In vivo diffusion magnetic resonance imaging of the white matter microstructure from dictionaries generated by Monte Carlo simulations: development and validation

Thèse N° 8497 (EPFL) et N° 737 (UCLouvain)

Présentée le 11 octobre 2019

À l'École polytechnique fédérale de Lausanne
à la Faculté des sciences et techniques de l'ingénieur
Laboratoire de traitement des signaux 5
Programme doctoral en génie électrique

et

À l'Université catholique de Louvain
Institut des technologies de l'information et de la communication,
électronique et mathématiques appliquées (ICTEAM)
École doctorale sciences de l'ingénieur et technologie

pour l'obtention du grade de Docteur ès Sciences

par

Gaëtan Olivier D. RENSONNET

Acceptée sur proposition du jury

Dr J.-M. Vesin, président du jury
Prof. J.-Ph. Thiran, Prof. B. Macq, directeurs de thèse
Dr M. Taquet, rapporteur
Prof. B. Scherrer, rapporteur
Dr I. Jelescu, rapporteuse

2019

UCLouvain

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Abstract

Non-invasive virtual histology of white matter tissues is the ultimate promise of diffusion-weighted MRI (DW-MRI). This hope is fueled by the exquisite sensitivity of DW-MRI to micrometer-scale displacements of diffusing water molecules, averaged over a whole voxel. Combined with adequate mathematical modeling, DW-MRI therefore has the potential to provide unprecedented, microscope-like insight into the microstructure of the brain, with important consequences in neuroscience, neurology and psychiatry.

Many closed-form analytical models have been proposed to relate the DW-MRI signal to morphological characteristics of the axons and glial cells of the white matter. These models generally make assumptions about the tissue and the diffusion processes which often depart from the biophysical reality, limiting their reliability and interpretability in practice. Monte Carlo simulations of the random walk of water molecules are widely recognized to provide near numerical groundtruth for DW-MRI signals. However, they have mostly been limited to the validation of simpler models rather than used for the estimation of microstructural properties.

This thesis proposes a general framework which leverages Monte Carlo simulations for the estimation of physically interpretable microstructural parameters such as indices of axon diameter and density, both in single and in crossing fascicles of axons, with no restriction on the data acquisition protocol. Monte Carlo simulations of DW-MRI signals, or *fingerprints*, are pre-computed for a large collection of microstructural configurations. At every voxel, the microstructural parameters are estimated by finding a sparse optimal combination of these fingerprints. The final complexity of the model is thus solely determined by the level of tissue detail incorporated in the Monte Carlo simulations. The parameter estimation requires no meta-parameter tuning.

The superposition approximation for DW-MRI signals in the presence of multiple fascicles of axons, often taken for granted in state-of-the-art approaches, was thoroughly verified. This enabled a dramatic reduction of the size of the dictionary and of the ensuing fingerprint matching. Our approach was then validated extensively on synthetic data as well as on a variety of *in vivo* and *ex vivo* datasets. It was shown to systematically provide more robust and physically interpretable tissue parameters than a range of traditional closed-form models claiming similar biophysical complexity.

Abstract

An accelerated method based on efficient convex estimation and a deep feed-forward neural network was developed to make our framework suitable for larger population studies. This fast two-stage procedure reduced the execution time by several orders of magnitude while maintaining a similar level of estimation accuracy.

All software tools related to this work will be shared upon publication of this thesis. They can easily be combined with, for instance, recently-released Monte Carlo simulators offering exciting new levels of tissue realism. This is intended to make microstructure imaging based on Monte Carlo simulations accessible to as broad an audience as possible, for future population studies and the general advancement of the field.

Résumé

La grande promesse de l'imagerie par résonance magnétique pondérée en diffusion (DW-MRI) est de fournir un outil d'histologie virtuelle non-invasive des tissus de la matière blanche. Cet espoir est entretenu par la très fine sensibilité du signal DW-MRI à des déplacements de l'ordre du micromètre des molécules d'eau en diffusion, moyennés sur tout le voxel. Combinée à des modèles mathématiques adéquats, la DW-MRI pourrait fournir des informations sur la microstructure cérébrale d'un niveau digne du microscope, avec d'importantes conséquences en neurosciences, en neurologie et en psychiatrie.

Parmi tous les modèles analytiques proposés pour lier le signal DW-MRI à des caractéristiques morphologiques des axones et des cellules gliales de la matière blanche, beaucoup reposent sur des hypothèses simplificatrices sur les tissus et les processus de diffusion qui sont contraires à la biophysique et limitent la fiabilité et l'interprétabilité de ces modèles en pratique. Les simulations Monte Carlo de diffusion des molécules d'eau sont largement reconnues comme le standard de référence pour les signaux DW-MRI. Cependant, elles ont principalement été utilisées pour la validation de modèles simples et non pour l'estimation directe de propriétés microstructurelles.

Cette thèse propose un cadre général qui tire parti de simulations Monte Carlo pour l'estimation de paramètres microstructurels interprétables physiquement tels que des indices de diamètre et densité axonaux, dans des faisceaux d'axones simples et croisés, sans restriction sur le protocole d'acquisition de données. Les signaux DW-MRI Monte Carlo, ou fingerprints, sont pré-simulés pour une large collection de configurations microstructurelles. En chaque voxel, les paramètres de microstructure sont estimés en cherchant une combinaison optimale parcimonieuse de ces fingerprints. La complexité finale du modèle est donc déterminée par le niveau de détails du tissu représenté dans les simulations. L'estimation de paramètres ne requiert l'ajustement d'aucun méta-paramètre.

L'approximation de superposition des signaux DW-MRI en présence de faisceaux d'axones multiples, souvent négligée dans la littérature, a été vérifiée en détail. Cela a permis de diminuer drastiquement la taille des dictionnaires et d'alléger la phase de reconnaissance de fingerprint. Notre approche a ensuite été validée extensivement sur des données synthétiques et sur des bases de données diverses, à la fois *in vivo* et *ex vivo*. Elle s'est montrée capable de fournir des paramètres des tissus plus interprétables et robustes que ceux de nombreux mo-

dèles traditionnels basés sur formules analytiques et d'un niveau de complexité biophysique présenté comme similaire.

Une méthode accélérée basée sur de l'optimisation convexe rapide et un réseau de neurones profond a été développée pour faire face à des études de plus grande ampleur. Cette procédure en deux temps a pu réduire le temps d'exécution de notre méthode de plusieurs ordres de grandeur tout en maintenant la qualité des estimations. Tous les outils logiciels développés dans ce travail seront partagés publiquement après publication de cette thèse. Ceux-ci s'intégreront par exemple facilement à de récents simulateurs permettant des simulations dans des tissus saisissants de réalisme. Cela dans le but de rendre accessible à l'audience la plus large possible l'imagerie de la microstructure basée sur des signaux Monte Carlo pour de futures études de population et l'avancement général du domaine.

Contents

List of Figures

List of Tables

Abbreviations

CSF	Cerebrospinal fluid
CUSP	Cube and sphere (diffusion MRI)
DCI	Diffusion compartment imaging
DIAMOND	Distribution of anisotropic microstructural environments with DWI
dMRI	diffusion MRI
DSI	Diffusion spectrum imaging
DTI	Diffusion tensor imaging
DWI	Diffusion-weighted imaging
DW-MRI	Diffusion-weighted magnetic resonance imaging
GM	Gray matter
GPD	Gaussian phase distribution
HARDI	High angular resolution diffusion imaging
HCP	Human connectome project
MAPL	Laplacian-regularized mean apparent propagator
MC	Monte Carlo
MCF	Multiple correlation function
MLP	Multi-layer perceptron
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MRF	Magnetic resonance fingerprinting
MMWMD	Minimal model of white matter diffusivity
NMR	Nuclear magnetic resonance
NODDI	Neurite orientation dispersion and density imaging
PGSE	Pulsed-gradient spin echo (sequence)
PD	Proton density
RF	Radio frequency
SE	Spin echo (sequence)
SGP	Short gradient pulse
SNR	Signal-to-noise ratio
STE	Stimulated spin echo (sequence)
TE	Echo time
WM	White matter
WMTI	White matter tract integrity

Symbols

A	normalized DW-MRI signal with NMR relaxation	/
b	b-value	$\mathrm{s\,m^{-2}}$
D	diffusivity	$\mathrm{m^2\,s^{-1}}$
E	DW-MRI signal attenuation	
f_{in} (f_{ex})	intra- (extra-) axonal volume fraction	
$\mathbf{F}$	single-fascicle sub-dictionary	/
$\mathbf{g} = G\hat{\mathbf{g}}$	magnetic-field gradient	$\mathrm{T\,m^{-1}}$
K	number of fascicles in a voxel	/
L_{step}	spins' fixed step length	m
M	number of sequences (measurements) in an acquisition protocol	/
M_0	net initial transverse magnetization	/
n	number of spatial dimensions	/
N	number of fingerprints in a single-fascicle dictionary	/
N_{spin}	number of spins	/
N_{step}	number of time steps	/
P	number of parameters describing a single fascicle of axons	/
$\mathscr{P}$	acquisition protocol	/
S	DW-MRI signal (not normalized)	/
S_0	non diffusion-weighted signal	/
$T1$	spin-lattice relaxation time	s
$T2$	spin-spin relaxation time	s
$\mathrm{Tr}\,\mathbf{A}$	trace of matrix $\mathbf{A}$	/
$\mathbf{u}$	orientation of a fascicle of axons	/
w	NMR apparent signal weight of a fascicle of axons	/
δ	duration of gradient pulse	s
Δ	separation between gradient pulses	s
ν	physical volume fraction of a fascicle of axons	/
ϕ	accumulated phase by a spin-bearing nucleus	/
ω_0	Larmor angular frequency	$\mathrm{rad\,s^{-1}}$
Ω	microstructural properties or diffusion environment	/
τ	time between RF pulses in spin-echo sequence	s

Physical Constants

Gyromagnetic ratio of hydrogen $\gamma = 2\pi\ 42.576 \times 10^6\ \mathrm{rad\ s^{-1}T^{-1}}$

1 Theoretical background

This Chapter provides the very basic tools for understanding the physics of diffusion-weighted magnetic resonance imaging, the models used for the diffusion of water in the white matter and Monte Carlo simulations, which this thesis heavily relies on. Reading of this Chapter is not strictly required for understanding the following Chapters as those were written to be as self-contained as possible.

1.1 Physics of diffusion-weighted MRI

1.1.1 Diffusion processes

Diffusion is a transport mechanism widely used in physics to describe the seemingly random movement of a substance in a given medium, for example water molecules in human body tissues. There are two widely-used approaches to study diffusion: a macroscopic approach and a microscopic approach.

In the macroscopic, also known as phenomenological or continuous approach, one considers the general equation for mass conservation at any given position $\mathbf{r}$ and time t

$$\frac{\partial c}{\partial t}(\mathbf{r}, t) = -\nabla \cdot \mathbf{J}(\mathbf{r}, t), \tag{1.1}$$

where c is the concentration of the substance and $\mathbf{J}$ its net flux, and Fick's phenomenological law

$$\mathbf{J}(\mathbf{r}, t) = -\mathbf{D} \cdot \nabla c(\mathbf{r}, t), \tag{1.2}$$

stating that a substance tends to diffuse from regions of high concentration to regions of low concentration and where $\mathbf{D}$ is a symmetric, positive-definite diffusion tensor. Combining (1.1) and (1.2) leads to the *diffusion equation*

$$\frac{\partial c}{\partial t}(\mathbf{r}, t) = \nabla \cdot (\mathbf{D} \cdot \nabla c\,(\mathbf{r}, t)) = D \Delta c\,(\mathbf{r}, t)\,, \tag{1.3}$$

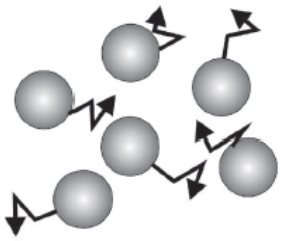

Figure 1.1: **Thermal agitation induces diffusion.** Diffusion resulting from seemingly random thermal motion of molecules (Price, 2009).

where the last equality only holds in the case of *isotropic diffusion*, with $D > 0$ being a scalar. This description of diffusion will be easily incorporated into the well-known Bloch equations and will lead to useful analytic results in diffusion-weighted imaging.

In the microscopic, also known as molecular or discrete approach, diffusion is treated as the motion of molecules colliding with each other under the effect of thermal agitation (Figure 1.1). The trajectory $\mathbf{r}(t)$ of each individual molecule is modeled as a discrete-time random walk on a discretized spatial grid, i.e. a succession of random "jumps" between adjacent positions on the spatial grid, each jump being independent from the past trajectory of the molecule (*Markov property*). The transition probabilities are generally taken to be spatially symmetric, such that $\mathbb{E}_{\mathbf{r}}[\mathbf{r} - \mathbf{r_0}] = 0$. The central limit theorem ensures that after a sufficiently long time t, the probability distribution of the displacement $\mathbf{r} - \mathbf{r_0}$ converges to a normal distribution of mean 0 and standard deviation proportional to $\sqrt{t}$, thus satisfying the requirements of Brownian motion. This description is particularly well-suited for numerical investigations and it will form the basis of the Monte Carlo simulations used in diffusion-weighted imaging.

The two above descriptions are closely related. It can be shown for instance that the continuous diffusion equation can be retrieved in the limit of finer time and space discretization of a random walk, i.e. when $\Delta t, \Delta x \to 0$, if the ratio $\frac{\Delta x^2}{\Delta t}$ is kept constant (Grebenkov, 2008).

1.1.2 Diffusion-weighted magnetic resonance imaging

Diffusion-weighted magnetic resonance imaging (DW-MRI) combines classical nuclear magnetic resonance (NMR) sequences of radio-frequency (RF) pulses such as the spin echo (SE) or stimulated echo (STE) sequences with applications of time-varying magnetic-field gradients $\mathbf{g}(t)$ leading to non-uniform magnetic fields and serving as a way to spatially encode the spin-bearing particles (*spins*) of the sample (Price, 2009, §2.1). A *diffusion sequence* is characterized by the gradient temporal profile $\mathbf{g}(t)$, or interchangeably by the vector quantity

$$\mathbf{q}(t) = \frac{\gamma}{2\pi} \int_0^t \mathbf{g}(s)\, ds \tag{1.4}$$

known as the *q-vector* and having units of $[\mathrm{m}^{-1}]$, where γ is the gyromagnetic ratio of protons and where the sign of each component of $\mathbf{g}(s)$ incorporates the effect of any refocusing 180° RF pulses.

The simplest diffusion sequence is known as the pulsed-gradient spin-echo (PGSE) or single diffusion encoding (SDE), the $\mathbf{g}$ and $\mathbf{q}$ temporal profiles of which are shown in Figure 1.2(a) and in the middle row of Figure 1.3. It consists in a constant gradient of duration δ, intensity G and direction $\hat{\mathbf{g}}$ turned on immediately after the initial RF excitation, followed by the same gradient applied a time Δ known as the *diffusion time* after the onset of the first pulse, which a refocusing 180° RF pulse made to have an opposite effective polarity $-\hat{\mathbf{g}}$. In double diffusion encoding (DDE) illustrated in Figure 1.2(b), two SDEs are performed successively using gradients $\mathbf{g}_1$ and $\mathbf{g}_2$ which traditionally have the same duration and intensity but different directions. Triple diffusion encoding (TDE) follows the same construction (see Figure 1.2(c)). We note a slight ambiguity arising when using the terms SDE, DDE and TDE as these do not indicate whether the underlying RF sequence was a SE or a STE. Such information must then be deduced contextually.

More general waveforms have been developed recently (Topgaard, 2017) which are referred to as q-vector trajectory encoding, isotropic encoding, spherical encoding or b-tensor encoding, as represented in Figure 1.2(d). These waveforms let each gradient component vary as a continuous function of time, unlike the previously described sequences in which a discrete number of gradient directions are probed. We finally mention the oscillating-gradient spin-echo (OGSE) sequence in which the direction of the gradient is fixed but its amplitude oscillates as depicted in Figure 1.3. The advantage of high-frequency OGSE sequences is often thought to be the ability to probe smaller length scales through shorter effective diffusion times (Drobnjak et al., 2016).

1.1.2.1 Mathematical descriptions of the DW-MRI signal

There are two convenient ways to mathematically apprehend the DW-MRI signal. The first description uses the random walk of each water molecule at the *microscopic* scale. The second one is a partial differential equation describing the diffusion and magnetic encoding at a larger, *macroscopic* scale.

Microscopic description Let us assume N_{spin} spin-bearing protons are evolving in Brownian motion in a given environment, such as a *voxel* of human white matter. At time $t = 0$ shortly after the initial 90° RF excitation, all N_{spin} spins are assumed to be precessing together in phase at the Larmor angular frequency. The dephasing ϕ_l of each spin at echo time TE under the external application of an effective gradient profile $\mathbf{g}(t)$ is

$$\phi_l(TE) = \gamma \int_0^{TE} \mathbf{g}(t) \cdot \mathbf{r_l}(t)\, dt \tag{1.5}$$

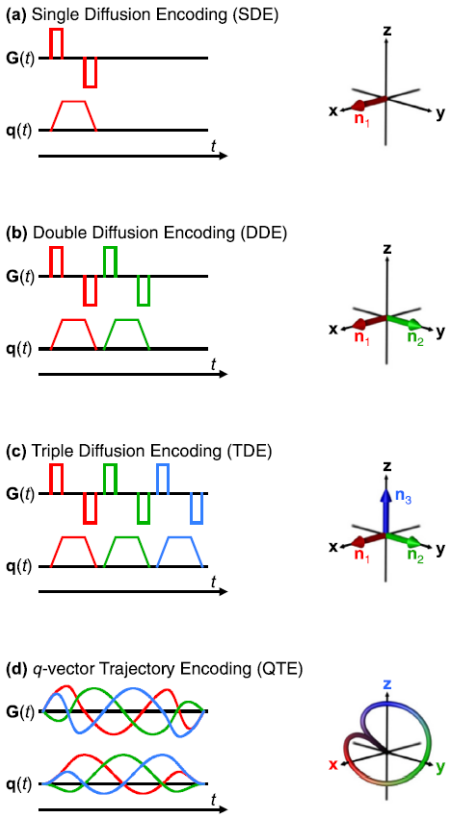

Figure 1.2: Time evolution of the x (red), y (green) and z (blue) components of the gradient **g** and q-vector **q** for the SDE, DDE, TDE and spherical encoding sequences. The right-hand-side column is a 3D representation of the time evolution of the q-vector. Figure taken from Topgaard (2017).

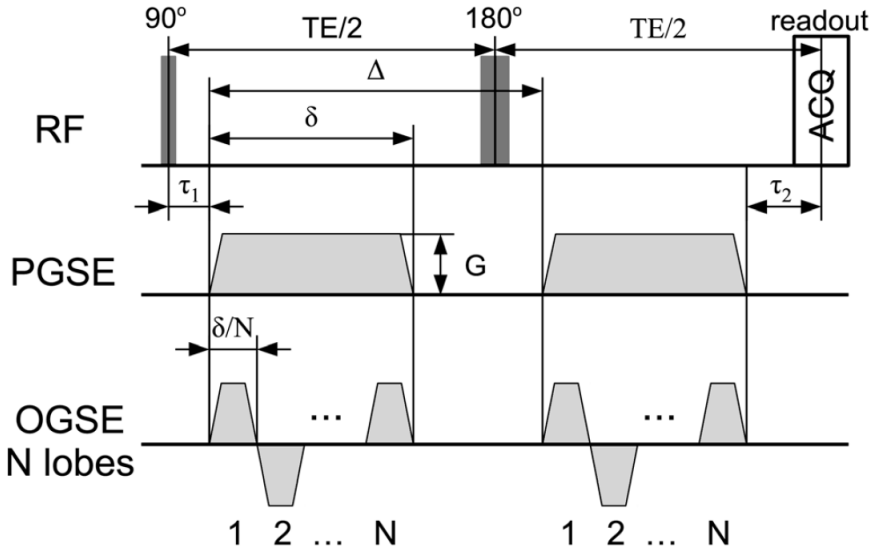

Figure 1.3: Comparison of the PGSE or SDE diffusion weighting (middle row) versus the OGSE sequence (bottom row), laid atop a traditional SE sequence (top row). Figure taken from Drobnjak et al. (2016).

where $\mathbf{g}(t)$ is the effective gradient with sign reflecting the effect of the refocusing 180° RF pulses, and where $\mathbf{r_l}(t)$ is the random trajectory of spin l, for $l = 1,\ldots,N_{\text{spin}}$, which is *completely independent from the applied magnetic gradient profile* and depends on the diffusivity D of the medium. The uniform magnetic field $\mathbf{B_0}$ contributes equally to the phase of each spin by a factor $\gamma B_0 \cdot TE$ and can therefore be included in an initial reference phase ϕ_0 that does not affect the measured signal.

Once the phase ϕ_l of each spin is known and assuming uniform T2-decay across the voxel, introducing the complex notation $S = M_x + iM_y$ for components of the net magnetization in the transverse plane, the attenuation E of the transverse magnetization resulting from the application of the magnetic-field gradient with profile $\mathbf{g}(t)$ in the considered diffusion environment is obtained as

$$E = \frac{S(TE)}{S_0} = \frac{e^{-\frac{TE}{T2}} \sum_{l=1}^{N_{\text{spin}}} \frac{M_0}{N_{\text{spin}}} e^{i\phi_l}}{M_0 e^{-\frac{TE}{T2}}} = \frac{1}{N_{\text{spin}}} \sum_{l=1}^{N_{\text{spin}}} e^{i\phi_l}, \tag{1.6}$$

where $M_0 = M_{x0} + iM_{y0}$ is the net magnetization in the xy-plane immediately after the first 90° RF pulse resulting from the equal contributions $\frac{M_0}{N_{\text{spin}}}$ of the N_{spin} spins, and where S_0 is the reference $T2$-weighted signal associated with that sample, when no external gradients are applied.

Given the enormous number of water molecules N_{spin} in a typical voxel of biological tissue[1],

[1] Considering pure water as a first approximation for brain tissues, a 1-mm^3 voxel contains 10^{-3} g of water,

Eq. (1.6) could be rewritten as a continuous expectation over ϕ

$$E = \mathbb{E}_{\phi}[e^{i\phi}]. \tag{1.7}$$

In the absence of diffusion barriers in the environment, the distribution of spin displacements $\mathbf{r}(t)$ is Gaussian by definition of Brownian motion, causing the distribution for the phase ϕ to be Gaussian too. This leads to the well-known decaying exponential for the signal attenuation E forming the theoretical basis of the *diffusion tensor* model (Basser et al., 1994), which most microstructural models are based on. In the general case where tissue structures hinder the free diffusion of water, closed-form expressions for the probability distribution $P_{\Phi}(\phi)$ are usually very difficult to obtain. A very common approximation known as the Gaussian phase distribution (GPD) is to simply assume a Gaussian distribution of spin displacements $\mathbf{r}(t)$, which is motivated by theoretical results in the long-time regime (Price, 2009, §2.3.3). The GPD approximation yields analytical solutions in simple geometries such as spheres and cylinders, which are very widely-used in the biophysical models of the microstructure described in Section 1.2.2 below.

Macroscopic description Combining the phenomenological Bloch equations (see e.g., (Price, 2009, §2.3.2)) with the classical diffusion equation yields the macroscopic Bloch-Torrey system of coupled linear partial differential equations for the evolution of the complex-valued transverse magnetization $M_{xy} = M_x + iM_y$, at every location $\mathbf{r}$ of the spatial domain and at every time $t > 0$ (Price, 2009, §2.3.2)

$$\frac{\partial M_{xy}}{\partial t}(\mathbf{r}, t) = \underbrace{\nabla \cdot \big(\mathbf{D} \cdot \nabla M_{xy}(\mathbf{r}, t)\big)}_{\text{diffusion}} \underbrace{- i\gamma \big(\mathbf{g}(t) \cdot \mathbf{r}\big) M_{xy}(\mathbf{r}, t)}_{\text{gradient encoding}}. \tag{1.8}$$

The $T2$-relaxation term was omitted since it only affects the measured signal by a known exponential factor. Similarly, the B_0 field only contributes an oscillating component $e^{i\gamma B_0 t} = e^{i\omega_0 t}$ and was therefore left out.

If the spatial domain is bounded, e.g. a sphere to model a glial cell or a cylinder to model an axon in the human brain, the conservation of mass is ensured through the general Robin boundary condition

$$\underbrace{\big(\mathbf{D} \cdot \nabla M_{xy}(\mathbf{r}, t)\big) \cdot \hat{\mathbf{n}}}_{\text{diffusion}} + \underbrace{\kappa M_{xy}(\mathbf{r}, t)}_{\text{permeability}} = 0, \tag{1.9}$$

where $\hat{\mathbf{n}}$ is the outward unit normal and κ denotes the permeability of the boundary, with units of $[\mathrm{m\,s^{-1}}]$. The most widely used boundary condition is the Neumann condition corresponding to perfectly reflecting cellular membranes, obtained setting $\kappa = 0$. The attenuation E of the

equivalent to $\frac{10^{-3}}{18.02}$ mol based on water's molar mass, $\frac{10^{-3} \times 6.022 \times 10^{23}}{18.02}$ water molecules using the Avogadro constant and $\frac{2 \times 10^{-3} \times 6.022 \times 10^{23}}{18.02} = 6.684 \times 10^{19}$ spin-bearing hydrogen nuclei.

transverse magnetization associated to the gradient profile $\mathbf{g}(t)$ in the diffusion environment Ω at the echo time TE is obtained as

$$E = \frac{1}{S_0}\int_\Omega M_{xy}(\mathbf{r}, TE)\, d\mathbf{r}, \tag{1.10}$$

where S_0 is the reference $T2$-weighted signal associated with that sample, i.e. when no gradients are applied. The Bloch-Torrey equation is in general difficult to solve for arbitrary domains, which makes the modeling of DW-MRI signals in arbitrary geometries a challenging task.

Free diffusion The Bloch-Torrey equations (1.8) can be solved in the case of so-called free, *Gaussian* diffusion where no diffusion barrier is present and the domain can be considered to be the whole n-dimensional space $\mathbb{R}^n$. Considering a diffusion process governed by the symmetric, positive-definite diffusion tensor $\mathbf{D}$, the following attenuation of the measured signal (Price, 2009, §2.3.2,§4.4.1) is obtained for an arbitrary gradient profile $\mathbf{g}(t)$

$$\begin{aligned} E(\mathbb{R}^n) &= \exp\left(-(2\pi)^2 \int_0^{TE} \mathbf{q}(t)^T \mathbf{D}\mathbf{q}(t)\, dt\right) \\ &= \exp\left(-\gamma^2 \int_0^{TE}\int_0^t\int_0^t \mathbf{g}(s_1)^T \mathbf{D}\mathbf{g}(s_2)\, ds_1 ds_2 dt\right) \end{aligned} \tag{1.11}$$

which in the case of the simple gradient profile of the PGSE experiment becomes

$$\begin{aligned} E(\mathbb{R}^n) &= e^{-b_{\text{pgse}}\hat{\mathbf{g}}^T\mathbf{D}\hat{\mathbf{g}}} && \text{(anisotropic case)}, \\ &= e^{-b_{\text{pgse}}D} && \text{(isotropic case)}, \end{aligned} \tag{1.12}$$

where $\hat{\mathbf{g}}$ is the unit gradient direction and $b_{pgse} = \left(\gamma\delta G\right)^2(\Delta - \delta/3)$. Solution (1.12) is equivalent to assuming the GPD in $\mathbb{R}^n$ in Eq. (1.7) and, as mentioned in the microscopic description of the signal above, is used in a variety of mathematical models used to approximate DW-MRI signals even in the case of *restricted* diffusion.

1.1.2.2 Statistical model for the acquisition noise

Diffusion-weighted MRI is marred by noise, motion artefacts and Eddy current distortions (Setsompop et al., 2013). In the PGSE sequence for instance, increasing the desired diffusion-weighting to increase the microstructural information contained in the signal necessarily comes at the cost of increased sequence duration (via Δ or δ) or increased gradient intensity G, which aggravates distortions (non-linearity, eddy currents due to rapid switching, etc.) and indirectly leads to longer sequences to give heated gradient systems time to cool down. To accelerate the final read-out at echo time, echo planar imaging (EPI) is often used, which is known to suffer from distortions (Chen and Wyrwicz, 1999; Price, 2009). DW-MRI scanning sessions are intrinsically long due to the requirement of applying multiple external magnetic gradients successively, which makes DW-MRI data more prone to motion between images.

All these artefacts are well understood and can be corrected to a certain extent in the post-acquisition phase (Sotiropoulos et al., 2013). In this paragraph, a statistical model is given for the random thermal noise which cannot be avoided or corrected.

Modern phased-array MRI systems possess various coils able to acquire multiple images in parallel. The detected signal S_l in each of the N coils can be accurately described by (Aja-Fernández and Tristán-Vega, 2012)

$$S_l = A_l + n_l, \qquad l = 1, \dots, N, \tag{1.13}$$

where A_l is the true signal value and n_l is a complex-valued white Gaussian noise process of variance σ_g^2 assumed identical for all N coils. A popular choice is to reconstruct a final *magnitude* or *Sum-of-Squares (SoS)* signal M as $M = \sqrt{\sum_{l=1}^{N} |S_l|^2}$. If there is no correlation between the coils, M can be shown to follow a non-central Chi distribution of $2N$ degrees of freedom with the following probability density function (Koay et al., 2009; Aja-Fernández and Tristán-Vega, 2012)

$$p_{\tilde{\chi}}\left(m|\eta, \sigma_g, N\right) = \frac{m^N}{\sigma_g^2 \eta^{N-1}} \exp\left(-\frac{m^2+\eta^2}{2\sigma_g^2}\right) I_{N-1}\left(\frac{m\eta}{\sigma_g^2}\right), \quad m > 0 \tag{1.14}$$

where $\eta = \sqrt{\sum_{l=1}^{N} |A_l|^2}$ is known as the underlying intensity and I_j is the j-th order modified Bessel function. It can be shown that the SoS reconstruction asymptotically (i.e. for large SNR) leads to an optimal SNR of η/σ_g, inducing a SNR gain of $\sqrt{N}$ compared to the ratio A_l/σ_g in each coil (Larsson et al., 2003).

A particular case of Eq. (1.14) is when $N = 1$. The distribution is then said to be Rician (Gudbjartsson and Patz, 1995). In practice, inter-coil correlation is impossible to avoid and has the effect of decreasing the effective numer of coils and increasing their effective variance (Aja-Fernández and Tristán-Vega, 2012). The Rician distribution is therefore often invoked as an effective noise model.

1.2 Modeling diffusion in the white matter

This section provides a very brief overview of the biology of the white matter and of the mathematical models aiming to describe it, classified into biophysical and signal models with a hint at numerical, computational models. Readers are referred to Jelescu and Budde (2017), Novikov et al. (2018a) or Alexander et al. (2019) for more thorough reviews of the literature.

1.2.1 Microstructure of the white matter

The microstructure of the white matter chiefly consists of long cables known as axons covered by a myelin sheath, usually bundled together into fascicles; and glial cells which among others

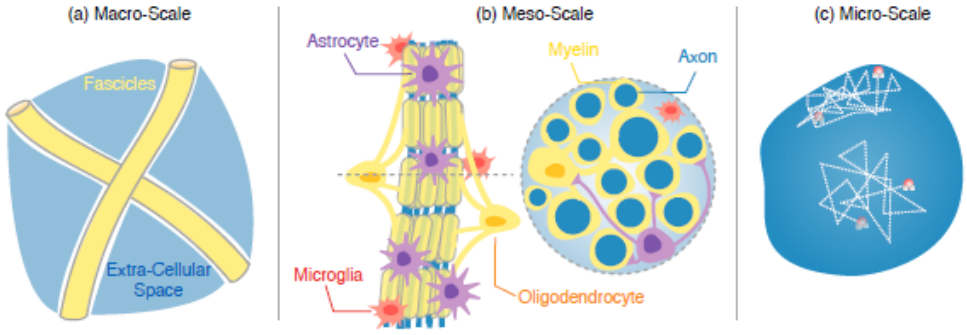

Figure 1.4: Schematic view of the white matter at three different length scales. Figure taken from Taquet (2013).

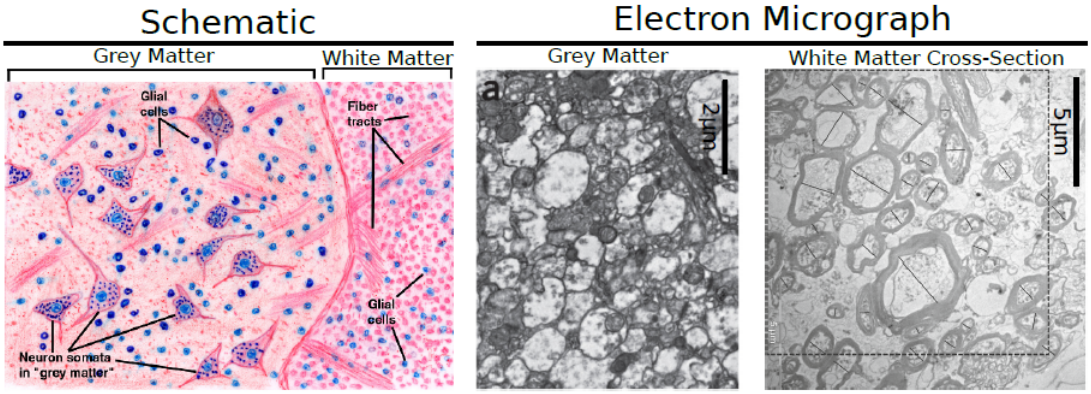

Figure 1.5: Differences between white and gray matter of nervous tissues. Figure taken from Fick (2017) adapted from original figures by Liewald et al. (2014) and Kay et al. (2013).

produce the myelin and nurture the axons (see Figure 1.4(b)). The spinal cord features single fascicles of mostly parallel axons. In the cerebrum and brain stem however, fascicles of axons exhibit various orientations and many areas contain complex arrangements of intersecting fascicles, as depicted in Figure 1.4(a). The main difference between the gray matter and the white matter is that the gray matter possesses neuron bodies or somas, as illustrated in Figure 1.5. Microstructural characteristics of interest within a voxel include but are not limited to: the number of fascicles and their respective orientations, the volume occupied by cerebrospinal fluid (CSF), the distribution of axon diameters within each fascicle, the intra-axonal volume fraction, the undulation of axons, the dispersion of their orientations, the myelin content, the permeability of the membranes and the volume occupied by glial cells.

Diffusivity of the white matter Water is assumed to diffuse in the intra- and the extra-axonal compartment and most of the mathematical modeling effort is devoted to describing diffusion in those two compartments. An important quantity is their intrinsic diffusivity D, which measures the speed of the diffusion process. However, the intrinsic diffusivity–at a given

temperature–of these compartments is difficult to measure in practice. The intra-axonal diffusivity was recently estimated to be $D = 2.25\,\mu m^2\,ms^{-1}$ *in vivo* (Dhital et al., 2019). Other recent experiments have suggested that the intra-axonal diffusivity was larger than the extra-axonal one (Kunz et al., 2018). Most modeling approaches assume equal diffusivities, which can have non-trivial consequences on the estimation of the model parameters (Jelescu et al., 2016).

NMR relaxation times in the white matter Quantities of similar interest for the accurate modeling of the white matter are the NMR relaxation times T1 and T2. Those quantities can greatly vary with the strength of the main B_0 field, which can have an important impact on preclinical experiments using high or ultra-high magnetic fields for small-animal studies, for instance. Figure 1.6 graphically summarizes the values reported in the literature for the T1 and T2 of the white matter and the CSF (which partially contaminates some voxels of white matter). T1 or T2 measurements of sub-voxel compartments such as the intra- or extra-axonal space are still extremely scarce. In the absence of a clear consensus, most models simply assume that the two compartments have equal relaxation times. However, experimental data has contradicted this hypothesis, suggesting a larger intra-axonal T2 (Peled et al., 1999; Wachowicz and Snyder, 2002; Bonilla and Snyder, 2007; Dortch et al., 2010; Veraart et al., 2018).

1.2.2 Biophysical models

This section presents an overview of biophysical models which attempt to directly and specifically reflect the biology of the tissue. First, an important assumption known as the superposition principle is discussed as most biophysical models rely on it to build complex models from elementary tissue compartments. These compartments are then described and two representative diffusion compartment models are presented in more detail.

1.2.2.1 Superposition principle

If the diffusion environment Ω can be expressed as the union of K mutually-disjoint compartments $\Omega_1, \ldots, \Omega_K$ with no inter-compartment water exchange (e.g., due to perfectly-reflecting boundaries), then the total signal attenuation $E(\mathbf{g};\Omega)$ associated to the gradient profile $\mathbf{g}(t)$ in the diffusion environment Ω can be expressed as

$$E(\mathbf{g};\Omega) = \sum_{i=1}^{K} f_i E(\mathbf{g};\Omega_i), \tag{1.15}$$

where $f_i = \frac{|\Omega_i|}{|\Omega|}$, for $i = 1,\ldots,K$, are the volume fractions of each compartment, possibly weighted by T2-relaxation or proton density if those characteristics are not identical for all compartments.

At the time scale of a DW-MRI acquisition, typically not more than 100 ms, the *slow-exchange*

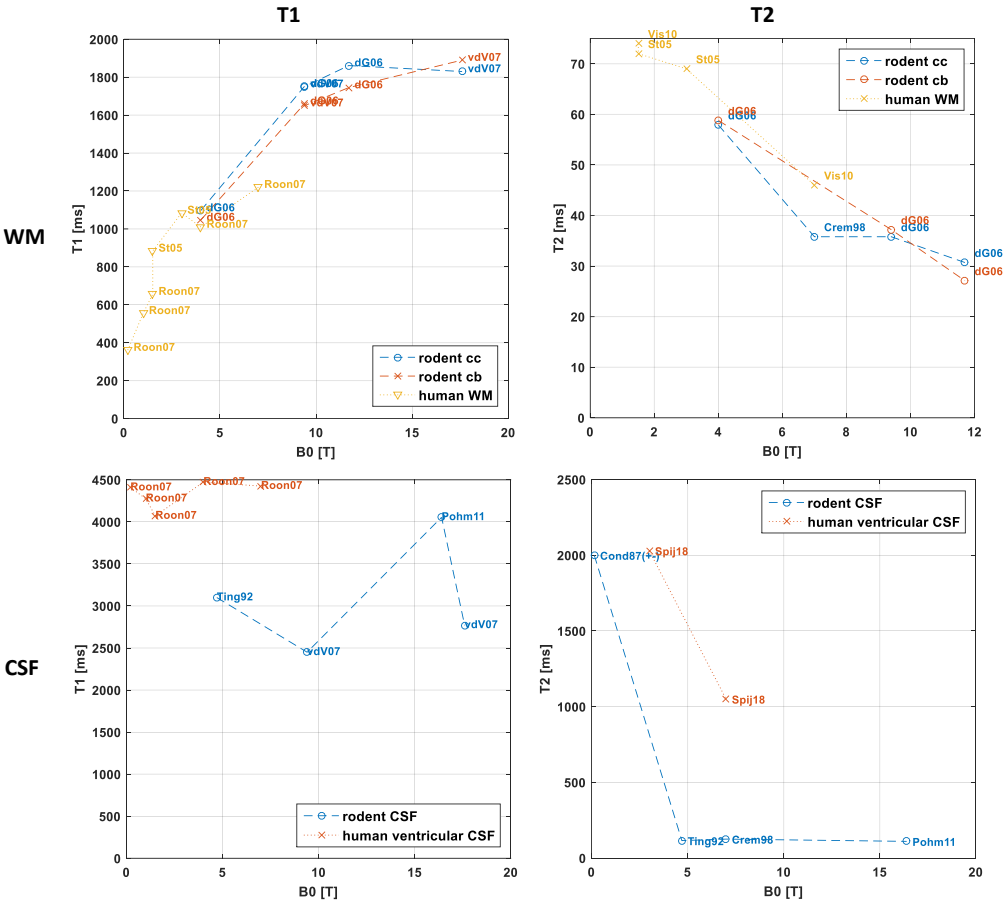

Figure 1.6: **T1 and T2 relaxation time measurements in the white matter (WM) and cerebrospinal fluid (CSF) in rodents and in humans.** Literature survey of the B_0 field-strength dependency for T1 and T2. Rodent values compiled in the corpus callosum (cc) and cerebellum (cb). Reference labels for rodent data: Cond87 (Condon et al., 1987), Ting92 (Ting and Bendel, 1992), Crem98 (Crémillieux et al., 1998), dG06 (de Graaf et al., 2006), vdV07 (van de Ven et al., 2007), Pohm11 (Pohmann et al., 2011). Reference labels for human data: St05 (Stanisz et al., 2005), Vis10 (Visser et al., 2010), Roon07 (Rooney et al., 2007), Spij18 (Spijkerman et al., 2018).

limit can be assumed meaning that little water transfer between compartments occurs and membranes can safely be considered impermeable. Consequently, the complex DW-MRI signal in a voxel can accurately be described as a weighted sum of the signals arising from individual compartments, supporting a paradigm known as diffusion compartment imaging (DCI) which encompasses the majority of state-of-the-art biophysical models of the microstructure.

1.2.2.2 Diffusion Compartment Imaging (DCI)

In compartment models of the microstructure or *diffusion compartment imaging* (DCI) and under the slow-exchange hypothesis, the total DW-MRI signal is described as a sum of the contributions arising from different independent compartments. Most models are based on two main compartments: the intra-axonal compartment capturing the diffusion of water trapped inside axons and usually represented by a stick or a cylinder model; and the extra-axonal compartment accounting for the water located outside of the axons in the tissue interstitium (Novikov et al., 2018a). Note that most of these models were primarily derived for the simple PGSE sequence.

Further refinements to the basic two-compartment model include modeling the fact that axons in a voxel of biological tissue are not perfectly parallel by a distribution of orientations on the sphere, either through a non-parametric decomposition into basis functions or by a fixed, parameterized distribution such as Bingham or Watson as in the NODDI model (Zhang et al., 2012) described below. The heterogeneity of axon diameters within a fascicle can be represented by a parameterized distribution of diameters such as the gamma distribution motivated by histological observations (Aboitiz et al., 1992) as in the AxCaliber framework (Assaf et al., 2008), although non-parametric approaches also exist (Benjamini et al., 2016). Ideally, the intra-axonal and extra-axonal compartments should be interdependent since the environment experienced by water molecules evolving in the extra-axonal space is directly affected by the shape and position of the axons modeled by the intra-axonal compartment. Heuristics to link intra- and extra-axonal compartments are generally referred to as tortuosity models (Stanisz et al., 2005).

In many cases, a so-called ball compartment represents partial volumes of CSF in a voxel. Similarly, the dot has been used as a way to model stationary water molecules like those trapped between lipid layers of the myelin sheath surrounding most axons or trapped in glial cells, although the relevance of this compartment has been challenged in recent *in vivo* experiments on human volunteers (Dhital et al., 2017).

ActiveAx The ActiveAx model (Alexander et al., 2010), also referred to as minimal model of white matter diffusivity (MMWMD) considers the signal contributions from up to 4 compartments

- intra-axonal (E_1) : DW-MRI signal attenuation due to water molecules confined within a cylinder of radius r calculated with the GPD approximation;
- extra-axonal I (E_2) : also known as the *hindered* compartment, models the contribution of molecules that evolve near cylindrical axons through a diffusion tensor **D** having its principal eigenvector parallel to the cylinders of the intra-axonal compartment;
- extra-axonal II (E_3) : CSF compartment (ball), corresponding to freely-diffusing molecules in an isotropic brain fluid, following $E_3 = e^{-bD}$;
- extra-axonal III (E_4) : stationary molecules (dot): $E_4 = 1$.

The slow-exchange hypothesis is assumed and the total signal attenuation of the ActiveAx model E_{AX} is calculated as

$$E_{AX} = \sum_{i=1}^{4} f_i E_i, \tag{1.16}$$

where the f_i are the volume proportions of each compartment within the voxel under study. This model thus only assumes one dominant fascicle of axons.

NODDI The Neurite Orientation Dispersion and Density Imaging (NODDI) model Zhang et al. (2012) captures the so-called dispersion of axons' orientations through a Watson distribution of orientations on the unit sphere.

To mitigate the additional complexity brought by the estimation of the parameters of the Watson distribution, simplifying assumptions are made such as assuming cylinders of zero radius or "sticks", considering a single population of axons in the voxel, making use of a simple tortuosity model or fixing the intra- and extra-axonal parallel diffusivities to a same value for all tissues.

This model has gained traction over the last few years, being used in a number of studies of normal human brain development (Jelescu et al., 2015; Kodiweera et al., 2016; Sato et al., 2017) as well as in disease (Adluru et al., 2014; Winston et al., 2014; Timmers et al., 2016; Schneider et al., 2017). It has also drawn criticism regarding the interpretation of the estimated parameters and how the model assumptions and simplifications may have an adversarial effect on the microstructural estimation (Jelescu et al., 2016; Lampinen et al., 2017; Jelescu and Budde, 2017).

1.2.3 Signal and phenomenological models

This section is concerned with models primarily designed to capture and reproduce the DW-MRI signal for any microstructural configuration. Very few assumptions about the underlying tissue microstructure are made as these models focus on a more general quantity known as the *average diffusion propagator*. The Fourier theory relating this quantity to the DW-MRI signal is first presented. Two sub-classes of methodologies are then described, which differ in the way they estimate the average diffusion propagator. A concluding paragraph discusses the

link with biophysical parameters of the tissue.

