2010). Principles and Limitations of Computational Algorithms in Clinical Diffusion Tensor MR Tractography. *American Journal of Neuroradiology*, 32(1):3–13.

Clark, K. A., Nuechterlein, K. H., Asarnow, R. F., Hamilton, L. S., Phillips, O. R., Hageman, N. S., Woods, R. P., Alger, J. R., Toga, A. W., and Narr, K. L. (2011). Mean diffusivity and fractional anisotropy as indicators of disease and genetic liability to schizophrenia. *Journal of Psychiatric Research*, 45(7):980–988.

Coelho, S., Baete, S. H., Lemberskiy, G., Ades-Aron, B., Barrol, G., Veraart, J., Novikov, D. S., and Fieremans, E. (2022). Reproducibility of the Standard Model of diffusion in white matter on clinical MRI systems. *NeuroImage*, 257:119290.

Coelho, S., Liao, Y., Szczepankiewicz, F., Veraart, J., Chung, S., Lui, Y. W., Novikov, D. S., and Fieremans, E. (2024). Assessment of precision and accuracy of brain white matter microstructure using combined diffusion MRI and relaxometry. *Human Brain Mapping*, 45(9):e26725.

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A. N., and Hajnal, J. V. (2019). Complex diffusion-weighted image estimation via matrix recovery under general noise models. *Neuroimage*, 200:391–404.

Daianu, M., Jacobs, R. E., Weitz, T. M., Town, T. C., and Thompson, P. M. (2015). Multi-Shell Hybrid Diffusion Imaging (HYDI) at 7 Tesla in TgF344-AD Transgenic Alzheimer Rats. *PLOS ONE*, 10(12):e0145205.

Dao Trong, P., Olivares, A., El Damaty, A., and Unterberg, A. (2023). Adverse events in neurosurgery: a comprehensive single-center analysis of a prospectively compiled database. *Acta Neurochirurgica*, 165(3):585–593.

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., and Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *NeuroImage*, 46(2):530–541.

Dela Haije, T., Özarslan, E., and Feragen, A. (2020). Enforcing necessary non-negativity constraints for common diffusion MRI models using sum of squares programming. *NeuroImage*, 209:116405.

Deleuze, G. (1994). *Difference and Repetition*. Columbia University Press, New York.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510.

Dewan, M. C., Rattani, A., Gupta, S., Baticulon, R. E., Hung, Y.-C., Punchak, M., Agrawal, A., Adeleye, A. O., Shrime, M. G., Rubiano, A. M., Rosenfeld, J. V., and Park, K. B. (2019). Estimating the global incidence of traumatic brain injury. *Journal of Neurosurgery*, 130(4):1080–1097.

Diamandi, J., Raimondo, C., Alizadeh, M., Flanders, A., Tjoumakaris, S., Gooch, M. R., Jabbour, P., Rosenwasser, R., and Mouchtouris, N. (2025a). Use of mean apparent propagator (MAP) MRI in patients with acute ischemic stroke: A comparative study with DTI and NODDI. *Magnetic Resonance Imaging*, 117:110290.

Diamandi, J., Raimondo, C., Piper, K., Roy, J., Serva, S., Alizadeh, M., Flanders, A., Tjoumakaris, S., Gooch, R., Jabbour, P., Rosenwasser, R., and Mouchtouris, N. (2025b). Use of multi-modal non-contrast MRI to predict functional outcomes after stroke: A study using DP-pCASL, DTI, NODDI, and MAP MRI. *NeuroImage: Clinical*, 45:103742.

Donoho, D. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306.

Dresler, M., Sandberg, A., Bublitz, C., Ohla, K., Trenado, C., Mroczko-Wąsowicz, A., Kühn, S., and Repantis, D. (2018). Hacking the Brain: Dimensions of Cognitive Enhancement. *ACS Chemical Neuroscience*, 10(3):1137–1148.

Duan, F., Zhao, T., He, Y., and Shu, N. (2015). Test–retest reliability of diffusion measures in cerebral white matter: A multiband diffusion MRI study. *Journal of Magnetic Resonance Imaging*, 42(4):1106–1116.

Eichner, C., Cauley, S. F., Cohen-Adad, J., Möller, H. E., Turner, R., Setsompop, K., and Wald, L. L. (2015). Real diffusion-weighted MRI enabling true signal averaging and increased diffusion contrast. *NeuroImage*, 122:373–384.

Esteban, O. (2024). *Standardized Preprocessing in Neuroimaging: Enhancing Reliability and Reproducibility*, page 153–179. Springer US.

Farag, S. and Kayser, O. (2017). *The Cannabis Plant: Botanical Aspects*. Elsevier.

Feng, G., Chen, R., Zhao, R., Li, Y., Ma, L., Wang, Y., Men, W., Gao, J., Tan, S., Cheng, J., He, Y., Qin, S., Dong, Q., Tao, S., and Shu, N. (2023). Longitudinal development of the human white matter structural connectome and its association with brain transcriptomic and cellular architecture. *Communications Biology*, 6(1).

Fick, R. H., Wassermann, D., Caruyer, E., and Deriche, R. (2016a). MAPL: Tissue microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP data. *NeuroImage*, 134:365–385.

Fick, R. H. J., Daianu, M., Pizzolato, M., Wassermann, D., Jacobs, R. E., Thompson, P. M., Town, T., and Deriche, R. (2017). Comparison of biomarkers in transgenic alzheimer rats using multi-shell diffusion MRI. *Mathematics and Visualization*, pages 187–199.

Fick, R. H. J., Pizzolato, M., Wassermann, D., Zucchelli, M., Menegaz, G., and Deriche, R. (2016b). A sensitivity analysis of q-space indices with respect to changes in axonal diameter, dispersion and tissue composition. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, page 1241–1244. IEEE.

Fieremans, E. et al. (2013). Novel white matter tract integrity metrics sensitive to Alzheimer disease progression. *AJNR. American journal of neuroradiology*, 34(11):2105–2112.

Fieremans, E., Jensen, J. H., Helpern, J. A., Kim, S., Grossman, R. I., Inglese, M., and Novikov, D. S. (2012). Diffusion Distinguishes Between Axonal Loss and Demyelination in Brain White Matter. In *ISMRM Annual Meeting*, Melbourne, Australia. ISMRM.

Foucault, M. (1972). *The Archaeology of Knowledge*. Pantheon Books, New York.

Frank, L. R. (2002). Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47(6):1083–1099.

Fugate, J. E. (2015). Complications of Neurosurgery. *CONTINUUM: Lifelong Learning in Neurology*, 21:1425–1444.

Galazzo, I. B., Brusini, L., Obertino, S., Zucchelli, M., Granziera, C., and Menegaz, G. (2018). On the viability of diffusion MRI-based microstructural biomarkers in ischemic stroke. *Frontiers in Neuroscience*, 12(FEB).

Gao, Z., Zhang, D., Yan, X., Shi, H., and Xian, X. (2022). Effects of ω-3 polyunsaturated fatty acids on coronary atherosclerosis and inflammation: A systematic review and meta-analysis. *Frontiers in Cardiovascular Medicine*, 9.

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S., Descoteaux, M., and Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, 8.

Grech-Sollars, M., Hales, P. W., Miyazaki, K., Raschke, F., Rodriguez, D., Wilson, M., Gill, S. K., Banks, T., Saunders, D. E., Clayden, J. D., et al. (2015). Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR in Biomedicine*, 28(4):468–485.

Grech-Sollars, M., Hales, P. W., Miyazaki, K., Raschke, F., Rodriguez, D., Wilson, M., Gill, S. K., Banks, T., Saunders, D. E., Clayden, J. D., Gwilliam, M. N., Barrick, T. R., Morgan, P. S., Davies, N. P., Rossiter, J., Auer, D. P., Grundy, R., Leach, M. O., Howe, F. A., Peet, A. C., and Clark, C. A. (2015). Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *NMR in Biomedicine*, 28(4):468–485.

Gudbjartsson, H. and Patz, S. (1995). The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34(6):910–914.

Han, B.-C. (2017). *Psychopolitics: Neoliberalism and the New Power of Control*. Verso, London.

Harville, D. A. (1997). *Matrix algebra from a statistician's perspective*. Springer.

Henriques, R. N., Ianuş, A., Novello, L., Jovicich, J., Jespersen, S. N., and Shemesh, N. (2023). Efficient PCA denoising of spatially correlated redundant MRI data. *Imaging Neuroscience*, 1:1–26.

Henriques, R. N., Jespersen, S. N., Jones, D. K., and Veraart, J. (2021). Toward more robust and reproducible diffusion kurtosis imaging. *Magnetic Resonance in Medicine*, 86(3):1600–1613.

Henriques, R. N., Tax, C. M. W., Shemesh, N., and Veraart, J. (2018). Biophysical modeling of the white matter: from theory towards clinical practice. In *Proceedings of the 27th ISMRM annual meeting, Montreal, CA, Abstract #1008*.

Herberthson, M., Boito, D., Haije, T. D., Feragen, A., Westin, C.-F., and Özarslan, E. (2021). Q-space trajectory imaging with positivity constraints (QTI+). *NeuroImage*, 238:118198.

Hobson, E. W. (2012). *The theory of spherical and ellipsoidal harmonics*. Cambridge University Press.

Hori, M., Maekawa, T., Kamiya, K., Hagiwara, A., Goto, M., Takemura, M. Y., Fujita, S., Andica, C., Kamagata, K., Cohen-Adad, J., and Aoki, S. (2022). Advanced Diffusion MR Imaging for Multiple Sclerosis in the Brain and Spinal Cord. *Magnetic Resonance in Medical Sciences*, 21(1):58–70.

Jahanshad, N., Kochunov, P. V., Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., Blangero, J., Brouwer, R. M., Curran, J. E., de Zubicaray, G. I., Duggirala, R., Fox, P. T., Hong, L. E., Landman, B. A., Martin, N. G., McMahon, K. L., Medland, S. E., Mitchell, B. D., Olvera, R. L., Peterson, C. P., Starr, J. M., Sussmann, J. E., Toga, A. W., Wardlaw, J. M., Wright, M. J., Hulshoff Pol, H. E., Bastin, M. E., McIntosh, A. M., Deary, I. J., Thompson, P. M., and Glahn, D. C. (2013). Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. *NeuroImage*, 81:455–469.

Jangwan, N. S., Ashraf, G., Ram, V., Singh, V., Alghamdi, B. S., Abuzenadah, A. M., and Singh, M. F. (2022). Brain augmentation and neuroscience technologies: current applications, challenges, ethics and future prospects. *Frontiers in Systems Neuroscience*, 16.

Jansen, J. F., Stambuk, H. E., Koutcher, J. A., and Shukla-Dave, A. (2010). Non-Gaussian analysis of diffusion-Weighted MR imaging in head and neck squamous cell carcinoma: A feasibility study. *American Journal of Neuroradiology*, 31(4):741–748.

Jelescu, I. O., de Skowronski, A., Geffroy, F., Palombo, M., and Novikov, D. S. (2022). Neurite Exchange Imaging (NEXI): A minimal model of diffusion in gray matter with inter-compartment water exchange. *NeuroImage*, 256:119277.

Jelescu, I. O. et al. (2016a). Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR in biomedicine*, 29(1):33–47.

Jelescu, I. O., Zurek, M., Winters, K. V., Veraart, J., Rajaratnam, A., Kim, N. S., Babb, J. S., Shepherd, T. M., Novikov, D. S., Kim, S. G., and Fieremans, E. (2016b). In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. *NeuroImage*, 132:104–114.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841.

Jenkinson, M. and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2):143–156.

Jensen, J. H., Helpern, J. A., 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.

Jespersen, S. N., Kroenke, C. D., Østergaard, L., Ackerman, J. J. H., and Yablonskiy, D. A. (2007). Modeling dendrite density from magnetic resonance diffusion measurements. *NeuroImage*, 34(4):1473–1486.

Jeurissen, B., Descoteaux, M., Mori, S., and Leemans, A. (2017). Diffusion MRI fiber tractography of the brain. *NMR in Biomedicine*, 32(4).

Jones, D., Horsfield, M., and Simmons, A. (1999). Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(3):515–525.

Jones, D. K. and Basser, P. J. (2004). Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine*, 52(5):979–993.

Jones, D. K. and Leemans, A. (2010). *Diffusion Tensor Imaging*, page 127–144. Humana Press.

Joseph, M., Mansour, S., Pisner, D., Richie-Halford, A., Lerma-Usabiaga, G., Keshavan, A., Dickie, E. W., Rokem, A., and Esteban, O. (2021). dMRIPrep: a robust preprocessing pipeline for diffusion MRI.

Kaden, E., Kruggel, F., and Alexander, D. C. (2016). Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic Resonance in Medicine*, 75(4):1752–1763.

Kanyongo, G. Y., Brook, G. P., Kyei-Blankson, L., and Gocmen, G. (2007). Reliability and Statistical Power: How Measurement Fallibility Affects Power and Required Sample Sizes for Several Parametric and Nonparametric Statistics. *Journal of Modern Applied Statistical Methods*, 6(1):81–90.

Karayumak, S. C., Bouix, S., Ning, L., James, A., Crow, T., Shenton, M., Kubicki, M., and Rathi, Y. (2019). Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *NeuroImage*, 184:180–200.

Kellner, E., Dhital, B., Kiselev, V. G., and Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic resonance in medicine*, 76(5):1574–1581.

Kellner, E. et al. (2015). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic Resonance in Medicine*.

Kerkelä, L., Henriques, R. N., Hall, M. G., Clark, C. A., and Shemesh, N. (2020). Validation and noise robustness assessment of microscopic anisotropy estimation with clinically feasible double diffusion encoding MRI. *Magnetic Resonance in Medicine*, 83(5):1698–1710.

Koay, C. and Basser, P. (2008). Simple harmonic oscillator based estimation and reconstruction for one-dimensional q-space MR. In *Proc. Intl. Soc. Mag. Reson. Med*, volume 16, page 35.

Koay, C. G. and Basser, P. J. (2006). Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *Journal of magnetic resonance*, 179(2):317–322.

Koen, B. V. (2003). *Discussion of the method*. Oxford University Press, New York, NY.

Kraus, M. F., Susmaras, T., Caughlin, B. P., Walker, C. J., Sweeney, J. A., and Little, D. M. (2007). White matter integrity and cognition in chronic traumatic brain injury: A diffusion tensor imaging study. *Brain*, 130(10):2508–2519.

Kubicki, M., Baxi, M., Pasternak, O., Tang, Y., Karmacharya, S., Chunga, N., Lyall, A. E., Rathi, Y., Eckbo, R., Bouix, S., Mortazavi, F., Papadimitriou, G., Shenton, M. E., Westin, C. F., Killiany, R., Makris, N., and Rosene, D. L. (2018). Lifespan Trajectories of White Matter Changes in Rhesus Monkeys. *Cerebral Cortex*, 29(4):1584–1593.

Laguna, P. A. L., Combes, A. J., Streffer, J., Einstein, S., Timmers, M., Williams, S. C., and Dell'Acqua, F. (2020). Reproducibility, reliability and variability of FA and MD in the older healthy population: A test-retest multiparametric analysis. *NeuroImage: Clinical*, 26:102168.

Lampinen, B., Szczepankiewicz, F., Novén, M., van Westen, D., Hansson, O., Englund, E., Mårtensson, J., Westin, C., and Nilsson, M. (2019). Searching for the neurite density with diffusion MRI: Challenges for biophysical modeling. *Human Brain Mapping*, 40(8):2529–2545.

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M.-Y., Gilmore, A. W., McDermott, K. B., Nelson, S. M., Dosenbach, N. U., Schlaggar, B. L., Mumford, J. A., Poldrack, R. A., and Petersen, S. E. (2015). Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron*, 87(3):657–670.

Lawrence, I. and Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, pages 255–268.

Leysen, S., Sunaert, S., Maes, F., and Christiaens, D. (2023). Voxel-level denoising of diffusion MRI using a rank-1 decomposition. *Proceedings International Society Magnetic Resonance Medicine*, 5016.

Lyall, A. E., Pasternak, O., Robinson, D. G., Newell, D., Trampush, J. W., Gallego, J. A., Fava, M., Malhotra, A. K., Karlsgodt, K. H., Kubicki, M., and Szeszko, P. R. (2017). Greater extracellular free-water in first-episode psychosis predicts better neurocognitive functioning. *Molecular Psychiatry*, 23(3):701–707.

Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N.-k., and Song, A. W. (2012). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1822(3):386–400.

Madden, D. J., Bennett, I. J., and Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: Contributions from diffusion tensor imaging. *Neuropsychology Review*, 19(4):415–435.

Maillard, P., Fletcher, E., Singh, B., Martinez, O., Johnson, D. K., Olichney, J. M., Farias, S. T., and DeCarli, C. (2019). Cerebral white matter free water: A sensitive biomarker of cognition and function. *Neurology*, 92(19).

Martín-Martín, C., Planchuelo-Gómez, Á., Guerrero, Á. L., García-Azorín, D., Tristán-Vega, A., de Luis-García, R., and Aja-Fernández, S. (2023). Viability of AMURA biomarkers from single-shell diffusion MRI in clinical studies. *Frontiers in Neuroscience*, 17:1106350.

McClymont, D., Teh, I., Carruth, E., Omens, J., McCulloch, A., Whittington, H. J., Kohl, P., Grau, V., and Schneider, J. E. (2016). Evaluation of non-Gaussian diffusion in cardiac MRI. *Magnetic Resonance in Medicine*, 78(3):1174–1186.

McDonald, B. A., Salzillo, T., Mulder, S., Ahmed, S., Dresner, A., Preston, K., He, R., Christodouleas, J., Mohamed, A. S., Philippens, M., van Houdt, P., Thorwarth, D., Wang, J., Shukla Dave, A., Boss, M., and Fuller, C. D. (2023). Prospective evaluation of in vivo and phantom repeatability and reproducibility of diffusion-weighted MRI sequences on 1.5 T MRI-linear accelerator (MR-Linac) and MR simulator devices for head and neck cancers. *Radiotherapy and Oncology*, 185:109717.

McGraw, K. O. and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1):30–46.

McQuarrie, D. and Simon, J. (1997). *Physical Chemistry: A Molecular Approach*. G - Reference,Information and Interdisciplinary Subjects Series. University Science Books.

Merisaari, H., Tuulari, J. J., Karlsson, L., Scheinin, N. M., Parkkola, R., Saunavaara, J., Lähdesmäki, T., Lehtola, S. J., Keskinen, M., Lewis, J. D., et al. (2019). Test-retest reliability of diffusion tensor imaging metrics in neonates. *NeuroImage*, 197:598–607.

Metzler-Baddeley, C., Jones, D. K., Steventon, J., Westacott, L., Aggleton, J. P., and O'Sullivan, M. J. (2012). Cingulum Microstructure Predicts Cognitive Control in Older Age and Mild Cognitive Impairment. *The Journal of Neuroscience*, 32(49):17612–17619.

Mirzaalian, H., de Pierrefeu, A., Savadjiev, P., Pasternak, O., Bouix, S., Kubicki, M., Westin, C.-F., Shenton, M. E., and Rathi, Y. (2015). Harmonizing diffusion MRI data across multiple sites and scanners. In *Lecture Notes in Computer Science*, pages 12–19. Springer International Publishing.

Mirzaalian, H., Ning, L., Savadjiev, P., Pasternak, O., Bouix, S., Michailovich, O., Grant, G., Marx, C. E., Morey, R. A., Flashman, L. A., George, M. S., McAllister, T. W., Andaluz, N., Shutter, L., Coimbra, R., Zafonte, R. D., Coleman, M. J., Kubicki, M., Westin, C. F., others, and Rathi, Y. (2016). Inter-site and inter-scanner diffusion MRI data harmonization. *NeuroImage*, 135:311–323.

Mitsea, E., Drigas, A., and Skianis, C. (2023). Digitally Assisted Mindfulness in Training Self-Regulation Skills for Sustainable Mental Health: A Systematic Review. *Behavioral Sciences*, 13(12):1008.

Moeller, S., Pisharady, P. K., Ramanna, S., Lenglet, C., Wu, X., Dowdle, L., others, and Akçakaya, M. (2021). NOise reduction with DIstribution Corrected (NORDIC) PCA in dMRI with complex-valued parameter-free locally low-rank processing. *NeuroImage*, 226:117539.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., Woods, R., et al. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an icbm template. *Neuroimage*, 40(2):570–582.

Muthupallai, R., Holder, C., Song, A., and Dixon, W. (1999). Navigator aided, multishot EPI diffusion images of brain with complete orientation and anisotropy information. In *Proceedings of the 7th Annual Meeting of ISMRM*, page 1825, Philadelphia, PA.

Nakaya, M., Sato, N., Matsuda, H., Maikusa, N., Shigemoto, Y., Sone, D., Yamao, T., Ogawa, M., Kimura, Y., Chiba, E., Ohnishi, M., Kato, K., Okita, K., Tsukamoto, T., Yokoi, Y., Sakata, M., and Abe, O. (2022). Free water derived by multi-shell diffusion MRI reflects tau/neuroinflammatory pathology in Alzheimer's disease. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 8(1).

Nedjati-Gilani, G. L., Schneider, T., Hall, M. G., Cawley, N., Hill, I., Ciccarelli, O., Drobnjak, I., Wheeler-Kingshott, C. A. M. G., and Alexander, D. C. (2017). Machine learning based compartment models with permeability for white matter microstructure imaging. *NeuroImage*, 150:119–135.

