Enabling cortical cell-specific sensitivity on diffusion MRI microstructure measurements using likelihood-free inference

Maëliss Jallais

université PARIS-SACLAY

Enabling Cortical Cell-Specific Sensitivity on Diffusion MRI Microstructure Measurements using Likelihood-Free Inference

Extension de la mesure de la microstructure par IRM de diffusion aux cellules corticales via inférence par simulation

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 580, Sciences et Technologies de l'Information et de la Communication (STIC)
Spécialité de doctorat : Traitement du signal et des images
Graduate School : Informatique et sciences du numérique, Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche Université Paris-Saclay, Inria, Inria Saclay-Île-de-France, 91120, Palaiseau, France, sous la direction de Dr. Demian Wassermann, chargé de recherche.

Thèse soutenue à Paris-Saclay, le 16 février 2022, par

Maëliss Jallais

Composition du jury

Derek Jones Professor, School of Psychology at Cardiff University and Director of CUBRIC	Président
Marco Palombo Associate Professor, School of Psychology at Cardiff University	Rapporteur & Examinateur
Jean-Philippe Thiran Professor, Ecole Polytechnique Fédérale de Lausanne (EPFL)	Rapporteur & Examinateur
Alexandre Gramfort Directeur de recherche, Université Paris-Saclay, Inria, CEA	Examinateur
Ileana Jelescu Associate Professor, Department of Radiology at Lausanne University Hospital (CHUV) and University of Lausanne (UNIL)	Examinateur
Dmitry Novikov Associate Professor, Department of Radiology at NYU Grossman School of Medicine	Examinateur
Julien Valette Directeur de recherche, CEA - Fontenay-aux-Roses	Examinateur
Demian Wassermann Associate Professor, Université Paris-Saclay, Inria, CEA	Directeur de thèse

THESE DE DOCTORAT
NNT : 2022UPASG012

Abstract

Neurodegenerative diseases, such as Alzheimer's or Huntington's disease, lead to the progressive and irreversible loss of mental functions. Dementia and cognitive deficits appear to be primarily related to neuronal and synaptic connectivity loss. Although these diseases' external impact and progression are readily observable, accessing microstructural changes in the brain remains a challenge, making it difficult to understand these diseases and develop treatments.

With technological advances, diffusion Magnetic Resonance Imaging (dMRI) has emerged as a novel method to study the microstructure of the brain non-invasively and in-vivo. This medical imaging technique is based on the study of random microscopic movements of water molecules, known as Brownian movements. In the brain, the movements of the molecules are constrained by cell membranes, making diffusion anisotropic. Each tissue component, such as somas (cell bodies) or neurites, has a distinct shape. The characteristics of the tissue thus modulate the diffusion brain signal obtained during an MRI acquisition.

My thesis aims to develop a method to infer a tissue microstructure from a dMRI acquisition in the gray matter (GM).

The solution to this inverse problem of estimating brain microstructure from dMRI is threefold:

1. **The definition of a biological model describing the GM tissues.** Existing microstructural models of white matter were proven not to hold in the GM. We adapted these models to take into account the abundance of somas in the GM.

2. **A mathematical modeling of the GM tissue.** We modeled each compartment of the tissue model by simple geometrical shapes, for which the diffusion signal is known. We developed a signal processing algorithm to synthesize the key information contained in the diffusion signal and relate it to a set of parameters describing the tissue (notably the size and density of neurons). This algorithm is based on a study of the statistical moments of the signal at different MRI gradient strengths. Unlike existing methods, no biological parameters are arbitrarily fixed, which allows for the best possible description of the cortical tissue of each subject.

3. **An inversion algorithm to estimate the tissue parameters that generated the acquisition signal.** Once the mathematical model relating tissue parameters to the diffusion signal is defined, the objective is to solve the inverse problem of estimating tissue microstructure from an observation. A limitation of current methods is their inability to identify all possible tissue configurations that can explain the same observed diffusion signal, making the interpretation of the proposed estimates difficult. We used a Bayesian deep-learning method called "likelihood-based inference" combined with neural networks to solve this problem. This method allows identifying and returning all possible tissue configurations along with their posterior distributions (probability given an observation), facilitating their interpretation.

We first validated this approach on simulations. Based on a few acquisition constraints, we then applied the global resolution method to the HCP MGH and HCP1200 databases of the Human Connectome Project. We developed a python library to study those simulated or acquired data. The obtained results were then compared with histological and cognitive studies to verify their validity.

Résumé

Les maladies neurodégénératives, telles que la maladie d'Alzheimer ou de Huntington, entraînent la perte progressive et irréversible des fonctions mentales. La démence et les déficits cognitifs semblent être principalement liés à une perte neuronale et de connectivité synaptique. Bien que l'impact et l'évolution extérieurs de ces maladies soient facilement observables, l'accès aux changements microstructuraux dans le cerveau reste un défi, rendant difficiles la compréhension de ces maladies et la mise au point de traitements.

Avec les avancées technologiques, l'Imagerie par Résonance Magnétique de diffusion (IRMd) s'impose comme une méthode novatrice pour étudier la microstructure du cerveau de manière non-invasive et in-vivo. Cette technique d'imagerie médicale est basée sur l'étude des mouvements microscopiques aléatoires des molécules d'eau, connus sous le nom de mouvements Browniens. Dans le cerveau, les mouvements des molécules sont contraints par les membranes des cellules, rendant la diffusion anisotropique. Chaque composant tissulaire, tel que les somas (corps des neurones) ou les neurites, possède une forme distincte. Le signal de diffusion du cerveau obtenu lors d'une acquisition IRM est ainsi modulé selon les caractéristiques du tissu.

L'objectif de ma thèse est de mettre en place une méthode qui permette d'inférer la microstructure d'un tissu à partir d'une acquisition d'IRM de diffusion dans la matière grise (MG).

La résolution de ce problème inverse d'estimation de la microstructure du cerveau à partir de l'IRMd s'organise autour de trois parties:

1. **La définition d'un modèle biologique décrivant les tissus de la MG.** Il a été prouvé que les modèles microstructuraux existants modélisant la matière blanche ne sont pas valides dans la MG. Nous avons commencé par adapter ces modèles en prenant en compte l'abondance des somas dans la MG.

2. **Une modélisation mathématique de la MG.** Chaque compartiment du modèle tissulaire a ensuite été modélisé par des formes géométriques simples, pour lesquelles le signal de diffusion est connu. Un algorithme de traitement du signal a été développé permettant de synthétiser les informations clés contenues dans le signal de diffusion et de les mettre en relation avec un

ensemble de paramètres décrivant le tissu (notamment la taille et la densité des neurones). Cet algorithme se base sur une étude des moments statistiques du signal pour différentes puissances de gradient de l'IRM. À la différence des méthodes existantes, aucun paramètre biologique n'est arbitrairement fixé, ce qui permet de décrire au mieux les tissus corticaux de chaque sujet.

3. **Un algorithme d'inversion permettant d'estimer les paramètres du tissu ayant généré le signal d'acquisition.** Une fois le modèle mathématique permettant de relier les paramètres du tissu au signal de diffusion défini, l'objectif est de résoudre le problème inverse d'estimation de la microstructure du tissu à partir d'une observation. Une limitation des méthodes actuelles est leur incapacité à identifier toutes les configurations du tissu possibles pouvant expliquer un même signal de diffusion observé, ce qui rend l'interprétation des estimations proposées difficile. Afin de résoudre ce problème, nous avons utilisé une méthode reposant sur des outils de l'analyse bayésienne et de deep learning appelée "likelihood-based inference", combinée à des réseaux de neurones. Celle-ci permet d'identifier et de retourner toutes les configurations possibles du tissu accompagnées de leurs distributions postérieures (probabilité étant donné une observation), ce qui facilite leur interprétation.

L'approche a tout d'abord été validée sur des simulations. Reposant sur peu de contraintes d'acquisition, la méthode de résolution globale a ensuite été appliquée sur les bases de données HCP MGH et HCP1200 du Human Connectome Project. Une bibliothèque python a été développée pour étudier ces données réelles ou simulées. Les résultats obtenus ont enfin été comparés avec des études histologiques et cognitives pour vérifier leur validité.

Acknowledgement

First and foremost, I would like to thank Demian. I am grateful for your patience, your advice, and your invaluable training. Thank you for always pushing me forward and supporting me at each step of this thesis. This thesis would not have been possible without you.

I thank Pr. Jean-Philippe Thiran, Dr. Marco Palombo, Pr. Julien Valette, Pr. Derek Jones, Dr. Dmitry Novikov, Dr. Ileana Jelescu, and Dr. Alexandre Gramfort for accepting to review this manuscript and be part of my defense jury.

I would like to extend my warm thanks to all people I collaborated with during this PhD and who made these works possible, including Pedro Rodrigues, Alexandre Gramfort, Thomas Meunier, and Chengran Fang. Special thanks go to Pedro, who was a great collaborator and friend. Thank you for your constant great mood and all the knowledge you shared with me.

I extend my thanks to all members of the Parietal team, former and actual, for these three years spent with you, particularly to the director of the team, Bertrand Thirion. I have a grateful thought for all the researchers who became my friends and made this PhD so enjoyable. It was a pleasure to travel, spend Parietal retreats, ping-pong games, coffee breaks, and after-works with such fun and diverse people. I especially want to thank Louis Rouillard, Raphaël Meudec, the best office composed of Hicham Janati, David Sabbagh and Valentin Iovene, Gaston Zanitti, Chengran Fang, Jonas Renault, Hugo Richard, Thomas Bazeille, Thomas Moreau (you are definitely a PI now!), Pierre Ablin, Patricio Cerda, Omar Chehab, Judith Abecassis, Alexis Thual, Guillaume Favellier, Marine Le Morvan and Majd Abdallah. I also want to thank Olivier Grisel, Gaël Varoquaux, Philippe Ciuciu, and Guillaume Lemaitre for exciting discussions, and Alain Couvreur for his support and offering me the opportunity to share my knowledge and passion with high school girls. I also thank Corinne Petitot, Valérie Berthou and Camille François for helping me navigate administrative processes.

I would like to address a special thank to my friends who were always there for me and made these three years incredible, and in particular Prince, Tran, Thomas, Alex, Laura and JB. I deeply thank Mathieu for being there, encouraging me, and supporting me every day.

Je voudrais terminer par remercier toute ma famille pour m'avoir toujours soutenue, et en particulier mes parents et mes soeurs. Merci de m'avoir donné toutes les chances de réussir et de m'avoir permis d'arriver où j'en suis. Une pensée toute particulière pour ma Mémé, Dédette et Tatou.

Contents

1 Introduction

The brain is the central human organ for taking decisions. Its alteration by neurodegeneratives diseases, such as Alzheimer's or Huntington's, can lead to dramatic consequences and considerably affect everyday life and social interactions (Monica Moore et al., 2021). Symptoms include a progressive and irreversible loss of mental functions, and particularly memory. The external impact and progression of such diseases are readily observable. However, accessing the changes at the microstructure level that engendered these functional losses remains a challenge. Quantifying changes at the neuronal scale seems yet crucial to understanding neuronal diseases, as dementia severity and cognitive deficits appear to be primarily related to neuronal loss and synaptic pathology (see e.g. Amor et al., 2010; Douaud et al., 2013). Histology is a technique for estimating a brain tissue microstructure by analyzing it under a microscope. However, this technique is highly invasive and can only be performed post-mortem. In addition, it provides only one time point and prevents us from having a global view of brain microstructure and connections. It is of utmost importance to develop techniques to measure tissue microstructure in-vivo and non-invasively to understand these diseases and develop treatments.

With technological advances, Magnetic Resonance Imaging (MRI) has emerged as a novel method to study the brain non-invasively and in-vivo. Three modalities have been designed from this imaging technique: anatomical MRI, functional MRI (fMRI), and diffusion MRI (dMRI). They provide complementary information on brain architecture, function, and development. Anatomical MRI corresponds to the most clinically widespread type of clinical MRI and allows to visualize contrasts between tissues. fMRI relies on blood flow changes to detect brain activation in response to functional or cognitive tasks. dMRI quantifies how water molecules diffuse in the brain. Indeed, water molecules undergo random microscopic movements, known as Brownian movements. These movements are notably restricted by the cell membranes. Studying these movements allows to deduct cell characteristics and, by extension, a tissue microstructure (Figure 1.1). Diffusion MRI is thus a perfectly adapted technique to study brain microstructure.

The mammalian brain can be decomposed into two main parts: white matter and gray matter. The former designates regions that contain mainly long-range myelinated axons and relatively few cell bodies. Axons are connections that cross the brain, bridging different parts of gray matter. This one mainly contains cell bodies called somas, and relatively few myelinated axons. Those somas are connected by

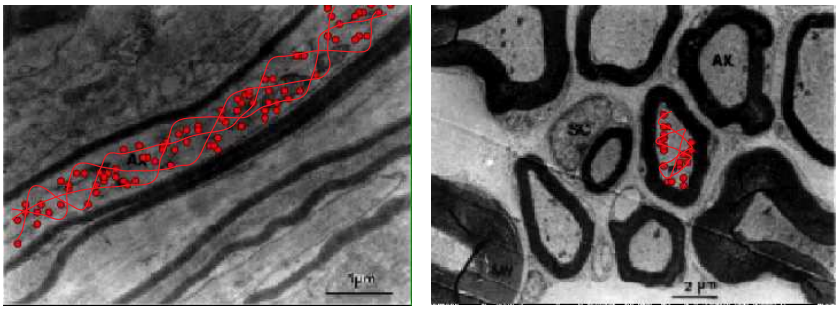


Fig. 1.1: Diffusion within brain tissue (here axons) is impeded by the cell's membranes (Purves, 2004).

neurites. Relevant parameters to estimate in the brain are the proportions of each component (such as somas, neurites, or axons), their size, and the diffusivity of the molecules (i.e., the rate at which molecules spread).

Two complementary approaches emerged for extracting this information about tissue microstructure from a diffusion MRI acquisition. The first one is known as signal representation (denomination from Novikov et al., 2018b) and is based on the decomposition of the diffusion signal on a mathematical basis to provide biological markers. An example of such an approach is diffusion tensor imaging (DTI) (Basser et al., 1994), which provides estimations of brain anisotropy and mean diffusivity of a tissue. These markers have been shown to be sensitive to microstructure changes between healthy and pathological subjects (see e.g. Mori and Zhang, 2006; Assaf and Pasternak, 2008). However, these markers present the limitation of being not specific. This created the need for a method that is both sensitive and specific to tissue changes: biophysical modeling. This technique relies on a schematic geometry of the tissue studied. Figure 1.2 illustrates the modeling of a neuron using simple geometric forms (spheres and cylinders). This method also depends on assumptions meant to simplify the complexity of a biological tissue, which allow keeping only relevant features that characterize the tissue and impact the diffusion signal acquired. However, a different model needs to be developed for each tissue, and the initial geometric assumption must be chosen to accurately capture all of the tissue features that effectively impact the diffusion signal in a given acquisition range (Novikov et al., 2018b). This model defines a relationship between the selected tissue features and the acquired diffusion signal. To estimate those tissue features, we need to be able to solve this inverse problem mathematically.

Current brain tissue models are predominantly based on the two-compartment Standard Model (SM) (Zhang et al., 2012; Novikov et al., 2018a). Recent evidence shows that the SM, mainly used in white matter, does not hold for gray matter microstructure analysis (Veraart et al., 2020). The definition of a model specific to gray matter tissue is thus necessary. In addition, to solve the inverse problem, current methods require demanding acquisition protocols and rely on non-linear

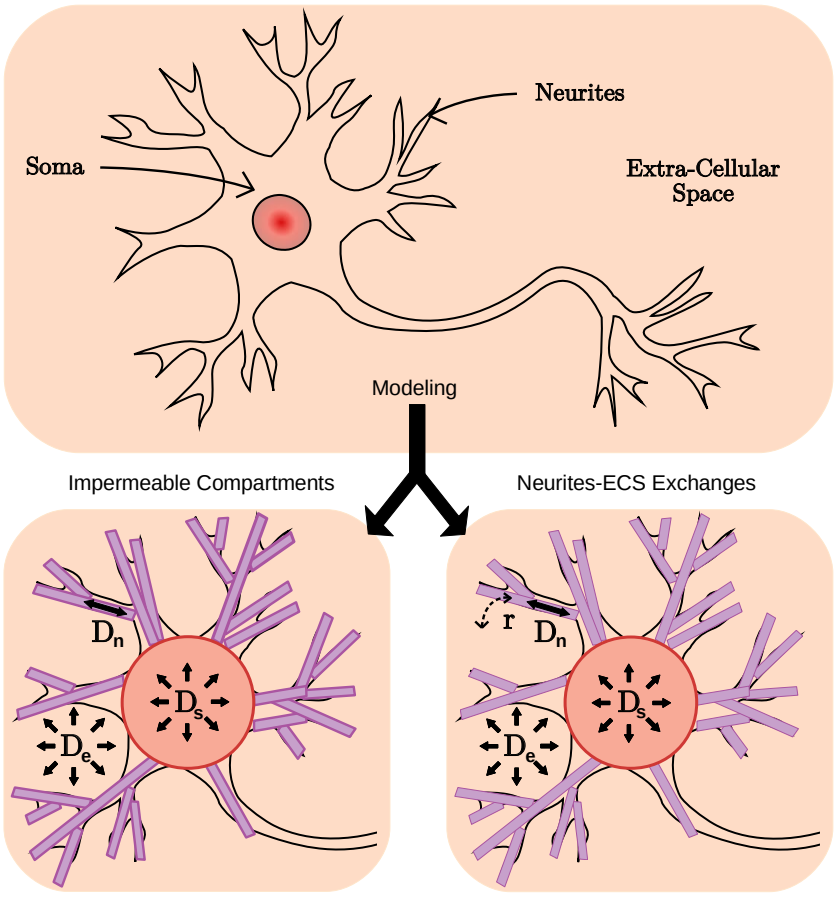


Fig. 1.2: Illustration of biophysical modeling. A neuron is modeled using simple geometric forms. Different assumptions allow defining diverse models (see Chapter 4 for more details about these models). This illustration is inspired from the publication by Olesen et al. (2021).

models for which several parameter values may yield the same observation, also known as parameter indeterminacy (Jelescu et al., 2016; Novikov et al., 2018c). This leads to numerically unstable solutions, which are hard to interpret. Another major challenge is the fact that quantifying brain tissue microstructure directly from the dMRI signal is an inherently difficult task because of the large dimensionality of the collected data and its stochastic nature (Fick et al., 2016).

Objective of the thesis

In this thesis, we focus on brain gray matter microstructure estimations using single encoding diffusion MRI. Our approach is based on **biophysical tissue modeling** to obtain features specific to brain gray matter. Several assumptions have been proposed to explain why the SM does not hold in gray matter. These include the abundance of cell bodies (somas) in gray matter (Shapson-Coe et al., 2021; Palombo et al., 2020), increased permeability in neurite membranes (Veraart et al., 2020; Jelescu et al., 2021; Olesen et al., 2021), or structural disorder (Lee et al., 2020). We decided to account for the presence of cell bodies in the gray matter by including

an impermeable soma compartment into the SM, following Palombo et al. (2020). We then extended this three-compartment model to also take into account exchange between neurites and the ECS, and propose other extensions to more accurately capture gray matter tissue characteristics.

Inferring parameters of the brain tissue model directly from dMRI signals has proven to be a challenging task and has motivated the development of new approaches to reduce the **dimensionality of the data** to be processed. For instance, Novikov et al. (2018c) proposes the LEMONADE system of equations based on a Taylor expansion of the diffusion signal and a set of rotationally invariant moments. In this work, we develop a framework to extract features from the observed signal for two different models. We call the resulting vector of features the 'summary statistics' of the dMRI signal.

Once the biological model and the set of equations relating the dMRI signal to brain tissue parameters are defined, a method is required to **solve the non-linear inverse problem**. A popular solution is NODDI (Zhang et al., 2012), which stabilizes the solution by imposing constraints on model parameters. However, these constraints were shown not to be biologically plausible, they bias the estimations, and the inverse problem remains largely degenerate (Jelescu et al., 2016; Novikov et al., 2018c). There has also been some interest in applying methods from the machine learning literature to solve the inverse problem. This is the case of the SANDI model (Palombo et al., 2020), in which the authors employ random forest regressors. Although the method provides a rather acceptable accuracy in real-case scenarios, it can only output one set of tissue parameters for a given observed dMRI signal, masking, therefore, other biologically plausible solutions that could generate the same observed signal. Furthermore, the parameter estimates are obtained following a deterministic approach, so no confidence interval description is available.

We propose a solution based on **Bayesian statistical inference**. The usual way of applying a Bayesian approach to solve non-linear inverse problems (Stuart, 2010) is to define two quantities: a prior distribution encoding initial knowledge of the parameter values (e.g. physiologically relevant intervals) and the likelihood function of the forward model being studied. One can then either obtain an analytic expression of the posterior distribution via Bayes' formula or use a Markov-Chain Monte Carlo (MCMC) procedure to numerically sample the posterior distribution (Gelman et al., 2013). However, the likelihood function of complex models such as the one we consider here is often very hard to obtain and makes the Bayesian approach rather challenging to use. **Likelihood-free inference** (LFI) bypasses this bottleneck by recurring to several simulations of the forward model using different parameters and learning an approximation to the posterior distribution from these examples (Cranmer et al., 2020).

The first contributions on LFI are known as approximate Bayesian computation (ABC) and have been applied to invert models from ecology, population genetics, and epidemiology (Sisson, 2018). Some of the limitations of these techniques include the large number of simulations required for the posterior estimation and the need to define a distance function to compare the results of two simulations. Recently, there has been a growing interest in the machine learning community in improving the limitations of ABC methods through deep generative modeling, i.e. neural network architectures specially tailored to approximate probability density functions from a set of examples (Goodfellow et al., 2016). **Normalizing flows** (Papamakarios et al., 2019) are a particular class of such neural networks that have demonstrated promising results for likelihood-free inference in different research fields (Cranmer et al., 2020; Gonçalves et al., 2020; Greenberg et al., 2019). In this thesis, we propose an approach that relies on normalizing flows in the context of LFI to solve the non-linear inverse problem of relating an acquired signal to the tissue microstructue. Such probabilistic approach allows to determine the full posterior distribution of the fitted parameters, i.e., all the probability densities of the parameters that could have generated the observed diffusion signal given this signal. These distributions are used to define a confidence factor, that indicates whether we can trust an estimation or not.

Organization of the Manuscript

The present manuscript is composed of five chapters.

In Chapter 2, we present an introduction to brain microstructure estimation using single encoding diffusion MRI. We start with the presentation of diffusion MRI basics and why this imaging technique is perfectly adapted for probing tissue microstructure. We then review two methods for estimating tissue microstructure: signal representation and biophysical modeling. An emphasis is made on the latter, which allows specific tissue features estimations. This chapter is partly based on a book chapter published in 2021:

> **M. Jallais**, and D. Wassermann. Single Encoding Diffusion MRI: A Probe to Brain Anisotropy. Anisotropy Across Fields and Scales, 171, 2021.

In Chapter 3, we present the theoretical foundations of statistical inference. We begin by introducing the general Bayesian formalism to solve an inverse problem, and show that the Bayes theorem can be used to estimate the probability density of the parameters that generated an observation given this observation, called the posterior, when the model's likelihood is available. We give a brief overview on two existing likelihood-based inference methods: Markov Chain Monte Carlo and Variational Inference. However, the likelihood of a simulator model is often hard to evaluate, making the likelihood-based methods hard to use. We present some

likelihood-free (or simulation-based) inference methods that have been developed to bypass this difficulty. Finally, we discuss two application-specific factors that need to be considered when solving an inverse problem: the amortization of a model and the utility of summary statistics.

In Chapter 4, we present two different biophysical brain gray matter models. We start with a three-compartment model for brain tissue composed of neurites, somas, and extra-cellular space, where each compartment is considered impermeable. We introduce a new parameter that jointly encodes soma radius and intracellular diffusivity without imposing constraints on these values. This new parameter reduces indeterminacies in the model and has relevant physiological interpretations. Second, we present a method to fit the model through summary features of the dMRI signal based on a large and small q-value analysis using boundary approximations. These rotationally-invariant features relate directly to the tissue parameters and enable us to invert the model without manipulating the raw dMRI signals. Such summary statistics ensure a stable solution of the parameter estimations, as opposed to the indeterminate models used in Zhang et al. (2012) and Palombo et al. (2020). This analysis in both large and small q-values was first presented at ISMRM in 2020:

> **M. Jallais**, and D. Wassermann. Indetermination-free cytoarchitecture measurements in brain gray matter via a forward diffusion MRI signal separation method. In ISMRM 2020-28th Annual Meeting & Exhibition.

Then, we extend this model to account for exchange between neurites and the ECS (Jelescu et al., 2021; Olesen et al., 2021), and propose a new pipeline for obtaining summary statistics, based on a large and small q-value analysis as well. This work was co-supervised with Demian Wassermann. Preliminary results were submitted to ISMRM 2022:

> T. Meunier, C. Fang, **M. Jallais***, and D. Wassermann*. Full posterior estimation of gray matter cytoarchitecture using a three-compartment model with exchange: a simulation-based study. Submitted to ISMRM 2022.

We suggest other extensions of the three-compartment model to more accurately describe brain gray matter tissue.

In Chapter 5, we employ LFI methods to solve our non-linear inverse problem under a probabilistic framework and determine the posterior distribution of the fitted parameters. Such approach offers a full description of the solution landscape and can point out degeneracies, as opposed to the usual deterministic least-squares based solution (Jelescu et al., 2016; Novikov et al., 2018c). This work was accepted as an oral presentation at ISMRM 2021 and awarded a "Magna Cum Laude" prize (abstract ranked in the top 15% within the diffusion MRI category). It also got accepted at

IPMI 2021 as an oral presentation. An extension of this work has been submitted to the journal MELBA.

> **M. Jallais**, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Diffusion MRI-Based Cytoarchitecture Measurements in Brain Gray Matter using Likelihood-Free Inference. In ISMRM 2021.
>
> **M. Jallais**, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Cytoarchitecture Measurements in Brain Gray Matter using Likelihood-Free Inference. In International Conference on Information Processing in Medical Imaging 2021 (pp. 191-202). Springer, Cham.
>
> **M. Jallais**, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Inverting brain gray matter models with likelihood-free inference: a tool for trustable cytoarchitecture measurements. Submitted to MELBA (arXiv preprint arXiv:2111.08693).

In Chapter 6, we present our concluding remarks and discuss future perspectives for what we have developed.

To foster reproducible research, `Python` code of our implementation to solve the inverse problem for both models is available online on the public repository:

`https://github.com/mjallais/SBI_dMRI`

Part I

Theoretical Background

2 Single Encoding Diffusion MRI: a Probe to Brain Microstructure

Contents

Acronyms used in the chapter

ADC	Apparent Diffusion Coefficient
CSF	Cerebropinal Fluid
CTI	Correlation Tensor MRI
DKI	Diffusion Kurtosis Imaging
dMRI	diffusion MRI
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Image
EAP	Ensemble Average Propagator
ECS	Extra-Cellular Space
FA	Fractional Anisotropy
GPD	Gaussian Phase Distribution
HARDI	High Angular Resolution Diffusion Imaging
KM	Kärger Model
LEMONADE	Linearly Estimated Moments provide Orientations of Neurites And their Diffusivities Exactly
MAP-MRI	Mean Apparent Propagator-MRI
MD	Mean Diffusivity
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
NODDI	Neurite Orientation Dispersion and Density Imaging
ODF	Orientation Distribution Function
PGSE	Pulsed Gradient Spin Echo
SANDI	Soma And Neurite Density Imaging
SM	Standard Model
SNR	Signal Noise Ratio
WMTI	White Matter Tract Integrity

Notations used in the chapter

δ	Duration of a gradient pulse [ms]
Δ	Duration which separates two gradient pulses [ms]
t	Diffusion time [ms]
b	b-value [$\mathrm{ms}\,\mu\mathrm{m}^{-2}$]
q	q-value [$\mu\mathrm{m}^{-1}$]
$\mathbf{g}$	Direction and strength [$\mathrm{T\,m}^{-1}$] of the applied diffusion weighting
$\hat{g}$	Unit vector of $\mathbf{g}$
S	Diffusion-weighted MRI signal
$\bar{S}$	Powder-average of the diffusion-weighted MRI signal
E	Attenuation of the diffusion-weighted MRI signal
D_X	Diffusivity within compartment X [$\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$]
K	Kurtosis
f_X	Signal fraction of compartment X
r_s	Soma radius [$\mu\mathrm{m}$]
$\mathcal{K}$	Response kernel
t_{ex}	Exchange time [ms]
$\mathcal{P}$	Fiber ODF
τ_X	Mean molecular lifetime in compartment X [ms]
r_X	Exchange rate from compartment X to another [$1/\mathrm{ms}$]

This chapter is meant as an introduction to this thesis, in the context of probing human brain microstructure using single encoded diffusion Magnetic Resonance Imaging (MRI). We will start by illustrating how diffusion MRI is a perfectly adapted technique to measure tissue microstructure in the human brain using water motion, followed by a biological presentation of human brain. The non-invasive imaging technique based on water motions known as diffusion MRI will be further presented, along with the difficulties that come with it. Within this context, we will first review and discuss methods based on signal representation that enable us to get an insight into microstructure anisotropy. Then, we will outline methods based on modeling, which are state-of-the-art methods to get parameter estimations of the human brain tissue, adapted for both white and gray matter.

2.1 Accessing Brain Microstructure using Diffusion MRI

2.1.1 Introduction

Brain diseases significantly impact its microstructure. Dementia and cognitive deficits appear to be primarily related to neuronal and synaptic connectivity loss (Douaud et al., 2013). Brain diseases such as ischemia, multiple sclerosis, trauma, or brain tumors, have also been reported to cause structural changes in the microstructure (see, e.g., Assaf and Pasternak, 2008; Soares et al., 2013). Panagiotaki et al. (2014) notably used neuron size growth as a marker to study the evolution of tumor cell size in response to a drug. Although these diseases' external impact and progression are readily observable, accessing microstructural changes in the brain remains a challenge, making it difficult to understand these diseases and develop treatments. With technological advances, diffusion MRI (dMRI) has emerged as a novel method to study brain microstructure non-invasively and in-vivo.

Molecules suspended in a fluid are constantly moving, following random movements due to their interactions with each other. If nothing impedes their movements, molecules can freely diffuse (i.e., isotropically). In the brain, molecules' movements get restricted by the boundaries of the underlying microstructure of tissues. Diffusion-weighted Magnetic Resonance (MR) imaging is a non-invasive technique that quantifies how water molecules diffuse in the brain. During a typical acquisition of tens of milliseconds, water molecules displace by diffusing up to tens of micrometers, making diffusion MRI sensitive to a wide range of microstructural and physiological parameters in the tissue. An ultimate goal of an MR diffusion theory is to relate these parameters quantitatively to the diffusion-weighted signal, i.e.,

estimate the microstructure of the brain given the molecules' displacements. However, this task appears to be complicated as deducing those parameters constitutes a complex inverse problem requiring careful modeling of the diffusion signal over a wide range of acquisition parameters.

2.1.2 Microstructure as Reflected by the Molecules' Motions

Particles suspended in a fluid are constantly undergoing small random movements, known as Brownian motion (Brown, 1828). Diffusion can be considered a macroscopic manifestation of this Brownian motion on the microscopic level (Einstein et al., 1905). When no barrier impedes diffusion preferentially in one direction over another, molecular displacements are equal in all directions. This is known as isotropic diffusion. However, molecules' movements are hindered in the brain, including by cell membranes. Diffusion is then not equal along all directions anymore and becomes anisotropic. The distance traveled by a water molecule depends on its interactions, in particular with other molecules and cell membranes. Certain geometric characteristics of the underlying structure at the microscopic level can therefore be inferred from the molecule movements (Kac, 1966).

2.1.3 Structural Brain Anisotropy

The mammalian brain can be decomposed in two main parts: white matter and gray matter (Figure 2.1). The former designates regions that contain mainly long-range myelinated axons and relatively few cell bodies. Axons are connections that cross the brain, bridging different parts of gray matter. This one contains mainly cell bodies called somas, connected by neurites and relatively few myelinated axons (Figure 2.2). Axons and neurites have cylindrical shapes, with typical diameters between 0.1-10 μm in humans (see e.g. Beaulieu, 2014; Veraart et al., 2020). Water molecules within those fibers move on average further along them than across them due to their small diameter. Soma radii range between 2 and 30 μm (Palombo et al., 2021).

The presence, or absence, of anisotropy in the brain provides key information about the tissue structure at the cellular level. The strong anisotropy in white matter, which is mainly due to the axons, encouraged its wide study over the past decades. A modeling technique of axonal connections in the brain, known as tractography, notably emerged (see e.g. Jeurissen et al., 2019). The more complex tissue structure and weaker anisotropy in gray matter make its analysis harder. The greater isotropy in gray matter can be explained by multiple factors. The presence of numerous somas whose shapes resemble spheres (see Section 2.4.1) can partly explain this isotropy

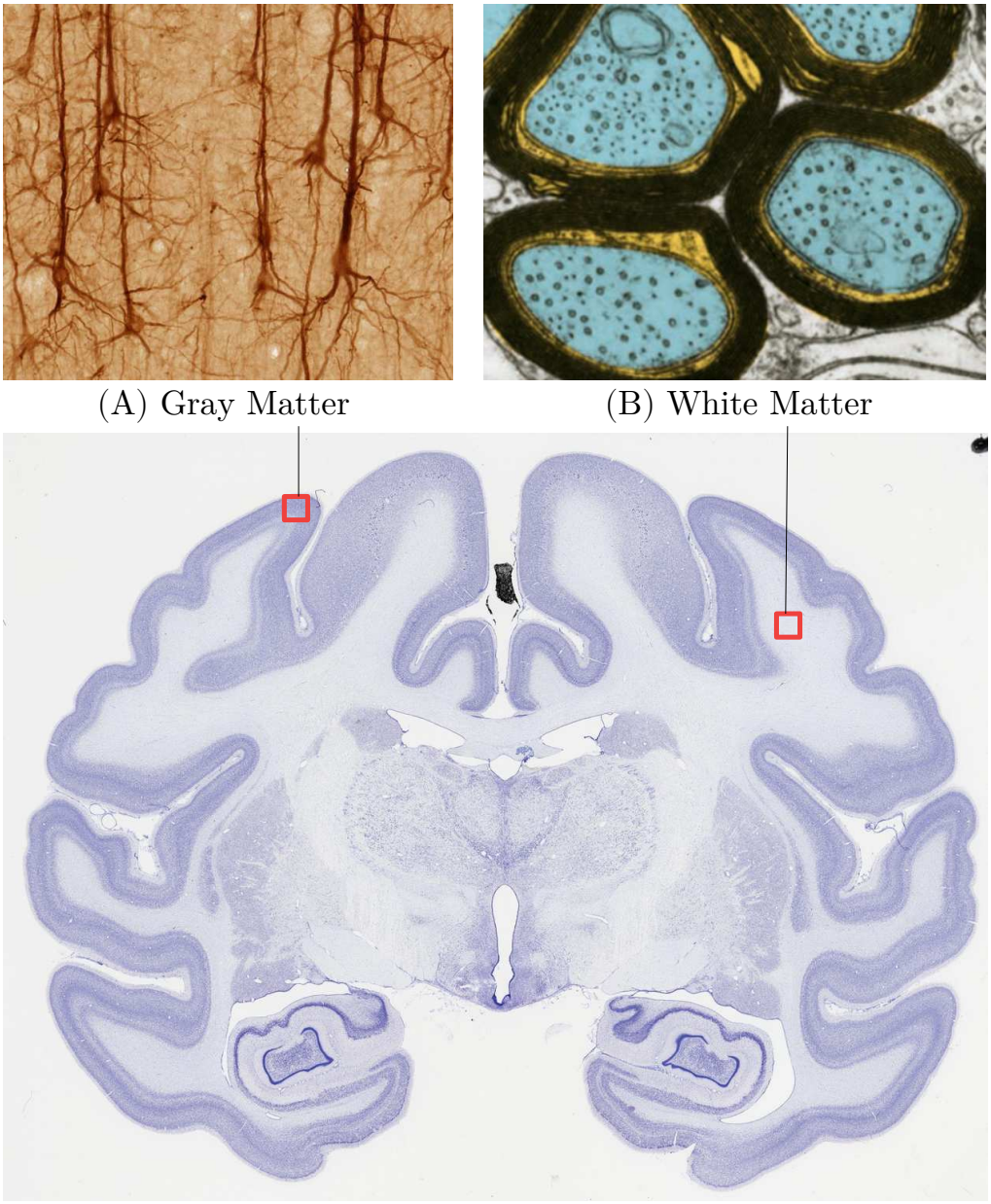


Fig. 2.1: White matter and gray matter decomposition of the brain. Top-left: SMI32-stained pyramidal neurons (brainmaps.org). Top-right: Transverse section of axons (Purves, 2004). Bottom: Nissl-stained coronal slice of human brain adult (brainmaps.org).

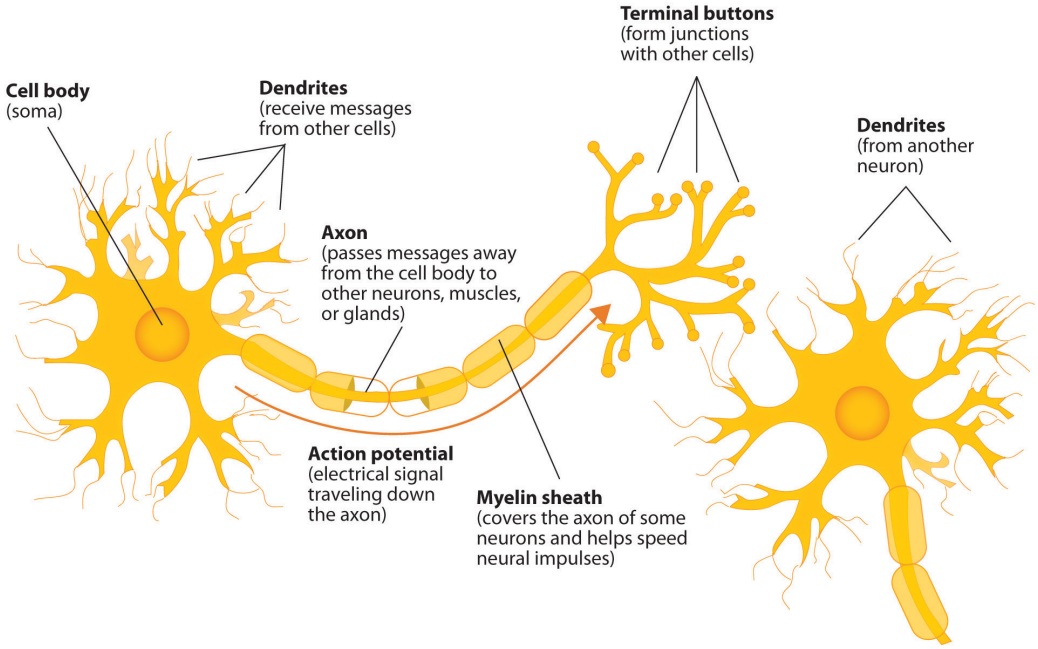


Fig. 2.2: Diagram of the components of a neuron (Walinga, 2010).

(Ligneul et al., 2019; Palombo et al., 2020). The microstructural organisation of neurites around somas is another possible factor (see Figure 2.1 A). Myelin also appears to modulate the degree of diffusion anisotropy between axons and processes (and so between white matter and gray matter) (Beaulieu and Allen, 1994). Indeed, myelin renders axons more impermeable, preventing molecules to "escape". The absence of myelin around most neurites makes exchanges between them and their surrounding environment possible, leading to a loss of anisotropy (Jelescu et al., 2021).

Note that anisotropy is not only a property of neural fibers. Anisotropy has also been observed in liquid crystals, muscles and other tissues, even in fruits and vegetables (Beaulieu, 2014).

2.2 Diffusion MRI: Introduction to a Non-Invasive Imaging Technique

Nuclear Magnetic Resonance Imaging allows to non-invasively study brain microstructure in-vivo. In this section, we present the basics of this technique, with a focus on the Pulsed Gradient Spin Echo (PGSE) sequence.

2.2.1 Diffusion MRI Acquisition Sequence

Consider an MRI acquisition sequence, as illustrated in Figure 2.3. After slice selection, all the nuclei on this plane are precessing at the same frequency. To obtain a diffusion MR image, two gradient pulses are added to the acquisition sequence, as presented in Figure 2.4. The first applied pulse is making the particles go off phase. We then apply a second gradient with the same strength in the opposite direction, during the same amount of time. If molecules stayed still between those two gradients, they would have all come back to their original phase, the two gradients canceling each other. However, after turning the first gradient on, molecules are moving randomly (Brownian motion). After a certain evolution time, if molecules are not at the same location, the second gradient causes destructive interference, which results in a loss of signal. The ratio between the signal obtained with diffusion gradients and the one without them quantifies the amount of ongoing diffusion. The further a molecule travels from its initial position during the time between the two diffusion gradients along the gradient direction, the greater signal attenuation we get. The time we let to the molecules for diffusing (called diffusion time, see Section 2.2.2) also impacts the resulting signal attenuation and the structural information we can measure. If the diffusion time was extremely short, only the local intrinsic diffusivity of the fluid, i.e. the rate at which particles can spread, would be measured. The hindrance effects would only become apparent at longer times. The objective is to deduce the structure of the medium where the water molecules are trapped in from those signal losses.

2.2.2 Mathematical Foundations

Stejskal and Tanner invented in 1965 the Pulsed Gradient Spin Echo (PGSE) sequence (Tanner and Stejskal, 1968) to measure diffusion (Figure 2.4). In this sequence, two opposite diffusion gradients are applied during a time δ, separated by an interval Δ. The diffusion-weighting is globally encoded by the b-value (Le Bihan and Breton, 1985), and reflects the strength and

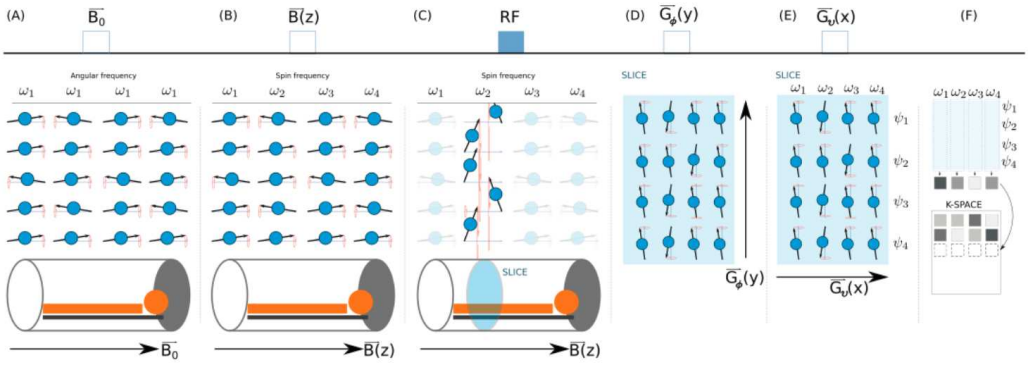


Fig. 2.3: Simplified scheme of a dMRI acquisition (Gallardo Diez, 2018). (A) A homogeneous magnetic field $\overrightarrow{B}_0$ is constantly emitted by an MRI scanner, making atomic nuclei precess around it. (B) A gradient magnetic field is emitted to change in a predictable way the frequency of the nuclei. (C) A radio frequency pulse is generated to select a plane, making certain nuclei resonate and moving their precession to the transversal plane. (D) A frequency encoding gradient is briefly applied to put out of phase nuclei from the selected plane in different vertical positions. (E) A phase encoding gradient is then turned on, to change in a predictable way the frequency in every horizontal position. (F) The obtained acquisition is defined in the k-space.

timing of the gradients used to generate the diffusion-weighted images. This factor is computed as follow:

$$b = \gamma^2 g^2 \delta^2 (\Delta - \delta/3) \, , \tag{2.1}$$

where γ ($\mathrm{MHz\,T^{-1}}$) is generally the nuclear gyromagnetic ratio of the water proton ^1H and g is the strength of the diffusion gradient ($\mathrm{T\,m^{-1}}$). In the following sections, $\mathbf{g}$ encodes the direction of the applied diffusion weighting in addition to its strength ($g = ||\mathbf{g}||$), and $\hat{\mathbf{g}}$ is the corresponding unit vector ($\hat{\mathbf{g}} = \mathbf{g}/||\mathbf{g}||$).