1.2.3.1 Fourier relationship between DW-MRI signal and average propagator

The PGSE experiment is considered and the duration δ of each gradient $\mathbf{g} = G\hat{\mathbf{g}}$ is assumed so short that the diffusion of spins during their application is negligible, an assumption known as the *short gradient pulse* (SGP) or narrow-pulse approximation. In particular, this requires, in an n-dimensional space, that $\delta \ll \frac{L^2}{2nD}$, where L is the characteristic length of the medium (e.g. the radius of an axon approximated by a cylinder) and D the diffusivity of the medium. However, it is assumed that the phase accumulated during each gradient pulse takes a finite, non-zero value, i.e. $\delta \to 0$ with $0 < \delta G < \infty$. The q-vector introduced in Eq. 1.4 is therefore well defined for any t in $]0, \Delta[$

$$\mathbf{q}(t) = \frac{1}{2\pi}\gamma\delta\mathbf{g}. \tag{1.17}$$

Under these assumptions, Eq. (1.7) gives the signal attenuation $E\left(\mathbf{q}, \Delta\right)$ associated with the diffusion time Δ and q-vector $\mathbf{q}$ in a diffusion environment Ω

$$\begin{aligned} E\left(\mathbf{q}; \Delta\right) &= \int_\Omega \int_\Omega \rho(\mathbf{r}_0)\mathscr{P}(\mathbf{r}_0, \mathbf{r}_f; \Delta) e^{i\gamma\delta\mathbf{g}\cdot(\mathbf{r}_0 - \mathbf{r}_f)}\, d\mathbf{r}_0 d\mathbf{r}_f \\ &= \int_\Omega \int_\Omega \rho(\mathbf{r}_0)\mathscr{P}(\mathbf{r}_0, \mathbf{r}_f; \Delta) e^{i2\pi\mathbf{q}\cdot(\mathbf{r}_0 - \mathbf{r}_f)}\, d\mathbf{r}_0 d\mathbf{r}_f, \end{aligned} \tag{1.18}$$

where $\rho(\mathbf{r}_0)$ is the equilibrium particle density and $\mathscr{P}(\mathbf{r}_0, \mathbf{r}_f; \Delta)$ is known as the diffusion propagator, interpreted as the probability of a molecule in Ω to travel from $\mathbf{r}_0$ to $\mathbf{r}_f$ between time 0 and Δ. Performing the change of variable $\mathbf{r} = \mathbf{r}_f - \mathbf{r}_0$, representing a vector displacement, and defining the *average diffusion propagator* P as

$$P(\mathbf{r}; \Delta) = \int_\Omega \rho(\mathbf{r}_0)\mathscr{P}(\mathbf{r}_0, \mathbf{r}_0 + \mathbf{r}; \Delta)\, d\mathbf{r}_0, \tag{1.19}$$

the elegant Fourier relationship is obtained between the measured signal attenuation E and the average propagator P (Price, 2009, §2.2.3)

$$E\left(\mathbf{q}; \Delta\right) = \int_\Omega P(\mathbf{r}; \Delta) e^{-i2\pi\mathbf{q}\cdot\mathbf{r}}\, d\mathbf{r} = \mathrm{FT}\left\{P(\cdot; \Delta)\right\}(\mathbf{q}), \tag{1.20}$$

from which quantities of interest can be computed such as the orientation distribution function (ODF) or the return-to-origin probability (RTOP).

All the techniques described in this section can be broadly branded as q-space imaging and essentially consist in sampling the signal in q-space to obtain the average diffusion propagator of a voxel via an inverse Fourier transform of the DW-MRI data $E\left(\mathbf{q}; \Delta\right)$, as originally developed by Callaghan (1991)

$$P(\mathbf{r}; \Delta) = \int_{\mathbb{R}^3} E\left(\mathbf{q}; \Delta\right) e^{i2\pi\mathbf{q}\cdot\mathbf{r}}\, d\mathbf{q}. \tag{1.21}$$

The spatial domain Ω is usually large enough with respect to the diffusion of particles over a length scale Δ so that Ω can be identified with $\mathbb{R}^3$.

In practice, many acquisition protocols relax the constraint on δ, with the consequence that $\mathbf{r}$ should be interpreted as the net displacement of a spin from its mean position within the time interval $[0,\delta]$ to its mean position in the interval $[\Delta,\Delta+\delta]$, a quantity known as the δ-averaged relative spin displacement (Wedeen et al., 2005). Remarkably, Eq. 1.20 and 1.21 hold irrespective of the actual underlying tissue microstructure. Making the reasonable assumption that the tissue and therefore the average propagator is symmetric with respect to the origin ensures that its Fourier transform, i.e. the DW-MRI signal, is real-valued under the application of uniform magnetic field gradients (Grebenkov, 2008).

1.2.3.2 Direct q-space imaging

Diffusion spectrum imaging (DSI), developed by Wedeen et al. (2005) is one of the most straightforward applications of the formalism described above. DW-MRI data is acquired on a lattice of the so-called q-space with fixed δ and Δ by letting the diffusion gradient $\mathbf{g}$ take values on a cubic lattice on grounds of the q-g relationship (1.17). A discretized version of the average diffusion propagator is obtained by inverse 3D Fourier transform of the acquired DW-MRI data, with integration performed in q-space, without resorting to any prior modeling.

The main drawback of direct q-space imaging is the high number of samples required for accurate inverse Fourier computation. In addition, integration of the average propagator along radial directions to obtain the ODF introduces inaccuracies if the data was acquired on a cartesian grid.

1.2.3.3 Functional bases for q-space imaging

Most of the shortcomings of direct q-space imaging hinted at in the previous paragraph can be addressed by decomposing the normalized DW-MRI signal E into a linear combination of well-chosen basis functions ϕ_i

$$E(\mathbf{q};\Delta) = \sum_i a_i(\Delta)\phi_i(\mathbf{q}), \tag{1.22}$$

where the dependence of the coefficients a_i on Δ is often omitted for clarity. The functions are usually taken to form a complete orthogonal basis for a functional space such as that of square-integrable functions L^2 in $\mathbb{R}^3$ or on the unit sphere S^2 in order to ensure adequate convergence of the series (1.22) to E. If the inverse Fourier transform φ_i of each basis function ϕ_i is available analytically (which is usually ensured by construction), it is straightforward to obtain the average diffusion propagator as

$$P(\mathbf{r};\Delta) = \sum_i a_i\varphi_i(\mathbf{r}), \tag{1.23}$$

from which quantities such as the ODF or RTOP can be analytically derived, irrespective of the sampling in q-space. In practice, heuristics or trial and error is needed to determine the number of terms after which the summations in Eq. (1.22) and (1.23) should be truncated.

In Mean Apparent Propagator MR imaging (MAP-MRI) (Özarslan et al., 2013), a three-dimensional functional basis is assembled based on the one-dimensional basis introduced in Simple Harmonic Oscillator based Reconstruction and Estimation (1D-SHORE) (Özarslan et al., 2008). A variety of scalar indices are mathematically expressed in terms of the estimated coefficients a_i, in addition to the RTOP index mentioned above, such as a return-to-the-axis probability (RTAP), return-to-the-plane (RTPP), propagator anisotropy (PA), non-gaussianity (NG) and a similarity metric to compare different apparent propagators, which were reported to provide useful contrasts on fixed monkey brain (Özarslan et al., 2013).

Many other functional bases are possible such as rotational and spherical harmonics in constrained spherical deconvolution (CSD (Tournier et al., 2004)), spherical harmonics alone (Descoteaux et al., 2007) or the Spherical Polar Fourier basis (Caruyer and Deriche, 2012). Additional regularization of either the reconstructed signal or the inferred propagator has also been shown to be beneficial, including smoothing of the laplacian (Fick et al., 2016) or imposing more stringent continuity constraints (Caruyer and Deriche, 2012). Recent extensions have included explicit dependence on the diffusion time Δ (Fick et al., 2015) or have accounted for non-symmetric diffusion propagators P, which lead to complex-valued signals E (Pizzolato et al., 2016).

1.2.3.4 Relating signal models to the tissue microstructure

A general weakness of signal models based on q-space imaging, whether they use functional basis decomposition or not, is the lack of direct link between indices based on the diffusion propagator and biophysical properties of the tissues. In order to estimate the intra-axonal volume fraction for instance, several indices were proposed such as the apparent fiber density (AFD) derived from the computed ODF (Raffelt et al., 2012) or an index based on RTAP from MAP-MRI (Fick, 2017). Another index based on RTAP was also proposed (Fick, 2017) for the estimation of the apparent axon diameter, a microstructural feature of utmost importance. However only limited success in correlating with histological measurements has been reported thus far (Fick, 2017, §4.4). Most of the q-space based methodologies in the literature report contrasts derived from signal models with the hope that it will lead to biomarkers of pathologies irrespective of what they exactly represent biologically. Table 1.1 summarizes the advantages and drawbacks of the two classes of models discussed so far.

1.2.4 Numerical models

Numerical methods provide the flexibility and the modeling complexity that the analytical models presented above lack at the cost of important computational requirements. Finite ele-

Table 1.1: **Trade-offs in selecting a model of microstructural diffusion.** Summary of the pros and cons of signals models versus biophysical models of the microstructure based on DW-MRI.

	+	−
Signal models	• suited for any microstructural configuration • signal interpolation, denoising • stable and efficient • contrasts may better generalize to abnormal WM	• not directly related to biophysical properties of the tissue • no closed-form formula to analyze effect of acquisition and tissue parameters • specific to PGSE with narrow pulses
Biophysical models	• more biologically realistic and interpretable • flexible choice of compartment models • usually provide analytical formulas for analysis	• complex and unstable non-linear estimation • model assumptions may break down in pathological cases • parameter interpretation should be exerted with caution

ments (Van Nguyen et al., 2014) benefit from strong theoretical guarantees and can accurately model complex tissue configurations and gradient waveforms. However, spatial meshing is a notoriously arduous task, non-trivial mathematics are required to set up a simulation and verify convergence, boundary conditions are difficult to implement, the complexity of the problems scales poorly with the size of the simulated voxel and simulations need to be run from the start for each new acquisition sequence. Monte Carlo simulations have emerged as the reference numerical model for the simulation of diffusion in the white matter owing to their relative ease of use, flexibility and physical soundness. However, they have largely been ignored as a intrinsic building tool for models of the microstructure, which was one of the motivations of this thesis. The next paragraph provides a brief description of the Monte Carlo framework and their use in practice.

1.3 Monte Carlo simulations of the diffusion

Monte Carlo simulations have become a modality of choice for the validation of closed-form analytical models such as described above as they are considered to provide a groundtruth against which these models can be assessed. This is because, for a fixed diffusion environment and given sufficient computation time, the output converges to the true solution of the Bloch-Torrey equation for the DW-MRI signal. Two steps are required: constructing a 3D geometry, and running the actual simulation of restricted Brownian motion and magnetic-field phase encoding.

1.3.1 Diffusion environment

Constructing realistic numerical phantoms of the tissue must strike the right balance between realism and computational complexity.

The simplest and most lightweight approach is to consider idealized tissue geometries such as cylinders for axons (Hall and Alexander, 2009; Rensonnet et al., 2018) or spheres for glia, possibly with mathematically-expressed refinements to simulate undulation or dispersion (Budde and Frank, 2010; Nilsson et al., 2010) and cellular processes such as spines and leaflets (Palombo et al., 2017) for instance.

For more complex tissue configurations, a spatial discretization stage is usually required. Traditionally, triangular meshes have been used to discretize 3D continuous environments (Hall and Alexander, 2009) and relatively complex tissue geometries have been replicated from electron microscopopy data (Panagiotaki et al., 2010; Xu et al., 2014). For more efficient GPU implementations, a few authors have considered simplified binary maps to label the intra- or extra-axonal space (Waudby and Christodoulou, 2011).

In recent years, a number of methods have been proposed to faithfully represent complex white matter configurations using just a limited number of parameters in a generative model based on idealized initial shapes (Palombo et al., 2016, 2019; Ginsburger et al., 2018). An example of a realistic white matter voxel is shown in Figure 1.7. In an approach based on the natural process of axon growth, high values of axon density and orientation dispersion have been obtained simultaneously (Callaghan et al., 2019). Voxel configurations with exquisite biological realism have also been achieved by threading small spheres together using an efficient GPU implementation (Ginsburger et al., 2019a).

Besides white matter, the modeling of gray matter tissues has also been progressing. As illustrated in Figure 1.8, generative models of complex soma and dendrite structures representing various types of cortical neurons and glial cells are becoming available (Palombo et al., 2019). Recent experiments have suggested that ellipsoids were satisfactory representations of glia while neurons could be modeled by spheres, which would decrease the overall complexity of generating synthetic environments (Gilani et al., 2019).

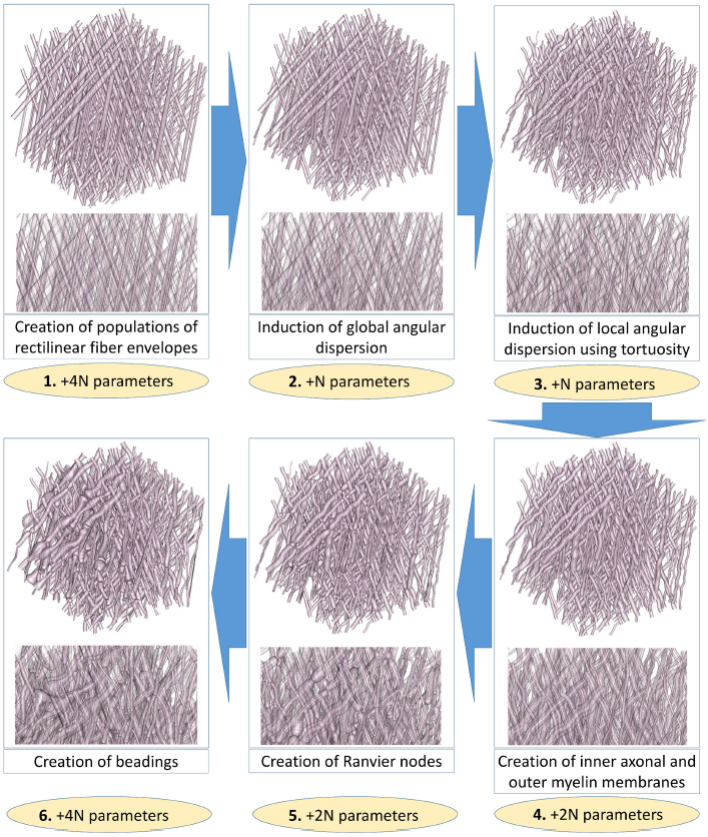

Figure 1.7: **Recent realistic white matter substrates for Monte Carlo simulations.** In Ginsburger et al. (2018), building from (1) simple straight cylinders representing intersecting populations of axons, additional features such as (2) dispersion, (3) tortuosity, (4) myelin sheaths, (5) Ranvier nodes and (6) beadings are gradually added.

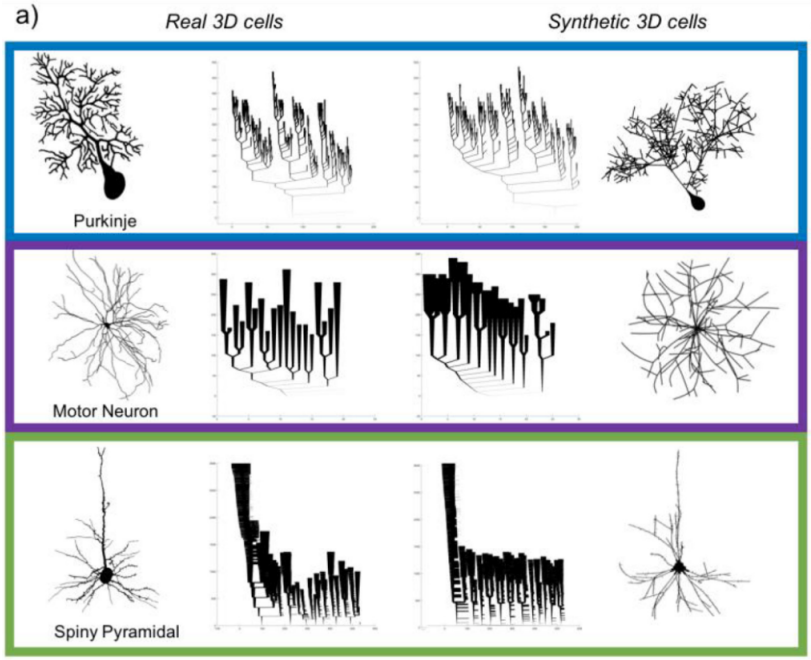

Figure 1.8: **Recent realistic gray matter substrates for Monte Carlo simulations.** Examples of three types of brain neurons along with their dendrograms synthetized with the generative parametric model presented in Palombo et al. (2019).

.

1.3.2 Spin dynamics

Monte Carlo simulations are directly based on the microscopic description of DW-MRI signals presented in Section 1.1.2.1 and mainly consist in generating the trajectories of N_{spin} spins or *random walkers*, initially uniformly distributed across the geometry, by means of a discrete-time random walk. Considering a partition of the time between the initial RF excitation pulse and the echo time TE into N_{step} intervals $\{0 = t_0, t_1, \ldots, t_{N_{\text{step}}-1}, t_{N_{\text{step}}} = TE\}$, then from a random initial position $\mathbf{r_0} = \mathbf{r}(t_0)$ in the domain Ω, each spin's trajectory is updated as

$$\mathbf{r}(t_{s+1}) = \mathbf{r}(t_s) + \Delta\mathbf{r}_s, \tag{1.24}$$

where each step $\Delta\mathbf{r}_s$ is of random orientation and of fixed length

$$L_{step} = \sqrt{2nD\delta t}, \tag{1.25}$$

with n the spatial dimension of the diffusion environment and $\delta t = t_s - t_{s-1}$, for $s = 1, \ldots, N_{\text{step}}$. As the steps sum up, the central limit theorem ensures that the distribution of all n components of the position $\mathbf{r}(t)$ at time t will converge to a Gaussian distribution of mean zero and variance $2Dt$ as expected from Brownian motion in the absence of reflection. Spins that encounter a barrier during one such step are elastically reflected if the membranes are considered perfectly impermeable, as illustrated in Figure 1.9. Membrane permeability is more tedious to

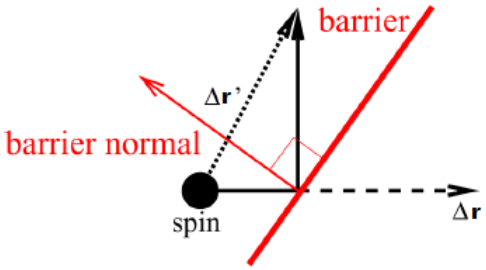

Figure 1.9: Spins that encounter a barrier during a step update $\Delta\mathbf{r}$ are elastically reflected and take an effective step $\Delta\mathbf{r}'$. Several consecutive reflections may occur. Figure adapted from Hall and Alexander (2009).

implement and requires the introduction of carefully-selected crossing probabilities (Regan and Kuchel, 2000; Fieremans et al., 2010; Nilsson et al., 2010; Lee et al., 2013). Non-elastic reflections can also be considered and might lead to similar results more efficiently (Xing et al., 2013).

The accumulated phase ϕ_l of each spin l is computed independently from the generation of the trajectory by computing the time integral (1.5) via an appropriate numerical method such as a rectangle quadrature

$$\phi_l = \gamma \cdot \delta t \cdot \sum_{s=1}^{N_{\text{step}}} \mathbf{g}(t_s) \cdot \mathbf{r}_l(t_s), \tag{1.26}$$

where $\mathbf{g}(t)$ is the profile of the external magnetic-field gradient, making the method suitable for any of the diffusion sequences presented in Section 1.1.2. The final signal attenuation E is computed as the sample mean of $e^{i\phi}$ as in Eq. (1.7)

$$E(\mathbf{g};\Omega) = \frac{1}{N_{\text{spin}}} \sum_{l=1}^{N_{\text{spin}}} e^{i\phi_l}.$$

Computing only the real part or the modulus of the signal is usually sufficient since it has been shown that DW-MRI signals are real and positive under relatively mild assumptions on the tissue microstructure, usually met in our experiments (Wedeen et al., 2005).

The computation time is dominated by the number of spins N_{spin}, which must be large enough to reduce statistical variance and by the number of time steps N_{step}, which should be large enough to avoid numerical bias. The choice of these two simulation parameters is not trivial and is often guided by practical heuristics in the absence of known groundtruth (Hall and Alexander, 2009; Rensonnet et al., 2015; Fieremans and Lee, 2018).

2 Validating the superposition approximation for crossing fascicles

This Chapter was based on a post-print version of Rensonnet et al. (2018). The introduction was updated with a reference to Schilling et al. (2017) suggesting that complex interweaving is the most likely pattern of fascicle crossings. Additional acquisition protocols were considered in Appendix A.3, with results briefly discussed in Section 2.3.1, strengthening a point made in the discussion (Section 2.4) about the validity of the superposition at higher b-values.

2.1 Introduction

A particularly challenging task in brain microstructure mapping based on diffusion-weighted magnetic resonance imaging (DW-MRI) is to estimate microstructural properties in regions of the white matter where multiple fascicles of axons intersect. In (Jeurissen et al., 2013), authors have reported that between 63% and 90% of voxels in the white matter at a resolution of $2.4\,\mathrm{mm}^3$ contain more than one fascicle, suggesting that regions of multiple fascicles actually make up the majority of the white matter voxels at common clinical resolutions. Recent evidence suggests that the prevalence of crossing-fascicle voxels might actually increase with smaller voxel sizes (Schilling et al., 2017). At the scale of a voxel, there can in theory be two types of crossing configurations: either each fascicle occupies its own portion of the voxel with axons that do not intermingle with the axons of the other fascicles; or axons from each fascicle abandon their tight, bundle-like organization and instead adopt an interwoven pattern at the intersection, evolving alongside axons from other fascicles (Axer et al., 2000; Schilling et al., 2017).

Most models of the microstructure rely on the *superposition approximation*, i.e. they consider that the signal arising from crossing fascicles is equal to the sum of the signals arising from each fascicle independently. This assumption is used in multi-tensor models (Tuch et al., 2002; Scherrer and Warfield, 2012) and their extensions to distributions of tensors (Scherrer et al., 2016, 2017), in models assuming various restricted water compartments with different directions (Assaf et al., 2004; Zhang et al., 2011), in dictionary-based methods (Ramirez-Manzanares et al., 2007; Auría et al., 2015a; Aranda et al., 2015) and also implicitely in spherical deconvolu-

tion frameworks (Tournier et al., 2004, 2007; Jeurissen et al., 2014; Canales-Rodríguez et al., 2015; Canales-Rodríguez et al., 2019). This hypothesis is expected to hold reasonably well when fascicles maintain their bundle-like structure and cross along separate pathways in the voxel, even though the interface between the two fascicles is ignored by the approximation. Indeed in a clinical voxel of a few cubic millimeters the diffusion of a vast majority of water molecules will not be affected by this micrometer-scale interface. In contrast the validity of the superposition approximation can be theoretically challenged in voxels where fascicles intersect in *interwoven patterns* since the diffusion of water molecules in the interstitium, or extracellular space, is then simultaneously hindered by all fascicles. Whether the approximation is appropriate and allows microstructural parameters of interest to be reliably estimated in such configurations with common acquisition sequences remains an open question that has so far received little attention in the literature and which we investigate numerically in this paper.

In this work, we assume that the groundtruth is a voxel in which fascicles intersect in interwoven planes, thus sharing the interstitium, and we evaluate the quality of the superposition approximation for fixed, clinically realistic acquisition protocols. Relying on simple yet representative two-compartment phantoms of the microstructure, we first examine the similarity between the DW-MRI signals arising from the groundtruth configuration and the signals arising from the superposition approximation at fixed, matching microstructural parameters. We then conduct a series of experiments to investigate whether microstructural parameters of interwoven fascicles can be accurately estimated from an approximate model consisting in the superposition of independent fascicles. All the DW-MRI signals are obtained using Monte Carlo simulations of the random walk of water molecules, which leverage the well-known physics of diffusion processes.

2.2 Methods

This section presents the synthetic phantoms used in this study, a formal definition of the superposition approximation and the Monte Carlo framework for the simulation of DW-MRI signals. It then details how we studied the impact of the superposition approximation on the DW-MRI signal at fixed microstructural configuration. Lastly, it describes the estimation experiments that were conducted to examine the impact of using the approximation to estimate microstructural parameters of interwoven fascicles.

2.2.1 Synthetic Phantoms of the Microstructure

Single fascicles. We modeled single fascicles of axons by an infinite array of hexagonally-packed straight, parallel and infinitely-long cylinders. A single fascicle is thus characterized by an orientation $\mathbf{u}$, a unique cylinder radius r interpreted as an axonal radius index and by a cylinder packing density f interpreted as an axonal density index (Alexander et al., 2010), which we denote by the quantity $\Omega_{\text{sing}} = \left(\mathbf{u}, r, f\right)$. The DW-MRI signal arising from such a

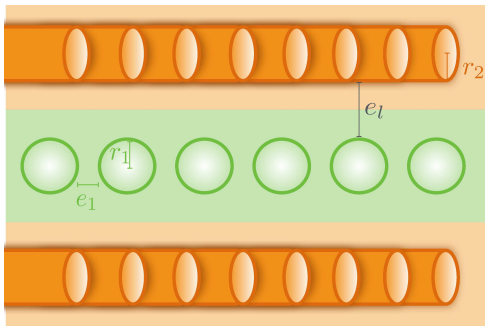

Figure 2.1: **Synthetic phantom of interwoven fascicles of axons.** The total environment is made up of a periodic arrangement where one layer of cylinders with radius r_1 and inter-cylinder spacing e_1 from Population 1 is followed by one layer of cylinders with radius r_2 and inter-cylinder spacing e_2 from Population 2, a distance e_l from the first layer. The green and orange shaded areas show how the water molecules of the interstitium are distributed among the populations, allowing us to define the respective fractions of occupancy v_1 and v_2 and of the population-specific cylinder packing densities f_1 and f_2.

configuration for an acquisition sequence of general parameters $\mathbf{p}$ is denoted by $S_{\text{sing}}\left(\Omega_{\text{sing}};\mathbf{p}\right)$ and forms the basis of the superposition approximation investigated in this work.

Interwoven fascicles. Interwoven fascicles were modeled by two populations of straight, parallel and infinitely-long cylinders of respective orientations $\mathbf{u}_1$ and $\mathbf{u}_2$, crossing in interleaved layers one cylinder thick with inter-layer spacing e_l in a pattern repeating periodically, as depicted in Figure 2.1. All cylinders within Population 1 (Population 2) have the same radius r_1 (r_2), interpreted as an axonal radius index, with inter-cylinder spacing e_1 (e_2). In practice it is often more intuitive to report the populations' intrinsic packing densities f_1 and f_2, interpreted as axonal density indices, and the populations' volume fractions of occupancy v_1 and v_2 rather than the spacing parameters. However, the interstitium is shared by both fascicles and any separation thereof as well as any definition of the quantities f_1, f_2, v_1 and v_2 is thus arbitrary. In the remainder of this work, we defined them based on a parceling of the extracellular space at a plane located a distance $e_l/2$ from both populations of axons (see Figure 2.1).

In general an interwoven-fascicle environment is thus completely characterized by the set of parameters $\Omega_{\text{intw}} = \left(\mathbf{u}_1, r_1, f_1, \mathbf{u}_2, r_2, f_2, v_1\right)$. Related useful quantities are readily obtained from Ω_{intw} such as $v_2 = 1 - v_1$, the crossing angle α as the smallest angle between $\mathbf{u}_1$ and $\mathbf{u}_2$ in the range $[0, 90°]$ and the global axonal density index $f_{\text{tot}} = v_1 f_1 + v_2 f_2$. The DW-MRI signal arising from such a configuration for an acquisition sequence of general parameters $\mathbf{p}$ is denoted by $S_{\text{intw}}\left(\Omega_{\text{intw}};\mathbf{p}\right)$ and was considered as the reference, groundtruth signal throughout this study. The impact of adding more cylinders to each population's layer in the interwoven configuration was studied in Appendix A.1 (Figures A.1 and A.2). The single-cylinder layers selected for our analyses were shown to least resemble the approximate superposed fascicles, thereby corresponding to a "worst-case" situation.

2.2.2 Definition of the Superposition Approximation

Given an acquisition sequence of parameters $\mathbf{p}$, we defined the superposition signal $S_{\text{appr}}\left(\Omega_{\text{appr}};\mathbf{p}\right)$, supposed to approximate the groundtruth $S_{\text{intw}}\left(\Omega_{\text{intw}};\mathbf{p}\right)$, as the linear combination

$$S_{\text{appr}}\left(\Omega_{\text{appr}};\mathbf{p}\right) = v_1 S_{\text{sing}}\left(\mathbf{u}_1, r_1, f_1;\mathbf{p}\right) + v_2 S_{\text{sing}}\left(\mathbf{u}_2, r_2, f_2;\mathbf{p}\right), \tag{2.1}$$

where the weights v_1 and v_2 are interpreted as the fractions of the total volume occupied by each fascicle, satisfying $v_1 + v_2 = 1$. Similarly to fascicles crossing in interleaved planes, the global axonal density index is obtained as $f_{\text{tot}} = v_1 f_1 + v_2 f_2$.

An approximate interwoven-fascicle environment is fully characterized by the set of parameters $\Omega_{\text{appr}} = \left(\mathbf{u}_1, r_1, f_1, \mathbf{u}_2, r_2, f_2, v_1\right)$ which all have a straightforward equivalent parameter in Ω_{intw} so that both parameter sets can easily be compared.

2.2.3 Signal Simulations

We relied on Monte Carlo simulations of the random walk of water molecules to obtain DW-MRI signals for the single and interwoven-fascicle configurations described above in order for our study to be as model-independent as possible and because no exact generative formulas exist for interwoven fascicles of axons.

Specifically, we followed the method described in (Rensonnet et al., 2015) where the exact intracellular signal S_{in} is obtained by the Multiple Correlation Function (MCF) formalism (Grebenkov, 2008) and the extracellular signal S_{ex} is computed from fixed-step Monte Carlo simulations of the random diffusive motion of water molecules in the extracellular space with perfectly elastic reflections at the membranes as described in (Hall and Alexander, 2009). Compared to Monte Carlo simulations performed in both the intra and extracellular compartments, this provides significant gains in precision at fixed computation time and equivalently, significant computational gains for a set precision (Rensonnet et al., 2015).

We used the same intrinsic diffusivity D in the intra and extracellular compartments and fixed its value to $D = 2.0 \times 10^{-9}\,\text{m}^2\,\text{s}^{-1}$, in agreement with values used in similar Monte Carlo settings (Hall and Alexander, 2009; Nilsson et al., 2009; Panagiotaki et al., 2010). As noted in (Fieremans et al., 2010), we stress that all the results obtained in this study can be obtained for another value of D by appropriately rescaling the spatial lengths L and magnetic gradient intensities G, since the diffusion signal is fully characterized by the two dimensionless parameters $p = DT/L^2$ and $q = \gamma GLT$ (Grebenkov, 2008), where γ is the gyromagnetic ratio of protons and T the characteristic time scale. All the simulations for the extracellular signal of single fascicles were performed using the Camino Diffusion MRI Toolkit (Hall and Alexander, 2009). The simulations in the extracellular space of interwoven fascicles utilized a new in-house software written in C/C++ and extending the capabilities of the Camino Toolkit to all the interwoven-fascicle configurations considered in this study. Care was taken to ensure that the new code reproduced the results of Camino on the simpler configurations handled by both softwares.

Throughout this study we focused on multi-shell high angular resolution diffusion-weighted imaging (HARDI) protocols (Tuch et al., 2002) based on the widely used pulsed-gradient spin-echo (PGSE) acquisition sequence. Table 2.1 summarizes the two clinically-realistic protocols that we used in our experiments: the 4-shell, 4-diffusion-times human protocol used with the ActiveAx model (Alexander et al., 2010) and the extended 4-shell protocol proposed with the NODDI estimation framework (Zhang et al., 2012), which we refer to as Protocols A and B, respectively. For completeness, additional HARDI shells spanning a larger range of acquisition parameters were considered in Appendix A.3.

It is worth noting that, given the way S_{sing} and S_{intw} were simulated in the absence of membrane permeability, the intracellular signal of the superposition approximation is by definition exactly identical to that of fascicles crossing in interwoven planes when the radius indices and the fascicles' orientations coincide. In our models, the discrepancies between the reference interwoven-fascicle signals and the approximate signals *at matching microstructural parameters* therefore solely arise from the extracellular signal contribution. The intra- and extracellular contributions are not independent however: as an example, changing the radius index at fixed inter-cylinder spacing directly impacts the intracellular signal but also affects the extracellular geometry and the global axonal density, thereby modifying the extracellular signal and the relative weight of each compartment's contribution. This impact is non-trivial and depends on whether or not the superposition approximation is used.

Table 2.1: **Multi-shell acquisition protocols.** The $\hat{\mathbf{g}}$ column contains the number of gradient directions in each shell (S) characterized by a gradient intensity G, duration δ and diffusion time Δ separating the onsets of the two gradient lobes, combining into a b-value $b := (\gamma G \delta)^2 (\Delta - \delta/3)$.

Acquisition protocol	$\hat{\mathbf{g}}$	G [mT m^{-1}]	δ [ms]	Δ [ms]	b [smm^{-2}]
Protocol A (ActiveAx (Alexander et al., 2010))	S1 : 90	57	5	87	496
	S2 : 90	60	13	20	682
	S3 : 90	46	15	77	2453
	S4 : 90	58	12	80	2635
Protocol B (NODDI (Zhang et al., 2012))	S1 : 30	31.9	17.5	37.8	711
	S2 : 30	37.8			1000
	S3 : 60	53.4			2000
	S4 : 60	63.8			2855

2.2.4 Impact of the Approximation on the DW-MRI Signal

This section aims at identifying the order of magnitude of the differences between the interwoven-fascicle signal S_{intw} and the approximate signal S_{appr} when the reference parameters Ω_{intw} and the parameters used in the approximation Ω_{appr} match. Without loss of generality, we compared DW-MRI signals of configurations featuring fascicles with identical microstructural properties, i.e. we fixed $r_1 = r_2 = r$, $f_1 = f_2 = f$ and $v_1 = v_2 = 0.5$.

Specifically, we considered 24 microstructural configurations resulting from the combination of 6 radius indices $r = [0.5, 1, 2, 3, 4, 5]\mu m$ and 4 crossing angles $\alpha = [22.5°, 45°, 67.5°, 90°]$ with density index $f = 0.5$ and we computed the root-mean-square (RMS) signal metric

$$\left\| S_{\text{appr}}\left(r, f, \alpha; \mathscr{P}\right) - S_{\text{intw}}\left(r, f, \alpha; \mathscr{P}\right) \right\|_{\text{RMS}} = \frac{1}{\sqrt{M}} \left(\sum_{i=1}^{M} \left(S_{\text{appr}}\left(r, f, \alpha; \mathbf{p}_i\right) - S_{\text{intw}}(r, f, \alpha; \mathbf{p}_i) \right)^2 \right)^{1/2},$$

over the M sequences of each HARDI shell $\mathscr{P}$ from Protocols A and B (Table 2.1).

2.2.5 Impact of the Approximation on Estimated Microstructural Parameters

The previous section seeks the order of magnitude characterizing the differences in DW-MRI signal between reference interwoven fascicles and the superposition approximation but it does not indicate how these signal discrepancies relate to underlying microstructural differences. In this section, we investigated whether we can accurately predict microstructural properties of interest in interwoven fascicles assuming only an approximate model made of the superposition of independent single fascicles, despite the signal differences incurred. Mathematically, for given reference microstructural properties Ω_{intw} and a fixed acquisition protocol $\mathscr{P}$, this consisted in solving the following microstructural estimation problem

$$\hat{\Omega}_{\text{appr}} = \underset{\Omega_{\text{appr}}}{\operatorname{argmin}}\, d\left(S_{\text{intw}}\left(\Omega_{\text{intw}}; \mathscr{P}\right); S_{\text{appr}}\left(\Omega_{\text{appr}}; \mathscr{P}\right) \right), \tag{2.2}$$

where $d(\cdot;\cdot)$ is a metric quantifying the discrepancy between the DW-MRI signals. We then examined the microstructural similarity between the reference and the estimated parameters $\hat{\Omega}_{\text{appr}} \approx \Omega_{\text{intw}}$ as an indicator of the quality of the superposition approximation.

Experiments I and II described below consisted in solving Problem (2.2) within varying subsets of the crossing-fascicle parameter space Ω_{cross}, with and without acquisition noise. Experiment III aimed at comparing the estimation errors due to the use of the superposition approximation to the errors caused by the presence of noise alone.

In our signal simulation framework, the effects of all parameters but v_1 and v_2 on the approximated signal S_{appr} are captured by Monte Carlo simulations, which lack a closed-form expression and prevent us from solving Problem (2.2) continuously. In Experiments I through III, we therefore resorted to a discrete, exhaustive search as illustrated in Figure 2.2, using a collection $\mathscr{D}_{\text{sing}}$ of 600 pre-simulated single-fascicle signals combining 20 radius indices r from 0.5µm to 10µm by increments of 0.5µm and 30 density indices f from 0.32 to 0.90 by steps of 0.02 along the fixed direction $\mathbf{u}_1$. Rotated-fascicle signals $S_{\text{sing}}(\mathbf{u}_2) = \mathscr{R}_\alpha\left[S_{\text{sing}}\right]$ along any direction $\mathbf{u}_2$ forming an angle α with $\mathbf{u}_1$ could then be obtained by interpolating the pre-computed signals on the inversely-rotated protocol $\mathscr{R}_{-\alpha}[\mathscr{P}]$, which we did for 119 values of α ranging from 1.5° to 178.5° by increments of 1.5°.

The use of Monte Carlo signals and this estimation strategy are however not representative of

most current microstructural estimation frameworks in which simplified closed-form formulas are fit to DW-MRI measurements using non-linear optimization, which is subject to pitfalls including the presence of multiple local minima (Jelescu et al., 2016) and sensitivity to the choice of an objective function or the fitting strategy (Sepehrband et al., 2016; Harms et al., 2017). In Experiment IV detailed below, we therefore studied the impact of the superposition approximation on the microstructural estimates of the DIAMOND model (Scherrer et al., 2016) which can handle multiple-fascicle configurations.

2.2.5.1 Experiment I: impact of the approximation on identical fascicles

We considered identical populations of axons both in the reference configurations Ω_{intw} and in the configurations Ω_{appr} over which the minimization in (2.2) was performed, which simplified the analysis while still providing general trends about the impact of the superposition approximation on the three main microstructural features characterizing crossing fascicles: the radius index, the density index and the crossing angle. This constraint was relaxed in Experiment II.

We selected the 24 reference interwoven-fascicle environments $\Omega_{\text{intw}} = \left(\mathbf{u}_1, r_{\text{ref}}, f_{\text{ref}}, \mathbf{u}_2 = \mathscr{R}_\alpha\left[\mathbf{u}_1\right], r_{\text{ref}}, f_{\text{ref}}, \nu_1 = 0.5\right)$ described in Section 2.2.4 and corrupted each interwoven-fascicle signal S_{intw} with Rician noise as $\tilde{S}_{\text{intw}} = \sqrt{\left(S_{\text{intw}}\left(\Omega_{\text{intw}}; \mathbf{p}_i\right) + \varepsilon_{I,i}\right)^2 + \varepsilon_{Q,i}^2}$, where $\varepsilon_{I,i}$ and $\varepsilon_{Q,i}$ are independent, Gaussian variables of zero mean and variance σ^2, for every acquisition i in the protocol, respectively modeling the acquisition noise in the in-phase (I) and quadrature (Q) channels. Additionally, 4 interwoven-fascicle configurations with larger radius index values $r = [6,7,8,9]$µm were also investigated at fixed density index $f = 0.5$ and crossing angle $\alpha = 67.5°$.

Problem (2.2) therefore becomes, for a fixed acquisition protocol $\mathscr{P}$,

$$(\hat{r}, \hat{f}, \hat{\alpha}) = \underset{r,f,\alpha}{\operatorname{argmin}}\, d\left(\tilde{S}_{\text{intw}}\left(r_{\text{ref}}, f_{\text{ref}}, \alpha_{\text{ref}}; \mathscr{P}\right); \underbrace{0.5 \cdot S_{\text{sing}}\left(\mathbf{u}_1, r, f; \mathscr{P}\right) + 0.5 \cdot S_{\text{sing}}\left(\mathscr{R}_\alpha\left[\mathbf{u}_1\right], r, f; \mathscr{P}\right)}_{S_{\text{appr}}(r,f,\alpha;\mathscr{P})}\right),$$

where the metric d is the negative log-likelihood of the Rician distribution. The problem was solved by exhaustive search over the $600 \times 191 = 114\,600$ pre-computed signals (Figure 2.2), where $\mathbf{u}_1$ and the plane in which $\mathbf{u}_1$ is rotated were assumed known. We defined the signal-to-noise ratio (SNR) as $1/\sigma$ and recorded the mean absolute error (MAE) on the estimated radius index $|\hat{r} - r_{\text{ref}}|$, density index $\left|\hat{f} - f_{\text{ref}}\right|$ and crossing angle $|\hat{\alpha} - \alpha_{\text{ref}}|$ over $N_{\text{reps}} = 100$ noise repetitions at each SNR level. The noise-free estimates corresponding to SNR $\to \infty$ were computed using the RMS metric to account for the asymptotic behavior of the Rician distribution (Gudbjartsson and Patz, 1995).

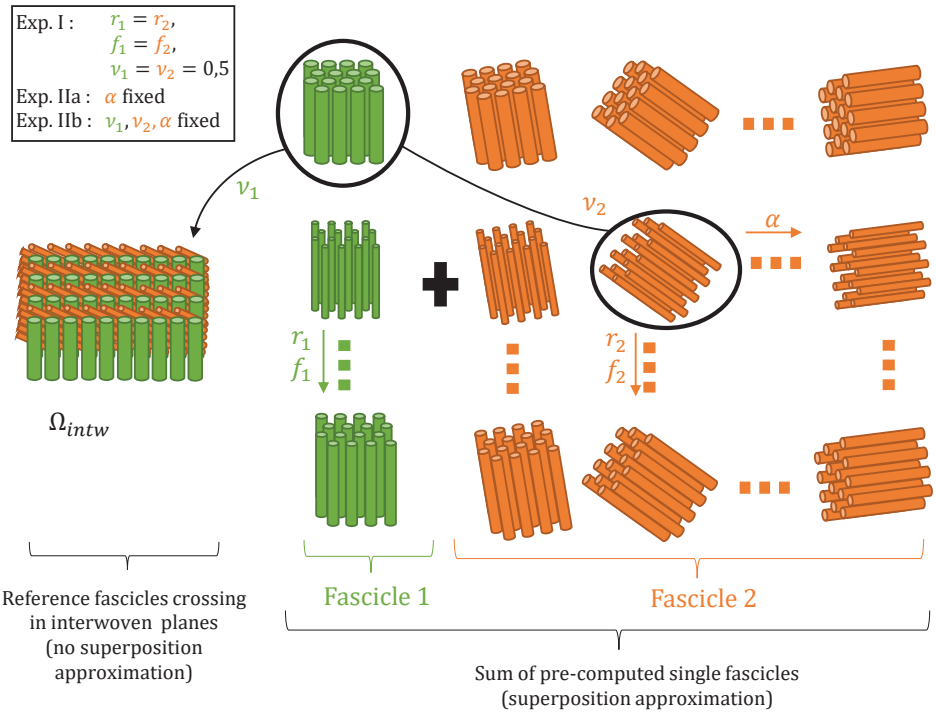

Figure 2.2: **Performing microstructural estimation with the superposition approximation using a discrete combinatorial search over pre-computed single-fascicle signals.** Solving Problem (2.2) is done by exhaustive search over the parameters $r_1, f_1, r_2, f_2, \alpha$ with constraints imposed at each experiment to reduce the scope of the exhaustive search. At fixed values of the latter parameters, continuous estimation over v_1 and v_2 is possible because they are the only parameters with an explicit, continuous contribution to the approximate signal S_{appr} via S_{sing}, as noted from Eq. (2.1). Note that $\mathbf{u}_1$ is assumed known throughout as well as the plane in which $\mathbf{u}_2$ lies.

2.2.5.2 Experiment II: impact of the approximation on dissimilar fascicles

In this experiment we relaxed the constraint that the crossing fascicles should be identical. To keep Problem (2.2) tractable, we first assumed in Experiment IIa that the crossing angle α was known, leaving 5 parameters to be estimated: $r_1, f_1, r_2, f_2, \nu_1$. In order to more fairly compare the results with Experiment I where the volume fractions were fixed, we then simplified the estimation in Experiment IIb by assuming that ν_1 and ν_2 were known as well.

In both cases, we conducted 8 noiseless microstructural estimation experiments on 8 reference interwoven-fascicle configurations Ω_{intw} in which the crossing angle was fixed to $\alpha = 67.5°$, the first population of axons had fixed parameters $r_1 = 1.0\,\mu\text{m}$ and $f_1 = 0.6$ and the microstructural properties of the second population were varied. In the first 4 configurations, we set $f_2 = f_1 = 0.6$ and let r_2 take on values in $[0.5, 1.0, 1.5, 2.0]\,\mu\text{m}$; in the last 4 configurations we set $r_2 = r_1 = 1.0\,\mu\text{m}$ and let f_2 vary in $[0.4, 0.5, 0.6, 0.7]$. The parameters e_1, e_2 and e_l were selected to ensure $f_{tot} = (f_1 + f_2)/2$ in all 8 reference configurations.