Nestor, P. G., Ohtani, T., Bouix, S., Hosokawa, T., Saito, Y., Newell, D. T., and Kubicki, M. (2014). Dissociating prefrontal circuitry in intelligence and memory: neuropsychological correlates of magnetic resonance and diffusion tensor imaging. *Brain Imaging and Behavior*, 9(4):839–847.

Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., Kriegeskorte, N., Milham, M. P., Poldrack, R. A., Poline, J.-B., Proal, E., Thirion, B., Van Essen, D. C., White, T., and Yeo, B. T. T. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience*, 20(3):299–303.

Nietzsche, F. (1997). *Human, all too human I / A book for free spirits.* The Complete Works of Friedrich Nietzsche. Stanford University Press, Palo Alto, CA.

Ning, L., Westin, C.-F., and Rathi, Y. (2015). Estimating diffusion propagator and its moments using directional radial basis functions. *IEEE transactions on medical imaging*, 34(10):2058–2078.

Novikov, D. S., Kiselev, V. G., and Jespersen, S. N. (2018a). On modeling. *Magnetic Resonance in Medicine*, 79(6):3172–3193.

Novikov, D. S., Veraart, J., Jelescu, I. O., and Fieremans, E. (2018b). Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *NeuroImage*, 174:518–538.

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A. W., Pike, G. B., Rosa-Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P. C. M., Mazziotta, J., and Mori, S. (2008). Human brain white matter atlas: Identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*, 43(3):447–457.

Oppenheim, A., Willsky, A., and Nawab, S. (1997). *Signals & Systems.* Prentice-Hall signal processing series. Prentice Hall.

Örzsik, B., Palombo, M., Asllani, I., Dijk, D.-J., Harrison, N. A., and Cercignani, M. (2023). Higher order diffusion imaging as a putative index of human sleep-related microstructural changes and glymphatic clearance. *NeuroImage*, 274:120124.

Osa García, A., Brambati, S. M., Desautels, A., and Marcotte, K. (2022). Timing stroke: A review on stroke pathophysiology and its influence over time on diffusion measures. *Journal of the Neurological Sciences*, 441:120377.

Ouyang, Y., Cui, D., Yuan, Z., Liu, Z., Jiao, Q., Yin, T., and Qiu, J. (2021). Analysis of Age-Related White Matter Microstructures Based on Diffusion Tensor Imaging. *Frontiers in Aging Neuroscience*, 13.

Özarslan, E., Koay, C. G., Shepherd, T. M., Blackband, S. J., and Basser, P. J. (2009). Simple Harmonic Oscillator Based Reconstruction and Estimation for Three-Dimensional Q-Space MRI. In *Proceedings of the 17th ISMRM annual meeting, Honolulu, HI, Abstract #1396*.

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

Özarslan, E., Shepherd, T., Vemuri, B., Blackband, S., and Mareci, T. (2005). The Diffusion Orientation Transform (DOT) for MR Imaging. In *APS Southeastern Section Meeting Abstracts*, volume 72, pages CE–022.

Özarslan, E., Shepherd, T. M., Vemuri, B. C., Blackband, S. J., and Mareci, T. H. (2006). Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31(3):1086–1103.

O'Connor, D., Potler, N. V., Kovacs, M., Xu, T., Ai, L., Pellman, J., Vanderwal, T., Parra, L. C., Cohen, S., Ghosh, S., Escalera, J., Grant-Villegas, N., Osman, Y., Bui, A., Craddock, R. C., and Milham, M. P. (2017). The Healthy Brain Network Serial Scanning Initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. *GigaScience*, 6(2).

Palmqvist, S., Janelidze, S., Quiroz, Y. T., Zetterberg, H., Lopera, F., Stomrud, E., Su, Y., Chen, Y., Serrano, G. E., Leuzy, A., Mattsson-Carlgren, N., Strandberg, O., Smith, R., Villegas, A., Sepulveda-Falla, D., Chai, X., Proctor, N. K., Beach, T. G., Blennow, K., Dage, J. L., Reiman, E. M., and Hansson, O. (2020). Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA*, 324(8):772.

Palombo, M., Ianus, A., Guerreri, M., Nunes, D., Alexander, D. C., Shemesh, N., and Zhang, H. (2020). SANDI: A compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI. *NeuroImage*, 215:116835.

París, G., Pieciak, T., Aja-Fernández, S., and Tristán-Vega, A. (2023a). Efficient estimation of propagator anisotropy and non-Gaussianity in multishell diffusion MRI with micro-structure adaptive convolution kernels and dual Fourier integral transforms. *Magnetic Resonance in Medicine*, 89(1):440–453.

París, G., Pieciak, T., Aja-Fernández, S., and Tristán-Vega, A. (2023b). Efficient estimation of the Generalised Cross-Validation regularisation weight: An application for HYDI-DSI-QP. In *ISMRM Iberian Chapter*, Valladolid, Spain. ISMRM.

Parkinson's Foundation (2025). Parkinson's disease statistics. Accessed: 3 April 2025.

Pasternak, O., Sochen, N., Gur, Y., Intrator, N., and Assaf, Y. (2009). Free water elimination and mapping from diffusion MRI. *Magnetic Resonance in Medicine*, 62(3):717–730.

Paudyal, R., Konar, A. S., Obuchowski, N. A., Hatzoglou, V., Chenevert, T. L., Malyarenko, D. I., Swanson, S. D., LoCastro, E., Jambawalikar, S., Liu, M. Z., Schwartz, L. H., Tuttle, R. M., Lee, N., and Shukla-Dave, A. (2019). Repeatability of Quantitative Diffusion-Weighted Imaging Metrics in Phantoms, Head-and-Neck and Thyroid Cancers: Preliminary Findings. *Tomography*, 5(1):15–25.

Pieciak, T., Aja-Fernandez, S., and Vegas-Sanchez-Ferrero, G. (2017). Non-Stationary rician noise estimation in parallel MRI using a single image: A Variance-Stabilizing approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(10):2015–2029.

Pieciak, T., París, G., Aja-Fernández, S., and de Luis-García, R. (2025). Repeatability and reliability of Diffusion Tensor Imaging in white matter rectified with free-water. In *ISMRM & ISMRT Annual Meeting & Exhibition*, Honolulu, Hawaii. ISMRM.

Pieciak, T., París, G., Tristán-Vega, A., and Aja-Fernández, S. (2024). Unbiasing the spherical variance of a diffusion-weighted MR signal. In *Proceedings of the ISMRM & ISMRT Annual Meeting & Exhibition, Singapore*, Singapore.

Pierpaoli, C. and Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine*, 36(6):893–906.

Pines, A. R., Cieslak, M., Larsen, B., Baum, G. L., Cook, P. A., Adebimpe, A., Dávila, D. G., Elliott, M. A., Jirsaraie, R., Murtha, K., Oathes, D. J., Piiwaa, K., Rosen, A. F., Rush, S., Shinohara, R. T., Bassett, D. S., Roalf, D. R., and Satterthwaite, T. D. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Developmental Cognitive Neuroscience*, 43:100788.

Pizzolato, M. et al. (2022). Axonal T2 estimation using the spherical variance of the strongly diffusion-weighted MRI signal. *Magnetic resonance imaging*, 86:118–134.

Planchuelo-Gómez, Á., García-Azorín, D., Guerrero, Á. L., de Luis-García, R., Rodríguez, M., and Aja-Fernández, S. (2020). Alternative microstructural measures to complement diffusion tensor imaging in migraine studies with standard MRI acquisition. *Brain Sciences*, 10(10):711.

Raven, E. P., Veraart, J., Kievit, R. A., Genc, S., Ward, I. L., Hall, J., Cunningham, A., Doherty, J., van den Bree, M. B. M., and Jones, D. K. (2023). In vivo evidence of microstructural hypo-connectivity of brain white matter in 22q11.2 deletion syndrome. *Molecular Psychiatry*, 28(10):4342–4352.

Reisert, M., Kellner, E., Dhital, B., Hennig, J., and Kiselev, V. G. (2017). Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. *NeuroImage*, 147:964–975.

Rodríguez-Galván, J. R., Merino Caviedes, S., Filgueiras Rama, D., Sánchez González, J., and Tristán Vega, A. (2023). Beyond DTI: Insights in Heart Infarcted Tissue. In *CASEIB 2023. Libro de Actas del XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica: Contribuyendo a la salud basada en valor*, pages 306–309. Universidad Politécnica de Cartagena.

Rodríguez-Galván, J. R., París, G., Tristán-Vega, A., and Alberola-López, C. (2023). Diffusion sampling schemes: A generalized methodology with nongeometric criteria. *Magnetic Resonance in Medicine*, 89(6):2270–2280.

Rosberg, A., Tuulari, J. J., Kumpulainen, V., Lukkarinen, M., Pulli, E. P., Silver, E., Copeland, A., Saukko, E., Saunavaara, J., Lewis, J. D., et al. (2022). Test–retest reliability of diffusion tensor imaging scalars in 5-year-olds. *Human Brain Mapping*, 43(16):4984–4994.

Rudrapatna, U., Parker, G. D., Roberts, J., and Jones, D. K. (2020). A comparative study of gradient nonlinearity correction strategies for processing diffusion data obtained with ultra-strong gradient MRI scanners. *Magnetic Resonance in Medicine*, 85(2):1104–1113.

Schwartz, S. C. (1967). Estimation of Probability Density by an Orthogonal Series. *The Annals of Mathematical Statistics*, 38(4):1261–1265.

Setsompop, K., Kimmlingen, R., Eberlein, E., Witzel, T., Cohen-Adad, J., McNab, J. A., Keil, B., Tisdall, M. D., Hoecht, P., Dietz, P., Cauley, S. F., Tountcheva, V., Matschl, V., Lenz, V. H., Heberlein, K., Potthast, A., Thein, H., Van Horn, J., Toga, A., others, and Wald, L. L. (2013). Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *NeuroImage*, 80:220–233.

Sexton, C. E., Mackay, C. E., Lonie, J. A., Bastin, M. E., Terrière, E., O'Carroll, R. E., and Ebmeier, K. P. (2010). MRI correlates of episodic memory in Alzheimer's disease, mild cognitive impairment, and healthy aging. *Psychiatry Research: Neuroimaging*, 184(1):57–62.

Sijbers, J., den Dekker, A. J., Scheunders, P., and Van Dyck, D. (1998). Maximum-likelihood estimation of rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357–361.

Sjölund, J., Eklund, A., Özarslan, E., and Knutsson, H. (2016). Gaussian process regression can turn non-uniform and undersampled diffusion MRI data into diffusion spectrum imaging.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., others, and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23:S208–S219.

Song, S.-K., Yoshino, J., Le, T. Q., Lin, S.-J., Sun, S.-W., Cross, A. H., and Armstrong, R. C. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. *NeuroImage*, 26(1):132–140.

Soni, N., Mehrotra, A., Behari, S., Kumar, S., and Gupta, N. (2017). Diffusion-tensor Imaging and Tractography Application in Pre-operative Planning of Intra-axial Brain Lesions. *Cureus*.

Spotorno, N., Strandberg, O., Vis, G., Stomrud, E., Nilsson, M., and Hansson, O. (2022). Measures of cortical microstructure are linked to amyloid pathology in alzheimer's disease. *Brain*, 146(4):1602–1614.

Sullivan, E. V. and Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neuroscience & Biobehavioral Reviews*, 30(6):749–761.

Tariq, M. et al. (2016). Bingham-NODDI: Mapping anisotropic orientation dispersion of neurites using diffusion MRI. *NeuroImage*, 133:207–223.

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

Thomas, C., Ye, F. Q., Irfanoglu, M. O., Modi, P., Saleem, K. S., Leopold, D. A., and Pierpaoli, C. (2014). Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. *Proceedings of the National Academy of Sciences*, 111(46):16574–16579.

Thompson, P. M., Stein, J. L., Medland, S. E., Hibar, D. P., Vasquez, A. A., Renteria, M. E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., et al. (2014). The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain imaging and behavior*, 8:153–182.

Tobisch, A., Stirnberg, R., Harms, R. L., Schultz, T., Roebroeck, A., Breteler, M. M. B., and Stöcker, T. (2018). Compressed Sensing Diffusion Spectrum Imaging for Accelerated Diffusion Microstructure MRI in Long-Term Population Imaging. *Frontiers in Neuroscience*, 12.

Tong, Q., He, H., Gong, T., Li, C., Liang, P., Qian, T., Sun, Y., Ding, Q., Li, K., and Zhong, J. (2019). Reproducibility of multi-shell diffusion tractography on traveling subjects: A multicenter study prospective. *Magnetic Resonance Imaging*, 59:1–9.

Tournier, J. D., Calamante, F., and Connelly, A. (2013). Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR in Biomedicine*, 26(12):1775–1786.

Tournier, J. D., Calamante, F., Gadian, D. G., and Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185.

Tournier, J. D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., and Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, 202:116137.

Tristán-Vega, A. and Aja-Fernández, S. (2021). Efficient and accurate EAP imaging from multi-shell dMRI with micro-structure adaptive convolution kernels and dual Fourier Integral Transforms (MiSFIT). *NeuroImage*, 227:117616.

Tristán-Vega, A., Aja-Fernández, S., and Westin, C.-F. (2012). Least squares for diffusion tensor estimation revisited: Propagation of uncertainty with rician and non-rician signals. *NeuroImage*, 59(4):4032–4043.

Tristán-Vega, A., París, G., de Luis-García, R., and Aja-Fernández, S. (2021). Accurate free-water estimation in white matter from fast diffusion MRI acquisitions using the spherical means technique. *Magnetic Resonance in Medicine*, 87(2):1028–1035.

Tristán-Vega, A., Pieciak, T., París, G., Rodríguez-Galván, J. R., and Aja-Fernández, S. (2023). HYDI-DSI revisited: Constrained non-parametric EAP imaging without q-space re-gridding. *Medical Image Analysis*, 84:102728.

Tristan-Vega, A., Vegas-Sanchez-Ferrero, G., and Aja-Fernandez, S. (2008). Local similarity measures for demons-like registration algorithms. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1087–1090.

Tristán-Vega, A. and Aja-Fernández, S. and París, G. (2022). dMRI-Lab: advanced diffusion MRI with Matlab. Accessed: 20 April 2025.

Tuch, D. S. (2002). *Diffusion MRI of complex tissue structure*. PhD thesis, Massachusetts Institute of Technology.

Tuch, D. S. (2004). Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372.

Tuch, D. S., Belliveau, J. W., and Wedeen, V. J. (1999). High angular resolution diffusion imaging of the human brain. *Proceedings of the International Society of Magnetic Resonance in Medicine*, page 321.

Uhl, Q., Pavan, T., Molendowska, M., Jones, D., Palombo, M., and Jelescu, I. (2024). Quantifying human gray matter microstructure using nexi and 300 mt/m gradients. *ISMRM Annual Meeting*.

Unterberg, A., Stover, J., Kress, B., and Kiening, K. (2004). Edema and brain trauma. *Neuroscience*, 129(4):1019–1027.

Varadarajan, D. and Haldar, J. P. (2015). A majorize-minimize framework for rician and non-central chi mr images. *IEEE transactions on medical imaging*, 34(10):2191–2202.

Venkatraman, V. K., Gonzalez, C. E., Landman, B., Goh, J., Reiter, D. A., An, Y., and Resnick, S. M. (2015). Region of interest correction factors improve reliability of diffusion imaging measures within and across scanners and field strengths. *Neuroimage*, 119:406–416.

Veraart, J., Cauter, S. V., Sunaert, S., and Sijbers, J. (2012). Conditional Least Squares Estimation of Diffusion MRI Parameters. In *Proceedings of the 20th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM)*, Melbourne, Australia. ISMRM.

Veraart, J., Fieremans, E., and Novikov, D. S. (2016a). Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, 76(5):1582–1593.

Veraart, J., Fieremans, E., Rudrapatna, U., Jones, D. K., and Novikov, D. S. (2018). Breaking the power law scaling of the dMRI signal on the Connectom scanner reveals its sensitivity to axon diameters. In *Proceedings of the 26th Annual Meeting of ISMRM, Paris, France.*

Veraart, J., Novikov, D. S., Christiaens, D., Ades-aron, B., Sijbers, J., and Fieremans, E. (2016b). Denoising of diffusion MRI using random matrix theory. *NeuroImage*, 142:394–406.

Veraart, J., Rajan, J., Peeters, R. R., Leemans, A., Sunaert, S., and Sijbers, J. (2013). Comprehensive framework for accurate diffusion MRI parameter estimation. *Magnetic resonance in medicine*, 70(4):972–984.

Veraart, J., Raven, E., Jones, D., and Palombo, M. (2023). Axon diameter mapping is confounded by glial cells. *ISMRM Annual Meeting*.

Veraart, J., Raven, E. P., Edwards, L. J., Weiskopf, N., and Jones, D. K. (2021). The variability of MR axon radii estimates in the human white matter. *Human Brain Mapping*, 42(7):2201–2213.

Veraart, J., Van Hecke, W., and Sijbers, J. (2011). Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magnetic Resonance in Medicine*, 66(3):678–686.

Vollmar, C., O'Muircheartaigh, J., Barker, G. J., Symms, M. R., Thompson, P., Kumari, V., Duncan, J. S., Richardson, M. P., and Koepp, M. J. (2010). Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. *Neuroimage*, 51(4):1384–1394.

Walsh, D. O., Gmitro, A. F., and Marcellin, M. W. (2000). Adaptive reconstruction of phased array MR imagery. *Magnetic Resonance in Medicine*, 43(5):682–690.

Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R. A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N., and Baneke, P. (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS. *Multiple Sclerosis Journal*, 26(14):1816–1821.

Wang, P., Weng, L., Xie, S., He, J., Ma, X., LI, B., Yuan, P., Wang, S., Zhang, H., Gao, Y., Wu, Q., and Niu, G. (2021a). Primary application of mean apparent propagator-MRI diffusion model in the grading of diffuse glioma. *European Journal of Radiology*, 138:109622.

Wang, P., Weng, L., Xie, S., He, J., Ma, X., Li, B., Yuan, P., Wang, S., Zhang, H., Niu, G., et al. (2021b). Primary application of mean apparent propagator-MRI diffusion model in the grading of diffuse glioma. *European journal of radiology*, 138:109622.

Ward, I. L., Raven, E. P., de la Rosa, S., Jones, D. K., Teufel, C., and von (2023). White matter microstructure in face and body networks predicts facial expression and body posture perception across development. *Human Brain Mapping*, 44(6):2307–2322.

Wedeen, V., Reese, T., Tuch, D., Weigel, M., Dou, J., Weiskoff, R., and Chessler, D. (2000). Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI. In *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, volume 8, page 82.

Wedeen, V. J., Hagmann, P., Tseng, W.-Y. I., Reese, T. G., and Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386.

Wiggins, G. C. et al. (2009). 96-Channel receive-only head coil for 3 Tesla: design optimization and evaluation. *Magnetic Resonance in Medicine*, 62(3):754–762.

World Health Organization (2025a). Dementia fact sheet. Accessed: 3 April 2025.

World Health Organization (2025b). Stroke (cerebrovascular accident). Accessed: 3 April 2025.

Wu, Y.-C. and Alexander, A. L. (2007). Hybrid diffusion imaging. *NeuroImage*, 36(3):617–629.

Wu, Y.-C., Field, A. S., and Alexander, A. L. (2008). Computation of Diffusion Function Measures in q-Space Using Magnetic Resonance Hybrid Diffusion Imaging. *IEEE Transactions on Medical Imaging*, 27(6):858–865.

Wu, Y.-C., Mustafi, S. M., Harezlak, J., Kodiweera, C., Flashman, L. A., and McAllister, T. W. (2018). Hybrid Diffusion Imaging in Mild Traumatic Brain Injury. *Journal of Neurotrauma*, 35(20):2377–2390.

Xu, T., Opitz, A., Craddock, R. C., Wright, M. J., Zuo, X.-N., and Milham, M. P. (2016). Assessing Variations in Areal Organization for the Intrinsic Brain: From Fingerprints to Reliability. *Cerebral Cortex*, 26(11):4192–4211.

Ye, W., Portnoy, S., Entezari, A., Blackband, S. J., and Vemuri, B. C. (2012). An Efficient Interlaced Multi-Shell Sampling Scheme for Reconstruction of Diffusion Propagators. *IEEE Transactions on Medical Imaging*, 31(5):1043–1050.

Yeh, F.-C., Irimia, A., Bastos, D. C. d. A., and Golby, A. J. (2021). Tractography methods and findings in brain tumors and traumatic brain injury. *NeuroImage*, 245:118651.

Yuan, J., Yeung, D. K. W., Mok, G. S., Bhatia, K. S., Wang, Y.-X. J., Ahuja, A. T., and King, A. D. (2014). Non-Gaussian analysis of diffusion weighted imaging in head and neck at 3T: a pilot study in patients with nasopharyngeal carcinoma. *PloS one*, 9(1):e87024.

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61(4):1000–1016.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57.

Zhong, J., Liu, X., Hu, Y., Xing, Y., Ding, D., Ge, X., Song, Y., Wang, S., Chen, L., Zhu, Y., Lu, W., Zhang, H., and Yao, W. (2023). Robustness of Quantitative Diffusion Metrics from Four Models: A Prospective Study on the Influence of Scan-Rescans, Voxel Size, Coils, and Observers. *Journal of Magnetic Resonance Imaging*, 60(4):1470–1483.