The quantity $E(b) = S(b)/S_0$ expresses, for each voxel, the attenuation of the diffusion-weighted signal along the selected gradient direction, S_0 being the image acquired without diffusion gradients. In the absence of restrictions (free diffusion), the signal attenuation can be expressed as:

$$E(b) = e^{-bD} \, , \tag{2.2}$$

with D the diffusion coefficient of the excited molecules.

If δ is assumed to be infinitely narrow, i.e. the diffusion during that time is negligible, the signal attenuation can be related to the Ensemble Aver-

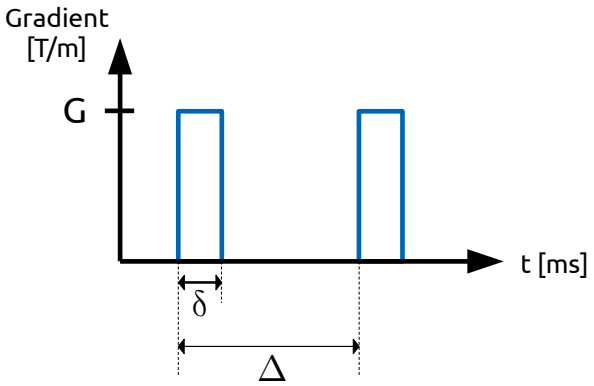


Fig. 2.4: Schematic of a PGSE sequence. Two pulses of gradient strength G are applied during δ ms in opposite directions, separated by Δ ms.

age Propagator (EAP) $P(\mathbf{r}, t)$ via a Fourier relationship under the q-space formalism (Tanner and Stejskal, 1968; Callaghan et al., 1991):

$$E(\mathbf{q}, t) = \frac{S(\mathbf{q}, t)}{S_0} = \int_{\mathbb{R}^3} P(\mathbf{r}, t) e^{-2\pi i \mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \,, \tag{2.3}$$

where $\mathbf{q}$ is the wave vector and t the diffusion time, which, for the PGSE sequence, are expressed as

$$\mathbf{q} = \gamma \delta g / 2\pi \text{ and } t = \Delta - \delta/3 \,. \tag{2.4}$$

The diffusion time t expresses the time interval during which spins are allowed to diffuse before measurement. By increasing the spatial frequency $q = ||\mathbf{q}||$ it is possible to achieve a higher spatial resolution of $P(\mathbf{r}, t)$ in the displacement space described by $\mathbf{r}$.

2.2.3 Acquisition Strategies

Experimental parameters, and especially q and t, influence the diffusion signal attenuation along different gradient directions, and therefore the estimation of diffusion anisotropy. Ideally, many gradient directions, q-values, and diffusion times would be required to completely characterize diffusion anisotropy in a tissue. In practice, the sampling strategy depends on the application and on the chosen signal representation. This way, only one shell of gradient directions is usually used in Diffusion Tensor Imaging (DTI) (see Section 2.3.1). Also using only one shell at a higher b-value and more directions, are the High Angular Resolution Diffusion Imaging (HARDI) schemes, which aim at increasing the angular resolution of the diffusion signal with the intent of resolving crossing tissue configurations (Tuch et al., 2002). Different diffusion-weightings signal acquisitions are also needed for some signal representations. In that case, multi-shell acquisitions are set up using

different q-shells with fixed diffusion time. Each shell represents a collection of samples in the three-dimensional space with the same q-value. An optimal spatial coverage is important to measure the diffusion signal as efficiently as possible. Expansions have been proposed such that all the acquired samples lie on different non-collinear directions (Caruyer et al., 2013). This multi-shell design can be extended to t-shells, called qt acquisitions (Fieremans et al., 2016) in order to exploit different values for both q and t. In that case, a complete q-shell scheme is acquired for each desired diffusion time. Ning et al. (2015) review and compare 16 reconstruction algorithms (single and multi-shells) to help determine an appropriate acquisition protocol (number of b-values) and the analysis method to use for a particular neuroimaging study.

2.2.4 Difficulties

A main drawback to take into consideration is inherent to the dMRI acquisition process. Due to the acquisition device limitations and the mesoscopic size of neurons, one voxel, at the macroscopic scale, includes thousands of somas and processes. This means that the acquired signal is an average of the signal coming from all those cells. Several issues have then to be considered.

First, the acquired signal in a voxel will be an average of the signal of all the diffusing molecules within this voxel, which could correspond to not less than 3000 axons in white matter. Features that will be computed from it, such as anisotropy, will be an average of all the components in the tissue. One needs to note that every tissue is made of several compartments and that the signals from each of these compartments where water molecules are present are averaged. Investigations using diffusion-weighted spectroscopy, an imaging technique with increased cellular specificity, are also led to trying and targeting specific compartment(s) (Palombo et al., 2016; Ligneul et al., 2019). This average problem leads to a second issue: a small change in anisotropy (or other features) can actually reflect greater pathological differences. It means that there needs to be a big change in the voxel to be able to detect it in the acquired diffusion signal. The third issue is that anisotropic cellular elements might be considered isotropic due to the tree pattern of processes within gray matter (Jespersen et al., 2007) or to crossing fibers in white matter. At least, as expected from an acquisition, the signal is noisy. Low concentration of water molecules in some tissue (and thus long scan times) can lead to a poor signal-to-noise ratio.

In addition to those issues, we must recall that the spacing between axons, axon diameter, myelin thickness, etc are all also variables, even within the

same tract, which adds to the complexity of the problem. The barriers to diffusion have also not a simple nor regular geometry. The correspondence between the biological features of the tissue and the non-invasive diffusion measure is therefore not straightforward.

2.2.5 Measuring Tissue Microstructure using Diffusion MRI

We aim at getting information about human brain structure using dMRI as a non-invasive probe. The acquired diffusion signal is a mixture of the diffusion signals coming from each compartment (Panagiotaki et al., 2012), and is therefore modulated by the geometry of the tissue microstructure. Relevant information to infer from it is soma diameters, soma and process densities, and diffusivities. Two complementary approaches have emerged for extracting this information about the tissue microstructure from the diffusion signal: signal representation and tissue modeling (denomination from Novikov et al., 2018b).

Signal representations are model-independent mathematical expressions. It can be coefficients of a decomposition on a specific basis for example. These coefficients do not carry any particular physical meaning. Representations can be used to store, compress or compare measurements. An infinite way to represent a continuous function exists. In practice, one chooses a representation according to the need of a particular neuroimaging study (Ning et al., 2015). Although signal representations are suited for all kinds of tissues, they lack specificity and provide only an indirect characterization of the microstructure.

Biophysical tissue models rely on a schematic geometry of the underlying tissue. They are pictures representing a physical reality relying on assumptions meant to simplify the complexity of a biological tissue. A good model only keeps relevant features that characterize the tissue and discards irrelevant degrees of freedom. The designed analytical expression is then fit to the diffusion data in order to estimate these relevant features of the microstructure. This advantage of providing greater specificity and interpretation of biologically-relevant parameters appears to be the weakness of the method. Indeed the initial geometric assumption must be chosen as to accurately capture all of the features of the tissue that effectively impact the diffusion signal in a given acquisition range (Novikov et al., 2018b), but we also must be able to mathematically solve this inverse problem. Model validations are important because a wrong model could lead to wrong interpretations of a physical phenomenon.

Techniques from these two approaches, signal representations and tissue modeling, will be reviewed respectively in Sections 2.3 and 2.4.

2.3 Quantifying Tissue Microstructure via Signal Representation

Signal representation is an indirect method that aims at describing the diffusion signal with no assumptions about the underlying structure. It can therefore be applied to healthy or diseased tissues. Several methods are described, with an emphasis made on the cumulant expansion, which is the most widespread signal representation.

2.3.1 Cumulant Expansion

Common signal representations are based on the cumulant expansion (Kiselev, 2010; Jensen et al., 2005), which corresponds to a development of the logarithm of the signal in polynomials up to a given order in b:

$$\ln\left(\frac{S(b)}{S_0}\right) = -bD + \frac{1}{6}(bD)^2K + \dots \tag{2.5}$$

where D is the diffusion coefficient and K the kurtosis. This formula can also be written in the tensor form :

$$\ln\frac{S(b)}{S_0} = -b\sum_{i,i=1}^{3} g_i g_j D_{ij} + \frac{1}{6}(b\bar{D})^2 \sum_{i,j,k,l=1}^{3} g_i g_j g_k g_l W_{ijkl} - \dots \ , \tag{2.6}$$

where $\mathbf{D}$ is the rank-2 diffusion tensor, $\mathbf{W}$ is the rank-4 kurtosis tensor, $\bar{D}$ is the mean diffusivity and $\mathbf{g}$ is the direction of the applied diffusion weighting (see Section 2.2.2).

An expansion in moments, which corresponds to a Taylor expansion of the signal, is also possible. While expansions in moments and in cumulants are mathematically equivalent, for a similar order truncation at some fixed (low) order, the cumulant series provides a more accurate estimation of the dMRI signal than a moment expansion. Moment expansion is more optimal for analytical treatments because contributions from different tissue compartments add up. A combinatorial relation exists between the two expansions (Mayer and Montroll, 1941; Kiselev, 2010). Computing the cumulant tensors and converting them into moments is promoted to be the most numerically stable methodology to adopt (Novikov et al., 2018c).

One of the most popular MRI techniques in brain research as well as in clinical practice is Diffusion Tensor Imaging (DTI) (Basser et al., 1994), based on the cumulant expansion up to the first order in b. This technique is valid for low diffusion weighting ($b \ll (DK)^{-1}$). Note that this technique does not assume that the medium is homogeneous with unrestricted diffusion (K=0),

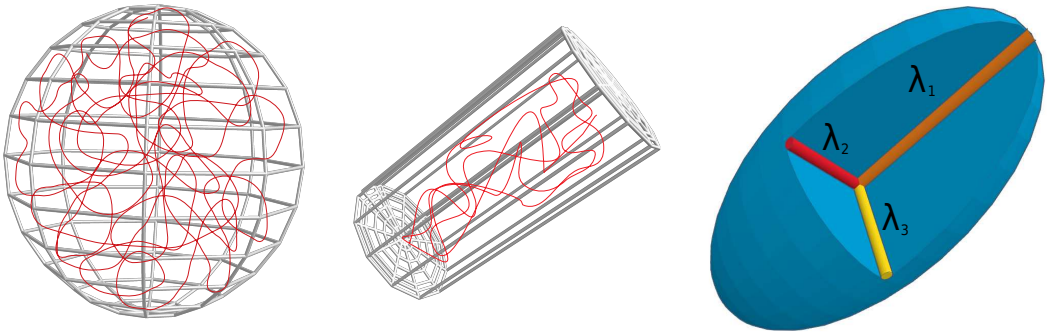


Fig. 2.5: Isotropic diffusion in somas can be modeled by a sphere (left). Anisotropic diffusion in neurites can be represented by an ellipsoid reflecting axial (λ_1) and radial (λ_2 and λ_3) diffusion. This image has been inspired by the book chapter written by Christian Beaulieu (Beaulieu, 2002).

which appears to be not true for most biological tissues, but that it follows a Gaussian law when $b \ll (DK)^{-1}$.

The three eigenvalues λ_1, λ_2 and λ_3 can be computed from the diffusion tensor **D**, that reflect axial and radial diffusivity of molecules within fibers and in the extra-cellular space (see Figure 2.5). They allow to define variables such as Mean Diffusivity (MD) or Fractional Anisotropy (FA) (Basser et al., 1994), as follows:

$$\begin{aligned} MD &= \frac{\lambda_1+\lambda_2+\lambda_3}{3} \\ FA &= \frac{\mathrm{std}(\lambda)}{\mathrm{rms}(\lambda)} = \sqrt{\frac{1}{2}}\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{\lambda_1^2+\lambda_2^2+\lambda_3^2}} \end{aligned} \tag{2.7}$$

Those two measures are complementary, as they bring different information to the comprehension of a tissue (Figure 2.6). Hofstetter et al. (2013) used MD to hypothesize the presence of bigger cells in the brain after a learning session. FA is seen as a potential biological marker for diseases affecting the brain (Strimbu and Tavel, 2010). Beaulieu (2002) also investigated anisotropy in the human brain gray matter using DTI. We refer the reader to Fick et al. (2017), who dedicated a review on more existing diffusion anisotropy metrics, including Generalized Fractional Anisotropy (GFA) (Tuch, 2004), Propagator Anisotropy (PA) (Özarslan et al., 2013), Orientation

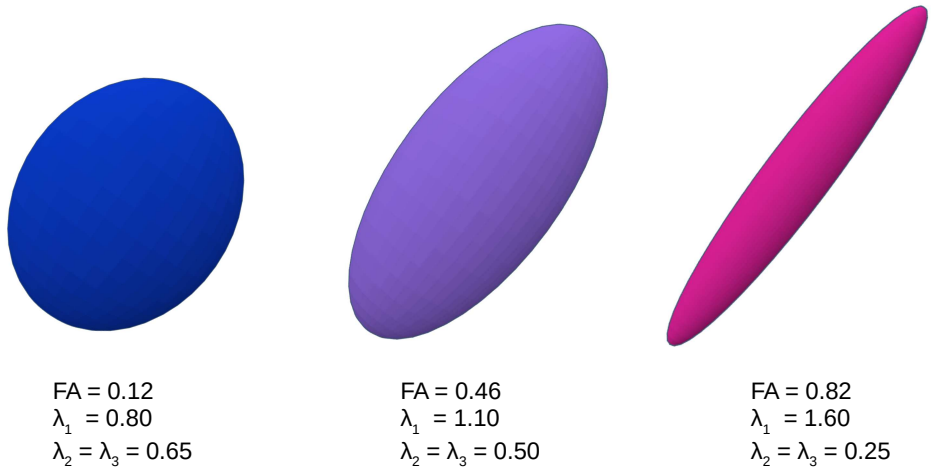


Fig. 2.6: Mean diffusivity (MD) and fractional anisotropy (FA) are two complementary measures. Here are three examples of ellipses ranging from isotropic to anisotropic that have the same mean diffusivity ($0.7\text{x}10^{-3}\text{mm}^2\text{s}^{-1}$). This image has been inspired by the book chapter written by Christian Beaulieu (Beaulieu, 2002).

Dispersion Index (ODI) (Zhang et al., 2012), and microscopic Fractional Anisotropy (μFA) (Kaden et al., 2016).

Diffusion Kurtosis Imaging (DKI) goes beyond DTI and its first-order expansion by also estimating the kurtosis of the diffusion probability distribution function (Jensen et al., 2005). The kurtosis quantifies the non-Gaussianity of a distribution. The information that it provides is complementary to DTI metrics. Fitting the kurtosis tensor significantly improves the accuracy of the diffusion tensor estimation (Veraart et al., 2011). In the same way, extending the series to the sixth order cumulant (in b^3) also increases the accuracy of the kurtosis estimation, albeit with a penalty on precision.

In order to estimate the six independent components of the diffusion tensor, the minimal required data is one b=0 (unweighted) image and six non-collinear directions on a single diffusion weighting, or "shell". The additional estimation of the 15 independent components of the kurtosis tensor requires a minimal acquisition of one b=0 image and one or two nonzero shells with 15 non-collinear gradient directions, so that a total of 22 diffusion-weighted images are acquired (Veraart et al., 2011). The choice of the shell b-values is a trade-off between accuracy and precision. The b-values should be as low as possible to respect the validity of the cumulant expansion, but slightly higher values enable to limit the impact of noise (Jones, 2010). Jelescu and Budde (2017a) suggest a typical value around b=1 ms μm^{-2} for DTI and 2

ms μm^{-2} for DKI in vivo. For further details on the optimization of acquisition parameters for precise measurement of diffusion in anisotropic systems, we invite the reader to have a look at the work of Jones et al. (1999).

Although not relying on single diffusion encoding, a work worth being reported is the Correlation Tensor MRI (CTI) approach developed by Henriques et al. (2020). Based on a double diffusion encoding acquisition, this technique pushes the kurtosis estimation further by proposing a method to disentangle the different non-Gaussian sources. The kurtosis is decoupled into three components, namely isotropic, anisotropic, and restricted kurtosis, which allow quantifying the degree of anisotropy in the tissue, the variance of mean diffusivities, and the degree of non-Gaussian effects induced by structural disorder and restricted diffusion in compartments with reflecting walls.

2.3.2 Other Representations

Yablonskiy et al. (2003) hypothesize that the acquired diffusion signal is a sum of signals originating from many spin packets, present in different cell types, at different positions. Each spin packet having then different trajectories and facing different hindrances, they make the assumption that they all have a different Apparent Diffusion Coefficient (ADC). Hence, they introduced a distribution of diffusion coefficients $\rho(D)$, and expressed the diffusion signal as follows:

$$\frac{S}{S_0} = \int_0^\infty \rho(D) e^{-bD} dD \tag{2.8}$$

Theoretically, the distribution of diffusion coefficients can be estimated using the inverse Laplace transform. In practice, some functional form needs to be assumed for $\rho(D)$ due to the mathematical ill-conditioning of the inverse Laplace transform. In addition, a very strong diffusion-weighted regime is needed for the estimated distribution to accurately reflect the tissue distribution of diffusion coefficients (Kiselev, 2017; Novikov and Kiselev, 2010).

Jian et al. (2007) propose a statistical method to infer connectivity patterns based on the characterization of the water molecule diffusion by a continuous distribution of diffusion tensors. They described the MR signal attenuation as the Laplace transform of this probability distribution defined on the manifold of symmetric positive-definite tensors. Combined with a spherical deconvolution approach, displacement probability functions and distinct fiber orientations can be estimated in each voxel in a HARDI dataset.

The multi-shell Mean Apparent Propagator-MRI (MAP-MRI) method, as proposed by Özarslan et al. (2013), expands the signal using harmonic oscillator basis functions. It represents the diffusion-weighted signal by an anisotropic Gaussian modulated by a series of Hermite polynomials. This method allows the estimation of three-dimensional EAP, where both restricted (non-Gaussian) diffusion and crossing axons can be represented. However, according to Ning et al. (2015), this method fails to estimate crossing angles correctly. The propagator anisotropy (PA) metric was derived from this method, which is a measure of dissimilarity between the reconstructed EAP and its closest isotropic approximation EAP.

Hanyga and Magin (2014) proposed a new space-fractional diffusion model based on an anomalous anisotropic diffusion equation that preserves positivity.

Other representations exist, but have not been included in this chapter.

2.3.3 Limitations

The validity and therefore the usefulness of the cumulant expansion depends on its convergence towards the acquisition signal, characterized by the convergence radius b_c (Kiselev and Il'yasov, 2007). If $b < b_c$, the series can be approximated using a couple of low terms in Equation 2.6, higher-order terms being flooded by the noise, i.e. small contributions to the signal can not be decoupled from the noise in experimental data. The number of parameters to estimate is then reduced, but a good accuracy does not assure its validity. Otherwise, if $b > b_c$, the series in Equation 2.6 diverges which means that the model function cannot be reduced to a polynomial. A good quality fitting gives then more credit to the underlying model.

Hutchinson et al. (2017) compare the DTI, DKI, MAP-MRI, and Neurite Orientation Dispersion and Density Imaging (NODDI) (see Section 2.4.2.1) methods in different experimental conditions to study the influence of noise and sampling (among others) on parameter estimations. All methods proved to be influenced by the acquisition parameters such as the b-values, the resolution, the Signal Noise Ratio (SNR), and the diffusion time. The need for DKI to fit a higher-order tensor explains its high sensitivity to noise.

Regional issues are also to be noted, related to crossing fibers, which can be detected as isotropic zones (Alexander and Seunarine, 2010). Indeed, several diffusion directions are possible in that case. The angular resolution needs to be high enough and the model designed to take this particular case into account.

2.4 Biophysical Modeling to Measure Tissue Microstructure

This second approach is based on a biophysical model designed for a particular tissue geometry. This model is fit to the diffusion signal acquired, which allows the estimation of the relevant parameters of the microstructure. While it can provide a greater specificity of biological parameters, the design of the model remains difficult, as it needs to accurately capture all the features that effectively and substantially impact the diffusion signal in a given acquisition range (the coarse-graining problem, see Novikov et al., 2018a).

Another big challenge of this approach comes from the number of unknowns to estimate after the definition of all effective parameters. To estimate them all we would need a lot of different b-values. This is unfeasible in clinical applications because first, the gradients used in clinical MRIs are not strong enough, and secondly, it would require a patient to stay in the MRI device for a very long time. Some methods rely on constraints to bypass this problem, as presented in Section 2.4.2.1.

2.4.1 Multi-compartmental Model

Tissue in the brain can generally be decomposed into four compartments. The first one corresponds to the somas, which are the central part of the neurons, mainly present in gray matter. They can be modeled as spheres of different diameters. Neurites, the second compartment, connect those neurons together, either in short distances in gray matter (they are called dendrites), or long-distance connections in white matter (axons). The diffusivity across the processes is considered zero due to the restriction implied by the fixed small diameter. Processes can therefore be modeled by cylinders with zero-radius ("sticks") (see Section 2.4.1.1 below). The orientation of a collection of processes within a voxel is characterized by an Orientation Distribution Function (ODF) (Tuch, 2004). The third compartment corresponds to Extra-Cellular Space (ECS) and is modeled as Gaussian anisotropic. Glial cells are usually comprised in the soma compartment, as done by Palombo et al. (2016). However, their possible high exchange rate with the extracellular space is still a matter of discussion and this argument would argue in favor of their better modeling in the ECS compartment (Fields et al., 2015). The last one is the Cerebropinal Fluid (CSF), which could contribute if a voxel contains part of a ventricle, and corresponds to free diffusing molecules. It is hence modeled as free diffusion.

The acquired water signal originates from these four compartments. Under the hypothesis that each compartment is perfectly impermeable (i.e., no molecule moves from one compartment to another), this signal is weighted according to their relative signal fraction f:

$$S(\mathbf{b}) = f_{\text{somas}} \cdot S_{\text{somas}}(\mathbf{b}) + f_{\text{neurites}} \cdot S_{\text{neurites}}(\mathbf{b}) + f_{\text{ECS}} \cdot S_{\text{ECS}}(\mathbf{b}) + f_{\text{CSF}} \cdot S_{\text{CSF}}(\mathbf{b}) \ , \tag{2.9}$$

with $f_{\text{somas}} + f_{\text{neurites}} + f_{\text{ECS}} + f_{\text{CSF}} = 1$. Remark that f_{somas}, f_{neurites}, f_{ECS} and f_{CSF} are not the relative volume fractions due to the T2 differences between the compartments (Novikov et al., 2018a). In a non-negligible exchange hypothesis, interactions between the compartments are more complex and require further modeling (see Section 2.4.3.2).

In the following models presented, a combination of those compartments is used to model particular tissues and keep only the relevant compartments.

2.4.1.1 Neurites as Sticks

Neurites have been modeled by zero-radius impermeable cylinders, characterized by their longitudinal diffusivity $D_n^{\parallel}$, the transverse diffusivity $D_n^{\perp}$ being usually considered zero. These neurites are called "sticks" and correspond to the most anisotropic Gaussian compartment possible (see e.g. Behrens et al., 2003; Kroenke et al., 2004; Jespersen et al., 2007).

The intra-neurite response function, i.e. the diffusion signal from water inside a stick of diffusivity $D_n = D_n^{\parallel}$ pointing in the unit direction $\hat{\mathbf{n}}$, is defined as :

$$G_{\hat{n}}(\hat{\mathbf{g}}, b) = \mathrm{e}^{-bD_n(\hat{\mathbf{g}}\cdot\hat{\mathbf{n}})^2} \ , \tag{2.10}$$

with $\hat{\mathbf{g}}$ being the unit gradient direction of the measurement. It is determined by $\cos\theta \equiv \hat{\mathbf{g}} \cdot \hat{\mathbf{n}}$, where θ is the angle between $\hat{\mathbf{n}}$ and $\hat{\mathbf{g}}$.

The signal, after being isotropically averaged over multiple gradient directions $\hat{\mathbf{g}}$ for large b-values, is the following (Callaghan et al., 1979; Joabsson et al., 1997; Jensen et al., 2016):

$$\bar{S} \simeq \beta \cdot b^{-1/2} \ , \tag{2.11}$$

with $\beta = \sqrt{\frac{\pi}{4}} \cdot f_{\text{neurites}} / (D_n)^{1/2}$.

In brain tissue, at sufficiently large b-values, the extra-axonal space signal is exponentially suppressed, its diffusivity being non-zero in any direction. The only remaining signal in the white matter comes from the axons (S_{neurites} in Equation 2.9), and follows the power law from Equation 2.11 (McKinnon et al., 2017; Veraart et al., 2019).

Veraart et al. (2020) proved that the radius of the axons can be estimated for very high b-values in white matter, where the transverse diffusivity $D_n^\perp$ is not considered null anymore. The direction-averaged Diffusion Weighted Image (DWI) signal then follows the following law:

$$\bar{S} \simeq \beta e^{-bD_n^\perp} \cdot b^{-1/2} \tag{2.12}$$

2.4.1.2 Somas as Spheres

Somas can be modeled as spheres of radius r_s and diffusivity D_s, following the Gaussian Phase Distribution (GPD) approximation (Balinov et al., 1993):

$$-\log \frac{\bar{S}_{\text{somas}}(\delta, \Delta, g)}{S_0} = \frac{2(\gamma g)^2}{D_s} \sum_{m=1}^{\infty} \frac{\alpha_m^{-4}}{\alpha_m^2 {r_s}^2 - 2} \cdot \left(2\delta - \frac{2 + e^{-\alpha_m^2 D_s(\Delta-\delta)} - e^{-\alpha_m^2 D_s \delta} - e^{-\alpha_m^2 D_s \Delta} + e^{-\alpha_m^2 D_s(\Delta+\delta)}}{\alpha_m^2 D_s} \right) ,$$

where α_m is the mth root of $(\alpha r_s)^{-1} J_{\frac{3}{2}}(\alpha r_s) = J_{\frac{5}{2}}(\alpha r_s)$, with $J_n(x)$ the Bessel functions of the first kind. In specific limit cases, this equation simplifies. In a narrow case pulse regime, i.e. in the limit of large r_s or small t, this equation reduces to a free diffusion case (Equation 2.2) (Murday and Cotts, 1968). In the Neuman (wide pulse) regime, i.e. when $D_s\Delta \gg r_s^2$ and $D_s\delta \ll 1$, the signal simplifies to (Balinov et al., 1993):

$$-\log \bar{S}_{\text{somas}}(\delta, g) = \frac{1}{5}(\gamma g \delta r_s)^2 . \tag{2.13}$$

2.4.1.3 Extra-Cellular Space as Free Diffusion

The extra-cellular space is approximated as Gaussian diffusion, i.e. an exponential diffusion signal with a scalar diffusion constant D_e, which reflects the molecular viscosity of the fluid. This diffusivity can be expressed with a perpendicular and a parallel component, noted $D_e^\perp$ and $D_e^\parallel$ respectively.

In some tissues, diffusivity in the ECS is not impeded in a particular direction, and is therefore considered as isotropic, i.e., $D_e^\perp = D_e^\parallel$. The powder-averaged approximation of the ECS signal is then expressed as:

$$-\log(\bar{S}_{\text{ECS}}(q)) = (2\pi q)^2 t D_e . \tag{2.14}$$

This approximation assumes that the ECS is fully connected.

2.4.2 Standard Model of Diffusion in Neural Tissue

The measured diffusion signal in brain is a sum of anisotropic compartments. It can be modeled as a convolution between a response kernel $\mathcal{K}$ from a

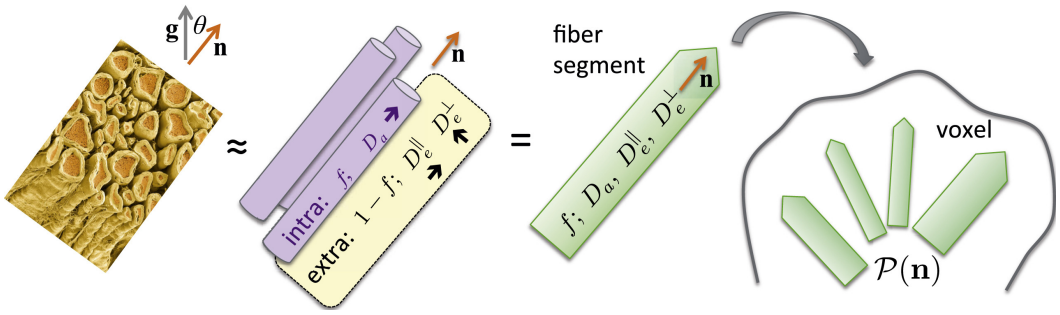


Fig. 2.7: Standard Model of diffusion in neuronal tissue. Two-compartment model (intra- and extra-neurite spaces) described by 4 independent parameters: f, D_n, $D_e^{\parallel}$ and $D_e^{\perp}$ and a fiber orientation distribution $\mathcal{P}(\hat{n})$. This figure is reproduced from Novikov et al. (2018c).

perfectly aligned fascicle pointing in the direction $\hat{\mathbf{n}}$ and the fiber ODF $\mathcal{P}(\hat{\mathbf{n}})$ normalized to $\int d\hat{\mathbf{n}}\mathcal{P}(\hat{\mathbf{n}}) \equiv 1$.

$$S_{\hat{\mathbf{g}}}(b) = \int_{|\hat{\mathbf{n}}|=1} \mathcal{P}(\hat{\mathbf{n}})\mathcal{K}(b, \hat{\mathbf{g}} \cdot \hat{\mathbf{n}})d\hat{\mathbf{n}} \,, \tag{2.15}$$

$\hat{\mathbf{g}}$ being defined in Section 2.2.2.

In the case of white matter, the kernel can be written as :

$$\mathcal{K}(b, \xi) = S_0 \left[f \mathrm{e}^{-bD_n\xi^2} + (1 - f - f_{\mathrm{CSF}})\, \mathrm{e}^{-bD_e^{\perp} - b\left(D_e^{||} - D_e^{\perp}\right)\xi^2} + f_{\mathrm{CSF}}\, \mathrm{e}^{-bD_{CSF}} \right] , \tag{2.16}$$

with $\xi = \hat{\mathbf{g}} \cdot \hat{\mathbf{n}}$. Those exponential contributions correspond to the intra-axonal space modeled by a stick compartment (Equation 2.10), the extra-axonal space modeled by an axially symmetric Gaussian compartment with transverse and longitudinal diffusivities $D_e^{\perp}$ and $D_e^{||}$, and the CSF compartment. All those compartments are represented in Figure 2.7. This decomposition has been widely used in white matter by the community. As a consequence, Novikov et al. suggested calling it the Standard Model (SM) (Novikov et al., 2018a). For the sake of reference, we will also refer to it as the SM in this chapter.

2.4.2.1 Standard Model Parameter Estimation using Constraints

In the previous sections, we presented the SM of diffusion in neural tissue as a sum of anisotropic Gaussian compartments, as defined by Novikov et al. (2018a). We will now introduce some methods based on the SM that rely on constraints to overcome the challenge of estimating many biological parameters of interest.

Neurite Orientation Dispersion and Density Imaging. In order to reduce the number of parameters that need to be estimated, Zhang et al. (2012)

proposed imposing restrictions on the intrinsic diffusivities. They introduced a method called NODDI, which relies on a three-compartment SM (intra-axonal space, extra-axonal space and CSF), described by seven parameters: volume fractions f_{intra} and f_{iso}, diffusivities $D_n^{\parallel}$, $D_e^{\parallel}$, $D_e^{\perp}$ and D_{iso}, and the orientation dispersion modeled by a Watson distribution of concentration parameter κ. By fixing the diffusivities to the following values:

$$D_n^{\parallel} = D_e^{\parallel} = 1.7\mu\mathrm{m}^2/\mathrm{ms} \tag{2.17}$$

$$D_e^{\perp} = (1 - f_{\text{intra}}) \cdot D_e^{\parallel} \tag{2.18}$$

$$D_{iso} = 3\mu m^2/\mathrm{ms} \tag{2.19}$$

only the two volume fractions and the orientation dispersion need to be estimated.

Although this method allows to estimate the parameters, the validity of those constraints needs to be questioned. To begin with, if we admit that the equalities are correct, they imply that a small deviation from the fixed values, as occurs in cerebral ischemia, will induce a non-negligible bias in the other parameters estimation, leading to false interpretations. However, studies using Diffusion-weighted spectroscopy MR which can quantify the diffusion of specific metabolites (e.g. Palombo et al., 2017), suggest, through the study of metabolites specifically found on different sections of the neurons and extra-cellular tissue, that such diffusivity is not constant across the whole brain. Whether and how these findings can be used to shed light on water diffusion in the brain, is an open question.

Another drawback of this method is that it leads to indetermination, which means that NODDI returns one possible result among a multiplicity of mathematical solutions by fixing $D_n^{\parallel} = D_e^{\parallel}$ (Jelescu et al., 2016; Novikov et al., 2018c). If we consider the case where all the parameter constraints are released and the CSF compartment neglected (called NODDIDA, which stands for NODDI with diffusivity assessment, Jelescu et al., 2015), two distinct solutions to the estimation problem exist: $D_n^{\parallel} > D_e^{\parallel}$ and $D_n^{\parallel} < D_e^{\parallel}$ (see Figures 8 and 9 in Jelescu et al., 2016). Both solutions lie within biologically plausible ranges, and determining which solution is biologically correct is an active field of research, although most studies are suggesting $D_n^{\parallel} > D_e^{\parallel}$. At least, the tortuosity approximation that relates $D_e^{\perp}$ and $D_e^{\parallel}$ has been invalidated in the case of tight packings of axons (Novikov and Fieremans, 2012).

White Matter Tract Integrity Metrics. Another approach to estimate the relevant features of interest in a tissue proposes to relate the scalar parameters to the DKI components. Called White Matter Tract Integrity (WMTI) (Fieremans

et al., 2011), it is a two-compartment SM that relies on the assumption that sticks are highly aligned within a voxel.

The tissue is described as a sum of two Gaussian compartments (intra- and extra-axonal space, Equation 2.9 with $f_{\text{somas}} = 0$ and $f_{\text{CSF}} = 0$), where axons are modeled as sticks embedded in a Gaussian anisotropic extra-axonal medium. Each compartment is characterized by a tensor (D_n and D_e) derived from the kurtosis tensors D and K. In any direction j:

$$D_j = f_{\text{intra}}\, D_{\text{n,j}} + (1 - f_{\text{intra}})\, D_{\text{e,j}} \,, \tag{2.20}$$

$$K_j = 3 f_{\text{intra}} \cdot (1 - f_{\text{intra}}) \frac{(D_{\text{e,j}} - D_{\text{n,j}})^2}{D_{\text{j}}^2} \tag{2.21}$$

We retrieve the two possible mathematical solutions mentioned before, as demonstrated by the square in Equation 2.21. The solution chosen in this method is $D_n^{\parallel} < D_e^{\parallel}$, which leads to:

$$f_{\text{intra}} = \frac{K_{\max}}{K_{\max} + 3} \,, \tag{2.22}$$

$$D_{\text{e,j}} = D_{\text{j}} \left[1 + \sqrt{\frac{K_j \cdot f_{\text{intra}}}{3\,(1 - f_{\text{intra}})}} \right] \,, \tag{2.23}$$

$$D_{\text{n,j}} = D_{\text{j}} \left[1 - \sqrt{\frac{K_{\text{j}}\,(1 - f_{\text{intra}})}{3 f_{\text{intra}}}} \right] \,. \tag{2.24}$$

Although WMTI enables to capture the changes of diffusivities, it has two main limitations. First, this approach is limited to regions of highly aligned single fiber bundles, which are only present in some white matter regions. Jespersen et al. suggested a method that alleviates this assumption by assuming a Watson distribution of the axons (like in NODDI) (Jespersen et al., 2018). Second, as it relies on the DKI decomposition, this method is only restricted to the low b-value regime, which could lead to some bias.

2.4.2.2 LEMONADE

As explained before, estimating both compartment diffusivities and orientation dispersion of neurites simultaneously is problematic and tends to be biased. Some methods suggest fixing some parameters such as NODDI or limiting its application to coherent fibers only as WMTI to work around these

problems. Releasing these constraints necessitates estimating a larger number of parameters.

A recent method in white matter estimates the scalar parameters of a two-compartment kernel separately from the ODF without any constraints. The method developed by Novikov et al. (2018c) is based on the modeling of the diffusion signal as a convolution of the ODF and the response kernel from a perfectly aligned fiber segment, as presented in Section 2.4.2. It can be decomposed into two steps. The first step solves an algebraic system of equations that relates the kernel parameters to the signal moments for low b-values. This part was called Linearly Estimated Moments provide Orientations of Neurites And their Diffusivities Exactly (LEMONADE), which stands for Linearly Estimated Moments provide Orientations of Neurites And their Diffusivities Exactly. It requires at least 3 non-zero b-shells inferior to 2.5 ms μm^{-2} and returns estimates for f_{intra}, $D_n^{\parallel}$, $D_e^{\parallel}$, $D_e^{\perp}$ and $p_2 = \frac{3\left\langle(\cos\psi)^2\right\rangle - 1}{2}$, which gives an estimate of the orientation dispersion. In a second step, a rotationally invariant energy function of the system is minimized exploiting all available data and using the first estimates as initialization values.

This method emphasizes the existence of the two mathematical solutions as introduced before, and shows that, in principle, the degeneracy can be avoided using measurements up to the 3rd order of b-values. However, due to noise in the data, the solution selection remains difficult in practice and individual validation should be carried out.

The assumptions made in this approach are, as in the other methods previously presented, the existence of only two compartments, the uniformity of diffusivities across all axons in the voxel, and axial symmetry of the kernel. These assumptions are also the limitations of the model used. Validation in the case of pathological tissue also needs to be investigated.

We refer the reader to Jelescu and Budde's review on the accuracy and validation of biophysical parameters of different diffusion models in white matter, which includes the ones presented before (Jelescu and Budde, 2017a).

2.4.3 Modeling Brain gray Matter: Moving away from the SM

Previously defined models focus on the study of the white matter, and were proven not to hold in gray matter, indicating that both tissues require different models in order to accurately capture their microstructure. Three main hypotheses have been elaborated to explain the different behavior of gray matter DWI signals.

Typically neglected in white matter due to their relatively small density (5-10% ex-vivo), the abundance of cell bodies (somas) in the gray matter could also be a possible explanation for the deviation from the stick power-law. Accounting for 10-20% of gray matter tissue volume (Shapson-Coe et al., 2021), the modeling of an extra impermeable isotropic compartment might be necessary (Palombo et al., 2019).

Increased permeability in cell membranes in neurites in the gray matter might also be the cause of an increased exponent in the stick model (McKinnon et al., 2017; Veraart et al., 2020). Indeed, in white matter tissue, axons are encompassed in myelin sheaths that guaranty their impermeability (i.e. negligible exchange between axons and ECS) over the NMR-relevant timescales. In gray matter, this myelin content being limited, water exchange across the neurite membranes might need to be accounted for.

Finally, neurite structural disorders, such as curvy projections, dendritic spines or neurite beading (Özarslan et al., 2018; Novikov et al., 2014; Lee et al., 2020) could lead to non-Gaussian diffusion along the effectively one-dimensional neurites.

Although ideally all three explanations should be taken into account, Jelescu et al. (2021) suggested that exchange effectively dominates over structural disorder. I will now present the two main modeling solutions that have been proposed in order to take into account soma and exchange.

2.4.3.1 Modeling Brain gray Matter: a Three-Compartment Model

A three-compartment model has been proposed to account for the presence of cell bodies, called Soma And Neurite Density Imaging (SANDI) (Palombo et al., 2020). It extends the SM by adding a soma compartment, relying on the assumption that, at short diffusion times ($\leq 20\text{ms}$) exchanges between the compartments can be considered negligible. The resulting direction-averaged signal is the following:

$$S(\mathbf{b}) = f_{\text{neurites}} \cdot S_{\text{neurites}}(\mathbf{b}) + f_{\text{somas}} \cdot S_{\text{somas}}(\mathbf{b}) + f_{\text{ECS}} \cdot S_{\text{ECS}}(\mathbf{b}) \ , \qquad (2.25)$$

where direction averaged-signals of all the compartments are considered normalized, and signal proportions sum to 1 ($f_{neurites} + f_{somas} + f_{\text{ECS}} = 1$). Water spins diffusing within cellular soma are considered restricted into spheres of apparent radius r_s with a fixed diffusivity $D_s = 3\,\mu\text{m}\,\text{ms}^{-2}$.

Using a random forest regression, Palombo et al. (2020) provide estimates for the signal proportions, neurites, and ECS diffusivities, and soma radius. Although the method provides a rather acceptable accuracy in real-case sce-

narios, it can only output one set of tissue parameters for a given observed dMRI signal, masking, therefore, other biologically plausible solutions that could generate the same observed signal. Furthermore, the parameter estimates are obtained following a deterministic approach, so no confidence interval description is available.

2.4.3.2 An Exchange Model: the Kärger Model

A two-compartment model with exchange has been developed in the pulsed gradient spin echo (PGSE) case, and is referred to as the Kärger Model (KM) (Kärger, 1985; Fieremans et al., 2010). The two exchanging compartments considered are the ECS and sticks with permeable barriers, which could be parallel to mimic axons in white matter (Fieremans et al., 2010), or isotropically oriented such as neurites in gray matter (Jelescu et al., 2021; Olesen et al., 2021). In order to keep the notation coherent within this manuscript, we will denote the sticks-related variables with the n underscore.

Two main hypotheses regulate this model. First, in each voxel volume, i.e. in each point $\mathbf{r}$, a two-component magnetization density exists:

$$\mathbf{m}(t,\mathbf{r}) = \begin{pmatrix} m_n(t,\mathbf{r}) \\ m_{\mathrm{ECS}}(t,\mathbf{r}) \end{pmatrix} \tag{2.26}$$

Both compartments are assumed to be present with proportions f_n and f_{ECS}, satisfying the relation $f_n + f_{\mathrm{ECS}} = 1$. At the diffusion time $t = 0$, the initial magnetization ratio is then equal to $\frac{f_n}{f_{\mathrm{ECS}}}$.

Secondly, the model assumes that diffusion within both compartments is Gaussian, with the diffusivities D_n and D_e. This hypothesis becomes asymptotically exact in the considered brain tissue at a coarse-grained scale. We consider here $D_n \equiv D_n^{\parallel}(\mathbf{g}\cdot\mathbf{n})^2$, which corresponds to the parallel diffusivity in the sticks, and $D_e \equiv D_e^{\parallel}(\mathbf{g}\cdot\mathbf{n})^2 + D_e^{\perp}\,(1-(\mathbf{g}\cdot\mathbf{n})^2)$.

The coupled Bloch-Torrey equations can then be used to describe the evolution of magnetization (Equation 2.26) (Kärger, 1985):

$$\frac{\partial \mathbf{m}}{\partial t} = \left(\mathcal{D}\nabla^2 - i\mathbf{g}(t)\cdot\mathbf{r}\right)\mathbf{m} - \mathcal{R}\mathbf{m}, \tag{2.27}$$

where $\mathcal{D} = \mathrm{diag}(D_n, D_e)$ is the diagonal matrix of the two compartment diffusivities and ∇^2 is the Laplacian. As exchange between the two compartments

is assumed to be independent from position and timing, it is modeled as a Poisson process regulated by the constant rate matrix:

$$\mathcal{R} = \begin{pmatrix} \tau_n^{-1} & -\tau_{\text{ECS}}^{-1} \\ -\tau_n^{-1} & \tau_{\text{ECS}}^{-1} \end{pmatrix} = \begin{pmatrix} r_n & -r_{\text{ECS}} \\ -r_n & r_{\text{ECS}} \end{pmatrix}, \tag{2.28}$$

with τ_n and τ_{ECS} the mean lifetimes in both compartments and r_n and r_{ECS} their exchange rates ($r_n = \frac{1}{\tau_n}$ and $r_{\text{ECS}} = \frac{1}{\tau_{\text{ECS}}}$). They satisfy the conservation of mass equation:

$$\frac{f_n}{\tau_n} = \frac{f_{\text{ECS}}}{\tau_{\text{ECS}}}. \tag{2.29}$$

Total exchange time t_{ex} can be defined as $t_{ex} = \frac{1}{r_n + r_{\text{ECS}}}$.