Experiment IIa. Problem (2.2) was solved by performing continuous optimization over the volume fractions for each possible combination of the 600 pre-computed single-fascicle configurations $\mathscr{D}_{\text{sing}}(k) \coloneqq S_{\text{sing}}(\mathbf{u}_1, r_k, f_k)$ with their rotated counterparts $\mathscr{R}_\alpha\left[\mathscr{D}_{\text{sing}}(l)\right]$,

$$(\hat{k}, \hat{l}) = \underset{1\le k,l\le 600}{\operatorname{argmin}} \quad \min_{x_1 \ge 0} \Big\| \underbrace{x_1 \cdot \mathscr{D}_{\text{sing}}(k) + (1 - x_1) \cdot \mathscr{R}_\alpha\left[\mathscr{D}_{\text{sing}}(l)\right]}_{S_{\text{appr}}} - S_{\text{intw}}(\Omega_{\text{intw}}) \Big\|_2^2 \,, \tag{2.3}$$

where each of the $600 \times 600 = 360\,000$ sub-problems admits a unique solution obtained with the MATLAB (MathWorks, MA, U.S.A.) routine `lsqlin`. We took ν_1 as the minimizer $\hat{x}_1$ of the optimal sub-problem $(\hat{k}, \hat{l})$ giving the lowest objective value and estimated r_1, f_1, r_2, f_2 as the microstructural properties $r_{\hat{k}}, f_{\hat{k}}, r_{\hat{l}}, f_{\hat{l}}$ of the corresponding optimal single-fascicle configurations $\mathscr{D}_{\text{sing}}(\hat{k})$ and $\mathscr{R}_\alpha\left[\mathscr{D}_{\text{sing}}(\hat{l})\right]$.

Experiment IIb. With ν_1 fixed, Eq. (2.3) becomes

$$(\hat{k}, \hat{l}) = \underset{1\le k,l\le 600}{\operatorname{argmin}} \Big\| \underbrace{\nu_1 \mathscr{D}_{\text{sing}}(k) + (1 - \nu_1) \mathscr{R}_\alpha\left[\mathscr{D}_{\text{sing}}(l)\right]}_{S_{\text{appr}}} - S_{\text{intw}}(\Omega_{\text{intw}}) \Big\|_2^2, \tag{2.4}$$

which was simply solved by exhaustive search over the 600×600 combinations of pre-computed single-fascicle configurations (see Figure 2.2).

2.2.5.3 Experiment III: impact of the approximation with noise compared to the impact of noise alone

Experiment I assessed the *combined* impact of the acquisition noise and of the "approximation noise" on the quality of the microstructural estimates. In this experiment we isolated the

impact of acquisition noise alone by performing a similar estimation but using the *true* model of interwoven fascicles for the estimation instead of the approximate model.

We focused on one particular reference interwoven-fascicle configuration with identical populations that led to a non-zero asymptotic error on at least one of the parameters r, f, α in Experiment I and we solved

$$(\hat{r}, \hat{f}, \hat{\alpha}) = \underset{r,f,\alpha}{\operatorname{argmin}}\, d\left(\tilde{S}_{\text{intw}}\left(r_{\text{ref}}, f_{\text{ref}}, \alpha_{\text{ref}}; \mathscr{P}\right); S_{\text{intw}}\left(r, f, \alpha; \mathscr{P}\right)\right), \tag{2.5}$$

where d is the negative log-likelihood of the Rician distribution and $\tilde{S}_{\text{intw}}$ is the reference interwoven-fascicle signal corrupted by noise. We compared the microstructural estimates obtained here to those obtained in Experiment I with the superposition approximation.

Similarly to the previous experiments, solving Problem (2.5) required a discrete, exhaustive search in the space (r, f, α). Pre-simulating the candidate signals is more computationally demanding than when using the superposition approximation because every new value of crossing angle α requires a complete new signal simulation. For the density index f and the crossing angle α, the same resolutions of 0.02 and 1.5° were kept, the ranges spanned by the exhaustive search were centered around the reference values f_{ref} and α_{ref} and their half-widths were selected to be slightly larger than the worst MAEs observed in Experiment I over all SNR values. For the radius index r we considered candidate values in the range $[0.5\,\mu\text{m}, 10\,\mu\text{m}]$ every $0.25\,\mu\text{m}$ for $|r - r_{\text{ref}}| \leq 0.5\,\mu\text{m}$, every $0.50\,\mu\text{m}$ for $1.0\,\mu\text{m} \leq |r - r_{\text{ref}}| \leq 4.0\,\mu\text{m}$ and every $1.0\,\mu\text{m}$ elsewhere.

2.2.5.4 Experiment IV: impact of the approximation on a closed-form model of the microstructure

We studied the impact of the superposition approximation on the parameters of the DIAMOND model (Scherrer et al., 2016) which represents the 3D-diffusivity of each voxel compartment j (e.g., a fascicle of axons) with a peak-shaped statistical distribution of diffusion tensors parameterized by a compartment heterogeneity index cHEI_j and a mean tensor $\mathbf{D}_{0,j}$. The contribution of compartment j to the total signal is weighted by an apparent volume fraction v_j. Compartment-specific diffusion characteristics such as the axial and radial diffusivities cAD_j and cRD_j are extracted from $\mathbf{D}_{0,j}$.

Specifically, the DIAMOND closed-form, continuous expression $S_{\text{DMD}}(\Omega_{\text{DMD}}; \mathscr{P})$ relating the DW-MRI signal S_{DMD} to the microstructural parameters Ω_{DMD} for a given acquisition protocol $\mathscr{P}$ was fitted to signals $S_{\text{intw}}(\Omega_{\text{cros}}; \mathscr{P})$ arising from reference interwoven configurations Ω_{cros} and to the signals $S_{\text{appr}}(\Omega_{\text{cros}}; \mathscr{P})$ arising from the superposition approximations of matching microstructural parameters. We obtained the model estimates $\hat{\Omega}_{\text{DMD}}\left(S_{\text{appr}}\right)$ and $\hat{\Omega}_{\text{DMD}}(S_{\text{intw}})$

by solving, for each of the 24 crossing-fascicle configurations Ω_{cros} previously described,

$$\hat{\Omega}_{\text{DMD}}\left(S_{\text{appr}}\right) = \underset{\Omega_{\text{DMD}}}{\text{argmin}}\, d\left(S_{\text{appr}}\left(\Omega_{\text{cros}};\mathscr{P}\right); S_{\text{DMD}}\left(\Omega_{\text{DMD}};\mathscr{P}\right)\right),$$
$$\hat{\Omega}_{\text{DMD}}\left(S_{\text{intw}}\right) = \underset{\Omega_{\text{DMD}}}{\text{argmin}}\, d\left(S_{\text{intw}}\left(\Omega_{\text{cros}};\mathscr{P}\right); S_{\text{DMD}}\left(\Omega_{\text{DMD}};\mathscr{P}\right)\right),$$

where d is a corrected least-squares metric and where the non-linear minimization was achieved with a customized BOBYQA algorithm (Scherrer et al., 2016). The impact of the signal discrepancies caused by the approximation was assessed by comparing $\hat{\Omega}_{\text{DMD}}\left(S_{\text{appr}}\right)$ to $\hat{\Omega}_{\text{DMD}}\left(S_{\text{intw}}\right)$ in all 24 cases for Protocols A and B.

The experiment was then repeated for the simpler NODDI model (Zhang et al., 2012), essentially designed to describe a single population of axons and which does not incorporate an explicit dependence on the axonal radius. More details are provided in Appendix A.4.2.

2.3 Results

This section first reports the observed DW-MRI signal differences between the reference interwoven fascicles and the superposition approximation at matching microstructural parameters. It then provides the results of the four microstructural estimation experiments assessing the impact of those signal differences on the underlying microstructural parameters.

2.3.1 Impact of the Approximation on the DW-MRI Signal

We found that the RMS difference between the approximate and groundtruth signal varied in $[0.0017, 0.024]$ over the 8 considered acquisition shells and 24 microstructural configurations, i.e. an order of magnitude comparable to the standard deviation of Gaussian noise in an MRI acquisition channel with SNR in $[1/0.024, 1/0.0017] \approx [42-603]$. The detailed shell-per-shell signal differences are available in Table A.1 in Appendix A.2. Figure 2.3 shows the DW-MRI signals of the particular scenario that yielded the highest RMS difference as a function of the direction of the applied magnetic gradient. The signal discrepancies seemed largest for gradients perpendicular to $\mathbf{u}_1$ and $\mathbf{u}_2$ but this varied from shell to shell, as illustrated in Figure A.3 in Appendix A.2. Little difference was found between the two protocols: the RMS metric averaged over all 24 microstructural configurations was 6.4×10^{-3} for Protocol A and 6.3×10^{-3} for Protocol B. As discussed in Appendix A.3, signal discrepancies obtained with the 50 additional shells were very similar to those obtained with Protocols A and B.

Figure 2.4 shows those same DW-MRI signals as a function of the microstructural parameters for a few selected acquisition sequences with magnetic gradient applied in the plane defined by the orientations of the two fascicles. It suggests that DW-MRI signals (both the reference and the approximation) exhibit heterogeneous degrees of sensitivity to the underlying microstructural parameters. In particular, the signals hardly varied at all in the range of smaller

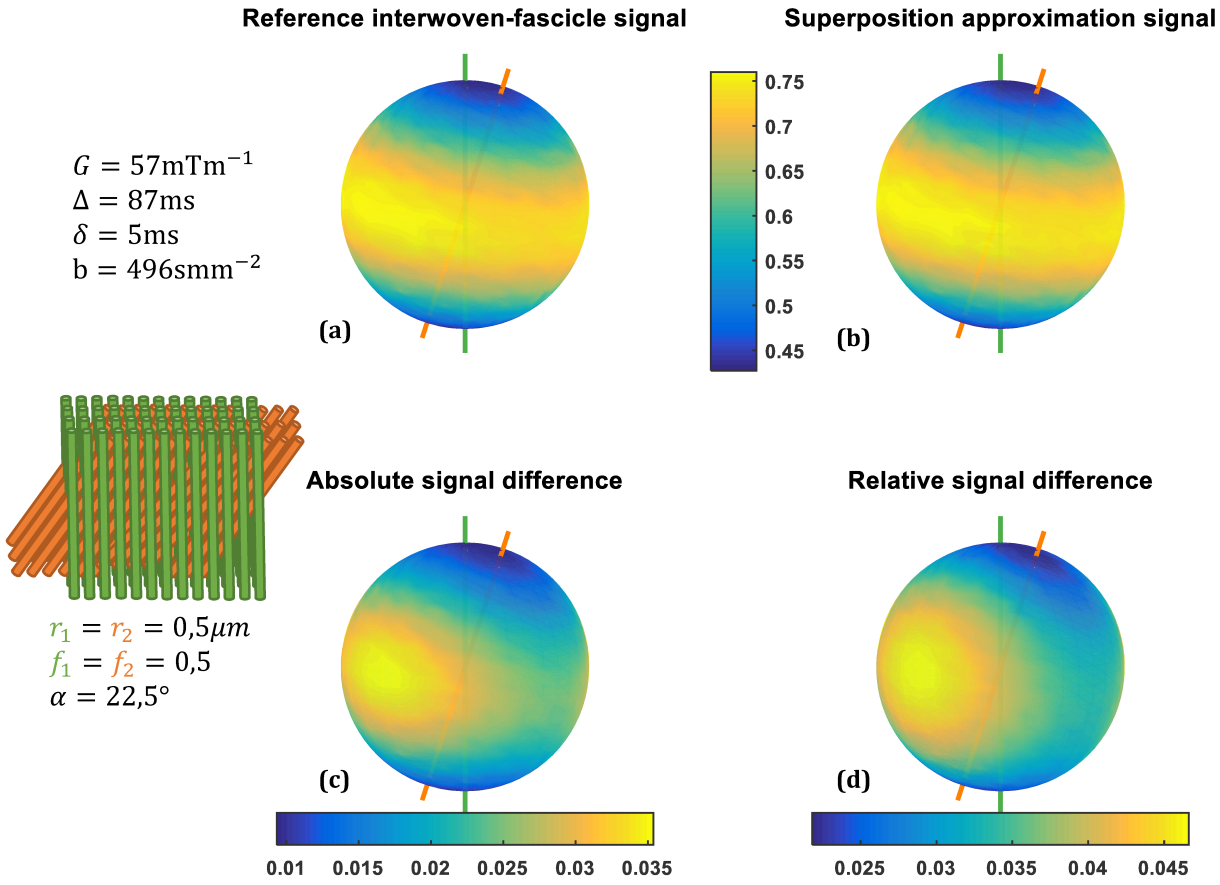

Figure 2.3: **DW-MRI signals from the superposition approximation closely match signals from interwoven fascicles.** Normalized DW-MRI signal attenuation ((a)-(b)) and differences ((c)-(d)) for the microstructural configuration and the HARDI shell that led to the highest RMS signal difference (RMS=2.4×10^{-2}), plotted on the 3D sphere as a function of the gradient direction $\hat{\mathbf{g}}$. Here the absolute (c) and the relative (d) differences were highest around the direction normal to the plane defined by the fascicles' orientations $\mathbf{u}_1$ and $\mathbf{u}_2$, where DW-MRI signals were highest too.

radius indices, which is a well-known limitation of the PGSE sequence (Clayden et al., 2015; Sepehrband et al., 2016; Drobnjak et al., 2016).

2.3.2 Impact of the Approximation on Estimated Microstructural Parameters

2.3.2.1 Experiment I: impact of the approximation on identical fascicles

Figure 2.5 shows the results of 6 of the 24 experiments with small reference radius indices, corresponding to the reference crossing angle $\alpha_{\text{ref}} = 67.5°$. The results for the other 3 angles were qualitatively similar and are provided in Figures A.5, A.6 and A.7 in Appendix A.4.1.

The superposition approximation led to a fast convergence with increasing SNR in the esti-

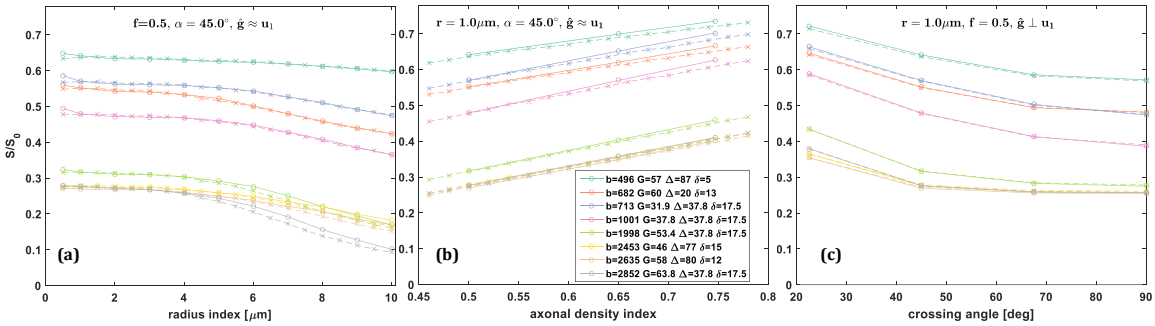

Figure 2.4: **DW-MRI signals from the superposition approximation closely match signals from interwoven fascicles but exhibit varying degrees of sensitivity to microstructural parameters.** Normalized DW-MRI signal attenuation of the reference interwoven-fascicle (continuous lines) and the superposition approximation (dashed lines) for one particular sequence extracted from each of the 8 HARDI shells reported in Table 2.1 with gradient direction $\hat{\mathbf{g}}$ selected in the plane defined by $\mathbf{u}_1$ and $\mathbf{u}_2$, parallel (a-b) or perpendicular (c) to $\mathbf{u}_1$. Units of b are in $\mathrm{s\,mm^{-2}}$, G in $\mathrm{mT\,m^{-1}}$ and Δ and δ in ms. We observed: (a) low signal sensitivity at low radius values, (b) generally high sensitivity across the whole range of density values and (c) generally high sensitivity across the whole range of crossing angle values.

mation of the crossing angle as the MAE over all 24 experiments consistently remained below 4.8° for all SNR ≥ 5 with Protocol A and below 7.1° with Protocol B, showing no sensitivity to the reference radius index. Given the resolution used in the discrete minimization, we deduced from the apparently zero asymptotic errors that the errors due to the superposition approximation in the absence of noise (SNR → ∞) were lower than 1.5°.

The convergence was fast as well for the density index as the worst-case MAE over all 24 reference configurations at SNR = 10 was 0.042 for Protocol A and 0.070 for Protocol B. Asymptotic errors attributable to the superposition approximation could be upper-bounded by 0.08 for reference configurations verifying $r_{\mathrm{ref}} \leq 1\,\mu\mathrm{m}$ and by 0.02 elsewhere.

The MAE on the estimated radius index was slower to level off to the asymptotic, noise-free errors. They were slightly larger for Protocol B at smaller radius indices, where they reached between 3.5 µm and 4.5 µm. The non-monotonicity of the red curves in Figure 2.5e-f corresponding to a reference radius index $r = 0.5\,\mu\mathrm{m}$ is due to both the absence of radius values smaller than 0.5 µm in the exhaustive minimization procedure (since the approximated signal hardly varies in that small-radius range) and to the asymptotic bias in the estimation, which artificially improved the estimation error of that parameter at low SNR values.

The variability of the microstructural estimates over the noise repetitions can help reveal the SNR regimes in which the errors due to the noise dominate the errors due to the use of the superposition approximation: a large variability compared to the asymptotic, noise-free error suggests that noise is the predominant source of error. The radius index estimates obtained with the superposition approximation using Protocol A for instance exhibited an interquartile range (IQR), defined as the difference between the 75th and the 25th percentile, that still

represented 50% or more of the asymptotic error in 17 of the 18 configurations yielding a non-zero asymptotic error at SNR=50, before dropping to 7 configurations at SNR=100 and 2 at SNR=200. This suggests a prevalence of noise-related errors for SNR levels up to about 50, as illustrated in Figure 2.5g-h for one particular experiment.

The results of the 4 configurations with larger radius index values are provided in Figure A.8 in Appendix A.4.1. The estimation of the crossing angle exhibited trends similar to those obtained with the smaller reference radii except for asymptotic, noise-free errors appearing with Protocol B. The estimation of the density index and the radius index generally benefited from larger reference radii.

2.3.2.2 Experiment II: impact of the approximation on dissimilar fascicles

The results of all 8 estimation experiments on dissimilar populations of axons are presented in Table 2.2 with (Experiment IIa, in blue) and without (Experiment IIb, in green) the estimation of the volume fractions, in the absence of noise.

Experiment IIa. The errors caused by the approximation without acquisition noise suggest an inter-dependence between the estimation of the volume fraction v and density index f of a fascicle: when v was overestimated, f was underestimated and conversely. This conflating effect seemed exacerbated at small radius index values but slightly less pronounced when the two fascicles occupied similar fractions of the voxel: with Protocol A for instance, the approximation yielded a maximum error on the estimated density index between 0.18 and 0.22 with the smallest radius indices while the largest error over the last 4 cases lay between 0.12 and 0.16 given the 0.02 resolution used in the discrete optimization. In general, the superposition approximation yielded microstructural errors larger than those observed without noise in Experiment I where the volume fractions were equal and known.

Experiment IIb. Fixing the volume fractions v of the fascicles a priori considerably improved the quality of the microstructural estimation using the superposition approximation, containing the error on the estimated density index within 0.04 given our 0.02 discrete resolution, in agreement with the noise-free results of Experiment I.

2.3.2.3 Experiment III: impact of the approximation with noise compared to the impact of noise alone

Figure 2.6 presents the results of the microstructural estimation of the reference interwoven-fascicle configuration characterized by $r_{\text{ref}} = 0.5\mu\text{m}, f_{\text{ref}} = 0.5, \alpha_{\text{ref}} = 45°$. The exhaustive search over the parameters of the groundtruth interwoven fascicles used 5202 pre-simulated signals and provided an estimate of the errors due to noise alone (green) to be compared to the results from Experiment I obtained using the superposition approximation (blue).

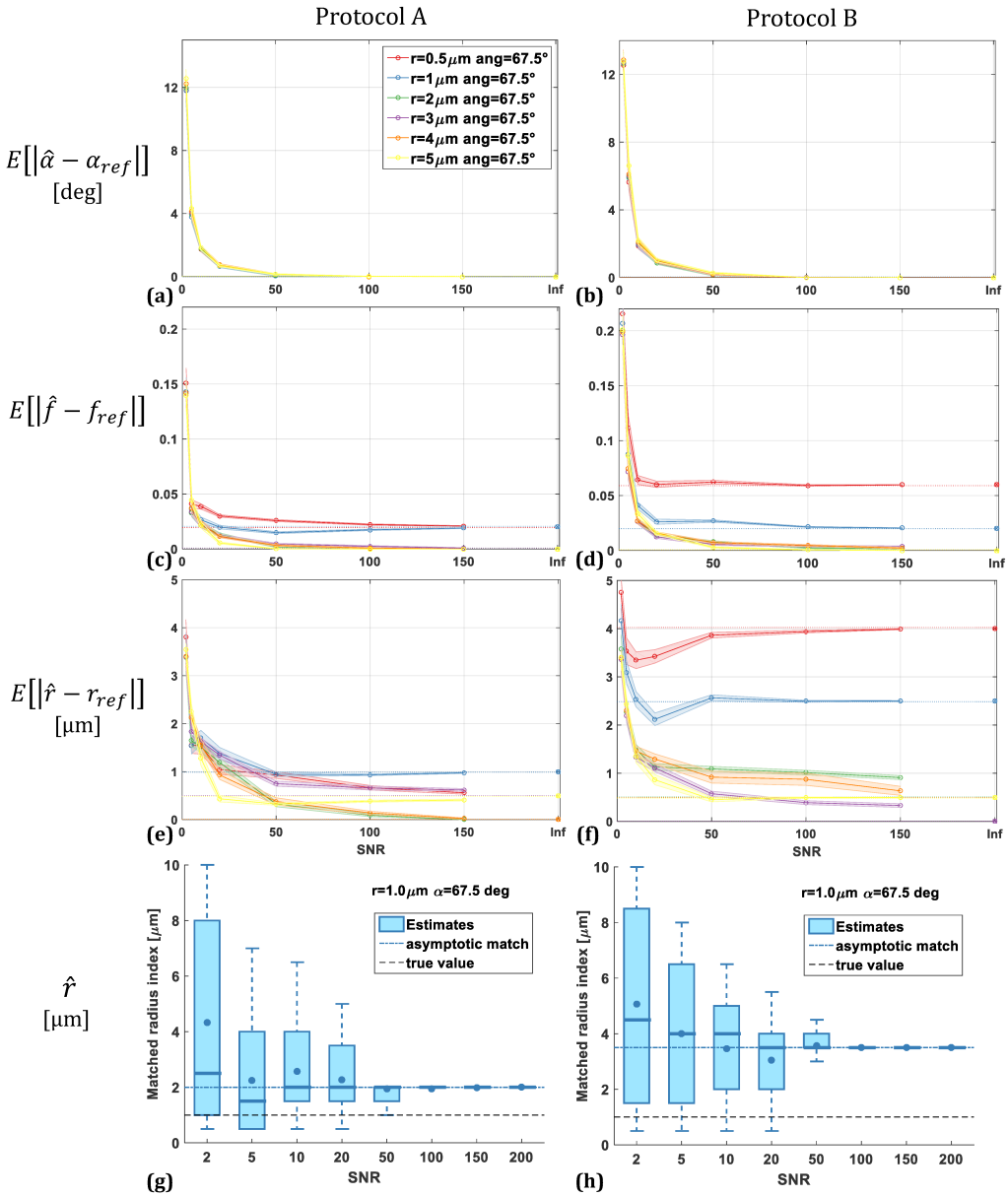

Figure 2.5: **The superposition approximation shows fast convergence with increasing SNR and small asymptotic errors for the estimation of the crossing angle and the density index but a slower convergence with larger asymptotic errors for the radius index.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. The dashed lines are the asymptotic, noise-free errors slightly offset around their true values for visualization purposes. Sub-figures (g) and (h) display box-plots of the radius index estimates corresponding to the blue curve in sub-figures (e)-(f), where large interquartile ranges relative to the asymptotic bias up to SNR=50 suggest that the acquisition noise is the predominant source of error rather than the use of the superposition approximation.

Table 2.2: **Microstructural estimation from the superposition approximation on dissimilar fascicles in the absence of noise.** The values in black are the reference values of the interwoven-fascicle configurations. The errors in Experiment IIa (in blue, obtained with the discrete-continuous optimization scheme (2.3)) are slightly larger than the asymptotic, noise-free errors observed in Experiment I. In Experiment IIb (green, with the volume fractions of the fascicles v_1 and v_2 known a priori following Eq. (2.4)), the estimated parameters are more in line with the asymptotic estimates of Experiment I, where the populations of axons had identical properties.

Protocol A

r_1 [µm]			f_1			r_2 [µm]			f_2			v_1			v_2		
Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb
1.0	2.5	1.0	0.60	0.52	0.60	0.5	1.5	3.0	0.60	0.80	0.58	0.65	0.74	(0.65)	0.35	0.26	(0.35)
1.0	2.0	0.5	0.60	0.52	0.60	1.0	0.5	1.0	0.60	0.70	0.60	0.50	0.57	(0.50)	0.50	0.43	(0.50)
1.0	0.5	1.0	0.60	0.70	0.58	1.5	2.0	1.0	0.60	0.54	0.60	0.41	0.36	(0.41)	0.59	0.64	(0.59)
1.0	0.5	3.0	0.60	0.70	0.58	2.0	2.0	1.0	0.60	0.54	0.60	0.35	0.29	(0.35)	0.65	0.71	(0.65)
1.0	1.0	1.0	0.60	0.74	0.58	1.0	1.5	0.5	0.40	0.34	0.40	0.50	0.41	(0.50)	0.50	0.59	(0.50)
1.0	0.5	1.0	0.60	0.70	0.58	1.0	3.0	0.5	0.50	0.44	0.50	0.50	0.43	(0.50)	0.50	0.57	(0.50)
1.0	2.0	0.5	0.60	0.52	0.60	1.0	0.5	1.0	0.60	0.70	0.60	0.50	0.57	(0.50)	0.50	0.43	(0.50)
1.0	2.5	1.0	0.60	0.54	0.60	1.0	0.5	0.5	0.70	0.80	0.70	0.50	0.56	(0.50)	0.50	0.44	(0.50)

Protocol B

r_1 [µm]			f_1			r_2 [µm]			f_2			v_1			v_2		
Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb	Ref	IIa	IIb
1.0	1.5	1.0	0.60	0.54	0.60	0.5	0.5	2.5	0.60	0.76	0.60	0.65	0.73	(0.65)	0.35	0.27	(0.35)
1.0	3.0	2.5	0.60	0.58	0.62	1.0	2.5	2.0	0.60	0.66	0.60	0.50	0.53	(0.50)	0.50	0.47	(0.50)
1.0	1.5	2.5	0.60	0.52	0.62	1.5	3.0	2.0	0.60	0.70	0.60	0.41	0.48	(0.41)	0.59	0.52	(0.59)
1.0	3.0	0.5	0.60	0.70	0.60	2.0	0.5	2.0	0.60	0.56	0.60	0.35	0.31	(0.35)	0.65	0.69	(0.65)
1.0	1.0	1.5	0.60	0.64	0.60	1.0	3.5	3.5	0.40	0.40	0.42	0.50	0.47	(0.50)	0.50	0.53	(0.50)
1.0	1.0	1.5	0.60	0.64	0.60	1.0	0.5	1.0	0.50	0.48	0.50	0.50	0.47	(0.50)	0.50	0.53	(0.50)
1.0	3.0	2.5	0.60	0.58	0.62	1.0	2.5	2.0	0.60	0.66	0.60	0.50	0.53	(0.50)	0.50	0.47	(0.50)
1.0	4.0	2.5	0.60	0.58	0.62	1.0	0.5	0.5	0.70	0.80	0.70	0.50	0.56	(0.50)	0.50	0.44	(0.50)

The true model and the approximate model yielded curves that could hardly be distinguished for the estimation of the crossing angle, as could be expected since no asymptotic error caused by the superposition approximation was found in Experiment I. The MAE of the two models on the density index differed by less than 0.022 for all SNR ≥ 5. The median estimates of the approximation systematically exceeded those of the true model by 0.02 across all SNR levels, consistent with the asymptotic errors detected in Experiment I. For the radius index, the two estimations were very close for SNR values up to 20 with a difference in MAEs representing

less than 25% of the asymptotic error. The bias due to the superposition approximation manifested itself after SNR=20 and became clearly apparent at SNR=50, in agreement with the observations made in Experiment I.

2.3.2.4 Experiment IV: impact of the approximation on a closed-form model of the microstructure

As shown in Figure 2.7, the DIAMOND parameters cHEI, cRD and cAD estimated from the superposition approximation were very close to those obtained from the interwoven-fascicle signals, with maximum differences of respectively 1.3×10^{-2}, $6.4 \times 10^{-6}\,\mathrm{mm^2\,s^{-1}}$ and $6.9 \times 10^{-5}\,\mathrm{mm^2\,s^{-1}}$ across all 48 experiments. The estimated apparent volume fraction v_1 differed by less than 0.0575 in 47 of the 48 tested cases and the errors on the crossing angles (not shown) were all below 2°. All the curves in Figure 2.7 exhibited similar heights as expected since both crossing-fascicle configurations had identical fascicles, irrespective of the crossing angle. Their relative flatness indicated low sensitivity to the reference radius index.

The results for the parameters of NODDI are shown and discussed in Appendix A.4.2 (see Figure A.9) and generally suggested little change associated with the use of the superposition approximation.

2.4 Discussion and Conclusions

This paper examined the validity of approximating the signal arising from fascicles crossing in interwoven planes by the superposition of signals arising from independent single fascicles through the use of Monte Carlo simulations.

The mean normalized signal differences between interwoven fascicles and the superposition approximation were reported in the range $10^{-3} - 10^{-2}$, suggesting that SNRs of the order of 100 would be necessary for those signal differences to become significant and to detect whether fascicles intermingle when they cross, which is not achievable with current MRI technology and clinically acceptable imaging times.

We considered two clinical protocols with b-values lower than $3000\,\mathrm{s\,mm^{-2}}$ which obtained very similar results: the average signal differences were within 1×10^{-4} of each other and the performances were nearly identical in all microstructural estimation experiments with $r > 1\,\mu\mathrm{m}$. As investigated through Monte Carlo simulations in (Raffelt et al., 2012), for larger b-values the signal differences are likely to be *even less* since the extracellular signal, the main source of signal discrepancy in our simulation setting, essentially decays away. This was confirmed in the additional experiments reported in Appendix A.3.

Our experiments have shown that the superposition approximation enables excellent estimation of the crossing angle between two populations of axons even in noisy settings, which supports the results obtained in the estimation of fascicle orientation relying on the super-

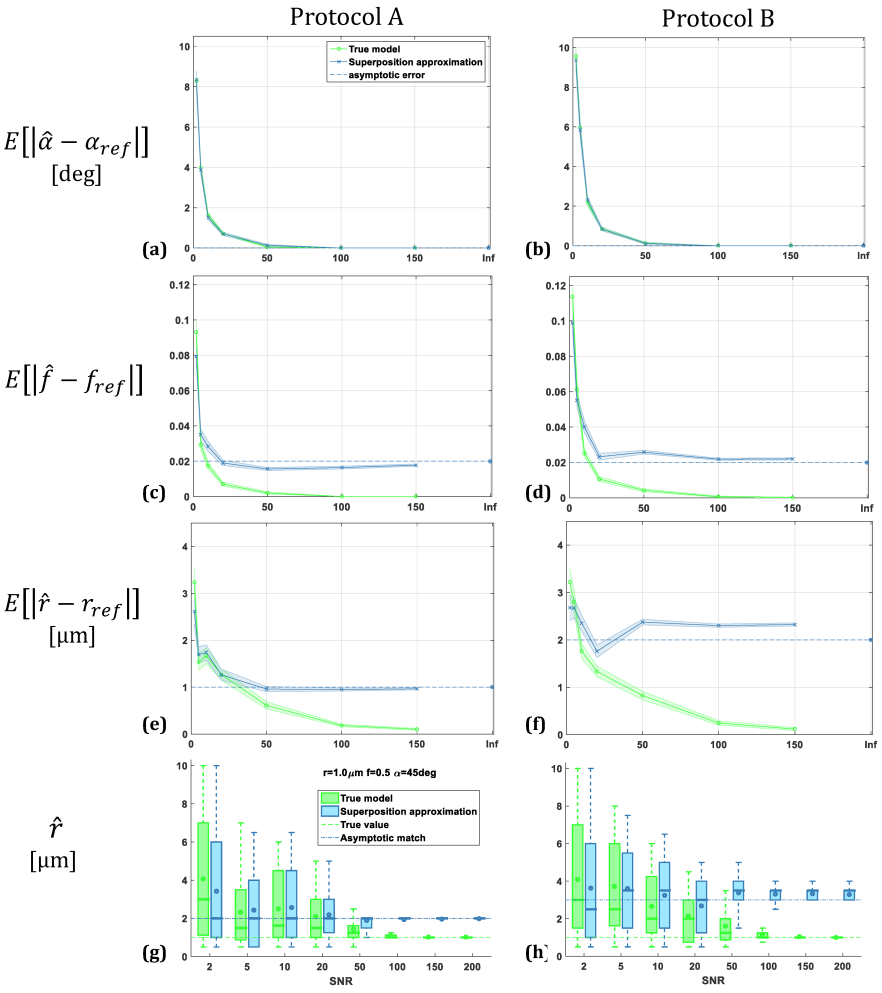

Figure 2.6: **The use of the superposition approximation has little impact at clinical SNR levels where the errors caused by noise dominate.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. Sub-figures (g) and (h) display detailed box-plots of the distribution of the radius index estimates over all the noise repetitions at each SNR level suggesting that the errors due to noise alone dominate those due to the use of the superposition approximation at lower SNR levels.

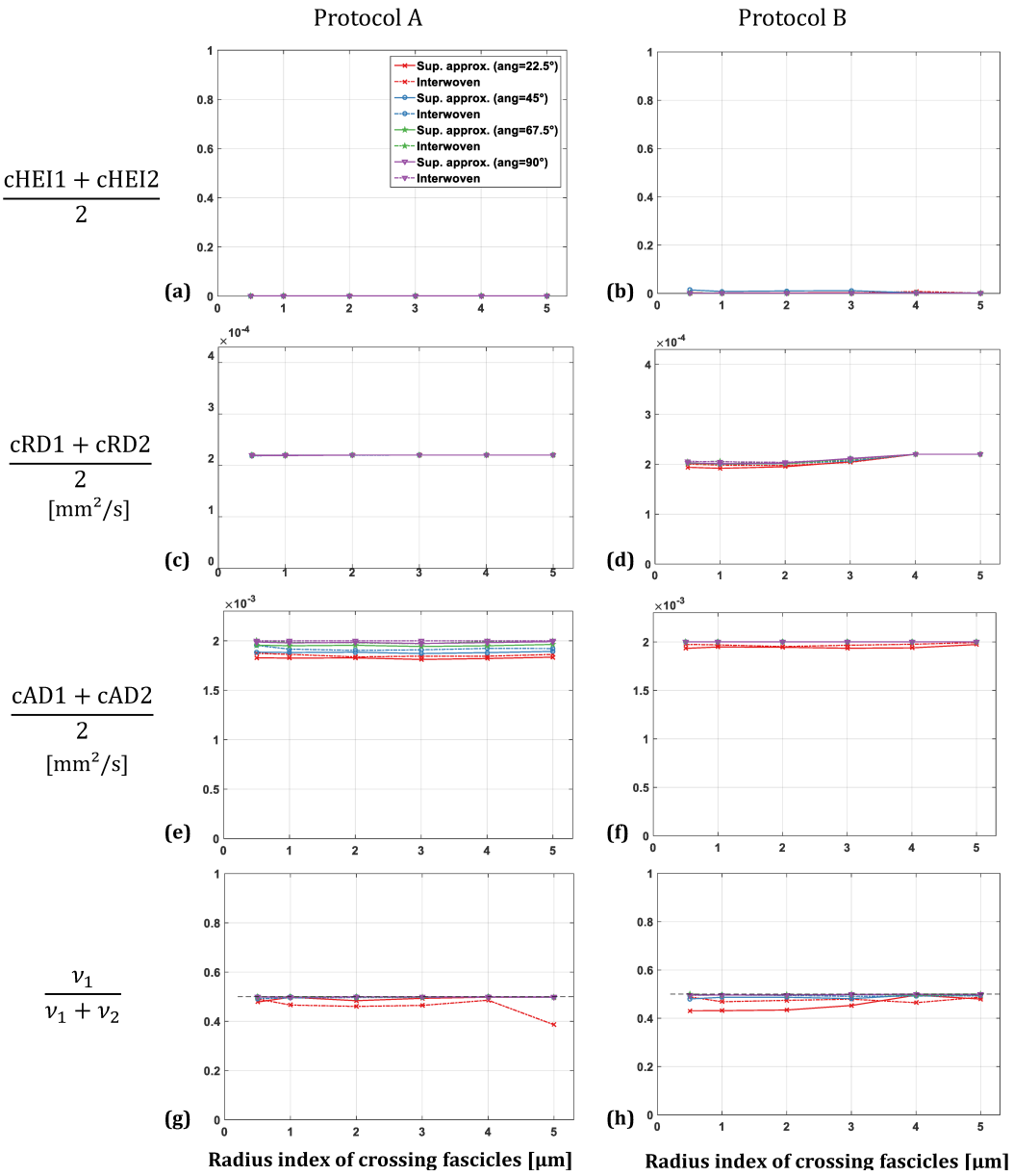

Figure 2.7: **The signal discrepancies of the superposition approximation have a limited impact on DIAMOND parameters.** Parameters obtained when fitting the DIAMOND model with 2 voxel compartments (one for each fascicle of axons) to the approximate signal (continuous lines) and the reference interwoven-fascicle signal (dashed lines). (a)-(b) Mean compartment heterogeneity index cHEI. (c)-(d) Mean compartment radial diffusivity cRD. (e)-(f) Mean compartment axial diffusivity cAD. (g)-(h) Normalized apparent volume fraction of occupancy of the first fascicle, where the horizontal dotted line represents the reference axonal density index $f_{\text{ref}} = 0.5$ of the crossing-fascicle configurations.

position hypothesis, such as sums of tensors (Tuch et al., 2002; Scherrer and Warfield, 2012; Scherrer et al., 2016) or tensors with zero radial diffusivity as in most spherical deconvolution frameworks (Tournier et al., 2004, 2007; Jeurissen et al., 2014; Canales-Rodríguez et al., 2015).

The use of the superposition approximation generally yielded high accuracy and small systematic biases on the axonal density index, also preserved with noise. Confounding effects were observed when simultaneously estimating the populations' axonal density indices and volume fractions (Experiment II), certainly because very similar signals can be reconstructed by assigning a higher weight v to a single-fascicle signal with lower intrinsic axonal density index f and conversely. This is reminiscent of the signal equivalence between multi-tensor models at a single b-value when changing the diffusivities of each tensor while correspondingly scaling the volume fractions (Scherrer and Warfield, 2012; Taquet et al., 2015). This mathematical redundancy may be inherent to crossing fascicles in general, whether an approximation is used or not.

Errors in the estimation of the radius index were recurrently observed throughout the experiments, especially at small radius indices. This was mainly driven by the notoriously low sensitivity of PGSE DW-MRI signals to small axonal radii (Clayden et al., 2015; Sepehrband et al., 2016; Drobnjak et al., 2016; Nilsson et al., 2017) which we observed in Figure 2.4 with and without the use of the approximation. The results of Experiments I and III suggested that the acquisition noise was the predominant source of error and that the use of the superposition approximation had little impact at low, and thus more realistic, SNR levels. At higher SNR and in noiseless scenarios however, the approximation showed a limitation since the error due to the use of the approximate model persisted while that due to noise alone vanished. Considering a simpler microstructural model without an explicit dependence on the radius in Experiment IV showed uniform estimates across radius indices.

Experiment IV suggested that more phenomenological diffusion models of the microstructure were hardly affected by the use of the superposition approximation, further supporting the quality of the approximation.

Limitations

The synthetic phantoms of the microstructure that we considered in this study did not take axonal membrane permeability into account, which would make the intracellular signal contribution another source of discrepancy between the reference and approximate signal. It is however still unclear whether this strongly affects DW-MRI signals obtained with diffusion times typically no longer than 100 ms, such as considered in our analyses. In (Nilsson et al., 2013), intracellular exchange times in the brain were reported between 25 and 620 ms while in vivo estimates in (Nedjati-Gilani et al., 2014) lay in 400 – 600 ms in the genu and splenium of the corpus callosum and in 300 – 500 ms in the corticospinal tract, for instance.

Another limitation of this work is the simplicity of the two-compartment model adopted at

the single-fascicle level. However, this study focused on evaluating the impact of using a *superposition* model rather than selecting the right model at the single-fascicle level. For reasonably simple interweaving patterns of axons in the groundtruth configurations, it should be possible to design a model of single fascicle such that the intracellular compartment of the superposition faithfully mirrors that of interwoven fascicles, as was done in this work. In the absence of membrane permeability, similar conclusions would therefore likely hold for models incorporating more realistic biological features such as myelin sheaths, radius heterogeneity, in-plane and out-of-plane axonal undulation or the presence of glial cells or other cellular compartments in the interstitium.

A myelin compartment in particular should be of limited impact due to rapid T_2-decay (Alonso-Ortiz et al., 2015) causing a roughly zero signal for the acquisition sequences considered in this work using standard MRI equipment ($B_0 \leq 3\,\mathrm{T}$). The reference as well as the approximate signal curves would consequently undergo identical downward translations. Since this would preserve the same signal differences and have little effect on the general sensitivity of signals to microstructural parameters, the conclusions of this study would likely hold.

Likewise, modeling fascicles with a fixed axonal radius index seems physically less realistic than considering a random packing with axonal radius heterogeneity, which would considerably impact the configuration of the extracellular space. However it is known that this index captures important properties of the underlying distribution of radii through its relation with the moments of that distribution (Alexander et al., 2010; Burcaw et al., 2015), simplifying our analyses while still providing meaningful results about the estimation of that microstructural feature. Future studies may have to specifically assess the impact of this simplification.

Conclusion

Overall, we have observed a strong correspondence between DW-MRI signals arising from configurations of interwoven fascicles and DW-MRI signals arising from a weighted sum of independent fascicles for a wide range of realistic microstructural indices and commonly-used acquisition protocols. Even though the approximation might negatively impact the estimation of small radius indices if high imaging SNR becomes available, our experiments have suggested that using a superposition model enables microstructural properties of interest to be accurately estimated in the presence of clinically-realistic levels of noise corruption, irrespective of the exact configuration of the axons.

3 Microstructure fingerprinting

This Chapter is for the most part a post-print version of Rensonnet et al. (2019). Some of the theoretical sections are redundant with the theory presented in Chapter 1; they were left to ensure each Chapter is self-contained.

3.1 Introduction

Adequate modeling of the diffusion-weighted magnetic resonance imaging (DW-MRI) signal holds the promise of characterizing white matter tissues at the micrometer level, including information about the principal orientation of axons, their average radius or packing density. Traditionally, a *forward signal model* is formulated relating the DW-MRI signal measured in a voxel to the externally-applied magnetic field gradient profile and the microstructural properties of the tissue. The estimation stage or *inverse problem* generally consists in minimizing a cost function incorporating the measured DW-MRI data and the closed-form forward formula.

Estimating microstructural features from closed-form mathematical expressions of the signal poses three major limitations. First, the formulations are generally so complex that simplifying assumptions need to be incorporated to obtain closed-form formulas. For example, representing intra-axonal restriction for molecules trapped inside a simple model of straight cylinders often requires assuming a Gaussian phase distribution (McCall et al., 1963; Vangelderen et al., 1994) or a short gradient pulse (Tanner and Stejskal, 1968; Callaghan, 1995) to obtain analytical formulas. Describing the complex diffusion in the extra-axonal space almost exclusively relies on the basic diffusion tensor (Basser et al., 1994), albeit with possible refinements such as considering a peak-shaped distribution of tensors (Scherrer et al., 2016, 2017) or a dependence on diffusion time (Burcaw et al., 2015; Ning et al., 2017) or gradient frequency (Xu et al., 2014) when more complex DW-MRI sequences are used. Coupling the intra- and extra-axonal models in a physically consistent way is usually addressed by tortuosity models (Whitaker, 1967; Szafer et al., 1995), the accuracy of which has been questioned (Lampinen et al., 2017). Second, generalizing models from the diffusion sequence they were originally designed for, often the pulse-gradient spin-echo (PGSE) (Stejskal and Tanner, 1965),

to sequences such as double diffusion encoding (DDE) (Cory et al., 1990; Callaghan and Manz, 1994) or oscillating-gradient spin-echo (OGSE) (Gross and Kosfeld, 1969) generally requires non-trivial modifications (Xu et al., 2014; Lam et al., 2015). Third, the cost function used in the inverse problem is often highly non-linear in the parameters to estimate, leading to the well-known pitfalls of non-linear optimization: convergence to local rather than global minima, multiple equivalent minima (Jelescu et al., 2016; Novikov et al., 2018b), long fitting times or sensitivity to the initialization strategy (Harms et al., 2017). These difficulties are usually overcome by making further simplifications to reduce complexity and stabilize the estimation, such as assuming axons with zero radius (Zhang et al., 2012) or neglecting fascicle crossings altogether (Alexander et al., 2010; Zhang et al., 2012). A recent alternative to such simplifications for stabilizing the estimation is to recast the inverse problem into a convex optimization program on a pre-computed dictionary. This was proposed by Daducci et al. (2015) and Sepehrband et al. (2016) for single fascicles and by Auría et al. (2015b) for multiple fascicles. It should be noted that some models target the ensemble average propagator (EAP) of water molecules rather than tissue properties directly (Özarslan et al., 2013; Ning et al., 2017). Those frameworks make little to no assumption about the tissue and the signal is linearly decomposed in a chosen functional basis. The estimation of the coefficients of the expansion is done efficiently through convex optimization under positiveness (Özarslan et al., 2013) or sparsity constraints (Merlet et al., 2013) for instance. These models provide indices that must then be interpreted in terms of specific tissue features (Avram et al., 2016; Zucchelli et al., 2016).