Zizek, S. (1991). *The truth arises from misrecognition*. Routledge, London, England.

Zucchelli, M., Deslauriers-Gauthier, S., and Deriche, R. (2020). A computational framework for generating rotation invariant features and its application in diffusion MRI. *Medical image analysis*, 60:101597.

Zuo, X.-N., Xu, T., and Milham, M. P. (2019). Harnessing reliability for neuroscience research. *Nature Human Behaviour*, 3(8):768–771.

Žižek, S. (2006). *Interview with Žižek*. The New York Times, New York.

A

Legendre Polynomials Integration

This section shows the resolution of scalar products between SH-spanned functions, since both PA and NG rely on it. First, take into account that the scalar product of two such functions f and g can be expressed in spherical coordinates as:

$$\langle f(\mathbf{q}), g(\mathbf{q})\rangle = \iiint_{\mathbb{R}^3} f(\mathbf{q})g(\mathbf{q})d\mathbf{q} = \iint_{\mathcal{S}}\int_0^\infty q^2 f(q\mathbf{u})g(q\mathbf{u})dqd\mathbf{u}. \tag{A.1}$$

The squared modulus of f results in $\|f(\mathbf{q})\|^2 = \int_0^\infty q^2 \iint_{\mathcal{S}} \left(f(q\mathbf{u})\right)^2 d\mathbf{u}dq$. Casting the attenuation signal $E(\mathbf{q})$, as in Eq. (3.7), into the previous result, we get:

$$\|E(q\mathbf{u})\|^2 = \int_0^\infty q^2 \iint_{\mathcal{S}} \left(\sum_{l,m} \hat{e}_l^{\text{ODF}}(q)\phi_l^m Y_l^m(\mathbf{u})\right)^2 d\mathbf{u}dq \tag{A.2}$$

$$= \int_0^\infty q^2 \sum_{l,m} \left(\hat{e}_l^{\text{ODF}}(q)\right)^2 (\phi_l^m)^2 dq = \sum_{l,m}(\phi_l^m)^2 \int_0^\infty q^2 \left(\hat{e}_l^{\text{ODF}}(q)\right)^2 dq. \tag{A.3}$$

To simplify, the integral in the variable q will be done separately by first replacing the convolution factors $\hat{e}_l^{\text{ODF}}(q) = 2\pi\int_{-1}^1 \Lambda_{\text{ODF}}(x,q)P_l(x)dx$, where Λ_{ODF} stands for the convolution kernel in a given shell, see Tristán-Vega and Aja-Fernández (2021):

$$E(q_i\mathbf{u}) = \iint_{\mathcal{S}} \Phi(\mathbf{v})\Lambda_{\text{ODF}}(\mathbf{u}^T\mathbf{v}; q_i)d\mathbf{v} \quad : \quad \Lambda_{\text{ODF}}(x,q_i) = \exp(-b_i\lambda_\perp)\exp(-bi\delta_\lambda x^2), \tag{A.4}$$

where $\delta_\lambda = \lambda_\parallel - \lambda_\perp$. Hence, we can continue with the computation in Eq. (A.3):

$$\int_0^\infty q^2(\hat{e}_l^{\text{ODF}}(q))^2 dq = \int_0^\infty q^2 \left(2\pi\int_{-1}^1 \Lambda_{\text{ODF}}(x,q)P_l(x)dx\right)^2 dq, \tag{A.5}$$

which results in:

$$\int_0^\infty q^2(\hat{e}_l^{\mathrm{ODF}}(q))^2 dq = 4\pi^2 \iint_{-1}^{1} \Big(q^2 \Lambda_{\mathrm{ODF}}(x_1,q)\Lambda_{\mathrm{ODF}}(x_2,q)dq\Big) P_l(x_1)P_l(x_2)dx_1dx_2 \tag{A.6}$$

$$= \frac{\pi}{(4\pi\tau\delta_\perp)^{3/2}} \iint_{-1}^{1} \frac{P_l(x_1)P_l(x_2)}{(2\rho_\lambda + x_1^2 + x_2^2)^{3/2}} dx_1dx_2 \tag{A.7}$$

$$= \frac{\pi}{(4\pi\tau\delta_\perp)^{3/2}} I_l(\rho_\lambda), \tag{A.8}$$

with $\rho_\lambda = \lambda_\perp/\delta_\lambda$. Although the previous equation does not provide a closed form expression, it shows that the computation of the norm of a function represented as the convolution with an SH-spanned ODF reduces to the numerical pre-calculation of the integral $I_l(\rho_\lambda)$ for a sufficiently large range of ρ_l and for $l = 0, 2, \ldots$ up to the desired order.

A.0.1 Propagator Anisotropy for EAP composite signal

In this section, the formulation of the Propagator Anisotropy, as defined in (Özarslan et al., 2013), will be developed for MiSFIT's composite signal. First, let us define $\mathrm{PA} = \gamma\left(\sin(\angle({}_cE(\mathbf{q}),{}_cO(q))), \epsilon\right)$, where $\sin(x)$ will be computed as $\sqrt{1-\cos^2(x)}$. The isotropic equivalent ${}_cO(q)$ is defined as the directional average of ${}_cE(\mathbf{q})$:

$${}_cO(q) = \frac{1}{4\pi}\int_{\mathcal{S}} {}_cE(q\mathbf{u})d\mathbf{u}. \tag{A.9}$$

The attenuation signal, in turn, is expressed in terms of SH basis functions as in Tristán-Vega and Aja-Fernández (2021):

$${}_cE(q\mathbf{u}) = (1-f)\cdot\exp\left(-4\pi^2\tau q^2 D_0\right) + f\cdot\sum_{l,m}\hat{e}_l^{\mathrm{ODF}}(q)\phi_l^m Y_l^m(\mathbf{u}), \tag{A.10}$$

so that ${}_cO(q)$ is trivially computed from the DC component of the expansion:

$${}_cO(q) = (1-f)\cdot\exp\left(-4\pi^2\tau q^2 D_0\right) + f\frac{1}{(4\pi)^{3/2}}\hat{e}_0^{\mathrm{ODF}}(q)\phi_l^m. \tag{A.11}$$

To begin with, the squared module of the composite signal is obtained by integrating, in spherical coordinates (unit sphere $\mathcal{S}$ and radial coordinate q), the squared ${}_cE(q\mathbf{u})$:

$$\|{}_cE(q\mathbf{u})\|^2 = \int_0^\infty q^2 \iint_{\mathcal{S}} \Big((1-f)\cdot\exp\left(-4\pi^2\tau q^2 D_0\right) + f\cdot\sum_{l,m}\hat{e}_l^{\mathrm{ODF}}(q)\phi_l^m Y_l^m(\mathbf{u})\Big)^2 d\mathbf{u}dq. \tag{A.12}$$

Applying a binomial expansion to Eq. (A.12), we get three terms, namely: isotropic, anisotropic and mixed. To ease the readability, the squared composite EAP will be expressed as $({}_cE(q\mathbf{u}))^2 =_c$

$E^2_{\text{iso}}(q\mathbf{u}) +_c E^2_{\text{ani}}(q\mathbf{u}) + 2_cE_{\text{mix}}(q\mathbf{u})$, so that we have:

$$
\begin{aligned}
{}_cE^2_{\text{iso}}(q\mathbf{u}) &= (1-f)^2 \exp\left(-8\pi^2\tau q^2 D_0\right) dq; && \text{(A.13)}\\
{}_cE^2_{\text{ani}}(q\mathbf{u}) &= f^2 \sum_{l,m} (\hat{e}_l^{\text{ODF}}(q))^2 (\phi_l^m)^2 (Y_l^m(\mathbf{u}))^2; && \text{(A.14)}\\
2_cE_{\text{mix}}(q\mathbf{u}) &= 2f(1-f)\exp\left(-4\pi^2\tau q^2 D_0\right) \sum_{l,m} \hat{e}_l^{\text{ODF}}(q)\phi_l^m Y_l^m(\mathbf{u}). && \text{(A.15)}
\end{aligned}
$$

For each of these operands, respectively:

$$
\begin{aligned}
\int_0^\infty q^2 \iint_{\mathcal{S}} {}_cE^2_{\text{iso}}(q\mathbf{u}) d\mathbf{u} dq &= (1-f)^2 4\pi \int_0^\infty q^2 \exp\left(-4\pi^2\tau q^2 D_0\right) dq\\
&= (1-f)^2 \frac{1}{(8\pi\tau D_0)^{3/2}}; && \text{(A.16)}\\
\int_0^\infty q^2 \iint_{\mathcal{S}} {}_cE^2_{\text{ani}}(q\mathbf{u}) d\mathbf{u} dq &= f^2 \sum_{l,m} (\phi_l^m)^2 \frac{\pi}{(4\pi\tau\delta_\lambda)^{3/2}} \mathcal{I}_l(\rho_\lambda); && \text{(A.17)}\\
\int_0^\infty q^2 \iint_{\mathcal{S}} 2_cE_{\text{mix}}(q\mathbf{u}) d\mathbf{u} dq &= \frac{2f(1-f)}{\sqrt{4\pi}} \int_0^\infty q^2 \exp\left(-4\pi^2\tau q^2 D_0\right) \hat{e}_0^{\text{ODF}}(q)\phi_0^0 dq, && \text{(A.18)}
\end{aligned}
$$

where the resolution of the second integral, referring to the purely-anisotropic term $\|{}_cE_{\text{ani}}\|^2$, is addressed in Appendix A. The last integral, referring to the integral of the mixed signal ${}_cE_{\text{mix}}(q\mathbf{u})$, is trivially solved since all SH basis, except for Y_0^0, have zero mean. Now, expanding the ODF 0-th SH coefficients in Eq. (A.18):

$$
\hat{e}_0^{\text{ODF}} = 2\pi \int_{-1}^{1} \Lambda_{\text{ODF}}(x,q) P_0(x) dx, \qquad \text{(A.19)}
$$

where $\Lambda = \exp(-b\lambda_\perp)\exp(-b\delta_\lambda x^2)$ results in:

$$
\begin{aligned}
\int_0^\infty q^2 \iint_{\mathcal{S}} 2E_{\text{mix}}(q\mathbf{u}) d\mathbf{u} dq &= 8\pi\sqrt{\pi} f(1-f)\phi_0^0 \int_{-1}^{1} \left(\int_0^\infty q^2 e^{-b(D_0+\lambda_\perp+\delta_\lambda x^2)} dq \right) dx\\
&= \frac{2\pi^2 f(1-f)}{(4\pi^2\tau)^{3/2}} \phi_0^0 \int_{-1}^{1} \frac{1}{(D_0+\lambda_\perp+\delta_\lambda x^2)^{3/2}} dx\\
&= \frac{2\pi^2 f(1-f)}{(4\pi^2\tau\delta_\lambda)^{3/2}} \phi_0^0 \frac{2}{(D_0/\delta_\lambda+\rho_\lambda)\sqrt{D_0/\delta_\lambda+\rho_\lambda+1}}\\
&= \frac{4\sqrt{\pi} f(1-f)}{(4\pi\tau)^{3/2}} \frac{1}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}} \phi_0^0. && \text{(A.20)}
\end{aligned}
$$

Finally, adding all the terms and simplifying the expression, we obtain the following result:

$$
\|{}_cE(q\mathbf{u})\|^2 = \psi \left(\frac{(1-f)^2}{(2D_0)^{3/2}} + f^2 \sum_{l,m} \frac{(\phi_l^m)^2 \pi}{\delta_\lambda^{3/2}} \mathcal{I}_l(\rho_\lambda) + \frac{4\sqrt{\pi} f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}} \phi_0^0 \right). \qquad \text{(A.21)}
$$

where $\psi = (4\pi\tau)^{-3/2}$. To compute the norm of ${}_cO(q)$, we use the same reasoning as in Eq. (A.11), and simply keep the term for $l = 0$ in Eq. (A.21):

$$\|{}_cO(q)\|^2 = \psi \left(\frac{(1-f)^2}{(2D_0)^{3/2}} + f^2 \frac{(\phi_0^0)^2 \pi}{\delta_\lambda^{3/2}} \mathcal{I}_0(\rho_\lambda) + \frac{4\sqrt{\pi} f(1-f)}{(D_0 + \lambda_\perp)\sqrt{D_0 + \lambda_\parallel}} \phi_0^0 \right). \tag{A.22}$$

Finally, note that:

$$\begin{aligned}
\langle {}_cE(q\mathbf{u}), {}_cO(q) \rangle &= \int_0^\infty q^2 \iint_{\mathcal{S}} {}_cE(q\mathbf{u}) {}_cO(q) d\mathbf{u} dq \\
&= \int_0^\infty q^2 \iint_{\mathcal{S}} {}_cE(q\mathbf{u}) \frac{1}{4\pi} \iint_{\mathcal{S}} {}_cE(q\mathbf{v}) d\mathbf{v} d\mathbf{u} dq \\
&= \frac{1}{4\pi} \int_0^\infty q^2 \left(\iint_{\mathcal{S}} {}_cE(q\mathbf{u}) d\mathbf{u} \right) \left(\iint_{\mathcal{S}} {}_cE(q\mathbf{v}) d\mathbf{v} \right) dq \\
&= 4\pi \int_0^\infty q^2 {}_cO^2(q) dq = \|{}_cO(q)\|^2 ,
\end{aligned} \tag{A.23}$$

hence the squared cosine between both two functions reads:

$$\begin{aligned}
\cos^2(\angle({}_cE(\mathbf{q}), {}_cO(q))) &= \frac{\langle {}_cE(q\mathbf{u}), {}_cO(q) \rangle^2}{\|{}_cE(q\mathbf{u})\|^2 \|{}_cO(q)\|^2} = \frac{\|{}_cO(q)\|^2}{\|{}_cE(q\mathbf{u})\|^2} = \\
&= \frac{\frac{(1-f)^2}{(2D_0)^{3/2}} + \frac{4\sqrt{\pi} f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}} \phi_0^0 + f^2 (\phi_0^0)^2 \pi \delta_\lambda^{-3/2} \mathcal{I}_0(\rho_\lambda)}{\frac{(1-f)^2}{(2D_0)^{3/2}} + \frac{4\sqrt{\pi} f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}} \phi_0^0 + f^2 \sum_{l,m} (\phi_l^m)^2 \pi \delta_\lambda^{-3/2} \mathcal{I}_l(\rho_\lambda)}.
\end{aligned} \tag{A.24}$$

B

Non-Gaussianity for EAP composite signal

As in the previous Section, we define the NG in terms of the sine between two signals, which is computed from the cosine:

$$\text{NG} = \sin(\angle({}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}}))) = \sqrt{1-\cos^2(\angle({}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})))}, \tag{B.1}$$

where ${}_cG$ is the DTI-like (Gaussian) propagator. The cosine between the signals, in turn, is computed as:

$$\cos(\angle({}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}}))) = \frac{\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle}{\|{}_cE(\mathbf{q})\|\|{}_cG(\mathbf{q},\widehat{\mathbf{D}})\|}. \tag{B.2}$$

First, we compute the numerator of Eq. (B.2) as follows:

$$\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle = \iiint_{\mathbb{R}^3} \left((1-f)e^{-4\pi^2\tau q^2 D_0} + fE(\mathbf{q})\right) e^{-4\pi^2\tau q^2 \mathbf{u}^T\widehat{\mathbf{D}}\mathbf{u}} d\mathbf{u}, \tag{B.3}$$

where $\widehat{\mathbf{D}}$ is the tensor associated to the DTI-like Gaussian propagator. Developing the multiplication of exponential functions, we obtain:

$$\begin{aligned}\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle = (1-f)\iiint_{\mathbb{R}^3} e^{-4\pi^2\tau q^2\mathbf{u}^T(\mathbf{D_0}+\widehat{\mathbf{D}})\mathbf{u}} d\mathbf{u} \\ + f\iint_{\mathcal{S}}\Phi(\mathbf{v})\iiint_{\mathbb{R}^3} e^{-4\pi^2\tau q^2\mathbf{u}^T(\widehat{\mathbf{D}}+\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}})\mathbf{u}} dq d\mathbf{u} d\mathbf{v},\end{aligned} \tag{B.4}$$

where $\mathbf{D_0} = D_0\mathbf{I}_3$, $\mathbf{I}_3$ is the 3×3 identity matrix and $\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}$ is an elemental diffusion tensor with eigenvectors $\lambda_{\parallel}$, $\lambda_{\perp}$, aligned with direction $\mathbf{v}\in\mathcal{S}$. Since the integral of the exponential over

$\mathbb{R}^3$ is trivial, we can obtain a simpler solution:

$$\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle = \frac{\psi(1-f)}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + f\iint_{\mathcal{S}} \frac{\psi\Phi(\mathbf{v})}{\sqrt{|\widehat{\mathbf{D}}+\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}|}}d\mathbf{v}, \tag{B.5}$$

where, as before, $\psi = (4\pi\tau)^{-3/2}$; and $|\cdot|$ stands for the determinant of a matrix. Now, expanding the ODF in the SH basis, as $\Phi(\mathbf{v}) = \sum_{l,m}\phi_l^m Y_l^m(\mathbf{v})$, we obtain the following expression:

$$\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}}))\rangle = \frac{\psi(1-f)}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + f\sum_{l,m}\phi_l^m \underbrace{\iint_{\mathcal{S}} \frac{\psi}{\sqrt{|\widehat{\mathbf{D}}+\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}|}} Y_l^m(\mathbf{v})d\mathbf{v}}_{\text{Projection onto SH basis}}. \tag{B.6}$$

The latter integral represents the projection of the corresponding function in the SH basis, hence, it can be accurately computed by LS fitting, for all l and m simultaneously, whenever it is properly sampled over the unit sphere. Such sampling strategy is described in depth in Appendix B.0.1 and results in:

$$\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle = \frac{\psi(1-f)}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + \psi f\sum_{l,m}\phi_l^m\xi_l^m, \tag{B.7}$$

where $|\widehat{\mathbf{D}}+\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}|^{-1/2} = \sum_{l,m}\xi_l^m Y_l^m(\mathbf{v})$. Finally:

$$\langle {}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\rangle = \psi\left(\frac{1-f}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + f\sum_{l,m}\phi_l^m\xi_l^m\right). \tag{B.8}$$

The EAP module was computed in Eq. (A.22), and the module of the DTI-like Gaussian propagator admits the closed form $\|\ {}_cG\| = \psi(|2\widehat{\mathbf{D}}|)^{-1/2}$. Hence:

$$\cos^2\left({}_cE(\mathbf{q}),\ {}_cG(\mathbf{q},\widehat{\mathbf{D}})\right) = \frac{\left(\frac{1-f}{\sqrt{|\mathbf{D_0}+\widehat{\mathbf{D}}|}} + f\sum_{l,m}\phi_l^m\xi_l^m\right)^2\sqrt{8|\widehat{\mathbf{D}}|}}{\frac{(1-f)^2}{(2D_0)^{3/2}} + f^2\sum_{l,m}\frac{(\phi_l^m)^2\pi}{\delta_\lambda^{3/2}}\mathcal{I}_l(\rho_\lambda) + \frac{4\sqrt{\pi}f(1-f)}{(D_0+\lambda_\perp)\sqrt{D_0+\lambda_\parallel}}\phi_0^0}. \tag{B.9}$$

B.0.1 Efficient Sampling of Spherical Function

Eq. (B.6) can be seen as a projection onto the SH basis, and can be solved efficiently by *properly* sampling $|\widehat{\mathbf{D}}+\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}|^{-1/2}$ in the variable $\mathbf{v}$. This is done by means of a spectral analysis at each voxel, by making use of the matrix determinant lemma (Harville, 1997):

$$|\mathbf{A}+\mathbf{u}\mathbf{v}^T| = (1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u})|\mathbf{A}|. \tag{B.10}$$

It can be applied by noticing that the elemental tensor $\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}$ can be written as:

$$\widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}} = \sqrt{\delta_\lambda}\mathbf{v}\sqrt{\delta_\lambda}\mathbf{v}^T + \lambda_\perp\mathbf{I}_3, \tag{B.11}$$

The proof is straightforward by noticing that both sides of the previous equation have the same eigenvalues and eigenvectors. Applying then the lemma in Eq. (B.10):

$$|\widehat{\mathbf{D}} + \widehat{\mathbf{D}}^{\lambda}_{\mathbf{v}}| = \left(1 + \delta_{\lambda}\mathbf{v}^T(\lambda_{\perp}\mathbf{I}_3 + \widehat{\mathbf{D}})\mathbf{v}\right)|\lambda_{\perp}\mathbf{I}_3 + \widehat{\mathbf{D}}|, \tag{B.12}$$

which reduces the problem to compute the determinant of $\lambda_{\perp}\mathbf{I}_3 + \widehat{\mathbf{D}}$ at each voxel, independent on the variable $\mathbf{v}$. Note that $\widehat{\mathbf{D}}$ is positive (semi)definite and the MiSFIT approach ensures $\lambda_{\perp}$ is greater than 0, so that Eq. (B.12) is always well defined and may be used to draw an arbitrary number of samples for $\mathbf{v} \in \mathcal{S}$. This means the coefficients ξ_l^m can be computed with arbitrary numerical precision up to any desired degree L, with the unique limit being that imposed by computational load constraints.