Taking the Fourier transform of the voxel-averaged spin-packet magnetization density (Equation 2.26) with the initial condition defined by proportions f_n and f_{ECS}, and considering the Gaussian diffusion assumption, the coupled Bloch-Torrey equations become:

$$S(q) = S_0 \begin{bmatrix} 1 & 1 \end{bmatrix} \exp\left(-(2\pi q)^2 t \cdot \mathcal{D} + \mathcal{R} \cdot t\right) \begin{bmatrix} f_n \\ f_{\text{ECS}} \end{bmatrix} \tag{2.30}$$

The resulting signal response is then the following:

$$\mathcal{K}(q, t, \mathbf{g} \cdot \mathbf{n}, D_n, D_e, \tau_n, \tau_{\text{ECS}}) = f'_n e^{-D'_n (2\pi q)^2 t} + f'_{\text{ECS}} e^{-D'_e (2\pi q)^2 t} \tag{2.31}$$

with

$$D'_{n/e} = \frac{1}{2}\Bigg[D_n + D_e + \frac{1}{(2\pi q)^2}\left(\frac{1}{\tau_n} + \frac{1}{\tau_{\text{ECS}}}\right) \tag{2.32}$$

$$\pm \sqrt{\left(D_e - D_n + \frac{1}{(2\pi q)^2}\left(\frac{1}{\tau_{\text{ECS}}} - \frac{1}{\tau_n}\right)\right)^2 + \frac{4}{(2\pi q)^4 \tau_n \tau_{\text{ECS}}}}\Bigg] \tag{2.33}$$

$$\begin{aligned} f'_n &= 1 - f'_{\text{ECS}} \\ f'_{\text{ECS}} &= \frac{f_n D_n + f_{\text{ECS}} D_e - D'_n}{D'_e - D'_n}. \end{aligned} \tag{2.34}$$

Note that in the case of no exchange, i.e. $\mathcal{R} = 0$, we retrieve a biexponential function corresponding to the sum of two Gaussian compartments with diffusivities D_n and D_e, weighted by their proportion f_n and f_{ECS} (similar to Equation 2.16 with $f_{\text{CSF}} = 0$).

2.5 Summary and Beyond

Diffusion MRI is a perfectly adapted technique to measure tissue microstructure in the human brain using water motion. Accessing tissue microstructure non-incisively provides a great insight into the understanding of brain functioning, and the study of its evolution can enlighten the progression of certain pathologies.

We have shown two main approaches to describe microstructure anisotropy using diffusion MRI: signal representation and biophysical modeling. While the former is general and makes no assumptions about the underlying tissue, models are designed for a particular tissue and therefore provide greater specificity and interpretation of the estimated biological parameters. The difficulties in modeling reside in accurately capturing the features that effectively and substantially impact the diffusion signal in a given acquisition range, and being able to correctly fit the model (inverse problem).

Although great progress has been made during the last decade, some questions remain unresolved. Notably, the estimation of brain gray matter microstructure remains difficult. The question of which model to use is at the core of diffusion MRI research. We can wonder to which extent we can consider compartments as non-exchanging. Diffusion time, brain region, and myelination of the tissue will most likely impact the answer to this question. The necessity to add a soma compartment is also unresolved. Defining the right model is however not enough. A robust method to inverse the problem and estimate tissue microstructure from an acquired diffusion signal is mandatory in order to get interpretable and trustable estimates. My thesis integrates this context of getting reliable estimates of the gray matter cytoarchitecture non-invasively using diffusion MRI.

3 Theoretical Background on Statistical Inference

Contents

Acronyms used in the chapter

ABC	Approximate Bayesian Computation
ELBO	Evidence Lower BOund
i.i.d.	independent and identically distributed
KL	Kullback-Leibler
LFI	Likelihood-Free Inference
MCMC	Markov Chain Monte Carlo
NF	Normalizing Flows
p.d.f.	probability density functions
SBI	Simulation-Based Inference
VI	Variational Inference

Notations used in the chapter

$\mathbf{x}$	Observation vector
$\boldsymbol{\theta}$	Parameter vector
$p(\boldsymbol{\theta})$	Prior distribution
$p(\boldsymbol{x_0}\|\boldsymbol{\theta})$	Likelihood distribution of an observation $\boldsymbol{x_0}$ given parameters $\boldsymbol{\theta}$
$p(\boldsymbol{\theta}\|\boldsymbol{x_0})$	Posterior distribution
$p(\boldsymbol{x_0})$	Evidence of the data
$\mathbb{E}_s[\boldsymbol{x}]$	Expectation of any random variable $\boldsymbol{x}$ w.r.t. its subscript s when needed
$\boldsymbol{u}$	Real vector
T	Diffeomorphism
$J_T(\boldsymbol{u})$	Jacobian of T
det$J_T(\boldsymbol{u})$	Determinant of the Jacobian matrix of T
τ	Transformer
c_i	i-th conditioner

This chapter presents the statistical inference background needed for a good understanding of the present thesis. We start by presenting the general Bayesian formalism in the context of statistical inference. We show that the Bayes' theorem can be applied to estimate the probability density of the parameters which generated an observation given this observation, called the posterior, when the model's likelihood is available. However, the main difficulty in solving the inference problem of a simulator model is the evaluation of the likelihood. We then present likelihood-free (or simulation-based) inference methods, that have been developed to bypass this difficulty. Finally, we discuss two factors that need to be considered: the amortization of a model and the utility of summary statistics.

3.1 Bayesian Framework for Statistical Inference

3.1.1 Introduction to the Inverse Problem

A challenging question in science is understanding how a signal has been generated, i.e., what parameters specifically influenced this acquired signal, whether it is an image, a MEG brain activation, or a cognitive task result.

It is possible to define complex and high fidelity mechanistic models relying on experts' knowledge that predict how systems will behave in multiple scenarios, that is, to generate an observation $\boldsymbol{x}$ given the parameters $\boldsymbol{\theta}$. These statistical models stochastically generate a data vector $\boldsymbol{x}$ whose distribution depends on $\boldsymbol{\theta}$.

Given an observation $\boldsymbol{x}$ and a statistical model, we want to solve the inverse problem and retrieve the parameters $\boldsymbol{\theta}$ that produced $\boldsymbol{x}$. Inferring parameters of interest given a statistical model of their relationship is called *statistical inference*. Two main approaches to statistical inference have been developed: Bayesian inference and frequentist inference. We will focus on the Bayesian framework in this work.

3.1.2 The Bayesian Formalism

We consider an observed datum $\boldsymbol{x_0} \in \mathcal{X}$, whose generation can be described by a statistical model $\mathcal{M}$, parametrized by an unknown vector $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$. A *prior distribution* is defined, denoted $p(\boldsymbol{\theta})$, that expresses the prior knowledge we have about $\boldsymbol{\theta}$. In practice, it usually corresponds to the physically possible intervals of the parameters. A conditional density model $p(\boldsymbol{x_0}|\boldsymbol{\theta})$, known as the *likelihood distribution*, expresses a statistical relationship between $\boldsymbol{x_0}$ and $\boldsymbol{\theta}$. The probability distribution of $\boldsymbol{\theta}$ given $\boldsymbol{x}$, i.e., the distribution modeling

how our beliefs about $\boldsymbol{\theta}$ should change in light of observing $\boldsymbol{x_0}$, can then be formulated using Bayes' theorem (Bishop, 2006):

$$p(\boldsymbol{\theta}|\boldsymbol{x} = \boldsymbol{x_0}) = \frac{p(\boldsymbol{x_0}|\boldsymbol{\theta})\ p(\boldsymbol{\theta})}{p(\boldsymbol{x_0})} \quad , \tag{3.1}$$

where $p(\boldsymbol{\theta}|\boldsymbol{x_0})$ is called the *posterior distribution* and

$$p(\boldsymbol{x_0}) = \int p\left(\boldsymbol{x_0}|\boldsymbol{\theta'}\right) p\left(\boldsymbol{\theta'}\right) \mathrm{d}\boldsymbol{\theta'}$$

is a normalizing constant, commonly referred to as the *evidence* of the data. We assume that this quantity is finite for every $\boldsymbol{x_0}$ so that the posterior density is always well-defined.

$p(\boldsymbol{\theta}|\boldsymbol{x_0})$ allows to express uncertainty over plausible parameters, rather than just point estimates. Providing information such as posterior expectations or credible intervals can help reach conclusions, or make more informed decisions.

3.1.3 Explicit and Implicit Models

The type of statistical model used to infer the parameters is of primary importance in choosing the resolution method. Indeed, as we will present in this section, the main difficulty for performing statistical inference is the availability of the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta})$.

Density models, also dubbed *explicit* or *prescribed models* (Cranmer et al., 2020), directly express the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta})$. This way, the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x} = \boldsymbol{x_0})$ can easily be evaluated (up to a normalizing constant) using Bayes' rule (Equation 3.1). Such methods are referred to as *likelihood-based inference* methods, as the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta})$ can be explicitly evaluated.

In contrast, *simulator models*, also called *implicit models*, describe how the data are generated. They take as input a vector of relevant parameters $\boldsymbol{\theta}$, produce samples of some latent variables (internal states) $\boldsymbol{z_i} \sim p_i\left(\boldsymbol{z_i}|\boldsymbol{\theta}, \boldsymbol{z_{<i}}\right)$, and finally produce a data vector $\boldsymbol{x} \sim p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{z})$ as output, that corresponds to the observations. The latent variables do not necessarily have a physical meaning and might be (or not) conveniently accessed through the simulator model. A direct evaluation of the likelihood, such as for density models, is typically not possible. Inversion methods that do not rely on the density evaluation are necessary to approximate the posterior distribution. Such methods are called as *Likelihood-Free Inference* (LFI) or *Simulation-Based Inference* (SBI) (Cranmer et al., 2020).

Simulator models directly describe a real-world phenomenon, which makes them easily interpretable. They have been used in many different scientific fields, such as ecology (Wood, 2010), cosmology (Alsing et al., 2019), population genetics (Beaumont et al., 2002), and MR fingerprinting (Boux et al., 2021). The inference problems tackled in this thesis lie in this category, where the likelihoods of the defined models are not accessible.

Likelihood-based inference represents the majority of inference problems. It is efficient, as it can directly rely on Bayes' theorem. Likelihood-free inference was introduced later to deal with models whose likelihood could not be evaluated. Another possibility to solve the inference problem concerning a simulator model where the likelihood is not available could be to do it in two steps. We could first compute the likelihood given the model's internal functioning and then use likelihood-based methods afterward. Reasons for the intractability of the likelihood are presented in the next section and argue in favor of likelihood-free methods. A more detailed explanation is available in Papamakarios (2019).

3.1.4 Likelihood Intractability

3.1.4.1 Integration through the Latent Space

The likelihood can be defined as the integral over all possible paths through the latent space (i.e., all possible executions of the simulator). That is:

$$p(\boldsymbol{x}|\boldsymbol{\theta}) = \int p(\boldsymbol{x}, \boldsymbol{z}|\theta)\mathrm{d}\boldsymbol{z} \tag{3.2}$$

where $p(\boldsymbol{x}, \boldsymbol{z}|\theta)$ is the joint probability density of data $\boldsymbol{x}$ and latent variables $\boldsymbol{z}$. In the case of real-life simulators with large latent spaces, the explicit computation of this integral is impossible, as it would require a very high computational cost.

3.1.4.2 Black Box Simulators

In practice, black-box simulators may be used, such as executable programs or even a lab experiments, where the latent variables are not accessible. In those cases, the likelihood evaluation will not be possible even if, theoretically, likelihood-based inference methods could be used. Likelihood-free inference methods become essential to solve inverse problems in those cases.

3.2 Likelihood-Based Inference Methods

This part will discuss two main methods to deal with the Bayesian inference problem in which the likelihood is accessible: *Markov Chain Monte Carlo* (MCMC), which is a sampling-based approach, and *Variational Inference* (VI), which is an approximation-based method. Although we rely on a likelihood-free inference technique in this thesis, Monte Carlo is a well-known and widely used technique in the diffusion community (see e.g. Fieremans et al., 2010; Novikov et al., 2011), and we wanted to put it back in its context. VI introduces the idea of defining a family of functions, which is linked to the method chosen to solve the inverse problem in this thesis, namely Normalizing Flows (NF). Moreover, VI allows us to introduce the idea of distance between two distributions, which is an essential notion in NF.

3.2.1 Markov Chain Monte Carlo

Markov Chain Monte Carlo algorithms aim at generating samples from a complex probability distribution that can be defined up to a factor (Andrieu et al., 2003). This property allows removing the dependence on the evidence in Bayes' formula, which cannot be evaluated.

The "Monte Carlo" part of the method's name refers to a technique for randomly sampling a probability distribution based on the assumption that samples can be efficiently drawn from the target distribution. The "Markov Chain" part relates to a stochastic method for generating a sequence of random variables, where the current variable probabilistically depends on the last variable in the chain.

Combining these two techniques, the idea is to approximate a probability density using random sampling, where each sample depends on the previous one. In other words, the goal is to build a Markov Chain whose stationary distribution is the posterior $p(\boldsymbol{\theta}|\boldsymbol{x})$. The generation of many samples is needed to reach a steady state, with a close enough approximation of the density function.

An inconvenience of MCMC algorithms is their sensitivity to their starting point. They begin with a burn-in phase, during which samples cannot be used as the Markov chain has not converged towards its stationary distribution yet.

Two of the common MCMC algorithms are called *Metropolis-Hastings* and *Gibbs Sampling*. We refer the reader to Robert and Casella (2004) for a deeper presentation of this method.

3.2.2 Variational Inference

Unlike MCMC methods that sample from a target distribution, Variational Inference methods look for the best approximation of the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ among a parametrized family. Finding the closest distribution $q(\boldsymbol{\theta})$ to the posterior $p(\boldsymbol{\theta}|\boldsymbol{x})$ from a variational family Q is cast as an optimization problem over the family Q (Blei et al., 2017).

From above statement, two criteria need to be carefully defined: the variational family Q and the distance quantifying how close the distribution $q(\boldsymbol{\theta})$ is to the posterior $p(\boldsymbol{\theta}|\boldsymbol{x})$. Both notions are presented in the next sections.

3.2.2.1 Distribution Family

To begin with, a variational family of distributions Q needs to be defined. This family will determine the space in which we search for the best approximation (the definition of what "best" means, in this case, will be defined in Section 3.2.2.2). The choice of this family controls the bias and complexity of the method. Indeed, a simple family generates a restrictive model that will present a high bias but a simple optimization process. On the contrary, a complicated family results in a nearly free model with smaller bias, but whose optimization process will be far harder if not intractable. A correct balance needs to be found to ensure a good quality of the fit while maintaining a tractable optimization process.

3.2.2.2 Kullback-Leibler Divergence

Once a variational family has been defined, the question of finding the best approximation among this distribution family needs to be addressed. The notion of closeness between the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ and a distribution $q(\boldsymbol{\theta})$ can be measured using the *Kullback-Leibler (KL) divergence*. This divergence being asymmetric, two types of KL divergences exist: the *forward* and *reverse* KL divergences.

Forward Kullback-Leibler divergence. The forward KL divergence between the target posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$ and a distribution $q(\boldsymbol{\theta}|\boldsymbol{x})$ is defined as follows:

$$\mathrm{KL}(p(\boldsymbol{\theta}|\boldsymbol{x})\|q(\boldsymbol{\theta}|\boldsymbol{x})) = -\mathbb{E}_{p(\boldsymbol{\theta}|\boldsymbol{x})}[\log q(\boldsymbol{\theta}|\boldsymbol{x})] + \text{const.} \tag{3.3}$$

This divergence requires to be able to sample from the posterior distribution. In practice, this is usually not possible. A workaround is to sample from the joint distribution $p(\boldsymbol{\theta}, \boldsymbol{x})$ in an amortized context instead, as presented in Section 3.3.3.3.

Reverse Kullback-Leibler divergence. The reverse KL divergence reflects the divergence of a distribution $q(\boldsymbol{\theta}|\boldsymbol{x})$ from the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$:

$$\begin{aligned} \mathrm{KL}(q(\boldsymbol{\theta}|\boldsymbol{x})\|p(\boldsymbol{\theta}|\boldsymbol{x})) &= \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{x})}[\log q(\boldsymbol{\theta}|\boldsymbol{x}) - \log p(\boldsymbol{\theta}|\boldsymbol{x})] \\ &= \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{x})}[\log q(\boldsymbol{\theta}|\boldsymbol{x})] - \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{x})}[\log p(\boldsymbol{\theta}, \boldsymbol{x})] + \log p(\boldsymbol{x}) \end{aligned} \tag{3.4}$$

This divergence is well-suited for situations where it is possible to evaluate the joint density $p(\boldsymbol{\theta}, \boldsymbol{x})$.

In the case of VI, we are looking for the best approximation of a posterior distribution among a variational family. Minimizing the reverse KL divergence is equivalent to finding the distribution the closest to the posterior distribution. However, the evaluation of this reverse KL divergence is generally not possible because it requires computing the logarithm of the evidence $\log p(\boldsymbol{x})$. A workaround is to optimize the *Evidence Lower BOund (ELBO)* instead, which removes this dependence on the evidence and only relies on the joint distribution (Blei et al., 2017):

$$\mathrm{ELBO}(q) = \mathbb{E}_{q(\boldsymbol{\theta}|\boldsymbol{x})}[\log p(\boldsymbol{x}, \boldsymbol{\theta}) - \log q(\boldsymbol{\theta}|\boldsymbol{x}))] \tag{3.5}$$

Maximizing the ELBO is equivalent to minimizing the KL divergence, as proven by the relation:

$$\mathrm{KL}(q(\boldsymbol{\theta}|\boldsymbol{x}))\|p(\boldsymbol{\theta}|\boldsymbol{x})) = \mathbb{E}_q[\log p(\boldsymbol{x}|\boldsymbol{\theta})] - \mathrm{ELBO}(q) \tag{3.6}$$

To maximize the ELBO and find the distribution the closest to the posterior distribution, a Monte Carlo procedure is usually used. Gradient descent can for example be used as an optimization scheme with automatic differentiation (Kucukelbir et al., 2017), provided that a modern automatic differentiation framework has been used (Dillon et al., 2017; Paszke et al., 2019). At each optimizer step, we first sample points $\boldsymbol{\theta_m}$ from $q(\boldsymbol{\theta})$, and then approximate the discretized ELBO:

$$ELBO(q) \simeq \frac{1}{M} \sum_{m=1}^{M} \log p\left(\boldsymbol{x}, \boldsymbol{\theta_m}\right) - \log q\left(\boldsymbol{\theta_m}|\boldsymbol{x_m}\right) \tag{3.7}$$

3.2.3 MCMC vs VI

As previously described, MCMC and VI are two methods that aim at solving the inference problem using the likelihood distribution. MCMC approximates the posterior distributions by sampling from a Markov chain, while VI ap-

proximates the posterior distribution by solving an optimization problem. Differences in their guarantees and speed can be noted.

First, MCMC methods provide guarantees of producing (asymptotically) exact samples from the target density (Robert and Casella, 2004), contrarily to VI which does not benefit from such guarantees. However, VI is generally faster than MCMC, as it can rely on efficient methods such as stochastic or distributed optimization (Blei et al., 2017). Another critical consideration to keep in mind when using VI is its general underestimation of the variance of the posterior distribution, which could be problematic for some tasks. Nevertheless, VI allows for better complex posterior geometry approximations (e.g., multi-modal) than MCMC.

As a result, VI is better suited for large datasets or complex posterior geometry and MCMC for smaller datasets, where we are willing to pay a computational cost to get more precise samples (Blei et al., 2017).

3.3 Likelihood-Free Inference Methods

In the case of simulator models, the likelihood is intractable, and methods described in 3.2 cannot be directly applied. Other methods have been proposed to overcome this problem and a general presentation of existing workflows is presented in Section 3.3.1. The first contributions on LFI are known as *Approximate Bayesian Computation* (ABC) and have been applied to invert models from ecology, population genetics, and epidemiology (Sisson, 2018). We present the most popular ABC methods in Section 3.3.2. Recently, there has been a growing interest in the machine learning community in improving the limitations of ABC methods through deep generative modeling, i.e., neural network architectures specially tailored to approximate probability density functions from a set of examples (Goodfellow et al., 2016). *Normalizing Flows* NF (Papamakarios et al., 2019) are a particular class of such neural networks that have demonstrated promising results for LFI in different research fields (Cranmer et al., 2020; Gonçalves et al., 2020; Greenberg et al., 2019). NF are presented in Section 3.3.3.

3.3.1 Workflows for Simulation-Based Inference

Likelihood-free inference techniques can be separated into two main groups (Cranmer et al., 2020). The first group corresponds to methods that use the simulator itself during inference (such as ABC methods, presented in Section 3.3.2) and output an estimate value $\boldsymbol{x}$ that can be compared with the observation $\boldsymbol{x}_0$. The second group constructs instead a surrogate model, whose output is then used for inference. NF correspond to this second group of inference methods. Different ways to address the intractability of the likelihood exist in each group. Some methods construct a tractable surrogate model for the likelihood function or the likelihood-ratio function, while others never use it explicitly. At least, the methods can provide access to samples from the approximated posterior or an approximation of the posterior distribution.

3.3.2 Approximate Bayesian Computation

Approximate Bayesian Computation (ABC) methods are simulation-based methods that replace the evaluation of the likelihood with a comparison between the observed D-dimensional data $\boldsymbol{x_0}$ and a simulated data $\boldsymbol{x}^*$. Three types of ABC methods will be presented in this section, along with their limitations: *rejection ABC*, *Markov Chain Monte Carlo ABC* and *Sequential Monte Carlo ABC*. We refer the reader to Sisson (2018) for a complete review of ABC methods.

3.3.2.1 Rejection ABC

We define $B_\epsilon(\boldsymbol{x_0})$ the neighbourhood of $\boldsymbol{x_0}$ as the set of data points whose distance from $\boldsymbol{x_0}$ is inferior or equal to ϵ:

$$B_\epsilon\left(\mathbf{x_0}\right) = \left\{\mathbf{x} : \|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon\right\}, \tag{3.8}$$

with $\|\cdot\|$ a norm in $\mathbb{R}^D$. For small ϵ, the likelihood $p(\boldsymbol{x_0}|\boldsymbol{\theta})$ can be approximated as the average conditional density inside $B_\epsilon(\boldsymbol{x_0})$ (Papamakarios, 2019):

$$p\left(\mathbf{x_0}|\boldsymbol{\theta}\right) \approx \frac{\Pr\left(\|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon \mid \boldsymbol{\theta}\right)}{\left|B_\epsilon\left(\mathbf{x_0}\right)\right|}, \tag{3.9}$$

where $\left|B_\epsilon\left(\mathbf{x_0}\right)\right|$ corresponds to the volume of $B_\epsilon(\boldsymbol{x_0})$. Using Bayes' rule (3.1), the posterior distribution can then be approximated as:

$$p\left(\boldsymbol{\theta}|\mathbf{x} = \mathbf{x_0}\right) \approx \frac{\Pr\left(\|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon \mid \boldsymbol{\theta}\right) p(\boldsymbol{\theta})}{\int \Pr\left(\|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon \mid \boldsymbol{\theta}'\right) p\left(\boldsymbol{\theta}'\right) \mathrm{d}\boldsymbol{\theta}'} = p\left(\boldsymbol{\theta} \mid \|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon\right) \tag{3.10}$$

$p\left(\boldsymbol{\theta} \mid \|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon\right)$ can be interpreted as the exact posterior under an alternative observation $\boldsymbol{x}$. As ϵ tends towards 0, $B_\epsilon\left(\mathbf{x_0}\right)$ becomes infinitely small, and the approximate posterior $p\left(\boldsymbol{\theta} \mid \|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon\right)$ tends to the exact posterior $p(\boldsymbol{\theta}|\boldsymbol{x} = \boldsymbol{x_0})$ (see e.g. Prangle, 2017, supplementary material, theorem 1 for a formal proof). In turn, as ϵ tends towards infinity, $B_\epsilon\left(\mathbf{x_0}\right)$ encompasses the space entirely, and the approximate posterior tends to the prior $p(\boldsymbol{\theta})$.

The simplest ABC algorithm is called *rejection ABC* (Pritchard et al., 1999), and is a rejection-sampling method for obtaining independent samples from the approximate posterior $p\left(\boldsymbol{\theta} \mid \|\mathbf{x} - \mathbf{x_0}\| \leq \epsilon\right)$. This method is divided into three steps:

(1) Sample a candidate parameter vector $\boldsymbol{\theta}^*$ from the prior distribution $\boldsymbol{\theta}$.

(2) Simulate $\boldsymbol{x}^*$ using the sampled parameter $\boldsymbol{\theta}^*$ as defined by the conditional probability $p(\boldsymbol{x}|\boldsymbol{\theta}^*)$ of the simulator.

(3) Compare the simulated data vector $\boldsymbol{x}^*$ with the observed data $\boldsymbol{x_0}$ using a distance function and a tolerance $\epsilon \geq 0$. Accept $\boldsymbol{\theta}^*$ if $\|\boldsymbol{x}^* - \boldsymbol{x_0}\| \leq \epsilon$.

The parameter ϵ defines the desired level of agreement between $\boldsymbol{x_0}$ and $\boldsymbol{x}^*$. As it becomes smaller, the accepted samples approach the exact posterior, but fewer samples are accepted. That is, ϵ controls the trade-off between approximation quality and computational efficiency.

A binding limitation of this technique is that a large number of simulations can be necessary for small ϵ, especially when the data is high-dimensional.

To illustrate it, let us consider the case where $\boldsymbol{x}$ is D-dimensional, and the distance is defined as the maximum norm (example taken from Papamakarios, 2019). The acceptance region $B_\epsilon(\boldsymbol{x_0})$ is then a cube of side 2ϵ centered at x_0. For small ϵ, the acceptance probability can be estimated as:

$$\Pr\left(\max(\mathbf{x}-\mathbf{x}_0) \leq \epsilon \mid \boldsymbol{\theta}\right) \approx p\left(\mathbf{x}_0|\boldsymbol{\theta}\right)\left|B_\epsilon\left(\mathbf{x}_0\right)\right| = p\left(\mathbf{x}_0|\boldsymbol{\theta}\right)(2\epsilon)^D, \tag{3.11}$$

As ϵ becomes smaller, the acceptance probability tends towards zero at a rate $\mathcal{O}\left(\epsilon^D\right)$. Concretely, for $p\left(\mathbf{x}_0|\boldsymbol{\theta}\right) = 1$, $\epsilon = 0.05$ and $D = 15$, each accepted sample would require on average a thousand trillion simulations, which is obviously extremely costly. Improvements to this algorithm have been proposed to reduce the number of simulations needed, such as the two methods presented below.

In addition, in practice, for high dimensional cases, the data is transformed into a small number of features to keep a reasonable computation time. These features are dubbed *summary statistics*. If they are sufficient for inferring $\boldsymbol{\theta}$ (see definition 9.32 of sufficient in Wasserman, 2004), then using summary statistics instead of the data induces no loss of information about $\boldsymbol{\theta}$. This idea is developed in Section 3.4.2.

3.3.2.2 Markov Chain Monte Carlo ABC

As presented in Section 3.3.2.1, rejection ABC samples from the prior distribution and only accepts parameters whose simulated data are close to the observed data. However, if the approximated posterior is significantly narrower than the prior distribution, as is generally the case, this method will discard many samples.

An improvement on the sampling strategy based on Markov Chain Monte Carlo in its Metropolis-Hastings form (see Section 3.2.1) has been proposed to accept more samples, and therefore reduce the computation needs (Marjoram et al., 2003). Instead of independently sample parameters $\boldsymbol{\theta}$ from the prior distribution, *Markov Chain Monte Carlo ABC* suggests slightly perturbing the previously accepted sample, so that this new sample will likely be accepted too.

A proposal distribution $q(\boldsymbol{\theta'}|\boldsymbol{\theta})$ is introduced that suggests a new parameter $\boldsymbol{\theta'}$ based on a previously accepted parameter $\boldsymbol{\theta}$. An *acceptance ratio* is computed as follows:

$$\alpha = \begin{cases} \frac{p(\boldsymbol{\theta'})q(\boldsymbol{\theta}|\boldsymbol{\theta'})}{p(\boldsymbol{\theta})q(\boldsymbol{\theta'}|\boldsymbol{\theta})} & \text{if } \|\mathbf{x}-\mathbf{x}_0\| \leq \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

This algorithm is described in Algorithm 1 for more clarity.

Algorithm 1 Markov Chain Monte Carlo ABC

Propose $\boldsymbol{\theta}' \sim q(\boldsymbol{\theta}'|\boldsymbol{\theta})$
Simulate $\mathbf{x}_n \sim p(\mathbf{x}|\boldsymbol{\theta}')$
if $\|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon$ **then**
    Calculate acceptance ratio $\alpha = \frac{p(\boldsymbol{\theta}')q(\boldsymbol{\theta}|\boldsymbol{\theta}')}{p(\boldsymbol{\theta})q(\boldsymbol{\theta}'|\boldsymbol{\theta})}$
    Sample $u \sim \mathcal{U}(0, 1)$
    if $u \leq \alpha$ **then**
        return $\boldsymbol{\theta}'$
    else
        return $\boldsymbol{\theta}$
    end if
else
    return $\boldsymbol{\theta}$
end if

Unlike rejection ABC, samples are not independently produced in a MCMC ABC approach. Accepted parameters $\boldsymbol{\theta}$ are randomly perturbed to obtain new parameters $\boldsymbol{\theta}'$ close to $\boldsymbol{\theta}$ that are therefore likely to be accepted. Although this reliance on previous iterations leads to a higher acceptance rate, it also reduces the effective sample size.

Moreover, due to the reliance of parameter $\boldsymbol{\theta}'$ on its predecessor $\boldsymbol{\theta}$, the Markov chain may be stuck at its initial state for a long time. A careful initialization of the chain is consequently required. In practice, rejection ABC can be run until one parameter $\boldsymbol{\theta}$ is accepted, and this $\boldsymbol{\theta}$ can be used to initialize the Markov chain in the MCMC ABC approach (Papamakarios, 2019).

3.3.2.3 Sequential Monte Carlo ABC

In the previous section, a solution to increase the acceptance rate has been introduced, based on perturbing accepted parameters using MCMC. We present an alternative approach based on importance sampling in this section.

To begin with, suppose parameters $\boldsymbol{\theta}$ are not sampled from the prior distribution $p(\boldsymbol{\theta})$ but from a second distribution $\tilde{p}(\boldsymbol{\theta})$, assumed to be non-zero in the support of the prior. Accepted samples, i.e., samples whose simulated data lie in the neighborhood of the observed data, follow the distribution:

$$\tilde{p}\left(\boldsymbol{\theta}|\, \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon\right) \propto \Pr\left(\|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon|\boldsymbol{\theta}\right) \tilde{p}(\boldsymbol{\theta}) \tag{3.13}$$

A population of N importance-weighted samples $\{(w_1, \boldsymbol{\theta_1}), ..., (w_N, \boldsymbol{\theta_N})\}$ is considered that approximates $p(\boldsymbol{\theta} | \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon)$. Weights $\{w_1, ..., w_N\}$ are normalized (i.e., $\sum_{n=1}^{N} w_n = 1$) and are defined as follow:

$$w_n \propto \frac{p(\boldsymbol{\theta}_n | \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon)}{\tilde{p}(\boldsymbol{\theta}_n | \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon)} \propto \frac{\Pr(\|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon | \boldsymbol{\theta}_n) p(\boldsymbol{\theta}_n)}{\Pr(\|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon | \boldsymbol{\theta}_n) \tilde{p}(\boldsymbol{\theta}_n)} = \frac{p(\boldsymbol{\theta}_n)}{\tilde{p}(\boldsymbol{\theta}_n)} \tag{3.14}$$

The steps of this procedure, dubbed *importance-sampling ABC*, are summarized in Algorithm 2. Note that if $\tilde{p}(\boldsymbol{\theta}) = p(\boldsymbol{\theta})$, this approach reduces to rejection ABC.

Algorithm 2 Importance-Sampling ABC

1: **for** $n = 1 : N$ **do**
2: **repeat**
3: Sample $\boldsymbol{\theta_n} \sim \tilde{p}(\boldsymbol{\theta})$
4: Simulate $\boldsymbol{x} \sim p(\boldsymbol{x}|\boldsymbol{\theta_n})$
5: **until** $\|\boldsymbol{x} - \boldsymbol{x_0}\| \leq \epsilon$
6: **end for**
7: Calculate importance weights $w_n \propto \frac{p(\theta_n)}{\tilde{p}(\theta_n)}$ such that they sum to 1
8: **return** $\{(w_1, \boldsymbol{\theta_1}), ..., (w_N, \boldsymbol{\theta_N})\}$

The proposal $\tilde{p}(\boldsymbol{\theta})$ needs to be carefully chosen as it directly impacts the efficiency of the importance-sampling ABC algorithm. If the proposal is too restrained, the importance weights will have a high variance. On the contrary, if the proposal is too broad, the acceptance rate will be low (Alsing et al., 2018).

In practice, a series of T importance-sampling ABCs can be performed, where ϵ is progressively decreased and a new proposal is constructed at each iteration. $p(\boldsymbol{\theta})$ can be used in the first round along with a large enough ϵ_1, so that the acceptance rate is not too low. A population of importance-weighted parameter samples $\{(w_1, \boldsymbol{\theta_1}), ..., (w_N, \boldsymbol{\theta_N})\}$ is obtained. We then choose ϵ_2 that complies with the condition $\epsilon_2 < \varepsilon_1$. Progressively decreasing ϵ allows to obtain increasingly accurate posterior samples. The proposal $\tilde{p}(\boldsymbol{\theta})$ for next round can be defined as a mixture distribution given previous population:

$$\tilde{p}(\boldsymbol{\theta}) = \sum_n w_n q(\boldsymbol{\theta}|\boldsymbol{\theta}_n). \tag{3.15}$$

$q(\boldsymbol{\theta}|\boldsymbol{\theta}_n)$ can be thought as a perturbation of parameters $\boldsymbol{\theta_n}$, similarly to MCMC ABC. A common choice is to use a Gaussian distribution centered at $\boldsymbol{\theta_n}$, which is equivalent to adding a zero-mean Gaussian noise to $\boldsymbol{\theta_n}$. This iterative procedure is known as *sequential Monte Carlo ABC*.

A known issue with sequential Monte Carlo is *sample degeneracy*. If most samples have negligible weights, the number of independent samples, i.e., the effective sample size, becomes significantly inferior to the population size N. A strategy to avoid this is to compute the effective sample size after each iteration and resample the population if the effective sample size is less than a threshold (e.g., $N/2$) (Nowozin, 2015).

Sequential MC ABC provides a higher acceptance rate than rejection ABC and is usually easier to tune than MCMC ABC. Performing several rounds of importance sampling ABC allows to obtain increasingly accurate posterior samples. However, the recurrent ABC problem of a small acceptance probability as ϵ tends towards 0 is still present. The required number of simulation samples increases at each round as ϵ decreases, leading to possibly a huge number of required samples per round.

3.3.2.4 Limitations

Although ABC methods have been applied to invert models from ecology, population genetics, and epidemiology (Sisson, 2018), they have some limitations that lead our choice not to use them in the following thesis. First, these techniques require a large number of simulations for posterior estimations, as many samples are discarded. Moreover, the definition of a distance function is necessary to compare the results of two simulations, as well as a maximum distance (i.e., rejection threshold) for defining a neighborhood of a sample (done via the ϵ parameter). A good trade-off needs to be found between approximation quality and computational efficiency.

3.3.3 Normalizing Flows

Recently, there has been a growing interest in the machine learning community in improving the limitations of ABC methods through deep generative modeling, i.e., neural network architectures specially tailored to approximate probability density functions (p.d.f.) from a set of examples (Goodfellow et al., 2016). NF (Papamakarios et al., 2019) are a particular class of such neural networks that have demonstrated promising results for LFI in different research fields (Cranmer et al., 2020; Gonçalves et al., 2020; Greenberg et al., 2019). These flows are invertible functions capable of transforming vectors generated by a simple base distribution (e.g. the standard multivariate Gaussian distribution) into an approximation of the true posterior distribution. An essential advantage of NF versus other p.d.f. approximators such as generative adversarial networks (GAN) (Goodfellow et al., 2014) and variational auto-encoders (VAE) (Kingma and Welling, 2014) is that it provides the likelihood of any sample point of interest, and it is also straightforward to sample new data points from it. Furthermore, certain classes

of NF can be shown to be universal approximators of probability density functions. We refer the reader to the work by Papamakarios et al. (2019) for a complete review on NF.

3.3.3.1 Definition

Normalizing flows allow to construct flexible probability distributions by applying a series of transformations to a simple density. In the following, we consider $\boldsymbol{\theta}$ and $\boldsymbol{u}$ two D-dimensional real vectors, and we want to define a joint posterior distribution over $\boldsymbol{\theta}$. A diffeomorphism T models the transformation of the real vector $\boldsymbol{u}$, sampled from $p_{\mathrm{u}}(\mathbf{u})$, to $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} = T(\mathbf{u}) \quad \text{where} \quad \mathbf{u} \sim p_{\mathrm{u}}(\mathbf{u}) \tag{3.16}$$

By definition, this transformation is invertible and both T and T^{-1} are differentiable. $p_{\mathrm{u}}(\mathbf{u})$ is referred to as the *base distribution* of the flow-based model. Both T and $p_{\mathrm{u}}(\mathbf{u})$ can have parameters of their own, denoted respectively ϕ and ψ.The density of $\boldsymbol{\theta}$ is, under these hypotheses, well-defined and can be obtained by a change of variables:

$$p_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = p_{\mathrm{u}}(\mathbf{u}) \left|\det J_T(\mathbf{u})\right|^{-1} = p_{\mathrm{u}}\left(T^{-1}(\boldsymbol{\theta})\right) \left|\det J_{T^{-1}}(\boldsymbol{\theta})\right|, \tag{3.17}$$

where $J_T(\boldsymbol{u})$ is the Jacobian of T:

$$J_T(\mathbf{u}) = \begin{bmatrix} \frac{\partial T_1}{\partial \mathrm{u}_1} & \cdots & \frac{\partial T_1}{\partial \mathrm{u}_D} \\ \vdots & \ddots & \vdots \\ \frac{\partial T_D}{\partial \mathrm{u}_1} & \cdots & \frac{\partial T_D}{\partial \mathrm{u}_D} \end{bmatrix}. \tag{3.18}$$

The computation of the Jacobian determinant can have a high time cost, and become even intractable for high dimensions. Therefore, the Jacobian matrix must be tractable, that is it must be fast to compute, and scale well to large D.

A useful property of invertible and diferentiable transformations is that they are *composable*, that is the composition of two transformations T_1 and T_2 is also differentiable and invertible:

$$(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1} \tag{3.19}$$

$$\det J_{T_2 \circ T_1}(\mathbf{u}) = \det J_{T_2}\left(T_1(\mathbf{u})\right) \cdot \det J_{T_1}(\mathbf{u}) \tag{3.20}$$

Complex transformations with more expressive power can therefore be constructed by composing K simple transformations, without losing the invertibility and differentiability properties, allowing to evaluate the density $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$.

3.3.3.2 Autoregressive Flows

Three main types of methods exist for constructing flow-based models in the case of finite compositions (in opposition to continuous-time transformations): *autoregressive flows*, *linear flows* and *residual flows*. This section will present an overview of the existing methods to construct autoregressive flows, which are used in this thesis. For a more thorough description, we invite the reader to refer to Papamakarios et al. (2019).

Autoregressive flows are universal approximators, making this class of models very popular. They have the following form (Huang et al., 2018; Jaini et al., 2019):

$$\boldsymbol{x}_i = \tau\left(\boldsymbol{\theta}_i; \boldsymbol{h}_i\right) \quad \text{where} \quad \boldsymbol{h}_i = c_i\left(\boldsymbol{\theta}_{<i}\right). \tag{3.21}$$

The *transformer* τ is a strictly monotonic function of $\boldsymbol{\theta}_i$ parametrized by $\boldsymbol{h}_i$. c_i is called the i-th *conditioner*, and determines the parameters of the transformer (terms by Papamakarios et al., 2019). With this formulation, every $\boldsymbol{h}_i$ (and each $\boldsymbol{x}_i$) can be computed independently and only depends on $\boldsymbol{h}_{<i}$ ($\boldsymbol{\theta}_{<i}$).The partial derivatives of $\boldsymbol{x}_i$ with respect to $\boldsymbol{\theta}_{<i}$ is then 0 whenever $j > i$. The Jacobian of such a transformation is lower-triangular, which makes the computation of its determinant tractable (equals the product of its diagonal elements). The implementation of autoregressive flows is based on the choice of a transformer and a conditioner. Any combination of transformer-conditioner can be chosen from the functions presented below.

Multiple transformer types exist:

- **Affine:** Transformers can be defined as *affine* functions, which have the advantage of being simple and analytically tractable (note that having an affine transformer does not mean the whole transformation is affine). Their expressivity is, however, limited.
- **Combination-based:** Non-affine transformers can be obtained by using *conic combinations* or *compositions* of monotonic functions (which are themselves monotonic). Although they can represent any monotonic function, they cannot generally be inverted analytically and require an iterative method, such as bisection search.
- **Integration-based:** A non-affine transformation can also be constructed by computing the *integral* of a positive function (Wehenkel and Louppe,

2019), but the use of integral can hamper the possibility of analytical tractability.

- **Spline-based:** A proposed solution to overcome this analytical intractability is to implement the transformer as a monotonic *spline*. This has the additional advantages of being rapidly inverted and evaluated.

To avoid expensive computation times in high dimensional cases, parameters are in practice shared between the conditioners c_i. Several conditioner types have been developed:

- **Recurrent:** One way to allow sharing parameters between the conditioners this is to implement the conditioners jointly using a *recurrent neural network* (Oliva et al., 2018; Kingma et al., 2016). The main drawback of such technique is its requirement to run sequentially, although in theory, each $\boldsymbol{h}_i$ can be computed independently and in parallel, which makes the method slow for high-dimensional data.

- **Masked:** Another technique is based on a single feedforward neural network, where all connections that do not comply with the condition that each $\boldsymbol{h}_i$ can only depend on $\boldsymbol{\theta}_{<i}$ are *masked* (Germain et al., 2015). This method allows evaluating the parameters efficiently, given that only one neural network pass is required to compute all $\boldsymbol{h}_i$, and that all $\boldsymbol{x}_i$ can then be computed in parallel. Masked autoregressive flows are, moreover, universal approximators. However, they are not as efficient to invert as to evaluate and can even become very expensive for high-dimensional data such as images or videos.

- **Coupling layers:** A technique called *coupling layers* has been proposed to evaluate and invert the model equally rapidly (Dinh et al., 2014; Dinh et al., 2016) but has reduced expressive power. An autoregressive flow with only one coupling layer is then no longer a universal approximator. Composing multiple layers allows to increase the expressivity of the flow.