Monte Carlo simulations of the random walk of water molecules have the potential to provide near ground-truth forward signals for any type of gradient profile and for any geometry of the cellular environment (Hall and Alexander, 2009; Balls and Frank, 2009). Because they provide numerical results rather than a closed-form formula for the signal, they are not well suited for parameter estimation through traditional continuous optimization. As a consequence, Monte Carlo simulations have been widely used in the validation of simpler closed-form models (Grebenkov, 2007; Fieremans et al., 2008; Nilsson et al., 2009; Hall and Alexander, 2009; Panagiotaki et al., 2010; Fieremans et al., 2010; Nilsson et al., 2010; Drobnjak et al., 2011; Raffelt et al., 2012; Dyrby et al., 2013; Pizzolato et al., 2015; Lam et al., 2015; Clayden et al., 2015; McHugh et al., 2015; Burcaw et al., 2015; Scherrer et al., 2016; Ianuş et al., 2016; Kakkar et al., 2017; Vellmer et al., 2017; Ning et al., 2017; Ginsburger et al., 2018; Mercredi and Martin, 2018). They have also been used to investigate the sensitivity of the DW-MRI signal to complex biophysical features (Hall and Alexander, 2009; Nilsson et al., 2012; Harkins and Does, 2016; Palombo et al., 2017; Rensonnet et al., 2018; Lin et al., 2018). However, their use for the direct estimation of microstructural properties has been scarce and thus far limited to areas containing single fascicles of axons. For instance, in Nilsson et al. (2010), nearest-neighbor matching from a collection of Monte Carlo signals was compared with an analytical model of diffusion considering permeable membranes. In Nedjati-Gilani et al. (2017), a similar analytical model was assessed against a random forest regressor used to learn a mapping between microstructural properties of single fascicles and orientationally-invariant features

extracted from their Monte Carlo signals. In the context of MR spectroscopy, Monte Carlo simulations were integrated in a computational model to assess the intracellular diffusion of cell-specific metabolites in rodent and primate brain (Palombo et al., 2016).

The framework proposed in this Chapter aims at exploiting the accuracy and interpretability of Monte Carlo simulations directly in the forward signal model and not solely as a validation tool. Single-fascicle DW-MRI signals or *fingerprints* are first pre-computed for a large collection of microstructural configurations using Monte Carlo simulations. The final multi-fascicle estimation then consists in selecting the optimal sparse combination of these fingerprints, which is done by solving many small convex sub-problems. The approach is applied to the estimation of an apparent axonal radius index and an axonal density index in single- as well as in crossing-fascicle configurations, using synthetic and *in vivo* data.

3.2 Theory

This section first presents our general multi-fascicle signal model incorporating signals from Monte Carlo simulations, the theory of which is reviewed in a second subsection. A third subsection presents mathematical properties that allow large collections of Monte Carlo signals to be obtained at a minimal computational cost. Finally, the inverse problem of estimating microstructural features is formulated as a structured sparse optimization problem in the last subsection.

3.2.1 Signal model

The DW-MRI signal S at echo time TE under the application of an effective magnetic field gradient profile $\mathbf{g}(t)$ $(0 \leq t \leq \mathrm{TE})$ in a voxel of white matter is assumed to arise from the independent contributions of K fascicles of axons with principal unit orientation $\mathbf{u}_1, \ldots, \mathbf{u}_K$ occupying fractions $v_1, \ldots, v_K$ of the physical volume of the voxel and of a partial volume v_{csf} of cerebrospinal fluid (CSF)

$$\begin{aligned} S &= M_0 \cdot \left[\sum_{k=1}^{K} v_k A_{\mathrm{fasc}}\left(\mathbf{\Omega}_k, \mathbf{T}_k, \mathbf{u}_k; \mathbf{g}\right) + v_{\mathrm{csf}} A_{\mathrm{csf}}\left(D_{\mathrm{csf}}, \mathbf{T}_{\mathrm{csf}}; \mathbf{g}\right) \right] \\ &= \sum_{k=1}^{K} w_k A_k + w_{\mathrm{csf}} A_{\mathrm{csf}}, \end{aligned} \tag{3.1}$$

where the scaling factor M_0 captures the net initial transverse magnetization of the voxel detected by the scanner and $w_k \coloneqq M_0 v_k$ is the NMR-apparent signal weight of the contribution of the k-th fascicle. The normalized DW-MRI signal $A_k \coloneqq A_{\mathrm{fasc}}\left(\mathbf{\Omega}_k, \mathbf{T}_k, \mathbf{u}_k; \mathbf{g}\right)$ of the k-th fascicle is modeled by a Monte Carlo simulation of the random self-diffusion of molecules in an environment characterized by the set of microstructural parameters $\mathbf{\Omega}_k$, typically featuring geometrical arrangements of cylinders representing axons. The set of parameters $\mathbf{T}_k$ captures NMR relaxation such as T1, T2 and proton density, which is generally assumed to occur

independently of diffusion and therefore only affects A_k via a multiplying factor. Water is assumed to diffuse freely and isotropically in the CSF compartment; its normalized DW-MRI contribution A_{csf} is therefore characterized by a scalar diffusivity D_{csf}. Since $v_1+\cdots+v_K+v_{\text{csf}}=1$, the physical volume fractions v_k have a one-to-one correspondence with the NMR-apparent weights w_k:

$$w_k = M_0 v_k \Leftrightarrow v_k = \frac{w_k}{\sum_{k=1}^{K+1} w_k}, \qquad k = 1,\ldots,K+1, \tag{3.2}$$

where the index $k = K+1$ refers to the CSF compartment.

The flexibility of Monte Carlo simulations allows any type of gradient profile $\mathbf{g}(t)$ to be used, including PGSE, DDE, OGSE or more general b-tensor encoding (Topgaard, 2017) without the need to mathematically derive a new signal model.

A vector $\left[A_{\text{fasc}}\left(\mathbf{\Omega},\mathbf{T},\mathbf{u};\mathbf{g}_i\right)\right]_{i=1}^{M}$ corresponding to the set of M gradient profiles $\left\{\mathbf{g}_i(t)\right\}_{i=1}^{M}$ making up the acquisition protocol is defined as a *fingerprint*. It uniquely relates to the particular microstructural parameters $\mathbf{\Omega}$ for given relaxation parameters $\mathbf{T}$ and orientation $\mathbf{u}$.

3.2.2 Monte Carlo simulations

Monte Carlo simulations consist in sampling the distribution of the phase $\phi\left(\mathbf{g}\right)$ accumulated at echo time TE by spin-bearing nuclei or *spins* undergoing random diffusion in a spatial domain Ω when a diffusion-sensitizing gradient profile $\mathbf{g}(t)$ $(0 \le t \le \text{TE})$ is applied. As described in Hall and Alexander (2009), a large set of N_{spin} Brownian trajectories $\mathbf{r}_l(t)$, $1 \le l \le N_{\text{spin}}$, are approximated in the environment Ω by discrete trajectories $\hat{\mathbf{r}}_l$ consisting of spatial jumps with random orientation and length

$$L_{\text{step}} = \sqrt{2nD\delta t} \tag{3.3}$$

where the time step $\delta t = \text{TE}/N_{\text{step}}$ is chosen in order to ensure a small bias, D is the local diffusivity and n $(1 \le n \le 3)$ is the intrinsic spatial dimension of the diffusion process of interest. If Ω consists of parallel straight cylinders for instance, diffusion is unrestricted along the cylinders and Monte Carlo simulations can be limited to the plane perpendicular to the cylinders, implying $n = 2$. Interactions with boundaries and obstacles in Ω such as cellular membranes are tested for at each spatial jump.

With $g_\alpha(t)$ and $r_{l\alpha}$ denoting the components of $\mathbf{g}$ and $\mathbf{r}_l(t)$ along the direction $\mathbf{e}_\alpha$ and γ the gyromagnetic ratio of the spin-bearing nuclei, the phase ϕ_l of spin l accumulated at echo time is defined as

$$\begin{aligned}\phi_l\left(\mathbf{g}\right) &\coloneqq \gamma\int_0^{\text{TE}} \mathbf{g}(t)\cdot\mathbf{r}_l(t)\,dt \\ &= \sum_{\alpha=1}^{n} \underbrace{\gamma\int_0^{\text{TE}} g_\alpha(t) r_{l\alpha}(t)\,dt}_{\coloneqq\phi_{l\alpha}(g_\alpha)}.\end{aligned} \tag{3.4}$$

It is approximated from the discrete trajectory $\hat{\mathbf{r}}_l$ by a numerical quadrature such as a rectangle rule

$$\begin{aligned}\phi_l(\mathbf{g}) \approx \hat{\phi}_l(\mathbf{g}) &= \gamma\delta t \sum_{s=0}^{N_{\text{step}}-1} \mathbf{g}(s\cdot\delta t)\cdot\hat{\mathbf{r}}_l(s\cdot\delta t) \\ &= \sum_{\alpha=1}^{n} \underbrace{\gamma\delta t \sum_{s=0}^{N_{\text{step}}-1} g_\alpha(s\cdot\delta t)\hat{r}_{l\alpha}(s\cdot\delta t)}_{:=\hat{\phi}_{l\alpha}(g_\alpha)}\end{aligned} \tag{3.5}$$

where $\hat{\phi}_{l\alpha}$ is the numerical approximation of the *directional phase* $\phi_{l\alpha}$ due to $g_\alpha(t)$. It should be noted that ϕ (equivalently, $\hat{\phi}$) is linear with respect to $\mathbf{g}(t)$:

$$\phi(\beta_1\mathbf{g}_1 + \beta_2\mathbf{g}_2) = \beta_1\phi(\mathbf{g}_1) + \beta_2\phi(\mathbf{g}_2) \qquad (\forall\beta_1,\beta_2\in\mathbb{R}). \tag{3.6}$$

The normalized diffusion attenuation $A(\mathbf{g})$ in Ω associated with the gradient profile $\mathbf{g}(t)$ is finally approximated as the empirical mean

$$A(\mathbf{g}) = \left\langle e^{j\phi(\mathbf{g})}\right\rangle \approx \left|\frac{1}{N_{\text{spin}}}\sum_{l=1}^{N_{\text{spin}}} e^{j\hat{\phi}_l(\mathbf{g})}\right|, \tag{3.7}$$

where j denotes the complex number.

3.2.3 Data augmentation

3.2.3.1 Augmenting sequences

This paragraph demonstrates how the generation of directional phases –rather than final DW-MRI signals– can make Monte Carlo simulations reusable for new gradient directions and make the simulation time almost independent of the number of sequences M, at the expense of a moderate increase in storage space.

The M gradient profiles $\mathbf{g}(t)$ of a DW-MRI protocol can be expressed as a linear combination of $M_\chi \le M$ basis temporal profiles $\chi_m(t)$

$$\mathbf{g}(t) = \sum_{m=1}^{M_\chi}\sum_{\alpha=1}^{n} \beta_{m\alpha}\chi_m(t)\mathbf{e}_\alpha, \tag{3.8}$$

where every $\beta_{m\alpha}\in\mathbb{R}$. If the $n\times N_{\text{spin}}\times M_\chi$ directional phases $\hat{\phi}_l(\chi_m\mathbf{e}_\alpha)$ are computed *and stored*, they can later be used to compute the accumulated phase $\hat{\phi}_l(\mathbf{g})$ of spin l arising from

the application of any gradient profile $\mathbf{g}(t)$

$$\begin{aligned}\hat{\phi}_l\left(\mathbf{g}\right) &= \hat{\phi}_l\left(\sum_{m=1}^{M_\chi}\sum_{\alpha=1}^{n}\beta_{m\alpha}\chi_m\mathbf{e}_\alpha\right)\\ &= \sum_{m=1}^{M_\chi}\sum_{\alpha=1}^{n}\beta_{m\alpha}\hat{\phi}_l\left(\chi_m\mathbf{e}_\alpha\right)\end{aligned} \tag{3.9}$$

based on the linearity of $\hat{\phi}$ expressed in Eq. (3.6). This quantity can then be used to compute the normalized diffusion attenuation $A\left(\mathbf{g}\right)$ using Eq. (3.7).

Most current clinical protocols use a small number of temporal profiles, making this strategy particularly useful. For instance, PGSE-based Multi-shell HARDI (Tuch et al., 2002) and CUSP protocols (Scherrer and Warfield, 2012) with gradient duration δ and separation Δ fixed across all shells have an optimal $M_\chi = 1$. Once computed, the $n \times N_{\text{spin}} \times M_\chi$ directional phases are used to simulate every sequence of the protocol. Among others, this approach can efficiently generate signals for patient-specific gradient maps corrected for motion. With a fixed protocol, the stored phases also provide DW-MRI signals for any microstructural orientation by considering an adequately-rotated new set of gradient directions with the same temporal profiles $\chi_m(t)$.

Compared with a strategy in which just the final DW-MRI signals are stored, the storage space increases by a factor $\frac{nN_{\text{spin}}M_\chi}{M}$. On the other hand, the computational complexity of the quadrature rule in Eq. (3.5) is reduced by a factor M/M_χ, e.g., a factor 100 for a protocol with 100 sequences using fixed Δ and δ.

3.2.3.2 Augmenting configurations

This paragraph demonstrates that many more microstructural configurations Ω can be obtained from a finite set of stored DW-MRI directional phases or signals by leveraging the scaling properties of Brownian diffusion.

Considering spin-bearing particles evolving in an environment with characteristic length scale L and homogeneous diffusivity D under the application of a magnetic gradient profile $\mathbf{g}(t)$ with characteristic gradient intensity G and time scale T, dimensional analysis shows that the DW-MRI signal $A(L, D; G, T)$ is entirely characterized by the dimensionless parameters (Grebenkov, 2008)

$$\begin{aligned}p_1 &= \frac{DT}{L^2},\\ p_2 &= \gamma GLT.\end{aligned} \tag{3.10}$$

The following equivalence

$$A(L, D; G, T) = A(\sqrt{\alpha}L, \alpha D; \frac{G}{\sqrt{\alpha}}, T) \tag{3.11}$$

therefore holds for any real number $\alpha > 0$, meaning that the same simulated signal can be interpreted as arising from configurations characterized by different spatial dimensions and diffusivity under the application of an adequately scaled magnetic gradient. For instance, the DW-MRI signal arising from an environment with 3µm-wide axons separated by 1µm gaps is identical to the signal arising from 6µm-wide axons separated by 2µm gaps if the diffusivity is 4 times as large and the gradient magnitude is twice as small.

This property is useful in the context of generating large collections of DW-MRI signals. Instead of running Monte Carlo simulations for all combinations (L_i, D_j) of a biological region of interest $[L_{\min}, L_{\max}] \times [D_{\min}, D_{\max}]$ in the space of microstructural parameters Ω, it is sufficient to run simulations for one fixed diffusivity D_{sim} in $[D_{\min}, D_{\max}]$ and a sampling $\{L_{\text{sim},i}\}_i$ spanning a slightly larger region $\left[\sqrt{\frac{D_{\text{sim}}}{D_{\max}}}L_{\min}, \sqrt{\frac{D_{\text{sim}}}{D_{\min}}}L_{\max}\right]$. By setting $\alpha = \frac{D_{\text{sim}}}{D}$ in Eq. (3.11), a collection of DW-MRI signals corresponding to a new arbitrary $D \in [D_{\min}, D_{\max}]$ is then directly obtained as

$$A\left(\sqrt{\frac{D}{D_{\text{sim}}}}L_{\text{sim},i}, D; G, T\right) = A\left(L_{\text{sim},i}, D_{\text{sim}}; \sqrt{\frac{D}{D_{\text{sim}}}}G, T\right), \tag{3.12}$$

where the left-hand-side corresponds to a sampling $\left\{\sqrt{\frac{D}{D_{\text{sim}}}}L_{\text{sim},i}\right\}_i$ covering the region of interest $L_{\min} \leq L \leq L_{\max}$. B.1 shows that the choice of D_{sim} does not affect the number of reference simulations N required to achieve sufficient resolution in L for all $D \in [D_{\min}, D_{\max}]$.

If the reference simulations with D_{sim} stored the directional phases as explained in the previous paragraph, then the right-hand-side of Eq. (3.12) can be exactly evaluated at $\sqrt{\frac{D}{D_{\text{sim}}}}G$. Otherwise, signal interpolation with respect to G must be performed. In order to avoid extrapolation outside of the range $[G_{\min}, G_{\max}]$ covered by the acquisition protocol, the reference simulations can be run for G spanning the slightly larger interval $\left[\sqrt{\frac{D_{\min}}{D_{\text{sim}}}}G_{\min}, \sqrt{\frac{D_{\max}}{D_{\text{sim}}}}G_{\max}\right]$.

3.2.4 Inverse problem

Our framework consists in pre-computing a dictionary of Monte Carlo DW-MRI fingerprints, each corresponding to a unique microstructural configuration. At runtime, for every voxel, our method then aims at finding the optimal combination of single-fascicle configurations $\hat{\mathbf{\Omega}}_1, \ldots, \hat{\mathbf{\Omega}}_K$ and volume fractions $\hat{v}_1, \ldots, \hat{v}_K$ for a vector $\mathbf{y} \in \mathbb{R}^M$ of M noisy DW-MRIs. More specifically, the process involves five steps: two during the pre-computing stage and three during runtime.

Step 1 (pre-computing) performs a discrete sampling of N points $\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_N$ of the space Ω of microstructural parameters with bounds and granularity justified by tissue biology and the expected resolution of the M diffusion-encoding gradient profiles $\mathbf{g}_1(t), \ldots, \mathbf{g}_M(t)$ of the protocol $\mathscr{P} = \{\mathbf{g}_i(t)\}_{i=1}^M$ at hand.

Step 2 (pre-computing) generates a canonical single-fascicle dictionary $\mathbf{F}^0 \in \mathbb{R}^{M\times N}$ containing the DW-MRI fingerprints of the N selected microstructural configurations for a single fascicle along a fixed direction $\mathbf{u}_0$

$$\mathbf{F}^0 = \left[\mathbf{A}^0_1, \ldots, \mathbf{A}^0_N\right], \tag{3.13}$$

where $\mathbf{A}^0_j = A_{\text{fasc}}\left(\mathbf{\Omega}_j, \mathbf{T}_0, \mathbf{u}_0; \mathscr{P}\right)$, $1 \le j \le N$. This the most time-consuming step involving up to N Monte Carlo simulations which, however, need only be performed once and can be done before any data is even acquired.

Step 3 (runtime) requires an external routine to estimate in each voxel the number K of fascicles and their orientations $\mathbf{u}_1, \ldots, \mathbf{u}_K$.

Step 4 (runtime) requires an efficient routine for rotating single-fascicle signals in order to obtain the single-fascicle dictionaries $\mathbf{F}^1, \ldots, \mathbf{F}^K \in \mathbb{R}^{M\times N}$ from $\mathbf{F}^0$ along the orientations $\mathbf{u}_1, \ldots, \mathbf{u}_K$ estimated in Step 3.

Step 5 (runtime) finally consists in solving the following sparse optimization problem

$$\begin{aligned} \hat{\mathbf{w}} = \quad & \underset{\mathbf{w}\ge 0}{\operatorname{argmin}} \quad \left\| \mathbf{y} - \left[\mathbf{F}^1 | \ldots | \mathbf{F}^K | \mathbf{A}_{\text{csf}}\right] \cdot \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_K \\ w_{csf} \end{bmatrix} \right\|^2_2 \\ & \text{subject to} \quad |\mathbf{w}_k|_0 = 1, \quad k = 1, \ldots, K, \end{aligned} \tag{3.14}$$

where the sparsity constraints on the sub-vectors $\mathbf{w}_k$ guarantee that only one fascicle configuration Ω_{j_k} per single-fascicle dictionary $\mathbf{F}^k$ contributes to the measured signal $\mathbf{y}$. Problem (3.14) is solved exactly by selecting the optimal solution out of N^K independent non-negative least-squares sub-problems of $(K+1)$ variables

$$(\hat{j}_1, \ldots, \hat{j}_K) = \underset{1\le j_1,\ldots,j_K\le N}{\operatorname{argmin}} \quad \min_{\mathbf{w}\ge 0} \left\| \mathbf{y} - \left[\mathbf{A}^1_{j_1} | \ldots | \mathbf{A}^K_{j_K} | \mathbf{A}_{\text{csf}}\right] \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_K \\ w_{csf} \end{bmatrix} \right\|^2_2 . \tag{3.15}$$

Each sub-problem is convex and is solved exactly by an efficient active-set algorithm (Lawson and Hanson, 1995, chap. 23, p. 161). The optimal microstructural parameters $\hat{\Omega}_k$ are taken as those of the optimal fingerprint $\hat{j}_k$ in each $\mathbf{F}^k$ and the volume fractions $\hat{v}_k$ are estimated from the corresponding optimal weights $\hat{w}_k$ using Eq. (3.2). Note that the weights w_k are not required to sum to one since the quantity $\sum_{k=1}^{K+1} \hat{w}_k$ should reflect the scale M_0 of the acquired signals $\mathbf{y}$.

3.3 Materials and Methods

This section presents the synthetic and *in vivo* experiments performed to validate the general estimation framework described in Section 3.2.4. The first part is concerned with voxels containing only single fascicles of axons while the second part considers voxels with crossing fascicles. All the synthetic experiments were designed to reproduce the experimental conditions of the *in vivo* acquisitions as closely as possible.

3.3.1 Validation on single fascicles

3.3.1.1 Diffusion protocol

The DW-MRI protocol used for the synthetic and *in vivo* validation on single-fascicle configurations, referred to as the *rodent protocol*, consisted of 6 PGSE shells of b-values 300, 700, 1500, 2800, 4500, 6000 s mm^{-2} with high gradient intensities $G \in [140, 628]$ mT m^{-1} and short gradient duration $\delta = 4.5$ ms and diffusion time $\Delta = 12$ ms, with TE = 23 ms. Each shell contained 36 non-collinear directions computed with the electrostatic repulsion method of Caruyer et al. (2013) and 3 unweighted or b0 images for a total of 234 images.

3.3.1.2 Implementation details of the dictionary estimation

In Step 1 of the procedure described in Section 3.2.4, hexagonal packing of straight, impermeable cylinders was selected to represent single fascicles of axons, characterized by a cylinder radius r, interpreted as an apparent axonal radius index, and a cylinder packing density f, interpreted as an axonal density index (Budde and Frank, 2010; Alexander et al., 2010). The space of microstructural parameters $\Omega = (r, f)$ was sampled at 34 values for r from 0.4 µm to 7 µm by steps of 0.2 µm and 23 values for f from 0.21 to 0.87 by increments of 0.03. This resulted in a canonical single-fascicle dictionary $\mathbf{F}^0$ containing the diffusion fingerprints of $N = 782$ microstructural configurations along a reference orientation. The groundtruth diffusivities of the intra- and extra-axonal or "white matter" space were kept equal and fixed to $D_{\text{wm}} = 2.0 \times 10^{-9}$ m^2 s^{-1} based on the intra-axonal estimate by Dhital et al. (2019). The diffusivity of CSF was set to $D_{\text{csf}} = 3.0 \times 10^{-9}$ m^2 s^{-1} (Xing et al., 1997). Since a PGSE protocol was considered, NMR relaxation consisted of T2 decay of the form $\exp(-\text{TE}/\text{T2})$, assumed independent of the diffusion process. In order to match our *in vivo* rat data set acquired at 11.7T, the groundtruth relaxation times were considered to be $\text{T2}_{\text{wm}} = 30$ ms for the intra- and extra-axonal space based on estimates at 11.7T in the rat corpus callosum and cerebellar white matter (de Graaf et al., 2006). For CSF, $\text{T2}_{\text{csf}} = 120$ ms was selected in line with rat brain estimates at 7.0T (Crémillieux et al., 1998) and at 11.7T (Pohmann et al., 2011).

The Monte Carlo simulations of Step 2 in our estimation procedure were performed using an in-house software for the extra-axonal signal of the gradient components perpendicular to the cylinders with a number of random walkers $N_{\text{spins}} = 150\,000$ and a time step δt between 1

and 5µs based on the distance between cylinders (Rensonnet et al., 2015, 2016). Exact intra-axonal signals were obtained using an efficient implementation[1] of the Multiple Correlation Function (MCF) approach (Grebenkov, 2008). As described in Section 3.2.3.1, one single set of directional phases ϕ_x and ϕ_y per fingerprint needed to be stored in order to compute all 216 diffusion-weighted acquisitions of the rodent protocol.

In Step 3, the number of fascicles in each voxel was limited to $K = 1$ and the orientations were obtained from a ball-and-sticks model estimated using a maximum a posteriori approach as described in Scherrer et al. (2016), using the CRKIT software[2] without spatial regularization across voxels.

Since a multi-shell PGSE protocol was considered, the rotation of DW-MRI signals required in Step 4 was done by simple linear spline interpolation separately on each shell.

3.3.1.3 Synthetic experiments

Two synthetic experiments were designed to validate the estimation method in a variety of controlled groundtruth configurations and to investigate the effect of uncertainties on fixed or pre-estimated parameters.

In both experiments, the results of our Monte Carlo based estimations were systematically compared with the output of the minimal model of white matter diffusivity (MMWMD) introduced in Alexander et al. (2010) (see Section 1.2.2.2). MMWMD was chosen for comparison as it provides direct indices of the radius and axonal density to which our method can be compared. More specifically, the quantity $\frac{f_{\text{cylinder}}}{f_{\text{cylinder}}+f_{\text{zeppelin}}}$ was used as an axonal density index, where f_{cylinder} and f_{zeppelin} are the fractions of signal modeled by the cylinder and the zeppelin-like diffusion tensor, respectively. All MMWMD fitting was performed with the Camino Diffusion MRI Toolkit (Cook et al., 2006) using Markov chain Monte Carlo (MCMC) fitting with 40 samples at intervals of 200 iterations after a burn-in of 2000 iterations as recommended in Alexander et al. (2010).

The signal-to-noise ratio (SNR) was defined as SNR = $0.5M_0/\sigma$, where σ is the standard deviation of the noise in an individual MRI detection coil. The factor 0.5 comes from the mid-point value of the unweighted b0 signal for 0 and 25% of the volume occupied by CSF, using the T2 values described above. The scaling parameter M_0 depends on the sensitivity of the MRI scanner in practice and was fixed to an arbitrary $M_0 = 1000$ in the groundtruth signals throughout all synthetic experiments.

3.3.1.3.1 Experiment 1.A. Single-fascicle groundtruth, effect of fixed diffusivity This experiment focused on the estimation of microstructural properties in single fascicles and

[1] Our code was based on publicly available scripts from the original author's web page https://pmc.polytechnique.fr/pagesperso/dg/MCF/MCF_e.htm.

[2] http://crl.med.harvard.edu/software/.

included no CSF contribution in the synthetic groundtruth and in the fitted models. Consequently Eq. (3.15) was solved with $w_{csf} = 0$ in our Monte Carlo dictionary estimation while the ball and dot compartments were ignored in the MMWMD fitting.

Equation (3.1) was used to generate reference groundtruth signals for 64 single-fascicle configurations obtained from the combinations of 8 groundtruth radius index values r from 0.6 to 4.8µm by steps of 0.6µm and 8 groundtruth density index values f from 0.42 to 0.84 by steps of 0.06. Each signal was corrupted by 10 independent simulations of Rician noise with 10 SNR levels varying from 5 to 150. Estimation was performed for all 4 combinations of groundtruth diffusivity $D^* = \{2.0, 3.0\} \times 10^{-9} \mathrm{m}^2\,\mathrm{s}^{-1}$ and model diffusivity $D = \{2.0, 3.0\} \times 10^{-9} \mathrm{m}^2\,\mathrm{s}^{-1}$ in order to investigate the effect of fixing D to an under- or overestimated value (with signals generated using data augmentation as explained in Section 3.2.3.2). This yielded a total of $64 \times 10 \times 10 \times 4 = 25\,600$ independent synthetic voxels.

3.3.1.3.2 Experiment 1.B. CSF partial volumes, effect of fixing T2 values In this experiment, partial volumes of CSF were added both to the groundtruth voxels and to the fitted models. The 64 single-fascicle configurations selected in Exp. 1.A were considered and isotropic CSF contributions with physical volume fractions $v_{csf} = \{0.0, 0.25, 0.50\}$ were successively added following Eq. (3.1). The levels of Rician noise corruption and number of repetitions were as in Exp. 1.A and the model diffusivities were set to the groundtruth values. Three configurations of T2 values were examined as detailed in Table 3.1. The "$T2_{\checkmark}$" scenario was the ideal case; in the "$T2_{\text{✗}1}$" setting the groundtruth values were unchanged but the models assumed uniform T2 at a typical 3T value of 70 ms; the "$T2_{\text{✗}2}$" case was more challenging as the groundtruth signals were generated with different T2 values in the intra- and extra-axonal space of the single fascicle. MMWMD fitting was performed using the full four-compartment model including a dot compartment.

Table 3.1: T2 values of the intra-axonal, extra-axonal and CSF compartments assumed in the groundtruth (GT) and the model in three different scenarios for Exp. 1.B.

	$T2_{\checkmark}$		$T2_{\text{✗}1}$		$T2_{\text{✗}2}$	
	GT	Model	GT	Model	GT	Model
$T2_{in}$	30	30	30	70	30	70
$T2_{ex}$	30	30	30	70	45	70
$T2_{csf}$	120	120	120	70	120	70

3.3.1.4 *In vivo* experiment

3.3.1.4.1 Animal model of Wallerian degeneration Three female Long Evans rats (Janvier Labs, Le Genest-Saint-Isle, France; weight 180 – 200 g) underwent laminectomy of verterbrae L2-L3 to expose the spinal cord. A left unilateral dorsal root axotomy was then performed, inducing Wallerian degeneration in the ipsilateral side of the gracile fasciculus of the spinal cord while leaving the contralateral side untouched. Two similar rats served as the control

group and underwent laminectomy at identical vertebral levels without the dorsal root axotomy, leaving the whole spinal cord unaffected. All rats were scanned 51 days after surgery (see next paragraph) and sacrificed immediately after the imaging session. The spinal cords were then extracted, frozen and sliced axially in 20-µm thick sections to perform SMI312 staining, used to expose neurofilaments and indicate the presence of axons. The entire protocol was approved by the local animal care and ethics committee at Université catholique de Louvain (2016/UCL/MD/011).

3.3.1.4.2 *In vivo* DW-MRI DW-MRI was performed on all five rats on an 11.7T Bruker BioSpec scanner (Bruker, Billerica, MA) using a 72-mm diameter transmitter volume coil and a 4-channel, 3 × 3 cm surface receiver coil covering the L4-T12 vertebral segments. Rats lay in the dorsal decubitus position on a custom-made bed, anesthetized with an isoflurane-air mixture (2.5% for induction and 1-1.5% for maintenance). Respiration and rectal temperature were continuously monitored and body temperature was kept stable at 37 °C using a circulating warm-water pad. Diffusion-weighted images were acquired with the PGSE parameters described in Section 3.3.1.1 and TR = 3 s, using 2-D echo planar imaging (EPI) with in-plane voxel resolution $0.1 \times 0.1\text{mm}^2$ (128×128 matrix) and slice thickness 1 mm for 16 contiguous axial slices, for a total acquisition time of about 2.5 h per rat. Correction for animal motion and Eddy current was achieved by affine registration of each scalar DW-MRI to the b0 images interleaved in the protocol. To this end, images were resampled to a resolution of $0.05 \times 0.05 \times 0.06$mm enabling improved multi-scale pyramidal registration.

3.3.1.4.3 Model fitting and statistical analysis The gracile fasciculus was manually segmented based on the hypo-intense signal of the anterior spinal vein on the $b0$ images and the high fractional anisotropy (FA) values of the corticospinal tract obtained from diffusion tensor fitting, while CSF voxels were manually removed at the periphery of the spinal cord.

Estimation was performed in the voxels of the gracile fasciculus of all five rats using our Monte Carlo dictionary approach, yielding estimates of r, f and v_{csf}. MMWMD was estimated as described in Exp. 1.B, providing an extra parameter v_{dot}. For further comparison, the four closed-form microstructural models DIAMOND (Scherrer et al., 2016), NODDI (Zhang et al., 2012), WMTI (Fieremans et al., 2011) and MAPL (Fick et al., 2016) were also fitted to the data. From DIAMOND, the compartment heterogeneity index cHEI, compartment radial diffusivity cRD, compartment axial diffusivity cAD and isotropic volume fraction v_{iso} were examined. From NODDI, the intra-neurite volume fraction f_{icvf}, orientation dispersion index ODI and isotropic volume fraction v_{iso} (see Section 1.2.2.2) were estimated using the NODDI MATLAB (The MathWorks, Inc, Natick, MA) Toolbox[3]. From WMTI, the axonal water fraction (AWF), intra-axonal radial diffusivity RD_{in}, extra-axonal radial diffusivity RD_{ex} and extra-axonal axial diffusivity AD_{ex} were estimated using the `reconst.dki_micro` module from the DIPY

[3] http://mig.cs.ucl.ac.uk/index.php?n=Tutorial.NODDImatlab

software project[4] (Garyfallidis et al., 2014). From MAPL, the Return-to-Origin Probability RTOP, Return-to-Axis Probability RTAP, Return-to-Plane Probability RTPP and mean squared displacement MSD (see Section 1.2.3.3) were estimated using the `reconst.mapmri` module from DIPY.

After model fitting, the following linear mixed-effect regression was estimated for each of the microstructural parameters described above using MATLAB's `fitlme` routine

$$y = \beta_0 + \beta_{\text{WD}} \times S \times I + \beta_{\text{surgery}} \times S + \beta_{\text{ipsilateral}} \times I, \tag{3.16}$$

where y is the microstructural property of interest, S and I are indicator variables respectively indicating surgery ($S = 1$) versus controls ($S = 0$) and ipsilateral ($I = 1$) versus contralateral ($I = 0$) sides. A subset of voxels in the gracile fasciculus corresponding to the original DW-MRI resolution were selected for the analysis in order to avoid artificially increasing our sample size and driving p-values to zero. In Eq. (3.16), the coefficient β_{WD} captures the effect of Wallerian degeneration. It should only be large for parameters *physically impacted by the surgery* and not by intrinsic differences between the control and the injured group (β_{surgery}) or rat-specific differences between the left and right sides of the spinal cord ($\beta_{\text{ipsilateral}}$). In this experiment, the quality of a model does not lie in its ability to detect significant group differences but rather in its ability to attribute the signal change to specific microstructural parameters in line with histological observations.

3.3.2 Validation on crossing fascicles

3.3.2.1 Diffusion protocol

All synthetic and *in vivo* experiments on crossing-fascicle configurations were carried out using the MGH-USC Adult Diffusion protocol of the Human Connectome Project (HCP) described in Setsompop et al. (2013). The protocol comprised 4 PGSE HARDI shells containing 64 gradient directions at $b = 1000\,\text{s}\,\text{mm}^{-2}$, 64 at $b = 3000\,\text{s}\,\text{mm}^{-2}$, 128 at $b = 5000\,\text{s}\,\text{mm}^{-2}$, 256 at $b = 10\,000\,\text{s}\,\text{mm}^{-2}$ and 40 $b0$ images interleaved throughout the protocol, for a total of 552 acquisitions. Gradients intensities reached $G = 219\,\text{mT}\,\text{m}^{-1}$ with $\delta/\Delta = 12.9/21.8\,\text{ms}$, enabling $\text{TE} = 57\,\text{ms}$.

3.3.2.2 Implementation details of the dictionary estimation

In Step 1 of the inverse problem (Section 3.2.4), the same single-fascicle model as in Section 3.3.1.2 was selected, along with the same sampling of the microstructural parameter space $\Omega = (r, f)$ and identical diffusivities. Groundtruth T2 values were set to typical human brain values at 3T with $\text{T2}_{\text{wm}} = 70\,\text{ms}$ based on Stanisz et al. (2005) and Smith et al. (2008), as-

[4] http://nipy.org/dipy/index.html

sumed identical in all fascicles. In CSF, $T2_{csf} = 1000\,ms$ was interpolated from 0.14 T estimates in humans (Condon et al., 1987) as well as rat estimates at 4.7 T (Ting and Bendel, 1992) and 7 T (Crémillieux et al., 1998).

Steps 2, 3, and 4 were performed as in the single-fascicle experiments except that the ball-and-sticks estimation in Step 3 was set to detect up to $K = 2$ fascicles. In step 5, $N^2 = 611\,524$ non-negative least squares sub-problems were solved in voxels containing two fascicles.

3.3.2.3 Synthetic experiments

Two synthetic experiments were designed to validate the estimation method in a variety of controlled crossing-fascicle configurations and to investigate the effect of misestimating fixed or estimated parameters.

As with the rodent protocol, the scanner-specific scaling parameter was fixed to $M_0 = 1000$ and the SNR was computed as $\text{SNR} = 0.5M_0/\sigma$ based on unweighted $b0$ signals without CSF contamination and with 25% of CSF in the voxel, using the above T2 values.

3.3.2.3.1 Experiment 2.A. Independent voxels, effect of orientation misestimation and crossing angle This experiment focused on the estimation of microstructural properties in crossing fascicles and included no CSF contribution in the synthetic groundtruth and in the Monte Carlo dictionary estimation, i.e. Eq. (3.15) was solved with $w_{csf} = 0$.

Equation (3.1) was used to generate reference groundtruth signals for fascicles with identical microstructural properties $r = r_1 = r_2$ and $f = f_1 = f_2$ with volume occupied by the first fascicle $v_1 = \{0.3, 0.4, 0.5\}$ and $v_2 = 1 - v_1$, for all 32 combinations of 4 radius index values $r = \{1, 2, 3, 4\}\,\mu m$ and 8 density index values f from 0.42 to 0.84 by steps of 0.06. Each signal was corrupted by 10 independent simulations of Rician noise with 10 SNR levels varying from 5 to 150.

To examine the effect of an incorrect estimation of the fascicle's orientations in Step 3 of the inverse problem, the estimation was performed by independently selecting $\mathbf{u}_1$ and $\mathbf{u}_2$ randomly on a cone with principal axis along the groundtruth orientations, forcing angular errors of $0°, 5°$ and $10°$ successively. Fascicles crossing at angles $\angle\mathbf{u}_1, \mathbf{u}_2 = \{30°, 60°, 90°\}$ were considered in order to study the effect of the groundtruth crossing angle, thereby yielding a total of $3 \times 32 \times 10 \times 10 \times 3 \times 3 = 86\,400$ independent voxel estimations.

3.3.2.3.2 Experiment 2.B. Synthetic 2D phantom, effect of dissimilar fascicles A synthetic phantom was designed containing three axonal tracts (see Fig. 3.7(a)). Each tract had a constant radius index r (respectively 1.2, 1.6 and 2.0 µm) and spatially-smooth variations of the fascicle-specific density index f ranging from 0.45 to 0.81. The complete 2D-slice featured $17 \times 17 = 289$ voxels including 119 voxels containing one single fascicle of axons, 34 voxels con-

taining a single fascicle with 25% of CSF and 86 voxels containing two fascicles intersecting at angles comprised in $[31.4°, 86.0°]$ with mean 62.7°. The crossing-fascicle configurations were more complex than in Exp. 2.A because the two crossing fascicles had different microstructural properties $r_1 \neq r_2$, $f_1 \neq f_2$ in general. Estimation was performed at SNR levels of 25, 50 and 100.

In the ball-and-sticks estimation of Step 3 of the Monte Carlo dictionary estimation, spatial regularization across voxels was enabled and sticks with directions separated by fewer than 15° were merged. Denoting by β_1 the largest weight attributed to a stick, all secondary sticks were then removed if their weight β verified either $\beta < \beta_1/2.5$ and $\beta < 0.20$ or just $\beta < 0.10$. The Monte Carlo dictionary estimation of Step 5 set $w_{\text{csf}} = 0$ when two fascicles were detected in Step 3.

3.3.2.4 *In vivo* experiment

3.3.2.4.1 *In vivo* DW-MRI One healthy subject was randomly selected from the MGH Adult Diffusion data release[5] (Setsompop et al., 2013).

3.3.2.4.2 Model fitting and statistical analysis Voxels containing white matter were identified based on the segmentation obtained with the FAST algorithm (Zhang et al., 2001) from the FMRIB Software Library (FSL)[6]. Monte Carlo dictionary estimation was performed independently in each voxel as described in the synthetic experiments.

Our experiment focused on extracting the distribution of apparent axonal radius index r and density index f of axons passing through the anterior, mid-anterior, central, mid-posterior and posterior sub-regions of the corpus callosum (CC) as identified by the subcortical segmentation tool of the FreeSurfer software (Fischl et al., 2002). As these axons cross other macroscopic tracts such as the corticospinal tract (CST) or the longitudinal fasciculus (LF), each fascicle or peak at the local voxel level had to be assigned to one or more of these macroscopic tracts. In order to do so, probabilistic tractography was first performed with 5 seeds per voxel and streamline segments constrained to follow the orientation of a detected peak in each voxel, using routines from DIPY. The streamlines were clustered into tracts using the white matter query language (Wassermann et al., 2016) and a local voxel fascicle was considered to belong to a macroscopic tract if at least 5% of all streamlines going through the local peak had been assigned to that tract.

[5] https://www.humanconnectome.org/study/hcp-young-adult/document/mgh-adult-diffusion-data-acquisition-details
[6] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL

3.4 Results

3.4.1 Validation on single fascicles

3.4.1.1 Experiment 1.A. Single-fascicle groundtruth, effect of fixed diffusivity

The mean absolute error (MAE) in the estimation of the radius and density index for our approach and for MMWMD are depicted in Fig. 3.1. When the correct diffusivity D was assumed (green curves), the MAE over all repetitions and over all configurations with our approach converged to zero with increasing SNR. Underestimating (resp. overestimating) D led to a systematic underestimation (resp. overestimation) of r and f as indicated by the blue (resp. yellow) estimate bars at SNR level 25. On the other hand, systematic errors persisted for MMWMD estimates at large SNR values even when the correct value for D was assumed in the model.

3.4.1.2 Experiment 1.B. CSF partial volumes, effect of fixing T2 values

Figure 3.2 suggests that our approach was able to provide accurate microstructural estimates in the presence of CSF. Errors on r and f exhibited a moderate upward trend as the fraction of CSF increased, likely due to a reduced relative signal (and hence a reduction in apparent SNR) arising from the fascicle of axons. The estimation errors for MMWMD were systematically larger than in Exp. 1.A. The MAE on v_{csf} hovered around 0.05 across SNR levels even in the $\text{T2}_{\checkmark}$ scenario and when there was no CSF in the groundtruth. As shown in Figure 3.3, the MMWMD estimates exhibited a larger variability than those of our Monte Carlo dictionary approach.

Notably, the estimates of r and f using the incorrect T2 values of the $\text{T2}_{\text{✗}1}$ scenario (cross markers in Fig. 3.2) were identical to those obtained with the groundtruth T2 values (circle markers in Fig. 3.2) for both models. The error in T2 was simply corrected by scaling the fascicle and CSF signals by an adjusted weight w, which then led to a misestimated volume fraction v_{csf}. This was no longer the case in the $\text{T2}_{\text{✗}2}$ scenario because the model and the groundtruth signals no longer differed by just a scaling constant.

3.4.1.3 *In vivo* experiment

Histological slices of the rats which underwent surgery revealed lighter SMI312 staining on the ipsilateral side, indicating axonal loss induced by Wallerian degeneration (Fig. 3.4(a)).

As depicted in Fig. 3.4(b-c), our Monte Carlo dictionary approach exhibited an important decrease in axonal density index f (-0.15, $p = 1.6 \times 10^{-5}$) and statistically non-significant changes in r and v_{csf}. In contrast, MMWMD detected no significant change in f and a large increase in r ($+1.9\,\mu\text{m}$, $p = 2.1 \times 10^{-2}$) while yielding no statistically-significant changes in v_{csf} and v_{dot}. NODDI found a non-significant increase in f_{icvf}, an increase in dispersion ($+0.09$, $p <$

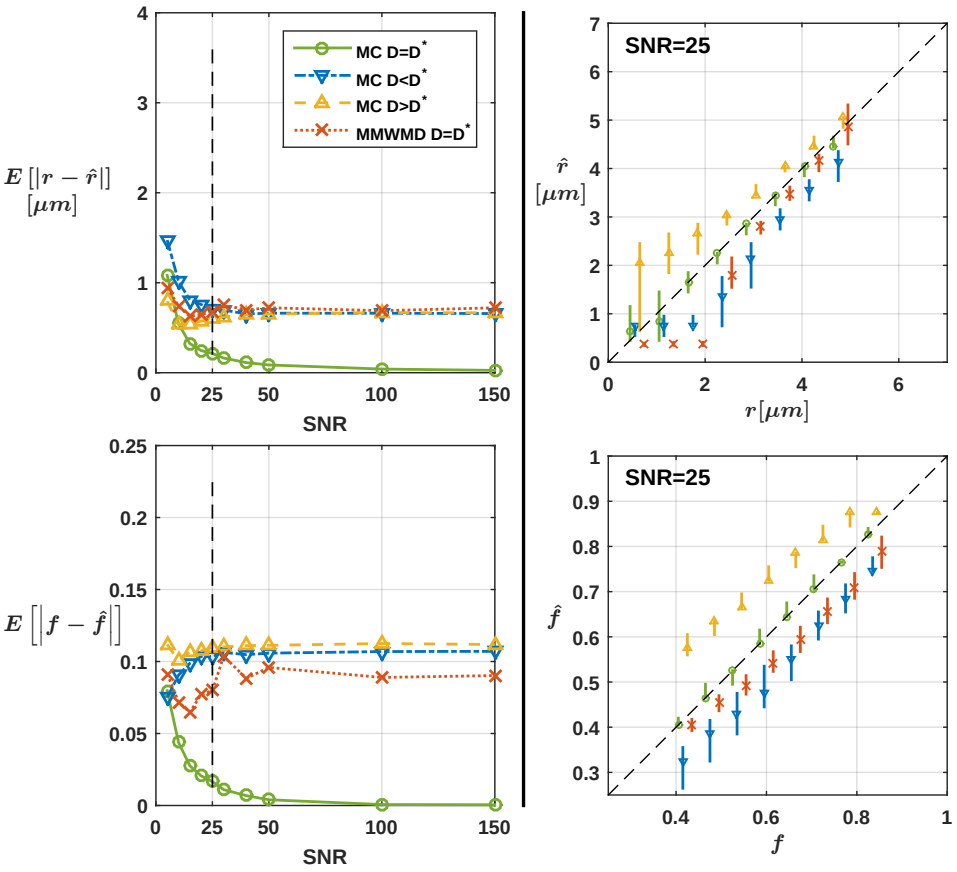

Figure 3.1: (Exp. 1.A) **Consistent microstructural estimates in single fascicles with Monte Carlo (MC) dictionary estimation.** The left column displays the mean absolute error on the radius index r (top) and density index f (bottom) as a function of signal-to-noise ratios (SNR). The right column provides a snapshot of estimates at SNR=25, with markers indicating the median over all noise repetitions and groundtruth configurations and bars ranging from the 25th to the 75th percentile of all estimates. Asymptotic errors remained for the minimal model of white matter diffusivity (MMWMD) even though it used the true diffusivity D.