C

Other results

This section shows some intermediate results that were not included in the document. Contrarily to the already seen appendices, these are not central to the development of the experiments but show some interesting results that support the steps finally taken.

Outlier Rejection Procedure:
The main goal of the outlier rejection is to remove those voxels placed in the edges of the WM regions; thus palliating the effect of misregistration. Fig. C.1 shows different methods for outlier removal. As it can be observed, some methods offer a more conservative approach and reject fewer outliers (i.e. standard rejection and mean-based rejection) than others (percentile-based one). The mean-based rejection is too conservative and there is almost no rejection. The two percentile approaches are probably too loose and reject many outliers that are placed in the middle of the WM labels. Therefore, the standard outlier rejection was chosen since it is the one more aligned with our needs: It successfully removes outliers placed on the edges of WM labels to palliate any possible misregistration. Fig. C.2 shows the placement of such outliers in several slices of the first session of the subject 1 of the MICRA dataset.

DTI-like Gaussian Approximation:
The current way of computing the DTI-like Gaussian propagator is motivated by the computational efficiency that MiSFIT offers. By the way it is formalized, MAPL is built upon successive refinements of the DTI approximation so that the Gaussian equivalent is the diffusion tensor that both (1) best describes the low b-value ($b < 2000\text{s/mm}^2$) regime and (2) mathematically minimizes the cost:

$$\min_{G(\mathbf{q})} \iiint_{\mathbb{R}^3} ||E(\mathbf{q}) - G(\mathbf{q})||^2 d\mathbf{q} \tag{C.1}$$

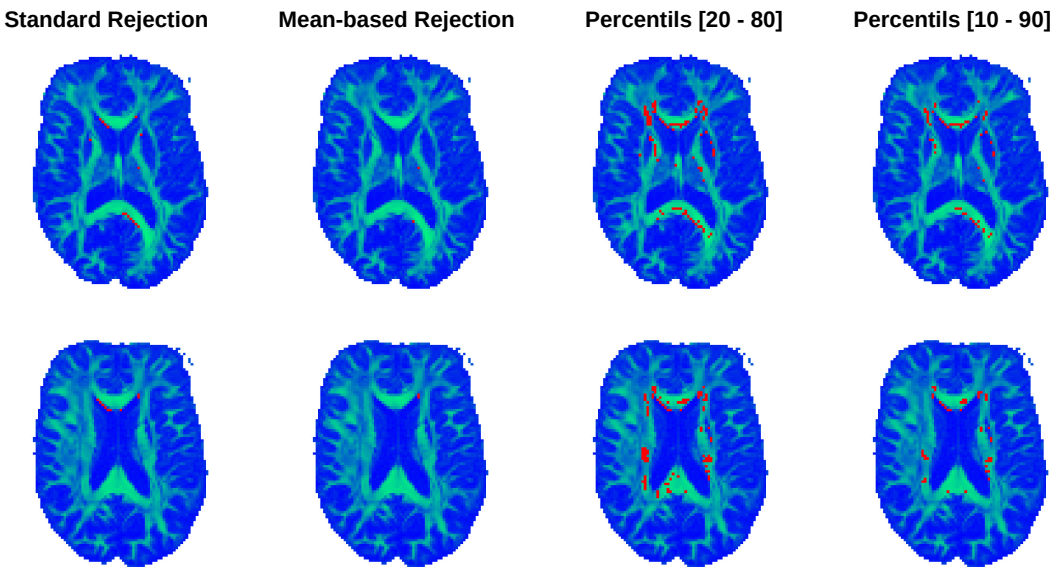


Figure C.1: Outlier Rejection Methods Comparison: Outliers (red) overlayed onto several slices of the FA map (Subject 1, session 1), using 4 different outlier rejection methodologies: Standard Rejection, Mean-based Rejection, Percentils [20-80] and Percentils [10-90].

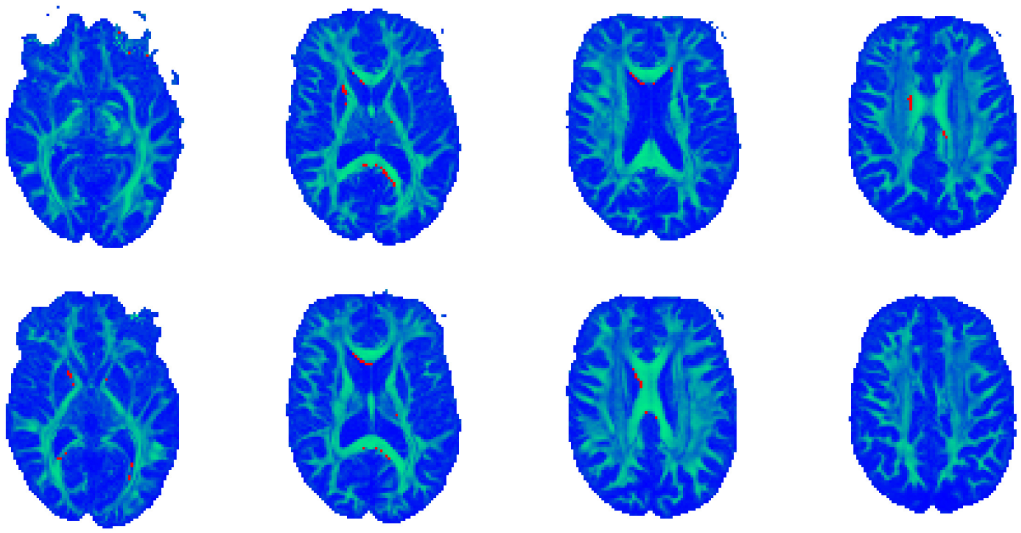


Figure C.2: Standard Outlier Rejection Method: Outliers (red) overlayed onto several slices of the FA map (Subject 1, session 1), using the standard outlier rejection approach. Note how the outliers are correctly placed in the edges of the WM regions.

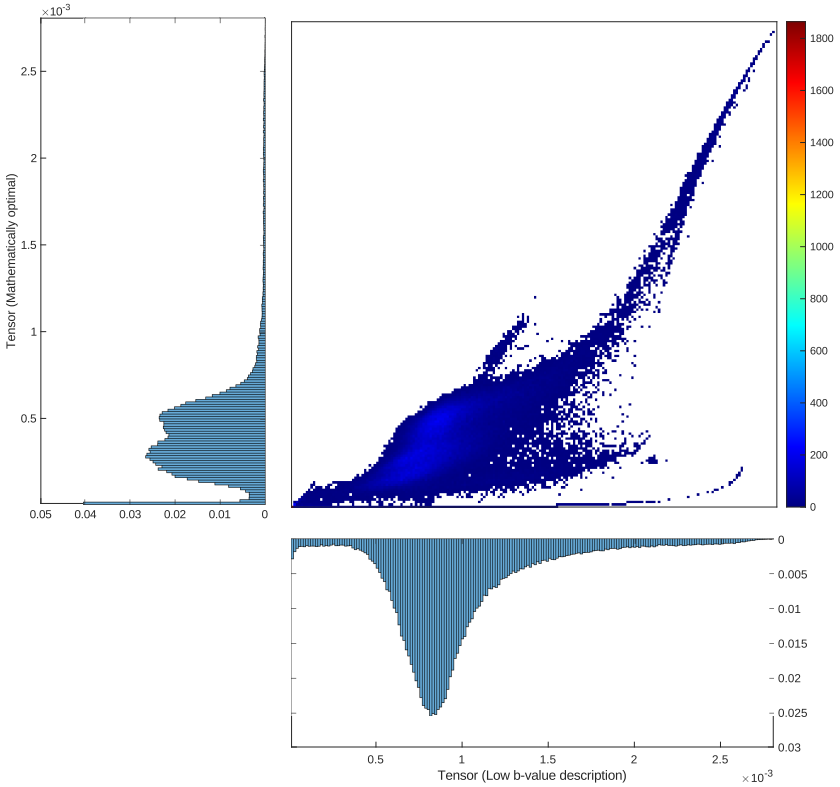


Figure C.3: 2D joint histograms of both DTI representations ($M \times N \times P \times 6$).

In MiSFIT, however, we can only aim at obtaining one of both representations. We could proceed by mimicking Eq. (C.1) via a discretization in a Cartesian lattice, so that the integral might be approximated with arbitrary precision, and the mathematically optimal Gaussian computed by means of (possibly constrained) numerical optimization. Nevertheless, this solution not only would compromise the anatomical interpretation of the NG, since we are no longer measuring deviations from the standard, low b-value DTI approach; but would also involve solving a non-linear non-convex optimization problem, where no optimal solution is assured. Therefore, MiSFIT currently aims at obtaining the Gaussian counterpart that best describes the low b-value regime, by minimizing:

$$\min_{D} \sum_{k=0}^{K} \| b\mathbf{g}_k^T \mathbf{D} \mathbf{g}_k - \ln\left(\frac{S_k}{S_0}\right) \|^2 \tag{C.2}$$

Notice the problem is reformulated in its logarithmic domain, where convexity—and thus, optimality—is assured. Fig. C.3 shows the 2D joint histogram of the tensorial signals ($M \times N \times P \times 6$) for both implementations. Also, Fig. C.4 show the 2D joint histogram of the NG computed with both implementations against the MAPL's map together with both Pearson coefficients. As it can be seen, there is more similarity between MAPL's NG and the current implementation of the NG than the mathematically optimal one.

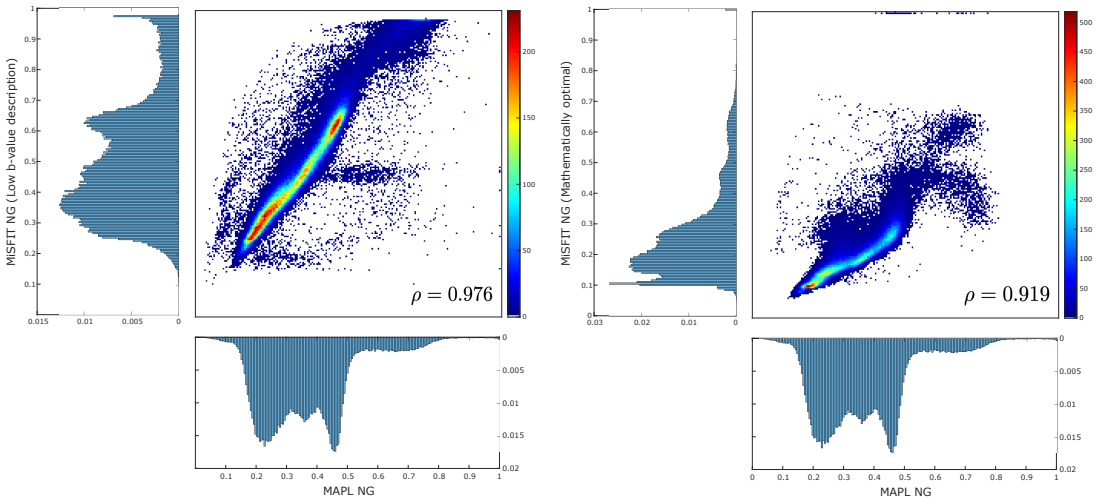


Figure C.4: 2D joint histograms of the NG maps obtained with the two implementations (current implementation, low b-value description, in the left; mathematically optimal, right) against the MAPL's NG. Pearson correlation coefficient is included.

D

Reliability computation from a standard two-way ANOVA

Let B be the number of subjects and let S be the number of sessions acquired for each subject. For the j-th session of the i-th subject ($1 \leq j \leq S$, $1 \leq i \leq B$), let R_{ij} be the number of measurements inside the tract of interest (e.g. all voxels of the RTOP map of the j-th session of the i-the subject inside the external capsule). We will denote $R_{i\cdot}$ the total amount of measurements for subject i:

$$R_{i\cdot} = \sum_{j=1}^{S} R_{ij}. \tag{D.1}$$

Conversely, the total amount of measurements corresponding to session j (gathering all subjects), $R_{\cdot j}$, is:

$$R_{\cdot j} = \sum_{i=1}^{B} R_{ij}, \tag{D.2}$$

so that the total amount of data, R, reads:

$$R = \sum_{i=1}^{B}\sum_{j=1}^{S} R_{ij} = \sum_{i=1}^{B} R_{i\cdot} = \sum_{j=1}^{S} R_{\cdot j}. \tag{D.3}$$

The samples of the dataset (e.g. all RTOP values inside the External Capsule of all subjects and sessions) will be indexed with the tuple (i, j, r): $\{y_{ijr}\}$, $1 \leq i \leq B$, $1 \leq j \leq S$, $1 \leq r \leq R_{ij}$. With this notation we can define, respectively, the per subject-session sample mean $\bar{y}_{ij\cdot}$, the per subject

sample mean $\bar{y}_{i\cdot\cdot}$, the per-session sample mean $\bar{y}_{\cdot j\cdot}$, and the overall sample mean $\bar{y}_{\cdots}$ as:

$$\bar{y}_{ij\cdot} = \frac{1}{R_{ij}} \sum_{r=1}^{R_{ij}} y_{ijr}, \quad 1 \leq i \leq B;\ 1 \leq j \leq S; \tag{D.4}$$

$$\bar{y}_{i\cdot\cdot} = \frac{1}{R_{i\cdot}} \sum_{j=1}^{S} \sum_{r=1}^{R_{ij}} y_{ijr} = \frac{1}{R_{i\cdot}} \sum_{j=1}^{S} R_{ij}\, \bar{y}_{ij\cdot}, \quad 1 \leq i \leq B; \tag{D.5}$$

$$\bar{y}_{\cdot j\cdot} = \frac{1}{R_{\cdot j}} \sum_{i=1}^{B} \sum_{r=1}^{R_{ij}} y_{ijr} = \frac{1}{R_{\cdot j}} \sum_{i=1}^{B} R_{ij}\, \bar{y}_{ij\cdot}, \quad 1 \leq j \leq S; \tag{D.6}$$

$$\begin{aligned} \bar{y}_{\cdots} &= \frac{1}{R} \sum_{i=1}^{B} \sum_{j=1}^{S} \sum_{r=1}^{R_{ij}} y_{ijr} = \frac{1}{R} \sum_{i=1}^{B} R_{i\cdot}\, \bar{y}_{i\cdot\cdot} = \frac{1}{R} \sum_{j=1}^{S} R_{\cdot j}\, \bar{y}_{\cdot j\cdot} \\ &= \frac{1}{R} \sum_{i=1}^{B} \sum_{j=1}^{S} R_{ij}\, \bar{y}_{ij\cdot\cdot} \end{aligned} \tag{D.7}$$

The standard two-way ANOVA test for unbalanced data splits the total sum of squares (SS_T) into four uncorrelated terms corresponding to subject variability (SS_{sb}), session variability (SS_{ss}), interaction (SS_{sb-ss}), and intra-class (i.e. unexplained) variability (SS_E).

$$\begin{aligned} \underbrace{\sum_{i=1}^{B} \sum_{j=1}^{S} \sum_{r=1}^{R_{ij}} y_{ijr}^2 \;-\; R\,\bar{y}_{\cdots}^2}_{\mathrm{SS}_T} &= \underbrace{\sum_{i=1}^{B} R_{i\cdot} \bar{y}_{i\cdot\cdot}^2 \;-\; R\,\bar{y}_{\cdots}^2}_{SS_{sb}} + \underbrace{\sum_{j=1}^{S} R_{\cdot j} \bar{y}_{\cdot j\cdot}^2 \;-\; R\,\bar{y}_{\cdots}^2}_{\mathrm{SS}_{ss}} \\ &+ \underbrace{\sum_{i=1}^{B} \sum_{j=1}^{S} R_{ij} \bar{y}_{ij\cdot}^2 \;-\; \sum_{i=1}^{B} R_{i\cdot} \bar{y}_{i\cdot\cdot}^2 \;-\; \sum_{j=1}^{S} R_{\cdot j} \bar{y}_{\cdot j\cdot}^2 \;+\; R\,\bar{y}_{\cdots}^2}_{\mathrm{SS}_{sb-ss}} \\ &+ \underbrace{\sum_{i=1}^{B} \sum_{j=1}^{S} \sum_{r=1}^{R_{ij}} y_{ijr}^2 \;-\; \sum_{i=1}^{B} \sum_{j=1}^{S} R_{ij} \bar{y}_{ij\cdot}^2}_{\mathrm{SS}_E}. \end{aligned} \tag{D.8}$$

Under the usual assumptions of Gaussian distributed, uncorrelated measurements, SS_T has $R-1$ DOF; SS_{sb} has $B-1$, SS_{ss} has $S-1$, SS_{sb-ss} has $(B-1)(S-1)$, and SS_E has $R-BS$ DOF. SS_T therefore follows a χ^2 distribution with $R-1$ DOF.

We will define $\mathrm{SS}_{sep} = \mathrm{SS}_{sb}$, and $\mathrm{SS}_{inc} = \mathrm{SS}_{ss} + \mathrm{SS}_{sb-ss}$ (which will have $S - 1 + (B-1)(S-1) = B(S-1)$ DOF); thereby resulting in:

$$\begin{aligned} \mathrm{SS}_T &= \underbrace{\sum_{i=1}^{B} R_{i\cdot} \left(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}\right)^2}_{\mathrm{SS}_{sep}} + \underbrace{\sum_{i=1}^{B} \sum_{j=1}^{S} R_{ij} \left(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot}\right)^2}_{\mathrm{SS}_{inc}} + \\ &+ \underbrace{\sum_{i=1}^{B} \sum_{j=1}^{S} \left(\sum_{r=1}^{R_{ij}} y_{ijr} - \bar{y}_{ij\cdot} \right)^2}_{\mathrm{SS}_E}. \end{aligned} \tag{D.9}$$

Once the terms in Eq. (D.9) are normalized by their respective numbers of DOF:

$$\mathrm{MS}_{sep} = \frac{\mathrm{SS}_{sep}}{B-1}; \quad \mathrm{MS}_{inc} = \frac{\mathrm{SS}_{inc}}{B(S-1)}; \quad \mathrm{MS}_{E} = \frac{\mathrm{SS}_{E}}{R - B\,S}, \tag{D.10}$$

$\mathrm{MS}_{\mathrm{sep}}$ and $\mathrm{MS}_{\mathrm{inc}}$ are metrics that portray, respectively, the deviation of the within-subject means with respect to the overall mean (i.e., separability across subjects), and the deviation of each session with respect to the corresponding within-subject mean (i.e., inconsistency across sessions).