3.3.3.3 Conditional Normalizing Flows

In the context of likelihood-free inference, the goal is to approximate the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x})$. We use a conditional NF, which defines a conditional density estimator, i.e., a family of conditional p.d.f. approximators $q_\phi(\boldsymbol{\theta}|\boldsymbol{x})$ parametrized by ϕ, given a parameter $\boldsymbol{\theta}$ and an observation $\boldsymbol{x}$. A different p.d.f. approximation is obtained for each observed $\boldsymbol{x}$. The forward Kullback-Leibler divergence (see Equation 3.3) is used to quantify the closeness between the approximator $q_\phi(\boldsymbol{\theta}|\boldsymbol{x})$ and the target density $p(\boldsymbol{\theta}|\boldsymbol{x})$. The

simulator used to generate an observation $\boldsymbol{x}$ from $\boldsymbol{\theta}$ allows us to sample from the joint distribution $p(\boldsymbol{x}, \boldsymbol{\theta})$ but not to evaluate it. Therefore, contrarily to VI (Section 3.2.2), we use the forward KL divergence instead of the reverse KL (Section 3.2.2.2). Finding the best posterior approximator is then obtained by minimizing its average KL divergence with respect to the conditional density estimator for different choices of $\boldsymbol{x}$, as expressed by (Papamakarios and Murray, 2016):

$$\underset{\phi}{\text{argmin}} \quad \mathcal{L}(\phi) \quad \text{with} \quad \mathcal{L}(\phi) = \mathbb{E}_{\boldsymbol{x} \sim p(\boldsymbol{x})} \left[D_{\text{KL}}(p(\boldsymbol{\theta}|\boldsymbol{x}) \| q_\phi(\boldsymbol{\theta}|\boldsymbol{x})) \right] \ , \tag{3.22}$$

which can be rewritten as

$$\begin{aligned} \mathcal{L}(\phi) &= \int D_{\text{KL}}(p(\boldsymbol{\theta}|\boldsymbol{x}) \| q_\phi(\boldsymbol{\theta}|\boldsymbol{x})) p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \ , \\ &= -\iint \log \Big(q_\phi(\boldsymbol{\theta}|\boldsymbol{x}) \Big) p(\boldsymbol{\theta}|\boldsymbol{x}) p(\boldsymbol{x}) \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{x} + \text{const} \ , \\ &= -\iint \log \Big(q_\phi(\boldsymbol{\theta}|\boldsymbol{x}) \Big) p(\boldsymbol{x}, \boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{x} + \text{const} \ , \\ &= -\mathbb{E}_{(\boldsymbol{x}, \boldsymbol{\theta}) \sim p(\boldsymbol{x}, \boldsymbol{\theta})} \Big[\log \Big(q_\phi(\boldsymbol{\theta}|\boldsymbol{x}) \Big) \Big] + \text{const} \ , \end{aligned} \tag{3.23}$$

where the constant does not depend on ϕ. The parametrization ϕ that we obtain by the end of the optimization procedure yields a posterior which is amortized for different choices of $\boldsymbol{x}$. Thus, for a specific observation $\boldsymbol{x}_0$ we may simply write $q_\phi(\boldsymbol{\theta}|\boldsymbol{x}_0)$ to get an approximation of $p(\boldsymbol{\theta}|\boldsymbol{x}_0)$.

Note that in practice an N-sample Monte-Carlo approximation of the loss function is used, given by:

$$\mathcal{L}(\phi) \approx \mathcal{L}^N(\phi) = -\frac{1}{N} \sum_{i=1}^{N} \log \Big(q_\phi(\boldsymbol{\theta}_i|\boldsymbol{x}_i) \Big) \ , \tag{3.24}$$

where the N data points $(\boldsymbol{\theta}_i, \boldsymbol{x}_i)$ are sampled from the joint distribution with $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$ and $\boldsymbol{x}_i \sim p(\boldsymbol{x}|\boldsymbol{\theta}_i)$. We can then use stochastic gradient descent to obtain a set of parameters ϕ which minimizes $\mathcal{L}^N$.

If the class of conditional density estimators is sufficiently expressive, it is possible to show that the minimizer of Equation 3.24 converges to $p(\boldsymbol{\theta}|\boldsymbol{y})$ when $N \to \infty$ (Greenberg et al., 2019).

An important consideration here is that the inference is dubbed *amortized*, which means the inference is run only once for all observed points $\boldsymbol{x}$ (Cremer et al., 2018). This amortization is more computationally expensive than non-amortized inference setups and leads to a posterior approximation further away from the true posterior (the amortization gap). However, this

amortization can be beneficial in terms of computational time if many aci.i.d. points are observed. Indeed the training is done only once, such that new observations $\boldsymbol{x}$ can be evaluated efficiently. The cost of inference is then amortized across values of $\boldsymbol{x}$.

3.4 Application-Specific Considerations

Among all the existing methods used for Bayesian inference, multiple criteria must be considered for choosing the most appropriate model for a given application.

3.4.1 Amortization

To begin with, the notion of amortization of the method can have a high impact on the computation cost, depending on if a single data point or multiple independent and identically distributed (i.i.d.) data points are observed. Consider the inference problem for a given observation $\boldsymbol{x} = \boldsymbol{x_0}$. In a non-amortized setup, inference has to be rerun for every new observed point $\boldsymbol{x}_{i>0}$. This is notably the case for MCMC or ABC methods. Solving the inverse problem for any $\boldsymbol{x}$ is called *amortized inference*. The latter is more computationally expensive and induces a bias in the posterior approximation. We could say that such a setup has an entrance cost that can be amortized if many observations need to be evaluated.

3.4.2 Summary Statistics

The curse of dimensionality is a recurrent problem in the methods presented before. In the context of Bayesian inference, a proposed solution to address the very expensive computational cost in high dimensional cases is to turn data into a smaller number of features, known as *summary statistics* (Blum et al., 2013). These low-dimensional summary statistics are defined to capture the full information in the observations about the parameters $\boldsymbol{\theta}$. The quality of the inference is then closely related to the choice of those summary statistics. If the summary statistics are sufficient to infer $\boldsymbol{\theta}$, there is no lost information about $\boldsymbol{\theta}$. On the other hand, if some information is discarded using summary statistics instead of the raw data, a bias will be introduced in the estimations.

Finding the correct summary statistics can be difficult. In practice, their definition is often domain-specific and chosen with respect to the inference task. However, several general methods have been developed in the context of ABC methods that aim at constructing summary statistics that reduce the dimensionality of the data while minimizing information loss (Blum et al., 2013; Prangle et al., 2014; Charnock et al., 2018; Fearnhead and Prangle, 2012).

To begin with, a class of methods suggests a *best subset selection* approach. Candidate subsets are created, evaluated, and ranked according to an information-based criterium. Joyce and Marjoram (2008) suggest using a measure of

sufficiency, while Nunes and Balding (2010) propose to use the entropy of the posterior distribution and Blum et al. (2013) a criterium based on Akaike and Bayesian information. According to the chosen measure, the subset giving the highest score (i.e., the subset describing the data the best) is kept for the inference procedure. A limitation of this approach is that one better-specified summary statistic not present in the raw data could provide the same or more information than several elements of one subset.

To overcome this problem, some techniques aim to combine several data elements via linear or non-linear transformations to create a more powerful set of summary statistics with lower dimensionality. These are dubbed *projection techniques*. They rely on the addition of a regression layer in the ABC framework (Beaumont et al., 2002; Blum and François, 2010). Examples are partial least squares regression (Wegmann et al., 2009), feed-forward neural networks (Blum and François, 2010), and regression guided by minimum expected posterior loss (Fearnhead and Prangle, 2012). These techniques have the advantage of scaling well with an increasing number of summary statistics, unlike best subset selection methods, and can deal with the interdependence of raw data. A drawback of these techniques is the difficulty of interpreting those newly constructed summary statistics. In addition, most of them rely on the setting of a hyperparameter that defines the number of projections to perform.

3.5 Conclusion

Statistical models help predict how systems will behave in multiple scenarios. They generate an observation $\boldsymbol{x}$ given a parameter $\boldsymbol{\theta}$. In this chapter, we addressed the problem of inferring the parameters $\boldsymbol{\theta}$ given some observations $\boldsymbol{x}$. Solving non-linear inverse problems with a Bayesian point of view is based on two quantities: a prior distribution encoding initial knowledge of the parameter values (e.g., physiologically relevant intervals) and the likelihood function of the forward model being studied. If the statistical model is a density model, then the likelihood is easily accessible. We can then either obtain an analytic expression of the posterior distribution via Bayes' formula, use a Markov-Chain Monte Carlo procedure to numerically sample the posterior distribution, or approximate the posterior distribution using variational inference. However, in the case of a simulator model, the likelihood function is often tough to obtain and makes the Bayesian approach rather challenging to use. Likelihood-free inference bypasses this bottleneck by recurring to several simulations of the forward model using different parameters and learning an approximation to the posterior distribution from these examples. Neural network architectures specially tailored to approximate probability density functions from a set of examples, such as normalizing flows, have been developed to improve limitations of the ABC methods, the first contributions on LFI. The choice of the inference method is also guided by the number of observations to infer, that is, if an amortized technique is relevant in this case or not, and by the number of dimensions of the data, i.e., if summary statistics need to be computed from the data.

Part II

Contributions

Brain Gray Matter Modeling

4

Contents

Acronyms used in the chapter

dMRI	diffusion MRI
ECS	Extra-Cellular Space
GPD	Gaussian Phase Distribution
PGSE	Pulsed Gradient Spin Echo
RTOP	Return-To-the-Origin Probability
SANDI	Soma And Neurite Density Imaging
SM	Standard Model

Notations used in the chapter

δ	Duration of a gradient pulse [ms]
Δ	Duration which separates two gradient pulses [ms]
t	Diffusion time [ms]
b	b-value [$\mathrm{ms}\,\mu\mathrm{m}^{-2}$]
q	q-value [$\mu\mathrm{m}^{-1}$]
$\mathbf{g}$	Direction and strength [$\mathrm{T\,m}^{-1}$] of the applied diffusion weighting
$\hat{g}$	Unit vector of $\mathbf{g}$
S	Diffusion-weighted MRI signal
$\bar{S}$	Powder-average of the diffusion-weighted MRI signal
D_X	Diffusivity within compartment X [$\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$]
K	Kurtosis
f_X	Signal fraction of compartment X
r_s	Soma radius [$\mu\mathrm{m}$]
C_s	Parameter encoding soma radius and diffusivity
$\mathcal{K}$	Response kernel
t_{ex}	Exchange time [ms]
τ_X	Mean molecular lifetime in compartment X [ms]
r_X	Exchange rate from compartment X to another [ms^{-1}]
V	Volume [$\mu\mathrm{m}^3$]
A	Area [$\mu\mathrm{m}^2$]
κ	Permeability of the cell membrane [$\mu\mathrm{m\,ms}^{-1}$]

4.1 Introduction

Obtaining quantitative brain gray matter microstructure measurements with a dedicated soma representation is a growing field of interest in the dMRI community (see e.g. Palombo et al., 2020; Jelescu et al., 2021; Olesen et al., 2021). Two main approaches have emerged for estimating tissue microstructure from a dMRI acquisition signal: signal representation and tissue modeling (see Chapter 2). In this thesis, we present an approach based on tissue modeling. Solving the inverse problem, in that case, requires the definition of a biological model that describes the brain gray matter tissue.

Current brain tissue models are predominantly based on the two-compartment SM (Zhang et al., 2012; Novikov et al., 2018a). Recent evidence shows that the SM, mainly used in white matter, does not hold for gray matter microstructure analysis (Veraart et al., 2020). Several assumptions have been proposed to explain this issue, such as the abundance of cell bodies in gray matter (Shapson-Coe et al., 2021; Palombo et al., 2020), increased permeability in neurite membranes (Veraart et al., 2020; Jelescu et al., 2021; Olesen et al., 2021), or structural disorder, leading to non-Gaussian diffusion along the neurites (Lee et al., 2020). In this thesis, we follow the hypothesis that the SM does not hold due to an abundance of cell bodies in gray matter. Our proposed biophysical model comprises three compartments, following Palombo et al. (2020): neurites, somas, and ECS. In Section 4.2, we present a three-compartment model which assumes that each compartment is impermeable, i.e., there are no exchanges between them, similar to the model SANDI (Soma And Neurite Density Imaging, Palombo et al., 2020). Despite the increased complexity of such a model, its main advantage is the possibility of estimating each compartment's characteristic features together. While defining this model, we introduce a new parameter that jointly encodes soma radius and inner diffusivity without imposing constraints on these values. This new parameter reduces indeterminacies in the model and has relevant physiological interpretations. In Section 4.3, we extend the SANDI model to account for exchange between neurites and the ECS, based on the publication by Jelescu et al. (2021), which shows that this exchange has a more significant impact on dMRI signal than structural disorder. Figure 4.1 illustrates the two models used in this thesis.

Inferring brain tissue model parameters directly from dMRI signals has proven to be a challenging task because of the large dimensionality of the collected data and its stochastic nature (Fick et al., 2016). The dimension of the dMRI signal for one voxel equals the number of b-values times the number of

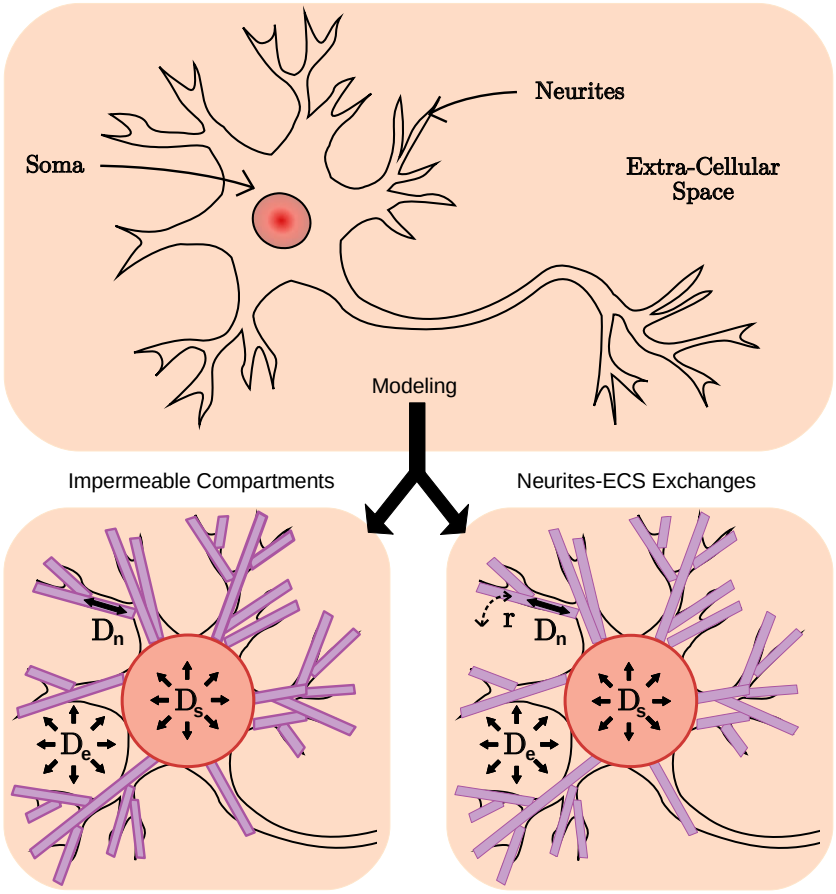


Fig. 4.1: Graphical representation of the considered gray matter models. Neurons are modeled by simple geometric forms (neurites as tubes and somas as spheres) and the ECS as isotropic Gaussian diffusion. The first model (on the left) considers cell membranes as impermeable, similar to the model SANDI (Palombo et al., 2020), while the second model also takes into account exchange between neurites and the ECS, similar to the model NEXI with a soma compartment (Jelescu et al., 2021). This illustrating is inspired from the publication by Olesen et al. (2021).

directions per b-value. For example, this corresponds to a vector of dimension 512 per voxel in the case of the HCP MGH database. Several general methods have been developed in statistical inference to reduce the dimensionality of the data to be processed (see Section 3.4.2). New approaches have also been developed within the dMRI community to address this problem. For instance, Novikov et al. (2018c) proposes the LEMONADE system of equations based on a Taylor expansion of the diffusion signal. In this chapter, we present a small and large q-value analysis for each model using boundary approximations to extract features from the observed signal. We call the resulting vectors of features the *summary statistics* of the dMRI signal. These rotationally-invariant features relate directly to the tissue parameters and enable us to solve the inverse problem without manipulating the raw dMRI signals.

In this chapter, we describe the two brain gray matter biological models mentioned before and present the signal analysis used to extract summary statistics, which allow reducing the dimensionality of the dMRI signal. We will present the inversion method and the results in Chapter 5. At last, Section 4.4 presents possible improvements to model brain gray matter more accurately.

4.1.1 Contributions

The content of this chapter is based on (and extends) the works on two papers:

> **M. Jallais**, and D. Wassermann. Indetermination-free cytoarchitecture measurements in brain gray matter via a forward diffusion MRI signal separation method. In ISMRM 2020-28th Annual Meeting Exhibition,

and

> T. Meunier, C. Fang, **M. Jallais***, and D. Wassermann*. Full posterior estimation of gray matter cytoarchitecture using a three-compartment model with exchange: a simulation-based study. Submitted to ISMRM 2022.

The first abstract presents gray matter tissue as a three-compartment model and introduces the large and small q-value approximations to extract summary statistics. In the second abstract, we extend this model to account for exchange between neurites and the ECS (Jelescu et al., 2021; Olesen et al., 2021), and propose a new pipeline for obtaining summary statistics, based on a large and small q-value analysis as well. This abstract contains preliminary results and was recently submitted to ISMRM 2022.

4.2 Impermeable Three-Compartment Model with Soma Sensitivity

4.2.1 Modeling Brain Gray Matter with a Three-Compartment Model

4.2.1.1 The Three-Compartment Model

Research in histology has shown that brain gray matter is composed of neurons embedded in a fluid environment (Shapson-Coe et al., 2021). Each neuron is composed of a soma, corresponding to the cellular body, surrounded by neurites connecting somas. Following this biophysical tissue composition, we model the gray matter tissue as three-compartmental (Palombo et al., 2020), moving away from the usual SM designed for white matter. We further assume that our acquisition protocol is insensitive to exchanges between the compartments, i.e., molecules moving from one compartment to another have a negligible influence on the signal (Palombo et al., 2020). Many approaches also include a dot compartment into the SM with zero apparent diffusivity and no exchange. However, we have not considered such an assumption because previous publications have considered its presence very unlikely in the gray matter (Veraart et al., 2019; Tax et al., 2020). The observed diffusion signal is considered as a convex mixture of signals arising from somas, neurites, and ECS. Unlike white matter-centric methods (see e.g. Jelescu and Budde, 2017b), we are not interested in the fiber orientation and only estimate orientation-independent parameters. This enables us to work on the direction-averaged dMRI signal, denoted $\bar{S}(q)$, known as the powder averaged signal (Callaghan and Soderman, 1983). This consideration mainly matters for neurites, as their signal is not isotropic, unlike the proposed model for somas and ECS. Our direction-averaged gray matter signal model is then:

$$\frac{\bar{S}(q)}{S(0)} = f_n \bar{S}_{\text{neurites}}(q, D_n) + f_s \bar{S}_{\text{somas}}(q, D_s, r_s) + f_{\text{ECS}} \bar{S}_{\text{ECS}}(q, D_e) \, . \tag{4.1}$$

In this equation, f_n, f_s, and f_{ECS} represent signal fractions for neurites, somas, and ECS respectively ($f_n + f_s + f_{\text{ECS}} = 1$). Note that the relative signal fractions do not correspond to the relative volume fractions of the tissue compartments as they are also modulated by different T2 values (Novikov et al., 2018a). D_n corresponds to axial diffusivity inside neurites, while D_s and D_e correspond to somas and extra-cellular diffusivities. r_s is the average soma radius within a voxel. This model is the same as the model SANDI proposed by Palombo et al. (2020), with $f_n = f_{ic} f_{in}$, $f_s = f_{ic} f_{is}$

and $f_{\text{ECS}} = f_{ec}$. We use q-values for more readability and harmonization throughout the chapter, but a direct conversion to b-values is also possible, using $b = (2\pi q)^2 t$ with $t = \Delta - \delta/3$.

We now review the model for each compartment to make explicit the impact of each parameter on the diffusion MRI signal.

4.2.1.2 Neurite Compartment

As in the SM, neurites are modeled as 0-radius impermeable cylinders ("sticks"), with effective diffusion along the parallel axis and a negligible radial intra-neurite diffusivity. Its direction averaged signal is, for large q-values (Callaghan et al., 1979; Veraart et al., 2019):

$$\bar{S}_{\text{neurites}}(q) \simeq \frac{1}{4\sqrt{\pi t D_n}} \cdot q^{-1} \, . \tag{4.2}$$

4.2.1.3 Soma Compartment

Somas are modeled as spheres, whose signal can be computed using the GPD approximation (Balinov et al., 1993):

$$-\log \bar{S}_{\text{somas}}(q) = C(r_s, D_s) \cdot q^2 \, . \tag{4.3}$$

We exploit this relation here to extract a parameter $C_s = C(r_s, D_s)[m^2]$ which, at fixed diffusivity D_s, is modulated by the radius of the soma r_s:

$$\begin{aligned} C(r_s, D_s) \quad = \quad & \tfrac{2}{D_s \delta^2} \textstyle\sum_{m=1}^{\infty} \tfrac{\alpha_m^{-4}}{\alpha_m^2 {r_s}^2 - 2} \\ & \cdot \left(2\delta - \tfrac{2 + e^{-\alpha_m^2 D_s(\Delta-\delta)} - e^{-\alpha_m^2 D_s \delta} - e^{-\alpha_m^2 D_s \Delta} + e^{-\alpha_m^2 D_s(\Delta+\delta)}}{\alpha_m^2 D_s} \right) \, , \end{aligned} \tag{4.4}$$

where α_m is the mth root of $(\alpha r_s)^{-1} J_{\frac{3}{2}}(\alpha r_s) = J_{\frac{5}{2}}(\alpha r_s)$, with $J_n(x)$ the Bessel functions of the first kind. Figure 4.2 (left) presents the relationship between soma radius, diffusivity, and the newly introduced parameter C_s. For a fixed diffusivity within somas, C_s is a proxy for soma radius.

In certain specific cases, C_s has a simpler and more interpretable expression. For instance, consider a narrow pulse regime, with small t or large r_s. In this case, molecules do not have enough time to hit the walls of the sphere and, as a result, C_s loses its dependence on r_s (Balinov et al., 1993). We obtain:

$$C_s = D_s t \, . \tag{4.5}$$

In the Neuman (wide pulse) regime, i.e., when $D_s \Delta \gg r_s^2$ and $D_s \delta \ll 1$, molecules hit the boundaries of the sphere many times, making it impossible

to estimate their diffusivity from a dMRI signal. C_s becomes thus only dependent on the soma radius (Murday and Cotts, 1968):

$$C_s = \frac{1}{5} r_s^2 \, . \tag{4.6}$$

When the acquisition requirements are met, these two approximations allow to better interpret the parameter C_s.

4.2.1.4 Extra-Cellular Space Compartment

The extra-cellular space is approximated as isotropic Gaussian diffusion, i.e., a mono-exponential diffusion signal with a scalar diffusion constant D_e, which reflects the molecular viscosity of the fluid. This approximation assumes that the ECS is fully connected. The approximation is therefore:

$$-\log(\bar{S}_{\text{ECS}}(q)) = (2\pi q)^2 t D_e \, . \tag{4.7}$$

Because of the geometry of the problem, we estimate D_e as equal to one-third of the diffusivity in the ventricles (considered as free diffusivity), given the same metabolic composition of the extra-cellular fluid and ventricles (Vincent et al., 2021).

4.2.2 dMRI Summary Statistics

The tissue model presented in Section 4.2.1 relates the dMRI signal with parameters representing gray matter tissue microstructure. However, solving the inverse problem directly from Equation 4.1 is a difficult task, leading to indeterminacies and poorly estimated parameters with large variability. Current methods addressing this issue have not studied its stability (Palombo et al., 2020), but simpler models with only two compartments have been shown to be indeterminate (Novikov et al., 2018a).

To produce a method that addresses this indeterminacy, we introduce rotationally invariant summary statistics to describe the dMRI signal. The goal is to reduce the dimensionality of the data at hand and represent all the relevant information for statistical inference with a few features. These dMRI-based summary statistics are extracted from our proposed model presented in Section 4.2.1 via the following analysis of the dMRI signal on the boundaries of large and small q-value cases.

4.2.2.1 Large q-value Approximation: RTOP

We compute a q-bounded Return-To-the-Origin Probability (RTOP), which measures the restrictions of the diffusing fluid molecule motion and provides information about the structure of the media (Mitra et al., 1995):

$$\mathrm{RTOP}(q) = 4\pi \int_0^q \frac{\bar{S}(\eta)}{S(0)} \eta^2 d\eta \ . \tag{4.8}$$

For q large enough, the RTOP in our three-compartment model in Equation 4.1 yields a soma and extra-cellular signal which converges towards a constant value in q, while the neurites' contribution becomes quadratic in q. In this case, RTOP becomes:

$$\mathrm{RTOP}(q)|_{q\to\infty} = \underbrace{f_s \left(\frac{\pi}{C_s}\right)^{3/2} + \frac{f_{\mathrm{ECS}}}{8(\pi t D_e)^{3/2}}}_{a_{\mathrm{fit}}} + \underbrace{\frac{f_n}{2} \cdot \sqrt{\frac{\pi}{tD_n}}}_{b_{\mathrm{fit}}} \cdot q^2 + \gamma q^3 + O\left(\frac{1}{q^4}\right) \ . \tag{4.9}$$

The last term of the equation, γq^3, is a nuisance parameter that describes a constant noise in the direction averaged signal. It has been added experimentally to capture excess nuisance. By accurately estimating the second derivative of $\mathrm{RTOP}(q)$ at q large enough, we can solve the coefficients of interest of the polynomial in Equation 4.9: a_{fit} and b_{fit}. We do this efficiently by casting it as an overdetermined ordinary least squares problem with a unique solution.

The influence of soma diameter onto the signal RTOP for a varying q_{max} is presented in Figure 4.2 (right). Estimating non-biased summary statistics from the RTOP approximation requires smaller q-values for big somas than small somas. Consequently, it will be easier for a given acquisition to get stable estimates from large somas.

4.2.2.2 Small q-value Approximation: Spiked LEMONADE

We propose a second approximation, based on a moment decomposition for small q-values (Novikov et al., 2018c):

$$\left.\frac{S_{\hat{\mathbf{g}}}(q)}{S(0)}\right|_{q\to 0^+} = 1 - b(q) M^{(2)}_{i_1 i_2} g_1 g_2 + \frac{b(q)^2}{2!} M^{(4)}_{i_1 \ldots i_4} g_1 \ldots g_4 - \ldots, \quad b(q) = (2\pi q)^2 t \ , \tag{4.10}$$

where i_k are the directional basis of the tensors M, $g_k = i_k \cdot \hat{\mathbf{g}} \in \mathbb{R}^3$, and $\hat{\mathbf{g}}$ the unit direction of the dMRI acquisition. From the moment tensors of this decomposition, Novikov et al. (2018c) extract rotational invariant scalar indices $M^{(i),j}$, $i, j \in \{0, 2, 4, \ldots\}$ and relate them to the SM parameters by

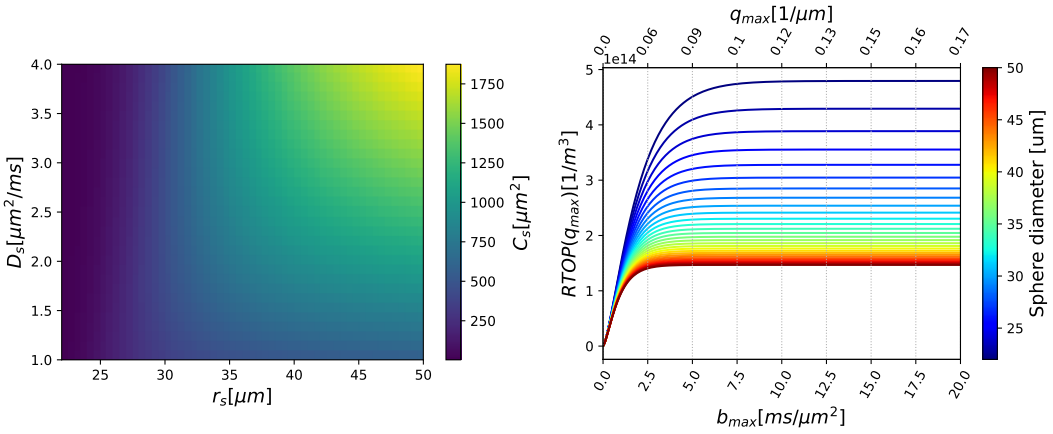


Fig. 4.2: Left: C_s values for multiple soma radius r_s and diffusivity D_s. Multiple combinations of (r_s, D_s) can generate the same C_s. Using C_s instead of r_s and D_s avoids indeterminacies and does not require to fix soma diffusivity to a predefined value. Right: Relationship between soma diameter and RTOP. RTOP(q_{max}) depends on C_s, itself directly related to the sphere diameter (for a fixed soma diffusivity). RTOP(q_{max}) converges faster towards a value that depends on C_s for big somas than small somas. The estimation of C_s for a fixed q_{max} will therefore be less biased for big somas.

plugging the two-compartment SM into Equation 4.10 (see Novikov et al., 2018c, app. C). They end up with a system of equations they call LEMONADE.

We extended LEMONADE to our three-compartment model presented in Section 4.2.1 by plugging Equation 4.1 into Equation 4.10 and performing tedious arithmetic. This results in the following system of equations, which now includes the soma parameter C_s, relating the dMRI signal to gray matter microstructure:

$$\begin{cases} M^{(2),0} = f_n D_n + 3 f_s \frac{C_s}{(2\pi)^2 t} + 3 f_{\text{ECS}} D_e \\ M^{(2),2} = f_n D_n p_2 \\ M^{(4),0} = f_n {D_n}^2 + 5 f_s \left(\frac{C_s}{(2\pi)^2 t}\right)^2 + 5 f_{\text{ECS}} D_e^2 \\ M^{(4),2} = f_n {D_n}^2 p_2 \end{cases} \tag{4.11}$$

where p_2 is a scalar measure of neurite orientation dispersion (Novikov et al., 2018c).

Note that this approximation is only valid for small b-values. As a consequence, only the shells with $b(q) \leq 2.5\,\text{ms}\,\mu\text{m}^{-2}$ are used, to get an unbiased estimation of the rotational invariant moments $M^{(2),0}$, $M^{(2),2}$, $M^{(4),0}$ and $M^{(4),2}$, following Novikov et al. (2018c).

4.2.2.3 Complete System

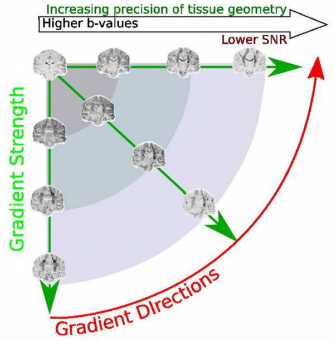
Combining Equations 4.9 and 4.11 and adding the constraint that the fractions for the three compartments sum to one, we obtain a non-linear system of 7 equations and 7 unknowns. Following Menon et al. (2020), we assume D_e nearly-constant per subject acquisition and estimate it as one-third of the mean diffusivity in the subject's ventricles (Vincent et al., 2021). This assumption allows us to drop an unknown from the system, use D_e as a reference diffusivity, and turn our system of equations unitless with $D_n^u = \frac{D_n}{D_e}$ and $C_s^u = \frac{C_s}{(2\pi)^2 t D_e}$, which gives:

Spiked LEMONADE Small q-values	**RTOP** Large q-values
$\begin{cases} \frac{M^{(2),0}}{D_e} = f_n D_n^u + 3 f_s C_s^u + 3 f_{\text{ECS}} \\ \frac{M^{(2),2}}{D_e} = f_n D_n^u \cdot p_2 \\ \frac{M^{(4),0}}{D_e^2} = f_n D_n^{u2} + 5 f_s C_s^{u2} + 5 f_{\text{ECS}} \\ \frac{M^{(4),2}}{D_e^2} = f_n D_n^{u2} \cdot p_2 \end{cases}$	$\begin{cases} a_{\text{fit}} \left(t D_e\right)^{3/2} = \frac{f_s}{8(\pi C_s^u)^{3/2}} + \frac{f_{\text{ECS}}}{8\pi^{3/2}} \\ b_{\text{fit}} \left(t D_e\right)^{1/2} = \frac{f_n}{2}\sqrt{\frac{\pi}{D_n^u}} \end{cases}$
$f_n + f_s + f_{\text{ECS}} = 1$	

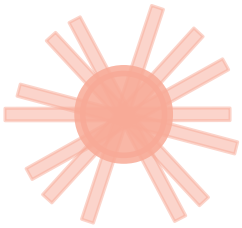
Figure 4.3 summarizes this approach.

In this section, we addressed the problem of the SM not holding in brain gray matter by explicitly modeling somas, more abundant in gray matter than white matter. We presented a three-compartment model describing brain gray matter based on the biological composition of the tissue. Another assumption explaining that gray matter tissue cannot be described using the SM is the presence of a non-negligible exchange between neurites and the ECS. The next section presents an expansion of this model, taking into account this hypothesis.

Multi shell MRI samples along different gradient directions and b-values



Sphere and tube modelisation of soma and dendrites



SPIKED LEMONADE

- Cumulant tensors
- Rotational Invariant moment tensors

RTOP(q_{max})

4.3 Three-Compartment Model with Exchange

In white matter, axons are surrounded by myelin sheaths that guarantee their impermeability over the NMR relevant timescales. However, these myelin sheaths are limited in gray matter, making the exchange between neurites and ECS non-negligible anymore for typical clinical diffusion times ($20 < t < 80\text{ms}$) (Jelescu et al., 2021). We hypothesize that exchange between somas and neurites is negligible, similar to SANDI. In addition, we suppose that exchange between somas and ECS can be neglected due to the small surface-to-volume ratio of somas with respect to neurites (Olesen et al., 2021). In the human cortex, the in-vivo exchange time has been estimated between 10 and $30\,\text{ms}$ (Veraart et al., 2020).

Current exchange methods are considering models without soma, such as NEXI (Jelescu et al., 2021), or propose unstable results (Olesen et al., 2021) (see Chapter 5). We present a new forward model that considers both water exchange and soma, along with summary statistics that allow solving the inverse problem with more stability.

4.3.1 Three-Compartment Model with Neurite Exchanges

Similar to the previous section describing the three-compartment model, we model neurites as sticks with effective diffusion along the parallel axis D_n and a negligible radial intra-neurite diffusivity. Somas are modeled as spheres whose diffusivity and radius are encoded in a parameter C_s introduced in Equation 4.4. Its signal follows the GPD approximation. The ECS is approximated as isotropic Gaussian diffusion with diffusivity D_e.

We define the powder average signal of the three-compartment model with exchange as a sum of the signal arising from the exchanging compartments, and non-exchanging ones:

$$\frac{\bar{S}(q,t)}{\bar{S}(0)} = f^{ex} \underbrace{\bar{S}_{ex}(q,t)}_{\text{Neurites and ECS}} + (1-f^{ex}) \underbrace{\bar{S}_{non\_ex}(q)}_{\text{Somas}} \tag{4.12}$$

In the considered model, we only have one impermeable compartment that is not exchanging with the other ones, that is:

$$\begin{aligned} \bar{S}_{non\_ex}(q) &= \bar{S}_{\text{somas}}(q) = e^{-C_s q^2}, \\ 1-f^{ex} &= f_s. \end{aligned} \tag{4.13}$$

The exchanging compartments correspond to the neurites and the ECS, with respective signal proportions f_n^{ex} and f_{ECS}^{ex} ($f_n^{ex} + f_{\text{ECS}}^{ex} = 1$). To approximate

this two-compartment model with exchange, we are using the Kärger model (see Section 2.4.3.2) with mean lifetimes τ_n and τ_{ECS}, exchange rates $r_n = \frac{1}{\tau_n}$ and $r_{\text{ECS}} = \frac{1}{\tau_{\text{ECS}}}$ and exchange time $t_{ex} = \frac{1}{r_n + r_{\text{ECS}}}$. The signal response of such a model is the following:

$$\mathcal{K}_{ex}(q, t, \mathbf{g}\cdot\mathbf{n}, D_n, D_e, \tau_n, \tau_{\text{ECS}}) = f'_n e^{-D'_n(2\pi q)^2 t} + f'_{\text{ECS}} e^{-D'_e(2\pi q)^2 t} \tag{4.14}$$

with

$$\begin{aligned} D'_{n/e} &= \frac{1}{2}\Bigg[D_n(\mathbf{g}\cdot\mathbf{n})^2 + D_e + \frac{1}{(2\pi q)^2}\left(\frac{1}{\tau_n} + \frac{1}{\tau_{\text{ECS}}}\right) \\ &\quad \pm \sqrt{\left(D_e - D_n(\mathbf{g}\cdot\mathbf{n})^2 + \frac{1}{(2\pi q)^2}\left(\frac{1}{\tau_{\text{ECS}}} - \frac{1}{\tau_n}\right)\right)^2 + \frac{4}{(2\pi q)^4 \tau_n \tau_{\text{ECS}}}}\Bigg], \\ f'_n &= 1 - f'_{\text{ECS}}, \\ f'_{\text{ECS}} &= \frac{f_n^{ex} D_n(\mathbf{g}\cdot\mathbf{n})^2 + f_{\text{ECS}}^{ex} D_e - D'_n}{D'_e - D'_n} \end{aligned} \tag{4.15}$$

Using the conservation of mass relation $\frac{f_n^{ex}}{\tau_n} = \frac{f_{\text{ECS}}^{ex}}{\tau_{\text{ECS}}}$, $D'_{n/e}$ can be rewritten as (Jelescu et al., 2021):

$$D'_{n/e} = \frac{1}{2}\left[D_n(\mathbf{g}\cdot\mathbf{n})^2 + D_e + \frac{1}{(2\pi q)^2 t_{ex}} \pm \sqrt{\left(D_e - D_n(\mathbf{g}\cdot\mathbf{n})^2 + \frac{2f_n^{ex} - 1}{(2\pi q)^2 t_{ex}}\right)^2 + \frac{4 f_n^{ex} f_{\text{ECS}}^{ex})}{(2\pi q)^4 t_{ex}^2}}\right]. \tag{4.16}$$

The powder-average signal of the exchange compartment is then the following:

$$\frac{\bar{S}_{ex}(q,t)}{\bar{S}(0)} = \int_0^1 \mathcal{K}_{ex}(q, t, \mathbf{g}\cdot\mathbf{n}, D_n, D_e, \tau_n, \tau_{\text{ECS}}) d(\mathbf{g}\cdot\mathbf{n}) \tag{4.17}$$

We aim at estimating all the tissue parameters, i.e., f_n^{ex}, f_{ECS}^{ex}, f_s, D_n, D_e, C_s and t_{ex}, without indeterminacy. Similar to previous section, we propose to extract summary statistics that relate the dMRI signal with the tissue parameters, using large and small q-value approximations.

4.3.2 dMRI Summary Statistics

4.3.2.1 Large q-value Approximation

Obtaining an analytical formula for the total signal $\bar{S}(q,t)$ is difficult, due to the integral of Equation 4.17. We therefore first restrict ourselves to the case

with large q-values. We re-write the signal response of the exchange part as follows:

$$\mathcal{K}_{ex}(q,t,\mathbf{g}\cdot\mathbf{n},D_n,D_e,\tau_n,\tau_{\mathrm{ECS}})|_{q\to\infty} = f'_{n,\infty}e^{-D'_{n,\infty}(2\pi q)^2 t} + f'_{\mathrm{ECS},\infty}e^{-D'_{e,\infty}(2\pi q)^2 t} \tag{4.18}$$

To obtain $D'_{n,\infty}$ and $D'_{n,\infty}$, we perform a Taylor expansion at $\frac{1}{q}\to 0$ ($q\to\infty$) of D'_n and D'_e. We obtain:

$$\begin{cases} D'_{n,\infty} = D_n(\mathbf{g}\cdot\mathbf{n})^2 + \frac{f_{\mathrm{ECS}}}{(2\pi q)^2 t_{ex}} - \frac{f_n f_{\mathrm{ECS}}}{(2\pi q)^4 t_{ex}^2 (D_e - D_n(\mathbf{g}\cdot\mathbf{n})^2)} + O\left(\frac{1}{q^5}\right) \\ D'_{e,\infty} = D_e + \frac{f_n}{(2\pi q)^2 t_{ex}} + \frac{f_n f_{\mathrm{ECS}}}{(2\pi q)^4 t_{ex}^2 (D_e - D_n(\mathbf{g}\cdot\mathbf{n})^2)} + O\left(\frac{1}{q^5}\right) \end{cases} \tag{4.19}$$

We then inject those formulas into Equation 4.15:

$$f'_{n,\infty} = f_n^{ex} + 2f_n^{ex}\frac{f_{\mathrm{ECS}}^{ex}}{(2\pi q)^2 t_{ex} D_e} + O\left(\frac{1}{q^3}\right) \tag{4.20}$$

Integrating over all directions, we obtain the following large q-value approximation of the exchange powder-average signal:

$$\begin{aligned} \bar{S}_{ex}(q,t)\Big|_{q\to\infty} \quad = \quad & \frac{1}{4q\sqrt{\pi t D_n}}\left(f_n + \frac{2f_n f_{\mathrm{ECS}}}{(2\pi q)^2 t_{ex} D_e} + O\left(\frac{1}{q^3}\right)\right) \\ & \cdot \exp\left[-f_{\mathrm{ECS}}\left(\frac{t}{t_{ex}}\right) + \frac{f_n f_{\mathrm{ECS}}}{(2\pi q)^2 t D_e}\left(\frac{t}{t_{ex}}\right)^2 + O\left(\frac{1}{q^3}\right)\right] \\ & + \left(1 - f_n - 2f_n\frac{f_{\mathrm{ECS}}}{(2\pi q)^2 t_{ex} D_e} + O\left(\frac{1}{q^3}\right)\right) \\ & \cdot \exp\left[-(2\pi q)^2 t D_e - f_n\left(\frac{t}{t_{ex}}\right) - \frac{f_n f_e}{(2\pi q)^2 t D_e}\left(\frac{t}{t_{ex}}\right)^2 + O\left(\frac{1}{q^3}\right)\right] \end{aligned} \tag{4.21}$$

We finally perform a second Taylor approximation in large q-values to simplify the exponentials, and add the soma compartment. The resulting total powder-average signal is:

$$\begin{aligned}\frac{\bar{S}(q,t)}{\bar{S}(0)}\Big|_{q\to\infty} &= (1-f_s)\\ &\cdot\left(\frac{f_n^{ex}}{4q\sqrt{\pi t D_n}}\exp\left[-f_{\text{ECS}}^{ex}\left(\frac{t}{t_{ex}}\right)\right]\left(1+\frac{2f_{\text{ECS}}^{ex}\left(\frac{t}{t_{ex}}\right)+f_n^{ex}f_{\text{ECS}}^{ex}\left(\frac{t}{t_{ex}}\right)^2}{(2\pi q)^2 t D_e}+O\left(\frac{1}{q^3}\right)\right)\right.\\ &+f_{\text{ECS}}^{ex}\exp\left[-f_n^{ex}\left(\frac{t}{t_{ex}}\right)\right]\left(1-\frac{2f_n^{ex}\left(\frac{t}{t_{ex}}\right)+f_n^{ex}f_{\text{ECS}}^{ex}\left(\frac{t}{t_{ex}}\right)^2}{(2\pi q)^2 t D_e}+O\left(\frac{1}{q^3}\right)\right)\\ &\left.\cdot\exp\left[-(2\pi q)^2 t D_e\right]\right)\\ &+f_s\exp\left[-C_s q^2\right]\end{aligned} \tag{4.22}$$

Using this equation, we illustrate the impact of exchange in a three-compartment model signal in Figure 4.4. The considered tissue parameters correspond to the characteristics of a pyramidal neuron (03b_pyramidal3aACC) present in the `NeuroMorpho` database: $f_s = 0.58$, $f_n^{ex} = 0.20$, $f_{\text{ECS}}^{ex} = 0.80$ and $r_s = 11.3\,\mu\text{m}$. Diffusivities are fixed as follows: $D_s = 2.0\,\mu\text{m}^2\,\text{ms}^{-1}$ (which leads to $C_s = 519\,\mu\text{m}^2$), $D_n = 2.0\,\mu\text{m}^2\,\text{ms}^{-1}$ and $D_{\text{ECS}} = 0.97\,\mu\text{m}^2\,\text{ms}^{-1}$. The acquisition times Δ and δ are chosen to mimic the HCP MGH acquition protocol, i.e. $\Delta = 21.8$ ms and $\delta = 12.9$ ms. A variation up to 30% for realistic cases with $t_{ex} \geq 20\text{ms}$ can be observed. The significant impact of exchange on the signal justifies the need of modeling brain gray matter with exchange.