2.9×10^{-22}) and no significant change in v_{iso}. DIAMOND obtained a non-significant change in cHEI, an increase in cRD ($+0.12\,\mu m^2\,ms^{-1}$, $p = 1.4 \times 10^{-4}$), a decrease in cAD ($-0.40\,\mu m^2\,ms^{-1}$,

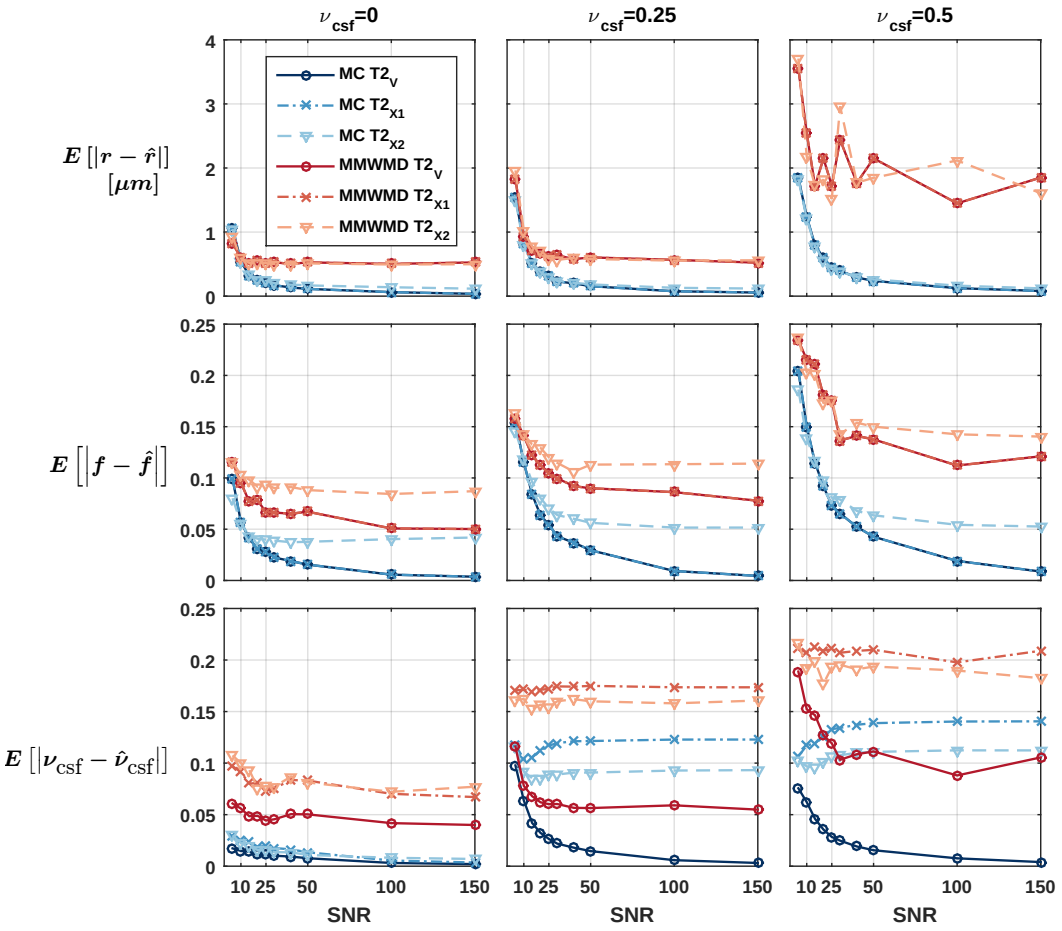

Figure 3.2: (Exp. 1.B) **Consistent microstructural estimates from Monte Carlo dictionary in the presence of CSF contamination.** Mean absolute errors on the radius index r (top), density index f (middle) and cerebrospinal fluid (CSF) volume fraction v_{csf} as a function of signal-to-noise ratio for various levels of CSF contamination. The MMWMD approach yielded generally larger errors. Assuming incorrect T2 values in the model did not affect the estimation of r and f in the the first two scenarios described in Table 3.1.

$p = 4.2 \times 10^{-4}$) and a slight increase in v_{iso} (+0.07, $p = 3.0 \times 10^{-2}$). WMTI saw a slight decrease in AWF (−0.04, $p = 4.8 \times 10^{-3}$), no statistically-significant change in RD_{in}, an increase in RD_{ex} ($+0.19\,\mu m^2\,ms^{-1}$, $p = 2.2 \times 10^{-3}$) and a decrease in AD_{ex} ($-0.57\,\mu m^2\,ms^{-1}$, $p = 4.4 \times 10^{-12}$). MAPL identified a decrease in RTAP ($-5.4 \times 10^{-3}\,\mu m^{-2}$, $p = 1.6 \times 10^{-2}$), an increase in RTPP ($+1.5 \times 10^{-2}\,\mu m^{-1}$, $p = 1.2 \times 10^{-8}$) and no statistically-significant changes in RTOP and MSD.

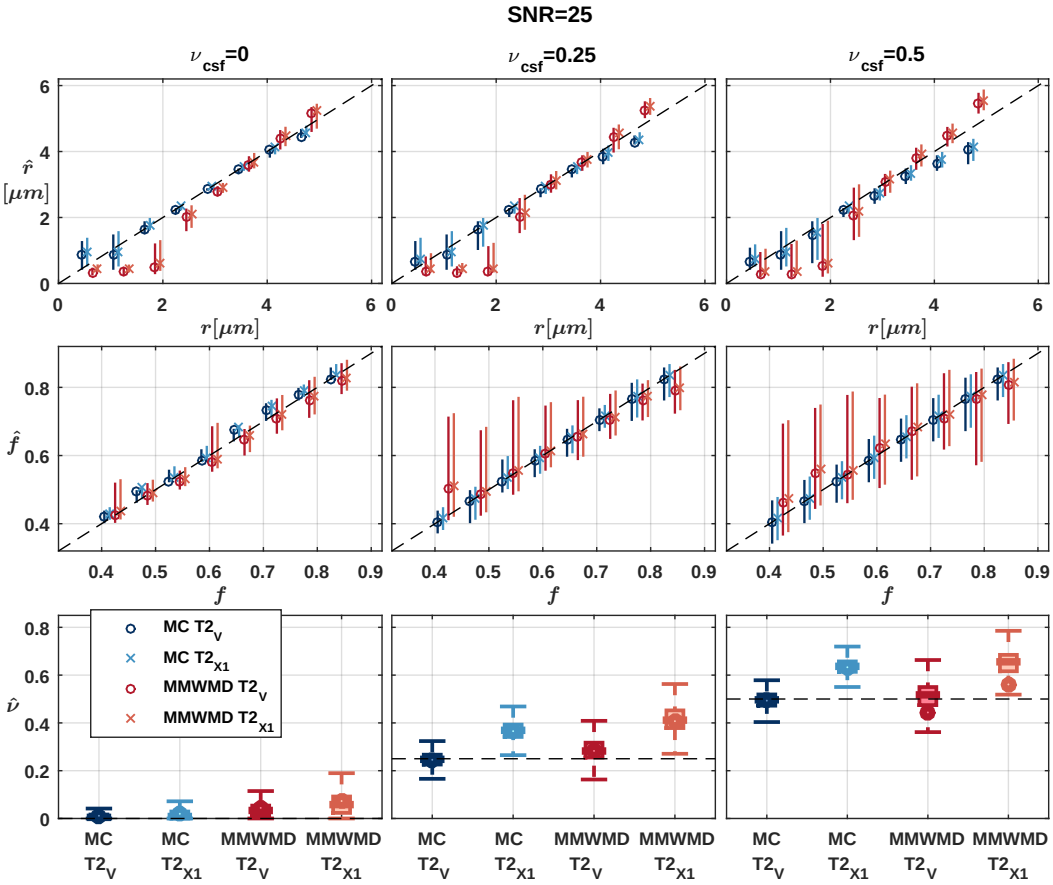

Figure 3.3: (Exp. 1.B) **Consistent microstructural estimates from Monte Carlo dictionary in the presence of CSF contamination.** Snapshot of the estimates of r, f and ν_{csf} at SNR=25 (see Fig. 3.2), with markers indicating the median over all noise repetitions and groundtruth configurations and bars ranging from the 25th to the 75th percentile of all estimates. The estimates of the T2$_{X2}$ scenario were left out for clarity. MMWMD had larger variability and more biased estimates.

3.4.2 Validation on crossing fascicles

3.4.2.1 Experiment 2.A. Independent voxels, effect of orientation misestimation and crossing angle

As suggested by Fig. 3.5, the crossing angle had a marginal impact on the estimation as all the curves are very close to one another. Figure 3.6 indicates that the MAE over all noise repetitions and groundtruth configurations converged to zero as the SNR increased when the orientation of each fascicle was perfectly estimated in Step 3 of our estimation procedure (blue curves). Misestimation of the orientation of fascicles introduced systematic errors in the microstructural estimation (red and yellow curves). The estimates were generally slightly better for the dominant groundtruth fascicle: at a crossing angle of 60° with $v_1 = 0.3$, the MAE

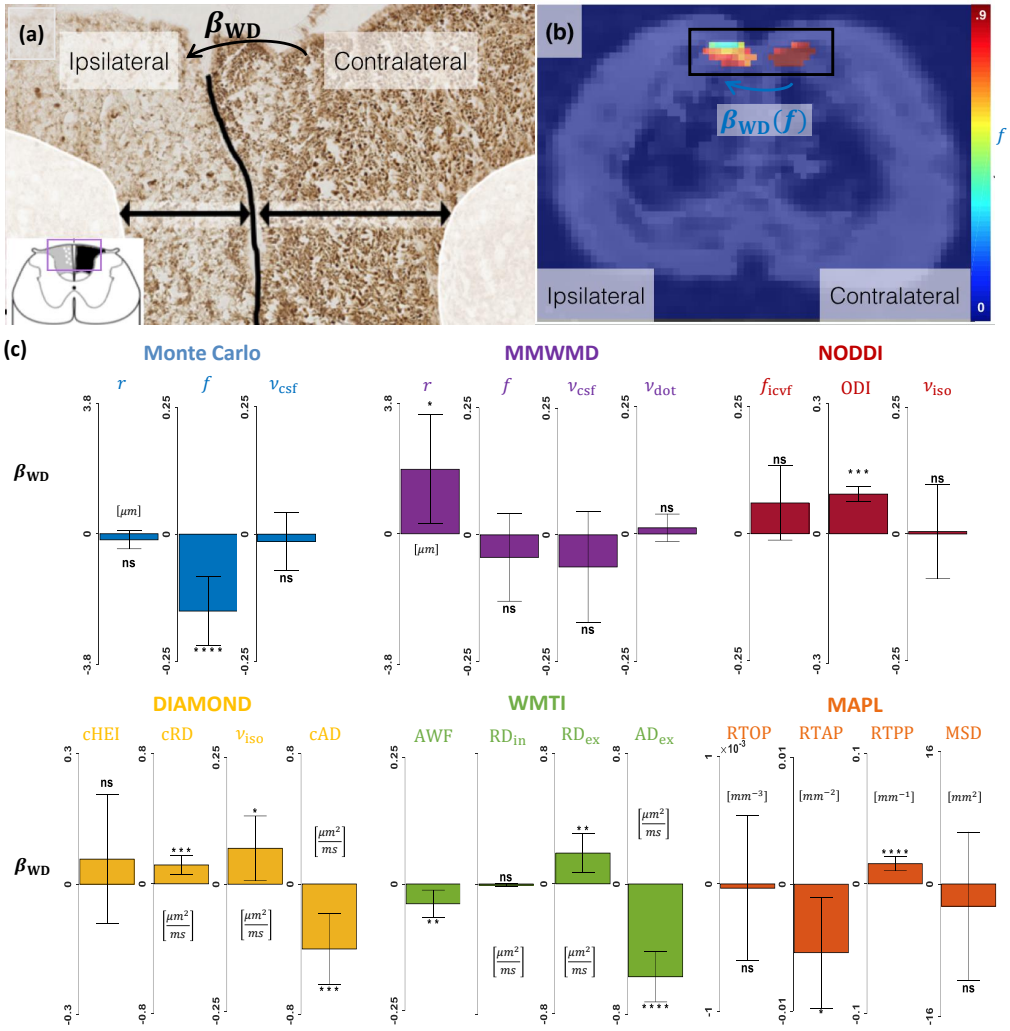

Figure 3.4: **Monte Carlo dictionary approach yields physically-interpretable parameters consistent with histology** (a) Histology of a slice in the spinal cord of a rat which underwent surgery, stained with SMI312 immunohistochemistry. Darker colors indicate more neurofilaments present inside axons. (b) Representative map of axonal density index f obtained with our Monte Carlo dictionary approach, laid atop a fractional anisotropy (FA) map. (c) Effect of Wallerian degeneration (WD) on selected parameters from our approach and from popular closed-form models of the microstructure. Vertical bars indicate the 95% confidence interval on β_{WD}. Non-significant (ns) corresponds to $p > 0.05$; * to $p \leq 0.05$; ** to $p \leq 0.01$; *** to $p \leq 0.001$; **** to $p \leq 0.0001$.

on r_1 exceeded the MAE on r_2 in 25 out of 30 cases with a mean signed difference of 0.23 µm; the MAE on f_1 exceeded the MAE on f_2 in all 30 cases, with a mean difference of 0.060.

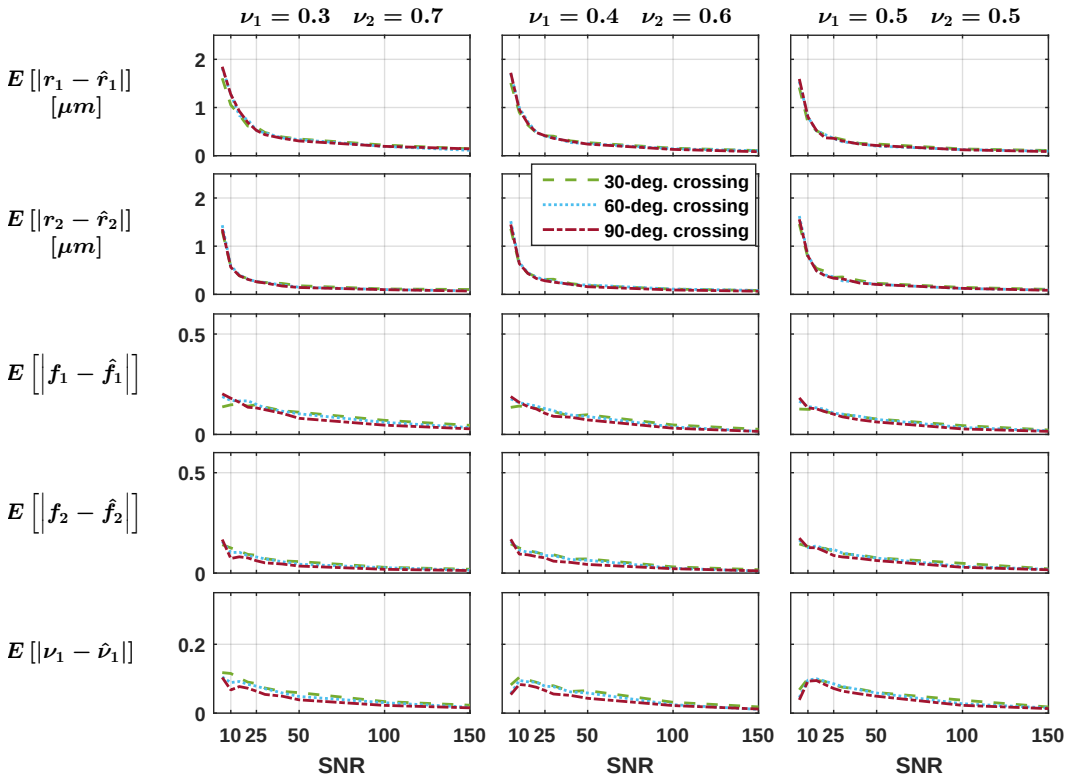

Figure 3.5: (Exp. 2.A) **The groundtruth crossing angle has limited impact on the estimation of fascicles' microstructural properties.** Mean absolute error on each fasicle's radius index r_1 and r_2, density index f_1 and f_2 and on the physical volume fraction occupied by the first fascicle v_1. The groundtruth volume fraction of the second fascicle decreases from left to right. The fascicles' orientations were perfectly estimated.

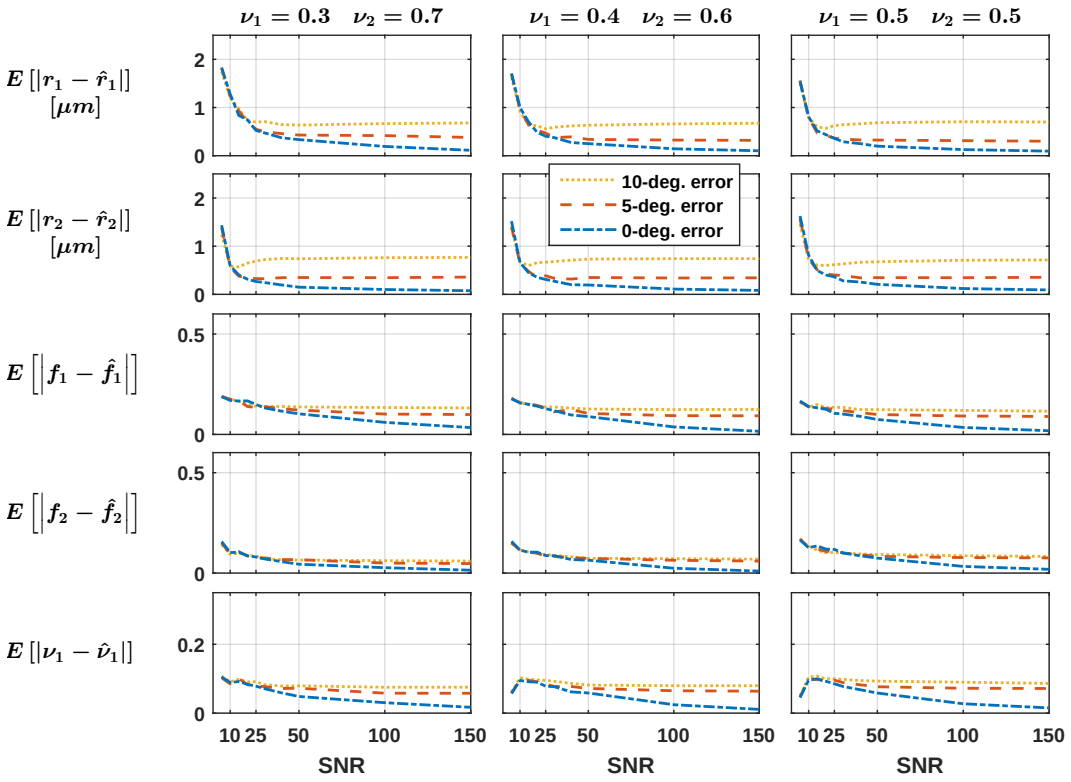

Figure 3.6: (Exp. 2.A) **Errors in the initial estimation of fascicles' orientations yield asymptotically-biased estimates.** Mean absolute error on each fasicle's radius index r_1 and r_2, density index f_1 and f_2 and on the physical volume fraction occupied by the first fascicle ν_1. The groundtruth volume fraction of the second fascicle decreases from left to right and the groundtruth crossing angle was fixed to 60°.

3.4.2.2 Experiment 2.B. Synthetic 2D phantom, effect of dissimilar fascicles

Figure 3.7 suggests that the errors on the estimated microstructural properties converged to zero as the SNR increased. At SNR=25, the average absolute error made on the radius index r expressed as a percentage of the groundtruth value was 33.0% in single fascicles, close to its value of 37.4% in voxels of crossing fascicles (Fig. 3.7(b)). The difference was more pronounced for the density index f with an average 4.94% error in single-fascicle voxels and 29.7% in crossing fascicles. As shown in Fig. 3.7(c), the fastest convergence with SNR in r occurred for Tract 3, which had the largest groundtruth radius index. Convergence for f was slightly faster in Tracts 1 and 2, which were less exposed to fascicle crossings and CSF contamination. The mean absolute angular error on the fascicles' orientations by the ball-and-sticks routine at SNR=25 was as low as 0.43° in single-fascicle voxels and 0.85° in crossing-fascicle voxels, suggesting that the impact of Step 3 on the final errors was minimal.

3.4.2.3 *In vivo* experiment

The top row in Fig. 3.8 suggests that the estimates of both the apparent axonal radius index r and density index f were spatially smooth. The distributions over all voxel-level fascicles (histograms in Fig. 3.8) were smoother for r than f. As reported in Table 3.2, both parameters exhibited a low-high-low trend, with lower mean values in axons passing through the anterior and posterior parts of the CC compared to axons of the mid-anterior, central and mid-posterior CC. The standard deviations for the two parameters were very similar across the five considered sub-regions.

Table 3.2: **Low-high-low trend in apparent radius and density index in CC streamlines.** Mean and standard deviation over all local fascicles or peaks assigned to streamlines passing through five sub-regions of the corpus callosum (CC). Units of apparent radius index in μm.

	apparent radius index	density index
anterior CC	4.06 ± 0.87	0.599 ± 0.12
mid-anterior CC	4.27 ± 0.73	0.636 ± 0.12
central CC	4.19 ± 0.90	0.627 ± 0.12
mid-posterior CC	4.22 ± 0.93	0.645 ± 0.13
posterior CC	3.87 ± 0.99	0.581 ± 0.14

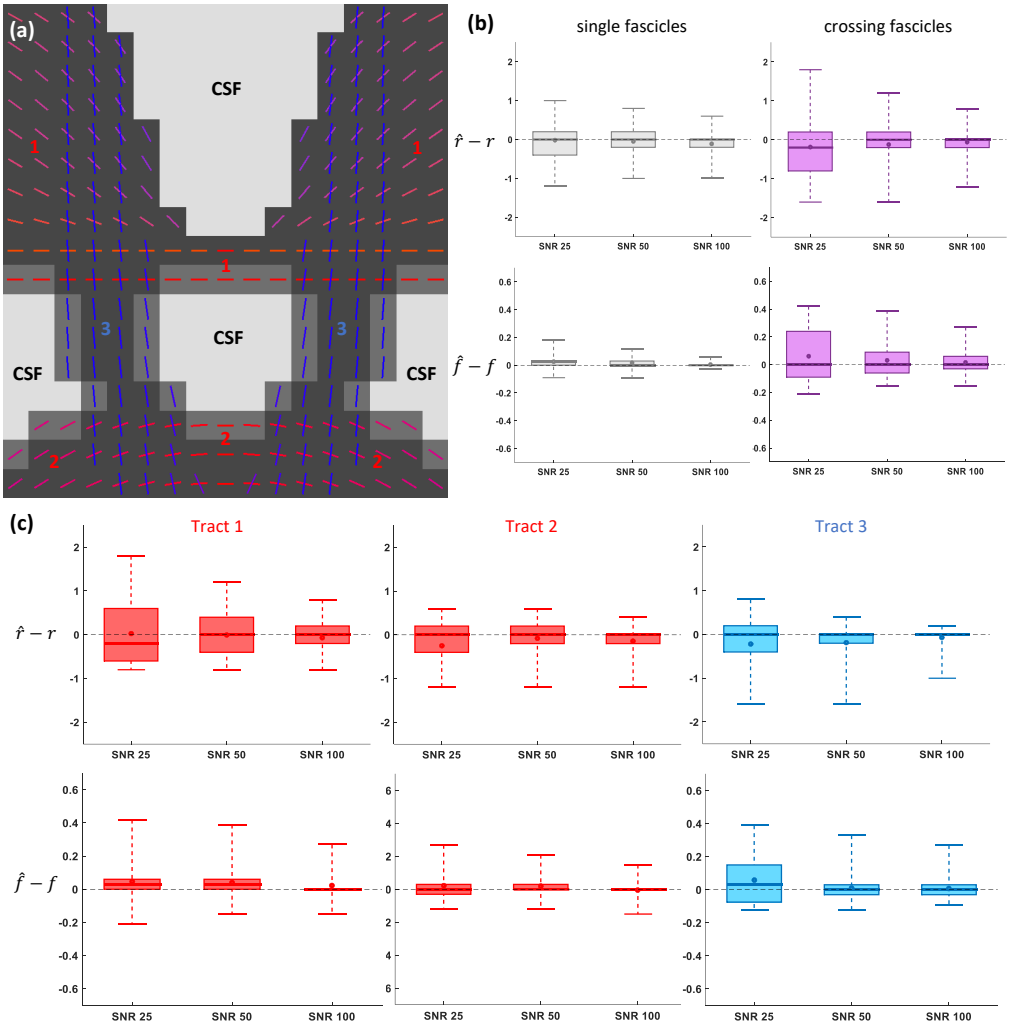

Figure 3.7: (Exp. 2.B) **Tract-specific microstructural estimated enabled by Monte Carlo dictionary estimation.** (a) Color-coded direction of all three axonal tracts with the gray-scale background indicating the level of CSF contamination in each voxel. (b) Signed error made on the estimated radius index r and density index f in regions of single (left) and crossing (right) fascicles. (c) Signed errors for the local voxel fascicles of each tract independently. In (b)-(c), the whiskers of the boxplots extend from the minimum to the maximum value of the data.

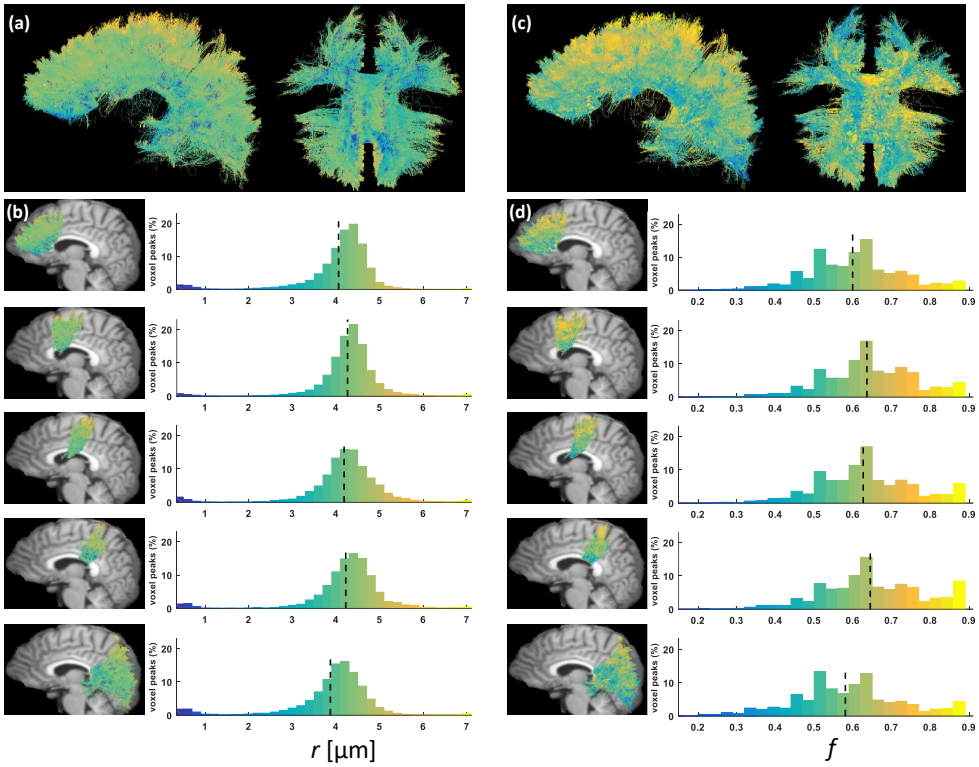

Figure 3.8: **Spatially-smooth estimates of apparent radius and density indices obtained with Monte Carlo dictionary estimation.** (a) Left and inferior view of all the axons passing through the corpus callosum (CC), color-coded by the apparent axonal radius index r. (b) Histograms of estimated apparent radius index in the five axonal tracts studied in the experiment, respectively the anterior, mid-anterior, central, mid-posterior and posterior CC tracts. The mid-sagittal maps of color-coded axons are laid atop a referential T1 image warped into the DW-MRI space. (c)-(d) Same as (a) and (b) for the axonal density index f.

3.5 Discussion

3.5.1 Advantages of Monte Carlo modeling

The main strength of Monte Carlo simulations is their ability to provide exact signals for any fixed tissue geometry. This advantage was made apparent throughout the comparison with MMWMD, which captures the same level of tissue complexity as the particular single-fascicle model selected in this study but differs in the way the signal is formulated. MMWMD for instance failed to consistently estimate r and f in Exp. 1.A, which can only be attributed to the formulation of the extra-axonal signal. In Exp. 1.B, a non-zero CSF weight for the full 4-compartment MMWMD persisted although no isotropic compartment was included in the groundtruth, suggesting that some of the extra-axonal signal may be captured by the isotropic compartment. MMWMD relies on a zeppelin-like diffusion tensor with perpendicular diffusivity proportional to the parallel diffusivity and the extra-axonal volume fraction through a tortuosity model. Monte Carlo simulations on the other hand naturally incorporate physical compatibility between the intra- and extra-axonal compartment.

Our *in vivo* rat model of spinal cord injury demonstrated that the parameters provided by Monte Carlo simulations are generally more specific and interpretable than those of closed-form continuous models (Figure 3.4). Wallerian degeneration is a process with dramatic consequences on the fascicles of the spinal cord which after several weeks is mainly characterized by severe axonal loss (Waller, 1850; George and Griffin, 1994). This was correctly captured by our Monte Carlo dictionary approach but not by MMWMD, which detected an increase in radius index. NODDI attributed all the signal differences to an increase in dispersion and not to a decrease in neurite density. This unexpected result is likely caused by the use of fixed parameters in NODDI which impairs our ability to interpret the remaining free parameters (Scherrer et al., 2016; Jelescu et al., 2016; Hutchinson et al., 2017). The increase in fascicle-specific RD and decrease in fascicle-specific AD detected by DIAMOND were in agreement with a large body of studies correlating diffusion tensor imaging with Wallerian degeneration (Song et al., 2003; Kim et al., 2007; Sun et al., 2008; Zhang et al., 2009; Liu et al., 2013). However no parameter in DIAMOND directly relates to axonal density. The WMTI model predicted a decrease in axonal density similar (although of smaller magnitude) to our Monte Carlo dictionary method as well as changes in extra-axonal RD and AD similar to DIAMOND. Taking the intra-axonal RD as a proxy for axonal radius then the non-significant change agrees with our own findings for r. The good agreement between our approach and WMTI will be investigated in the future. In MAPL, a decrease in RTAP has been shown to represent an increase in mean apparent axonal radius (Fick et al., 2016), similar to what MMWMD detected. Alternatively, considering that Wallerian degeneration causes the number of diffusion barriers to decrease in the extra-axonal space, RTAP could be interpreted as inversely proportional to the extra-axonal RD, which would be in agreement with DIAMOND and WMTI. Similarly, if diffusion is assumed unhindered along the axons, the increase in RTPP can be interpreted as a decrease in AD.
This highlights that signal models such as DIAMOND and MAPL do capture group differences;

however these must be carefully interpreted a posteriori. Our approach outperforms geometric models such as MMWMD and NODDI by ascribing the signal difference to the correct variation in microstructural features.

3.5.2 Whole-brain estimation

The whole-brain HCP experiment was intended to showcase the ability of our method to extract microstructural properties *in vivo* in each voxel locally in areas of crossing fascicles, which is still an open issue in the field. The MIX optimization technique (Farooq et al., 2016) provides a faster and more stable algorithm to fit multi-fascicle extensions of analytical compartment-based models but these models still intrinsically rely on approximate analytical formulas. The multi-fascicle extension of Amico (Daducci et al., 2015) to Amico-X (Auría et al., 2015b) uses a simple diffusion tensor for the extra-cellular signal and does not impose the geometrical compatibility between the intra- and extra-axonal signal that is naturally enforced in Monte Carlo simulations. DIAMOND was also formulated as a multi-fascicle model from the onset but its fascicle-specific parameters are tensor-related quantities and are therefore surrogate measurements of tissue properties.

The low-high-low trend in apparent axonal radius r observed in the human CC (Aboitiz et al., 1992) was found to extend to all callosal axons in our experiment. Remarkably, this pattern was obtained by independent estimations of our model at each voxel and was therefore not a consequence of spatial regularization. In Girard et al. (2017), a similar trend for r was observed on most of the 34 HCP subjects, albeit on DW-MRI data upsampled for tractography analysis. In the same work, the streamline-specific estimates of apparent fiber density computed over all 34 subjects were found to be lowest in the anterior CC, intermediate in the mid-anterior and posterior CC and largest in the central and mid-posterior CC. This roughly coincides with the estimates of f obtained in our particular subject. It should be noted that the values of radius index reported in the experiment ($\approx 4\,\mu m$) are considerably larger than actual measurements of axonal radius in the human brain ($\approx 0.5-1\,\mu m$), as discussed in the next paragraph.

3.5.3 Limitations

In our framework, the tissue geometry selected for the Monte Carlo simulations at the single-fascicle level determines the complexity of the final model. To compare our approach with models of similar tissue complexity such as MMWMD, the simple hexagonal packing geometry was selected in this work. In particular, the use of the single scalar parameter r to characterize the whole intra-axonal signal is known to considerably overestimate actual axonal radii (Alexander et al., 2010; Dyrby et al., 2013) and as such should only be considered as an *index* of apparent axonal radius.

In order to fully exploit the potential of Monte Carlo simulations, future tissue geometries will need to be more realistic and include randomly-located axons with a distribution of radii (Hall et al., 2017), a myelin sheath around axons (Harkins and Does, 2016), axonal undulation (Nilsson et al., 2012), cells with complex morphology such as glia and neurons (Palombo et al., 2016, 2017) or tissue geometries directly obtained from histological slices (Xu et al., 2014). Incorporating axonal orientation dispersion in synthetic substrates for Monte Carlo simulations requires the careful configuration and location of all axons in order to avoid unrealistic intersections. Axonal oscillations with large periodicity may help achieve apparent orientation dispersion (Nilsson et al., 2012).

It is worth recalling that the only simplifying hypothesis made in our framework is that no water exchange occurs *between* fascicles during the acquisition, which allowed us to write the fundamental equation (3.1) as a simple superposition (Rensonnet et al., 2018). This however does not prevent the incorporation of membrane permeability and water exchange *within* a fascicle.

In theory, the sparsity constraints used in Eq. (3.14) do not allow mixtures of fingerprints to reconstruct the signal arising from a single fascicle of axons. This could be a limitation for fascicles consisting of several well-delimited sub-regions exhibiting distinct microstructural properties (e.g., one half with a high and one half with a low axonal density). As discussed in B.2, with the DW-MRI protocols used in this study, the signals of non-uniform voxels were very similar to the fingerprints of uniform voxels with a density index f precisely corresponding to the average packing density of the non-uniform configuration. If non-uniform configurations were not distinguishable from uniform, "average" configurations, a possible workaround would be to add fingerprints arising from non-uniform configurations to the single-fascicle dictionary (Steps 1 and 2 in Section 3.2.4).

The traditional PGSE sequence used in our experiments has been shown to have limited sensitivity to microstructural features such as the axonal radius (Dyrby et al., 2013). Improved sensitivity may be obtained using other diffusion-encoding sequences such as OGSE (Drobnjak et al., 2016; Mercredi and Martin, 2018), STEAM-DTI with varying diffusion times (Fieremans et al., 2016) or b-tensor encoding (Topgaard, 2017). One major advantage of the presented framework is precisely its ability to promptly integrate such extensions, which will be considered in future work.

3.5.4 Fixed parameters and external routines

A number of parameters such as the intrinsic diffusivity and the T2 values of the intra-axonal, extra-axonal and CSF compartments were fixed a priori based on literature values rather than estimated from the data in order to simplify the estimation. Our synthetic experiments have

shown however that the model is quite resilient to slightly misestimated parameter values, in particular for T2 relaxation which in some situations can be exactly compensated by the free weights w in Eq. (3.14). As seen in Exp. 1.B, care must be taken in case multiple T2 values should be present within a single fascicle, as has recently been suggested (Veraart et al., 2018).

It was shown in Experiment 2.A (Figure 3.6) that errors in the orientation of fascicles (estimated with an external routine) could bias the microstructural estimates. To overcome this issue, orientation-estimation methods based on rotationally-invariant dictionary learning (Reisert et al., 2014; Christiaens et al., 2017) could use the estimated fingerprints as their fiber orientation response and retroactively refine the estimated orientations, eventually leading to more accurate microstructural estimates.

3.5.5 Efficiency

The pre-computing stage described as Step 2 of our estimation procedure in Section 3.2.4 may come with a high computational cost. The two canonical single-fascicle dictionaries used in this study required about 35 days worth of computation time on a standard laptop i5 core. In practice this was reduced to about 1-2 days using computing clusters at Université catholique de Louvain depending on cluster load and availability. Algorithmic improvements such as proposed by Hall et al. (2017), optimized implementation and mathematical properties such as presented in Section 3.2.3 should further help assuage the burden of massive Monte Carlo simulations in the future.

Runtime efficiency is not an issue in single-fascicle voxels, where Eq. (3.15) takes less than a second to solve for dictionary sizes N such as used in our experiments. For voxels containing $K > 1$ fascicles, exactly solving N^K convex problems leads to longer computation times and is likely to become a more important issue with larger dictionaries (resulting from a finer resolution in microstructural parameters for instance). Different techniques can be used to reduce the size of the problem, such as a multi-scale optimization approach starting with a coarse-grained dictionary or initial dictionary pruning using sparsity-enforcing methods (Canales-Rodríguez et al., 2015; Canales-Rodríguez et al., 2019).

3.6 Conclusion

A framework was proposed for the estimation of microstructural features incorporating Monte Carlo simulations known for their accurate modeling of the DW-MRI signal. The inverse estimation problem was formulated as a sparse optimization problem on a large pre-computed dictionary and decomposed into many independent convex and easy-to-solve sub-problems. Owing to their unique correspondence with a microstructural configuration, the selected diffusion fingerprints provided the microstructural parameters for each fascicle of axons in

each voxel.

In single-fascicle voxels, our approach demonstrated more accurate, consistent and interpretable results than popular closed-form microstructural models of the literature in many simulation settings as well as in the analysis of an *in vivo* dataset of rat spinal cord. When extended to crossing fascicles, our framework achieved consistent estimates of apparent axonal radius and density indices in synthetic experiments and on whole-brain HCP data. Future work will focus on using a more realistic model at the single-fascicle level and generalizing the microstructural trends found in one HCP subject to larger cohorts.

This work paves the way for microstructure fingerprinting in which Monte Carlo simulations are used as the building blocks of a model of the diffusion signal which directly relate to the underlying microstructure. Our framework offers new opportunities for whole-brain quantitative and interpretable microstructure imaging. Such a capability may prove critical for studies exploring the pathogenesis of neurological and psychiatric disorders as well as in the assessment of responses to treatments.

4 Microstructure fingerprinting with heterogeneous axon diameters

4.1 Introduction

The dictionary matching approach introduced in Chapter 3 was designed to easily accommodate model changes and refinements at the single-fascicle level. The tissue geometry in the Monte Carlo simulations just needs to be updated and the rest of the estimation can proceed seamlessly. In this Chapter, the simple hexagonal-packing model used in the experiments of Chapter 3 was replaced by a more complex single-fascicle tissue configuration, featuring a random packing of straight cylinders with diameter heterogeneity.

This model may help unveil characteristics of the axon diameter distribution (ADD) including its mean (μ_d) and standard deviation (σ_d), as well as estimates of the axon density in each voxel. Such information is key to the understanding of the pathogenesis of diseases like multiple sclerosis (Shintaku et al., 1988; Lovas et al., 2000; Evangelou et al., 2001; DeLuca et al., 2004), amyotrophic lateral sclerosis (Sasaki and Maruyama, 1992; Al-Chalabi and Miller, 2003), Parkinson's disease (Al-Chalabi and Miller, 2003; Burke and O'malley, 2013), Alzheimer's disease (Joyashiki et al., 2011) or degeneration following traumatic injury (Nashmi and Fehlings, 2001; Song et al., 2003; Payne et al., 2011; Maxwell et al., 2015). Non-invasive monitoring of those microstructural properties with diffusion-weighted magnetic resonance imaging (DW-MRI) therefore has great value in the assessment of response to treatments (Horsfield and Jones, 2002; Joyashiki et al., 2011; Burke and O'malley, 2013). Accurate mapping of the ADD is also crucial for the finer study of conduction delays in the mammalian brain (Innocenti et al., 2018; Drakesmith and Jones, 2018; Berman et al., 2019; Deslauriers-Gauthier and Deriche, 2019).

Our extended model, referred to as *heterogeneous fingerprinting* in the rest of this Chapter, was systematically compared to the simpler fingerprinting approach considering hexagonal packing of Chapter 3, referred to as *homogeneous fingerprinting*. The dictionary-based Accelerated Microstructure Imaging via Convex Optimization (AMICO) framework (Daducci et al., 2015) was also included for comparison as it can be used to estimate the apparent ADD and axon density. Unlike our approach, AMICO relies on closed-form analytical expressions

for the signal and separates the contributions of water inside and outside the axons. The three approaches were tested in single-fascicle configurations on a public *ex vivo* dataset of cat spinal cord acquired with an experimental protocol specifically designed to increase the sensitivity of the DW-MRI signal to the ADD (Duval et al., 2016). Segmented histological maps provided a groundtruth to compare the microstructural parameter estimates against. The whole dataset including the *ex vivo* MRI as well as the segmented and unsegmented histological images were downloaded from the White Matter Microscopy Database[1].

4.2 Methods

4.2.1 *Ex vivo* cat spinal cord dataset

The data was originally acquired and presented by Duval et al. (2016). It is described here for ease of reading. A cervical segment of cat spinal cord was collected post-mortem, perfused and post-fixed with paraformaldehyde 4%, and two contiguous pieces 1 cm in length were extracted.

MRI preparation The first piece of tissue was washed in PBS for 5 days at 4 °C before MRI scanning on an Agilent 7T animal scanner able to deliver gradient intensities of up to 600 mT m^{-1} independently in each direction. In all MRI experiments, the tissue was enclosed in a glass tube filled with buffered water and a custom-made solenoid coil was used for transmission and reception (S11 $\approx$ 40 dB). One axial slice of spinal cord was acquired with matrix size fixed to 64x64. In-plane resolution was 0.16 mm $\times$ 0.16 mm and the slice thickness was 0.20 mm.

Diffusion-weighted MRI A single shot EPI sequence was used with BW = 250 kH, TR = 2 s. Two diffusion protocols were considered, corresponding to a 2D and a 3D sampling of the q-space.

In the *2D protocol* summarized in Table 4.1, diffusion was probed at multiple time scales by including 9 pairs of gradient duration δ and separation Δ, inducing time-dependence in the DW-MRI signal (Burcaw et al., 2015). For each pair, 4 b0 images were acquired and gradients were applied along two orthogonal directions in the xy plane perpendicular to the axis of the spinal cord sample. Gradient intensities G were varied along the two directions in respectively 100 and 95 increments. The maximum intensity reached 849 mT m^{-1}, which is expected to enhance the sensitivity of the DW-MRI signal to the axon diameter (Dyrby et al., 2013; Sepehrband et al., 2016). The total number of images was 1791.

In the *3D protocol*, a traditional 4-shell HARDI protocol was employed with $\Delta/\delta = 30/3$ ms, $G = 47.1, 100.8, 300, 600$ mT m^{-1}, $b = 41, 190, 1680, 6725$ s mm^{-2}. The protocol was mainly used for a Diffusion Tensor Imaging (DTI, Basser et al. (1994)) fit in order to correct for slight

[1] hosted on OSF: https://osf.io/abqtz/

Table 4.1: **2-D sampling of the q-space for enhanced sensitivity to axon diameter distribution.** For each of the nine combinations of (Δ, δ), 100 magnetic gradients were applied along the line of direction $\left[1/\sqrt{2}, 1/\sqrt{2}, 0\right]^T$ and 95 along the orthogonal direction $\left[-1/\sqrt{2}, 1/\sqrt{2}, 0\right]^T$ with intensity G gradually incremented. The whole protocol comprises 1755 diffusion-weighted images and 36 b0 acquisitions (1791 images in total).