E

Reliability – repeatability Tables: MICRA

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	**.505**	.638	.885	*.943*	.907	*.501*	.958	**.998**	.996	.995	*.332*	.892	**.979**	.935	.954
PCT	*.892*	.744	**.725**	.880	.782	.987	*.905*	**.994**	.988	.992	.893	*.709*	**.977**	.906	.962
GCC	**.490**	.735	*.862*	.785	.816	*.451*	.952	.996	.997	**.997**	*.295*	.840	.975	**.984**	.983
BCC	.839	.892	*.892*	.758	**.659**	.997	*.975*	**.998**	.997	.997	.983	*.811*	.984	.989	**.992**
SCC	**.565**	.648	*.762*	.642	.678	*.844*	.973	**.998**	.995	.995	*.702*	.926	**.990**	.986	.985
FRX	**.354**	.471	*.543*	.456	.470	.995	*.852*	.949	.996	**.996**	.992	*.753*	.894	.992	**.993**
CST-R	.857	.578	.743	*.887*	.821	.981	*.885*	**.994**	.975	.980	.881	*.764*	**.976**	.815	.900
CST-L	**.436**	.747	.633	*.886*	.795	*.420*	.892	**.993**	.980	.978	*.290*	.677	**.981**	.852	.902
ML-R	.563	*.634*	**.488**	.620	.586	.961	*.895*	.965	**.972**	.967	.915	*.756*	**.933**	.929	.924
ML-L	*.694*	.635	**.545**	.623	.559	.979	*.924*	.970	**.982**	.975	.936	*.817*	.937	**.954**	.946
ICP-R	*.770*	.623	**.548**	.741	.577	.962	*.881*	**.974**	.971	.959	.853	*.735*	**.945**	.897	.908
ICP-L	.712	*.770*	**.554**	.704	.583	**.974**	*.938*	.966	.969	.960	.915	*.778*	**.928**	.902	.909
SCP-R	.560	**.502**	*.632*	.570	.557	.922	*.797*	**.975**	.936	.942	.839	*.661*	**.934**	.863	.878
SCP-L	.593	*.641*	**.396**	.549	.492	.961	*.880*	**.977**	.961	.957	.909	*.724*	**.963**	.918	.919
CP-R	.763	**.616**	*.824*	.769	.779	*.942*	.958	**.997**	.979	.984	*.794*	.898	**.981**	.915	.933
CP-L	.520	**.518**	.722	.703	*.732*	*.523*	.950	**.997**	.984	.988	*.344*	.902	**.990**	.948	.956
ALIC-R	*.801*	**.604**	.637	.746	.715	.987	*.915*	**.994**	.987	.990	.939	*.811*	**.983**	.953	.965
ALIC-L	*.804*	**.629**	.685	.750	.662	.974	*.955*	**.990**	.988	.988	*.882*	.887	**.968**	.952	.966
PLIC-R	**.397**	.641	.695	*.805*	.719	*.488*	.962	**.996**	.990	.990	*.365*	.901	**.987**	.952	.964
PLIC-L	.809	.553	**.493**	*.875*	.810	.978	*.953*	**.995**	.987	.985	*.896*	.900	**.990**	.903	.927
RPIC-R	*.769*	**.509**	.750	.760	.699	.990	*.894*	.980	.993	**.994**	.958	*.806*	.925	.972	**.980**
RPIC-L	*.900*	.658	.608	.861	.767	**.992**	*.952*	.990	.992	.991	.929	*.872*	**.974**	.946	.962
ACR-R	*.942*	**.761**	.931	.903	.899	**.999**	*.955*	.999	.999	.999	.987	*.834*	.985	**.988**	.986
ACR-L	*.958*	**.760**	.945	.935	.919	**.999**	*.945*	.999	.999	.999	.983	*.804*	.982	**.984**	.983
SCR-R	*.914*	**.730**	.876	.883	.875	**.999**	*.921*	.998	.999	.999	.986	*.758*	.983	.989	**.989**
SCR-L	*.946*	**.716**	.898	.930	.905	**.998**	*.960*	.997	.998	.998	.970	*.873*	.974	.971	**.979**
PCR-R	*.850*	**.513**	.727	.813	.727	**.994**	*.690*	.992	.993	.993	.962	*.520*	.971	.963	**.976**
PCR-L	*.855*	.636	.722	.819	.768	**.997**	*.899*	.994	.996	.996	.978	*.764*	.977	.977	**.981**
PTR-R	*.854*	**.546**	.778	.789	.779	.996	*.821*	**.997**	.992	.994	.970	*.675*	**.985**	.965	.975
PTR-L	*.889*	.583	.783	.812	.790	**.998**	*.826*	.997	.996	.997	.980	*.664*	**.989**	.981	.987
SS-R	*.858*	**.571**	.816	.812	.802	.995	*.907*	**.996**	.991	.992	.964	*.807*	**.978**	.955	.961
SS-L	*.908*	**.644**	.800	.843	.794	.994	*.943*	**.994**	.988	.988	.939	*.854*	**.972**	.926	.943
EC-R	*.869*	.656	.704	.809	.630	**.996**	*.913*	.996	.994	.994	.968	*.783*	**.985**	.970	.983
EC-L	*.895*	**.621**	.729	.838	.745	**.993**	*.862*	.988	.991	.983	.935	*.703*	**.958**	.949	.935
CGG-R	*.918*	.720	**.535**	.807	.703	**.995**	*.953*	.995	.995	.994	.947	*.850*	**.990**	.974	.981
CGG-L	*.930*	**.634**	.684	.865	.805	.995	*.968*	**.996**	.994	.994	.931	*.917*	**.986**	.957	.968
CGH-R	*.808*	.583	**.568**	.768	.679	.977	*.907*	**.980**	.964	.972	.890	*.802*	**.954**	.862	.917
CGH-L	*.866*	**.471**	.605	.810	.712	**.993**	*.949*	.993	.990	.989	.950	*.907*	**.981**	.948	.963
FX/ST-R	*.722*	.647	**.498**	.583	.580	.982	*.926*	.957	**.984**	.979	.939	*.815*	.917	**.961**	.952
FX/ST-L	*.781*	.649	**.643**	.708	.644	**.975**	*.940*	.967	.974	.974	.894	*.846*	.912	.917	**.929**
SLF-R	*.909*	**.662**	.896	.819	.826	.999	*.940*	**.999**	.999	.999	.986	*.840*	.989	**.992**	.992
SLF-L	*.956*	**.768**	.924	.866	.885	.999	*.962*	**.999**	.998	.998	.974	*.855*	.985	**.987**	.986
SFOF-R	*.748*	**.470**	.685	.719	.738	.847	*.765*	**.982**	.907	.941	*.582*	.633	**.946**	.733	.808
SFOF-L	*.809*	.585	**.539**	.749	.698	.936	.916	**.976**	.891	*.780*	.735	.819	**.949**	.672	*.517*
UF-R	.714	*.716*	.546	.605	**.508**	.973	*.929*	**.984**	.966	.971	.912	*.789*	**.966**	.918	.943
UF-L	*.775*	**.586**	.629	.696	.644	**.982**	*.812*	.961	.977	.971	.924	*.640*	.901	**.927**	.922
TP-R	.705	**.535**	.583	*.715*	.713	.983	*.509*	.985	.984	**.985**	.946	*.326*	**.964**	.945	.951
TP-L	*.737*	**.476**	.619	.730	.714	.982	*.491*	.980	**.984**	.982	.934	*.336*	**.949**	.942	.941

Table E.1: Region-based statistics for measurement: RTOP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	**.505**	.631	*.884*	.864	.852	*.501*	.956	**.998**	.996	.995	*.332*	.890	**.983**	.969	.966
PCT	.831	**.718**	.746	*.839*	.764	.988	*.882*	**.994**	.987	.987	.933	*.678*	**.978**	.923	.948
GCC	**.490**	.718	*.845*	.686	.694	*.453*	.945	**.997**	.994	.995	*.297*	.829	.979	.982	**.983**
BCC	.716	*.886*	.822	.744	**.633**	.995	*.975*	**.998**	.995	.995	.984	*.816*	**.990**	.983	.987
SCC	**.570**	.658	*.756*	.658	.665	*.844*	.969	**.998**	.997	.997	*.700*	.914	**.992**	.991	.990
FRX	**.342**	.464	*.528*	.390	.446	.994	*.865*	.983	.994	**.996**	.991	*.774*	.964	.991	**.992**
CST-R	.770	**.622**	.724	*.794*	.754	.980	*.881*	**.992**	.969	.968	.917	*.737*	**.972**	.867	.881
CST-L	**.436**	.730	.653	*.819*	.728	*.420*	.876	**.992**	.977	.978	*.290*	.657	**.976**	.885	.922
ML-R	**.393**	*.636*	.500	.460	.543	.961	*.859*	.966	**.969**	.968	.938	*.689*	.935	**.943**	.932
ML-L	.542	*.598*	.476	.535	**.458**	.977	*.931*	.978	**.980**	.974	.952	*.844*	**.959**	.958	.954
ICP-R	.530	*.587*	**.466**	.561	.535	.952	*.866*	**.975**	.957	.957	.904	*.727*	**.954**	.908	.911
ICP-L	.508	*.752*	.527	.566	**.452**	.960	*.939*	**.970**	.961	.945	.922	*.793*	**.939**	.915	.903
SCP-R	.470	.509	*.616*	.479	**.411**	.927	*.789*	**.977**	.917	.918	.871	*.647*	**.943**	.852	.868
SCP-L	.409	*.640*	.366	.404	**.357**	.969	*.865*	**.982**	.964	.951	.948	*.698*	**.972**	.941	.926
CP-R	.736	**.626**	*.793*	.667	.696	*.933*	.951	**.996**	.960	.965	*.787*	.879	**.980**	.889	.894
CP-L	.520	**.508**	*.685*	.582	.612	*.522*	.946	**.997**	.970	.971	*.344*	.896	**.989**	.931	.929
ALIC-R	*.704*	**.603**	.654	.666	.670	.985	*.897*	**.994**	.981	.983	.950	*.775*	**.983**	.946	.952
ALIC-L	*.747*	.619	.643	.720	**.605**	.963	*.943*	**.991**	.981	.982	.868	*.863*	**.974**	.937	.955
PLIC-R	**.396**	.626	.732	*.747*	.675	*.491*	.952	**.997**	.984	.979	*.368*	.882	**.988**	.940	.939
PLIC-L	.758	**.541**	.629	*.810*	.755	.977	*.943*	**.996**	.983	.965	.912	.884	**.989**	.916	*.872*
RPIC-R	.717	**.524**	*.760*	.652	.604	.992	*.880*	.989	.993	**.994**	.972	*.778*	.955	.981	**.985**
RPIC-L	*.856*	**.635**	.672	.831	.761	**.993**	*.939*	.993	.992	.991	.952	*.848*	**.978**	.952	.963
ACR-R	.879	**.728**	*.885*	.859	.804	.997	*.944*	**.998**	.997	.996	.979	*.822*	.981	**.982**	.981
ACR-L	.904	**.708**	.907	*.913*	.861	.998	*.933*	.998	**.999**	.997	**.984**	*.802*	.980	.983	.982
SCR-R	.833	**.709**	*.861*	.855	.793	**.998**	*.917*	.998	.998	.998	.990	*.762*	.987	.988	**.991**
SCR-L	.861	**.678**	*.886*	.848	.855	.998	*.953*	.997	**.998**	.997	.984	*.867*	.972	**.985**	.982
PCR-R	*.781*	**.478**	.698	.775	.705	.987	*.640*	.987	**.990**	.988	.942	*.482*	.957	.957	**.960**
PCR-L	.765	**.625**	.701	*.773*	.692	**.996**	*.875*	.991	.995	.995	.981	*.725*	.971	.980	**.984**
PTR-R	.651	**.524**	*.636*	*.656*	.580	.992	*.846*	**.995**	.991	.990	.978	*.724*	**.987**	.974	.977
PTR-L	*.754*	**.570**	.731	.737	.710	.995	*.828*	**.997**	.995	.994	.981	*.674*	**.988**	.981	.981
SS-R	*.782*	**.535**	.718	.742	.677	.989	*.892*	**.992**	.988	.976	.950	*.793*	**.971**	.954	.930
SS-L	*.798*	**.600**	.735	.779	.652	.986	*.924*	**.986**	.985	.974	.933	*.829*	**.949**	.937	.928
EC-R	.732	.634	.651	*.783*	**.573**	.992	*.890*	.993	**.994**	.989	.970	*.746*	**.979**	.971	.974
EC-L	.760	**.586**	.692	*.794*	.670	.990	*.806*	.964	**.994**	.988	.959	*.632*	.893	**.972**	.965
CGG-R	*.857*	.682	**.575**	.833	.711	.993	*.949*	**.994**	.994	.993	.951	*.856*	**.986**	.965	.975
CGG-L	*.888*	**.646**	.669	.864	.805	.993	*.964*	**.994**	.993	.992	.938	*.904*	**.983**	.951	.961
CGH-R	.605	**.567**	.593	*.610*	.595	.967	*.876*	**.982**	.964	.968	.921	*.753*	**.956**	.914	.925
CGH-L	*.708*	**.511**	.596	.705	.632	.988	*.940*	**.991**	.986	.984	.960	*.885*	**.978**	.954	.957
FX/ST-R	*.637*	.591	**.528**	.577	.531	.981	*.886*	.959	**.985**	.984	.950	*.761*	.916	.965	**.966**
FX/ST-L	.665	.625	.585	*.667*	**.515**	.975	*.925*	.965	.973	**.977**	.929	*.822*	.920	.924	**.954**
SLF-R	.815	**.643**	*.835*	.749	.718	.998	*.934*	**.999**	.998	.998	.991	*.834*	.992	**.993**	.993
SLF-L	*.899*	**.746**	.866	.853	.776	.997	*.958*	**.998**	.997	.997	.975	*.853*	.985	.982	**.985**
SFOF-R	.708	**.461**	*.726*	.704	.696	.910	*.731*	**.978**	.913	.911	.747	*.593*	**.925**	.758	.758
SFOF-L	*.750*	**.580**	.669	.731	.724	.903	.888	**.943**	.930	*.800*	.700	.769	**.846**	.781	*.525*
UF-R	.493	*.670*	.488	.515	**.388**	.965	*.911*	**.973**	.960	.963	.933	*.773*	**.949**	.922	.941
UF-L	.632	**.555**	.584	*.658*	.566	.981	*.765*	.965	**.981**	.975	**.950**	*.592*	.920	.947	.943
TP-R	.670	**.544**	.648	*.686*	.655	.972	*.518*	**.980**	.977	.977	.921	*.329*	**.945**	.930	.936
TP-L	.723	**.469**	.656	*.728*	.656	**.988**	*.470*	.985	.986	.985	.958	*.320*	.958	.949	**.958**

Table E.2: Region-based statistics for measurement: RTAP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	.940	**.601**	.776	*.963*	.900	**.995**	*.943*	.992	.994	.982	.929	.869	**.964**	.857	*.846*
PCT	.812	**.639**	.759	*.816*	.780	.983	*.907*	**.985**	.936	.915	.915	.780	**.939**	.729	*.702*
GCC	**.597**	.727	*.936*	.807	.759	.990	*.973*	**.998**	.996	.992	.975	*.909*	.975	**.980**	.969
BCC	.870	.785	*.879*	.869	**.710**	*.971*	.992	**.999**	.998	.994	*.815*	.966	**.994**	.985	.980
SCC	.809	**.528**	.758	*.842*	.778	.992	*.857*	**.999**	.994	.996	.961	*.739*	**.995**	.964	.980
FRX	.654	**.571**	.689	.670	*.707*	.999	*.984*	.994	.997	**.999**	**.996**	*.963*	.981	.991	.996
CST-R	.776	.559	.687	*.838*	.706	.965	*.864*	**.988**	.965	.952	.859	*.737*	**.963**	.816	.854
CST-L	.696	.547	.635	*.832*	.592	.971	*.853*	**.987**	.967	.956	.912	*.725*	**.965**	.832	.899
ML-R	.652	.627	**.487**	*.702*	.566	.953	*.863*	**.969**	.909	.962	.875	*.701*	**.941**	.749	.916
ML-L	.724	.576	**.443**	*.792*	.613	.929	.880	**.966**	*.875*	.928	.783	.756	**.941**	*.593*	.833
ICP-R	.720	.620	.643	*.791*	**.617**	.886	.851	**.935**	.863	*.317*	.685	.685	**.837**	.569	*.151*
ICP-L	*.800*	**.599**	.627	.774	.714	.959	*.889*	**.959**	.955	.953	.823	*.762*	**.897**	.828	.852
SCP-R	.581	.576	*.668*	.637	**.574**	.931	*.867*	**.977**	.887	.922	.850	*.734*	**.933**	.740	.834
SCP-L	.730	**.585**	.586	*.782*	.697	.963	.907	**.976**	*.801*	.938	.875	.801	**.944**	*.466*	.821
CP-R	.753	**.584**	*.817*	.676	.767	.945	.976	**.997**	*.944*	.986	*.808*	.943	**.986**	.845	.941
CP-L	*.837*	**.584**	.769	.826	.779	*.931*	.976	**.998**	.955	.990	*.687*	.944	**.991**	.786	.954
ALIC-R	.709	**.602**	.805	*.823*	.800	.982	*.958*	**.996**	.966	.979	.941	.901	**.978**	*.833*	.905
ALIC-L	**.503**	.568	.713	*.785*	.722	*.953*	.957	**.995**	.978	.980	.910	.907	**.981**	*.906*	.931
PLIC-R	*.846*	**.618**	.765	.760	.723	.984	*.968*	**.997**	.973	.982	.907	.920	**.987**	*.897*	.939
PLIC-L	*.878*	**.631**	.702	.787	.753	.992	*.966*	**.996**	.985	.986	.936	*.912*	**.987**	.934	.946
RPIC-R	**.575**	.644	*.724*	.711	.685	**.990**	*.944*	.974	.963	.977	**.978**	*.858*	.911	.884	.931
RPIC-L	.816	**.651**	.673	*.836*	.678	**.987**	*.949*	.982	.953	.959	.935	.867	**.946**	*.770*	.882
ACR-R	.832	**.726**	.822	*.885*	.832	.997	*.989*	.998	**.998**	.998	.985	*.962*	.989	.986	**.989**
ACR-L	.785	**.776**	.828	*.909*	.850	.995	*.986*	.998	**.998**	.998	.979	*.939*	**.987**	.977	.984
SCR-R	*.911*	.791	**.718**	.905	.795	.997	*.987*	**.998**	.996	.998	.969	*.940*	**.992**	.956	.989
SCR-L	.942	.844	.839	*.950*	**.836**	.998	*.989*	.998	.997	**.998**	.961	*.932*	.987	.939	**.989**
PCR-R	.734	.688	.644	*.823*	**.598**	.992	*.958*	.977	**.994**	.992	.972	*.876*	.938	.967	**.981**
PCR-L	.684	.678	**.442**	*.828*	.648	.992	*.962*	.991	**.995**	.994	.974	*.891*	**.984**	.969	.982
PTR-R	.754	.706	**.576**	*.887*	.733	.993	*.976*	.988	**.994**	.992	**.974**	*.922*	.972	.950	.972
PTR-L	.802	.682	**.565**	*.880*	.684	.990	.959	*.903*	**.994**	.985	.950	.880	*.802*	.954	**.955**
SS-R	.739	**.608**	.650	*.842*	.724	.990	*.957*	.990	.989	**.992**	.961	*.897*	**.971**	.932	.970
SS-L	.793	.665	**.625**	*.898*	.814	.989	*.955*	**.989**	.986	.988	.948	.876	**.971**	*.875*	.939
EC-R	.771	**.644**	*.791*	.739	.723	.974	.975	**.996**	*.971*	.992	*.896*	.933	**.982**	.898	.973
EC-L	.822	**.665**	.789	*.867*	.804	.990	*.977*	**.997**	.985	.994	.945	.934	**.986**	*.898*	.972
CGG-R	.693	.602	.694	*.811*	**.477**	.983	*.946*	**.992**	.989	.982	.947	*.875*	**.975**	.944	.967
CGG-L	.821	.663	**.594**	*.868*	.722	.974	*.966*	**.995**	.985	.981	*.869*	.906	**.987**	.894	.935
CGH-R	.763	**.567**	.724	*.776*	.753	.992	*.960*	**.993**	.987	.991	.966	*.913*	**.976**	.946	.963
CGH-L	.745	**.570**	.672	*.809*	.662	**.991**	*.956*	.990	.985	.988	.964	*.904*	**.970**	.924	.964
FX/ST-R	.599	**.520**	.615	*.681*	.579	.976	*.918*	**.985**	.963	.976	.942	*.843*	**.962**	.893	.945
FX/ST-L	**.514**	.551	.546	*.782*	.633	.935	*.925*	**.982**	.958	.964	.875	.846	**.961**	*.832*	.908
SLF-R	**.543**	.666	.712	*.860*	.641	**.996**	*.973*	.988	.992	.988	**.991**	*.925*	.959	.948	.966
SLF-L	.770	**.702**	.715	*.930*	.802	.996	*.980*	.996	.996	**.997**	.982	*.936*	**.986**	.951	.986
SFOF-R	.579	.564	**.553**	*.665*	.555	.974	*.907*	**.986**	.971	.959	.940	*.810*	**.969**	.919	.913
SFOF-L	*.804*	**.611**	.658	.735	.771	.986	*.935*	**.992**	.982	.989	.933	*.848*	**.976**	.936	.955
UF-R	.793	**.638**	.744	*.806*	.763	.967	*.911*	**.985**	.954	.973	.860	*.788*	**.944**	.801	.895
UF-L	.737	**.642**	.743	*.821*	.760	.980	*.891*	**.987**	.957	.984	.928	*.746*	**.951**	.800	.938
TP-R	.597	.556	**.515**	*.738*	.563	.979	*.917*	**.983**	.979	.982	.949	*.831*	**.965**	.924	.959
TP-L	.595	.521	**.447**	*.714*	.662	**.984**	*.903*	.952	.976	.983	**.961**	*.817*	.917	.919	.951