An interesting comparison between the large q-values approximations used in NEXI (Jelescu et al., 2021) or SMEX (Olesen et al., 2021) and Equation 4.22 can be made. Their large q-value approximation of the neurites and ECS compartments is equal to the second line of Equation 4.22. The q^{-1} term corresponds to the power-law scaling of neurites (Veraart et al., 2020), and the q^{-3} term is due to slow exchange (Olesen et al., 2021). The third and fourth lines of Equation 4.22 arise from the ECS, which is considered negligible for large q-values in NEXI and SMEX. We explored the impact of this term in Figure 4.5, using the same parameters as Figure 4.4. For b-values $\geq 6\ \text{ms}\,\mu\text{m}^{-2}$ (red vertical line), we can see that the relative error is closed to 0, indicating that the additional term expressing the ECS can therefore be neglected, as done by Jelescu et al. (2021) and Olesen et al. (2021). However, it is not always possible to respect the constraint of using only b-values $\geq 6\ \text{ms}\,\mu\text{m}^{-2}$ when fitting the large q-value approximation in practice. Indeed, datasets do not always contain a lot of b-values superior

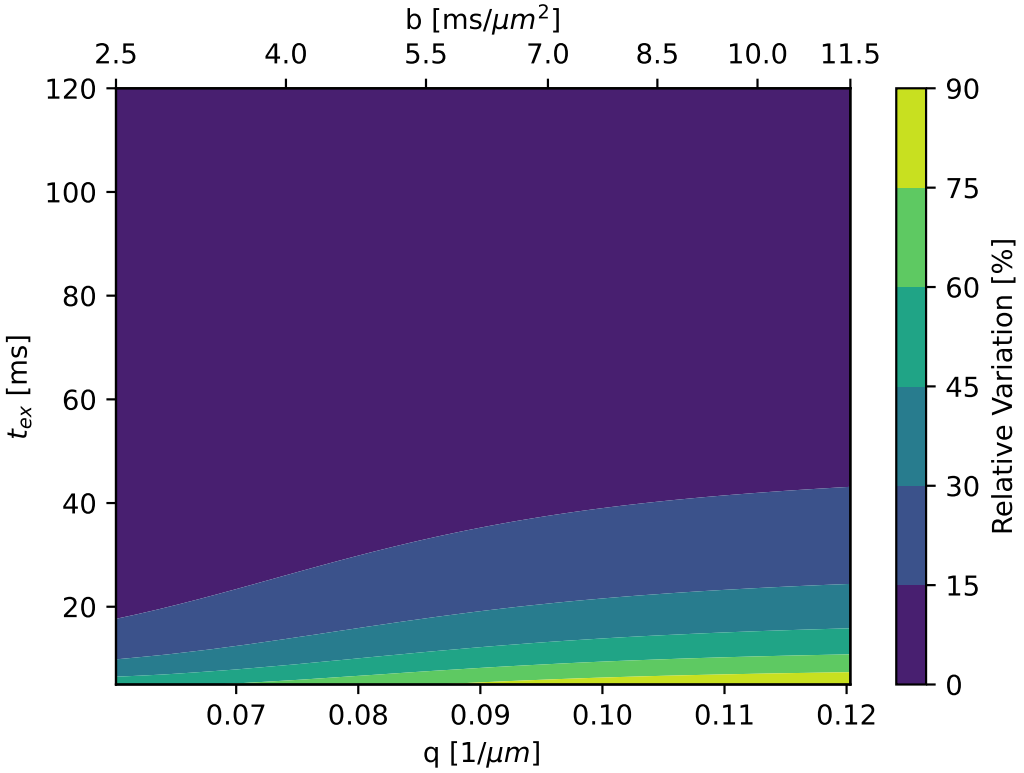


Fig. 4.4: Relative variation of the powder-averaged signal approximated for large q-values with regard to the direction-averaged signal computed for $t_{ex} = 120\text{ms}$. Exchange is considered undetectable at such high diffusion exchange times, and this regime is close to the SANDI model (Palombo et al., 2020). A clear dependence of t_{ex} on the signal is observed, with a variation up to 30% for realistic cases with $t_{ex} \geq 20$ ms.

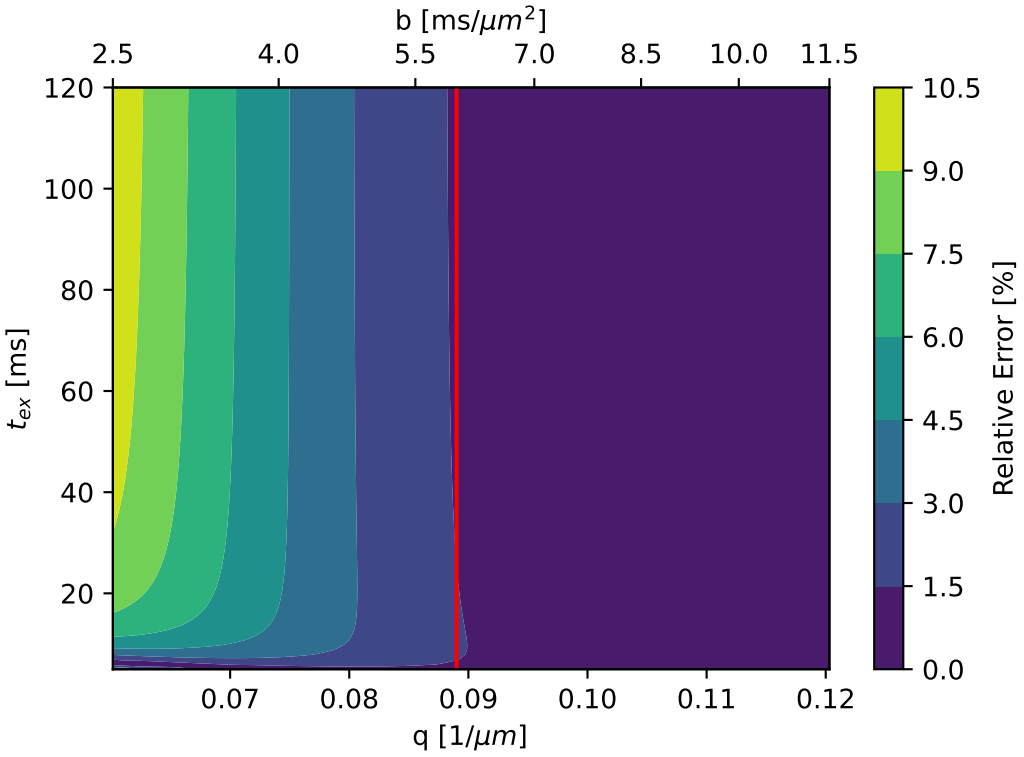


Fig. 4.5: Relative error between the SMEX model (Olesen et al., 2021), equivalent to the NEXI (Jelescu et al., 2021) model with soma compartment, and Equation 4.22. Parameters are defined identically to Figure 4.4. The ECS compartment can be neglected for b-values superior to $6\ \mathrm{ms}\mu\mathrm{m}^{-2}$ (vertical red line). However, its contribution to the total signal can be considered as non-negligible when fitting the approximation to datasets with a limited number of b-values.

to $6\ \mathrm{ms}\mu\mathrm{m}^{-2}$. For example, the HCP MGH dataset only contains signal acquisitions for $b = 0, 1, 3, 5$ and $10\ \mathrm{ms}\mu\mathrm{m}^{-2}$. Using the three-largest b-values, a relative error up to 7.5% could be introduced in that case. For that reason, we decided to keep this term in the definition of the large q-values approximation.

Given the difficulty of inverting Equation 4.22, we introduced a "truncated RTOP" for large q-values, benefiting from the numerical stability of the integral operator:

$$\begin{aligned} RTOP(q_{min}, q)\big|_{q\to\infty} &= 4\pi \int_{q_{min}}^{q} \frac{\bar{S}(\eta,t)}{S(0)}\Big|_{q\to\infty} \eta^2 d\eta \\ &= a_{fit} + b_{fit} q^2 + c_{fit} \log q \end{aligned} \tag{4.23}$$

with

$$\begin{cases} a_{fit} &= f_s\left(\left(\frac{\pi}{C_s}\right)^{\frac{3}{2}}\text{erfc}\left(\sqrt{C_s}q_{min}\right)+2\pi q_{min}\frac{e^{-C_s {q_m}in^2}}{C_s}\right) \\ &\quad -(1-f_s)f_n^{ex}e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}}\sqrt{\frac{\pi}{tD_n}}\left(\frac{q_{min}^2}{2}+\frac{\eta_{ex}}{(2\pi)^2tD_e}\log(q_{min})\right) \\ &\quad +(1-f_s)f_{\text{ECS}}^{ex}e^{-f_n^{ex}\frac{t}{t_{ex}}}\left(\frac{1}{(4\pi tD_e)^{3/2}}(1-2\widetilde{\eta}_{ex})+\frac{(2\pi)^2q_{min}e^{-(2\pi)^2tD_eq_{min}^2}}{tD_e}\right), \\ b_{fit} &= (1-f_s)f_n^{ex}e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}}\sqrt{\frac{\pi}{4tD_n}}, \\ c_{fit} &= (1-f_s)f_n^{ex}e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}}\sqrt{\frac{\pi}{tD_n}}\frac{\eta_{ex}}{(2\pi)^2tD_e}, \end{cases} \tag{4.24}$$

where $\eta_{ex}=2f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}+f_n^{ex}f_{\text{ECS}}^{ex}(\frac{t}{t_{ex}})^2$ and $\widetilde{\eta}_{ex}=2f_n^{ex}\frac{t}{t_{ex}}+f_n^{ex}f_{\text{ECS}}^{ex}(\frac{t}{t_{ex}})^2$.

Those a_{fit}, b_{fit} and c_{fit} are the first three summary statistics that we will use to invert the system.

Comparing Equation 4.24 to Equation 4.9, the γq^3 error term that was experimentally added in Section 4.2.2.1 might be capturing exchange between the neurites and the ECS.

4.3.2.2 Small q-value Approximation

The small q-value approximation is based on a cumulant decomposition of the powder-averaged signal. For more readability of the formula in this section, we will write the equations with respect to the b-values instead of the q-values, given $b=(2\pi q)^2t$. The approximation for the exchange part of Equation 4.12 is given by Jelescu et al. (2021) as follows:

$$\bar{S}_{ex}(b,t)\Big|_{b\to 0}=\exp\left[-b\bar{D}+\frac{b^2\bar{D}^2\bar{K}}{6}+O(b^3)\right] \tag{4.25}$$

with

$$\begin{aligned} \bar{D} &= \tfrac{1}{3}\left[f_n^{ex}D_n+3f_{ECS}^{ex}D_e\right], \\ \bar{K} &= K_0F(\tfrac{t}{t_{ex}})+K_\infty, \end{aligned} \tag{4.26}$$

and

$$\begin{aligned} K_0 &= \frac{3f_n^{ex}f_{ECS}^{ex}\left[D_{ECS}^2-\frac{2}{3}D_{ECS}D_n+\frac{1}{5}D_n^2\right]}{\bar{D}^2}, \\ K_\infty &= \tfrac{4}{15}\left(\frac{f_n^{ex}D_n}{\bar{D}}\right)^2, \\ F(\tfrac{t}{t_{ex}}) &= \frac{2\left(e^{-\frac{t}{t_{ex}}}-1+\frac{t}{t_{ex}}\right)}{\left(\frac{t}{t_{ex}}\right)^2}. \end{aligned} \tag{4.27}$$

The total signal approximation for small b-values is then the following:

$$\left.\frac{\bar{S}(q,t)}{S(0)}\right|_{b\to 0} = (1-f_s)\exp\left[-b\bar{D}+\frac{1}{6}b^2\bar{D}^2\bar{K}+O\left(b^3\right)\right]+f_s\exp\left[-\frac{C_s}{(2\pi)^2 t}b\right] \tag{4.28}$$

which, after a Taylor expansion, can be reduced to:

$$\left.\frac{\bar{S}(q,t)}{S(0)}\right|_{b\to 0} = \exp(-b\widetilde{D}+\frac{1}{6}b^2\widetilde{D}^2\widetilde{K}+O\left(b^3\right)) \tag{4.29}$$

with

$$\begin{cases} \widetilde{D} &= f_s\frac{C_s}{(2\pi)^2 t}+(1-f_s)\,\bar{D} \\ \widetilde{K} &= 3\frac{(1-f_s)}{\widetilde{D}^2}\left[f_s\bar{D}^2-2f_s\frac{\bar{D}}{D_e}\frac{C_s}{(2\pi)^2 t}+\frac{\bar{D}^2\bar{K}}{3}+f_s\left(\frac{C_s}{(2\pi)^2 t}\right)^2\right] \end{cases} \tag{4.30}$$

$\widetilde{D}$ and $\widetilde{K}$ correspond to the two additional summary statistics of our model.

4.3.2.3 Complete System

Similar to the previous model presented in Section 4.2, we consider D_e nearly constant per subject acquisition and estimate it as one-third of the mean diffusivity in the subject's ventricles (Vincent et al., 2021). Combining Equations 4.24 and 4.30 from the high and small q-values approximations and the constraint $f_n^{ex}+f_{\text{ECS}}^{ex}=1$, we obtain a system of 6 equations for 6 unknowns.

Cumulant decomposition
Small q-values
$\begin{cases} \frac{\widetilde{D}}{D_e} &= f_s C_s^u+(1-f_s)\,\frac{\bar{D}}{D_e} \\ \widetilde{K} &= 3\frac{(1-f_s)}{\widetilde{D}^2}\left[f_s\bar{D}^2-2f_s\frac{\bar{D}}{D_e}\frac{C_s}{(2\pi)^2 t}+\frac{\bar{D}^2\bar{K}}{3}+f_s\left(\frac{C_s}{(2\pi)^2 t}\right)^2\right] \end{cases}$
q-bounded RTOP
Large q-values
$\begin{cases} a_{fit}(tD_e)^{3/2} &= f_s\left(\left(\frac{1}{(4\pi C_s^u)}\right)^{\frac{3}{2}}\text{erfc}\left(\sqrt{C_s}q_{min}\right)+\frac{\sqrt{tD_e}}{2\pi}C_s^u q_{min}e^{-C_s q_m in^2}\right) \\ &\quad -(1-f_s)f_n^{ex}\sqrt{\frac{\pi}{D_n^u}}(tD_e)e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}}\left(\frac{q_{min}^2}{2}+\frac{\eta_{ex}}{(2\pi)^2 tD_e}\log(q_{min})\right) \\ &\quad +(1-f_s)f_{\text{ECS}}^{ex}e^{-f_n^{ex}\frac{t}{t_{ex}}}\left(\frac{(1-2\widetilde{\eta}_{ex})}{(4\pi)^{3/2}}+\sqrt{tD_e}(2\pi)^2 q_{min}e^{-(2\pi)^2 tD_e q_{min}^2}\right) \\ b_{fit}(tD_e)^{1/2} &= (1-f_s)f_n^{ex}\sqrt{\frac{\pi}{4D_n^u}}e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}} \\ c_{fit}(tD_e)^{3/2} &= (1-f_s)f_n^{ex}\frac{\eta_{ex}}{(2\pi)^2}e^{-f_{\text{ECS}}^{ex}\frac{t}{t_{ex}}}\sqrt{\frac{\pi}{D_n^u}} \end{cases}$
$f_n^{ex}+f_{\text{ECS}}^{ex}=1$

with $D_n^u=\frac{D_n}{D_e}$ and $C_s^u=\frac{C_s}{(2\pi)^2 tD_e}$.

4.4 Towards a More Complex Model

Section 4.2 presented a three-compartment model where each compartment was considered impermeable. Section 4.3 expanded this model to account for exchange between neurites and the ECS. A following interrogation is whether this model is complex enough to represent brain gray matter correctly.

A proposed approach to answer this question is to compare the diffusion signal of a neuron with the approximations obtained using our model. To this end, we generated the diffusion signals of several neurons available in `NeuroMorpho.org`. We reconstructed the surface mesh given the provided neuron morphology for each neuron and computed a tight-wrapped envelop around it to model ECS. We then used the open-source SpinDoctor software, which includes a matrix formalism method with permeable interfaces (Agdestein et al., 2021), to generate the diffusion signal of the neurons. The minimum length scale of the eigenfunctions has been set to $1\mu\mathrm{m}$, which gives a relative error inferior to 0.1% compared to the finite element method. Our method assumes that the somas are always impermeable. The SpinDoctor simulations assume an equal permeability for the entire neuron membranes.

Figure 4.6 presents the results obtained for one pyramidal neuron (03b_pyramidal3aACC). These neuron parameters are identical to Figure 4.4. We can visualize the pyramidal neuron on the left, encompassed in a green envelope representing the ECS. On the right, we plotted the powder-averaged signal obtained from the SpinDoctor simulations and the large q-values approximations (Equation 4.22) for different exchange times t_{ex}. For $t_{ex} \geq 40\mathrm{ms}$, simulations and large q-approximations seem to be in accordance, with a maximal relative error of 11% at $b = 10\ \mathrm{ms\,m^{-2}}$. However, for $t_{ex} < 40\mathrm{ms}$, an importance bias can be noted. A possible explanation for this bias could be the too short diffusion time to capture this exchange. Another explanation could be that exchange between somas and the ECS can no longer be considered negligible for fast exchange. Indeed, in that case, the soma volume-to-area ratio equals 3.28, while the neurites volume-to-area ratio equals 0.35 (approximately ten times smaller). The volume-to-area ratio and the exchange time can be related using the following equation (Fieremans et al., 2010):

$$t_{ex} = f_{\mathrm{ECS}} \frac{V}{A} \frac{1}{\kappa}, \tag{4.31}$$

with κ the permeability of the boundary (the cell membrane) between the intracellular space and the ECS. If we consider an equal permeability for the

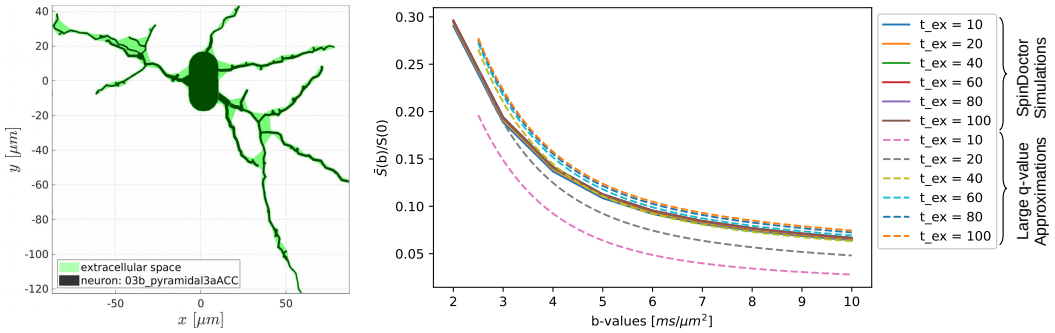


Fig. 4.6: Comparison between SpinDoctor simulations and large q-values approximations from the three-compartment model with exchange between neurites and ECS (Section 4.3), with $\Delta = 21.8$ ms and $\delta = 12.9$ ms. Simulations and large q-values approximations seem to agree for $t_{ex} \geq 40$ ms. A bias is introduced in the large q-value approximation for fast exchange times, likely due to invalid assumptions for fast exchange.

neurites and soma membranes, then the soma exchange time becomes ten times superior to the neurites exchange time (for a fixed ECS fraction). For a high neurites exchange time, the soma exchange time can be regarded as infinite, that is, the soma can be considered as impermeable. However, for fast exchange between neurites and the ECS, the exchange time between somas and ECS could become non-negligible anymore. The inclusion of exchange between somas and the ECS in the case of fast exchange between neurites and ECS (small t_{ex}) could therefore be a way to improve the model.

Other improvements could also be included in the model to achieve a more accurate approximation. For instance, the ECS is modeled as isotropic Gaussian diffusion in the current model. However, the geometry of ECS is very complex and tortuous, and diffusion signals have been proven to deviate from a mono-exponential behavior (Vincent et al., 2021). Another improvement could be to deviate from modeling the ECS diffusion with a simple isotropic Gaussian and consider more complex geometric representations instead. Exchange between somas and neurites has also not been considered.

Although we aim to define a model as close to gray matter tissue as possible, it is important to keep in mind that the estimation of additional tissue parameters also makes the inverse problem harder to solve with a PGSE experiment.

4.5 Conclusion

In this chapter, we presented a three-compartment model describing brain gray matter based on the biological composition of the tissue. The three compartments correspond to somas, neurites, and extra-cellular space. We modeled each compartment by simple geometrical shapes, for which the diffusion signal is known. Sections 4.2 and 4.3 present this three-compartment model following two different assumptions. Section 4.2 supposes that each compartment is impermeable (i.e., no exchange between the compartments), while Section 4.3 considers exchange between neurites and ECS non-negligible. A signal processing algorithm for each model has been developed to synthesize the diffusion signal key information and relate it to a set of parameters describing the tissue (notably the size and density of neurons). These summary statistics are obtained from a large and small q-value approximation of the dMRI signal. Unlike existing methods, no biological parameters are arbitrarily fixed, which allows for the best possible description of the cortical tissue of each subject.

The necessity to define a more complex model is still an open question in the dMRI community. As an example, Jelescu et al. (2021) argue that unmyelinated neurites do not need to be considered in the definition of a model because their contribution is dominated by exchange, while Olesen et al. (2021) suggest a model dubbed eSANDIX that explicitly incorporates them.

In Chapter 5, we propose an approach based on likelihood-free inference to solve the inverse problem of estimating gray matter microstructure parameters using the extracted summary statistics. Results are obtained using the impermeable three-compartment model. Applications of the three-compartment model that accounts for exchange between neurites and the ECS onto real datasets are still in progress. Preliminary results based on experiments will be presented in Section 5.6.2.

5 Gray Matter Microstructure Estimation using an Impermeable Three-Compartment Model via Likelihood-Free Inference

Contents

Acronyms used in the chapter

ABC	Approximate Bayesian Computation
ADC	Apparent Diffusion Coefficient
CSF	Cerebropinal Fluid
CTI	Correlation Tensor MRI
DKI	Diffusion Kurtosis Imaging
dMRI	diffusion MRI
DTI	Diffusion Tensor Imaging
DWI	Diffusion Weighted Image
EAP	Ensemble Average Propagator
ECS	Extra-Cellular Space
ELBO	Evidence Lower BOund
FA	Fractional Anisotropy
GPD	Gaussian Phase Distribution
HARDI	High Angular Resolution Diffusion Imaging
KL	Kullback-Leibler
KM	Kärger Model
LEMONADE	Linearly Estimated Moments provide Orientations of Neurites And their Diffusivities Exactly
LFI	Likelihood-Free Inference
MAP-MRI	Mean Apparent Propagator-MRI
MCMC	Markov Chain Monte Carlo
MD	Mean Diffusivity
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
NF	Normalizing Flows
NODDI	Neurite Orientation Dispersion and Density Imaging
ODF	Orientation Distribution Function
p.d.f.	probability density functions
PGSE	Pulsed Gradient Spin Echo
RTOP	Return-To-the-Origin Probability
SANDI	Soma And Neurite Density Imaging
SM	Standard Model
SNR	Signal Noise Ratio
VI	Variational Inference
WMTI	White Matter Tract Integrity

Notations used in the chapter

δ	Duration of a gradient pulse [ms]
Δ	Duration which separates two gradient pulses [ms]
t	Diffusion time [ms]
b	b-value [$\mathrm{ms}\,\mu\mathrm{m}^{-2}$]
q	q-value [$\mu\mathrm{m}^{-1}$]
S	Diffusion-weighted MRI signal
D_X	Diffusivity within compartment X [$\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$]
f_X	Signal fraction of compartment X
r_s	Soma radius [$\mu\mathrm{m}$]
C_s	Parameter encoding soma radius and diffusivity
τ_X	Mean molecular lifetime in compartment X [ms]
r_X	Exchange rate from compartment X to another [$1/\mathrm{ms}$]
$\mathbf{x}$	Observation vector
$\boldsymbol{\theta}$	Parameter vector
$p(\boldsymbol{\theta})$	Prior distribution
$p(\boldsymbol{x_0}\|\boldsymbol{\theta})$	Likelihood distribution of an observation $\boldsymbol{x_0}$ given parameters $\boldsymbol{\theta}$
$p(\boldsymbol{\theta}\|\boldsymbol{x_0})$	Posterior distribution
$p(\boldsymbol{x_0})$	Evidence of the data
$\mathbb{E}_s[\boldsymbol{x}]$	Expectation of any random variable $\boldsymbol{x}$ w.r.t. its subscript s when needed

5.1 Introduction

Obtaining quantitative measurements of brain gray matter microstructure with a dedicated soma representation is a growing field of interest in the dMRI community (see e.g. Palombo et al., 2020; Jelescu et al., 2021; Olesen et al., 2021). The large dimensionality of the collected data makes this task very difficult and leads to indeterminacies, that is, the existence of multiple mathematical solutions (Novikov et al., 2018a). It can then be complicated to disentangle the correct parameter combination that generated the acquired diffusion signal (see for example the proposed solutions in Fieremans et al., 2011 and Novikov et al., 2018a). To tackle this problem, some methods suggest to stabilize the solution by imposing constraints on model parameters, such as the model NODDI (Zhang et al., 2012, see Section 2.4.2.1). However, these constraints were shown not to be biologically plausible, and the inverse problem remains largely degenerate (Jelescu et al., 2016; Novikov et al., 2018c). Palombo et al. (2020) address this issue by employing random forest regressors to solve the inverse problem. Although an acceptable accuracy is obtained in real-case scenarios, this technique can only deliver one possible set of tissue parameters, masking, therefore, other biologically plausible solutions that could generate the same observed signal. Furthermore, this machine learning approach is deterministic, so no description in terms of confidence into the output parameter estimates is returned. Some methods propose to reduce the dimensionality of the data to be processed, such as the LEMONADE system of equations, which is based on a Taylor expansion of the diffusion signal leading to a set of rotationally invariant moments (Novikov et al., 2018c). Some contrasting methods take a totally different approach by relying on very demanding acquisitions with many q-shells (equivalently b-shells), including very large q-values, to estimate microstructure parameters using deterministic least-squares based solutions (Olesen et al., 2021). However, these requirements seem hardly feasible for human acquisitions.

To overcome such limitations, we rely on the use of summary statistics combined with modern tools from Bayesian analysis known as Likelihood-Free Inference (LFI) (see Chapter 3). These rotationally-invariant features were introduced in Chapter 4, and derived for two gray matter models: a three-compartment model with impermeable membranes, and a three-compartment model with non-negligible exchange between neurites and the ECS. Equations relating the summary statistics to the tissue microstructure are obtained from a signal analysis based on large and small q-values approximations. These summary statistics allow to reduce the dimensionality of

the data at hand and to represent all the relevant information for statistical inference with a few features. In this chapter, we focus on solving the inverse problem in the case of the impermeable model. Models including exchange between neurites and ECS have only been introduced very recently in the dMRI community. Promising preliminary results are presented in Section 5.6.2, but more experiments are in progress and will be submitted soon.

Given the equations relating the summary statistics and the microstructure parameters (Section 4.2.2.3), we use a probabilistic framework to solve the non-linear inverse problem efficiently using Bayesian inference, as described in Chapter 3, and determine the posterior distribution of the fitted parameters. Such approach offers a full description of the solution landscape and can point out degeneracies, as opposed to the usual deterministic least-squares based solution (Jelescu et al., 2016; Novikov et al., 2018c). The general framework of our approach is presented in Figure 5.1.

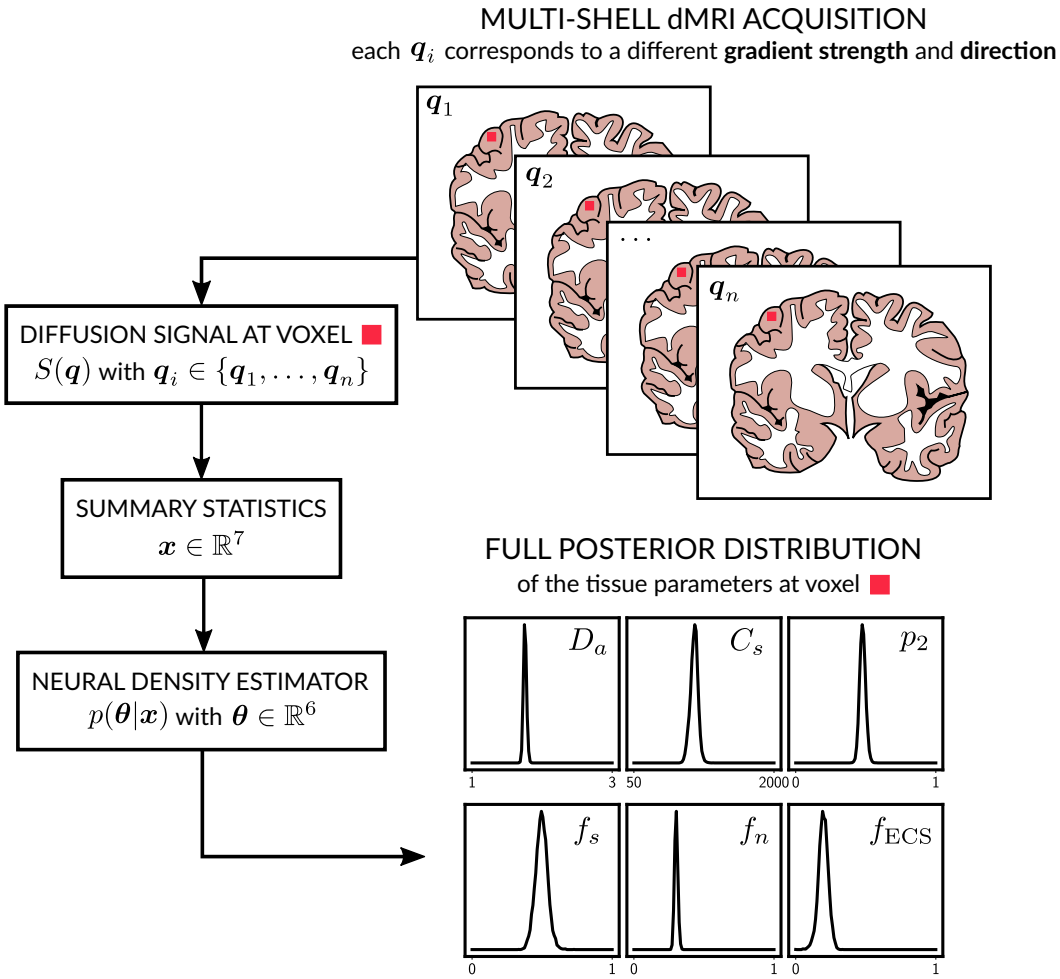


Fig. 5.1: Visual abstract. On the top right we illustrate a multi-shell dMRI acquisition. Based on the proposed 3-compartment model, we then extract summary statistics. Applying a neural density estimator allows to estimate the tissue parameters and their full posterior distribution.

The usual way of applying a Bayesian approach to solve non-linear inverse problems (Stuart, 2010) is to define two quantities: a prior distribution encoding initial knowledge of the parameter values (e.g. intervals which are physiologically relevant) and the likelihood function of the forward model being studied. One can then either obtain an analytic expression of the posterior distribution via Bayes' formula or use a Markov-Chain Monte Carlo (MCMC) procedure to numerically sample the posterior distribution (Gelman et al., 2013). However, the likelihood function of complex models such as the one that we consider here is often very hard to obtain and makes the Bayesian approach rather challenging to use. Likelihood-free inference (LFI) bypasses this bottleneck by recurring to several simulations of the forward model using different parameters and learning an approximation to the posterior distribution from these examples (Cranmer et al., 2020).

The first contributions on LFI are known as approximate Bayesian computation (ABC) (see Section 3.3.2) and have been applied to invert models from ecology, population genetics, and epidemiology (Sisson, 2018). Some of the limitations of these techniques include the large number of simulations required for the posterior estimations and the need of defining a distance function to compare the results of two simulations. Recently, there has been a growing interest in the machine learning community in improving the limitations of ABC methods through deep generative modeling, i.e. neural network architectures specially tailored to approximate probability density functions from a set of examples (Goodfellow et al., 2016). NF (see Section 3.3.3) are a particular class of such neural networks that have demonstrated promising results for likelihood-free inference in different research fields (Cranmer et al., 2020; Gonçalves et al., 2020; Greenberg et al., 2019).

5.1.1 Contributions

The content of this chapter is based on (and extends) the works of three papers:

> **M. Jallais**, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Diffusion MRI-Based Cytoarchitecture Measurements in Brain Gray Matter using Likelihood-Free Inference. In ISMRM 2021.

> **M. Jallais**, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Cytoarchitecture Measurements in Brain Gray Matter using Likelihood-Free Inference. In International Conference on Information Processing in Medical Imaging 2021 (pp. 191-202). Springer, Cham.

M. Jallais, P. L. C. Rodrigues, A. Gramfort, and D. Wassermann. Inverting brain gray matter models with likelihood-free inference: a tool for trustable cytoarchitecture measurements. Submitted to MELBA (arXiv preprint arXiv:2111.08693).

Our contributions have been to employ LFI methods to solve our non-linear inverse problem under a probabilistic framework and determine the posterior distribution of the fitted parameters. Such approach offers a full description of the solution landscape and can point out degeneracies, as opposed to the usual deterministic least-squares based solution (Jelescu et al., 2016; Novikov et al., 2018c). This work was accepted as oral presentations at ISMRM 2021 and IPMI 2021. An extension of this work has been submitted to the journal MELBA. Similar application of LFI methods to the three-compartment model with exchange between neurites and ECS presented in Section 4.3 is currently being studied. Preliminary results are presented in Section 5.6.2. More experiments are in progress and will be submitted soon.

To foster reproducible research, `Python` code of our implementation to solve the inverse problem for both models is available online on the public repository:

`https://github.com/mjallais/SBI_dMRI`

5.2 Inverting the model with Bayesian inference

We focus on the three-compartment model presented in Section 4.2. Our main goal is to determine the values of the parameter vector

$$\boldsymbol{\theta} = (D_n, C_s, p_2, f_s, f_n, f_{\text{ECS}}) \in \mathbb{R}^6$$

that best explain a given observed dMRI signal. Because of the high dimensionality of this kind of signal and the difficulties in obtaining stable estimates of $\boldsymbol{\theta}$ directly from it, we recur to the set of summary features defined in Section 4.2.2. We consider D_e nearly-constant per subject acquisition following Menon et al. (2020), and estimate it as one-third of the mean diffusivity in the subject's ventricles (Vincent et al., 2021). This assumption allows to drop an unknown from the system, use D_e as a reference diffusivity and turn our system of equations unitless. The obtained equations relating those summary statistics to the microstructure are the following:

$$\begin{cases} \frac{M^{(2),0}}{D_e} = f_n D_n^u + 3 f_s C_s^u + 3 f_{\text{ECS}} \\ \frac{M^{(2),2}}{D_e} = f_n D_n^u \cdot p_2 \\ \frac{M^{(4),0}}{D_e^2} = f_n D_n^{u2} + 5 f_s C_s^{u2} + 5 f_{\text{ECS}} \\ \frac{M^{(4),2}}{D_e^2} = f_n D_n^{u2} \cdot p_2 \\ a_{\text{fit}} \left(tD_e\right)^{3/2} = \frac{f_s}{8(\pi C_s^u)^{3/2}} + \frac{f_{\text{ECS}}}{8\pi^{3/2}} \\ b_{\text{fit}} \left(tD_e\right)^{1/2} = \frac{f_n}{2}\sqrt{\frac{\pi}{D_n^u}} \\ f_n + f_s + f_{\text{ECS}} = 1, \end{cases} \tag{5.1}$$

with $D_n^u = \frac{D_n}{D_e}$ and $C_s^u = \frac{C_s}{(2\pi)^2 t D_e}$.

This set of summary statistics will be further denoted $\boldsymbol{x}$:

$$\boldsymbol{x} = \left(\frac{M^{(2),0}}{D_e}, \frac{M^{(2),2}}{D_e}, \frac{M^{(4),0}}{D_e^2}, \frac{M^{(4),2}}{D_e^2}, a_{\text{fit}}(tD_e)^{3/2}, b_{\text{fit}}\sqrt{tD_e} \right) \in \mathbb{R}^7 \ ,$$

We make the assumption that it carries all the information necessary for determining the $\boldsymbol{\theta}_0$ having generated a given dMRI signal $\boldsymbol{S}_0$. We denote the relation between these quantities as

$$\boldsymbol{x} = \mathcal{M}(\boldsymbol{\theta}) + \boldsymbol{n} \ , \tag{5.2}$$

where $\mathcal{M} : \mathbb{R}^6 \to \mathbb{R}^7$ is a multivariate function that implements the system of equations defined in Equation 5.1 and $\boldsymbol{n}$ is an additive noise capturing

the imperfections of our modeling procedure, the limitations of the summary statistics, and the measurement noise. Note that noise statistics of different elements of $\boldsymbol{x}$ are different, due for example to Rician floor.

5.2.1 The Bayesian formalism.

We interpret the inverse problem of inferring the parameters that best describe a given observed summary feature vector $\boldsymbol{x}_0$ as that of determining the posterior distribution of $\boldsymbol{\theta}$ given an observation $\boldsymbol{x}_0$ (see Chapter 3). By first choosing a prior distribution $p(\boldsymbol{\theta})$ describing our initial knowledge of the parameter values, we may use Bayes' theorem to write

$$p(\boldsymbol{\theta}|\boldsymbol{x}_0) = \frac{p(\boldsymbol{x}_0|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{x}_0)} \ , \tag{5.3}$$

where $p(\boldsymbol{x}_0|\boldsymbol{\theta})$ is the likelihood of the observed data point and $p(\boldsymbol{x}_0)$ is a normalizing constant, commonly referred to as the evidence of the data. Note that such a probabilistic approach returns not only which $\boldsymbol{\theta}$ best fits the observed data (i.e. the parameter that maximizes the posterior distribution), but the full posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x}_0)$. The latter can be possibly multi-modal or flat, which would indicate the difficulty of summarizing it with a unique maximum.

5.2.2 Bypassing the likelihood function.

Despite its apparent simplicity, it is usually difficult to use Equation 5.3 to determine the posterior distribution, since the likelihood function for data points generated by complex non-linear models is often hard to write. To avoid such difficulty, we directly approximate the posterior distribution using a conditional density estimator, i.e. a family of conditional p.d.f. approximators $q_\phi(\boldsymbol{\theta}|\boldsymbol{x})$ parametrized by ϕ and that takes $\boldsymbol{\theta}$ (the parameter) and $\boldsymbol{x}$ (the observation) as input arguments. Our posterior approximation is then obtained by minimizing its average KL divergence with respect to the conditional density estimator for different choices of $\boldsymbol{x}$, as per (Papamakarios and Murray, 2016)

$$\min_{\boldsymbol{\phi}} \quad \mathcal{L}(\phi) \quad \text{with} \quad \mathcal{L}(\phi) = \mathbb{E}_{\boldsymbol{x}\sim p(\boldsymbol{x})}\left[D_{\mathrm{KL}}(p(\boldsymbol{\theta}|\boldsymbol{x})\|q_\phi(\boldsymbol{\theta}|\boldsymbol{x}))\right] \ , \tag{5.4}$$

which can be rewritten as (see Section 3.3.3.3):

$$\mathcal{L}(\phi) \quad = \quad -\mathbb{E}_{(\boldsymbol{x},\boldsymbol{\theta})\sim p(\boldsymbol{x},\boldsymbol{\theta})}\left[\log\left(q_\phi(\boldsymbol{\theta}|\boldsymbol{x})\right)\right] + C \ , \tag{5.5}$$

where C is a constant that does not depend on ϕ. Note, however, that in practice we actually consider a N-sample Monte-Carlo approximation of the loss function given by

$$\mathcal{L}(\phi) \approx \mathcal{L}^N(\phi) = -\frac{1}{N}\sum_{i=1}^{N} \log\left(q_\phi(\boldsymbol{\theta}_i|\boldsymbol{x}_i)\right) , \tag{5.6}$$

where the N data points $(\boldsymbol{\theta}_i, \boldsymbol{x}_i)$ are sampled from the joint distribution with $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$ and $\boldsymbol{x}_i \sim p(\boldsymbol{x}|\boldsymbol{\theta}_i)$. We can then use stochastic gradient descent to obtain a set of parameters ϕ which minimizes $\mathcal{L}^N$.

If the class of conditional density estimators is sufficiently expressive, it is possible to show that the minimizer of Equation 5.6 converges to $p(\boldsymbol{\theta}|\boldsymbol{y})$ when $N \to \infty$ (Greenberg et al., 2019). Note, also, that the parametrization ϕ that we obtain by the end of the optimization procedure yields a posterior which is amortized for different choices of $\boldsymbol{x}$. Thus, for a specific observation $\boldsymbol{x}_0$ we may simply write $q_\phi(\boldsymbol{\theta}|\boldsymbol{x}_0)$ to get an approximation of $p(\boldsymbol{\theta}|\boldsymbol{x}_0)$.

5.2.3 Neural density estimators.

In this work, our conditional p.d.f. approximators belong to a class of neural networks called Normalizing Flows (NF) (Papamakarios et al., 2019). These flows are invertible functions capable of transforming vectors generated by a simple base distribution (e.g. the standard multivariate Gaussian distribution) into an approximation of the true posterior distribution (see Section 3.3.3). An important advantage of normalizing flow versus other p.d.f. approximators such as generative adversarial network (GAN, Goodfellow et al., 2014) and variational auto-encoders (VAE, Kingma and Welling, 2014) is that it provides both the likelihood of any sample point of interest and it is also straightforward to sample new data points from it. Furthermore, certain classes of NF can be shown to be universal approximators of probability density functions. We refer the reader to Papamakarios et al. (2019) for more information on the different types of normalizing flows.

5.3 Materials and methods

This section presents the technical details on how we have implemented our theoretical contributions and describes the simulated and real datasets used in the numerical illustrations.

5.3.1 The likelihood-free inference setup

Using a likelihood-free inference approach for inverting the 3-compartment model relating tissue parameters $\boldsymbol{\theta}$ and dMRI summary statistics $\boldsymbol{x}$ relies on four important aspects:

(1) **The forward model.** As explained in Section 5.2, we obtain an approximation of the amortized posterior distribution using a dataset containing several paired examples of a parameter $\boldsymbol{\theta}_i$ and its corresponding summary statistics $\boldsymbol{x}_i$, related by Equations 5.1. In what follows, we adopt the usual assumptions from the inverse problems literature and consider the additive noise $\boldsymbol{n}$ from Equation 5.2 small enough to be ignored in the data generation, so that we have $\boldsymbol{x}_i \approx \mathcal{M}(\boldsymbol{\theta}_i)$.

(2) **Prior distribution.** The simplest way of defining a prior distribution $p(\boldsymbol{\theta})$ in our setting is to use an uniform distribution with limits within physiologically relevant intervals for each parameter. From Section 4.2, we know that the fractions f_s, f_n, and f_{ECS} have values between 0 and 1 and all sum up to one. To encode this information in $p(\boldsymbol{\theta})$, we define new parameters k_1 and k_2 and relate them with the fractions by

$$f_n = k_2\sqrt{k_1}, \quad f_s = \sqrt{k_1}(1-k_2), \quad f_{\text{ECS}} = 1-\sqrt{k_1}\,. \tag{5.7}$$

In this way, whenever we want to generate a prior sample for $(f_n, f_s, f_{\text{ECS}})$ we first generate a sample of $k_1, k_2 \sim \mathcal{U}([0,1])$ and then transform them according to Equation 5.7 to get a set of fractions which is uniformly sampled in the region $\{f_n, f_s, f_{\text{ECS}} \in [0,1] \ : f_n + f_s + f_{\text{ECS}} = 1\}$. We follow the usual assumption that the diffusivity of the compartments are inferior or equal to the self-diffusion coefficient of free water, which is $3\,\mu\text{m}^2\,\text{ms}^{-1}$ (Li et al., 2016). We fix, therefore, the interval for neurite diffusivity (D_n) as between 10^{-5} and $3\,\mu\text{m}^2\,\text{ms}^{-1}$. The newly introduced parameter C_s is parametrized by soma radius and diffusivity. To account for a soma radius comprised between 2 and 30 μm (Palombo et al., 2021) and a diffusivity range as previously defined, we used the C_s interval [50, 2500] μm^2. Parameter p_2, which measures the dispersion of neurites orientation, is comprised in the interval [0,1], with 1 indicating an anisotropic orientation distribution function.