	Δ[ms]	δ[ms]	TE [ms]	G[mT m^{-1}]	b[s mm^{-2}]	#b0
1	7	3	36	9-849	0-2783	4
2	12	8	46	9-849	0-30780	4
3	15	8	46	9-849	0-40674	4
4	20	8	46	9-849	0-57163	4
5	25	8	47	9-849	0-73653	4
6	30	8	52	9-849	0-90142	4
7	35	8	57	9-849	0-106631	4
8	40	3	57	9-849	0-18087	4
9	40	8	62	9-849	0-123121	4

misalignment of the imaged spine.

Gibbs unringing was performed (Kellner et al., 2016) on both DW-MRI datasets, after which 15 DW images were discarded from each dataset based on visual inspection.

Macromolecular Tissue Volume imaging Macromolecular Tissue Volume (MTV), a proxy for myelin content, was measured using the procedure described in Mezer et al. (2013). In order to derive Myelin Volume from MTV, a scaling factor of 1.65 evaluated from data acquired on monkey corpus callosum was used (Stikov et al., 2015).

Histology and segmentation The second piece of spinal cord was stained with osmium 4%, dehydrated, embedded in paraffin, cut in 4 µm slices and imaged using an optical 20x whole slice microscope (Hamamatsu NanoZoomer 2.0-HT) with resolution 230nm/px. Axon segmentation was performed automatically with the publicly available software AxonSeg (Zaimi et al., 2016)[2]. After segmentation, the resolution of the images was decreased and the axon properties averaged over squares 150 µm × 150 µm and registered to the MRI space. This produced histological maps of number-weighted (NW) axon diameter mean (μ_d) and standard deviation (σ_d), intracellular volume fraction (icvf), g-ratio, myelin volume fraction (mvf) and axon count (N_{ax}). Additionally, a volume-weighted (VW) axon diameter mean map was computed from the NW map and a smoothed version of the icvf map ($\tilde{\text{icvf}}$) was created.

[2] https://github.com/neuropoly/axonseg

4.2.2 Data analysis

Heterogeneous fingerprinting Monte Carlo simulations were performed in geometries consisting of $N_{\text{cyl}} = 1000$ randomly-packed straight cylinders (Figure 4.1, left) with diameters drawn from 109 different gamma distributions of pdf $\Gamma\left(\mu_d, \sigma_d\right)$ with (number-weighted) mean $\mu_d \in [0.4 : 0.4 : 8]\,\mu\text{m}$ and standard deviation σ_d such that $\sigma_d / \mu_d \in [0.2 : 0.1 : 0.8]$, and intracellular volume fraction icvf $\in [0.02 : 0.02 : 0.80]$. After cylinder packing, geometries containing one or more cylinders with a diameter over $20\,\mu\text{m}$ were discarded. This left $N = 4120$ single-fascicle configurations for the dictionary, with each fingerprint related to the microstructural parameters $\Omega = (\mu_d, \sigma_d, \text{icvf})$ along a given orientation, from which the volume-weighted diameter mean μ_d could also be computed. The cylinder packing and Monte Carlo simulations of the random walk of molecules were performed by our in-house C/C++ code using spatial gridding optimization for efficient intersection checking, similarly to Nedjati-Gilani et al. (2017) and Hall et al. (2017). Both the 2D and the 3D diffusion protocols were considered.

In the case of a single fascicle of axons, the inverse problem of Section 3.2.4 becomes

$$\begin{aligned} \mathbf{w}^* = \quad & \underset{\substack{\mathbf{w}_{\text{fasc}} \geq 0 \\ w_{CSF} \geq 0}}{\text{argmin}} \quad \left\| \mathbf{y} - \left[\mathbf{A}_{\text{fasc}}(\Omega_1, \mathbf{u}_{\text{fasc}}), \ldots, \mathbf{A}_{\text{fasc}}(\Omega_N, \mathbf{u}_{\text{fasc}}) | \mathbf{A}_{CSF} \right] \cdot \begin{bmatrix} \mathbf{w}_{\text{fasc}} \\ w_{CSF} \end{bmatrix} \right\|_2^2 \\ & \text{subject to} \quad |\mathbf{w}_{\text{fasc}}|_0 = 1, \end{aligned} \tag{4.1}$$

where $\mathbf{u}_{\text{fasc}}$ is the estimated fascicle orientation. Equation (4.1) was solved exactly in each voxel through dictionary look-up with non-negativity constraints.

Homogeneous fingerprinting A slightly larger sampling of $N = 986$ combinations $\Omega_i = \left(\mu_{d,i}, \text{icvf}_i\right)$ was performed compared to the hexagonal-packing dictionaries used in Chapter 3, with icvf ranging from 0.03 to 0.87 in steps of 0.03. Estimation was similar to Eq. (4.1).

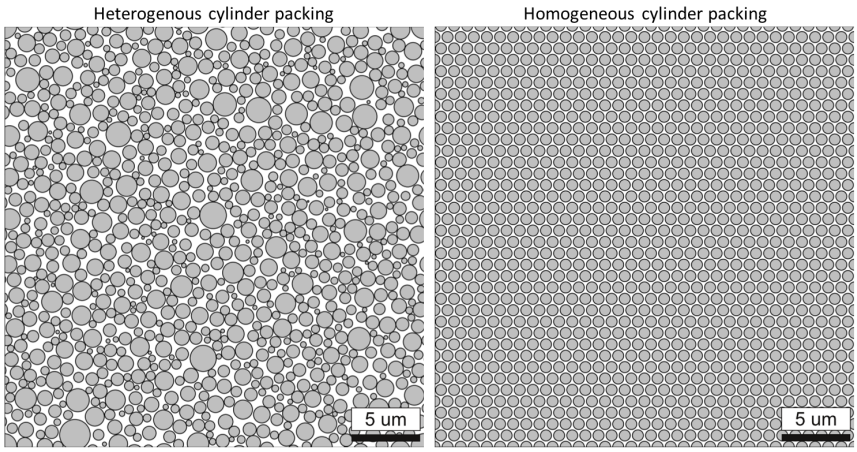

Figure 4.1: **Realistic Monte Carlo simulations for microstructure estimation.** Example of a random packing of straight cylinders with heterogeneous diameters (left) and of a regular hexagonal packing of identical cylinders (right). The Monte Carlo simulations of the self-diffusion of water molecules performed in those geometries produced the diffusion signatures or *fingerprints* used in the estimation.

AMICO In AMICO (Daducci et al., 2015), the intra-axonal contribution is assumed to arise from a combination of straight parallel cylinders with diameters sampled at N_{in} values $d_1,\ldots,d_{N_{\text{in}}}$, with $d_{N_{\text{in}}} = 20\mu\text{m}$ to match the heterogeneous fingerprinting approach. The extra-axonal signal is separately modeled by a zeppelin-like diffusion tensor with perpendicular diffusivity $D_\perp$ sampled at N_{ex} values $D_{\perp 1},\ldots,D_{\perp N_{\text{ex}}}$, where the maximum value $D_{\perp N_{\text{ex}}}$ depends on the intrinsic parallel diffusivity D via a tortuosity model. AMICO solves, in each voxel,

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}\geq 0} \left\| \mathbf{y} - \left[\mathbf{A}_{\text{cyl}}(d_1),\ldots,\mathbf{A}_{\text{cyl}}(d_{N_{\text{in}}}) | \mathbf{A}_{\text{zep}}(D_{\perp 1}),\ldots,\mathbf{A}_{\text{zep}}(D_{\perp N_{\text{ex}}}) \right] \cdot \begin{bmatrix} \mathbf{w}_{\text{in}} \\ \mathbf{w}_{\text{ex}} \end{bmatrix} \right\|_2^2 + \lambda \left\| \mathbf{w} \right\|_2^2,$$

where A_{cyl} and A_{zep} are the signal contributions of individual cylinders and zeppelins respectively and where λ is a regularization parameter (compare to Eq. (4.1)). From $\hat{\mathbf{w}} = [\hat{\mathbf{w}}_{\text{in}}, \hat{\mathbf{w}}_{\text{ex}}]^T$, one obtains the icvf = $|\hat{\mathbf{w}}_{\text{in}}|_1 \,/\, |\hat{\mathbf{w}}|_1$ of the voxel, the volume-weighted mean axon diameter $\mu_d = \sum_{j=1}^{N_{\text{in}}} \hat{w}_{\text{in},j} d_j$ and a standard deviation σ_d after converting to the number-weighted ADD $\hat{w}_{\text{in},j} \leftarrow \frac{\hat{w}_{\text{in},j}/d_j^2}{\sum_{k=1}^{N_{\text{in}}} \hat{w}_{\text{in},k}/d_k^2}$.

Model comparison A region of interest (ROI) in the white matter was defined as those voxels in which the histology-derived axon count was non-zero and the myelin volume fraction was over 0.20. All three estimated models ignored any CSF contribution ($w_{\text{csf}} = 0$) and used an intrinsic diffusivity D computed from an initial DTI fit in each voxel using the 3D protocol and averaged over all voxels of the ROI. The fascicle orientations were estimated from the DTI fit too but were kept individually in each voxel. Voxel-wise correlations and mean absolute errors

(MAE) were computed between histology-derived μ_d, σ_d, icvf and the corresponding model indices.

4.3 Results

Our heterogeneous microstructure fingerprinting was the only model to produce statistically-significant correlations for the three axonal indices, as shown in Figures 4.2 and 4.3.

The axon diameter indices μ_d of the fingerprinting approaches exhibited slightly lower MAEs and higher correlations than AMICO. The estimates of the heterogeneous fingerprinting approach were more evenly distributed around the histological values while the homogeneous fingerprinting and AMICO underestimated larger diameters. In voxels where the VW distribution mean computed from segmented histology was greater than 5µm, our heterogeneous fingerprinting approach yielded a MAE of 0.726µm, significantly lower than the homogeneous fingerprinting and AMICO approaches, which respectively obtained 1.075µm ($p = 5.52 \times 10^{-6}$) and 1.45µm ($p < 1 \times 10^{-16}$), where the p-values were obtained from a two-sample, single-tailed t-test.

Our approach yielded estimates of σ_d with higher correlation and lower MAE than AMICO, in which all the σ_d estimates neared zero, suggesting a degenerate optimum with one single dominant cylinder j with diameter d_j for the intracellular signal: $|\hat{\mathbf{w}}_{\text{in}}|_1 \approx \hat{w}_{\text{in},j}$. AMICO obtained a stronger correlation for the icvf index but a larger MAE due to systematically underestimating the histological values.

For the heterogeneous fingerprinting approach, the MAEs indicated that the error on the estimated mean diameter was of the order of half a micron, the error on the standard deviation of the order of a third of a micron and the error on the icvf was about 0.045.

Table 4.2 and Figure 4.4 further investigate the specificity of heterogeneous fingerprinting by displaying the correlations, within our white matter ROI, between histological measurements and the indices obtained by the heterogeneous fingerprinting approach. The highest correlation between a histological variable and our NW μ_d occurred with its NW μ_d counterpart; for our NW σ_d it occurred with the VW μ_d, closely followed by its NW σ_d counterpart; for our VW μ_d it occurred with its VW μ_d counterpart; and for our icvf it occurred with the smoothed icvf from histology (see boldfaced values in Table 4.2).

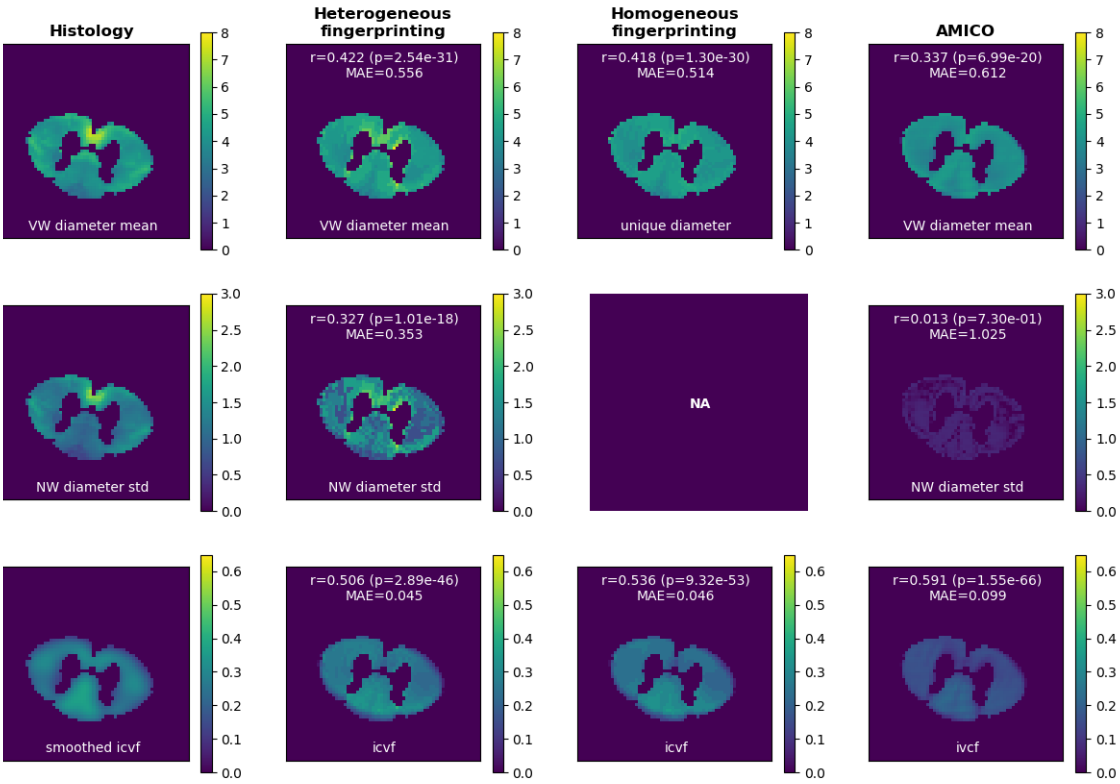

Figure 4.2: **Quantitative estimation of microstructural ADD and density enabled by heterogeneous fingerprinting.** Maps of histological measurements (1st column), Monte Carlo fingerprinting with heterogeneous diameters (2nd column), Monte Carlo fingerprinting with hexagonally-packed identical cylinders (3rd column) and AMICO (4th column) on the 2D slice of fixed cat spinal cord. The reported Pearson correlation coefficient (r), associated p-value (not corrected for multiple comparisons) and mean absolute error (MAE) are between the model indices and the reference histology values for the volume-weighted (VW) axon diameter mean μ_d (1st row, in µm), the axon diameter standard deviation σ_d (2nd row, in µm) and the intra-axonal volume fraction icvf (3rd row).

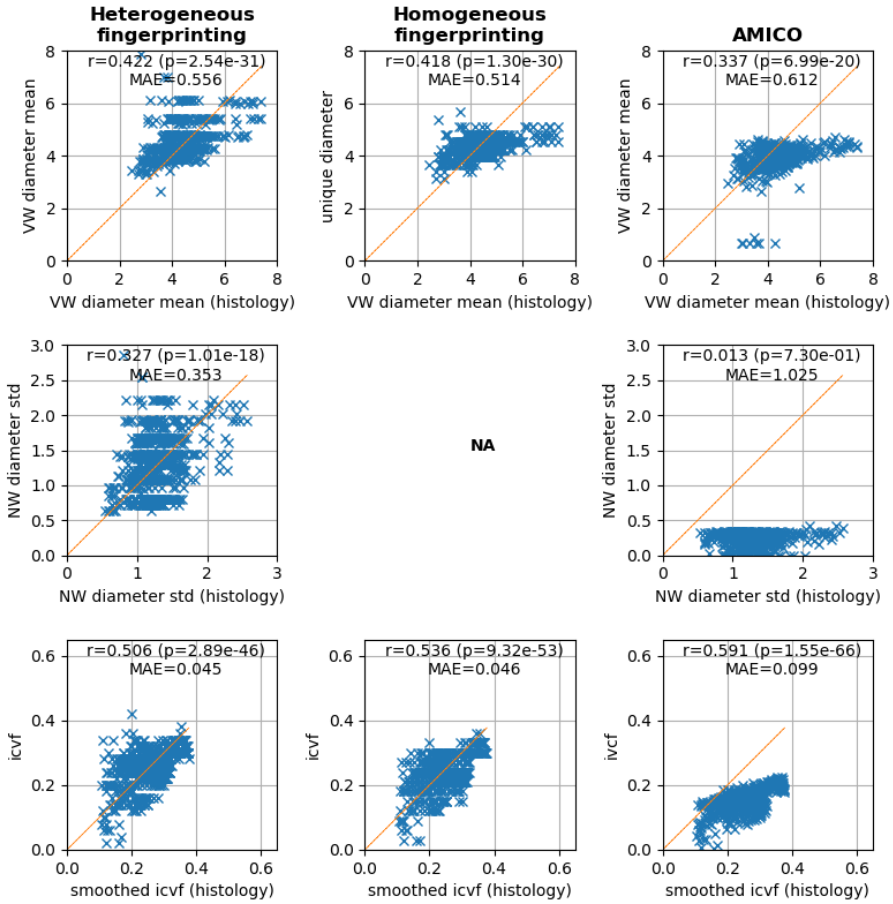

Figure 4.3: **Quantitative estimation of microstructural ADD and density enabled by heterogeneous fingerprinting.** Agreement between the reference histological measurements (x-axes) and the estimated model parameters (y-axes) for the volume-weighted (VW) diameter mean μ_d (1 st row, in µm), standard deviation σ_d (2nd row, in µm) and intra-axonal volume fraction icvf (3rd row). The voxel-wise Pearson correlation coefficient (r), associated p-value (not corrected for multiple comparisons) and mean absolute error (MAE) are reported for each comparison.

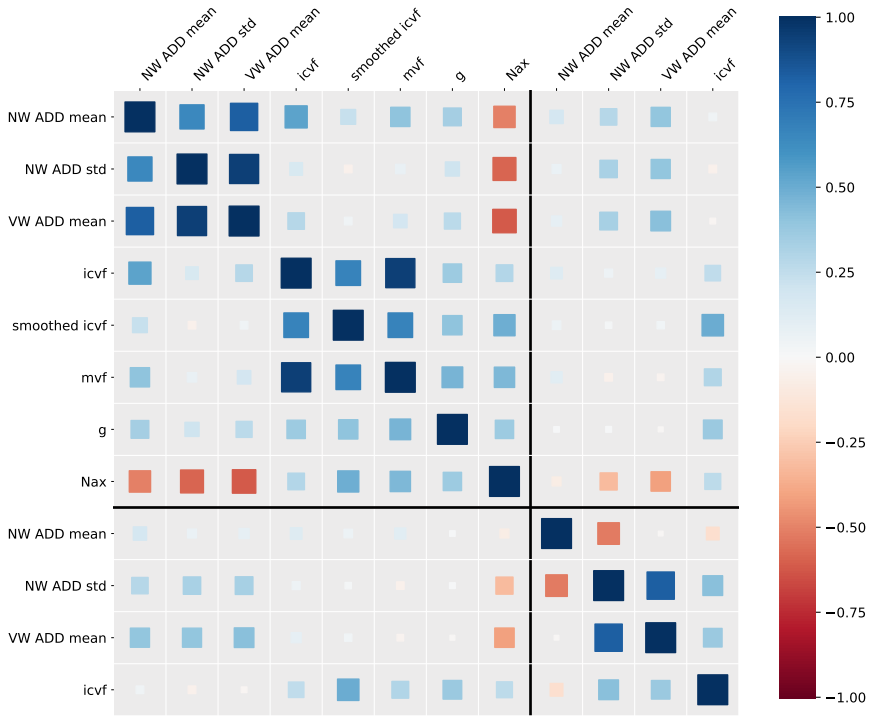

Figure 4.4: **Heterogeneous fingerprinting estimates show high correlations with biologically relevant parameters.** Pearson correlation coefficients within histological measurements (top left), within the indices of our heterogeneous fingerprinting approach (bottom right) and between histology and the proposed approach (top right or bottom left). See Table 4.2 for exact values. NW and VW indicate number- and volume-weighted respectively.

Table 4.2: **Heterogeneous fingerprinting estimates show high correlations with biologically relevant parameters.** Pearson correlation coefficients within and between histological measurements and the estimates of the proposed approach. Boldfaced values highlight the highest correlations, for each of the four fingerprinting indices, with histological variables. NW and VW indicate number- and volume-weighted respectively.

	Histology								Heterogeneous fingerprinting			
	NW μ_d	NW σ_d	VW μ_d	icvf	$\tilde{\text{icvf}}$	mvf	g	N_{ax}	NW μ_d	NW σ_d	VW μ_d	icvf
NW μ_d	1.000	0.648	0.830	0.540	0.232	0.406	0.342	-0.501	**0.183**	0.289	0.401	0.046
NW σ_d		1.000	0.951	0.163	-0.046	0.081	0.209	-0.582	0.071	0.327	0.397	-0.047
VW μ_d			1.000	0.292	0.046	0.187	0.275	-0.610	0.097	**0.339**	**0.422**	-0.017
icvf				1.000	0.673	0.953	0.368	0.296	0.135	0.053	0.096	0.256
$\tilde{\text{icvf}}$					1.000	0.672	0.407	0.493	0.064	0.026	0.040	**0.506**
mvf						1.000	0.471	0.454	0.120	-0.045	-0.029	0.304
g							1.000	0.365	0.015	0.018	-0.007	0.380
N_{ax}								1.000	-0.071	-0.315	-0.411	0.270
NW μ_d									1.000	-0.514	-0.005	-0.165
NW σ_d										1.000	0.833	0.427
VW μ_d											1.000	0.377
icvf												1.000

4.4 Discussion

Estimation of the ADD AMICO has adopted a non-parametric approach to estimate a discretized version of the whole ADD while heterogeneous fingerprinting a priori imposed a smooth Gamma distribution on the diameters of the cylinders used in the Monte Carlo simulations. The results obtained in our experiment suggested the potential of using such a prior parameterization of the ADD for regularizing the model fitting. Despite its theoretical promise to reconstruct any distribution non-parametrically, AMICO consistently obtained degenerate distributions with one single cylinder contribution having a non-zero weight. This led to zero standard deviations in almost all voxels, which was inconsistent with histological measurements. Homogeneous fingerprinting was by definition unable to capture diameter heterogeneity but its estimates of the mean of the ADD were very close to those of the heterogeneous approach, which highlights the value of its unique diameter index.

Estimation of the intracellular volume fraction One of the main differences between AMICO and the fingerprinting approaches is the separation of the intra- and extra-axonal signal contributions. AMICO consistently underestimated the icvf index and suffered from artifactual transfer of weight between $\mathbf{w}_{\text{in}}$ and $\mathbf{w}_{\text{in}}$. Preventing this requires careful tuning of the regularization parameter λ prior to estimation whereas in fingerprinting approaches, the relative weights of the intra- and extra-axonal contributions are constrained in the Monte Carlo simulations directly. AMICO may also have suffered from the use of diffusion tensors for the

extra-axonal signal. Tensors fundamentally ignore diffusion hindrances and time-dependence effects(Burcaw et al., 2015), which become especially visible when multiple diffusion times are used such as in our experiments. Diffusion hindrance and time-dependence are naturally handled by MC simulations. This could be improved in the AMICO framework in the future by incorporating more advanced analytical model of extra-axonal diffusion based on generalized higher order tensors(Romascano et al., 2019).

Limitations The automatic segmentation of the histological images was subject to algorithm-specific bias and was ultimately limited by the precision of the optic microscope and the thickness of the tissue slice. The marked difference between the icvf and the smoothed icvf correlations indicated that the results were sensitive to that aspect of the data processing.

The reference histological values were all within fairly narrow ranges, which is typical of healthy tissue but prevented us from testing our models on wide range of possible parameter values. For instance, as apparent in Figure 4.3, all histologically-computed icvf values were below 0.40. In a small range, correlations can be severely impacted even at moderate absolute errors, which suggests that the reported correlation coefficients may have been underestimated.

Our heterogeneous fingerprinting approach ignored important microstructural features such as axonal undulation and myelin sheaths. If accounted for, these features may help further improve the estimation of the remaining tissue parameters.

4.5 Conclusion

On *ex vivo* cat spinal cord data imaged at 7T with ultra-high gradient intensities, indices of mean and standard deviation of the ADD and of axon density obtained by our new heterogeneous fingerprinting approach positively correlated with histological measurements. Our method outperformed a simpler fingerprinting approach assuming a homogeneous packing of identical cylinders. It also compared favorably with the AMICO framework, which relies on approximate analytical expressions for the signal.

Future work will focus on incorporating a myelin sheath in our single-fascicle tissue model. Prior information on myelin volume fraction obtained from MTV imaging, for instance, could be used to improve the estimation of the ADD and intra-axonal fraction. Hexagonal-packing with released intrinsic diffusivity or random packings of identical cylinders would also be interesting to compare to heterogeneous fingerprinting, as those approaches would have an indentical number of parameters. Finally, performing the same experiments with oscillating-gradient waveforms able to probe smaller length scales than PGSE acquisitions will likely lead to interesting observations (Drobnjak et al., 2016; Mercredi and Martin, 2018).

5 Accelerated microstructure fingerprinting using neural networks

5.1 Introduction

The dictionary matching approach presented in Chapters 3 and 4 offered the benefit of increased robustness in the inverse problem compared to traditional continuous optimization while leveraging the superior modeling ability of Monte Carlo simulations. As hinted at in Section 3.5.5 however, an important drawback of the exhaustive fingerprint search presented in Step 5 of the inverse problem (see Eq. (3.15) in Section 3.2.4) is its poor scaling with the size of the dictionary, especially in brain voxels containing crossing fascicles. If N represents the number of fingerprints in a single-fascicle dictionary and K the number of fascicles in a voxel, the complexity of the runtime estimation was shown to be $\mathcal{O}(N^K)$. Since N grows exponentially in the number of parameters P describing a single fascicle of axons, the computational requirements can quickly become intractable for crossing fascicles when more complex models are used such as the one selected in Chapter 4. This will eventually become a bottleneck for future large-scale population studies.

Deep neural networks emerge as natural candidates to overcome this issue. They are able to learn complex mathematical mappings given enough training data, generalize well to unseen data and have a very efficient forward evaluation pass once training has been performed. Their major disadvantage is their black-box like behavior as their inner workings are sometimes difficult to interpret in terms of the data (Schmidhuber, 2015).

In the field of Magnetic Resonance Fingerprinting, the dictionaries initially used to estimate two tissue parameters (T1, T2) have been replaced by new dictionaries several orders of magnitude larger in order to include more complex tissue representations. Speed-up factors between 300 and 5000 have been reported by Cohen et al. (2018) using an extremely simple two-layer perceptron with 300 hidden units per layer, trained on a pre-computed dictionary of 69 000 fingerprints. A similar network architecture with the addition of 1-D batch normalization layers was used by Zhang et al. (2019) and trained on synthetic data artificially corrupted by noise. In their preliminary results, a speed-up factor of 7500 was reported for the estimation of 4 tissue parameters compared to exhaustive matching using a dictionary of 6 149 000 entries.

Data-driven approaches have also been investigated with networks trained with conventional MRI maps such as T1, T2 or ADC used as training labels. The task is then to learn the mapping from the fast, potentially under-sampled temporal sequences used in MR fingerprinting and those conventional labels, taken as groundtruth tissue properties (Pirkl1 et al., 2019).

In microstructure imaging based on diffusion-weighted MRI (DW-MRI), data-driven deep learning has been used to learn the parameters of the NODDI model from under-sampled data and thereby reduce the required scan time while accelerating the estimation (Golkov et al., 2016; Ye, 2017; Gibbons et al., 2019). More traditional machine learning techniques such as random forest regressors have also been used to learn a mapping from a dictionary of Monte Carlo signals to the tissue parameters (Nedjati-Gilani et al., 2017; Hill et al., 2019; Ginsburger et al., 2019b). In those approaches, Diffusion Tensor Imaging (DTI) and spherical harmonics features were first extracted from the DW-MRI signal to avoid having to generate a new dictionary and retrain the regressor for every new acquisition protocol. This came at the cost of partially losing the advantage of modeling the signal with Monte Carlo simulations however. In addition, spherical harmonics usually require HARDI-like acquisition protocols, which can prove restrictive.

In this chapter, a 2-stage estimation procedure is introduced, designed to increase efficiency and preserve accuracy while maintaining sufficient interpretability. The first stage is a simple and efficient non-negative (linear) least squares (NNLS) estimation on the multiple-fascicle dictionary used in the usual estimation procedure described in Section 3.2.4. Its output vector of KN non-negative entries can be interpreted as a feature vector in the space of fingerprints, indicative of the relative weight of each Monte Carlo fingerprint in the DW-MRI signal. In the second stage, the NNLS output (the feature vector) is fed to a neural network with an architecture reflective of the multi-fascicle nature of the signal and trained on a synthetic dataset of 500 000 DW-MRI fingerprints. The final microstructural parameters of each fascicle are then estimated almost instantaneously with a single forward pass of the network.

The framework was tested on a variety of unseen data after training, both synthetic and *in vivo*. Its performance both in parameter accuracy and in execution time was systematically compared to the reference dictionary matching procedure presented in Chapter 3.

5.2 Methods

5.2.1 Datasets

The three following test datasets were used throughout our experiments.

Synthetic HCP dataset This is the dataset of 86 400 synthetic crossing-fascicle voxels described in Experiment 2.A. of Section 3.3.2.3 for validation of the reference dictionary matching method in areas of crossing fascicles. The two fascicles in each voxel were hexagonally-

packed straight cylinders described by an identical radius index $r_1 = r_2 = r$ and density index $f_1 = f_2 = f$, with volume fractions v_1 and v_2 summing to 1, crossing at an angle 30, 60 or 90°. The DW-MRI signals were simulated using the multi-shell HARDI, high-gradient MGH-HCP protocol (Setsompop et al., 2013) described in Section 3.3.2.1, which contains $M = 512$ DW-MRI measurements. The signals were corrupted with artificial Rician noise with SNR levels varying between 5 and 150. The single-fascicle dictionaries provided to the NNLS estimation in each voxel had size $N = 782$ and were intentionally deviated from the groundtruth orientations to assess the effect of angular errors of 0, 5 and 10 degrees.

Synthetic Rodent dataset This dataset was simulated exactly like the Synthetic HCP dataset, except that the DW-MRI signals were simulated using the multi-shell HARDI, high-gradient Rodent protocol described in Section 3.3.1.1, which was originally designed for spinal cord imaging. Note that T2 relaxometry values for a B_0 field strength of 11.7T were used, as detailed in Sections 1.2.1 and 3.3.2.2.

In vivo HCP dataset The same healthy subject as in Section 3.3.2.4 was selected from the MGH Adult Diffusion data release[1] (Setsompop et al., 2013).

5.2.2 First stage: blind NNLS estimation

Given a vector $\mathbf{y} \in \mathbb{R}^M$ of DW-MRI measurements and assuming K fascicles with orientations $\mathbf{u}_1, \ldots, \mathbf{u}_K$ have been detected in Step 3 of our estimation procedure (Section 3.2.4), the following NNLS problem was solved:

$$\hat{\mathbf{w}} = \underset{\mathbf{w} \geq 0}{\operatorname{argmin}} \quad \left\| \mathbf{y} - \mathscr{D} \cdot \mathbf{w} \right\|_2^2,$$

where the optimization completely ignored the structure of the dictionary $\mathscr{D} = \left[\mathbf{F}^1 | \ldots | \mathbf{F}^K \right]$ (with optionally a column for isotropic CSF contribution) in which each $\mathbf{F}^k$ is the pre-computed single-fascicle dictionary rotated along the estimated orientation $\mathbf{u}_k$, with $1 \leq k \leq K$.

The minimization was performed via the well-known active-set algorithm (Lawson and Hanson, 1995), using an efficient in-house version which notably avoids computing the costly matrix-matrix product $\mathscr{D}^T \mathscr{D}$. In this algorithm, the KN variables of the problem are either forced to zero (the *active set*, because they make the inequality constraints active) or let free (the *free set*). At each iteration, the (zero) variable from the active set having the most negative partial derivative is moved to the free set and an unconstrained least squares problem is solved considering just the current subset of free variables. If the new solution is not acceptable (i.e., some free variables turned negative), an intermediate iterate is found between the previous solution (which was acceptable) and the current, non-acceptable one. The algorithm stops

[1] https://www.humanconnectome.org/study/hcp-young-adult/document/mgh-adult-diffusion-data-acquisition-details

when all the zero variables in the active set have positive or zero partial derivatives.

Since the number of variables is finite and the objective decreases at each iteration, the algorithm always converges. Because the problem is convex, the optimum is always global. If the solution $\hat{\mathbf{w}}$ is sparse (i.e., the optimal free set has size $p << KN$), the algorithm is made particularly efficient by selecting $\hat{\mathbf{w}} = \mathbf{0}$ as initial iterate, which is always an acceptable solution and corresponds to an empty free set. There is no theoretical guarantees on the number of iterations until convergence to the optimal free set; in fact the worst-case complexity $\sum_{i=1}^{KN} \binom{KN}{i}$ behaves exponentially in N. However, the optimal free set is usually found quickly in practice (Slawski and Hein, 2011), which is reminiscent of the Simplex algorithm for linear optimization programs (Dantzig et al., 1955). Furthermore, the free set typically contains p variables or fewer, making the unconstrained least squares problems very fast to solve at each iteration.

5.2.2.1 Experiment 1.A. Sparsity of the NNLS output on synthetic data

NNLS estimation was performed on the 86 400 synthetic crossing-fascicle configurations of the Synthetic HCP dataset described above. The number of non-zero weights of the NNLS output $\hat{\mathbf{w}}$ was averaged over all noise repetitions and groundtruth parameter values and reported for each fascicle separately, as a function of increasing SNR.

5.2.2.2 Experiment 1.B. Sparsity of the NNLS output on *in vivo* data

NNLS estimation was performed in those 217 373 white matter voxels of the selected HCP subject in which the ball-and-stick estimation (Step 3 of the estimation procedure in Section 3.2.4) detected two crossing fascicles. Histograms of the number of non-zero weights was reported for each fascicle separately.

5.2.2.3 Experiment 1.C. Complexity of first-stage NNLS vs dictionary size

This experiment aimed at investigating the runtime complexity of the NNLS estimation as a function of the size N of the single-fascicle dictionary. Because of its large size $N = 5378$, the dictionary $\mathbf{F}^0 = [\mathbf{A}_1^0, \ldots, \mathbf{A}_N^0]$ incorporating a random packing of cylinders with diameter heterogeneity and based on the 3D, 4-shell HARDI protocol from Chapter 4 was considered. That protocol contains $M = 781$ DW-MRI measurements. NNLS estimation was performed for N ranging from 782 (the size used in Chapter 3 and in the remainder of this Chapter) to 5378 by randomly selecting subsets of the dictionary columns. For larger values of N, a dummy dictionary with entries uniformly distributed in $[0, 1]$ was generated, independently for each repetition. Experiments up to $N = 55\,000$ were performed.

For each dictionary size N, the following procedure was repeated 50 times. The dictionary was rotated along two directions $\mathbf{u}_1$ and $\mathbf{u}_2$ randomly selected on the 3D sphere with a crossing

angle fixed to 30°, one single fingerprint was randomly selected ($1 \leq j_1, j_2 \leq N$) from each of the two rotated single-fascicle dictionaries $\mathbf{F}^1$ and $\mathbf{F}^2$. Superposition was performed as $\mathbf{y} = v_1\mathbf{A}^1_{j_1} + v_2\mathbf{A}^2_{j_2}$ where the volume fraction v_1 was uniformly distributed in $[0.2, 0.8]$ and $v_2 = 1 - v_1$. The final synthetic DW-MRI signal $\mathbf{y}$ was obtained after corruption by Rician noise with signal-to-noise ratio (SNR) uniformly distributed in $[5, 50]$. The execution time of the NNLS estimation was recorded and averaged over the 50 repetitions.

5.2.3 Second stage: neural network regression

In the second stage of the accelerated estimation, the output $\hat{\mathbf{w}}$ of the first-stage NNLS estimation was passed through the neural network depicted in Figure 5.1. Its architecture exploited the multi-fascicle nature of the problem: each sub-vector $\hat{\mathbf{w}}_{\mathbf{F}_k}$ of $\hat{\mathbf{w}}$ was first processed by a "split" independent multi-layer perceptron (MLP) containing N input units followed by $\left(L_{\text{split}} - 1\right)$ fully-connected layers of H_{split} hidden units each (blue in the Figure). Splitting the input had the advantage of dramatically reducing the number of model parameters while accelerating the learning of fascicle-specific features by preventing coadaptation of the model weights (Hinton et al., 2012). The outputs of the split networks were merged and rescaled in a 1-dimensional batch normalization layer (Ioffe and Szegedy, 2015). A final MLP with L_{final} layers containing H_{final} hidden units each (green in Figure 5.1) performed the final regression task and computed the microstructural features of each fascicle. The output of all fully-connected layers of all sub-networks was systematically passed through a rectified linear unit (ReLU) layer (not shown in Figure 5.1). Before splitting, $\hat{\mathbf{w}}$ was made to have zero mean and unit variance, which is known to improve the back-propagation procedure used to train deep neural networks (Glorot and Bengio, 2010). The network possessed 208 006 free parameters, which had to be learned during training.

5.2.3.1 Network training

Training set A total of 500 000 synthetic voxel acquisitions were simulated from the hexagonal-packing Monte Carlo dictionary $\mathbf{F}^0 = \left[\mathbf{A}^0_1, \ldots, \mathbf{A}^0_N\right]$ containing $N = 782$ fingerprints used in Chapter 3, generated with the HCP protocol. For each voxel, two dictionary fingerprints $\mathbf{A}^0_{j_1}$ and $\mathbf{A}^0_{j_2}$ were selected randomly ($j_1, j_2 \in \{1 \ldots N\}$) and rotated along unit directions $\mathbf{u}_1$ and $\mathbf{u}_2$ uniformly distributed on the 3D sphere with a minimum crossing angle of 15°. The rotated signals were weighted by volume fractions v_1 uniformly distributed in $[0.15, 0.85]$ and $v_2 = 1 - v_1$ and corrupted by Rician noise with SNR uniformly distributed in $[4, 100]$. Those voxels were therefore similar in their construction to the Synthetic HCP dataset described above; however they were simulated completely independently from it with broad parameter ranges informed by the physics of the problem rather than by the properties of that specific dataset. The first-stage NNLS estimation was performed on each of the simulated DW-MRI acquisitions, with a dictionary $\mathscr{D}$ containing $\mathbf{F}^0$ rotated along the groundtruth orientations $\mathbf{u}_1$ and $\mathbf{u}_2$ (i.e., the orientations were not estimated from the data). The NNLS outputs $\hat{\mathbf{w}} \in \mathbf{R}^{2N}$ served as training data samples.

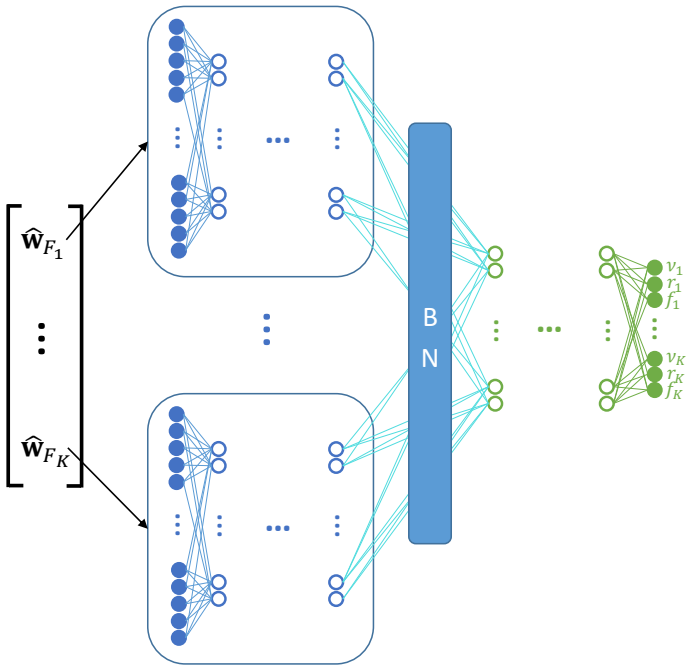

Figure 5.1: **Deep split-layer perceptron to learn microstructural properties from an NNLS feature vector.** The feature vector $\hat{\mathbf{w}}$ from the first-stage NNLS estimation is split into the K sub-vectors corresponding to each voxel fascicle, which are fed to independent multi-layer perceptrons. The outputs are recombined and rescaled in a 1-D batch normalization (BN) layer. A global multi-layer perceptron performs the final regression, outputting the volume fraction v, radius index r and density index f of each fascicle. Solid fills indicate input or output units.

The training targets associated with each sample were the groundtruth microstructural properties of each fascicle $[v_1, r_1, f_1, v_2, r_2, f_2]^T$ rescaled so that each entry had a range centered around zero.

Training strategy Network training was performed with the open-source PyTorch library[2] in Python 3. Stochastic gradient descent using minibatch size 200 was performed for 5 epochs with gradient steps of lengths commanded by the AdaGrad optimizer(McMahan and Streeter, 2010; Duchi et al., 2011) with initial learning rate 0.01. The loss function was the mean squared error (MSE) on the $3K = 6$ output targets. Based on performance of the network on an independent validation set of 3000 samples, the architecture of the network was fixed to $L_{\text{split}}, L_{\text{final}} = 2$ and $H_{\text{split}}, H_{\text{final}} = 100$. All training was performed on a standard Intel (Santa Clara, CA) Core i7 processor.

[2] https://pytorch.org/

5.2.3.2 Experiment 2.A. Estimation of microstructural properties on synthetic data

Generalization of the neural network to unseen synthetic data was tested on the Synthetic HCP dataset described above. The input to the network was the output of the first-stage NNLS estimation performed in Experiment 1.A. The errors between the 2-stage estimates and the groundtruth parameter values were computed and compared to the errors of the reference exhaustive dictionary matching.

5.2.3.3 Experiment 2.B. Estimation of microstructural properties on *in vivo* data

Generalization of the neural network to *in vivo* data, after training on synthetic data, was tested on the *in vivo* HCP dataset described above. The input to the network was the output of the first-stage NNLS estimation performed in Experiment 1.B. The estimates of apparent radius and density index were collected along all axons traversing the corpus callosum (CC), as described in Section 3.3.2.4. The network estimates were compared to the estimates produced by the reference exhaustive dictionary matching.

5.2.3.4 Experiment 2.C. Generalization to unseen experimental conditions

This experiment aimed at demonstrating the ability of the neural network to perform protocol transfer, i.e. to generalize to data acquired with a different protocol than in training. It also assesses the relevance of the NNLS weights as feature vectors, verifying that the NNLS weights are related to the microstructural configurations in the dictionary rather than to the specific DW-MRI protocol used to generate the dictionary.

The Synthetic Rodent test set described above was used for that purpose. It was fed to the first-stage NNLS estimation, which produced an output feature vector given to the neural network pre-trained independently on the HCP-like training set data. The errors between the 2-stage estimates and the groundtruth parameter values were computed and compared to the errors of the reference exhaustive dictionary matching.

5.3 Results

5.3.1 First stage: blind NNLS estimation

5.3.1.1 Experiment 1.A. Sparsity of the NNLS output on synthetic data

Figure 5.2 shows the mean number of non-zero weights detected by the first-stage NNLS estimation for fixed $v_1 = 0.3, v_2 = 0.7$, crossing angle $\alpha = 30^\circ$ and no angular mismatch between the orientations of the groundtruth fascicles and the rotated dictionaries supplied to the optimizer. Varying the reference crossing angle had virtually no effect. Increasing the systematic angular error shifted the curves downward (not shown), similarly to the decrease

observed with increasing SNR. Over all 86 400 independent experiments, the N entries in $\hat{\mathbf{w}}$ corresponding to a given fascicle were never identically zero, i.e., no fascicle was ever missed. The NNLS solutions were very sparse with a mean number of non-zero weights in $\hat{\mathbf{w}}$ consistently below 5 (whereas this number could have theoretically been any value up to $N = 782$). The number of positive weights was always greater for the fascicle with larger groundtruth volume fraction.

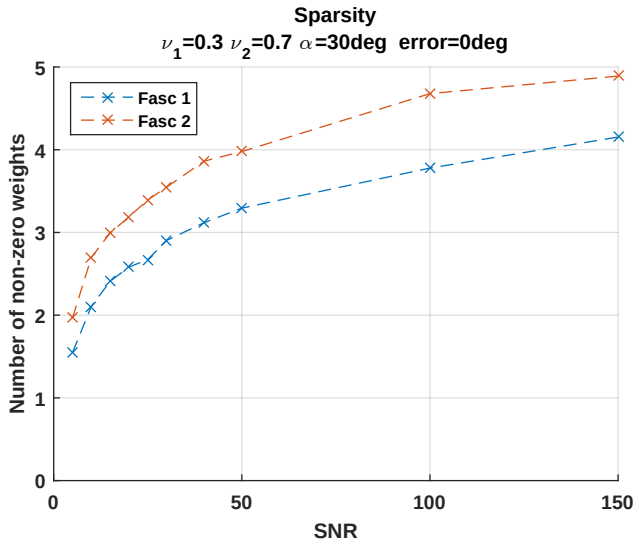

Figure 5.2: **A small number of fingerprints are sufficient to optimally reconstruct noisy synthetic DW-MRI signals.** Number of non-zero weights found by NNLS estimation vs signal-to-noise ratio (SNR). Each data point is an average over all 32 combinations of groundtruth values for r and f and 10 noise repetitions.