Table E.3: Region-based statistics for measurement: RTPP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	—	**.896**	.954	*.959*	.916	—	*.988*	**.999**	.997	.994	—	*.897*	**.969**	.933	.937
PCT	—	**.803**	.857	*.875*	.839	—	*.966*	**.997**	.993	.984	—	*.847*	**.977**	.946	.909
GCC	—	.826	.929	*.940*	**.536**	—	.980	**.998**	.998	*.517*	—	.893	**.977**	.961	*.332*
BCC	—	.896	.910	*.961*	**.474**	—	.992	**.999**	.996	*.466*	—	.925	**.989**	.911	*.314*
SCC	—	.819	.796	*.901*	**.419**	—	.988	**.998**	.992	*.649*	—	.937	**.988**	.927	*.518*
FRX	—	.591	.521	*.698*	**.488**	—	.978	.995	**.999**	*.540*	—	.948	.990	**.998**	*.375*
CST-R	—	.754	.715	*.800*	**.653**	—	.963	**.992**	.979	*.947*	—	.866	**.973**	.904	*.862*
CST-L	—	.807	.747	*.831*	**.436**	—	.960	**.992**	.988	*.420*	—	.821	**.967**	.932	*.290*
ML-R	—	*.666*	.624	.642	**.439**	—	.917	**.978**	.930	*.865*	—	.788	**.943**	.827	*.783*
ML-L	—	*.677*	.650	.669	**.623**	—	.934	**.982**	.968	*.900*	—	.821	**.951**	.910	*.773*
ICP-R	—	*.733*	**.630**	.662	.634	—	*.938*	**.983**	.970	.944	—	*.802*	**.954**	.916	.859
ICP-L	—	*.766*	.626	.646	**.528**	—	.961	**.980**	.970	*.958*	—	*.852*	**.947**	.919	.915
SCP-R	—	.603	*.644*	.555	**.520**	—	.885	**.977**	.851	*.757*	—	.754	**.938**	.717	*.600*
SCP-L	—	*.611*	**.441**	.583	.471	—	.922	**.980**	.896	*.867*	—	.822	**.965**	.783	*.776*
CP-R	—	.747	*.799*	.752	**.495**	—	.980	**.996**	.970	*.634*	—	.927	**.980**	.890	*.466*
CP-L	—	.708	*.731*	.723	**.447**	—	.979	**.997**	.978	*.523*	—	.932	**.988**	.924	*.377*
ALIC-R	—	*.764*	.761	.759	**.735**	—	*.972*	**.996**	.990	.984	—	*.891*	**.984**	.960	.943
ALIC-L	—	.765	**.685**	*.807*	.768	—	*.970*	**.996**	.994	.991	—	*.885*	**.987**	.971	.964
PLIC-R	—	.767	.806	*.813*	**.424**	—	.983	**.997**	.975	*.519*	—	.930	**.986**	.881	*.383*
PLIC-L	—	.767	**.722**	*.839*	.813	—	.982	**.997**	.992	*.970*	—	.929	**.989**	.950	*.860*
RPIC-R	—	.773	*.805*	.749	**.738**	—	*.968*	**.985**	.980	.971	—	*.871*	**.930**	.925	.899
RPIC-L	—	.770	**.663**	*.859*	.836	—	.980	**.986**	*.957*	.967	—	.917	**.961**	*.756*	.826
ACR-R	—	.937	**.933**	*.958*	.945	—	*.979*	**.999**	.998	.997	—	*.747*	**.980**	.950	.948
ACR-L	—	**.944**	.945	*.960*	.952	—	*.986*	**.999**	.998	.998	—	*.799*	**.976**	.959	.953
SCR-R	—	**.935**	.938	*.971*	.949	—	.992	**.998**	.995	*.990*	—	.885	**.965**	.862	*.840*
SCR-L	—	.934	**.932**	*.966*	.960	—	.995	**.997**	.995	*.985*	—	.923	**.957**	.867	*.719*
PCR-R	—	.841	**.762**	.843	*.871*	—	*.981*	**.992**	.987	.986	—	*.893*	**.966**	.924	.900
PCR-L	—	.837	**.702**	*.891*	.878	—	*.982*	**.993**	.986	.986	—	.898	**.976**	*.884*	.895
PTR-R	—	**.769**	.782	*.897*	.851	—	*.957*	.997	**.997**	.994	—	*.837*	**.986**	.973	.959
PTR-L	—	**.777**	.788	*.928*	.895	—	*.957*	**.996**	.995	.993	—	*.832*	**.982**	.939	.935
SS-R	—	.788	**.749**	*.820*	.789	—	*.965*	**.996**	.990	.982	—	*.855*	**.984**	.949	.922
SS-L	—	**.767**	.836	*.910*	.883	—	*.973*	**.995**	.990	.980	—	.895	**.967**	.897	*.853*
EC-R	—	*.921*	**.711**	.716	.747	—	.990	**.996**	.992	*.987*	—	*.884*	**.986**	.972	.952
EC-L	—	*.900*	.698	**.690**	.706	—	.991	**.993**	.993	*.985*	—	*.915*	**.977**	.976	.951
CGG-R	—	.778	**.764**	*.867*	.861	—	.979	**.993**	.978	*.975*	—	.911	**.969**	.857	*.841*
CGG-L	—	**.759**	.785	*.903*	.876	—	.984	**.995**	.977	*.974*	—	.935	**.979**	*.801*	.820
CGH-R	—	*.757*	**.677**	.736	.721	—	*.973*	**.993**	.990	.987	—	*.896*	**.977**	.962	.954
CGH-L	—	*.714*	.620	.652	.688	—	*.975*	**.993**	.986	.983	—	*.918*	**.981**	.961	.948
FX/ST-R	—	*.733*	**.579**	.604	.682	—	.943	**.965**	.949	*.932*	—	.814	**.920**	.881	*.813*
FX/ST-L	—	*.740*	**.701**	.737	.736	—	.962	**.978**	.944	*.910*	—	.868	**.930**	.815	*.729*
SLF-R	—	**.878**	.888	*.911*	.879	—	*.986*	**.997**	.991	.996	—	*.899*	**.978**	.908	.965
SLF-L	—	**.864**	.940	*.962*	.958	—	*.989*	**.997**	.995	.996	—	.926	**.958**	*.883*	.912
SFOF-R	—	**.706**	.706	*.799*	.792	—	.948	**.982**	.954	*.918*	—	.844	**.941**	.808	*.700*
SFOF-L	—	.738	**.661**	*.806*	.758	—	.961	**.976**	.950	*.922*	—	.866	**.933**	.786	*.740*
UF-R	—	*.644*	**.571**	.631	.609	—	*.902*	**.984**	.942	.931	—	*.765*	**.964**	.856	.842
UF-L	—	.662	**.655**	*.666*	.656	—	.903	**.950**	*.783*	.798	—	.758	**.868**	*.546*	.577
TP-R	—	**.478**	.538	.582	*.659*	—	*.808*	**.981**	.944	.970	—	*.688*	**.959**	.875	.918
TP-L	—	**.592**	.624	*.686*	.631	—	*.877*	**.976**	.973	.965	—	*.744*	**.939**	.920	.910

Table E.4: Region-based statistics for measurement: NG. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	*.988*	**.947**	.985	.979	.976	.997	*.993*	**.998**	.998	.996	*.818*	.876	.873	**.897**	.863
PCT	*.937*	**.868**	.915	.915	.920	*.974*	.985	**.989**	.980	.986	*.703*	**.895**	.887	.808	.854
GCC	.936	.911	**.898**	.934	*.945*	.999	*.995*	.997	.998	**.999**	**.984**	*.950*	.968	.971	.981
BCC	.979	.956	**.943**	*.981*	.957	**.999**	*.997*	.998	.998	.999	.947	.946	.962	*.914*	**.973**
SCC	.941	**.872**	.884	.931	*.947*	.997	*.992*	**.998**	.995	.996	.959	.944	**.980**	.937	*.931*
FRX	*.714*	.649	.551	**.513**	.649	**.000**	*.992*	.999	.995	.999	**.999**	*.979*	.998	.990	.998
CST-R	*.933*	**.845**	.907	.881	.926	**.973**	.971	.973	*.937*	.946	.709	**.837**	.773	.641	*.565*
CST-L	.923	**.848**	*.930*	.909	.918	*.926*	.971	.982	.976	**.983**	*.491*	**.836**	.793	.786	.821
ML-R	*.837*	.750	.761	**.739**	.766	.950	.939	**.971**	*.914*	.945	.756	.793	**.888**	*.735*	.800
ML-L	*.881*	**.769**	.844	.823	.831	.949	*.949*	**.985**	.970	.975	*.689*	.812	**.909**	.851	.868
ICP-R	*.905*	**.783**	.871	.844	.839	.962	*.947*	**.966**	.966	.961	*.709*	.794	.788	**.816**	.799
ICP-L	*.900*	**.775**	.878	.815	.856	.984	*.958*	**.986**	.978	.984	.861	*.836*	.894	.892	**.901**
SCP-R	*.745*	.694	.626	**.615**	.725	.915	**.944**	.896	*.886*	.940	*.733*	**.837**	.764	.750	.811
SCP-L	*.824*	**.696**	.807	.770	.790	*.859*	**.950**	.947	.905	.886	*.517*	**.853**	.776	.686	.620
CP-R	*.904*	**.809**	.841	.890	.813	.985	.987	**.992**	.976	*.964*	.863	.938	**.952**	*.814*	.833
CP-L	*.906*	**.781**	.794	.841	.806	.990	.987	**.994**	*.974*	.979	.902	.945	**.969**	*.855*	.899
ALIC-R	*.917*	**.822**	.893	.843	.902	.992	*.985*	**.993**	.986	.988	.907	.921	**.935**	.915	*.888*
ALIC-L	*.899*	**.802**	.860	.860	.877	.993	*.987*	.992	.994	**.995**	*.933*	.936	.942	**.961**	.957
PLIC-R	*.933*	**.853**	.872	.896	.926	.991	.988	**.993**	*.961*	.988	.882	.925	**.947**	*.718*	.858
PLIC-L	*.955*	**.867**	.920	.933	.954	.993	*.991*	.994	.991	**.995**	*.865*	**.935**	.929	.880	.893
RPIC-R	.777	.767	.763	**.702**	*.806*	**.988**	*.979*	.985	.984	.986	**.948**	*.917*	.938	.947	.933
RPIC-L	.930	**.829**	.873	.928	*.930*	**.989**	*.978*	.985	.981	.983	.863	.883	**.893**	*.792*	.806
ACR-R	.958	**.942**	.947	*.961*	.961	**.000**	*.998*	.999	.999	.999	**.989**	*.960*	.979	.977	.987
ACR-L	.971	.953	**.946**	.974	*.977*	**.999**	*.997*	.999	.999	.999	**.981**	*.950*	.976	.965	.974
SCR-R	.960	.950	**.935**	*.974*	.968	**.999**	*.996*	.999	.998	.999	.975	.932	**.980**	*.925*	.964
SCR-L	.978	.961	**.945**	*.979*	.971	.999	*.996*	.998	.998	**.999**	.947	.913	**.966**	*.897*	.962
PCR-R	.892	.895	**.794**	*.916*	.876	**.997**	*.989*	.994	.995	.996	.972	*.908*	.970	.943	**.972**
PCR-L	.910	.898	**.827**	*.940*	.926	**.997**	*.989*	.994	.996	.997	**.971**	*.905*	.968	.943	.961
PTR-R	.930	**.859**	.932	*.940*	.935	.997	*.992*	.995	**.998**	.997	.960	.949	*.936*	**.962**	.953
PTR-L	.939	**.875**	.916	*.958*	.938	**.998**	*.993*	.997	.998	.998	**.971**	.945	.962	*.945*	.962
SS-R	*.910*	.848	**.828**	.884	.893	**.997**	*.989*	.990	.994	.994	**.964**	*.933*	.946	.948	.947
SS-L	*.957*	**.882**	.921	.948	.955	**.997**	*.989*	.990	.995	.994	**.936**	.912	*.883*	.904	.888
EC-R	*.925*	**.853**	.866	.858	.918	**.996**	.991	.995	*.989*	.995	.950	.941	**.962**	*.927*	.945
EC-L	*.942*	.885	**.884**	.887	.937	.994	*.991*	.991	.995	**.997**	*.911*	.928	.930	**.955**	.953
CGG-R	.940	.871	**.761**	*.943*	.905	.996	.989	.995	*.988*	**.997**	.933	.922	**.979**	*.831*	.964
CGG-L	.948	.887	**.740**	*.955*	.931	.994	*.987*	.994	.992	**.995**	.890	.898	**.976**	*.848*	.926
CGH-R	*.902*	.781	**.731**	.842	.873	**.993**	.979	*.979*	.990	.990	.933	*.911*	.926	**.942**	.926
CGH-L	*.921*	**.770**	.866	.811	.875	**.996**	*.984*	.990	.990	.994	.952	.936	*.927*	.948	**.954**
FX/ST-R	*.789*	.738	**.701**	.787	.717	**.983**	*.953*	.977	.974	.980	.924	*.843*	.926	.887	**.933**
FX/ST-L	.829	**.761**	.782	.838	*.850*	.947	.955	**.957**	*.932*	.934	.756	**.836**	.828	.690	*.679*
SLF-R	*.936*	.902	.918	.919	**.900**	**.998**	.994	.998	*.994*	.998	.976	.943	**.979**	*.929*	.978
SLF-L	.971	**.935**	.966	*.978*	.967	**.999**	*.996*	.999	.998	.999	**.973**	.947	.961	*.903*	.964
SFOF-R	.772	**.718**	.777	*.806*	.759	**.964**	.939	.959	*.926*	.951	**.861**	.812	.838	*.707*	.825
SFOF-L	*.880*	**.772**	.777	.869	.870	.983	*.961*	.986	.962	**.989**	.876	.849	**.941**	*.767*	.921
UF-R	*.917*	**.741**	.847	.787	.877	**.984**	.948	.976	*.922*	.970	.836	.825	**.863**	*.715*	.800
UF-L	*.927*	.816	**.769**	.849	.909	**.987**	.939	.967	*.906*	.979	.851	.738	**.872**	*.591*	.809
TP-R	.736	.684	.590	.723	*.746*	.983	*.950*	**.984**	.970	.975	.938	*.858*	**.962**	.898	.907
TP-L	.783	**.725**	.750	.777	*.801*	**.976**	.949	.976	*.908*	.975	.898	.835	**.910**	*.688*	.884

Table E.5: Region-based statistics for measurement: MSD. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	*.954*	.852	**.676**	.953	.792	.996	.988	.984	**.996**	*.982*	.924	.923	**.953**	.926	*.920*
PCT	*.898*	.817	.863	.896	**.764**	.989	*.980*	**.997**	.989	.992	.902	*.899*	**.979**	.900	.967
GCC	.812	*.878*	.864	.817	**.790**	.996	*.992*	**.996**	.996	.995	**.977**	*.938*	.973	.977	.975
BCC	.893	*.938*	.832	.896	**.733**	.997	*.994*	**.997**	.997	.995	.970	*.918*	**.984**	.969	.983
SCC	.722	*.850*	**.717**	.726	.717	.997	*.989*	.995	**.997**	.996	**.989**	*.930*	.982	.989	.986
FRX	.741	**.683**	.738	*.742*	.737	.000	*.994*	.000	**.000**	.000	.000	*.980*	.999	**.000**	.999
CST-R	*.876*	.743	.633	**.617**	.800	**.983**	.958	.939	*.831*	.978	.874	.855	.850	*.654*	**.898**
CST-L	.873	.752	**.647**	*.879*	.738	.973	*.958*	.965	**.974**	.973	.822	.848	**.907**	*.817*	.906
ML-R	.536	*.621*	.618	**.528**	.560	.954	*.933*	**.974**	.954	.955	.906	*.840*	**.934**	.908	.904
ML-L	.689	.684	.631	*.689*	**.507**	.976	*.952*	**.981**	.977	.962	.927	*.861*	**.950**	.929	.925
ICP-R	.747	.700	.568	*.751*	**.558**	.963	*.941*	**.975**	.963	.960	.869	*.827*	**.945**	.866	.914
ICP-L	.648	.647	.563	*.662*	**.523**	.974	*.945*	**.976**	.974	.961	.930	*.859*	**.946**	.928	.921
SCP-R	.646	**.625**	*.669*	.636	.641	.907	*.870*	**.948**	.912	.921	.776	*.715*	**.857**	.790	.808
SCP-L	.566	*.644*	.484	.554	**.479**	.956	*.903*	.956	**.958**	.949	.905	*.767*	**.918**	.911	.907
CP-R	.798	.739	.744	*.803*	**.726**	.978	*.972*	**.987**	.980	.975	*.900*	.902	**.951**	.906	.916
CP-L	.800	**.697**	.718	*.802*	.743	.987	*.975*	**.990**	.989	.981	.940	*.922*	**.967**	.945	.931
ALIC-R	.783	.766	.750	*.787*	**.732**	.988	*.974*	**.993**	.987	.989	.945	*.899*	**.973**	.944	.961
ALIC-L	.819	.764	**.597**	*.823*	.618	.987	*.983*	**.993**	.987	.986	.931	.930	**.982**	*.929*	.964
PLIC-R	.793	.761	.777	*.795*	**.701**	.988	*.977*	**.995**	.988	.984	.943	*.910*	**.979**	.943	.948
PLIC-L	.841	.762	**.749**	*.846*	.795	.989	.982	**.994**	.989	*.977*	.936	.929	**.978**	.935	*.899*
RPIC-R	.698	.755	*.796*	.697	**.647**	.992	*.975*	.990	.992	**.995**	.975	*.905*	.951	.974	**.985**
RPIC-L	.901	.814	.752	*.903*	**.751**	**.992**	*.972*	.986	.992	.989	.922	*.865*	.947	.921	**.957**
ACR-R	.933	.933	.909	*.935*	**.885**	.999	*.996*	.999	**.999**	.998	**.988**	*.947*	.986	.987	.983
ACR-L	.953	.941	.919	*.954*	**.908**	.999	*.997*	.998	**.999**	.998	**.984**	*.948*	.980	.984	.979
SCR-R	.903	*.930*	.886	.905	**.860**	**.999**	*.995*	.998	.999	.998	.987	*.933*	.983	.987	**.988**
SCR-L	.937	*.939*	.904	.939	**.895**	**.998**	*.994*	.997	.998	.998	.976	*.915*	.970	.975	**.980**
PCR-R	.843	*.852*	.739	.846	**.729**	**.994**	*.981*	.992	.994	.993	.962	*.886*	.971	.961	**.976**
PCR-L	.861	*.874*	**.720**	.863	.752	.997	*.986*	.994	**.997**	.996	.977	*.900*	.980	.977	**.983**
PTR-R	.836	.818	.806	*.837*	**.773**	.995	*.989*	**.996**	.995	.994	.968	*.944*	**.981**	.968	.974
PTR-L	.856	.849	.806	*.859*	**.792**	.997	*.990*	**.998**	.997	.997	.982	*.936*	**.989**	.982	.985
SS-R	.852	.821	**.761**	*.854*	.767	**.994**	*.984*	.994	.994	.989	.960	*.917*	**.975**	.959	.954
SS-L	.898	.851	.790	*.901*	**.775**	**.993**	*.980*	.990	.993	.982	.937	*.880*	**.952**	.935	.924
EC-R	.847	.828	.678	*.852*	**.591**	.995	*.989*	**.996**	.995	.993	.970	*.942*	**.986**	.969	.982
EC-L	.874	.852	**.671**	*.877*	.699	**.993**	.990	.989	.993	*.979*	.945	.938	**.967**	.943	*.932*
CGG-R	.929	.880	**.691**	*.932*	.753	**.995**	*.984*	.987	.995	.991	.929	*.884*	.961	.926	**.963**
CGG-L	.931	.882	**.777**	*.934*	.813	.995	*.982*	.991	**.995**	.990	.930	*.865*	**.963**	.927	.948
CGH-R	.787	.737	.609	*.791*	.655	.971	*.968*	**.990**	.971	.973	.877	.889	**.974**	*.877*	.927
CGH-L	.844	.701	**.693**	*.845*	.716	**.993**	*.977*	.990	.993	.987	.959	*.927*	**.970**	.958	.957
FX/ST-R	.676	*.694*	.494	.685	**.463**	**.984**	.955	*.942*	.984	.974	.952	*.867*	.891	.950	**.952**
FX/ST-L	.768	.769	.565	*.773*	**.538**	**.979**	*.953*	.971	.978	.971	.915	*.823*	.936	.909	**.939**
SLF-R	.891	.889	.803	*.895*	**.795**	**.999**	*.992*	.997	.999	.998	.987	*.933*	.985	.986	**.991**
SLF-L	.948	.931	**.840**	*.951*	.867	.999	*.994*	.997	**.999**	.998	.975	*.914*	.979	.973	**.984**
SFOF-R	.733	.705	**.649**	*.737*	.697	.883	.913	**.964**	*.880*	.933	.669	.757	**.904**	*.659*	.808
SFOF-L	.809	.735	**.665**	*.813*	.712	.928	.923	**.933**	.928	*.450*	.710	.762	**.822**	.707	*.191*
UF-R	.701	.673	.545	*.706*	**.502**	.975	*.920*	**.981**	.975	.970	.922	*.791*	**.960**	.919	.941
UF-L	.782	.736	.674	*.786*	**.658**	.981	*.937*	.965	**.981**	.963	**.917**	*.796*	.899	.916	.898
TP-R	.776	**.647**	.758	.773	*.786*	.978	*.948*	.980	.977	**.982**	.907	*.866*	.923	.908	**.923**
TP-L	*.822*	**.665**	.817	.821	.816	.974	*.964*	.973	.974	**.975**	.870	**.901**	*.867*	.872	.879

Table E.6: Region-based statistics for measurement: QIV. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