(3) **Posterior approximator.** We use an autoregressive architecture for normalizing flows implemented via the masked autoencoder for distribution estimation (MADE) (Germain et al., 2015). We follow the same setup from Greenberg et al. (2019) for LFI problems, stacking five MADEs, each with two hidden layers of 50 units, and a standard normal base distribution. This choice provides a sufficiently flexible function capable of approximating complex posterior distributions.

(4) **Training procedure.** The parametrization of our normalizing flow is obtained by minimizing the loss function (Equation 5.6) using the ADAM optimizer (Kingma and Ba, 2017) with default parameters, a learning rate of 5.10^{-4} and a batch size of 100. Except for a few validation experiments, we have used $N = 10^5$ simulated data points to approximate the posterior distribution.

5.3.2 Simulated dMRI data

We first validate our proposed method using simulated dMRI data. For this, we fix a parameter vector $\boldsymbol{\theta}_0$ based on a plausible biophysical configuration (Palombo et al., 2021) and generate a simulated observation $\boldsymbol{x}_0$ associated to it. Our goal, then, is to check whether the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{x}_0)$ concentrates around the ground truth parameter, i.e. if it is peaked around the true values of the parameters in $\boldsymbol{\theta}_0$. If this is the case, we can assert that the LFI procedure is capable of inverting our non-linear model successfully.

The simplest way of generating an observation from the ground truth parameter would be to use the forward model defined in Section 5.3.1, which yields very good results, since the posterior approximation is trained on data points generated in the same way. We have also considered a more challenging situation, in which the dMRI signals are simulated following a setup that is closer to what we would expect from real experimental experiments. This is based on two steps. Firstly, we use the `dmipy` simulator (Fick et al., 2019) to simulate the three-compartment model described in Section 4.2 and obtain a dMRI signal $\boldsymbol{S}_0$. Then, we calculate the summary statistics of this signal as defined in Section 4.2.2 to reduce the dimensionality of the observation and obtain a feature vector $\boldsymbol{x}_0$.

We have carried out our simulations on `dmipy` considering three different kinds of acquisition setup. They all have b-shells with 128 uniformly distributed directions, but they differ in their b-values and acquisition times:

– Setup `Ideal` corresponds to a rather "confortable" case with 10 b-values between 0 and 10 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$. We use $\delta/\Delta = 12.9/21.8$ ms as in the HCP MGH database.

– Setup `HCP MGH` reproduces the setup from the HCP MGH dataset, with 5 b-values: 0, 1, 3, 5, and $10\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$ and $\delta/\Delta = 12.9/21.8$ ms. Since the Spiked LEMONADE approximation (4.2.2.2) requires at least three b-values inferior to $2.5\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$, we extrapolated an extra b-shell at $0.1\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$ using `MAPL` (Fick et al., 2016), a method for modeling multi-shell q-space signals.

– Setup `HCP 1200` reproduces the setup from the HCP 1200 dataset, with only 4 b-values: 0, 1, 2, and 3 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$ and $\delta/\Delta = 10.6/43.1$ ms An extra b-shell at $2.8\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$ has been interpolated to be used in the RTOP approximation.

Note that in the simulations with all setups we have used the three b-shells with the lowest b-values for the small q-value approximation (Spiked LEMONADE), and the three largest q-values for the RTOP approximation.

5.3.3 HCP MGH dataset

After validating our proposal on different simulated settings, we carried out our analysis on two publicly available databases. Our goal was to estimate the tissue parameters for each voxel in the dMRI acquisitions corresponding to the gray matter and determine how these parameters vary. We segmented the brain gray matter using `FreeSurfer` before applying our pipeline to the selected voxels. Because of the probabilistic framework that we use, these estimates are accompanied of credible intervals that can be used to inform our degree of confidence of these estimates.

Our first analysis was on the HCP MGH Adult Diffusion database (Setsompop et al., 2013). This database is composed of 35 subjects with $\delta/\Delta = 12.9/21.8\,\mathrm{ms}$ and b = 1, 3, 5, 10 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$. We used the 3, 5 and 10 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$ b-values for the RTOP approximation (i.e. the large q-value analysis), and 0, 0.1 and 1 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$ for the Spiked LEMONADE approximation. We used `MAPL` with the 0, 1 and 3 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$ b-values to reduce noise and interpolate a b-value of $0.1\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$ to improve the estimations. D_e was estimated as $1/3$ of the mean diffusivity in the ventricles.

The spatial distribution of the estimated parameters were mapped to the MNI template, averaged over all subjects, and then projected onto an inflated cortical surface using `FreeSurfer`. We have then evaluated whether such distributions seemed physiologically plausible by using the Brodmann

atlas (Brodmann, 1909; Zilles, 2018), which is a parcellation of the brain based on cytoarchitecture features. In addition, we compared qualitatively the results of soma estimations with Nissl-stained histological images of cytoarchitecture (Allman et al., 2010; Amunts et al., 1999; Geyer et al., 1999).

5.3.4 HCP 1200

We proceeded with our analysis on real data using a more challenging database, in which the dMRI signals were acquired with only a few small b-values. Our goal was to demonstrate that the credible regions obtained via the posterior approximation can be used to inform which parameters remain possible to estimate even in very challenging situations. Note that this unlocks the door to the analysis of any dMRI database, since the estimates always come with a "quality certificate".

We applied our pipeline to a subset of the HCP 1200 database. We randomly picked 30 subjects, to have an identical number of subjects to that in our analysis of the HCP MGH database. The data were acquired for b-values equal to 1, 2 and 3 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$, with $\delta/\Delta = 10.6/43.1\,\mathrm{ms}$. Using `MAPL`, we interpolated a b-shell at $2.8\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$ to improve the computation of the summary statistics. We used all the three lowest b-values for the Spiked LEMONADE approximation, and $b = 2$, 2.8 and 3 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$ for the high b-value approximation based on RTOP. Similarly to the HCP MGH dataset, we have averaged the parameter estimations in a common space (MNI) and then projected the resulting parametric maps onto an inflated cortical surface.

5.3.5 Software

All our experiments have been implemented with Python (Python Software Fundation, 2017) using several scientific packages: dMRI signals were simulated with the package `dmipy` (Fick et al., 2019) and processed using `dipy` (Garyfallidis et al., 2014) or custom implementations based on `numpy` (Harris et al., 2020). We used the `sbi` (Tejero-Cantero et al., 2020) and `nflows` (Durkan et al., 2020) packages for carrying out the LFI procedures and combined them with data structures and functions from `pyTorch` (Paszke et al., 2019). The figures of results on real experimental data were generated with `mayavi` (Ramachandran and Varoquaux, 2011). The code used for solving the inverse problem is available online on the following public repository:

`https://github.com/mjallais/SBI-dMRI`

5.4 Results

In this section, we describe our results obtained on simulated data and two real datasets.

5.4.1 Simulated data

5.4.1.1 Validating the LFI pipeline.

In this first round of experiments, our aim was to check whether the LFI pipeline worked correctly on a setting where we knew the true values of the parameter $\boldsymbol{\theta}_0$ (ground truth) generating the observed data $\boldsymbol{x}_0$. Furthermore, we wanted to confirm whether the use of summary statistics for the dMRI signal actually conveyed any improvements in the parameter estimation. We have considered, therefore, three different cases:

- **Case 1.** Generate $\boldsymbol{x}_0$ directly from $\boldsymbol{\theta}_0$ using the forward model defined in Section 5.3.1. This is a rather favorable case for our posterior approximation, since it is applied on a data point generated in the same way as the dataset in which it was trained.

- **Case 2.** Generate a dMRI signal $\boldsymbol{S}_0$ from $\boldsymbol{\theta_0}$ using `dmipy` and then obtain $\boldsymbol{x}_0$ by calculating the summary statistics presented in Section 4.2.2. We use the same posterior approximator from **Case 1**, meaning that the data point in inference time is generated differently from those for the training procedure. This is the actual realistic case that interests us the most.

- **Case 3.** Generate a dMRI signal $\boldsymbol{S}_0$ from $\boldsymbol{\theta_0}$ using `dmipy` and do not use any summary statistics for the model inversion, i.e. consider $\boldsymbol{x}_0 = \boldsymbol{S}_0$. Note that the posterior approximator has to be trained on a dataset with observations consisting of dMRI signals, so it is different from the approximators in **Case 1** and **Case 2**. Depending on how the LFI pipeline behaves on this case, the use of summary statistics can be justified or not.

All simulations were carried out with setup `Ideal`, which corresponds to an ideal dMRI acquisition scheme, and the posterior approximators were trained with $N = 10^5$ simulated data points. While we have validated the LFI pipeline on multiple choices of physiologically relevant ground truth parameters $\boldsymbol{\theta}_0$, Figure 5.2 portrays the results only for

$$\boldsymbol{\theta}_0 = (D_n, C_s, p_2, f_s, f_n, f_{\text{ECS}}) = (2.5\,\mu\text{m}^2\,\text{ms}^{-1}, 617\,\mu\text{m}^2, 0.50, 0.15, 0.45, 0.40)\,.$$

This choice of parameters represent a sample tissue containing pyramidal neurons of radius $12\,\mu\mathrm{m}$ (Palombo et al., 2021) and diffusivity $3\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$. The three compartment proportions were chosen in accordance to reported values from histology of human gray matter tissue (Shapson-Coe et al., 2021). Results are presented in Figure 5.2. We see that in **Case 1** the marginals of the posterior distribution are well concentrated around the values of the ground truth parameter $\boldsymbol{\theta}_0$. This confirms that the posterior approximator successfully inverts the non-linear model using the examples in the training set. We also see that the parameter estimation in **Case 2** captures very well most of the true values of the parameters, indicating both that our posterior approximator is robust to observed data generated differently from its training set and that our summary statistics are descriptive enough to synthesize each tissue configuration. Finally, the poor results in the estimation for **Case 3** reflect the largely indeterminate system of equations that results from trying to infer the tissue parameters directly from the dMRI signals. This behavior was expected, as similar degeneracies were shown for a simpler model in Novikov et al. (2018a).

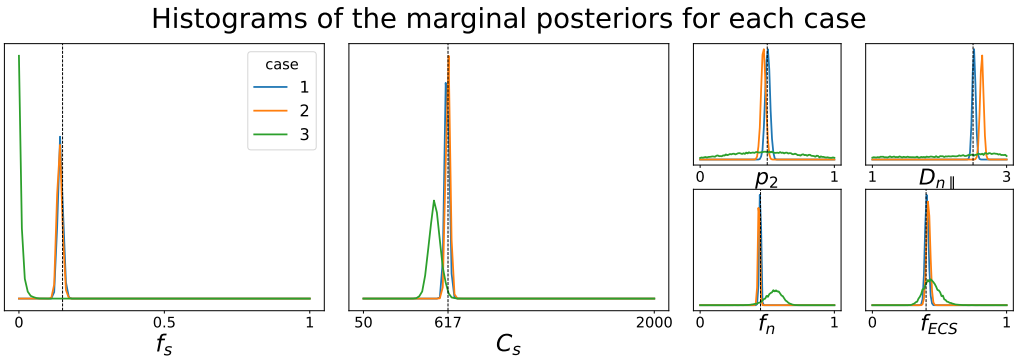


Fig. 5.2: Histograms of 10^4 samples of the approximate posterior distribution in three different base cases (see text for details). Vertical black dashed lines represent ground truth values of $\boldsymbol{\theta}_0$ which generated the observed dMRI signals. We observe that the marginals tend to concentrate around the ground truth parameters when the observed summary statistics are obtained directly from the parameters (**Case 1**) and have a small bias when the signals are generated using `dmipy` (**Case 2**). The figure also shows that inverting the model directly from the dMRI signals leads to rather poor results (**Case 3**).

Focusing on **Case 1**, we want to make sure the system of equations 5.1 can estimate any realistic microstructure parameters. We simulated summary statistics using many combinations of parameters and estimated them using the LFI pipeline presented before. Figure 5.3 presents the results in the form of a Bland Altman plot, which shows the mean difference between the estimations and the ground truth values for all combinations. A linear bias is observed for k_1 and k_2, indicating a slight overestimation of compartment fractions for $k_1 < 0.4$ and $k_2 < 0.45$, and an underestimation otherwise. D_a

presents a very low bias (inferior to 3%) for every considered value. The estimation of C_s is biased for small soma radii with a relative error of up to 11% and provides otherwise reasonable estimates with relative errors inferior to 2.5%. Overall, this figure indicates that the proposed method relying on LFI allows solving the inverse problem of retrieving the microstructure parameters of a tissue using the set of summary statistics introduced in Section 4.2.2 and provides reasonable estimations in realistic cases.

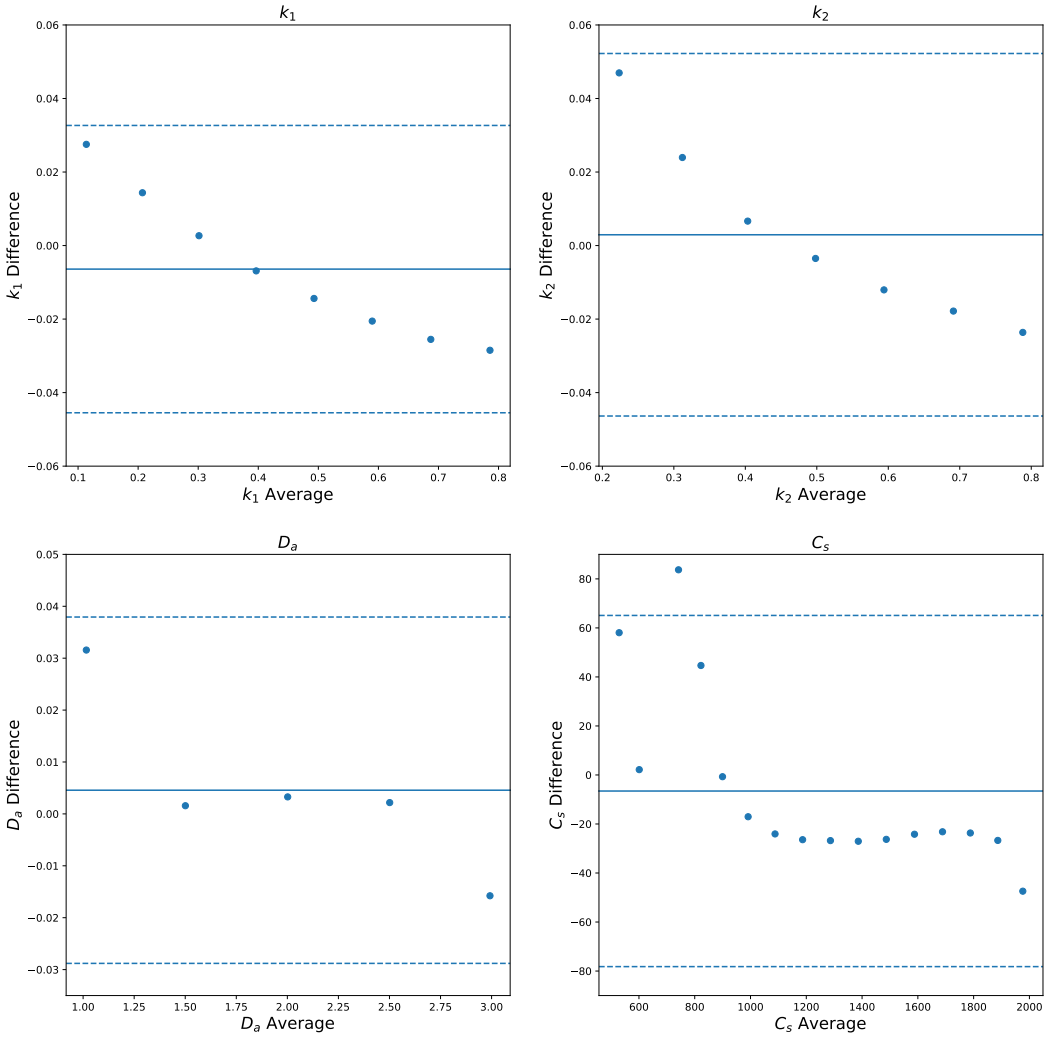


Fig. 5.3: Bland-Altman plot illustrating the difference between estimations obtained from the LFI pipeline in noiseless simulations and their ground truth values. Results show the capacity of the proposed pipeline to solve the inverse problem of estimating tissue microstructure from summary statistics.

5.4.1.2 Influence of the number of b-values.

We have also investigated how different choices of b-values in the acquisition scheme affect the quality of the parameter estimation using the posterior approximation. Note that these choices have no influence on how the posterior approximator is obtained, since it is trained on data generated directly from the relations between tissue parameters and the diffusion summary statis-

tics, in which the b-values do not interact. In fact, the different acquisition setups that we consider have only an impact over the observed data point generated via `dmipy`. Figure 5.4 portrays the marginal posterior distributions for each tissue parameter in setups `Ideal`, `HCP MGH`, and `HCP 1200`. We see that estimations in the `Ideal` setup (equivalent to Case 2 in Figure 5.2) are very much concentrated around the true values of the parameters. For `HCP MGH` and `HCP 1200` the estimations of the soma-related parameters are rather good, but a small bias is present for the other parameters in `HCP MGH` and even more for `HCP 1200`. Note that the `Ideal` and `HCP MGH` ground truth value of C_s (vertical black dashed line) is different from the one of the `HCP 1200` setup (vertical green dashed line), because of their different diffusion times. These simulations allow us to have a fair confidence in the estimations on real data, presented in Sections 5.4.2 and 5.4.3.

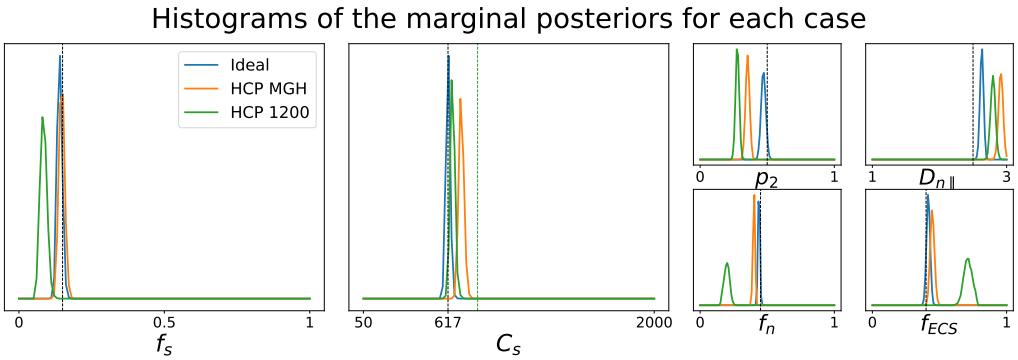


Fig. 5.4: Histograms of 10^4 samples of the approximate posterior distribution with observed dMRI signals generated under three acquisition setups (see text for details). Vertical black dashed lines represent ground truth values of $\boldsymbol{\theta}_0$ which generated the observed signals (the green dashed line indicates the C_s value expected for the `HCP 1200` setup). We see that while the `Ideal` case delivers very good estimates, the results for the two other setups are only reliable for a subset of the tissue parameters.

5.4.1.3 Our new parameter avoids model indeterminacy.

In Chapter 4, we introduced the parameter C_s, which serves as a proxy of the soma radius and provides key information on the soma compartment. In this experiment, we illustrate the results of our model inversion if we had not defined parameter C_s.

Figure 5.5 presents the marginal posterior distributions of r_s and D_s as well as their joint distribution using the setup `Ideal`. To obtain these results, we have altered our LFI pipeline so to consider an extended parameter vector including r_s and D_s. The prior distribution reflects our assumption that $r_s \in [10^{-5}, 30]\ \mu\text{m}$ and $D_s \in [1, 3]\ \mu\text{m}^2\,\text{ms}^{-1}$ and we consider ground truth parameters $r_s = 17.5\ \mu\text{m}$ and $D_s = 2.3\ \mu\text{m}^2\,\text{ms}^{-1}$. We note that in addition to larger marginal posterior distributions for each parameter, the joint posterior

has a valley of large values for the (r_s, D_s) pair, including the ground truth parameters. This result is typical of non-injective models, i.e. models for which several input parameters may yield the same output observation, and is an important asset of a probabilistic framework such as ours.

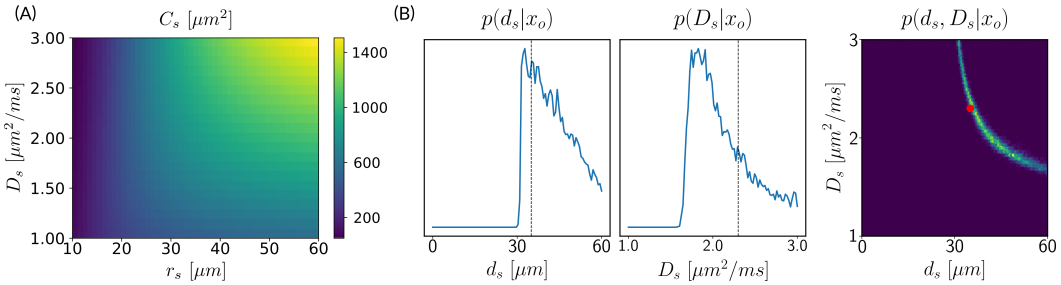


Fig. 5.5: (A) C_s dependence on soma radius r_s and diffusivity D_s. We see that there are several values of (d_s, D_s) that yield the same C_s. (B) and (C) Histograms of 10^4 samples from the marginal and joint posterior distributions of $d_s = 2r_s$ and D_s. The ridge in the joint distribution indicates that there are several possible values for the pair (d_s, D_s) with high probability, which are those yielding the same C_s. Estimating C_s directly bypasses this indeterminacy.

5.4.1.4 Assessing the variances of estimated parameters.

Deriving the posterior distributions of the parameter vectors allows us to report the values of the most likely tissue parameters for a given observation, along with our certitude regarding our inference. Figure 5.6 presents the logarithm of the standard deviation of the marginal posterior samples for different ground truth parameter choices (varying f_s and f_n) under setup `Ideal`. These values indicate how sharp a posterior distribution is and, therefore, quantify the quality of the fit. We observe larger standard deviations in the absence (or weak presence) of soma compartments in the mixture signal, e.g., the standard deviation of C_s is large when few or no somas are present ($f_s \approx 0$). This is to be expected, since the lack of contribution from the somas in the diffusion signal makes it difficult to estimate parameters related to them.

5.4.2 HCP MGH

Although we manage to invert very well the brain tissue parameter on settings for which the dMRI signal is obtained with several b-values, our results on simulated data show that the estimates for more realistic settings are less robust and demand a more subtle analysis. Indeed, we have observed that for both setups `HCP MGH` and `HCP 1200` the soma parameters seem to be rather well estimated without too much bias, which has lead us to consider mainly these parameters in our interpretations of the results.

Figure 5.7 presents the results on the HCP MGH dataset. The inference takes approximately one hour per subject using 20 CPUs, and the estimation of

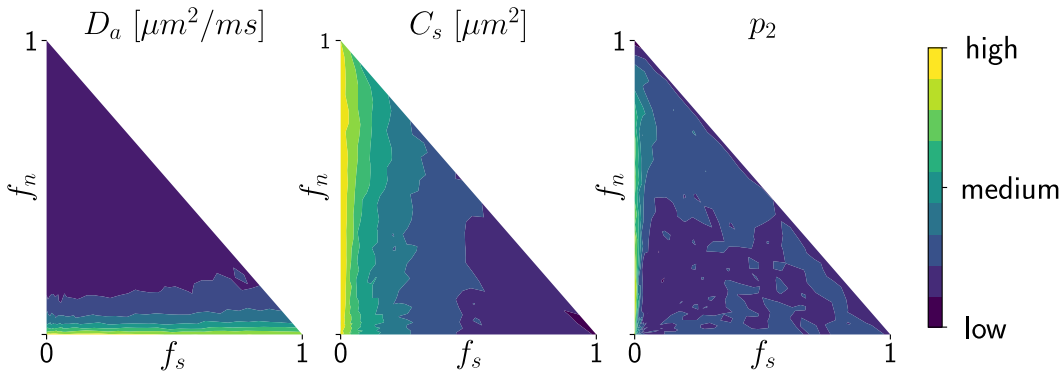


Fig. 5.6: Logarithm of the standard deviations for the marginal posterior distribution of D_n, C_s, and p_2 with different choices of ground truth parameters (varying f_s and f_n). Since the ranges of values for each plot were quite different, we labeled the colorbar in terms of {'high', 'medium', 'low'} values of standard deviations so to provide mainly qualitative information to the reader. We see that when the signal fraction of somas decreases ($f_s \to 0$) the standard deviation of the C_s-estimation increases; and when less neurites are present ($f_n \to 0$) the standard deviation of p_2 and D_n increase.

the seven parameters for every voxel in the gray matter is about 3 hours per subject, when computed in parallel on 20 CPUs. We have masked our results so to show only areas where parameters were deemed stable, i.e. when the values were larger than 2 times the LFI-obtained standard deviations of the fitted posterior, indicating that the posterior distribution is narrow and centered around its mean value. We observe a lack of stability on small sections including the auditory cortex and the precentral gyrus fundus. Our figure assesses qualitatively the results on soma size by comparing with Nissl-stained histological studies (Allman et al., 2010; Amunts et al., 1999; Geyer et al., 1999). Our comparison shows good agreement between different cortical areas and the parameter C_s, which, under nearly-constant intra-soma diffusion D_s, is modulated by soma size. Note that we modeled brain gray matter as homogeneous per voxel. That is, we suppose only one type of neuron is present in each voxel. The most probable one is returned by the LFI approach.

Interestingly, most regions of Figure 5.8 in which the parameter estimation has low confidence are located in the fundus of the sulci. Two main hypothesis could explain such behavior. Firstly, brain regions which are very curved may be more prone to mixing between tissue layers and CSF, which generates noisier signals. Thus, the estimation of summary statistics becomes more biased and the posterior distributions tend to be wider. Another possible explanation, based on cytoarchitecture considerations, points out the fact that the fundus of sulci is where sharp changes in cellular populations happen (Brodmann, 1909; Pandya et al., 2015). A mixing of several types of neurons

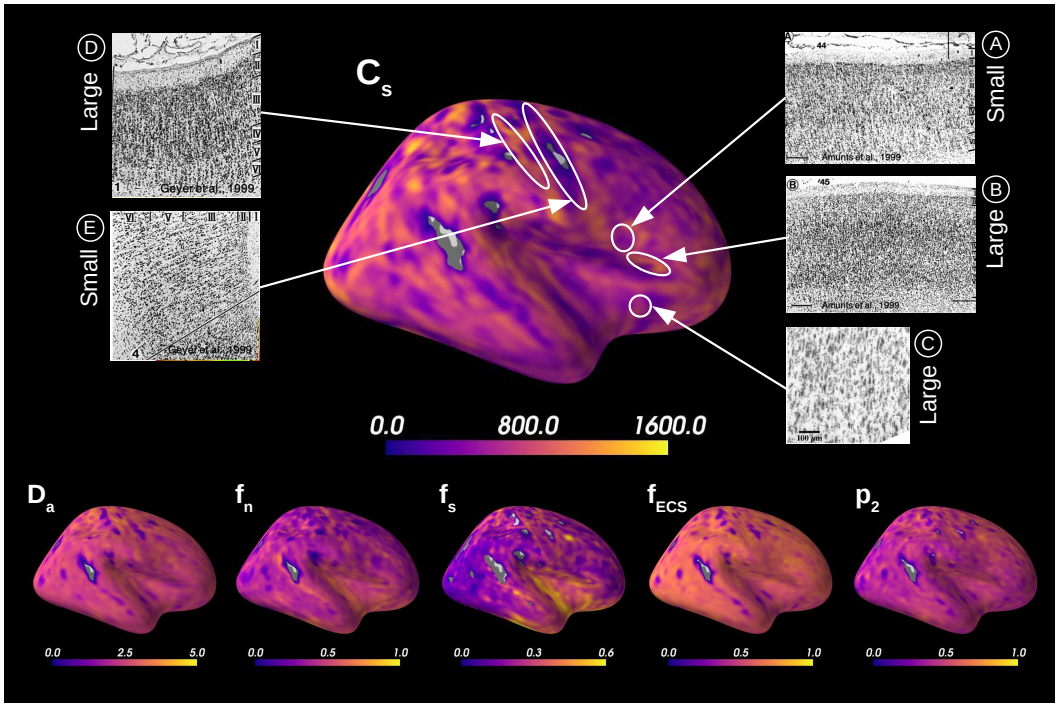


Fig. 5.7: Microstructural measurements averaged over 31 HCP MGH subjects. We deemed stable measurements with a z-score larger than 2, where the standard deviation on the posterior estimates was estimated through our LFI fitting approach. In comparing with Nissl-stained cytoarchitectural studies we can qualitatively evaluate our parameter C_s: Broadmann area 44 (A) has smaller soma size in average than area 45 (B) (Amunts et al., 1999); large von Economo neurons predominate the superior anterior insula (C) (Allman et al., 2010); precentral gyrus (E) shows very small somas while post-central (D) larger ones (Geyer et al., 1999).

in one voxel could generate multi-modal posterior distributions, and hence a region with large variance.

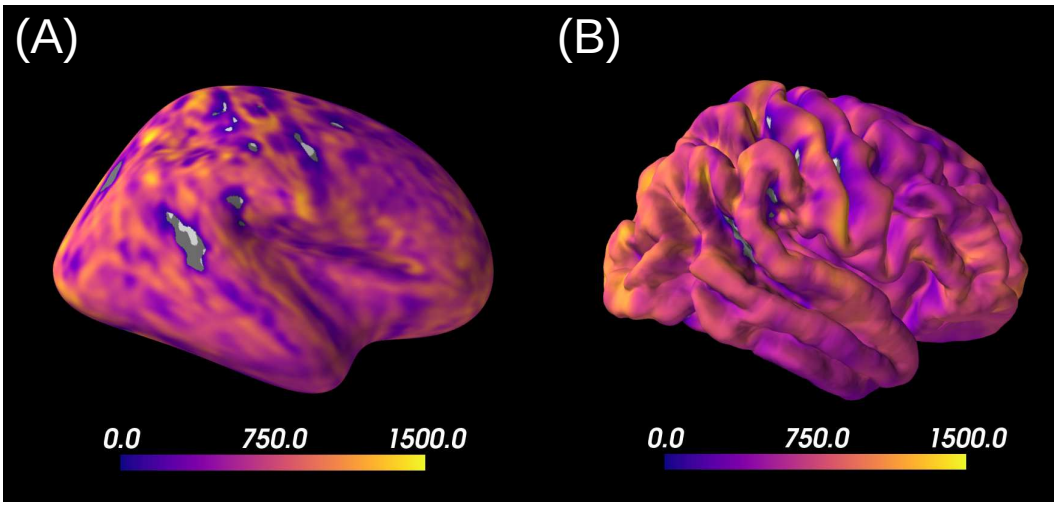


Fig. 5.8: C_s estimations averaged over 31 HCP MGH subjects, with unstable results masked, projected onto inflated (A) and pial (B) surfaces. Interestingly, low confidence areas correspond to the fundus of the sulci.

Figure 5.9 A reports the soma proportion (parameter f_s) averaged over 31 HCP MGH subjects, masking unstable results. Mean soma signal proportion in the cortex equals 0.22 (mean computed in trusted estimations only). These results are coherent with the mean volume occupancy of $10-20\%$ observed in gray matter (Shapson-Coe et al., 2021). To interpret the results at a region-based level, we have superimposed the soma proportion estimations with the Brodmann atlas. We observe a general agreement between the estimations and the atlas, and more particularly in the somatosensory and Broca's areas. A clear difference in soma proportion can be observed in the 12 regions, as presented in the barplot.

Despite the C_s parameter being useful for avoiding indeterminacies in the model inversion, its biological interpretation remains difficult. With the goal of relating our C_s estimation with physiological insight, we estimated soma radius r_s by fixing soma diffusivity. Similarly to the SANDI method, we fixed soma diffusivity D_s to the value of the self-diffusion coefficient of free water ($3\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$). Note that this value could be adjusted for each voxel, and is only used here in a matter of comparison and interpretation. Using a fixed-point method, we computed the soma radius r_s from the averaged C_s map obtained from our posterior distribution. The results are portrayed in Figure 5.9 B. We see that the estimated soma radius vary between 8 and $14\,\mu\mathrm{m}$, which is in accordance with histology (Palombo et al., 2021). Mean C_s values are presented in the barplot beneath the soma radius estimations in the different Brodmann regions.

5.4.3 HCP 1200

Figure 5.10 shows the results obtained on a database with only three b-shells. We see that 55 % of the C_s estimations on brain gray matter is considered unstable and is, therefore, masked. Indeed, our q-bounded RTOP approximation relies on high b-values, where the signal is expected to have converged towards a value that depends on the radius of the soma. The larger the soma, the sooner the q-bounded RTOP converges. However, the largest b-value contained in this database equals $3\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$, which is not enough for the signal to have converged. Thus, the poor quality of the summary statistics estimation leads to rather wide posterior distributions, resulting in unreliable results, as shown in simulations. Note, however, that estimations of the superior temporal gyrus for example are not masked, and both data sets seem to indicate large neurones in that region. The estimation of D_a is considered as unstable for 98.7% of the data set. This behavior was expected, given the results presented in Figure 5.4, for which a large bias is observed.

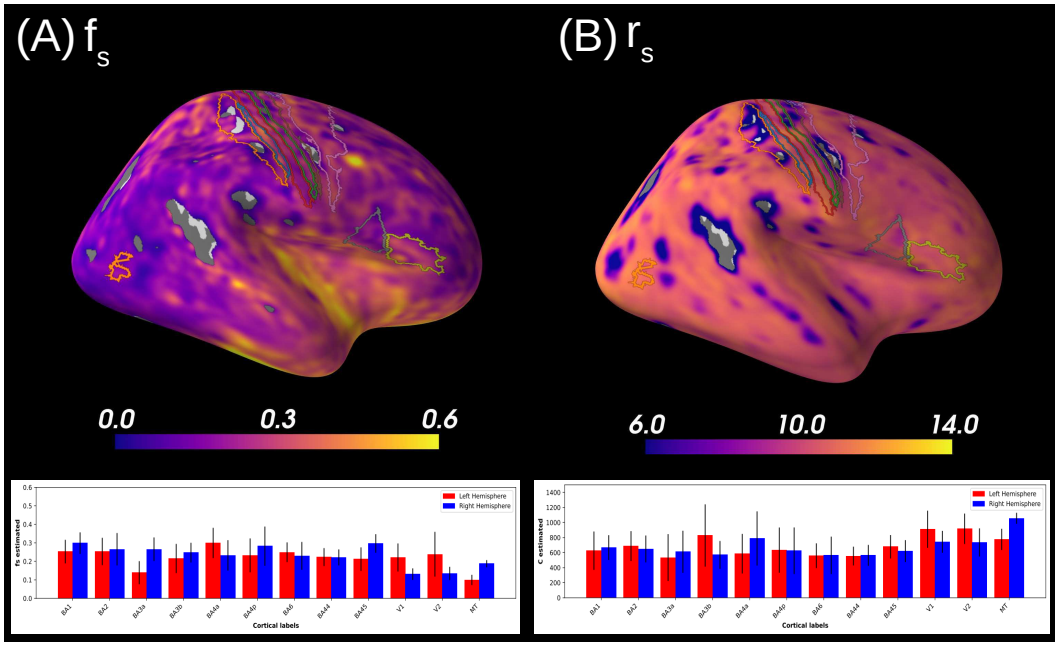


Fig. 5.9: (Left) The average soma proportion (f_s) over the 31 HCP MGH subjects is projected onto an inflated cortical surface, with unstable results masked; see text for details. We also show the main Brodmann areas available on FreeSurfer. The mean values of f_s over these regions (using only reliable estimations) are reported on the bar plots below. (Right) Soma radius map computed from C_s with soma diffusivity fixed to $D_s = 3\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$, averaged on 31 subjects. The bar plots below report the C_s mean values on main Brodmann areas. We observe a good agreement between our reported values and the Brodmann areas.

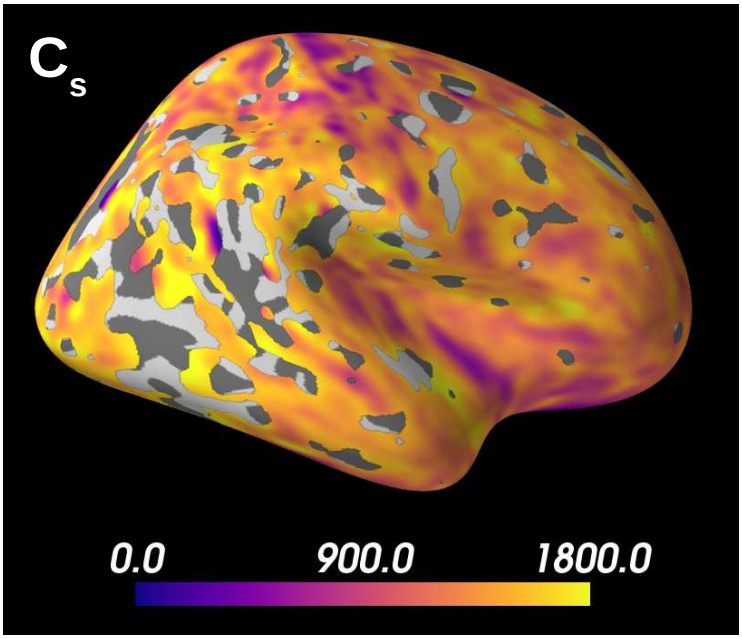


Fig. 5.10: Spatial distribution of the C_s estimations averaged over 30 HCP 1200 subjects and projected onto an inflated cortical surface, with unstable estimations masked. Due to the scarce and rather low b-values ($\leq 3\,\mathrm{ms}\,\mu\mathrm{m}^{-2}$) used in the database, the results are very unstable and, therefore, many voxels are discarded.

5.5 Discussion

5.5.1 Validation simulated data

An important aspect of our work is the thorough validation that we have carried out on simulated data, using different acquisition setups and ways of generating diffusion signals. Part of this validation had the goal of ensuring that our method gave consistent results on simple standard cases, as confirmed by the results in Figure 5.2. Additionally, we have demonstrated the benefits of using summary statistics for describing dMRI signals, attaining better parameter estimates when using them instead of directly manipulating the diffusion signals. Such results are very encouraging and should push other researchers into using these summary statistics for processing their dMRI signals.

Another relevant byproduct of our validation was observing that the quality of the parameter estimations depends heavily on the distribution of b-values used to acquire and simulate the dMRI signals. Indeed, if only small b-values are available, the summary statistics of dMRI signals are poorly estimated and the parameter estimation too. Based on this observation, we were able to apply our LFI pipeline to real datasets HCP MGH and HCP 1200 knowing in advance the limitations of our methodology; for example, we knew from which parameters we could expect good estimates (mostly soma-related ones) and which ones should not be taken into account in our analysis.

5.5.2 C_s: A proxy to soma size

Estimating both soma radius (r_s) and diffusivity (D_s) with diffusion MRI is a challenging task. When trying to estimate them separately, we can expect a 'banana-shape' in their joint posterior distribution as shown in Figure 5.5. This indicates that several values of the pair (r_s, D_s) can explain the observed signal with high probability and, therefore, one is confronted with model indeterminacy.

The new parameter C_s that we introduce in this paper is modulated both by the soma radius and its diffusivity. Thus, estimating it directly avoids problems of indeterminacy, as shown in Figure 5.2, for example. Note, however, that avoiding such indeterminacy comes with the price of losing specificity and, therefore, physiological interpretations. Fortunately, acquisitions in the narrow or wide pulse regimes allow us to better interpret estimations of C_s, as it only depends on r_s or D_s (see Section 4.2.1). The HCP MGH and HCP 1200 datasets do not correspond to one of those cases, leading to the impossibility to disentangle soma radius and diffusivity uniquely.

5.5.3 Can I apply this approach to my data?

One of the main benefits of a probabilistic framework is that it can be applied to any data set or acquisition setup without too much hesitation, since the estimates always come with a "quality certificate" described by the credible regions derived from the posterior distribution. We have benefited from this feature when creating all figures related to databases HCP MGH and HCP 1200, since they allow us to mask regions for which the variance of the parameter estimation is too high. We can also use it as a proxy to identify regions for which the three-compartment model is adequate or not, or assessing whether the distribution of b-values used to acquire the observed diffusion signal is sufficiently informative. Note, however, that it would not be realistic to expect that our method should give acceptable results on every database to which it is applied. Indeed, the distribution of b-values used to acquire the data under study is a key predictor of whether the estimated posterior distribution will be useful for inverting the three-compartment model or not. For instance, if the b-values are too small, then the RTOP summary statistics will be poorly estimated in most voxels, leading to unreliable estimations; only regions with larger somas will be correctly estimated. Similarly, if the acquisition uses only b-values greater than 2.5 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$, the Spiked LEMONADE moments will be biased, which also leads to neurite estimations with large variability. These observations were useful when analyzing the `HCP 1200` database, since the distribution of b-values are concentrated at low values and, therefore, our estimations are prone to bias and have large variances.

5.6 Limitations and perspectives

There are many extensions that we could envision for the method that we propose. We present preliminary results in this section, for which two publications are in preparation.

5.6.1 Application to other tissues

The proposed approach has been designed and applied to brain gray matter, but one could also want to apply it to brain white matter (see Supplementary Material, Figure 5.16). Results on the HCP MGH dataset indicate that the axons distribution is more anisotropic in white matter than gray matter ($p_{2,WM} > p_{2,GM}$), with values coherent to the ones obtained with the LEMONADE framework (Novikov et al., 2018c). Somas are also less present in white matter than gray matter, as indicated by a lower signal fraction f_s, which is expected from histology (Shapson-Coe et al., 2021). However, the ECS model used in the three-compartment model defines its diffusivity as isotropic. While this assumption seems to hold in brain gray matter, a tensor representation is usually preferred for the ECS diffusivity, as it is the case in the SM (Novikov et al., 2018a). The application of a similar LFI pipeline based on a model designed for brain white matter (such as the SM) could help improve the estimations of tissue parameters and better interpret its output thanks to the posterior distributions. For example, this pipeline could be a new way to solve the LEMONADE system of equations proposed by Novikov et al. (2018c) within a probabilistic framework, and help obtaining radius distributions of axons in the white matter (Veraart et al., 2020).

We also decided to apply the proposed method to the cerebellum, which is divided into white matter and cortex (Eccles, 2013). Parameters estimations averaged over all HCP MGH subjects are presented in Figure 5.11 in coronal, axial and lateral slices of the brain in MNI space. Only voxels deemed stable in each subject were used to compute the mean estimation per voxel. Voxels in the MNI space were less than five subjects presented stable estimations were themselves considered as unstable and left empty. Cerebellum white matter is segmented in green, and cerebellum cortex in magenta. The associated parameter distributions in cerebellum white matter and cortex are presented in Figure 5.12. We recall that each voxel was estimated independently, without priors on their spatial position. The obtained estimations of C_s indicate the presence of big neurons in the cerebellum, agreeing with the presence of Purkinje cells, some of the largest neurons in the human brain (Marr and Thach, 1991). A difference in neurite diffusivity between estimations within the white matter and the cortex can be observed. A higher

diffusivity in the cerebellum cortex with respect to the cerebellum white matter seems coherent with the estimations presented in Figure 5.16. The ECS proportion in the cerebellum cortex appears to be surprisingly high, while a denser concentration of neurons is expected compared to brain gray matter (Trepel, 2021). The method also seems to struggle with the estimation of the compartment proportions in the cerebellum white matter, where the somas and neurites proportions are (almost) all considered unstable. A strongest isotropy is observed in white matter compared to gray matter, maybe due to the presence of many axonal intersections.

5.6.2 Model Improvements

The three-compartment model that we use approximates well the intracellular signal in brain gray matter tissue by adding a sphere compartment to account for soma presence (Palombo et al., 2020). However, this is a rather simplified model and it could be improved; for instance, the geometry of ECS is very restricted and tortuous, and diffusion signals have been proven to deviate from a mono-exponential behavior (Vincent et al., 2021). A first improvement could be, as mentioned before, to deviate from modeling the ECS diffusion with a simple isotropic Gaussian and consider more complex geometric representations.