5.3.1.2 Experiment 1.B. Sparsity of the NNLS output on *in vivo* data

Similarly to the synthetic results of Experiment 1.A, the results on the *in vivo* HCP data shown in Figure 5.3 suggest that the DW-MRI signal is naturally sparse in the space of fingerprints. A fascicle initially detected by the ball-and-stick estimation was only missed in one of the 217 373 crossing-fascicle voxels. The number of non-zero NNLS weights never exceeded 8.

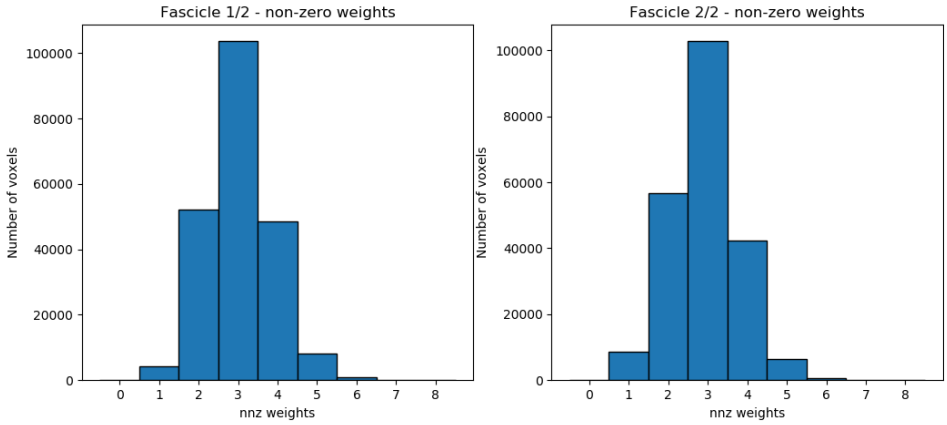

Figure 5.3: **A small number of fingerprints are sufficient to optimally reconstruct *in vivo* DW-MRI signals.** Number of non-zero weights found in each single-fascicle sub-dictionary by the NNLS estimation. Data for voxels from the HCP-MGH subject selected in Section 3.3.2.4 in which fascicle crossings were detected by the ball-and-stick estimation (in Step 3 of the estimation procedure, see Section 3.2.4). Fascicle 1 is the fascicle which was assigned the largest weight during the ball-and-stick estimation.

5.3.1.3 Experiment 1.C. Complexity of first-stage NNLS vs dictionary size

As suggested in Figure 5.4, the computational complexity of the NNLS estimation appeared to be linear in the number of entries N in a single-fascicle dictionary. At $N = 782$ (the size used in all other experiments), the average runtime was 11.1 ms and 7.4 ms per voxel respectively, depending on whether random dummy dictionaries were used or not. Note that this execution time was also affected linearly by the number of measurements M in the protocol ($M = 781$ in this particular experiment).

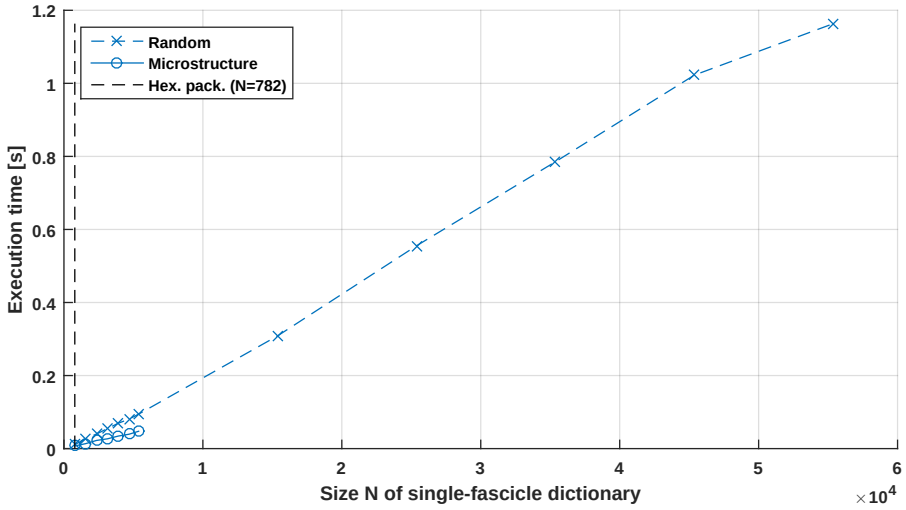

Figure 5.4: **NNLS complexity scales linearly with dictionary size N in crossing-fascicle settings.** Mean NNLS complexity appears consistently $\mathcal{O}(N)$ through single-fascicle dictionary sizes up to $N \approx 50\,000$. The black dashed vertical line indicates the size of the hexagonal-packing single-fascicle dictionary used in Chapter 3 on the rat spinal cord and the HCP datasets.

5.3.2 Second stage: neural network regression

5.3.2.1 Experiment 2.A. Estimation of microstructural properties on synthetic data

Figure 5.5 shows the mean absolute error (MAE) in estimated microstructural features and Figure 5.6 the distribution of the estimates at SNR=25, for a fixed crossing angle $\alpha = 30°$. The results for the other crossing angles were similar. Exhaustive dictionary matching only outperformed the fast 2-stage procedure for the estimation of the apparent radius index r. This was due to an overestimation of the small radius index values by the 2-stage method in fascicles with low groundtruth volume fraction (see top left of Figure 5.6). The difference between the two methods diminished in fascicles with larger groundtruth volume fraction. Remarkably, the 2-stage estimation outperformed the reference dictionary matching in the estimation of the apparent density index f and volume fraction v_1 in most scenarios with SNR ≤ 100. In that noise regime, the MAE on f was lower for the 2-stage procedure in 475/486 cases; in 239/243 cases for the MAE on v_1. Training took about 18 minutes on a standard Intel Core i7 processor.

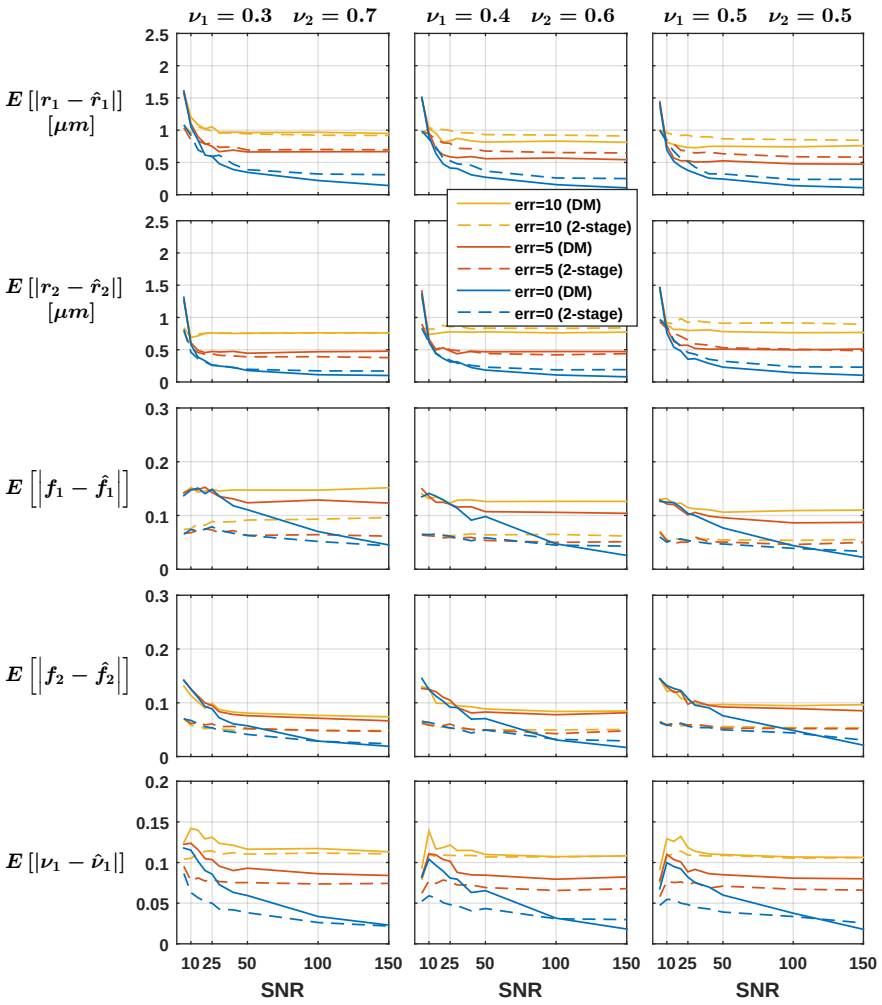

Figure 5.5: **Our fast two-stage estimation provides accurate microstructural estimates on unseen synthetic data.** Mean absolute error on each fasicle's radius index r_1 and r_2, density index f_1 and f_2 and on the physical volume fraction occupied by the first fascicle ν_1. The groundtruth volume fraction of the second fascicle decreases from left to right and the groundtruth crossing angle was fixed to 30°. Exhaustive dictionary matching (DM) is compared to our accelerated 2-stage procedure.

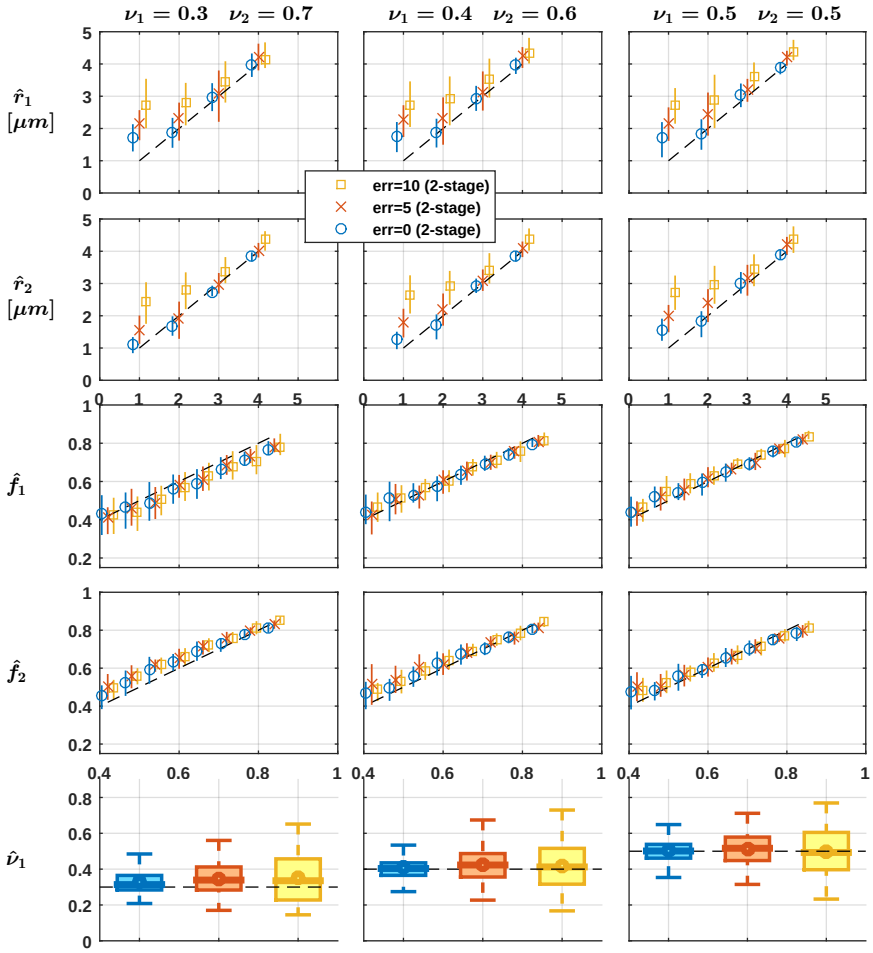

Figure 5.6: **Our fast two-stage estimation provides accurate microstructural estimates on unseen synthetic data.** Snapshot at SNR=25 of the mean (markers) and interquartile range (vertical lines) for the neural network estimates of each fasicle's radius index r_1 and r_2 and density index f_1 and f_2. In the box-plots for the estimates of v_1, the physical volume fraction occupied by the first fascicle, the whiskers indicate the minimum and maximum values in the data. The groundtruth volume fraction of the second fascicle decreases from left to right and the groundtruth crossing angle was fixed to 30°.

5.3.2.2 Experiment 2.B. Estimation of microstructural properties on *in vivo* data

As reported in Table 5.1, the discrepancies between the estimates of the exhaustive dictionary matching and those of the fast 2-stage procedure were larger on the *in vivo* dataset than they were on the synthetic data of Experiment 2.A. Although a systematic overestimation was observed, the method was able to detect the same trends in apparent axon radius and density indices, which has been discussed in Section 3.5.2.

Table 5.1: **Our fast 2-stage estimation detects the low-high-low trend in apparent radius and density index in CC streamlines.** Mean and standard deviation over all local fascicles or peaks assigned to streamlines passing through five sub-regions of the corpus callosum (CC), for the apparent axonal radius index r (in µm) and density index f. The exhaustive dictionary matching (DM) estimates are compared to those of the fast 2-stage procedure.

	r (DM)	r (2-stage)	f (DM)	f (2-stage)
anterior CC	4.06 ± 0.87	5.48 ± 0.83	0.599 ± 0.12	0.695 ± 0.10
mid-anterior CC	4.27 ± 0.73	5.56 ± 0.75	0.636 ± 0.12	0.731 ± 0.09
central CC	4.19 ± 0.90	5.48 ± 0.76	0.627 ± 0.12	0.712 ± 0.10
mid-posterior CC	4.22 ± 0.93	5.35 ± 0.83	0.645 ± 0.13	0.719 ± 0.10
posterior CC	3.87 ± 0.99	5.21 ± 0.94	0.581 ± 0.14	0.649 ± 0.11

5.3.2.3 Experiment 2.C. Generalization to unseen experimental conditions

The second-stage neural network exhibited good protocol transfer qualities. Similarly to Experiment 2.A on the HCP protocol, an overestimation of small apparent radius index values was observed, especially for fascicles with lower groundtruth volume fraction (see top left of Figure 5.7). With SNR ≤ 100 however, the 2-stage procedure outperformed the reference dictionary matching, exhibiting smaller MAEs for f in 461/486 cases and for v_1 in 227/243 cases, without having ever seen any data related to the Rodent protocol.

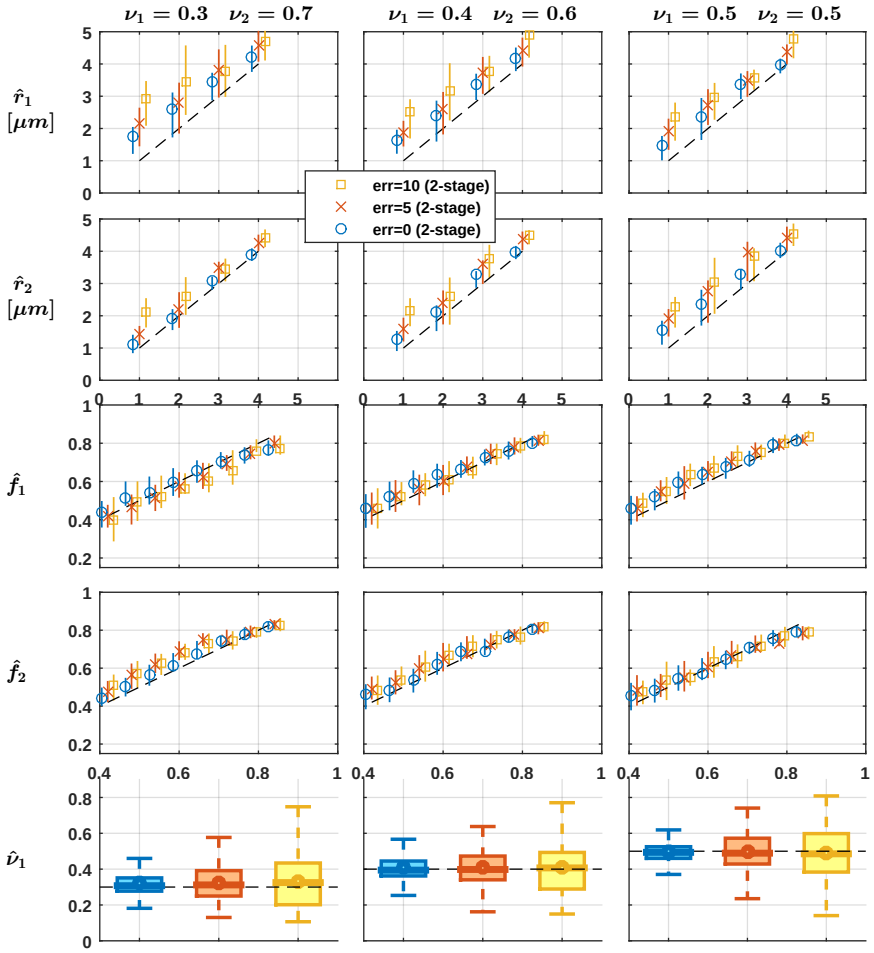

Figure 5.7: **Our neural network generalizes to DW-MRI data from a different protocol.** Results of our pre-trained neural network on synthetic data simulating a different scanner with a different protocol. Snapshot at SNR=25 of the mean (markers) and interquartile range (vertical lines) for the neural network estimates of each fasicle's radius index r_1 and r_2 and density index f_1 and f_2. In the box-plots for the estimates of v_1, the physical volume fraction occupied by the first fascicle, the whiskers indicate the minimum and maximum values in the data. The groundtruth volume fraction of the second fascicle decreases from left to right and the groundtruth crossing angle was fixed to 30°.

5.4 Discussion

5.4.1 Efficiency

Since the time taken for a forward pass of the second-stage neural network is extremely short, the computational complexity of the proposed 2-stage method is dominated by the first-stage NNLS. Our implementation was shown to be very efficient owing to the natural sparsity of the

data in the space of fingerprint weights (Experiments 1.A and A.B). Looking at the dictionary matching step alone, our method provided a speed-up factor of about 145 (Table 5.2) at fixed single-dictionary size $N = 782$, for $K = 2$ crossing fascicles. As suggested in Experiment 1.C, the linear complexity scaling of NNLS is likely to provide considerably larger speed-up factors as N further grows.

In our experiments, the rotation routine required in Step 4 of the estimation actually became the bottleneck. Taking rotations into account decreased our speed-up factor to about 26, as indicated in Table 5.2. However, the complexity of rotating a dictionary is simply proportional to the number of fingerprints in the dictionary; the speed-up factors will therefore keep increasing as N increases. An end-to-end deep learning solution where a mapping is learned from the DW-MRI signal directly to the microstructural parameters could be considered to entirely bypass the rotation stage.

Table 5.2: **Our 2-stage estimation enables speed-ups even at moderate dictionary sizes.** Average execution time in a voxel of the *in vivo* HCP dataset containing two crossing fascicles for the exhaustive dictionary matching (DM) and the proposed 2-stage acceleration. The rotation routine is identical in the two methods. The estimation time of the 2-stage is the sum of the NNLS estimation and the neural network (NN) forward pass. Quantities expressed in seconds and measured on a standard laptop equipped with an Intel Core i7 processor.

	exh. DM	2-stage	Speed-up
Rotations	0.02638	0.02638	1
Estimation	0.79158	0.00534 (NNLS) + $9.79e-5$ (NN)	145.65
Rot. + Est.	0.81795	0.03181	25.71

5.4.2 Accuracy

The neural network trained on a synthetic dataset of HCP-like acquisitions proved to generalize well to unseen synthetic data (Experiment 2.A), including on data simulated using completely different protocol and experimental parameters (Experiment 2.C). Remarkably, the network even outperformed the reference dictionary matching in a majority of realistic settings. This suggests that the feature vectors stemming from the NNLS estimation contained enough information about their underlying microstructure. If our model were perfect, only one diffusion fingerprint per fascicle would be needed to reconstruct the signal. The sparsity observed in Experiments 1.A and 1.B was therefore a strong indication of the relevance of the selected single-fascicle model. More marked discrepancies were observed between the full dictionary matching and the fast 2-stage estimator in the HCP data, even though the general trends were conserved. Evaluations on larger datasets will be required to fully validate the accelerated method for large-scale population studies.

5.4.3 Limitations

Training on synthetic data Using synthetic data for training the second-stage neural network implies that the simulation of a dictionary of Monte Carlo signals cannot be avoided during the pre-computing stage. However, given the good generalization performance of the network, the granularity of the underlying single-fascicle parameters can likely be coarser and the size of the reference dictionary reduced.

Size of the neural network As N increases, the size of the network will be largely dominated by the first fully-connected layer of each split MLP. To circumvent this issue, a reduced number of input units could be considered, such as for instance the non-zero weights of the vector $\hat{\mathbf{w}}$ produced by the NNLS along with their associated microstructural parameters. The maximum number of input units could be determined from a sparsity analysis on a subset of pre-acquired data, as in Experiment 1.B. Recurrent neural networks adapted to variable input size such as Long-Short Term Memory (LSTM) units (Hochreiter and Schmidhuber, 1997) will also be considered.

Estimated orientations Given the design of the Synthetic HCP dataset and the way training was performed, the network was only confronted to angular errors of 0, 5 and 10 degrees during training. This may account for some of the differences observed on the *in vivo* HCP dataset compared with the reference dictionary matching.

5.5 Conclusion

As summarized in Figure 5.8, our accelerated fingerprint method exhibits $\mathcal{O}(KN)$ runtime complexity compared to the $\mathcal{O}\left(N^K\right)$ of the reference method while preserving similar accuracy in the estimated microstructural parameters. A reasonable level of interpretability was maintained by relying on meaningful feature vectors in the space of dictionary fingerprints.

Future work will consider new architectures for the neural network of the second stage and will investigate the benefits and drawbacks of a full, end-to-end deep learning solution bypassing the feature extraction stage. Training a separate network for each output microstructural parameter will be considered as well.

This work will ultimately enable us to perform large-scale population studies and advance our knowledge of neurological and psychiatric disorders.

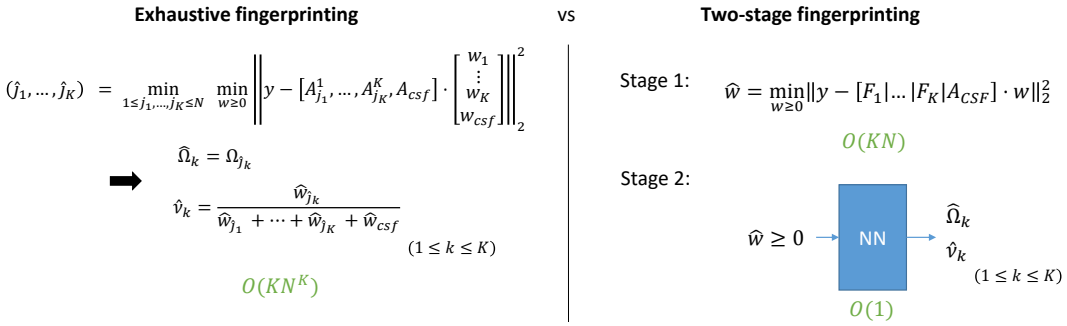

Figure 5.8: $\mathcal{O}(KN)$ **versus** $\mathcal{O}\left(N^K\right)$ **complexity enabled by our accelerated fingerprinting method.**

Conclusion

A general framework was proposed for the estimation of microstructural properties of white matter tissues based on Monte Carlo simulations, generally recognized as a reference standard in diffusion-weighted magnetic resonance imaging (DW-MRI).

The limitations of Monte Carlo simulations, i.e., their numerical nature and their computational cost, were circumvented by selecting a robust dictionary matching approach.

Compared to traditional approaches based on closed-form formulas for the DW-MRI signal, our approach was shown to provide more reliable and physically interpretable tissue parameters in a wide range of synthetic, *in vivo* and *ex vivo* experiments.

Steps have been taken to keep our framework tractable with increased model complexity and therefore dictionary sizes by verifying the superposition approximation for crossing fascicles and by proposing an accelerated method based on deep neural networks.

Limitations Important microstructural properties such as membrane permeability and axonal undulation (or dispersion) have been ignored in our experiments. However, the advantage of our framework is that these features can be readily incorporated. Recently-developed Monte Carlo simulators now seem to offer the possibility of realistically simulating such features (Palombo et al., 2019; Ginsburger et al., 2019a; Callaghan et al., 2019), paving the way for exciting new developments.

Similarly, glial cells have not been incorporated into our models yet. They could either be incorporated into the Monte Carlo simulations and part of the pre-simulated fingerprints or they could be modeled separately by a simpler compartment such as a closed sphere. Columns containing the diffusion signature of such spheres could be appended to the pre-simulated dictionary.

Due to a lack of available additional data, this work only considered the simple PGSE sequence with its known limitations on the sensitivity to certain microstructural features. It is likely that more advanced diffusion sequences such as those presented in Chapter 1 will be needed in the future for the finer study of the brain and spinal cord white matter. Combining DW-MRI with other MRI modalities such as multi-echo T2 and myelin water imaging in general will

also be considered as future developments.

A systematic and reliable method for the selection of the meta-parameters for Monte Carlo simulations (e.g., number of spins or number of time steps) remains an open issue in the field. It will need to be addressed to convince more users to incorporate Monte Carlo in their models and benefit from their great physical accuracy.

Perspectives The method presented in this thesis was designed from the start to be used for large-scale population studies in normal development and pathology. Clinical neurology and psychiatry would largely benefit from novel microstructural markers to relate symptoms to brain neuroanatomy. Microstructure imaging constantly requires validation to assess the specificity of its markers. Reliable validation techniques such as advanced histology and physical phantoms will continue to be explored methodically as soon as they become available. Large datasets of high-quality data such as in the HCP study will keep growing and will become an invaluable source of investigations and developments. Hardware improvement in MRI scanners and new diffusion-weighted sequences are also promising and will help DW-MRI become the modality of choice for non-invasive brain histology.

A Validating the superposition approximation for crossing fascicles

A.1 Impact of the number of cylinder layers in the groundtruth configuration

We considered a slightly refined interweaving pattern for the groundtruth crossing-fascicle configuration where an n_1-cylinder thick layer of Population 1 is followed by an n_2-cylinder thick layer of Population 2 instead of the one-cylinder-thick layers described in Section 2.2.1, as depicted in Figure 2.1. Below we show that the 1-cylinder-layer configuration that we considered in our study actually provides a "worst-case" groundtruth scenario for the superposition approximation and that increasing the thickness of each layer in the groundtruth configuration is apt to improve the behavior of the superposition approximation.

For simplicity, we considered DW-MRI signals of configurations featuring fascicles with identical microstructural properties, i.e. we fixed $r_1 = r_2 = r$, $f_1 = f_2 = f$, $n_1 = n_2 = n$ and $\nu_1 = \nu_2 = 0.5$. We compared the signal S_{appr} arising from the superposition of two identical fascicles of fixed radius index $r = 1.0\mu$m and density index $f = 0.5$ crossing at an angle $\alpha = 67.5°$ to a series of signals $S_{\text{intw}}(n)$ arising from reference interwoven-fascicle environments with varying number n of cylinders per layer but with matching parameters $r_1 = r_2 = r$, $f_1 = f_2 = f$, $\nu_1 = \nu_2 = 0.5$ and same crossing angle α. We selected the worst-case scenario for the superposition approximation by finding the value of n maximizing the root-mean-square (RMS) signal metric

$$\left\| S_{\text{appr}}(\mathscr{P}) - S_{\text{intw}}(n;\mathscr{P}) \right\|_{\text{RMS}} = \frac{1}{\sqrt{M}} \left(\sum_{i=1}^{M} \left(S_{\text{appr}}\left(\mathbf{p}_i\right) - S_{\text{intw}}(n;\mathbf{p}_i) \right)^2 \right)^{1/2},$$

over all M measurements of each tested protocol $\mathscr{P}$, namely protocols A and B.

Figure A.2 shows that the difference in normalized DW-MRI signal attenuation is of the order of 10^{-3} and is largest when the number of cylinders per population layer n is smallest. This confirms the intuition that as the number of cylinders n in a population's layer grows, one can expect the interwoven-fascicle configuration to be more similar to the superposition approxi-

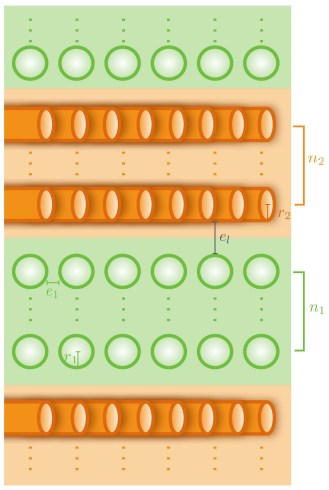

Figure A.1: **Refined synthetic phantom of interwoven fascicles of axons.** The total environment is made up of a succession of n_1 cylinders of radius r_1 and inter-cylinder spacing e_1 from Population 1 followed by n_2 cylinders of radius r_2 and inter-cylinder spacing e_2 from Population 2 (compare with the single-cylinder population layers of Figure 2.1). The packing within each layer is hexagonal and e_l is the spacing between layers. The green and orange shaded areas show how the water molecules of the interstitium are distributed among the populations, allowing us to unequivocally define the fractions of occupancy v_1 and v_2 as well as the population-specific cylinder packing densities f_1 and f_2.

mation since the extracellular trajectories of most water molecules will be fully contained in a single fascicle. We therefore set the layer thickness to $n = 1$ all the experiments presented in Chapter 2.

A.2 Impact of the Approximation on the DW-MRI Signal

This section provides additional data related to the computed signal differences between the interwoven-fascicle signal and the approximate signal when the reference parameters and the parameters used in the approximation match (see Section 2.3.1). Table A.1 provides the root-mean-square (RMS) differences obtained for each microstructural configuration and each HARDI acquisition shell described in Table 2.1 in Section 2.2.3. Figure A.3 illustrates the DW-MRI signals as a function of the gradient's orientation for one particular scenario where the relative signal discrepancies seemed largest for gradients parallel to the bisector of the two fascicles' orientations.

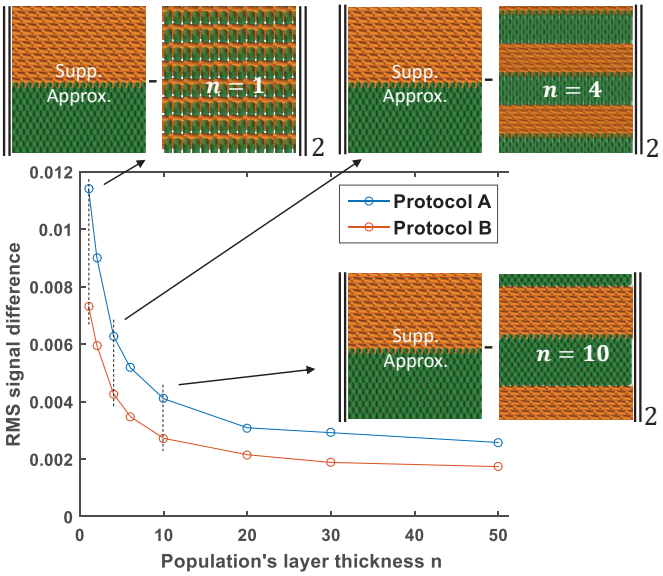

Figure A.2: **Reference interwoven fascicles with one-cylinder thick layers least resemble superposed fascicles.** The superposition of independent fascicles is schematically represented by the three images on the left-hand side of each inset even though, strictly speaking, the interface between the two fascicles appearing on these images is *not* modeled in the superposition approximation. Fixed parameters are $r = 1.0\,\mu\text{m}$, $f = 0.5$ and $\alpha = 67.5°$.

Table A.1: **Shell-per-shell RMS differences between interwoven-fascicle and approximate DW-MRI signals.** The RMS differences are computed over all the gradient directions contained in each of the HARDI shells comprised in Protocols A and B (see Table 2.1 in Section 2.2.3) for 24 crossing-fascicle configurations characterized by a radius index r, density index f and crossing angle α, as specified in the left-hand-side columns. Units of b are in s mm^{-2}. The boldfaced values correspond to the detailed illustrations in Figures 2.3 and A.3. The data suggest that the superposition approximation led to small DW-MRI signal differences compared to current clinical noise levels.

r [µm]	f	α [deg]	A-s1 (b=496)	A-s2 (b=682)	A-s3 (b=2453)	A-s4 (b=2635)	B-s1 (b=711)	B-s2 (b=1000)	B-s3 (b=2000)	B-s4 (b=2855)
0.5	0.5	22.5	**2.4e-02**	2.2e-02	9.6e-03	8.6e-03	2.2e-02	2.1e-02	1.1e-02	5.6e-03
0.5	0.5	45	2.2e-02	2.1e-02	9.0e-03	8.0e-03	2.2e-02	2.1e-02	1.1e-02	5.7e-03
0.5	0.5	67.5	2.1e-02	2.0e-02	7.4e-03	6.6e-03	2.0e-02	1.9e-02	1.0e-02	5.1e-03
0.5	0.5	90	2.1e-02	2.0e-02	7.8e-03	6.8e-03	2.0e-02	1.9e-02	1.0e-02	5.1e-03
1	0.5	22.5	1.1e-02	9.1e-03	3.3e-03	2.7e-03	9.0e-03	8.3e-03	3.7e-03	2.1e-03
1	0.5	45	1.1e-02	7.7e-03	2.9e-03	**2.6e-03**	8.7e-03	8.2e-03	4.3e-03	2.4e-03
1	0.5	67.5	1.1e-02	8.2e-03	3.3e-03	3.0e-03	8.8e-03	8.0e-03	3.6e-03	2.1e-03
1	0.5	90	1.1e-02	8.9e-03	3.5e-03	3.1e-03	7.7e-03	7.2e-03	3.8e-03	2.1e-03
2	0.5	22.5	3.2e-03	3.5e-03	1.9e-03	1.9e-03	4.1e-03	3.7e-03	1.9e-03	1.7e-03
2	0.5	45	3.8e-03	4.6e-03	2.5e-03	2.4e-03	4.0e-03	4.0e-03	2.7e-03	2.1e-03
2	0.5	67.5	2.8e-03	4.4e-03	2.1e-03	2.0e-03	4.2e-03	4.4e-03	3.1e-03	2.2e-03
2	0.5	90	3.7e-03	4.2e-03	2.7e-03	2.6e-03	4.0e-03	4.2e-03	3.4e-03	2.5e-03
3	0.5	22.5	4.1e-03	4.5e-03	2.7e-03	2.4e-03	4.4e-03	4.5e-03	3.1e-03	2.1e-03
3	0.5	45	3.7e-03	5.5e-03	2.4e-03	2.1e-03	4.5e-03	4.8e-03	3.8e-03	3.0e-03
3	0.5	67.5	3.5e-03	3.9e-03	2.7e-03	2.3e-03	4.4e-03	4.6e-03	3.1e-03	2.2e-03
3	0.5	90	3.4e-03	4.4e-03	3.5e-03	3.2e-03	3.7e-03	4.0e-03	3.5e-03	2.8e-03
4	0.5	22.5	3.8e-03	5.3e-03	2.8e-03	2.4e-03	4.5e-03	5.0e-03	5.0e-03	4.5e-03
4	0.5	45	4.0e-03	4.8e-03	3.3e-03	3.1e-03	4.7e-03	5.0e-03	4.1e-03	3.6e-03
4	0.5	67.5	4.4e-03	4.8e-03	2.6e-03	2.4e-03	4.7e-03	5.0e-03	4.5e-03	3.9e-03
4	0.5	90	3.8e-03	5.2e-03	2.0e-03	1.7e-03	3.9e-03	4.4e-03	4.5e-03	4.0e-03
5	0.5	22.5	3.5e-03	6.1e-03	3.8e-03	3.5e-03	4.5e-03	5.5e-03	7.7e-03	8.6e-03
5	0.5	45	3.3e-03	6.5e-03	3.8e-03	3.5e-03	5.4e-03	6.5e-03	7.3e-03	7.5e-03
5	0.5	67.5	4.0e-03	5.8e-03	3.4e-03	3.0e-03	5.1e-03	6.0e-03	6.9e-03	7.3e-03
5	0.5	90	3.9e-03	6.4e-03	4.1e-03	3.7e-03	5.7e-03	6.7e-03	7.7e-03	8.0e-03

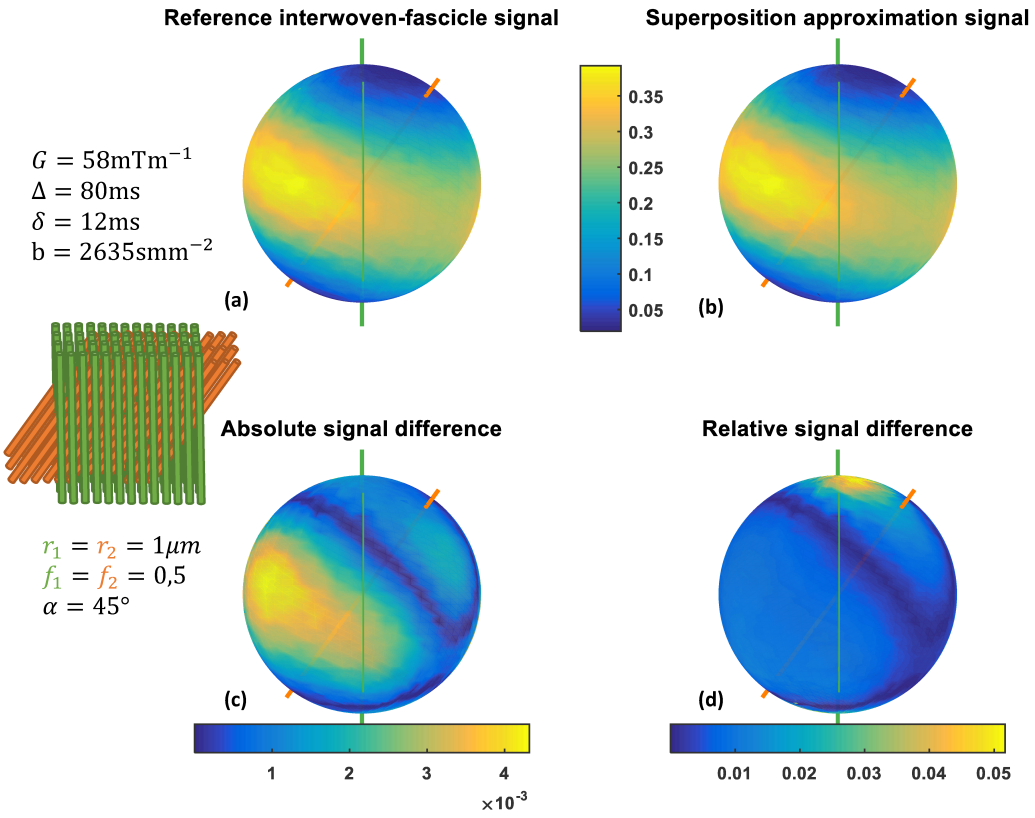

Figure A.3: **DW-MRI signals from the superposition approximation closely match signals from interwoven fascicles.** Normalized DW-MRI signal attenuation ((a)-(b)) and differences ((c)-(d)) for the HARDI shell of Protocol A with highest b-value plotted as a function of the gradient direction $\hat{\mathbf{g}}$, which yielded an RMS difference of 2.59×10^{-3}. Here the absolute signal differences (c) are highest around the direction normal to the plane defined by the fascicles' orientations $\mathbf{u}_1$ and $\mathbf{u}_2$, where the DW-MRI signals are largest. However the relative differences (d) are largest around the bisector of $\mathbf{u}_1$ and $\mathbf{u}_2$, where the signals are smallest.

A.3 Impact of PGSE parameters on the validity of the superposition approximation

This section investigates the impact of PGSE acquisition parameters on the validity of the approximation, by computing signal discrepancies between the reference interwoven-fascicle signal and the approximate signal for a wider range of parameters than considered in Protocols A and B (Table 2.1).

A.3.1 Methods

HARDI shells with a number of gradient directions uniformly distributed on the 3D sphere fixed to 90 were considered. The gradient intensity G, gradient separation Δ and gradient duration δ were selected by trying out every combination of candidate values $G = [40, 60, 70, 80, 120, 140]$ mT m^{-1},

$\Delta = [15, 20, 38, 66, 80, 100]$ ms and $\delta = [2, 15, 30]$ ms, subject to the following constraints:

- limited echo time (TE) to limit T2 decay by ensuring $\Delta + \delta \leq 180$ ms;
- realistic diffusion weighting by ensuring $200 \leq b \leq 20\,000$, with the b-value defined by Eq. (1.12), with units of [s mm^{-2}].

The root-mean-square (RMS) signal difference metric defined in Eq. (2.2.4) was computed for the same 24 microstructural configurations as in Section 2.2.4 for all the HARDI shells satisfying the above constraints.

A.3.2 Results

A total of 50 HARDI shells met the imposed constraints with b-values in the range $[209, 19\,253]$ s mm^{-2}. Figure A.4 suggests that RMS differences did not vary significantly between the 8 shells used in our validation of the superposition approximation (left of the dashed vertical line) and the 50 additional shells (right of the dashed vertical line) spanning a broader range of acquisition parameters. The mean RMS over all microstructural configurations and shells was 6.1×10^{-3} for the 8 shells of Protocols A and B and 4.9×10^{-3} for the other 50 shells. A correlation of -0.67 was found between the RMS averaged over the 24 microstructural configurations and the b-value, across all 58 shells, mainly due to the fact that signal values were lower at stronger b-values.

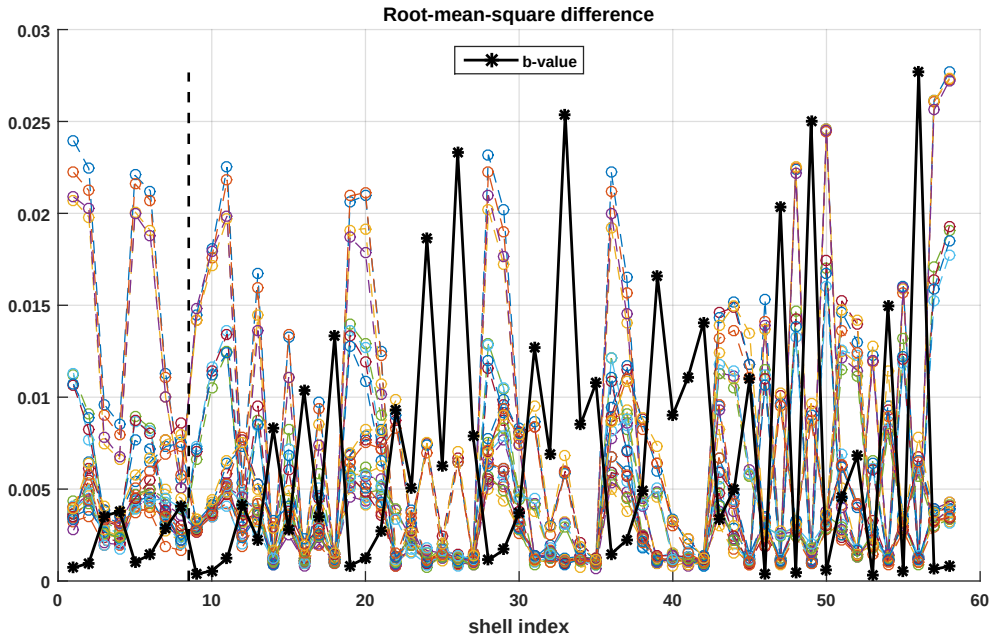

Figure A.4: **DW-MRI signals from the superposition approximation closely match signals from interwoven fascicles across PGSE parameters.** Root-mean-square difference between the reference and approximate signal for each HARDI shell. Each colored line represents one of the 24 microstructural configurations considered in our experiment. The black vertical dashed line separates the 8 shells of Protocols A and B (left, following the order in Table 2.1) from the 50 new shells considered in this section (right, no particular order). The thick black line indicates the b-value of each shell, scaled by an arbitrary constant for visualization purposes.

A.4 Results of the microstructural estimation experiments

This section complements the results presented in Section 2.3.2.

A.4.1 Experiment I: impact of the approximation on identical fascicles

Figures A.5, A.6 and A.7 present the results of Experiment I corresponding to reference crossing angles $\alpha = 22.5°, 45°$ and $90°$ respectively, which were qualitatively similar to the case $\alpha = 67.5°$ discussed in Section 2.3.2.1.

Figure A.8 displays the results of estimation experiments similar to those realized in Experiment I but considering the larger reference radius indices $r = [6,7,8,9]\mu m$, at fixed density index $f = 0.5$ and crossing angle $\alpha = 67.5°$. The estimation of the crossing angle exhibited trends similar to those obtained with smaller reference radii except that asymptotic, noise-free errors started to appear for $r = [7,8,9]\mu m$ with Protocol B. The DW-MRI signal arising from large pores is indeed more isotropic and less directional, which might have negatively affected the estimation of the crossing angle, especially for Protocol B which contains half the number of gradient directions present in Protocol A. The estimation of the density index and the radius index generally benefited from larger reference radii, with MAE respectively upper bounded by 0.06 and 1.25µm for SNR levels 20 and over.