F

Reliability – repeatability Tables: ZJU

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MiSFIT*	*MAPL*	*DSI*
MCP	.614	**.463**	.958	.932	*.962*	*.364*	.846	.995	**.998**	.995	*.181*	.747	.892	**.970**	.887
PCT	*.967*	**.920**	.951	.954	.948	**.993**	*.953*	.984	.987	.992	.834	*.621*	.757	.775	**.859**
GCC	.867	**.769**	*.964*	.872	.944	*.925*	.954	.993	**.996**	.996	*.623*	.827	.837	**.970**	.927
BCC	**.503**	.805	.945	*.957*	.801	*.494*	.509	**.990**	.965	.978	*.327*	*.168*	.844	.539	**.898**
SCC	.708	**.516**	.950	*.955*	.923	.917	*.797*	.956	.961	**.986**	.764	.656	*.519*	.524	**.845**
FRX	.607	**.540**	*.781*	.641	.681	.975	*.521*	.956	.980	**.984**	.938	*.333*	.825	.945	**.952**
CST-R	**.499**	.678	*.939*	.860	.866	*.494*	.655	.986	**.989**	.958	*.329*	.380	.814	**.927**	.755
CST-L	*.939*	**.399**	.933	.933	.874	**.985**	*.301*	.974	.961	.969	**.804**	*.205*	.719	.622	.799
ML-R	.772	**.354**	*.864*	.779	.847	.938	*.784*	**.986**	.929	.891	.774	.702	**.905**	.742	*.556*
ML-L	.856	**.599**	*.869*	.854	.832	.931	*.848*	**.988**	.960	.923	*.662*	.690	**.917**	.779	.668
ICP-R	.703	**.510**	*.920*	.686	.784	.889	*.511*	**.974**	.907	.795	.704	*.339*	.747	**.753**	.456
ICP-L	.874	**.535**	*.934*	.830	.899	**.957**	.771	.928	.948	*.717*	.737	.610	.459	**.756**	*.204*
SCP-R	*.823*	**.485**	.797	.808	.744	.736	*.348*	.742	**.863**	.691	.331	*.215*	.369	**.548**	.364
SCP-L	.808	**.390**	*.856*	.661	.785	.949	*.479*	**.977**	.854	.897	.780	*.359*	**.859**	.665	.652
CP-R	.404	**.400**	*.774*	.715	.628	*.586*	.899	**.994**	.990	.991	*.458*	.843	.976	.965	**.977**
CP-L	**.387**	.804	*.888*	.782	.853	*.444*	.760	**.993**	.984	.992	*.329*	.383	.942	.930	**.945**
ALIC-R	*.926*	.878	.898	**.875**	.898	.968	.954	.971	**.975**	*.950*	.691	.717	.774	**.831**	*.660*
ALIC-L	.864	*.884*	.841	**.775**	.795	**.994**	*.984*	.992	.993	.987	.960	*.879*	.952	**.970**	.941
PLIC-R	*.961*	.932	.954	.960	**.906**	.994	*.984*	**.996**	.994	.994	.867	*.809*	.924	.867	**.935**
PLIC-L	.919	**.439**	.921	*.947*	.775	.995	*.464*	**.997**	.995	.997	.938	*.327*	.963	.911	**.985**
RPIC-R	*.939*	.910	.921	.907	**.902**	.989	.990	.987	*.978*	**.991**	.844	.902	.860	*.809*	**.919**
RPIC-L	*.955*	.915	.932	.926	**.827**	**.996**	*.990*	.992	.995	.991	.911	*.890*	.897	.932	**.952**
ACR-R	*.986*	.967	.958	.958	**.732**	**.996**	.995	.993	*.989*	.990	*.767*	.874	.852	.785	**.965**
ACR-L	*.983*	.962	.973	.959	**.879**	**.996**	.990	.992	*.978*	.995	.827	.791	.759	*.651*	**.957**
SCR-R	*.992*	.980	.984	.972	**.927**	**.997**	.982	*.972*	.993	.992	.710	.514	*.351*	.790	**.899**
SCR-L	*.987*	.975	.974	.972	**.939**	**.997**	.993	*.975*	.995	.995	.820	.789	*.497*	.848	**.930**
PCR-R	.977	*.979*	.968	**.961**	.973	.996	.992	*.986*	**.996**	.996	.852	.710	*.686*	**.912**	.862
PCR-L	.956	.962	*.974*	.899	**.797**	.997	*.992*	.993	**.997**	.996	.933	.829	*.784*	.972	**.983**
PTR-R	*.962*	.931	.932	.938	**.891**	**.992**	.976	.970	.960	*.928*	**.824**	.733	.688	.602	*.585*
PTR-L	*.947*	**.542**	.922	.883	.740	**.998**	*.542*	.998	.998	.992	.971	*.351*	.976	**.982**	.970
SS-R	*.949*	.905	**.868**	.899	.919	**.978**	*.923*	.972	.949	.958	.692	*.530*	**.819**	.653	.647
SS-L	.897	*.918*	.789	.768	**.632**	**.987**	.979	.978	*.963*	.985	.884	*.796*	.905	.857	**.960**
EC-R	*.967*	.858	.857	.918	**.800**	.995	*.989*	.994	**.995**	.991	*.869*	.929	**.958**	.943	.957
EC-L	*.966*	.833	.761	.929	**.681**	**.997**	*.993*	.996	.997	.996	*.921*	.960	.985	.957	**.988**
CGG-R	.977	.947	*.979*	.963	**.918**	.994	*.991*	.991	**.996**	.993	.791	.858	*.710*	.910	**.919**
CGG-L	*.958*	.932	.944	.939	**.837**	.996	.991	*.989*	**.997**	.994	.923	.882	*.839*	.950	**.964**
CGH-R	.845	*.864*	.761	.861	**.686**	*.803*	**.969**	.954	.955	.912	*.387*	.808	**.832**	.748	.766
CGH-L	*.818*	.809	**.241**	*.579*	.783	**.991**	*.980*	.985	.991	.986	.953	*.905*	**.980**	.978	.937
FX/ST-R	*.921*	**.805**	.848	.896	.848	.917	.942	.956	**.973**	*.914*	*.467*	.760	.769	**.790**	.616
FX/ST-L	*.938*	**.630**	.650	.831	.736	.996	*.618*	.994	**.997**	.995	.944	*.374*	**.983**	.981	.980
SLF-R	*.989*	.968	.971	.982	**.897**	.997	.996	.997	**.998**	*.996*	*.800*	.891	.911	.894	**.958**
SLF-L	*.990*	**.933**	.946	.983	.951	.979	**.986**	*.758*	.928	.982	.326	**.827**	*.144*	.179	.728
SFOF-R	*.881*	.843	.857	.829	**.366**	*.034*	**.936**	.930	.645	.863	*.004*	.696	.654	.237	**.799**
SFOF-L	**.830**	*.930*	.927	.894	.870	**.965**	.964	.927	.882	*.804*	**.825**	.655	.479	.441	*.347*
UF-R	**.489**	*.803*	.672	.573	.708	**.963**	.931	*.803*	.953	.889	**.930**	.727	*.572*	.897	.699
UF-L	.840	.872	**.693**	*.875*	.869	.982	*.958*	.968	**.983**	.959	.897	*.745*	**.903**	.877	.754
TP-R	.769	.789	.803	*.813*	**.697**	**.989**	*.956*	.977	.988	.972	**.956**	*.821*	.894	.939	.912
TP-L	*.910*	.820	**.625**	.824	.713	**.987**	.968	.972	.975	*.946*	.873	.847	**.929**	.871	*.833*

Table F.1: Region-based statistics for measurement: RTOP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	.612	**.511**	*.912*	.862	.862	*.392*	.802	.994	**.996**	.996	*.200*	.665	.933	**.974**	.969
PCT	*.930*	.824	.910	**.760**	.824	**.991**	.955	*.697*	.990	.958	.889	.789	*.171*	**.958**	.800
GCC	.860	**.770**	*.946*	.835	.875	*.923*	.956	.995	.996	**.997**	*.627*	.832	.913	**.976**	.973
BCC	**.503**	.801	.903	*.956*	.938	.494	*.487*	**.988**	.813	.935	.327	*.159*	**.888**	.161	.474
SCC	.714	**.550**	.928	*.951*	.901	.917	*.816*	.829	.984	**.987**	.759	.666	*.258*	.748	**.884**
FRX	.518	.495	*.734*	**.385**	.637	.971	*.543*	.964	.962	**.979**	.941	*.375*	.876	.940	**.945**
CST-R	**.499**	.719	*.885*	.637	.774	*.494*	.636	.974	**.990**	.967	*.329*	.329	.813	**.973**	.867
CST-L	*.890*	**.432**	.881	.635	.768	**.986**	*.139*	.928	.970	.981	.885	*.084*	.607	.921	**.922**
ML-R	.741	**.350**	*.803*	.685	.657	.680	.684	**.978**	*.593*	.778	.354	.584	**.896**	*.315*	.546
ML-L	.814	**.577**	*.841*	.608	.676	.830	.848	**.985**	*.793*	.856	*.475*	.702	**.912**	.600	.659
ICP-R	**.491**	.510	*.847*	.545	.540	.955	*.512*	.967	.958	**.967**	.916	*.339*	.816	.912	**.931**
ICP-L	.858	**.484**	*.897*	.797	.745	.968	*.754*	.839	**.976**	.962	.811	.613	*.351*	**.893**	.865
SCP-R	.748	**.498**	.740	*.750*	.666	.807	.345	*.159*	**.907**	.859	.512	.209	*.047*	**.708**	.670
SCP-L	.714	.414	.777	**.404**	.559	.903	.483	**.941**	*.385*	.815	.727	.354	**.780**	*.272*	.660
CP-R	.399	**.374**	.433	*.741*	.712	*.590*	.888	**.992**	.969	.982	*.463*	.832	**.986**	.892	.941
CP-L	**.380**	.792	*.844*	.797	.645	*.447*	.754	**.987**	.849	.950	*.333*	.388	**.925**	.533	.870
ALIC-R	.816	.832	*.853*	.648	**.592**	*.740*	.950	**.956**	.889	.882	*.343*	**.762**	.760	.737	.752
ALIC-L	.851	*.909*	**.790**	.863	.846	**.993**	*.980*	.993	.991	.987	.957	*.815*	**.967**	.938	.918
PLIC-R	.945	**.881**	*.946*	.908	.892	.971	.957	**.995**	*.198*	.945	.648	.725	**.915**	*.022*	.651
PLIC-L	.900	**.439**	*.913*	.907	.775	.984	*.464*	**.996**	.752	.988	.860	.327	**.958**	*.219*	.950
RPIC-R	.898	.908	*.932*	.917	**.891**	.924	.969	**.988**	*.743*	.941	.554	.744	**.848**	*.194*	.634
RPIC-L	.927	.885	*.929*	.856	**.692**	.994	*.986*	**.994**	.992	.989	.928	*.893*	.927	.949	**.964**
ACR-R	*.972*	.898	.946	.953	**.729**	.974	.983	.987	**.989**	*.769*	.518	**.856**	.803	.807	*.475*
ACR-L	*.968*	.915	.965	.954	**.762**	**.987**	.978	*.976*	.984	.981	.713	.787	*.584*	.742	**.926**
SCR-R	*.982*	.922	.979	.974	**.771**	**.995**	*.941*	.972	.986	.988	.771	.555	*.420*	.651	**.949**
SCR-L	*.977*	**.936**	.965	.969	.941	**.993**	.987	*.964*	.988	.986	.769	**.833**	*.483*	.722	.801
PCR-R	.951	.951	*.966*	**.875**	.923	.990	*.983*	.991	.991	**.992**	.836	*.739*	.798	**.933**	.903
PCR-L	.929	.889	*.966*	.904	**.856**	.995	*.987*	.994	.994	**.995**	.930	.891	*.856*	.942	**.969**
PTR-R	.919	**.876**	*.926*	.905	.889	**.985**	.972	.966	.959	*.927*	**.844**	.809	.675	.690	*.586*
PTR-L	.887	**.542**	*.902*	.836	.591	.998	*.541*	**.998**	.997	.993	.982	*.351*	**.985**	.983	.984
SS-R	*.900*	**.829**	.863	.840	.881	**.972**	*.928*	.969	.958	.942	.775	.687	**.812**	.783	*.658*
SS-L	.837	*.864*	**.685**	.742	.686	.982	.979	**.983**	*.868*	.977	.899	.864	**.947**	*.628*	.931
EC-R	*.954*	.870	.865	.939	**.802**	.993	*.984*	**.995**	.994	.990	*.874*	.890	**.962**	.907	.951
EC-L	*.951*	**.804**	.840	.933	.876	.995	*.990*	**.997**	.994	.994	*.903*	.951	**.980**	.916	.951
CGG-R	.963	**.879**	*.976*	.951	.880	.995	*.990*	.994	**.996**	.996	.885	.927	*.811*	.927	**.966**
CGG-L	.939	**.884**	.938	*.946*	.892	.996	*.988*	.993	**.997**	.995	.939	.909	*.900*	.943	**.960**
CGH-R	.689	*.879*	.737	.738	**.456**	*.964*	.964	.970	.966	**.977**	.893	*.765*	.896	.881	**.959**
CGH-L	.793	*.898*	**.459**	.834	.583	**.990**	*.975*	.989	.988	.984	.955	*.801*	**.980**	.931	.962
FX/ST-R	.893	**.787**	.859	*.895*	.874	.962	*.948*	.963	**.980**	.973	*.730*	.796	.787	**.834**	.818
FX/ST-L	*.902*	.644	**.625**	.703	.734	**.995**	*.632*	.995	.995	.995	.956	*.380*	**.988**	.984	.983
SLF-R	.984	**.940**	.964	*.985*	.959	.995	*.992*	**.998**	.996	.995	*.768*	.886	**.937**	.798	.890
SLF-L	*.986*	**.907**	.946	.983	.972	.978	**.983**	*.926*	.947	.973	.383	**.846**	.405	*.232*	.502
SFOF-R	.755	.739	*.818*	.767	**.729**	.405	.917	**.922**	*.166*	.741	.143	**.743**	.682	*.044*	.437
SFOF-L	**.870**	.907	*.931*	.892	.920	.929	**.956**	.898	.903	*.807*	.631	**.670**	.375	.500	*.250*
UF-R	**.486**	*.849*	.647	.566	.659	.967	.915	.885	**.968**	*.844*	**.938**	*.621*	.731	.929	.648
UF-L	.831	*.871*	**.714**	.857	.838	**.985**	*.959*	.979	.985	.979	.918	*.750*	**.931**	.904	.885
TP-R	.826	**.765**	.810	*.843*	.771	**.983**	*.928*	.980	.980	.954	**.910**	*.751*	.905	.884	.825
TP-L	*.920*	.766	**.685**	.850	.813	.968	.923	**.975**	.958	*.841*	.709	.736	**.926**	.774	*.497*

Table F.2: Region-based statistics for measurement: RTAP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	.964	**.925**	*.971*	.958	.952	.995	.996	**.998**	*.988*	.996	.882	.947	**.950**	*.782*	.930
PCT	*.972*	**.905**	.939	.959	.966	**.998**	.982	*.980*	.993	.998	**.941**	.839	*.755*	.852	.936
GCC	.967	**.804**	*.984*	.908	.963	**.995**	*.941*	.947	.975	.987	**.876**	.758	*.218*	.782	.742
BCC	.991	.989	*.994*	**.985**	.988	.995	.993	**.997**	*.987*	.992	.645	.628	**.710**	*.534*	.589
SCC	*.979*	.956	.977	**.939**	.971	**.998**	.992	.997	.988	*.986*	**.911**	.852	.884	.838	*.666*
FRX	.828	.794	*.853*	.812	**.786**	.978	.965	**.982**	.978	*.900*	.886	.850	.892	**.892**	*.660*
CST-R	.901	.849	.922	*.927*	**.840**	.978	*.926*	.972	.941	**.981**	.817	.655	.734	*.539*	**.892**
CST-L	*.948*	**.716**	.885	.887	.893	**.960**	.921	.923	*.695*	.960	.556	**.768**	.578	*.205*	.720
ML-R	.917	**.850**	.889	.870	*.931*	.987	.979	.991	*.964*	**.991**	.861	.875	**.923**	*.778*	.883
ML-L	**.768**	.854	*.888*	.881	.811	.957	.964	**.989**	*.917*	.970	.837	.797	**.911**	*.568*	.861
ICP-R	.821	.798	*.951*	.886	**.752**	.986	.968	**.988**	*.967*	.984	.927	.859	.801	*.771*	**.939**
ICP-L	.871	**.748**	*.952*	.877	.875	.975	.975	**.986**	*.952*	.983	.835	**.907**	.778	*.708*	.876
SCP-R	**.564**	.646	*.893*	.747	.629	.953	.952	**.970**	*.872*	.963	.899	.876	.774	*.633*	**.905**
SCP-L	.801	.798	*.869*	.792	**.742**	.977	.978	**.992**	*.893*	.970	.893	.901	**.941**	*.633*	.894
CP-R	*.940*	.832	.876	**.825**	.892	.994	.985	**.996**	*.984*	.993	*.906*	.919	**.966**	.917	.937
CP-L	*.936*	**.832**	.913	.904	.856	**.996**	*.989*	.996	.992	.996	.940	.939	.954	*.920*	**.971**
ALIC-R	*.978*	**.918**	.943	.972	.958	**.996**	*.985*	.985	.991	.996	.844	.844	.790	*.753*	**.904**
ALIC-L	**.803**	.908	.920	*.951*	.918	.988	.984	**.993**	*.963*	.988	**.943**	.850	.923	*.560*	.870
PLIC-R	.933	.932	*.970*	.947	**.850**	**.999**	.995	.998	*.993*	.998	.979	.930	.943	*.876*	**.988**
PLIC-L	.917	**.845**	.918	*.954*	.902	.997	*.994*	**.998**	.996	.997	.970	.963	**.978**	*.926*	.969
RPIC-R	.909	.835	*.912*	.893	**.726**	.996	.993	**.996**	*.987*	.995	.953	.960	.961	*.892*	**.984**
RPIC-L	.939	**.890**	*.950*	.927	.917	*.971*	.990	**.996**	.976	.982	*.674*	.916	**.917**	.748	.816
ACR-R	.951	.964	*.968*	**.946**	.955	**.999**	*.996*	.998	.998	.998	**.971**	*.896*	.954	.970	.965
ACR-L	.953	**.941**	*.976*	.959	.959	**.999**	*.997*	.999	.998	.999	**.978**	*.948*	.956	.957	.970
SCR-R	**.969**	.979	*.994*	.980	.982	.993	.987	**.994**	*.984*	.990	**.813**	.617	*.520*	.550	.651
SCR-L	.964	.975	*.987*	**.964**	.976	**.999**	.995	.998	*.992*	.998	**.962**	.824	.875	*.817*	.917
PCR-R	.967	.968	*.985*	**.947**	.974	.995	*.983*	**.996**	.991	.991	**.873**	*.657*	.780	.851	.750
PCR-L	.961	**.957**	*.980*	.973	.969	.993	.987	**.996**	.988	*.983*	**.854**	.767	.838	.690	*.647*
PTR-R	.935	.899	*.945*	.945	**.780**	.953	.971	**.981**	.961	*.952*	*.570*	.775	.736	.574	**.815**
PTR-L	.918	.882	*.935*	.807	**.794**	.959	.996	**.999**	.992	*.945*	*.659*	.967	**.980**	.960	.779
SS-R	*.913*	.879	.834	.891	**.744**	.942	.989	**.992**	*.911*	.952	.587	.918	**.956**	*.529*	.835
SS-L	.876	*.894*	.837	.891	**.825**	*.928*	.989	**.994**	.972	.971	*.614*	.908	**.962**	.789	.853
EC-R	.867	.834	**.766**	.834	*.870*	*.972*	.991	**.996**	.975	.981	*.821*	.947	**.982**	.865	.870
EC-L	*.891*	.859	.763	**.662**	.877	*.985*	.993	**.998**	.993	.991	*.875*	.956	**.993**	.979	.930
CGG-R	.971	.966	*.989*	**.953**	.977	.993	.977	**.995**	*.838*	.989	**.814**	.585	.662	*.197*	.685
CGG-L	.973	.964	*.978*	**.955**	.959	.990	*.979*	**.994**	.980	.991	.724	*.633*	.795	.689	**.809**
CGH-R	.894	.851	*.897*	**.762**	.876	**.988**	.957	*.600*	.974	.985	.896	.769	*.134*	**.899**	.891
CGH-L	*.937*	**.796**	.817	.847	.886	*.255*	.975	**.992**	.889	*.619*	*.021*	.890	**.957**	.551	*.157*
FX/ST-R	.807	.850	.819	*.871*	**.502**	**.987**	.976	*.972*	.972	.986	.936	.858	.863	*.820*	**.972**
FX/ST-L	*.877*	.799	.777	.866	**.756**	*.947*	.982	**.996**	.984	.978	*.687*	.916	**.981**	.889	.914
SLF-R	.975	**.961**	.975	.981	*.981*	.982	.980	**.996**	.975	*.961*	.578	.661	**.875**	.433	*.315*
SLF-L	.976	**.949**	.970	.968	*.983*	**.993**	*.988*	.990	.991	.991	.779	**.806**	.752	.773	*.651*
SFOF-R	.867	.897	*.947*	.828	**.822**	*.779*	.801	**.965**	*.723*	*.709*	*.319*	*.293*	**.592**	*.310*	*.302*
SFOF-L	.872	**.857**	*.927*	.916	.867	.889	**.916**	.913	*.477*	.831	.507	**.610**	.434	*.072*	.395
UF-R	**.692**	.830	*.704*	*.925*	.761	.978	.958	.975	*.951*	**.983**	.931	.795	.921	*.594*	**.933**
UF-L	.782	.859	.789	*.921*	**.176**	**.984**	.940	*.759*	*.762*	.972	.931	.688	*.400*	*.202*	**.966**
TP-R	**.608**	.709	*.821*	.812	.805	.974	*.923*	**.980**	.933	.928	**.937**	.778	.899	.722	*.717*
TP-L	**.674**	.800	.795	*.864*	*.698*	**.993**	*.967*	.990	.972	.985	**.980**	.853	.954	*.826*	.951