Another improvement would be to consider exchange between the neurites and the ECS, as very recently suggested by Jelescu et al. (2021) and Olesen et al. (2021). We therefore extended the proposed LFI approach to the second model presented in Chapter 4 (see Section 4.3). Figure 5.13 presents the updated framework used. Simulations have been generated with $f_s = 0.33$, $f_n^{ex} = 0.5$, $D_n = 2\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$, $C_s = 1000\,\mu\mathrm{m}^2$, which corresponds to a soma radius $r_s = 16.7\,\mu\mathrm{m}$ and diffusivity $D_s = 3\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$, and $D_e = 1\,\mu\mathrm{m}^2\,\mathrm{ms}^{-1}$. We set $\Delta = 21.8$ ms and δ =12.9 ms, similar to the HCP MGH acquisition protocol. We investigated the utility of the introduced summary statistics by comparing the posterior distributions obtained when directly fitting the powder-averaged signal, as done for the SMEX model in Olesen et al. (2021). We generated a diffusion signal following their high q-value approximation (corresponding to the second line of equation 4.22 in Chapter 4), with five b-values comprised between 6 and 10 $\mathrm{ms}\,\mu\mathrm{m}^{-2}$. This interval was chosen to respect the minimal b-value for which this equation is considered valid while keeping a reasonable range of values. Two exchange cases have been considered: first, we considered a relatively high exchange time (t_{ex}) of 60 ms, where the influence of exchange can be considered negligible on the signal (Figure 5.14), and then a small t_{ex} of 10 ms, i.e. a high exchange rate between the neurites and ECS compartments (Figure 5.15). Results for a

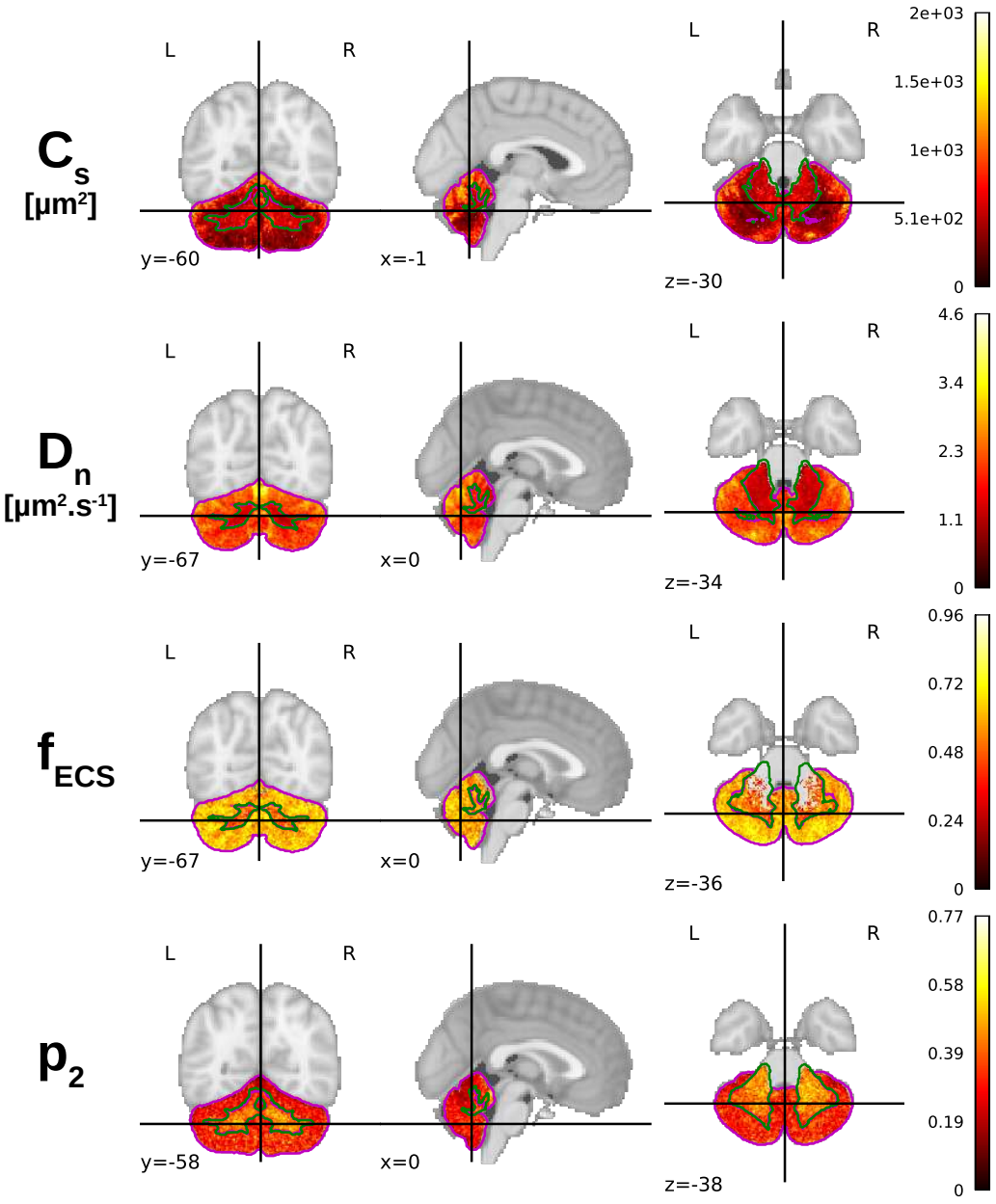


Fig. 5.11: Microstructure parameters estimation in the cerebellum, with estimations considered unstable masked. We can observe the presence of big somas in the cortex, a smaller neurite diffusivity along with a stronger isotropy and more unstable ECS fraction estimations in the white matter compared to the cerebellum cortex.

different simulated neuron can be found in the Appendix (Figure 5.17 and Figure 5.18). Plots in the diagonal correspond to the posterior distributions of each estimated parameter, the red line indicating ground truth values used to generate the signal. Upper-diagonal plots correspond to their joint posterior distributions. Indeterminacies are identifiable by the large green regions in the joint posterior distributions. Each green point indicates a possible parameter combination that could have generated the observed summary statistics (A) or the diffusion signal (B). Figure 5.14 A and Figure 5.14 A

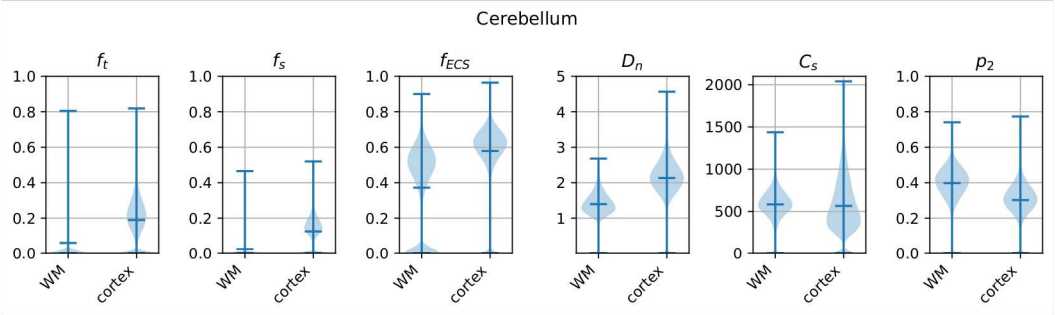


Fig. 5.12: Microstructure parameters estimation in the cerebellum. Most estimations of signal fractions are unstable in cerebellum white matter (WM). We can observe the presence of big somas in the cortex, a smaller neurite diffusivity and a stronger isotropy in the white matter compared to the cerebellum cortex.

show the capacity of our model to estimate the tissue parameters from the summary statistics with good confidence. Exchange between the neurites and ECS seems to influence the estimation of D_n, whose posterior distribution is wider for high exchange. In comparison, the direct estimation from the powder-averaged signal indicates large indeterminacy regions, as can be seen in the joint posterior distributions of Figure 5.14 B and Figure 5.15 B. This indeterminacy means that multiple tissue parameters can produce similar observed diffusion signal. Using summary statistics instead allows to greatly improve the robustness of the estimation.

5.6.3 Summary Statistics Extraction

In this method, we require to have a signal with large q-values to extract summary statistics using RTOP. This constraint limits the number of possible applications, as many dMRI datasets do not contain such high q-values. A solution proposed by Golkov et al. (2016) is to rely on deep learning to learn a mapping between a limited number of DWIs and some microstructure scalars, assuming that all the necessary information is contained in the low-dimensional acquisition. Following this q-space deep learning method, an improvement of the proposed method could be to learn the summary statistics from a dataset with less q-shells using deep learning, and then apply our likelihood-free inference method to solve the inverse problem.

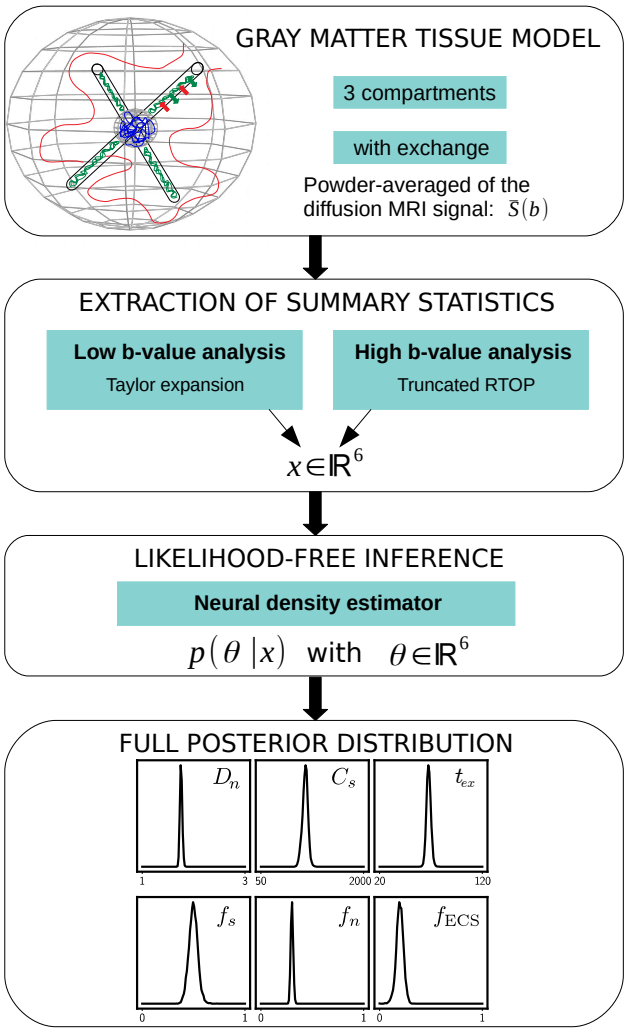


Fig. 5.13: Visual schematic of the gray matter tissue microstructure estimation pipeline. We start by modeling brain gray matter with three compartments: somas, neurites and extra-cellular space, with exchange between neurites and the ECS (see Section 4.3). We then extract summary statistics from the powder-averaged signal using a low and high b-values analysis (Section 4.3.2). Applying a likelihood-free inference algorithm allows to estimate both the parameters and their full posterior distribution.

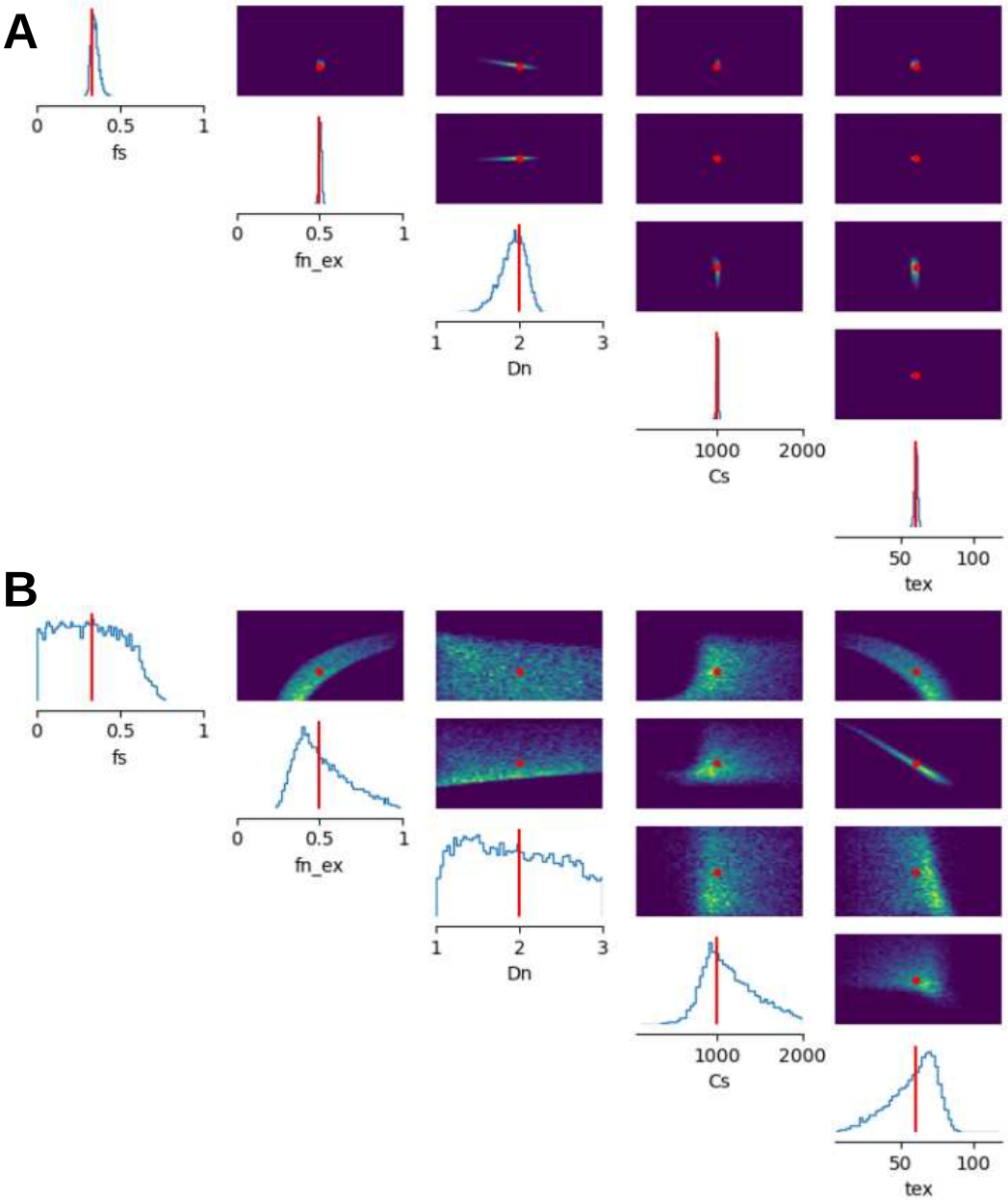


Fig. 5.14: Estimation of $\boldsymbol{\theta}$ with a relatively small exchange time using a LFI method. Diagonal: posterior distribution of each tissue parameters. Upper-triangle: joint posterior distributions. A: Estimations using summary statistics extracted from a large and small q-values analysis of the diffusion signal. B: Estimations using the powder-averaged signal directly, similarly to Olesen et al. (2021). The use of summary statistics allows to greatly reduce the indeterminacies in the parameter estimations.

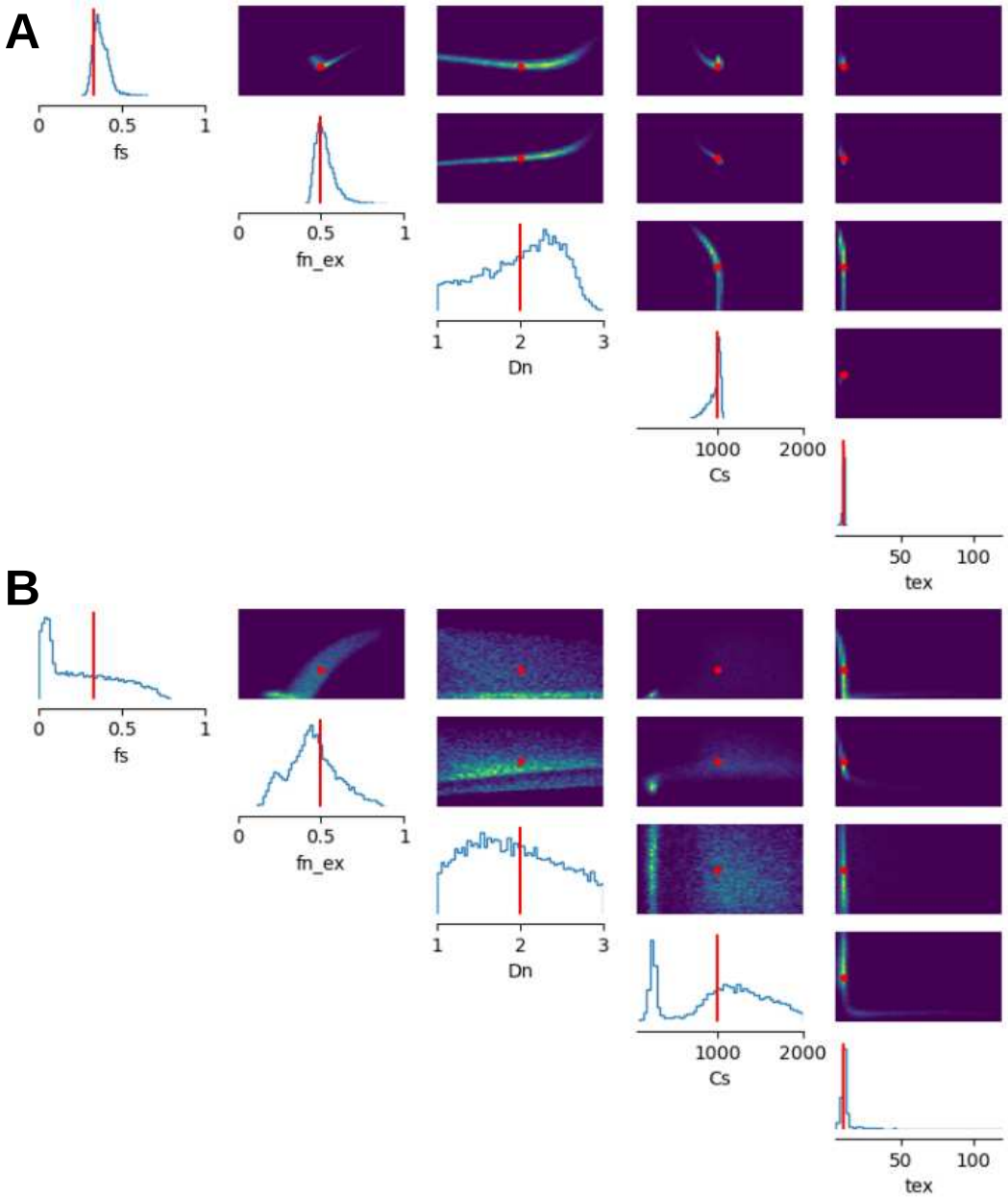


Fig. 5.15: Estimation of $\boldsymbol{\theta}$ with a fast exchange time using a LFI method. Diagonal: posterior distribution of each tissue parameters. Upper-triangle: joint posterior distributions. A: Estimations using summary statistics extracted from a large and small q-values analysis of the diffusion signal. Estimation of the diffusivity within the neurites is impacted by its exchange with the ECS. B: Estimations using the powder-averaged signal directly, similarly to Olesen et al. (2021). The use of summary statistics allows to greatly reduce the indeterminacies in the parameter estimations.

5.7 Appendix

White matter results

In this appendix we present additional experimental results of brain white matter microstructure estimations using the proposed LFI method.

Figure 5.16 A presents a comparison of the mean estimations of white matter and gray matter parameters among all HCP MGH subjects, keeping only trusted voxels. We can observe a lower proportion of somas in white matter ($f_{s,\text{WM}} < f_{s,\text{GM}}$) along with smaller soma size compared with gray matter ($C_{s,\text{WM}} < C_{s,\text{GM}}$). This weak presence of small somas in white matter is in accordance with histology. These results also indicate a more anisotropic distribution of axons in white matter than neurites in gray matter ($p_{2,\text{WM}} > p_{2,\text{GM}}$). Obtained p_2 values are coherent to the ones obtained with the LEMONADE framework (Novikov et al., 2018c). Figure 5.16 B presents the mean p_2 values in white matter averaged over all subjects, keeping only trusted estimations. The center of brain white matter appears more anisotropic than at the frontier with gray matter.

In this paper we approximate the ECS with an isotropic diffusion. Models specific to brain white matter, such as the SM, usually represent ECS diffusivity as a non-isotropic tensor (Novikov et al., 2018a). Results obtained from this model should then be taken with caution, because this model does not reflect white matter tissue properly.

SANDI model with neurites-ECS exchange

Experiments similar to those in Section 5.6.2 have been performed to assess the quality of the proposed approach estimations and investigate the utility of the introduced summary statistics. Simulations have been generated with $f_s = 0.4$, $f_n^{ex} = 0.33$, $D_n = 1.5\,\mu\text{m}^2\,\text{ms}^{-1}$, $C_s = 500\,\mu\text{m}^2$, which corresponds to a soma radius $r_s = 10.8\,\mu\text{m}$ and diffusivity $D_s = 3\,\mu\text{m}^2\,\text{ms}^{-1}$, and $D_e = 1\,\mu\text{m}^2\,\text{ms}^{-1}$. We compare the obtained estimations with a diffusion signal generated following Olesen et al. (2021)'s high q-value approximation, with five b-values comprised between 6 and 10 $\text{ms}\,\mu\text{m}^{-2}$ (see the interval justification in Section 5.6.2). Figure 5.17 and Figure 5.18 present the results for exchange times equal to 60 ms and 10 ms respectively.

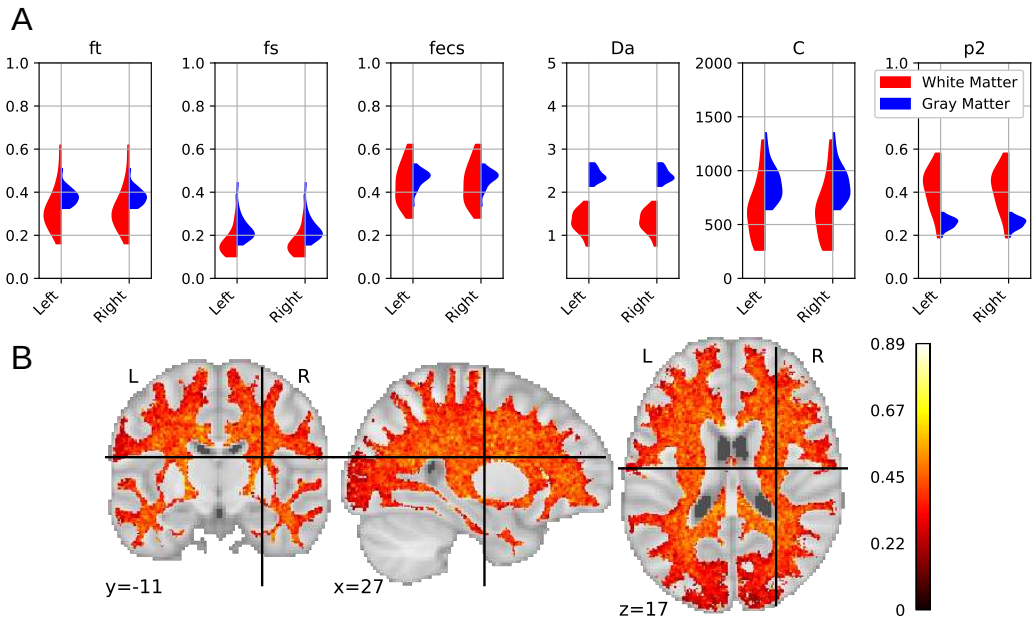


Fig. 5.16: Mean estimation over HCP MGH subjects of brain white matter microstructure parameters using LFI methods on a three-compartment model. A. Higher p_2 values are observed in brain white matter compared to gray matter, indicating a more anisotropic axon distribution. Soma proportion is also reduced compared to gray matter estimations, along with smaller soma size. These estimations, although encouraging, should be taken with caution, because the ECS model used here is not suited for white matter tissue. B. p_2 estimations in brain white matter, 0 indicating an isotropic distribution of axons, and 1 an anisotropic distribution (i.e. perfectly aligned fibers).

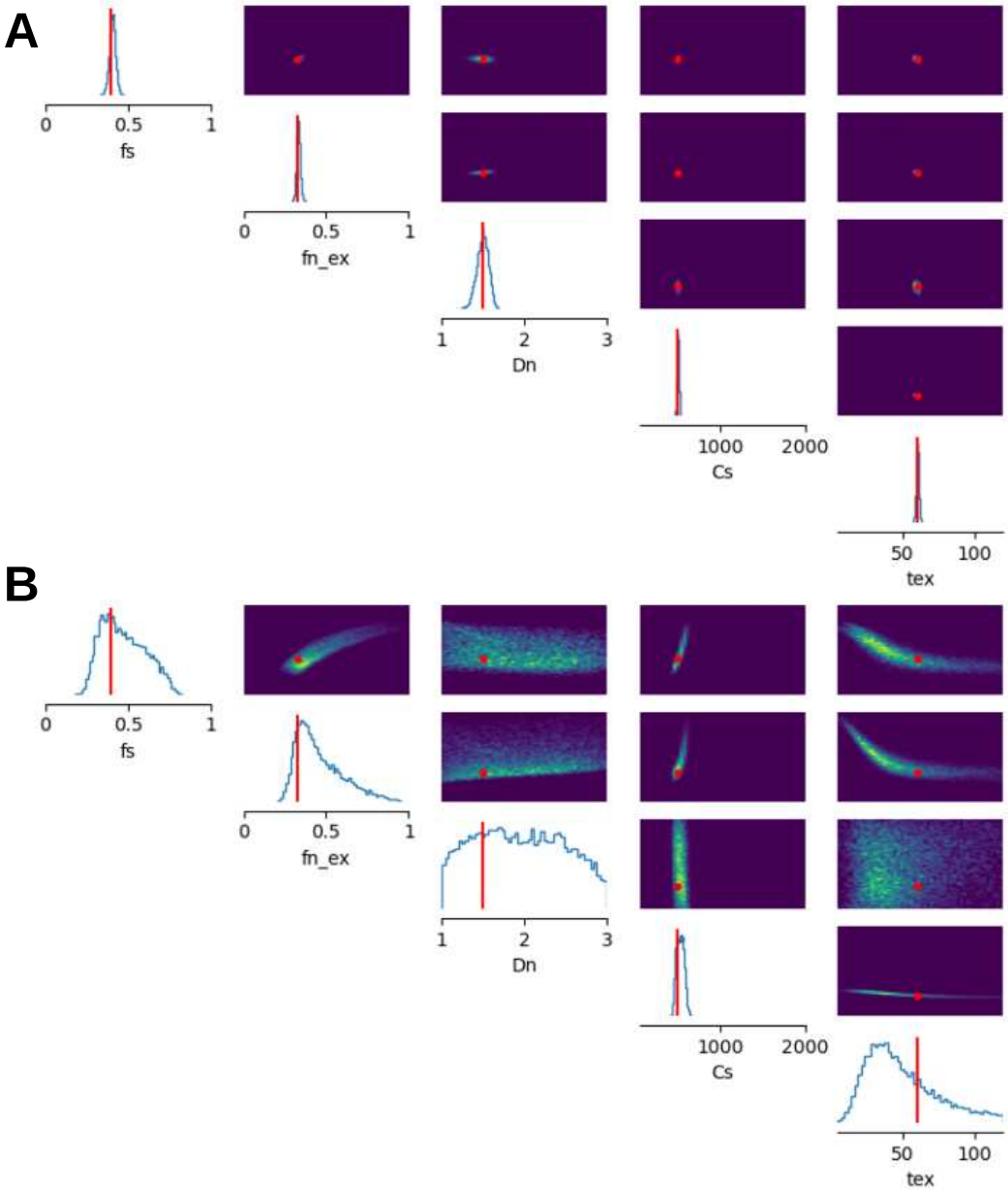


Fig. 5.17: Estimation of $\boldsymbol{\theta}$ with a relatively small exchange time using a LFI method. Diagonal: posterior distribution of each tissue parameters. Upper-triangle: joint posterior distributions. A: Estimations using summary statistics extracted from a large and small q-values analysis of the diffusion signal. B: Estimations using the powder-averaged signal directly, similarly to Olesen et al. (2021). The use of summary statistics allows to greatly reduce the indeterminacies in the parameter estimations.

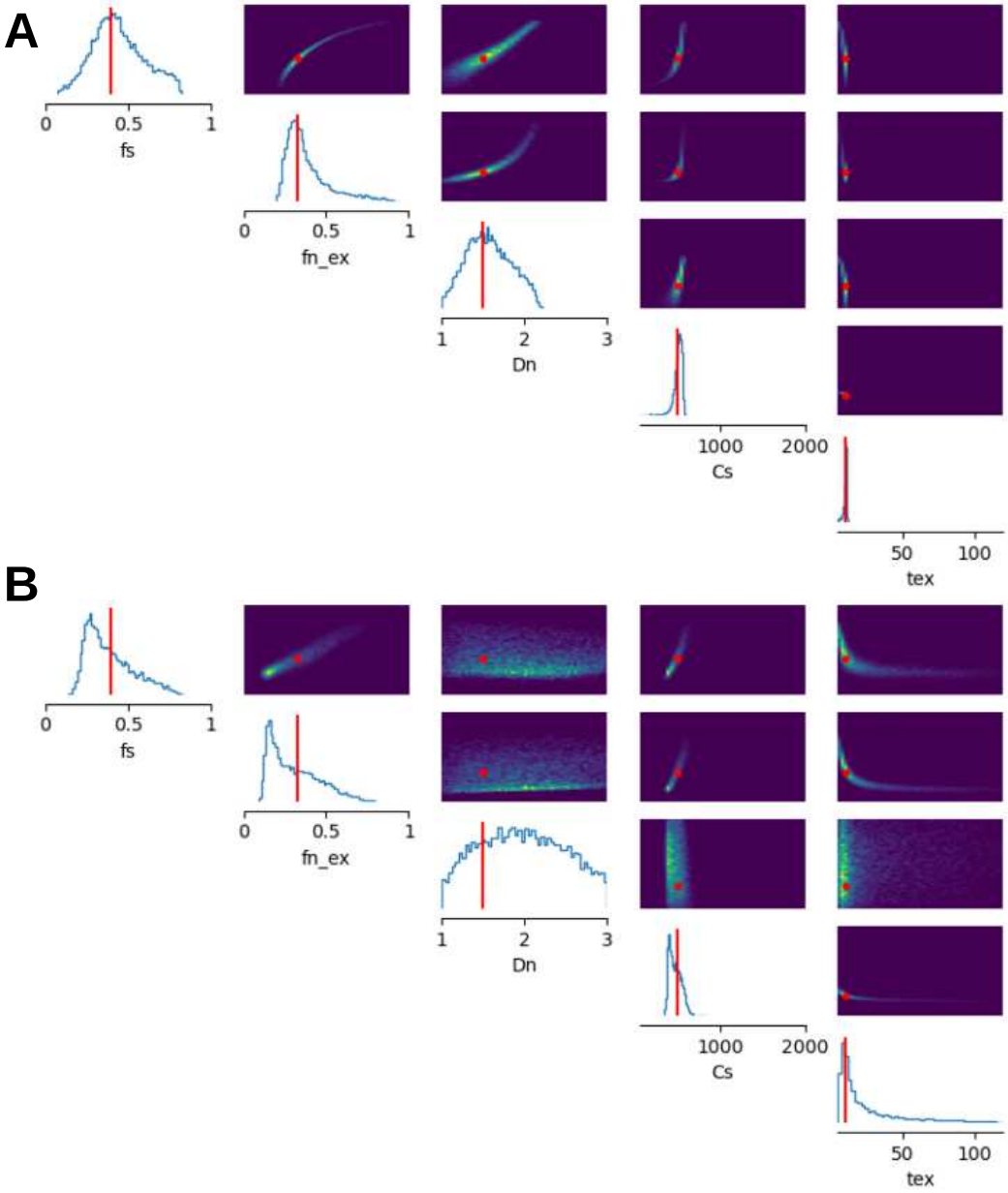


Fig. 5.18: Estimation of $\boldsymbol{\theta}$ with a fast exchange time using a LFI method. Diagonal: posterior distribution of each tissue parameters. Upper-triangle: joint posterior distributions. A: Estimations using summary statistics extracted from a large and small q-values analysis of the diffusion signal. Estimation of the diffusivity within the neurites is impacted by its exchange with the ECS. B: Estimations using the powder-averaged signal directly, similarly to Olesen et al. (2021). The use of summary statistics allows to greatly reduce the indeterminacies in the parameter estimations.

6 Conclusion

Quantifying brain gray matter tissue composition is challenging. In Chapter 4, we presented two gray matter models. The first one comprises three compartments: neurites, somas, and ECS. It considers all membranes impermeable, similar to the model SANDI (Palombo et al., 2020). Then, we extended this model to account for exchanges between neurites and the ECS based on very recent publications arguing for its non-negligible impact on the diffusion MRI signal (Jelescu et al., 2021; Olesen et al., 2021). We presented preliminary results for this new model, which requires further analysis. We proposed a signal analysis based on large and small q-values approximations to extract summary statistics for both models, allowing us to reduce the dimensionality of the data and the indeterminacies in the parameter estimations. Then, in Chapter 5, we solved the non-linear inverse problem of relating the summary statistics of the impermeable three-compartment model to tissue microstructure parameters using a Likelihood-Free Inference (LFI) method. Such an approach estimates posterior distributions of the fitted parameters. This rich description provides many useful tools, such as assessing the quality of the parameter estimation or characterizing regions in the parameter space where it is harder to invert the model. The inclusion of such a "quality certificate" accompanying our parameter estimation is very useful in practice and allows one to apply the pipeline on any database and know to which degree one can trust the results. Moreover, our proposal alleviates limitations from current methods in the literature by not requiring physiologically unrealistic constraints on the parameters and avoiding indeterminacies when estimating them (see Section 2.4). We presented some limitations of this model, such as the requirement to have large b-shells. Initial simulation results on the three-compartment model with exchanges between neurites and the ECS have also been presented.

Future directions

In the following paragraphs, we list a few perspectives for the works developed in this thesis.

To begin with, the tissue microstructure estimations obtained from the LFI method based on the three-compartment impermeable model could be used

in cognitive studies. The obtained parameter estimations show characteristic values for different cytoarchitectural regions, as illustrated in Figure 6.1. We computed the mean estimates across subjects in the HCP MGH dataset in regions defined by the Jülich atlas (Amunts et al., 2020), keeping only trusted estimations (see definition Section 5.4.2). We also added the condition to keep a region estimation for each parameter if it contained at least 20 voxels and the mean was computed on at least five subjects. This estimation framework could be used, for example, to follow the progression of Alzheimer's disease (Whitehouse et al., 1982) or study neuron sizes in the cerebellum in autism patients (Fatemi et al., 2012). We hope that other researchers will find our approach helpful for their applications.

Other improvements can be developed by going beyond the two biological models presented in Chapter 4. Indeed, brain gray matter modeling is still an open question in the diffusion MRI community. The addition of somas, as well as the consideration of exchanges between neurites and the ECS in gray matter models, are very recent (see e.g. Palombo et al., 2020; Jelescu et al., 2021; Olesen et al., 2021). A consensus has not been reached yet concerning, for example, the inclusion of non-exchanging neurites, such as the model eSANDIX (Olesen et al., 2021), or the need to model structural disorder (Lee et al., 2020). The ECS might also require a more complex modelization than an isotropic Gaussian diffusion (Vincent et al., 2021). More experiments are necessary to understand the impact of each possible model improvement. The challenge is to account for physical phenomena that have a relevant impact on the acquired dMRI signal while keeping as few parameters as possible (Novikov et al., 2018b).

Another interesting line of work would be the study of the posterior distributions, which allow having a full description of the solution landscape. In the current framework, we have not taken advantage of the full potential of these posterior distributions, as we are only using the mean and standard deviation to define a confidence factor. This definition of confidence in an estimation presents several limitations. First, it is only valid for Gaussian distributions, which we do not always obtain. Consider, for example, the posterior distribution of the parameter C_s in Figure 6.2. Our current confidence factor will define such estimation as not trusted. However, such bi-modality could originate from the presence of somas with two different radii and/or diffusivities in the tissue. A second limitation comes from the difficulty of interpreting the obtained posterior distributions. Indeed, we mentioned before the possibility that the posterior distribution in Figure 6.2 comes from a tissue containing somas with two different characteristics. Another explanation could be that

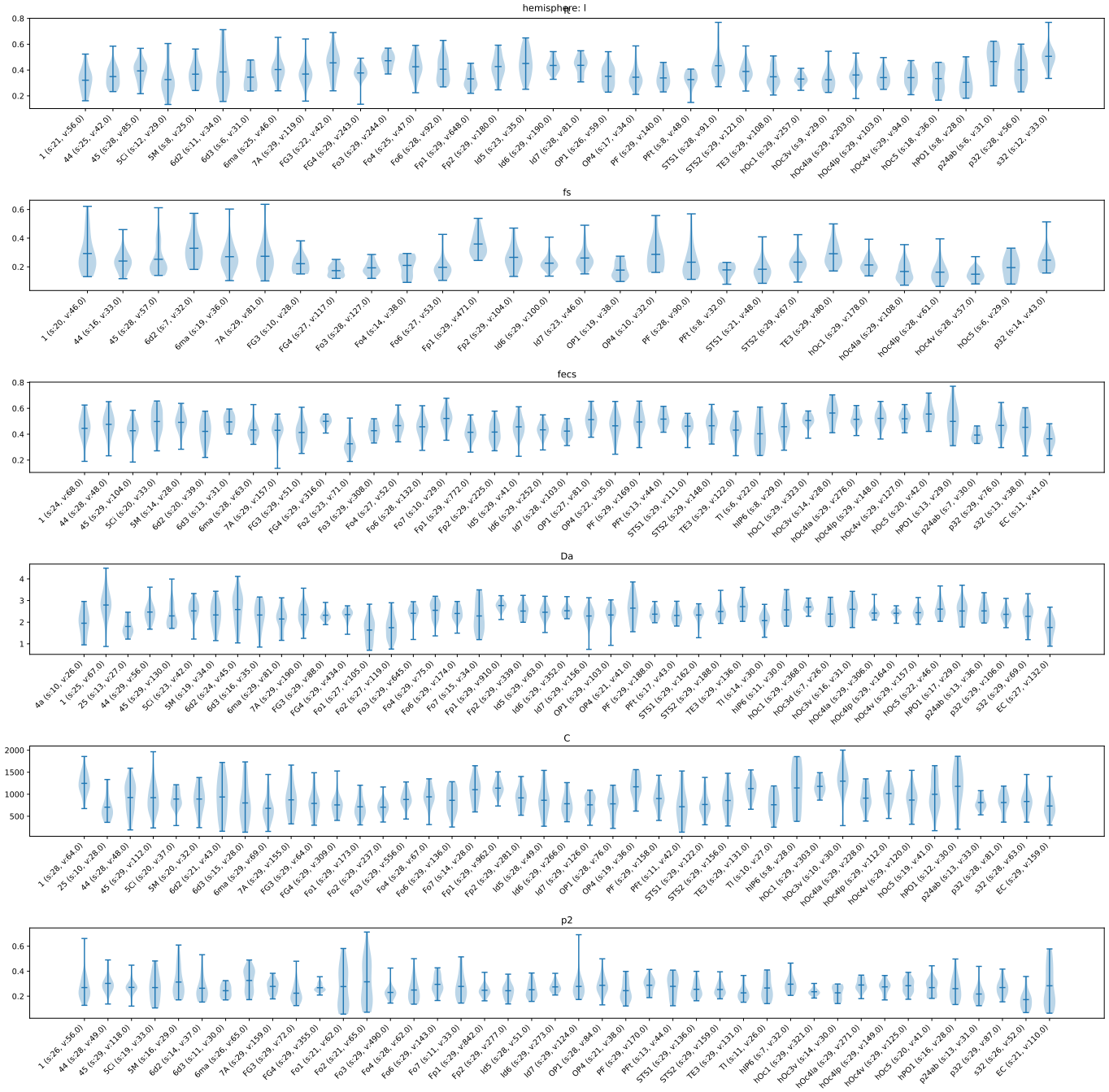


Fig. 6.1: Estimations of brain tissue microstructure parameters of the HCP MGH subjects for regions of the Jülich atlas (Amunts et al., 2020) (left hemisphere). Mean estimations per region are computed per subject using only trusted voxels. We only kept regions with at least five subjects whose mean estimation was based on at least 20 voxels. For each region, the number of subjects kept for computing the mean and the mean number of voxels per subject are indicated in parenthesis with the letters s and v, respectively.

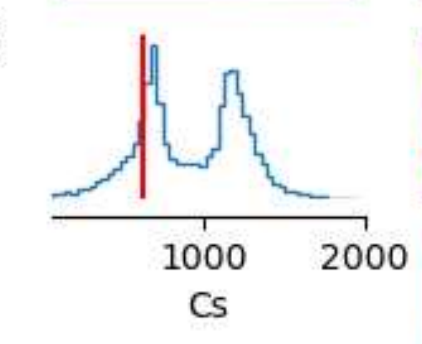


Fig. 6.2: Multi-modal posterior distribution of C_s, a proxy to soma radius. Such a posterior could describe either the presence of neurons with two different soma radii or that two parameter distributions can equally the acquired diffusion signal.

two sets of parameters explain the acquired diffusion signal equally well. Choosing between those two explanations is so far not possible.

To overcome these limitations, we envision two different solutions. The definition of a better confidence measure is first required for taking into account as many posterior distribution shapes as possible. Secondly, we need to define summary statistics that do not allow indeterminacies in the inverse problem resolution. The summary statistics proposed in this thesis are extracted from a large and small q-value analysis and seem to reduce indeterminacies compared to existing solutions considerably (see e.g. Jelescu et al., 2016). However, mathematical proof has not been derived from them to ensure it.

It would also be interesting to develop an automatic method to extract summary statistics from a diffusion signal to help find summary statistics that capture all the signal information. That method could be applicable on datasets with a limited number of b-shells and even in clinically-feasible scenarios, similar to the work by Golkov et al. (2016). Summary statistics would be specifically defined for a given biological model without requiring defining features that capture the characteristics of the acquired diffusion MRI signal using a signal analysis pipeline. Following this idea, an ideal protocol for data acquisition could be defined for tissue microstructure estimations. We could envision that this protocol relies on a few b-shells only, with clinically-accessible b-values. Diffusion time could also be chosen to fall into the narrow pulse regime or the Neuman regime, to estimate soma diffusivity or radius directly (see Section 4.2.1).

In addition, this thesis focuses on single encoding diffusion acquisitions. Extensions to different pulse sequences could be envisioned.

To conclude, we believe that our approach based on Bayesian inference with modern tools from neural networks is a promising one that can easily be applied to other applications in the medical imaging field: one only needs to define a sufficiently rich model describing a certain phenomenon of interest, and the LFI pipeline will manage to invert it and provide a related posterior distribution. We expect, therefore, that other researchers will find this contribution valuable for their own applications and see such a probabilistic approach more often used in the literature.

Bibliography

Agdestein, Syver Døving, Try Nguyen Tran, and Jing-Rebecca Li (2021). "Practical computation of the diffusion MRI signal based on Laplace eigenfunctions: permeable interfaces". In: *NMR in Biomedicine*, e4646 (cit. on p. 84).

Alexander, D. C and K. K Seunarine (2010). "Mathematics of crossing fibers". In: *Diffusion MRI: theory, methods, and application. Oxford*, pp. 451–464 (cit. on p. 26).

Allman, John M., Nicole A. Tetreault, Atiya Y. Hakeem, et al. (June 2010). "The von Economo Neurons in Frontoinsular and Anterior Cingulate Cortex in Great Apes and Humans." English. In: *Brain Structure and Function* 214, pp. 495–517 (cit. on pp. 100, 106, 107).