Table F.3: Region-based statistics for measurement: RTPP. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	—	**.839**	.913	.934	*.962*	—	.975	.985	*.847*	**.995**	—	.864	.853	*.268*	**.887**
PCT	—	**.832**	.932	.850	*.948*	—	.978	.969	*.917*	**.992**	—	**.881**	.682	*.624*	.859
GCC	—	**.915**	*.967*	.950	.944	—	.989	.988	*.973*	**.996**	—	.885	.725	*.643*	**.927**
BCC	—	.964	*.979*	.967	**.801**	—	*.973*	**.995**	.985	.978	—	*.568*	.806	.684	**.898**
SCC	—	**.922**	.947	*.965*	.923	—	.983	*.979*	**.988**	.986	—	.822	*.712*	.744	**.845**
FRX	—	.771	*.818*	**.552**	*.681*	—	*.902*	.980	.975	**.984**	—	*.677*	.900	.945	**.952**
CST-R	—	**.817**	*.921*	.906	.866	—	.951	**.986**	*.950*	.958	—	.781	**.851**	*.642*	.755
CST-L	—	.814	**.775**	.784	*.874*	—	*.913*	.963	.933	**.969**	—	*.660*	**.854**	.752	.799
ML-R	—	.779	.820	**.708**	*.847*	—	.937	**.977**	.896	*.891*	—	.767	**.885**	.716	*.556*
ML-L	—	**.785**	.846	*.919*	.832	—	.939	**.987**	.956	*.923*	—	.767	**.920**	*.639*	.668
ICP-R	—	**.661**	*.909*	.840	.784	—	.942	**.973**	.971	*.795*	—	**.847**	.769	.843	*.456*
ICP-L	—	**.602**	*.926*	.883	.899	—	.897	**.942**	.876	*.717*	—	**.775**	.547	.453	*.204*
SCP-R	—	*.792*	.779	**.700**	.744	—	.848	.894	**.964**	*.691*	—	.538	.652	**.888**	*.364*
SCP-L	—	.689	*.851*	**.662**	.785	—	*.803*	**.970**	.879	.897	—	*.560*	**.826**	.711	.652
CP-R	—	.691	.759	*.833*	**.628**	—	*.960*	.984	.982	**.991**	—	*.883*	.938	.902	**.977**
CP-L	—	**.658**	.886	*.937*	.853	—	*.945*	.984	.973	**.992**	—	.855	.873	*.691*	**.945**
ALIC-R	—	**.793**	*.900*	.864	.898	—	*.926*	**.974**	.947	.950	—	.721	**.787**	.709	*.660*
ALIC-L	—	*.895*	.853	.813	**.795**	—	*.964*	**.994**	.987	.987	—	*.738*	**.961**	.932	.941
PLIC-R	—	**.871**	.963	*.975*	.906	—	*.962*	**.996**	.977	.994	—	.763	.898	*.510*	**.935**
PLIC-L	—	.869	*.939*	.896	**.775**	—	*.931*	.994	.967	**.997**	—	*.637*	.911	.752	**.985**
RPIC-R	—	**.780**	*.952*	.891	.902	—	*.948*	**.994**	.992	.991	—	*.799*	.897	**.929**	.919
RPIC-L	—	.889	*.946*	.919	**.827**	—	.952	**.995**	*.908*	.991	—	.690	.917	*.443*	**.952**
ACR-R	—	.894	*.979*	.948	**.732**	—	.976	**.996**	*.975*	.990	—	.814	.840	*.667*	**.965**
ACR-L	—	.942	*.984*	.952	**.879**	—	.988	**.996**	*.956*	.995	—	.822	.785	*.509*	**.957**
SCR-R	—	.967	*.993*	.975	**.927**	—	*.875*	.984	.968	**.992**	—	*.189*	.295	.429	**.899**
SCR-L	—	.972	*.990*	.964	**.939**	—	*.978*	.981	.987	**.995**	—	.563	*.334*	.731	**.930**
PCR-R	—	**.930**	*.985*	.983	.973	—	*.953*	.996	.985	**.996**	—	.588	.775	*.524*	**.862**
PCR-L	—	.951	*.985*	.959	**.797**	—	*.960*	**.997**	.991	.996	—	*.539*	.809	.821	**.983**
PTR-R	—	**.827**	*.969*	.850	.891	—	**.977**	.961	*.672*	.928	—	**.878**	.436	*.235*	.585
PTR-L	—	.860	*.921*	.907	**.740**	—	.994	**.999**	*.702*	.992	—	.959	**.985**	*.180*	.970
SS-R	—	.883	.876	**.827**	*.919*	—	**.974**	.969	*.796*	.958	—	**.812**	.794	*.403*	.647
SS-L	—	.871	*.732*	*.963*	**.632**	—	*.982*	.982	**.986**	.985	—	.873	.935	*.731*	**.960**
EC-R	—	.825	*.857*	.839	**.800**	—	*.979*	**.994**	.988	.991	—	*.890*	.957	.929	**.957**
EC-L	—	*.901*	.899	.814	*.681*	—	*.979*	**.997**	.993	.996	—	*.821*	.971	.963	**.988**
CGG-R	—	**.911**	*.984*	.932	.918	—	*.977*	**.996**	.995	.993	—	*.789*	.791	**.929**	.919
CGG-L	—	.896	*.968*	.936	**.837**	—	.982	**.994**	*.910*	.994	—	.849	.840	*.394*	**.964**
CGH-R	—	.784	*.736*	*.821*	**.686**	—	*.884*	.973	**.986**	.912	—	*.622*	.905	**.928**	.766
CGH-L	—	*.797*	**.704**	.771	.783	—	*.924*	**.991**	.984	.986	—	*.712*	**.971**	.933	.937
FX/ST-R	—	.819	.841	**.584**	*.848*	—	.940	.964	**.980**	*.914*	—	.738	.807	**.954**	*.616*
FX/ST-L	—	*.872*	**.417**	*.558*	*.736*	—	.978	.993	*.905*	**.995**	—	.850	**.988**	*.808*	.980
SLF-R	—	.970	*.985*	.935	**.897**	—	*.971*	**.998**	.993	.996	—	*.498*	.864	.905	**.958**
SLF-L	—	.955	*.961*	**.888**	.951	—	.970	*.584*	.972	**.982**	—	.598	*.052*	**.795**	.728
SFOF-R	—	.832	*.908*	*.699*	**.366**	—	*.469*	**.969**	.928	.863	—	*.130*	.741	.794	**.799**
SFOF-L	—	**.869**	.945	*.954*	.870	—	.951	.888	**.967**	*.804*	—	**.719**	*.303*	.574	*.347*
UF-R	—	*.808*	.757	.789	**.708**	—	**.913**	*.812*	*.403*	.889	—	.668	*.512*	*.125*	**.699**
UF-L	—	.880	**.853**	*.903*	.869	—	.937	**.977**	*.933*	.959	—	.642	**.863**	*.575*	.754
TP-R	—	.782	*.853*	*.733*	**.697**	—	*.816*	**.985**	.869	.972	—	*.492*	.905	.640	**.912**
TP-L	—	.811	.757	*.822*	**.713**	—	.958	**.990**	*.725*	.946	—	.812	**.961**	*.319*	.833

Table F.4: Region-based statistics for measurement: NG. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	*.992*	.967	**.966**	.988	.975	**.999**	*.995*	.997	.995	.998	.861	.873	**.929**	*.714*	.908
PCT	*.981*	.958	.940	.973	**.935**	**.997**	.980	*.959*	.994	.970	**.863**	.680	*.587*	.816	.680
GCC	*.989*	.965	.951	.954	**.940**	**.997**	*.985*	.995	.996	.988	.800	*.696*	.901	**.912**	.828
BCC	*.996*	.979	**.962**	.963	.984	**.998**	.997	*.994*	.995	.996	*.702*	**.889**	.872	.874	.802
SCC	*.988*	.962	**.929**	.988	.951	**.998**	*.980*	.983	.997	.992	**.863**	*.658*	.805	.794	.856
FRX	.707	*.796*	**.564**	.641	.660	**.987**	.976	.973	*.960*	.975	**.958**	*.893*	.939	.897	.931
CST-R	*.969*	.929	**.901**	.968	.927	**.988**	.982	.987	*.958*	.985	.714	.792	**.885**	*.421*	.824
CST-L	*.974*	.924	.884	.959	**.851**	**.982**	.975	.956	.978	*.922*	*.581*	**.750**	.717	.646	.639
ML-R	*.916*	.902	.874	.877	**.866**	**.989**	.978	.947	*.864*	.976	**.887**	.814	.693	*.437*	.847
ML-L	*.917*	.891	.834	**.800**	.846	**.981**	.980	.928	*.912*	.950	.809	**.844**	.681	*.675*	.747
ICP-R	*.927*	.872	.788	.926	**.491**	.970	**.975**	*.749*	.832	.923	.701	.834	.387	*.268*	**.860**
ICP-L	*.904*	.896	.844	.754	**.746**	**.961**	.939	.912	.954	*.830*	.702	.613	.619	**.835**	*.554*
SCP-R	.878	.853	**.840**	*.883*	.865	.923	**.955**	*.852*	.915	.872	.594	**.758**	*.480*	.560	.480
SCP-L	*.915*	.814	.699	.888	.820	**.988**	.971	*.921*	.972	.941	**.880**	.863	.778	.795	*.741*
CP-R	.944	.889	**.609**	*.968*	.839	**.996**	.989	.988	.992	*.987*	.934	.909	**.970**	*.791*	.927
CP-L	.960	.922	**.888**	*.971*	.898	**.997**	.991	.992	*.989*	.994	.933	.893	.935	*.730*	**.942**
ALIC-R	*.982*	.938	**.912**	.970	.942	**.997**	*.987*	.990	.992	.997	.864	.822	.897	*.799*	**.944**
ALIC-L	.906	.894	.875	*.968*	**.841**	**.996**	.990	.987	*.982*	.990	**.954**	.917	.907	*.633*	.938
PLIC-R	.972	.956	**.675**	*.977*	.900	**.999**	.995	.997	*.985*	.999	.961	.897	**.991**	*.604*	.985
PLIC-L	.955	.909	.874	*.956*	**.796**	**.998**	.995	.997	*.986*	.997	.966	.946	.978	*.754*	**.985**
RPIC-R	.958	.916	**.847**	*.974*	.937	**.996**	.993	.991	*.970*	.988	.902	.924	**.944**	*.458*	.843
RPIC-L	.963	.917	**.740**	*.975*	.864	**.996**	.992	.994	*.972*	.994	.899	.910	**.977**	*.466*	.959
ACR-R	.990	.976	**.903**	*.991*	.927	**.999**	.998	.998	*.994*	.998	.906	.913	**.981**	*.604*	.979
ACR-L	*.988*	.973	**.781**	.984	.924	**.999**	.997	.997	*.996*	.998	.932	.904	**.988**	*.807*	.972
SCR-R	*.995*	.989	**.964**	.995	.967	**.998**	.992	*.991*	.995	.992	.755	.578	.794	*.502*	**.806**
SCR-L	.991	.981	**.835**	*.995*	.907	**.999**	.996	.997	*.996*	.997	.905	.837	**.983**	*.580*	.968
PCR-R	*.985*	.983	**.886**	.980	.928	**.998**	.994	.995	*.989*	.996	.881	.745	**.959**	*.638*	.942
PCR-L	.970	.975	.723	*.982*	.881	**.998**	.994	*.992*	.993	.995	.943	.808	**.971**	*.703*	.963
PTR-R	*.984*	.964	**.909**	.980	.959	**.992**	.984	.983	*.982*	.986	.678	.684	**.840**	*.519*	.737
PTR-L	.975	.936	**.807**	*.977*	.810	**.998**	.997	.998	*.985*	.997	.933	.962	**.988**	*.606*	.987
SS-R	.974	.938	**.867**	*.974*	.913	**.973**	.970	.683	.970	*.584*	.482	**.666**	.223	.453	*.108*
SS-L	.950	*.960*	**.867**	.958	.938	.987	**.989**	.988	.977	*.972*	.797	.784	**.913**	*.641*	.683
EC-R	.969	**.823**	.836	*.981*	.909	**.994**	.990	.987	*.947*	.986	.842	**.946**	.928	*.257*	.869
EC-L	.977	.877	**.876**	*.983*	.897	**.998**	.995	.995	*.942*	.996	.915	.960	.959	*.214*	**.967**
CGG-R	*.986*	.961	**.920**	.970	.963	**.992**	.990	.987	.979	*.971*	.622	.797	**.858**	.580	*.549*
CGG-L	*.980*	.958	.688	.967	.885	.996	.992	**.996**	*.952*	.994	.815	.844	**.986**	*.393*	.950
CGH-R	*.911*	.859	**.765**	.824	.770	*.884*	.967	.911	**.984**	.962	*.405*	.805	.705	**.915**	.855
CGH-L	.908	.788	.498	*.931*	.698	**.990**	.982	.945	*.900*	.949	.906	**.920**	.897	*.383*	.849
FX/ST-R	.926	.812	.728	*.945*	.838	.946	.956	.804	**.974**	*.777*	.564	**.804**	.527	.671	*.361*
FX/ST-L	.950	**.819**	.854	*.955*	.893	.996	.990	**.996**	*.984*	.996	.922	.945	**.976**	*.727*	.964
SLF-R	.993	.979	**.956**	*.993*	.959	**.998**	.997	.997	*.990*	.998	.798	.875	.942	*.390*	**.953**
SLF-L	*.991*	.965	.869	.989	**.862**	.996	.994	**.997**	*.985*	.996	.690	.863	**.975**	*.436*	.974
SFOF-R	*.951*	.897	**.806**	.920	.908	*.773*	.930	*.793*	**.969**	*.771*	*.142*	.577	*.426*	**.717**	*.236*
SFOF-L	.762	.918	.857	*.960*	.733	.987	*.975*	.985	**.988**	.985	**.948**	*.762*	.903	.768	.947
UF-R	**.640**	.774	*.814*	.805	.763	.975	.957	.909	*.827*	**.980**	**.933**	.834	.651	*.483*	.921
UF-L	.876	*.883*	.700	.881	.767	**.976**	.964	.963	*.710*	.973	.835	.757	.885	*.225*	**.892**
TP-R	.483	.797	.543	*.813*	**.321**	**.991**	.964	.976	*.959*	.984	**.983**	.843	.949	*.814*	.977
TP-L	.790	*.868*	.727	.861	.725	**.995**	.982	.970	*.945*	.983	**.975**	.881	.898	*.706*	.942

Table F.5: Region-based statistics for measurement: MSD. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).

	REPEATABILITY					SEPARABILITY					RELIABILITY				
	DTI	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*	*DTI*	*AMURA*	*MISFIT*	*MAPL*	*DSI*
MCP	.977	.947	.713	*.978*	.968	**.998**	.992	*.959*	.998	.995	**.902**	.871	.871	.899	*.863*
PCT	.958	.935	**.803**	*.959*	.955	.992	.964	*.909*	.992	**.993**	.846	*.635*	.664	.840	**.868**
GCC	.949	**.924**	*.954*	.950	.949	.997	*.989*	.992	**.997**	.995	**.946**	.874	*.850*	.946	.912
BCC	.951	.921	*.968*	.952	**.841**	.943	.988	**.993**	*.943*	.970	.447	**.864**	.831	*.443*	.838
SCC	.972	.929	**.910**	*.973*	.933	.970	*.926*	.949	**.973**	.941	.481	*.471*	**.625**	.493	.516
FRX	.659	.646	*.744*	.654	.686	.986	*.973*	.989	.986	**.990**	.961	*.926*	.959	.962	**.968**
CST-R	.955	**.899**	.927	*.957*	.916	**.985**	*.964*	.977	.985	.969	.750	.733	**.757**	.737	*.722*
CST-L	.926	.914	.588	*.927*	.893	.985	.933	*.861*	**.985**	.953	**.828**	*.544*	.718	.825	.683
ML-R	.741	**.733**	.819	.748	*.873*	.909	.911	**.959**	.912	*.906*	.721	.733	**.808**	.722	*.552*
ML-L	.870	**.826**	.857	*.872*	.833	.858	.953	**.979**	*.841*	.950	.441	.778	**.868**	*.405*	.760
ICP-R	.718	*.861*	.837	.727	.785	.883	**.962**	.950	.886	*.798*	.681	**.778**	.757	.681	*.459*
ICP-L	**.843**	.846	.881	.844	*.902*	.951	.847	.807	**.952**	*.587*	.752	.461	.333	**.755**	*.122*
SCP-R	.824	.803	.735	*.831*	**.711**	.881	**.899**	.844	.882	*.767*	.566	**.637**	.589	.558	*.487*
SCP-L	.820	**.709**	.806	*.826*	.790	.925	*.892*	.910	**.928**	.901	.691	**.707**	.663	.692	*.657*
CP-R	.930	.892	**.649**	*.934*	.683	.992	.987	*.960*	**.993**	.992	.901	*.892*	.893	.898	**.976**
CP-L	.838	.813	.713	.840	*.840*	.985	.982	*.966*	.986	**.990**	.914	.912	*.892*	.917	**.941**
ALIC-R	.897	**.869**	.874	*.900*	.883	.953	.952	**.973**	.953	*.892*	.675	.722	**.820**	.671	*.489*
ALIC-L	.912	.850	.551	*.917*	.803	.995	*.983*	.995	**.995**	.987	.945	*.895*	**.988**	.943	.937
PLIC-R	.963	.928	.923	*.965*	**.908**	.986	*.972*	**.992**	.988	.991	.726	*.716*	.905	.734	**.908**
PLIC-L	*.921*	.912	.913	.918	**.803**	.994	*.986*	.986	.994	**.996**	.930	*.863*	.863	.934	**.982**
RPIC-R	.927	.918	.911	*.928*	**.889**	.988	*.988*	**.995**	.988	.990	.860	.873	**.945**	*.858*	.916
RPIC-L	.935	.930	.878	*.938*	**.746**	.995	*.990*	**.995**	.995	.991	.927	*.868*	.963	.924	**.964**
ACR-R	.986	.963	.955	*.987*	.718	.995	.994	**.997**	.995	*.987*	.737	.870	.933	*.730*	**.954**
ACR-L	.983	.956	.968	*.984*	**.842**	.995	*.984*	**.996**	.995	.994	.769	*.724*	.891	.775	**.965**
SCR-R	.992	.975	.988	*.993*	**.920**	.996	*.982*	.986	**.996**	.988	.659	.582	*.456*	.655	**.871**
SCR-L	.987	.970	.970	*.988*	**.950**	.997	*.993*	.994	**.997**	.996	.811	.815	.841	*.807*	**.924**
PCR-R	.973	.976	*.979*	.974	**.964**	.996	*.991*	**.997**	.996	.995	.864	*.736*	.876	.861	**.889**
PCR-L	.948	.950	*.961*	.949	**.824**	.998	*.993*	**.998**	.998	.997	.958	*.880*	.952	.958	**.982**
PTR-R	.963	.928	.957	*.964*	**.887**	.992	.976	.990	**.992**	*.945*	**.818**	.746	.810	.816	*.660*
PTR-L	.947	.855	.824	*.949*	**.754**	.999	.996	**.999**	.999	*.991*	.975	.972	**.993**	.974	*.963*
SS-R	.934	.882	**.810**	*.938*	.900	.964	*.832*	.875	**.966**	.958	.637	*.369*	.570	.635	**.696**
SS-L	.918	*.943*	.704	.920	.741	.982	.976	*.974*	**.982**	.979	.814	*.695*	.918	.816	**.923**
EC-R	.965	.864	.570	*.968*	.816	.994	*.988*	.993	**.994**	.990	.852	.921	**.983**	*.844*	.950
EC-L	.966	.850	.823	*.967*	**.786**	.997	*.993*	.997	**.997**	.995	.911	.956	**.981**	*.907*	.978
CGG-R	.974	.928	.961	*.975*	**.911**	.995	*.993*	**.997**	.995	.993	.840	.908	.917	*.834*	**.930**
CGG-L	.963	.932	.942	*.964*	**.850**	.997	*.990*	.994	**.997**	.994	.915	*.877*	.912	.912	**.961**
CGH-R	.817	*.868*	.708	.820	.681	.906	.963	**.971**	*.891*	.958	.640	.777	**.907**	*.595*	.880
CGH-L	.795	*.838*	.716	.801	.777	.991	*.978*	.990	**.991**	.985	.956	*.876*	**.966**	.955	.936
FX/ST-R	.910	**.783**	.829	*.912*	.854	.907	**.940**	.929	.896	*.874*	.469	**.775**	.691	*.432*	.503
FX/ST-L	.916	.798	.680	*.919*	**.547**	.996	*.988*	.993	**.996**	.994	.953	*.945*	.977	.952	**.986**
SLF-R	.988	.958	.947	*.989*	**.936**	.997	.995	**.998**	.997	*.995*	.796	.900	**.963**	*.787*	.928
SLF-L	.988	.927	**.917**	*.988*	.964	.971	**.985**	*.882*	.971	.979	.282	**.827**	.383	*.282*	.629
SFOF-R	.862	.823	*.883*	.867	**.446**	*.781*	.928	**.953**	.788	.914	*.331*	.695	.702	.332	**.855**
SFOF-L	**.822**	*.929*	.911	.824	.915	.966	.956	*.509*	**.968**	.836	.836	.607	*.084*	**.844**	.303
UF-R	.664	*.832*	.743	.673	.771	.954	.926	.913	**.956**	*.895*	.875	.677	.730	**.877**	*.662*
UF-L	**.863**	.887	*.892*	.866	.869	**.986**	*.960*	.984	.986	.964	**.904**	*.731*	.871	.902	.777
TP-R	.780	*.826*	.748	.786	.725	.990	*.952*	.989	**.990**	.972	.955	*.776*	**.958**	.955	.905
TP-L	.905	.831	.499	*.906*	.725	.985	.967	**.992**	.985	*.941*	.859	.834	**.985**	.858	*.814*

Table F.6: Region-based statistics for measurement: QIV. The smaller the repeatability, the better. The larger the separability and the reliability, the better. Values in red correspond to p-values over 0.05; values in orange correspond to p-values over 0.01. Bold-face is used to highlight the best performer in each case (italics are used for the worst performer).