Alsing, Justin, Tom Charnock, Stephen Feeney, and Benjamin Wandelt (2019). "Fast likelihood-free cosmology with neural density estimators and active learning". In: *Monthly Notices of the Royal Astronomical Society* 488.3, pp. 4440–4458 (cit. on p. 43).

Alsing, Justin, Benjamin D Wandelt, and Stephen M Feeney (2018). "Optimal proposals for approximate Bayesian computation". In: *arXiv preprint arXiv:1808.06040* (cit. on p. 52).

Amor, Sandra, Fabiola Puentes, David Baker, and Paul Van Der Valk (2010). "Inflammation in neurodegenerative diseases". In: *Immunology* 129.2, pp. 154–169 (cit. on p. 1).

Amunts, Katrin, Hartmut Mohlberg, Sebastian Bludau, and Karl Zilles (2020). "Julich-Brain: A 3D probabilistic atlas of the human brain's cytoarchitecture". In: *Science* 369.6506, pp. 988–992 (cit. on pp. 124, 125).

Amunts, Katrin, Axel Schleicher, Uli Bürgel, et al. (1999). "Broca's region revisited: Cytoarchitecture and intersubject variability". In: *Journal of Comparative Neurology* 412.2, pp. 319–341 (cit. on pp. 100, 106, 107).

Andrieu, Christophe, Nando De Freitas, Arnaud Doucet, and Michael I Jordan (2003). "An introduction to MCMC for machine learning". In: *Machine learning* 50.1, pp. 5–43 (cit. on p. 44).

Assaf, Y. and O. Pasternak (Jan. 2008). "Diffusion Tensor Imaging (DTI)-based White Matter Mapping in Brain Research: A Review". en. In: *Journal of Molecular Neuroscience* 34.1, pp. 51–61 (cit. on pp. 2, 12).

Balinov, B., B. Jonsson, P. Linse, and O. Soderman (1993). "The NMR Self-Diffusion Method Applied to Restricted Diffusion. Simulation of Echo Attenuation from Molecules in Spheres and between Planes". In: *Journal of Magnetic Resonance, Series A* 104.1, pp. 17–25 (cit. on pp. 29, 70).

Basser, P.J., J. Mattiello, and D. Lebihan (1994). “Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo”. In: *Journal of Magnetic Resonance, Series B* 103.3, pp. 247 –254 (cit. on pp. 2, 22, 23).

Beaulieu, C. (2002). “The basis of anisotropic water diffusion in the nervous system – a technical review”. In: *NMR in Biomedicine* 15.7-8, pp. 435–455. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.782 (cit. on pp. 23, 24).

Beaulieu, C. and P. S. Allen (1994). “Determinants of anisotropic water diffusion in nerves”. In: *Magnetic Resonance in Medicine* 31.4, pp. 394–400. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.1910310408 (cit. on p. 15).

Beaulieu, Christian (2014). “Chapter 8 - The Biological Basis of Diffusion Anisotropy”. In: *Diffusion MRI (Second Edition)*. Ed. by Heidi Johansen-Berg and Timothy E.J. Behrens. Second Edition. San Diego: Academic Press, pp. 155–183 (cit. on pp. 13, 15).

Beaumont, Mark A, Wenyang Zhang, and David J Balding (2002). “Approximate Bayesian computation in population genetics”. In: *Genetics* 162.4, pp. 2025–2035 (cit. on pp. 43, 60).

Behrens, T.E.J., M.W. Woolrich, M. Jenkinson, et al. (Nov. 2003). “Characterization and propagation of uncertainty in diffusion-weighted MR imaging”. en. In: *Magnetic Resonance in Medicine* 50.5, pp. 1077–1088 (cit. on p. 28).

Bishop, Christopher M (2006). “Pattern recognition”. In: *Machine learning* 128.9 (cit. on p. 42).

Blei, David M, Alp Kucukelbir, and Jon D McAuliffe (2017). “Variational inference: A review for statisticians”. In: *Journal of the American statistical Association* 112.518, pp. 859–877 (cit. on pp. 45–47).

Blum, Michael GB and Olivier François (2010). “Non-linear regression models for Approximate Bayesian Computation”. In: *Statistics and computing* 20.1, pp. 63–73 (cit. on p. 60).

Blum, Michael GB, Maria Antonieta Nunes, Dennis Prangle, and Scott A Sisson (2013). “A comparative review of dimension reduction methods in approximate Bayesian computation”. In: *Statistical Science* 28.2, pp. 189–208 (cit. on pp. 59, 60).

Boux, Fabien, Florence Forbes, Julyan Arbel, Benjamin Lemasson, and Emmanuel L Barbier (2021). “Bayesian inverse regression for vascular magnetic resonance fingerprinting”. In: *IEEE Transactions on Medical Imaging* (cit. on p. 43).

Brodmann, Korbinian (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth (cit. on pp. 100, 106).

Brown, Robert (1828). “XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies”. In: *The philosophical magazine* 4.21, pp. 161–173 (cit. on p. 13).

Callaghan, P. T., A. Coy, D. MacGowan, K. J. Packer, and F. O. Zelaya (June 1991). “Diffraction-like effects in NMR diffusion studies of fluids in porous solids”. en. In: *Nature* 351.6326, pp. 467–469 (cit. on p. 18).

Callaghan, P.T., K.W. Jolley, and J. Lelievre (Oct. 1979). “Diffusion of water in the endosperm tissue of wheat grains as studied by pulsed field gradient nuclear magnetic resonance”. en. In: *Biophysical Journal* 28.1, pp. 133–141 (cit. on pp. 28, 70).

Callaghan, PT and O Soderman (1983). “Examination of the lamellar phase of aerosol OT/water using pulsed field gradient nuclear magnetic resonance”. In: *The Journal of Physical Chemistry* 87.10, pp. 1737–1744 (cit. on p. 69).

Caruyer, E., C. Lenglet, G. Sapiro, and R. Deriche (2013). “Design of multishell sampling schemes with uniform coverage in diffusion MRI”. In: *Magnetic Resonance in Medicine* 69.6, pp. 1534–1540. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.24736 (cit. on p. 19).

Charnock, Tom, Guilhem Lavaux, and Benjamin D Wandelt (2018). “Automatic physical inference with information maximizing neural networks”. In: *Physical Review D* 97.8, p. 083004 (cit. on p. 59).

Cranmer, Kyle, Johann Brehmer, and Gilles Louppe (2020). “The frontier of simulation-based inference”. In: *Proceedings of the National Academy of Sciences* 117.48, pp. 30055–30062 (cit. on pp. 4, 5, 42, 48, 53, 92).

Cremer, Chris, Xuechen Li, and David Duvenaud (2018). “Inference Suboptimality in Variational Autoencoders”. In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. PMLR, pp. 1078–1086 (cit. on p. 57).

Dillon, Joshua V, Ian Langmore, Dustin Tran, et al. (2017). “Tensorflow distributions”. In: *arXiv preprint arXiv:1711.10604* (cit. on p. 46).

Dinh, Laurent, David Krueger, and Yoshua Bengio (2014). “Nice: Non-linear independent components estimation”. In: *arXiv preprint arXiv:1410.8516* (cit. on p. 56).

Dinh, Laurent, Jascha Sohl-Dickstein, and Samy Bengio (2016). “Density estimation using real nvp”. In: *arXiv preprint arXiv:1605.08803* (cit. on p. 56).

Douaud, Gwenaëlle, Ricarda A. L. Menke, Achim Gass, et al. (2013). “Brain Microstructure Reveals Early Abnormalities more than Two Years prior to Clinical Progression from Mild Cognitive Impairment to Alzheimer’s Disease”. In: *Journal of Neuroscience* 33.5, pp. 2147–2155 (cit. on pp. 1, 12).

Durkan, Conor, Artur Bekasov, Iain Murray, and George Papamakarios (Nov. 2020). “nflows: normalizing flows in PyTorch”. Version v0.14. In: (cit. on p. 100).

Eccles, John C (2013). *The cerebellum as a neuronal machine*. Springer Science & Business Media (cit. on p. 112).

Einstein, Albert et al. (1905). “On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat”. In: *Annalen der physik* 17.549-560, p. 208 (cit. on p. 13).

Fatemi, S Hossein, Kimberly A Aldinger, Paul Ashwood, et al. (2012). “Consensus paper: pathological role of the cerebellum in autism”. In: *The Cerebellum* 11.3, pp. 777–807 (cit. on p. 124).

Fearnhead, Paul and Dennis Prangle (2012). “Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 74.3, pp. 419–474 (cit. on pp. 59, 60).

Fick, R. H. J., M. Pizzolato, D. Wassermann, and R. Deriche (2017). “Diffusion MRI Anisotropy: Modeling, Analysis and Interpretation”. en. In: *Modeling, Analysis, and Visualization of Anisotropy*. Ed. by T. Schultz, E. Özarslan, and I. Hotz. Series Title: Mathematics and Visualization. Cham: Springer International Publishing, pp. 203–228 (cit. on p. 23).

Fick, R. H. J., D. Wassermann, and R. Deriche (Oct. 2019). “The Dmipy Toolbox: Diffusion MRI Multi-Compartment Modeling and Microstructure Recovery Made Easy”. en. In: *Frontiers in Neuroinformatics* 13 (cit. on pp. 98, 100).

Fick, R. H.J., D. Wassermann, E. Caruyer, and R. Deriche (July 2016). “MAPL: Tissue microstructure estimation using Laplacian-regularized MAP-MRI and its application to HCP data”. en. In: *NeuroImage* 134, pp. 365–385 (cit. on pp. 3, 66, 99).

Fields, R. D., D. H. Woo, and P. J. Basser (2015). “Glial Regulation of the Neuronal Connectome through Local and Long-Distant Communication”. In: *Neuron* 86.2, pp. 374 –386 (cit. on p. 27).

Fieremans, E., J. H. Jensen, and J. A. Helpern (2011). “White matter characterization with diffusional kurtosis imaging”. In: *NeuroImage* 58.1, pp. 177 –188 (cit. on pp. 31, 90).

Fieremans, Els, Lauren M. Burcaw, Hong-Hsi Lee, et al. (2016). “In vivo observation and biophysical interpretation of time-dependent diffusion in human white matter”. In: *NeuroImage* 129, pp. 414–427 (cit. on p. 19).

Fieremans, Els, Dmitry S. Novikov, Jens H. Jensen, and Joseph A. Helpern (Sept. 2010). “Monte Carlo study of a two-compartment exchange model of diffusion”. en. In: *NMR in Biomedicine* 23.7, pp. 711–724 (cit. on pp. 35, 44, 84).

Gallardo Diez, Guillermo Alejandro (2018). “Construction et comparaison de parcellisations structurelles cérébrale par imagerie de diffusion”. Thèse de doctorat dirigée par Deriche, Rachid et Wassermann, Demian Automatique et traitement du signal et des images Université Côte d’Azur (ComUE) 2018. PhD thesis (cit. on p. 17).

Garyfallidis, Eleftherios, Matthew Brett, Bagrat Amirbekian, et al. (2014). “Dipy, a library for the analysis of diffusion MRI data”. In: *Frontiers in Neuroinformatics* 8, p. 8 (cit. on p. 100).

Gelman, A., J.B. Carlin, H.S. Stern, et al. (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis (cit. on pp. 4, 92).

Germain, Mathieu, Karol Gregor, Iain Murray, and Hugo Larochelle (2015). “MADE: Masked Autoencoder for Distribution Estimation”. In: *Proceedings of the 32nd International Conference on Machine Learning*. Vol. 37. PMLR, pp. 881–889 (cit. on pp. 56, 98).

Geyer, Stefan, Axel Schleicher, and Karl Zilles (July 1999). “Areas 3a, 3b, and 1 of Human Primary Somatosensory Cortex”. en. In: *NeuroImage* 10.1, pp. 63–83 (cit. on pp. 100, 106, 107).

Golkov, Vladimir, Alexey Dosovitskiy, Jonathan I Sperl, et al. (2016). “Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans”. In: *IEEE transactions on medical imaging* 35.5, pp. 1344–1351 (cit. on pp. 115, 126).

Gonçalves, Pedro J, Jan-Matthis Lueckmann, Michael Deistler, et al. (Sept. 2020). “Training deep neural density estimators to identify mechanistic models of neural dynamics”. In: *eLife* 9 (cit. on pp. 5, 48, 53, 92).

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press (cit. on pp. 5, 48, 53, 92).

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, et al. (2014). *Generative Adversarial Networks*. arXiv: 1406.2661 [stat.ML] (cit. on pp. 53, 96).

Greenberg, David, Marcel Nonnenmacher, and Jakob Macke (2019). "Automatic Posterior Transformation for Likelihood-Free Inference". In: *Proceedings of the 36th International Conference on Machine Learning*. Vol. 97. PMLR, pp. 2404–2414 (cit. on pp. 5, 48, 53, 57, 92, 96, 98).

Hanyga, A. and R. L. Magin (2014). "A new anisotropic fractional model of diffusion suitable for applications of diffusion tensor imaging in biological tissues". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2170, p. 20140319. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.2014.0319 (cit. on p. 26).

Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, et al. (Sept. 2020). "Array programming with NumPy". In: *Nature* 585.7825, pp. 357–362 (cit. on p. 100).

Henriques, Rafael Neto, Sune Nørhøj Jespersen, and Noam Shemesh (May 2020). "Correlation tensor magnetic resonance imaging". en. In: *NeuroImage* 211, p. 116605 (cit. on p. 25).

Hofstetter, S., I. Tavor, S. Tzur Moryosef, and Y. Assaf (2013). "Short-Term Learning Induces White Matter Plasticity in the Fornix". In: *Journal of Neuroscience* 33.31, pp. 12844–12850. eprint: https://www.jneurosci.org/content/33/31/12844.full.pdf (cit. on p. 23).

Huang, Chin-Wei, David Krueger, Alexandre Lacoste, and Aaron Courville (2018). "Neural autoregressive flows". In: *International Conference on Machine Learning*. PMLR, pp. 2078–2087 (cit. on p. 55).

Hutchinson, E. B., A. V. Avram, M. O. Irfanoglu, et al. (2017). "Analysis of the effects of noise, DWI sampling, and value of assumed parameters in diffusion MRI models". In: *Magnetic Resonance in Medicine* 78.5, pp. 1767–1780. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.26575 (cit. on p. 26).

Jaini, Priyank, Kira A Selby, and Yaoliang Yu (2019). "Sum-of-squares polynomial flow". In: *International Conference on Machine Learning*. PMLR, pp. 3009–3018 (cit. on p. 55).

Jelescu, I. O. and M. D. Budde (Nov. 2017a). "Design and Validation of Diffusion MRI Models of White Matter". en. In: *Frontiers in Physics* 5 (cit. on pp. 24, 33).

Jelescu, I. O., J. Veraart, V. Adisetiyo, et al. (2015). "One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI?" In: *NeuroImage* 107, pp. 242 –256 (cit. on p. 31).

Jelescu, I. O., J. Veraart, E. Fieremans, and D. S. Novikov (Jan. 2016). "Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue: Degeneracy in Model Parameter Estimation of Diffusion in Neural Tissue". en. In: *NMR in Biomedicine* 29.1, pp. 33–47 (cit. on pp. 3, 4, 6, 31, 90, 91, 93, 126).

Jelescu, Ileana O. and Matthew D. Budde (Nov. 2017b). "Design and Validation of Diffusion MRI Models of White Matter". In: *Frontiers in Physics* 5 (cit. on p. 69).

Jelescu, Ileana O., Alexandre de Skowronski, Marco Palombo, and Dmitry S. Novikov (2021). *Neurite Exchange Imaging (NEXI): A minimal model of diffusion in gray matter with inter-compartment water exchange*. arXiv: 2108.06121 [physics.med-ph] (cit. on pp. 3, 6, 15, 34, 35, 66–68, 76, 77, 79, 81, 82, 86, 90, 113, 123, 124).

Jensen, J. H., G. Russell Glenn, and J. A. Helpern (2016). "Fiber ball imaging". In: *NeuroImage* 124, pp. 824 –833 (cit. on p. 28).

Jensen, J. H., J. A. Helpern, A. Ramani, H. Lu, and K. Kaczynski (2005). "Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging". In: *Magnetic Resonance in Medicine* 53.6, pp. 1432–1440. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.20508 (cit. on pp. 22, 24).

Jespersen, S. N., C. D. Kroenke, L. Østergaard, J. J.H. Ackerman, and D. A. Yablonskiy (2007). "Modeling dendrite density from magnetic resonance diffusion measurements". In: *NeuroImage* 34.4, pp. 1473 –1486 (cit. on pp. 19, 28).

Jespersen, S. N., J. L. Olesen, B. Hansen, and N. Shemesh (2018). "Diffusion time dependence of microstructural parameters in fixed spinal cord". In: *NeuroImage* 182. Microstructural Imaging, pp. 329 –342 (cit. on p. 32).

Jeurissen, B., M. Descoteaux, S. Mori, and A. Leemans (2019). "Diffusion MRI fiber tractography of the brain". In: *NMR in Biomedicine* 32.4. e3785 NBM-17-0045.R2, e3785. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.3785 (cit. on p. 13).

Jian, B., B. C. Vemuri, E. Özarslan, P. R. Carney, and T. H. Mareci (2007). "A novel tensor distribution model for the diffusion-weighted MR signal". In: *NeuroImage* 37.1, pp. 164 –176 (cit. on p. 25).

Joabsson, F., M. Nydén, P. Linse, and O. Söderman (1997). "Pulsed Field Gradient NMR Studies of Translational Diffusion in Cylindrical Surfactant Aggregates". In: *The Journal of Physical Chemistry B* 101.47, pp. 9710–9716. eprint: https://doi.org/10.1021/jp971890g (cit. on p. 28).

Jones, DK. (2010). "Precision and accuracy in diffusion tensor magnetic resonance imaging". In: *Top Magn Reson Imaging* 21, pp. 87–99 (cit. on p. 24).

Jones, D.K., M.A. Horsfield, and A. Simmons (1999). "Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging". In: *Magnetic Resonance in Medicine* 42.3, pp. 515–525 (cit. on p. 25).

Joyce, Paul and Paul Marjoram (2008). "Approximately sufficient statistics and Bayesian computation". In: *Statistical applications in genetics and molecular biology* 7.1 (cit. on p. 59).

Kac, M. (1966). "Can One Hear the Shape of a Drum?" In: *The American Mathematical Monthly* 73.4P2, pp. 1–23. eprint: https://doi.org/10.1080/00029890.1966.11970915 (cit. on p. 13).

Kaden, E., F. Kruggel, and D. C. Alexander (2016). "Quantitative mapping of the per-axon diffusion coefficients in brain white matter". In: *Magnetic Resonance in Medicine* 75.4, pp. 1752–1763. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.25734 (cit. on p. 24).

Kärger, Jörg (1985). "NMR self-diffusion studies in heterogeneous systems". In: *Advances in Colloid and Interface Science* 23, pp. 129–148 (cit. on p. 35).

Kingma, Diederik P. and Jimmy Ba (2017). *Adam: A Method for Stochastic Optimization*. arXiv: 1412.6980 [cs.LG] (cit. on p. 98).

Kingma, Diederik P and Max Welling (2014). *Auto-Encoding Variational Bayes*. arXiv: 1312.6114 [stat.ML] (cit. on pp. 53, 96).

Kingma, Durk P, Tim Salimans, Rafal Jozefowicz, et al. (2016). "Improved variational inference with inverse autoregressive flow". In: *Advances in neural information processing systems* 29, pp. 4743–4751 (cit. on p. 56).

Kiselev, V. G. (2017). "Fundamentals of diffusion MRI physics". In: *NMR in Biomedicine* 30.3. e3602 nbm.3602, e3602. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.3602 (cit. on p. 25).

Kiselev, V. G. and K. A. Il'yasov (2007). "Is the "biexponential diffusion" biexponential?" In: *Magnetic Resonance in Medicine* 57.3, pp. 464–469. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.21164 (cit. on p. 26).

Kiselev, VG (2010). "The cumulant expansion: an overarching mathematical framework for understanding diffusion NMR". In: *Diffusion MRI: Theory, Methods, and Applications*. Ed. by Jones DK. Oxford University Press, pp. 152 –68 (cit. on p. 22).

Kroenke, C. D., J. J.H. Ackerman, and D. A. Yablonskiy (2004). "On the nature of the NAA diffusion attenuated MR signal in the central nervous system". In: *Magnetic Resonance in Medicine* 52.5, pp. 1052–1059. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.20260 (cit. on p. 28).

Kucukelbir, Alp, Dustin Tran, Rajesh Ranganath, Andrew Gelman, and David M Blei (2017). "Automatic differentiation variational inference". In: *The Journal of Machine Learning Research* 18.1, pp. 430–474 (cit. on p. 46).

Le Bihan, D. and E. Breton (Dec. 1985). "Imagerie de diffusion in-vivo par résonance magnétique nucléaire". In: *Comptes-Rendus de l'Académie des Sciences* 93.5, pp. 27–34 (cit. on p. 16).

Lee, Hong-Hsi, Antonios Papaioannou, Dmitry S. Novikov, and Els Fieremans (2020). "In vivo observation and biophysical interpretation of time-dependent diffusion in human cortical gray matter". In: *NeuroImage* 222, p. 117054 (cit. on pp. 3, 34, 66, 124).

Li, Hua, Xiaoyu Jiang, Jingping Xie, John C. Gore, and Junzhong Xu (June 2016). "Impact of transcytolemmal water exchange on estimates of tissue microstructural properties derived from diffusion MRI". In: *Magnetic Resonance in Medicine* 77.6, pp. 2239–2249 (cit. on p. 97).

Ligneul, C., M. Palombo, E. Hernández-Garzón, et al. (May 2019). "Diffusion-weighted magnetic resonance spectroscopy enables cell-specific monitoring of astrocyte reactivity in vivo". en. In: *NeuroImage* 191, pp. 457–469 (cit. on pp. 15, 19).

Marjoram, Paul, John Molitor, Vincent Plagnol, and Simon Tavaré (2003). "Markov chain Monte Carlo without likelihoods". In: *Proceedings of the National Academy of Sciences* 100.26, pp. 15324–15328 (cit. on p. 50).

Marr, David and W Thomas Thach (1991). "A theory of cerebellar cortex". In: *From the Retina to the Neocortex*. Springer, pp. 11–50 (cit. on p. 112).

Mayer, J. E. and E. Montroll (1941). "Molecular Distribution". In: *The Journal of Chemical Physics* 9.1, pp. 2–16. eprint: https://doi.org/10.1063/1.1750822 (cit. on p. 22).

McKinnon, E. T., J. H. Jensen, G. R. Glenn, and J. A. Helpern (Feb. 2017). “Dependence on b-value of the direction-averaged diffusion-weighted imaging signal in brain”. en. In: *Magnetic Resonance Imaging* 36, pp. 121–127 (cit. on pp. 28, 34).

Menon, Vinod, Guillermo Gallardo, Mark A Pinsk, et al. (2020). “Microstructural organization of human insula is linked to its macrofunctional circuitry and predicts cognitive control”. In: *elife* 9, e53470 (cit. on pp. 74, 94).

Mitra, Partha P, LL Latour, Robert L Kleinberg, and Christopher H Sotak (1995). “Pulsed-field-gradient NMR measurements of restricted diffusion and the return-to-the-origin probability”. In: *Journal of Magnetic Resonance, Series A* 114.1, pp. 47–58 (cit. on p. 72).

Monica Moore, MSG, Mirella Díaz-Santos, and Keith Vossel (2021). “Alzheimer’s Association 2021 Facts and Figures Report”. In: (cit. on p. 1).

Mori, Susumu and Jiangyang Zhang (2006). “Principles of diffusion tensor imaging and its applications to basic neuroscience research”. In: *Neuron* 51.5, pp. 527–539 (cit. on p. 2).

Murday, J S and R M Cotts (1968). “Self-Diffusion Coefficient of Liquid Lithium”. en. In: *The Journal of Chemical Physics* 48, p. 8 (cit. on pp. 29, 71).

Ning, L., F. Laun, Y. Gur, et al. (2015). “Sparse Reconstruction Challenge for diffusion MRI: Validation on a physical phantom to determine which acquisition scheme and analysis method to use?” In: *Medical Image Analysis* 26.1, pp. 316 –331 (cit. on pp. 19, 20, 26).

Novikov, D. S and E. Fieremans (2012). “Relating extracellular diffusivity to cell size distribution and packing density as applied to white matter”. In: *Proceedings of the 20th Annual Meeting of ISMRM*, p. 1829 (cit. on p. 31).

Novikov, D. S., E. Fieremans, S. N. Jespersen, and V. G. Kiselev (Oct. 2018a). “Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation: Brain microstructure with dMRI: Theory and parameter estimation”. en. In: *NMR in Biomedicine*, e3998 (cit. on pp. 2, 27, 28, 30, 66, 69, 71, 90, 102, 112, 119).

Novikov, D. S., J. H. Jensen, J. A. Helpern, and E. Fieremans (Apr. 2014). “Revealing mesoscopic structural universality with diffusion”. en. In: *Proceedings of the National Academy of Sciences* 111.14, pp. 5088–5093 (cit. on p. 34).

Novikov, D. S. and V. G. Kiselev (2010). “Effective medium theory of a diffusion-weighted signal”. In: *NMR in Biomedicine* 23.7, pp. 682–697. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nbm.1584 (cit. on p. 25).

Novikov, D. S., V. G. Kiselev, and S. N. Jespersen (June 2018b). “On modeling”. en. In: *Magnetic Resonance in Medicine* 79.6, pp. 3172–3193 (cit. on pp. 2, 20, 124).

Novikov, D. S., J. Veraart, I. O. Jelescu, and E. Fieremans (July 2018c). “Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI”. en. In: *NeuroImage* 174, pp. 518–538 (cit. on pp. 3, 4, 6, 22, 30, 31, 33, 68, 72, 73, 90, 91, 93, 112, 119).

Novikov, Dmitry S, Els Fieremans, Jens H Jensen, and Joseph A Helpern (2011). “Random walks with barriers”. In: *Nature physics* 7.6, pp. 508–514 (cit. on p. 44).

Nowozin, Sebastian (2015). *Effective Sample Sizein Importance Sampling*. http://www.nowozin.net/sebastian/blog/effective-sample-size-in-importance-sampling.html (cit. on p. 53).

Nunes, Matthew A and David J Balding (2010). “On optimal selection of summary statistics for approximate Bayesian computation”. In: *Statistical applications in genetics and molecular biology* 9.1 (cit. on p. 60).

Olesen, Jonas L, Leif Østergaard, Noam Shemesh, and Sune N Jespersen (2021). “Diffusion time dependence, power-law scaling, and exchange in gray matter”. In: *arXiv preprint arXiv:2108.09983* (cit. on pp. 3, 6, 35, 66–68, 76, 79, 81, 86, 90, 113, 117–119, 121–124).

Oliva, Junier, Avinava Dubey, Manzil Zaheer, et al. (2018). “Transformation autoregressive networks”. In: *International Conference on Machine Learning*. PMLR, pp. 3898–3907 (cit. on p. 56).

Özarslan, E., C. G. Koay, T. M. Shepherd, et al. (2013). “Mean apparent propagator (MAP) MRI: A novel diffusion imaging method for mapping tissue microstructure”. In: *NeuroImage* 78, pp. 16 –32 (cit. on pp. 23, 26).

Özarslan, E., C. Yolcu, M. Herberthson, H. Knutsson, and C.-F. Westin (2018). “Influence of the Size and Curvedness of Neural Projections on the Orientationally Averaged Diffusion MR Signal”. In: *Frontiers in Physics* 6, p. 17 (cit. on p. 34).

Palombo, M., D. C. Alexander, and H. Zhang (Mar. 2019). “A generative model of realistic brain cells with application to numerical simulation of the diffusion-weighted MR signal”. en. In: *NeuroImage* 188, pp. 391–402 (cit. on p. 34).

Palombo, M, DC Alexander, and H Zhang (2021). “Large-scale analysis of brain cell morphometry informs microstructure modelling of gray matter”. In: (cit. on pp. 13, 97, 98, 102, 108).

Palombo, M., A. Ianus, M. Guerreri, et al. (2020). “SANDI: A compartment-based model for non-invasive apparent soma and neurite imaging by diffusion MRI”. In: *NeuroImage* 215, p. 116835 (cit. on pp. 3, 4, 6, 15, 34, 66, 67, 69, 71, 80, 90, 113, 123, 124).

Palombo, M., C. Ligneul, C. Najac, et al. (2016). “New paradigm to assess brain cell morphology by diffusion-weighted MR spectroscopy in vivo”. In: *Proceedings of the National Academy of Sciences* 113.24, pp. 6671–6676. eprint: https://www.pnas.org/content/113/24/6671.full.pdf (cit. on pp. 19, 27).

Palombo, M., C. Ligneul, and J. Valette (Jan. 2017). “Modeling diffusion of intracellular metabolites in the mouse brain up to very high diffusion-weighting: Diffusion in long fibers (almost) accounts for non-monoexponential attenuation: Modeling Diffusion of Brain Metabolites in Vivo up to Very High Diffusion Weighting”. en. In: *Magnetic Resonance in Medicine* 77.1, pp. 343–350 (cit. on p. 31).

Panagiotaki, E., T. Schneider, B. Siow, et al. (2012). “Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison”. In: *NeuroImage* 59.3, pp. 2241 –2254 (cit. on p. 20).

Panagiotaki, E., S. Walker-Samuel, B. Siow, et al. (2014). “Noninvasive Quantification of Solid Tumor Microstructure Using VERDICT MRI”. In: *Cancer Research* 74.7, pp. 1902–1912. eprint: https://cancerres.aacrjournals.org/content/74/7/1902.full.pdf (cit. on p. 12).

Pandya, Deepak, Michael Petrides, and Patsy Benny Cipolloni (2015). *Cerebral cortex: architecture, connections, and the dual origin concept*. Oxford University Press (cit. on p. 106).

Papamakarios, George (2019). “Neural density estimation and likelihood-free inference”. In: *arXiv preprint arXiv:1910.13233* (cit. on pp. 43, 49–51).

Papamakarios, George and Iain Murray (2016). "Fast ϵ-free Inference of Simulation Models with Bayesian Conditional Density Estimation". In: *Advances in Neural Information Processing Systems*. Vol. 29. Curran Associates, Inc., pp. 1028–1036 (cit. on pp. 57, 95).

Papamakarios, George, Eric T. Nalisnick, Danilo Jimenez Rezende, Shakir Mohamed, and Balaji Lakshminarayanan (2019). "Normalizing Flows for Probabilistic Modeling and Inference". In: *ArXiv* abs/1912.02762 (cit. on pp. 5, 48, 53–55, 96).

Paszke, Adam, Sam Gross, Francisco Massa, et al. (2019). "PyTorch: An Imperative Style, High-Performance Deep Learning Library". en. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vancouver, BC, Canada, p. 12 (cit. on pp. 46, 100).

Prangle, Dennis (2017). "Adapting the ABC distance function". In: *Bayesian Analysis* 12.1, pp. 289–309 (cit. on p. 49).

Prangle, Dennis, Paul Fearnhead, Murray P Cox, Patrick J Biggs, and Nigel P French (2014). "Semi-automatic selection of summary statistics for ABC model choice". In: *Statistical applications in genetics and molecular biology* 13.1, pp. 67–82 (cit. on p. 59).

Pritchard, Jonathan K, Mark T Seielstad, Anna Perez-Lezaun, and Marcus W Feldman (1999). "Population growth of human Y chromosomes: a study of Y chromosome microsatellites." In: *Molecular biology and evolution* 16.12, pp. 1791–1798 (cit. on p. 49).

Purves D., Augustine G. J. Fitzpatrick D. Hall W. C. LaMantia A.-S. McNamara J. O. Williams S. M. (2004). *Neuroscience (3rd ed.)* Sinauer Associates (cit. on pp. 2, 14).

Python Software Fundation (2017). *Python Language Reference, Version 3.6* (cit. on p. 100).

Ramachandran, P. and G. Varoquaux (2011). "Mayavi: 3D Visualization of Scientific Data". In: *Computing in Science & Engineering* 13.2, pp. 40–51 (cit. on p. 100).

Robert, Christian P. and George Casella (2004). *Monte Carlo statistical methods*. Vol. 2. Springer (cit. on pp. 44, 47).

Setsompop, K., R. Kimmlingen, E. Eberlein, et al. (2013). "Pushing the limits of in vivo diffusion MRI for the Human Connectome Project". In: *NeuroImage* 80, pp. 220 –233 (cit. on p. 99).

Shapson-Coe, Alexander, Michał Januszewski, Daniel R. Berger, et al. (2021). "A connectomic study of a petascale fragment of human cerebral cortex". In: *bioRxiv*. eprint: https://www.biorxiv.org/content/early/2021/05/30/2021.05.29.446289.full.pdf (cit. on pp. 3, 34, 66, 69, 102, 108, 112).

Sisson, Scott A. (Sept. 2018). *Handbook of Approximate Bayesian Computation*. Chapman and Hall/CRC (cit. on pp. 5, 48, 53, 92).

Soares, J., P. Marques, V. Alves, and N. Sousa (2013). "A hitchhiker's guide to diffusion tensor imaging". In: *Frontiers in Neuroscience* 7, p. 31 (cit. on p. 12).

Strimbu, Kyle and Jorge A Tavel (2010). "What are biomarkers?" In: *Current Opinion in HIV and AIDS* 5.6, p. 463 (cit. on p. 23).

Stuart, A. M. (2010). "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19, 451–559 (cit. on pp. 4, 92).

Tanner, J. E. and E. O. Stejskal (1968). "Restricted Self-Diffusion of Protons in Colloidal Systems by the Pulsed-Gradient, Spin-Echo Method". In: *The Journal of Chemical Physics* 49.4, pp. 1768–1777. eprint: https://doi.org/10.1063/1.1670306 (cit. on pp. 16, 18).

Tax, Chantal M.W., Filip Szczepankiewicz, Markus Nilsson, and Derek K. Jones (2020). "The dot-compartment revealed? Diffusion MRI with ultra-strong gradients and spherical tensor encoding in the living human brain". In: *NeuroImage* 210, p. 116534 (cit. on p. 69).

Tejero-Cantero, Alvaro, Jan Boelts, Michael Deistler, et al. (2020). "sbi: A toolkit for simulation-based inference". In: *Journal of Open Source Software* 5.52, p. 2505 (cit. on p. 100).

Trepel, Martin (2021). *neuroanatomie: Struktur und Funktion*. Elsevier Health Sciences (cit. on p. 113).

Tuch, D. S. (2004). "Q-ball imaging". In: *Magnetic Resonance in Medicine* 52.6, pp. 1358–1372. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.20279 (cit. on pp. 23, 27).

Tuch, D. S., T. G. Reese, M. R. Wiegell, et al. (2002). "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity". In: *Magnetic Resonance in Medicine* 48.4, pp. 577–582. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.10268 (cit. on p. 18).

Veraart, J., E. Fieremans, and D. S. Novikov (Jan. 2019). "On the scaling behavior of water diffusion in human brain white matter". en. In: *NeuroImage* 185, pp. 379–387 (cit. on pp. 28, 69, 70).

Veraart, J., D. Nunes, U. Rudrapatna, et al. (2020). "Noninvasive quantification of axon radii using diffusion MRI". In: *eLife* 9. Ed. by Floris P de Lange, Birte Forstmann, Birte Forstmann, Saad Jbabdi, and Robert Mulkern, e49855 (cit. on pp. 2, 3, 13, 29, 34, 66, 76, 79, 112).

Veraart, J., D. H. J. Poot, W. Van Hecke, et al. (2011). "More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging". In: *Magnetic Resonance in Medicine* 65.1, pp. 138–145. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.22603 (cit. on p. 24).

Vincent, Mélissa, Mylène Gaudin, Covadonga Lucas-Torres, et al. (2021). "Characterizing extracellular diffusion properties using diffusion-weighted MRS of sucrose injected in mouse brain". In: *NMR in Biomedicine* 34.4, e4478 (cit. on pp. 71, 74, 83, 85, 94, 113, 124).

Walinga, J. (2010). *Introduction to Psychology: 1st Canadian Edition*. Online access: Center for Open Education Open Textbook Library. BCcampus (cit. on p. 15).

Wasserman, Larry (2004). *All of statistics: a concise course in statistical inference*. Vol. 26. Springer (cit. on p. 50).

Wegmann, Daniel, Christoph Leuenberger, and Laurent Excoffier (2009). "Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood". In: *Genetics* 182.4, pp. 1207–1218 (cit. on p. 60).

Wehenkel, Antoine and Gilles Louppe (2019). "Unconstrained monotonic neural networks". In: *Advances in Neural Information Processing Systems* 32, pp. 1545–1555 (cit. on p. 55).

Whitehouse, Peter J, Donald L Price, Robert G Struble, et al. (1982). "Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain". In: *Science* 215.4537, pp. 1237–1239 (cit. on p. 124).

Wood, Simon N (2010). “Statistical inference for noisy nonlinear ecological dynamic systems”. In: *Nature* 466.7310, pp. 1102–1104 (cit. on p. 43).

Yablonskiy, D. A., G. L. Bretthorst, and J. J.H. Ackerman (2003). “Statistical model for diffusion attenuated MR signal”. In: *Magnetic Resonance in Medicine* 50.4, pp. 664–669. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.10578 (cit. on p. 25).

Zhang, H., T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander (July 2012). “NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain”. en. In: *NeuroImage* 61.4, pp. 1000–1016 (cit. on pp. 2, 4, 6, 24, 30, 66, 90).

Zilles, Karl (Oct. 2018). “Brodmann: a pioneer of human brain mapping—his impact on concepts of cortical organization”. In: *Brain* 141.11, pp. 3262–3278 (cit. on p. 100).

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Extension de la mesure de la microstructure par IRM de diffusion aux cellules corticales via inférence par simulation

Mots clés : IRM de diffusion, Problèmes inverses, Machine Learning, Neurosciences

Résumé : Les maladies neurodégénératives, telles que la maladie d'Alzheimer ou de Huntington, entraînent la perte progressive et irréversible des fonctions mentales. La démence et les déficits cognitifs semblent être grandement liés à une perte neuronale. Bien que l'impact et l'évolution extérieurs de ces maladies soient facilement observables, l'accès aux changements microstructuraux dans le cerveau reste un défi, rendant difficiles la compréhension de ces maladies et la mise au point de traitements.

Avec les avancées technologiques, l'Imagerie par Résonance Magnétique de diffusion (IRMd) s'impose comme une méthode novatrice pour étudier la microstructure du cerveau de manière non-invasive et in-vivo. Cette technique d'imagerie médicale est basée sur l'étude des mouvements microscopiques aléatoires des molécules d'eau, connus sous le nom de mouvements Browniens. Dans le cerveau, les mouvements des molécules sont contraints par les membranes des cellules, rendant la diffusion anisotropique. Chaque composant tissulaire, tel que les somas (corps des neurones) ou les neurites, possède une forme distincte. Le signal de diffusion du cerveau obtenu lors d'une acquisition IRM est ainsi modulé selon les caractéristiques du tissu.

L'objectif de ma thèse est de mettre en place une méthode qui permette d'inférer la microstructure d'un tissu à partir d'une acquisition d'IRM de diffusion dans la matière grise (MG).

La résolution de ce problème inverse d'estimation de la microstructure du cerveau à partir de l'IRMd s'organise autour de trois parties :

1. La définition d'un modèle biologique décrivant les tissus de la MG. Il a été prouvé que les modèles microstructuraux existants modélisant la matière blanche ne sont pas valides dans la MG. Nous avons commencé par adapter ces modèles en prenant en compte l'abondance des somas dans la MG.

2. Une modélisation mathématique de la MG. Chaque compartiment du modèle tissulaire a ensuite été modélisé par des formes géométriques simples, pour lesquelles le signal de diffusion est connu. Un algorithme de traitement du signal a été développé permettant de synthétiser les informations clés contenues dans le signal de diffusion et de les mettre en relation avec un ensemble de paramètres décrivant le tissu (notamment la taille et la densité des neurones). Cet algorithme se base sur une étude des moments statistiques du signal pour différentes puissances de gradient de l'IRM. À la différence des méthodes existantes, aucun paramètre biologique n'est arbitrairement fixé, ce qui permet de décrire au mieux les tissus corticaux de chaque sujet.

3. Un algorithme d'inversion permettant d'estimer les paramètres du tissu ayant généré le signal d'acquisition. Une fois le modèle mathématique permettant de relier les paramètres du tissu au signal de diffusion défini, l'objectif est de résoudre le problème inverse d'estimation de la microstructure du tissu à partir d'une observation. Une limitation des méthodes actuelles est leur incapacité à identifier toutes les configurations du tissu possibles pouvant expliquer un même signal de diffusion observé, ce qui rend l'interprétation des estimations proposées difficile. Afin de résoudre ce problème, nous avons utilisé une méthode reposant sur des outils de l'analyse bayésienne et de deep learning appelée "likelihood-based inference", combinée à des réseaux de neurones. Celle-ci permet d'identifier et de retourner toutes les configurations possibles du tissu accompagnées de leurs distributions postérieures (probabilité étant donné une observation), ce qui facilite leur interprétation.

L'approche a tout d'abord été validée sur des simulations. Reposant sur peu de contraintes d'acquisition, la méthode de résolution globale a ensuite été appliquée sur les bases de données HCP MGH et HCP1200 du Human Connectome Project. Une bibliothèque python a été développée pour étudier ces données réelles ou simulées. Les résultats obtenus ont enfin été comparés avec des études histologiques et cognitives pour vérifier leur validité.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Title : Enabling Cortical Cell-Specific Sensitivity on Diffusion MRI Microstructure Measurements using Likelihood-Free Inference

Keywords : Diffusion MRI, Inverse problems, Machine Learning, Neuroscience

Abstract : Neurodegenerative diseases, such as Alzheimer's or Huntington's disease, lead to the progressive and irreversible loss of mental functions. Dementia and cognitive deficits appear to be primarily related to neuronal and synaptic connectivity loss. Although these diseases' external impact and progression are readily observable, accessing microstructural changes in the brain remains a challenge, making it difficult to understand these diseases and develop treatments.

With technological advances, diffusion Magnetic Resonance Imaging (dMRI) has emerged as a novel method to study the microstructure of the brain non-invasively and in-vivo. This medical imaging technique is based on the study of random microscopic movements of water molecules, known as Brownian movements. In the brain, the movements of the molecules are constrained by cell membranes, making diffusion anisotropic. Each tissue component, such as somas (cell bodies) or neurites, has a distinct shape. The characteristics of the tissue thus modulate the diffusion brain signal obtained during an MRI acquisition.

My thesis aims to develop a method to infer a tissue microstructure from a dMRI acquisition in the gray matter (GM).

The solution to this inverse problem of estimating brain microstructure from dMRI is threefold :

1. **The definition of a biological model describing the GM tissues.** Existing microstructural models of white matter were proven not to hold in the GM. We adapted these models to take into account the abundance of somas in the GM.
2. **A mathematical modeling of the GM tissue.** We modeled each compartment of the tissue model by simple geometrical shapes, for which the diffusion signal is known. We developed a signal processing algorithm to synthesize the key information contained in the diffusion signal and relate it to a set of parameters describing the tissue (notably the size and density of neurons). This algorithm is based on a study of the statistical moments of the signal at different MRI gradient strengths. Unlike existing methods, no biological parameters are arbitrarily fixed, which allows for the best possible description of the cortical tissue of each subject.
3. **An inversion algorithm to estimate the tissue parameters that generated the acquisition signal.** Once the mathematical model relating tissue parameters to the diffusion signal is defined, the objective is to solve the inverse problem of estimating tissue microstructure from an observation. A limitation of current methods is their inability to identify all possible tissue configurations that can explain the same observed diffusion signal, making the interpretation of the proposed estimates difficult. We used a Bayesian deep-learning method called "likelihood-based inference" combined with neural networks to solve this problem. This method allows identifying and returning all possible tissue configurations along with their posterior distributions (probability given an observation), facilitating their interpretation.

We first validated this approach on simulations. Based on a few acquisition constraints, we then applied the global resolution method to the HCP MGH and HCP1200 databases of the Human Connectome Project. We developed a python library to study those simulated or acquired data. The obtained results were then compared with histological and cognitive studies to verify their validity.