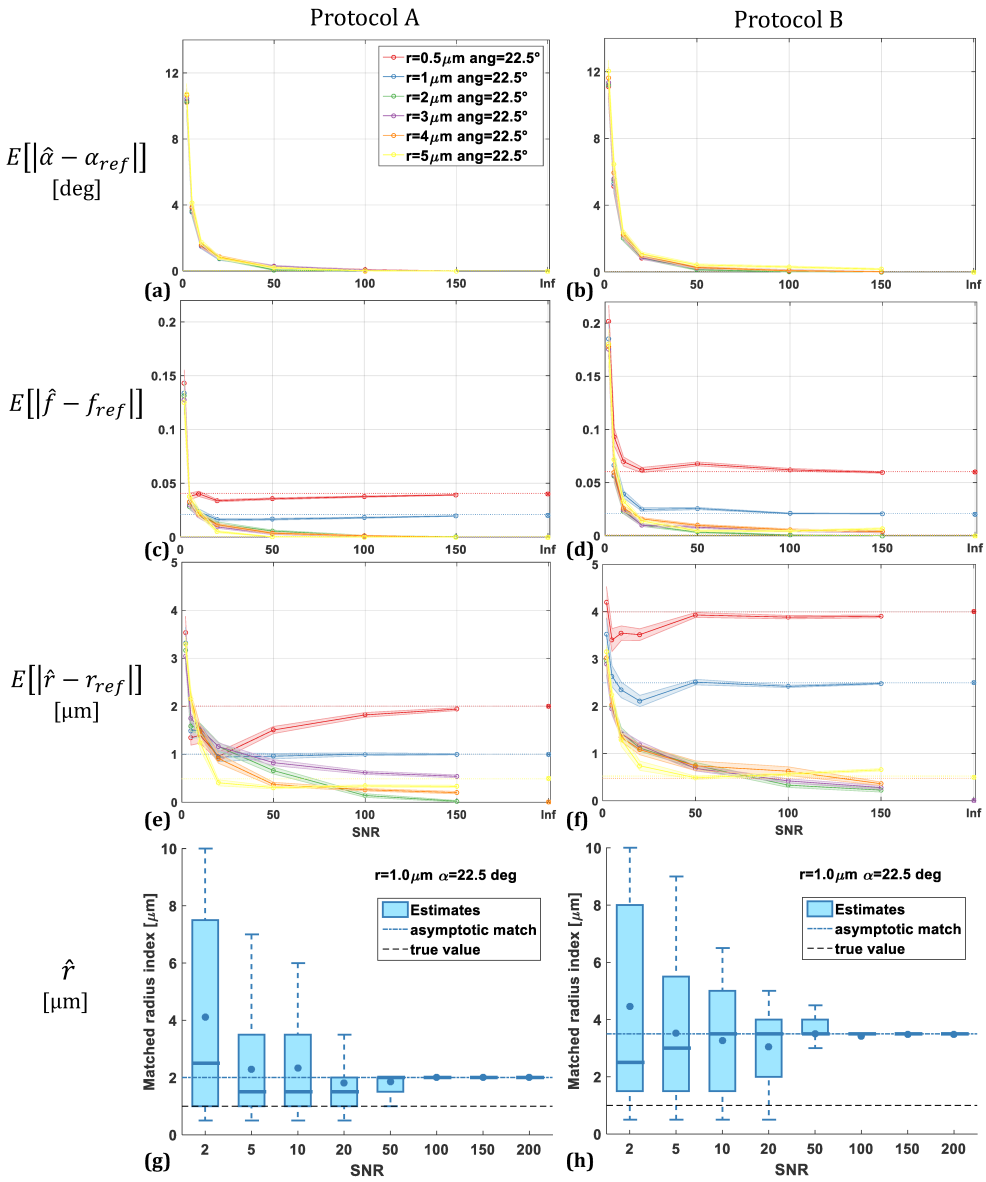

Figure A.5: **The superposition approximation shows fast convergence with increasing SNR and small asymptotic errors for the estimation of the crossing angle and the density index but a slower convergence with larger asymptotic errors for the radius index.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. The dashed lines are the asymptotic, noise-free errors slightly offset around their true values for visualization purposes. Sub-figures (g) and (h) display box-plots of the radius index estimates corresponding to the blue curve in sub-figures (e)-(f), where large interquartile ranges relative to the asymptotic bias up to SNR=50 suggest that the acquisition noise is the predominant source of error rather than the use of the superposition approximation.

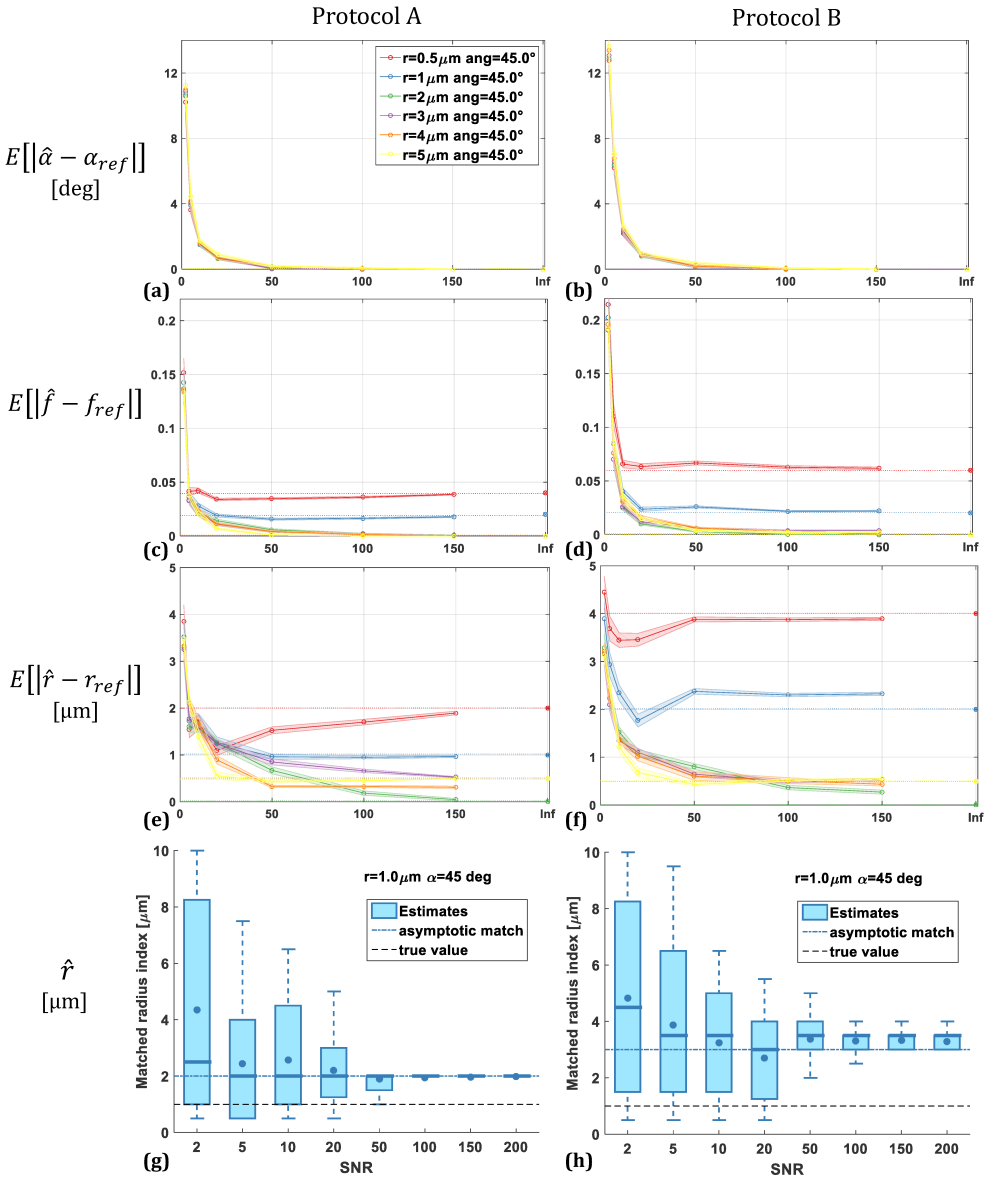

Figure A.6: **The superposition approximation shows fast convergence with increasing SNR and small asymptotic errors for the estimation of the crossing angle and the density index but a slower convergence with larger asymptotic errors for the radius index.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. The dashed lines are the asymptotic, noise-free errors slightly offset around their true values for visualization purposes. Sub-figures (g) and (h) display box-plots of the radius index estimates corresponding to the blue curve in sub-figures (e)-(f), where large interquartile ranges relative to the asymptotic bias up to SNR=50 suggest that the acquisition noise is the predominant source of error rather than the use of the superposition approximation.

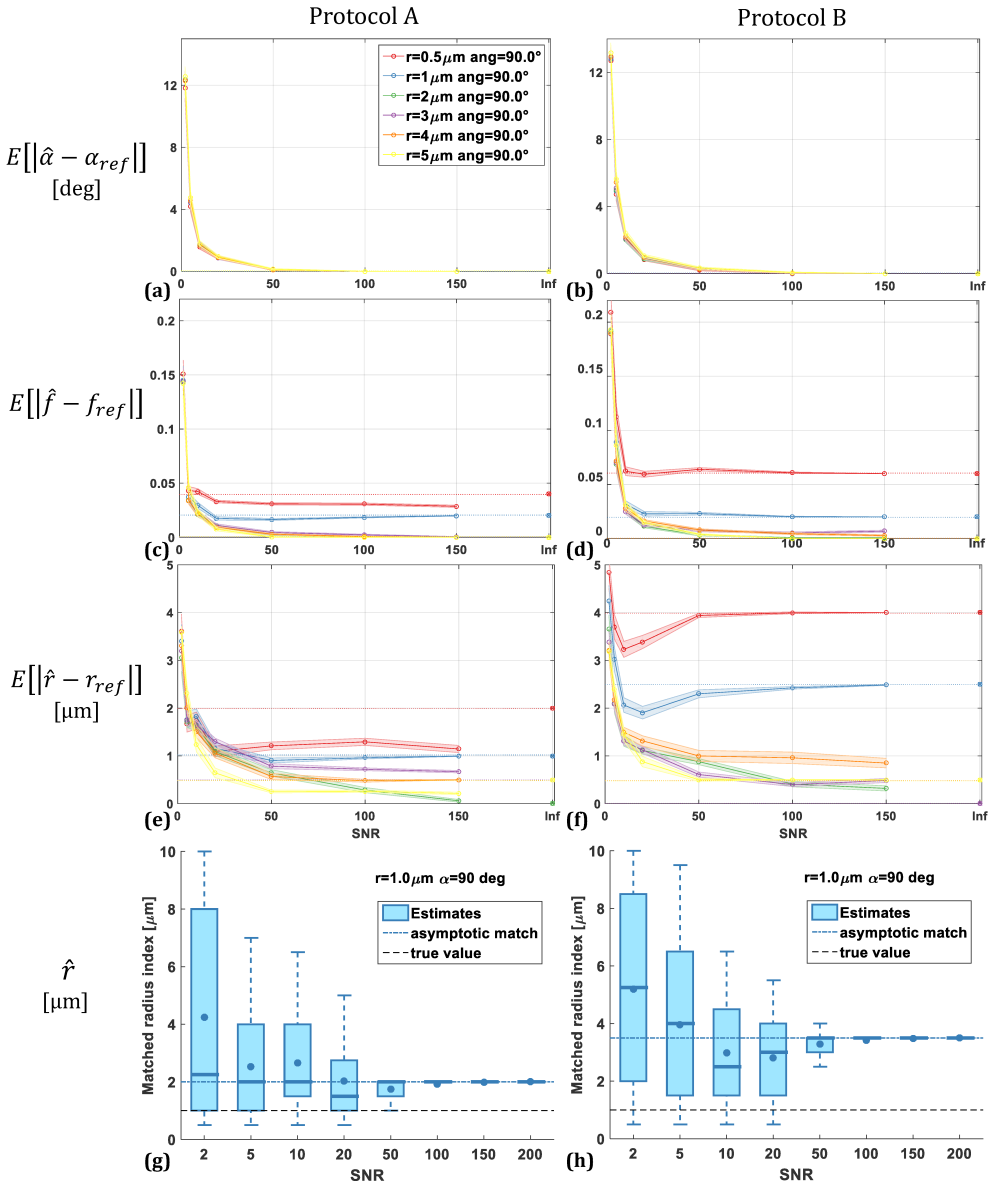

Figure A.7: **The superposition approximation shows fast convergence with increasing SNR and small asymptotic errors for the estimation of the crossing angle and the density index but a slower convergence with larger asymptotic errors for the radius index.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. The dashed lines are the asymptotic, noise-free errors slightly offset around their true values for visualization purposes. Sub-figures (g) and (h) display box-plots of the radius index estimates corresponding to the blue curve in sub-figures (e)-(f), where large interquartile ranges relative to the asymptotic bias up to SNR=50 suggest that the acquisition noise is the predominant source of error rather than the use of the superposition approximation.

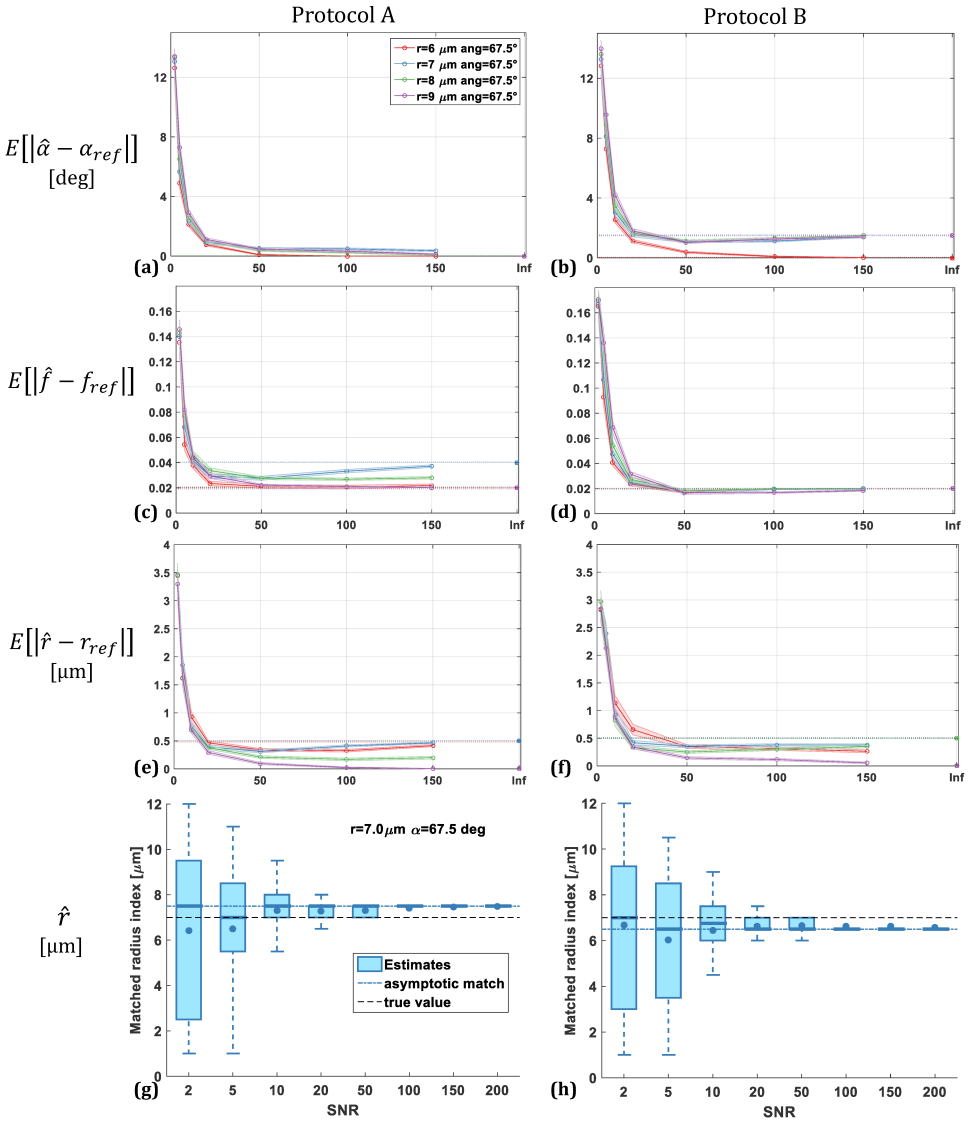

Figure A.8: **The superposition approximation shows fast convergence with increasing SNR and small asymptotic errors for the estimation of the crossing angle, the density index and the radius index at larger reference radius indices.** Mean absolute error (continuous lines) and standard errors (shaded areas) obtained with Protocol A (left) and B (right) on (a)-(b) the crossing angle α, (c)-(d) the density index f, (e)-(f) the radius index r. The dashed lines are the asymptotic, noise-free errors slightly offset around their true values for visualization purposes. Sub-figures (g) and (h) display box-plots of the radius index estimates corresponding to the blue curve in sub-figures (e)-(f), where large interquartile ranges relative to the asymptotic bias up to SNR=50 suggest that the acquisition noise is the predominant source of error rather than the use of the superposition approximation.

A.4.2 Experiment IV: impact of the approximation on a closed-form model of the microstructure

We studied the impact of the superposition approximation on the estimates of the NODDI model (Zhang et al., 2012), which is a simpler microstructural model essentially devised to describe a single population of axons and which does not explicitely incorporate a dependence on the axonal radius. In NODDI, a non-tissue compartment occupies a fraction v_{iso} of the volume of the voxel alongside a tissue compartment comprising an extracelluar space as well as neurites modeled as sticks and occupying a fraction v_{in} of the tissue compartment (Zhang et al., 2012). The orientations of the neurites follow a Watson distribution characterized by an orientation dispersion index odi and a principal direction **u**. A diffusion tensor is used to model the signal originating from the tissue space outside the neurites.

Specifically, the NODDI closed-form, continuous expression $S_{\text{NDI}}(\Omega_{\text{NDI}};\mathscr{P})$ relating the DW-MRI signal S_{NDI} to the microstructural parameters Ω_{NDI} for a given acquisition protocol $\mathscr{P}$ was fitted to signals $S_{\text{intw}}(\Omega_{\text{cros}};\mathscr{P})$ arising from reference interwoven configurations Ω_{cros} and to the signals $S_{\text{appr}}(\Omega_{\text{cros}};\mathscr{P})$ arising from the superposition approximations of matching microstructural parameters. We obtained the model estimates $\hat{\Omega}_{\text{NDI}}\left(S_{\text{appr}}\right)$ and $\hat{\Omega}_{\text{NDI}}(S_{\text{intw}})$ by solving, for each of the 24 crossing-fascicle configurations Ω_{cros} described in the Methods section, considering Protocols A and B,

$$\hat{\Omega}_{\text{NDI}}\left(S_{\text{appr}}\right) = \underset{\Omega_{\text{NDI}}}{\operatorname{argmin}}\, d\left(S_{\text{appr}}(\Omega_{\text{cros}};\mathscr{P});S_{\text{NDI}}(\Omega_{\text{NDI}};\mathscr{P})\right),$$
$$\hat{\Omega}_{\text{NDI}}(S_{\text{intw}}) = \underset{\Omega_{\text{NDI}}}{\operatorname{argmin}}\, d\left(S_{\text{intw}}(\Omega_{\text{cros}};\mathscr{P});S_{\text{NDI}}(\Omega_{\text{NDI}};\mathscr{P})\right),$$

where d is the negative log-likelihood of the Rician distribution and where the non-linear minimization was achieved using the 2-stage fitting routine provided in the NODDI toolbox (http://mig.cs.ucl.ac.uk) for MATLAB R2015a (MathWorks, Natick, MA, U.S.A.). The impact of the signal discrepancies caused by the approximation was assessed by comparing $\hat{\Omega}_{\text{NDI}}\left(S_{\text{appr}}\right)$ to $\hat{\Omega}_{\text{NDI}}(S_{\text{intw}})$.

The results of the fittings are shown in Supporting Figure A.9 and in general suggested little dependence on whether the model was estimated from the reference of the approximate signals. The maximum difference across the two protocols and all 24 microstructural configurations was 0.062 for v_{iso}, 0.026 for odi and 0.023 for v_{in}. As could be expected, the main fascicle direction **u** (not shown in Supporting Figure A.9) was close to the bisector of the two fascicles' orientations for $\alpha = [22.5°, 45°, 67.5°]$ and was aligned with either one of the two orientations when the reference crossing angle was 90°.

The only notable differences could be observed at $r = 0.5\,\mu\text{m}$ where v_{iso} and v_{in} estimated from the reference interwoven-fascicle signals were slightly lower than the estimates obtained from

the superposition approximation, which could be attributed to larger signal discrepancies at small radii. However increasing v_{iso} tends to *decrease* the total signal whereas increasing v_{in} generally *increases* the signal. The simultaneous over- or underestimation of these two parameters could therefore also be related to an inherent mathematical equivalence in the model leading to two plausible solutions between which the non-linear fitting routine was unable to discriminate.

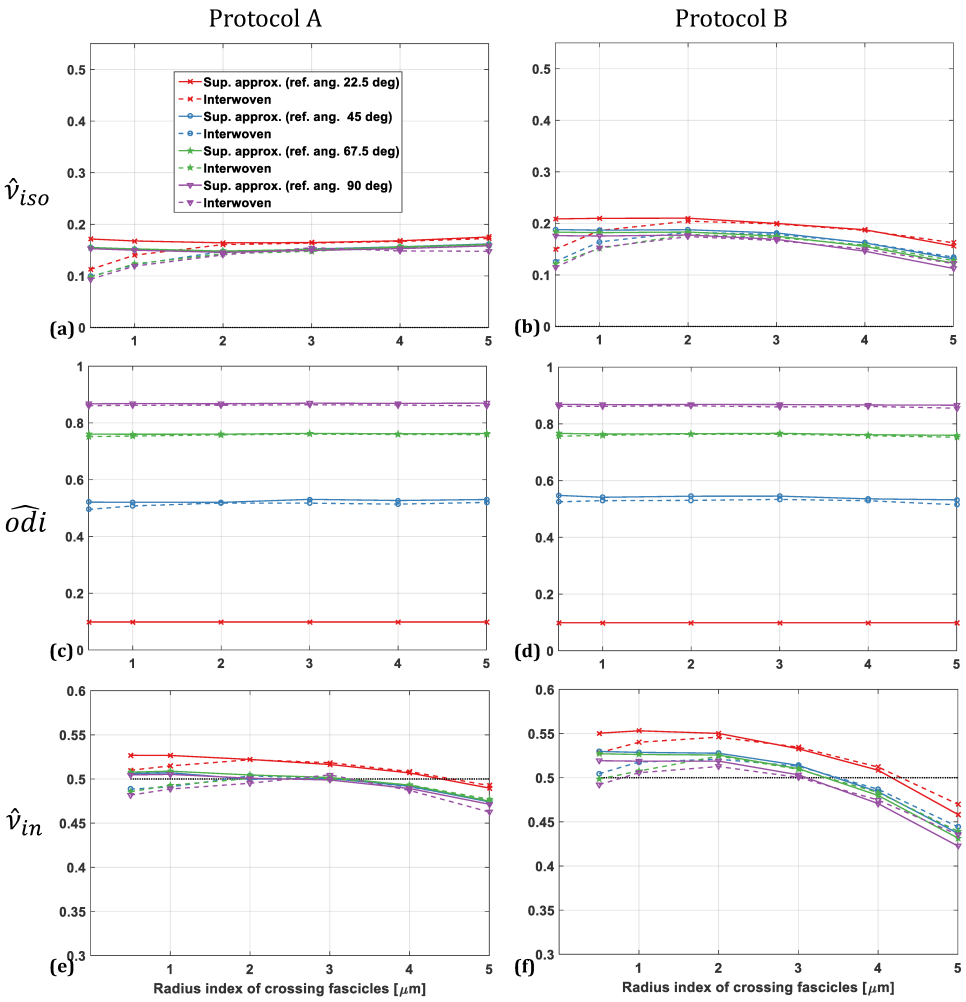

Figure A.9: **The signal discrepancies of the superposition approximation have a limited impact on NODDI parameters.** Parameters obtained when fitting the model to the approximate (continuous lines) and the reference interwoven-fascicle signal (dashed lines). (a)-(b) Isotropic or non-tissue volume fraction v_{iso}. (c)-(d) Orientation dispersion index (odi). (e)-(f) Intracellular or intra-neurite volume fraction v_{in}, with the horizontal line representing the axonal density index $f_{\text{ref}} = 0.5$ of the crossing-fascicle configurations.

B Microstructure Fingerprinting

B.1 Choice of the reference diffusivity for data augmentation

This appendix discusses the relationship between the number N of reference Monte Carlo simulations required to simulate configurations with different diffusivities and the fixed diffusivity $D_{\text{sim}} \in [D_{\min}, D_{\max}]$ used during those simulations, as described in Section 3.2.3.2. Specifically, we show that N is independent of D_{sim} and only depends on the desired resolution δ_L in the length scale parameter L, assuming a uniform sampling of L.

Let δ_L^{sim} be the step size separating consecutive values of L for which reference simulations using a diffusivity D_{sim} were performed. Given $D \neq D_{\text{sim}}$, using Eq. (3.12) yields a sampling $\left\{\sqrt{\frac{D}{D_{\text{sim}}}} L_{\text{sim},i}\right\}_{i=1}^{N}$ of the parameter L. The corresponding step size therefore becomes $\delta_L(D) = \sqrt{\frac{D}{D_{\text{sim}}}}\delta_L^{\text{sim}}$, which admits the upper bound $\delta_L(D) \leq \sqrt{\frac{D_{\max}}{D_{\text{sim}}}}\delta_L^{\text{sim}}$ for all $D \in [D_{\min}, D_{\max}]$.

Let $\bar{\delta}_L$ denote the largest acceptable step size in L for all values of D in $[D_{\min}, D_{\max}]$ (or equivalently, the coarsest acceptable granularity of a dictionary). To ensure that $\delta_L(D) \leq \bar{\delta}_L \quad \forall D$, it is necessary to guarantee that

$$\sqrt{\frac{D_{\max}}{D_{\text{sim}}}}\delta_L^{\text{sim}} \leq \bar{\delta}_L \Leftrightarrow \delta_L^{\text{sim}} \leq \sqrt{\frac{D_{\text{sim}}}{D_{\max}}}\bar{\delta}_L. \tag{B.1}$$

Recalling that the sampling of L should cover the interval $\left[\sqrt{\frac{D_{\text{sim}}}{D_{\max}}} L_{\min}, \sqrt{\frac{D_{\text{sim}}}{D_{\min}}} L_{\max}\right]$ during the reference simulations using D_{sim}, the lower bound on the total number of simulations N

becomes

$$\begin{aligned} N-1 &= \frac{\sqrt{\frac{D_{\text{sim}}}{D_{\text{min}}}}L_{\text{max}} - \sqrt{\frac{D_{\text{sim}}}{D_{\text{max}}}}L_{\text{min}}}{\delta_L^{\text{sim}}} \\ &\geq \frac{\sqrt{\frac{D_{\text{sim}}}{D_{\text{min}}}}L_{\text{max}} - \sqrt{\frac{D_{\text{sim}}}{D_{\text{max}}}}L_{\text{min}}}{\sqrt{\frac{D_{\text{sim}}}{D_{\text{max}}}}\bar{\delta}_L} = \frac{L_{\text{max}}/\sqrt{D_{\text{min}}} - L_{\text{min}}/\sqrt{D_{\text{max}}}}{\bar{\delta}_L/\sqrt{D_{\text{max}}}}, \end{aligned} \tag{B.2}$$

where inequality (B.1) was used. Equation (B.2) shows that the minimum number of simulations does not depend on D_{sim} and is entirely determined by the biophysical constraints of the problem.

B.2 Effect of non-uniform voxels on the fitting

The sparsity constraints in Eq. (3.14) assume that a single fingerprint can explain the signal of a fascicle of axons. With the hexagonal-packing model selected at the single-fascicle level for our experiments, all the fingerprints of our dictionary arise from voxels containing uniform arrangements of axons. In practice however, a voxel may contain sub-regions exhibiting different microstructural properties. The experiment described below examined how the fitted parameters behaved when the groundtruth voxels consisted of axons with spatially heterogeneous packing.

Experiment In this experiment, groundtruth voxels were split into two parts with different axonal packing density f. The DW-MRI signal S_{mixed} for such a non-uniform configuration was obtained as

$$S_{\text{mixed}}\left(\delta f_{\text{low}}, \delta f_{\text{high}}\right) = \nu_{\text{low}} S_{\text{unif}}\left(r^*, f^* - \delta f_{\text{low}}\right) + \nu_{\text{high}} S_{\text{unif}}\left(r^*, f^* + \delta f_{\text{high}}\right),$$

where S_{unif} is the fingerprint of a uniform configuration and where $\nu_{\text{low}} = \delta f_{\text{high}} / \left(\delta f_{\text{low}} + \delta f_{\text{high}}\right)$ and $\nu_{\text{high}} = \delta f_{\text{low}} / \left(\delta f_{\text{low}} + \delta f_{\text{high}}\right)$ ensure that the average packing density of the non-uniform voxel is f^*.

First, non-uniform groundtruth voxels with symmetric halves were considered by setting $\delta f_{\text{low}} = \delta f_{\text{high}} \coloneqq \delta f$ (leading to $\nu_{\text{low}} = \nu_{\text{high}} = 0.5$) and letting δf vary from 0.03 to 0.18 in 6 equal increments. Second, asymmetric halves were studied by setting $\delta f_{\text{high}} = 0.06$ and letting δf_{low} vary from 0.03 to 0.18 in 6 equal increments. In both cases, the fitting procedure of Section 3.2.4 was performed setting $w_{\text{csf}} = 0$ for 3 values of reference radius index $r^* = \{0.8, 1.6, 3.0\}\,\mu\text{m}$ and 5 values of average density index f^* equally spaced between 0.36 and 0.72, provided that $f^* - \delta f_{\text{low}} > 0.20$ and $f^* + \delta f_{\text{high}} < 0.90$.

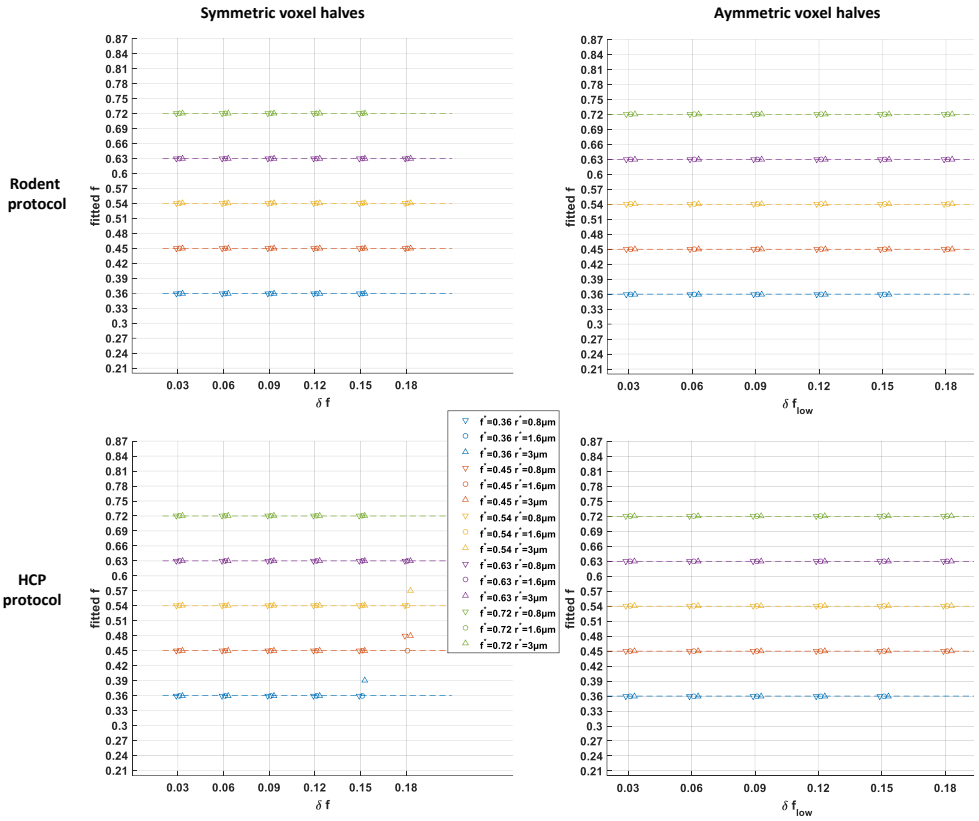

Figure B.1: **Fingerprints assuming uniform microstructural configurations capture the average axonal packing density of non-uniform configurations.** Fitted axonal density index $\hat{f}$ for groundtruth voxels containing two sub-regions of different axonal packing density, in a symmetric or asymmetric configuration, using the rodent protocol (Section 3.3.1.1) or the HCP protocol (Section 3.3.2.1). The missing data points correspond to scenarios in which one of the local density indices fell outside the range $]0.20, 0.90[$.

Results The results are displayed in Figure B.1. In all 168 cases with the rodent protocol, the fitted axonal density index $\hat{f}$ precisely matched the average density index f^* of the non-uniform voxels. This implied that $\left|\hat{f} - f^*\right| \leq 0.03$ given the granularity in f of the single-fascicle dictionary used for the estimation (see Section 3.3.1.2). With the HCP protocol, the bound reached $\left|\hat{f} - f^*\right| \leq 0.06$ in only 4 out of 168 cases, systematically for the largest density spread δf. Root-mean-square differences between the normalized groundtruth DW-MRI signals and their best fits were of the order of 1×10^{-4}.

Even though our Monte Carlo fingerprints arise solely from voxels with a uniform packing density, our estimation procedure systematically explained the signal by selecting a packing density corresponding to precisely the average density of the non-uniform voxel.

Bibliography

Aboitiz, F., Scheibel, A. B., Fisher, R. S., and Zaidel, E. (1992). Fiber composition of the human corpus callosum. *Brain research*, 598(1-2):143–153.

Adluru, G., Gur, Y., Anderson, J. S., Richards, L. G., Adluru, N., and DiBella, E. V. (2014). Assessment of white matter microstructure in stroke patients using noddi. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 742–745. IEEE.

Aja-Fernández, S. and Tristán-Vega, A. (2012). Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction. *Magnetic Resonance in Medicine*, 67(2):580–585.

Al-Chalabi, A. and Miller, C. C. (2003). Neurofilaments and neurological disease. *Bioessays*, 25(4):346–355.

Alexander, D. C., Dyrby, T. B., Nilsson, M., and Zhang, H. (2019). Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR in Biomedicine*, 32(4):e3841.

Alexander, D. C., Hubbard, P. L., Hall, M. G., Moore, E. A., Ptito, M., Parker, G. J., and Dyrby, T. B. (2010). Orientationally invariant indices of axon diameter and density from diffusion MRI. *Neuroimage*, 52(4):1374–1389.

Alonso-Ortiz, E., Levesque, I. R., and Pike, G. B. (2015). MRI-based myelin water imaging: A technical review. *Magnetic resonance in medicine*, 73(1):70–81.

Aranda, R., Ramirez-Manzanares, A., and Rivera, M. (2015). Sparse and Adaptive Diffusion Dictionary (SADD) for recovering intra-voxel white matter structure. *Medical image analysis*, 26(1):243–255.

Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., and Basser, P. J. (2008). Axcaliber: a method for measuring axon diameter distribution from diffusion mri. *Magnetic resonance in medicine*, 59(6):1347–1354.

Assaf, Y., Freidlin, R. Z., Rohde, G. K., and Basser, P. J. (2004). New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5):965–978.

Auría, A., Daducci, A., Thiran, J.-P., and Wiaux, Y. (2015a). Structured sparsity for spatially coherent fibre orientation estimation in diffusion MRI. *NeuroImage*, 115:245–255.

Auría, A., Romascano, D., Canales-Rodriguen, E., Wiaux, Y., Dirby, T., Alexander, D., Thiran, J.-P., and Daducci, A. (2015b). Accelerated microstructure imaging via convex optimisation for regions with multiple fibres (AMICOx). In *Image Processing (ICIP), 2015 IEEE International Conference on*, pages 1673–1676. IEEE.

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

Axer, H., Berks, G., and Keyserlingk, D. G. V. (2000). Visualization of nerve fiber orientation in gross histological sections of the human brain. *Microscopy research and technique*, 51(5):481–492.

Balls, G. T. and Frank, L. R. (2009). A simulation environment for diffusion weighted MR experiments in complex media. *Magnetic resonance in medicine*, 62(3):771–778.

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267.

Benjamini, D., Komlosh, M. E., Holtzclaw, L. A., Nevo, U., and Basser, P. J. (2016). White matter microstructure from nonparametric axon diameter distribution mapping. *NeuroImage*, 135:333–344.

Berman, S., Filo, S., and Mezer, A. A. (2019). Modeling conduction delays in the corpus callosum using MRI-measured g-ratio. *Neuroimage*, 195:128–139.

Bonilla, I. and Snyder, R. E. (2007). Transverse relaxation in rat optic nerve. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo*, 20(2):113–120.

Budde, M. D. and Frank, J. A. (2010). Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke. *Proceedings of the National Academy of Sciences*, 107(32):14472–14477.

Burcaw, L. M., Fieremans, E., and Novikov, D. S. (2015). Mesoscopic structure of neuronal tracts from time-dependent diffusion. *NeuroImage*, 114:18–37.

Burke, R. E. and O'malley, K. (2013). Axon degeneration in Parkinson's disease. *Experimental neurology*, 246:72–83.

Callaghan, P. T. (1991). *Principles of nuclear magnetic resonance microscopy*, volume 3. Clarendon Press Oxford.

Callaghan, P. T. (1995). Pulsed-gradient spin-echo NMR for planar, cylindrical, and spherical pores under conditions of wall relaxation. *Journal of magnetic resonance, Series A*, 113(1):53–59.

Callaghan, P. T. and Manz, B. (1994). Velocity exchange spectroscopy. *Journal of Magnetic Resonance, Series A*, 106(2):260–265.

Callaghan, R., Alexander, D. C., Zhang, H., and Palombo, M. (2019). Contextual fibre growth to generate realistic axonal packing for diffusion mri simulation. In *International Conference on Information Processing in Medical Imaging*, pages 429–440. Springer.

Canales-Rodríguez, E. J., Daducci, A., Sotiropoulos, S. N., Caruyer, E., Aja-Fernández, S., Radua, J., Mendizabal, J. M. Y., Iturria-Medina, Y., Melie-García, L., Alemán-Gómez, Y., et al. (2015). Spherical deconvolution of multichannel diffusion MRI data with non-Gaussian noise models and spatial regularization. *PloS one*, 10(10):e0138910.

Canales-Rodríguez, E. J., Legarreta, J. H., Pizzolato, M., Rensonnet, G., Girard, G., Rafael-Patino, J., Barakovic, M., Romascano, D., Aleman-Gomez, Y., Radua, J., et al. (2019). Sparse wars: A survey and comparative study of spherical deconvolution algorithms for diffusion MRI. *NeuroImage*, 184:140–160.

Caruyer, E. and Deriche, R. (2012). Diffusion MRI signal reconstruction with continuity constraint and optimal regularization. *Medical image analysis*, 16(6):1113–1120.

Caruyer, E., Lenglet, C., Sapiro, G., and Deriche, R. (2013). Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic resonance in medicine*, 69(6):1534–1540.

Chen, N.-k. and Wyrwicz, A. M. (1999). Correction for EPI distortions using multi-echo gradient-echo imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 41(6):1206–1213.

Christiaens, D., Sunaert, S., Suetens, P., and Maes, F. (2017). Convexity-constrained and nonnegativity-constrained spherical factorization in diffusion-weighted imaging. *NeuroImage*, 146:507–517.

Clayden, J. D., Nagy, Z., Weiskopf, N., Alexander, D. C., and Clark, C. A. (2015). Microstructural parameter estimation in vivo using diffusion MRI and structured prior information. *Magnetic resonance in medicine*.

Cohen, O., Zhu, B., and Rosen, M. S. (2018). MR fingerprinting deep reconstruction network (DRONE). *Magnetic resonance in medicine*, 80(3):885–894.

Condon, B., Patterson, J., Jenkins, A., Wyper, D., Hadley, D., Grant, R., Rowan, J., and Teasdale, G. (1987). MR relaxation times of cerebrospinal fluid. *Journal of computer assisted tomography*, 11(2):203–207.

Cook, P., Bai, Y., Nedjati-Gilani, S., Seunarine, K., Hall, M., Parker, G., and Alexander, D. (2006). Camino: open-source diffusion-MRI reconstruction and processing. In *14th scientific meeting of the international society for magnetic resonance in medicine*, volume 2759. Seattle WA, USA.

Cory, D., Garroway, A., and Miller, J. (1990). Applications of spin transport as a probe of local geometry. In *Abstracts of papers of the American Chemical Society*, volume 199, pages 105–poly. American Chemical Society.

Crémillieux, Y., Ding, S., and Dunn, J. F. (1998). High-resolution in vivo measurements of transverse relaxation times in rats at 7 Tesla. *Magnetic resonance in medicine*, 39(2):285–290.

Daducci, A., Canales-Rodríguez, E. J., Zhang, H., Dyrby, T. B., Alexander, D. C., and Thiran, J.-P. (2015). Accelerated microstructure imaging via convex optimization (AMICO) from diffusion MRI data. *NeuroImage*, 105:32–44.

Dantzig, G. B., Orden, A., Wolfe, P., et al. (1955). The generalized simplex method for minimizing a linear form under linear inequality restraints. *Pacific Journal of Mathematics*, 5(2):183–195.

de Graaf, R. A., Brown, P. B., McIntyre, S., Nixon, T. W., Behar, K. L., and Rothman, D. L. (2006). High magnetic field water and metabolite proton T1 and T2 relaxation in rat brain in vivo. *Magnetic resonance in medicine*, 56(2):386–394.

DeLuca, G., Ebers, G., and Esiri, M. (2004). Axonal loss in multiple sclerosis: a pathological survey of the corticospinal and sensory tracts. *Brain*, 127(5):1009–1018.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. (2007). Regularized, fast, and robust analytical Q-ball imaging. *Magnetic resonance in medicine*, 58(3):497–510.

Deslauriers-Gauthier, S. and Deriche, R. (2019). Estimation of axonal conduction speed and the inter hemispheric transfer time using connectivity informed maximum entropy on the mean. In *Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 10953, page 109530E. International Society for Optics and Photonics.

Dhital, B., Kellner, E., Kiselev, V. G., and Reisert, M. (2017). The absence of restricted water pool in brain white matter. *NeuroImage*.

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

Dortch, R. D., Apker, G. A., Valentine, W. M., Lai, B., and Does, M. D. (2010). Compartment-specific enhancement of white matter and nerve ex vivo using chromium. *Magnetic resonance in medicine*, 64(3):688–697.

Drakesmith, M. and Jones, D. K. (2018). Mapping axon conduction delays in vivo from microstructural MRI. *bioRxiv*, page 503763.

Drobnjak, I., Zhang, H., Hall, M. G., and Alexander, D. C. (2011). The matrix formalism for generalised gradients with time-varying orientation in diffusion NMR. *Journal of Magnetic Resonance*, 210(1):151–157.

Drobnjak, I., Zhang, H., Ianuş, A., Kaden, E., and Alexander, D. C. (2016). PGSE, OGSE, and sensitivity to axon diameter in diffusion MRI: Insight from a simulation study. *Magnetic resonance in medicine*, 75(2):688–700.

Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159.

Duval, T., Perraud, B., Vuong, M.-T., Lopez Rios, N., Stikov, N., and Cohen-Adad, J. (2016). Validation of quantitative MRI metrics using full slice histology with automatic axon segmentation. In *Proceedings of the 24th Annual Meeting of ISMRM (Singapore)*.

Dyrby, T. B., Hall, M. G., Ptito, M., Alexander, D., et al. (2013). Contrast and stability of the axon diameter index from microstructure imaging with diffusion MRI. *Magnetic resonance in medicine*, 70(3):711–721.

Evangelou, N., Konz, D., Esiri, M., Smith, S., Palace, J., and Matthews, P. (2001). Size-selective neuronal changes in the anterior optic pathways suggest a differential susceptibility to injury in multiple sclerosis. *Brain*, 124(9):1813–1820.

Farooq, H., Xu, J., Nam, J. W., Keefe, D. F., Yacoub, E., Georgiou, T., and Lenglet, C. (2016). Microstructure imaging of crossing (MIX) white matter fibers from diffusion MRI. *Scientific reports*, 6:38927.

Fick, R. (2017). *Advanced dMRI signal modeling for tissue microstructure characterization*. PhD thesis, Université Côte d'Azur; Inria Sophia Antipolis.

Fick, R., Wassermann, D., Pizzolato, M., and Deriche, R. (2015). A unifying framework for spatial and temporal diffusion in diffusion MRI. In *International Conference on Information Processing in Medical Imaging*, pages 167–178. Springer.

Fick, R. H., Wassermann, D., Caruyer, E., and Deriche, R. (2016). MAPL: Tissue microstructure estimation using laplacian-regularized MAP-MRI and its application to HCP data. *NeuroImage*, 134:365–385.

Fieremans, E., Burcaw, L. M., Lee, H.-H., Lemberskiy, G., Veraart, J., and Novikov, D. S. (2016). In vivo observation and biophysical interpretation of time-dependent diffusion in human white matter. *NeuroImage*, 129:414–427.

Fieremans, E., De Deene, Y., Delputte, S., Özdemir, M. S., D'Asseler, Y., Vlassenbroeck, J., Deblaere, K., Achten, E., and Lemahieu, I. (2008). Simulation and experimental verification of the diffusion in an anisotropic fiber phantom. *Journal of Magnetic Resonance*, 190(2):189–199.

Fieremans, E., Jensen, J. H., and Helpern, J. A. (2011). White matter characterization with diffusional kurtosis imaging. *Neuroimage*, 58(1):177–188.

Fieremans, E. and Lee, H.-H. (2018). Physical and numerical phantoms for the validation of brain microstructural MRI: A cookbook. *Neuroimage*, 182:39–61.

Fieremans, E., Novikov, D. S., Jensen, J. H., and Helpern, J. A. (2010). Monte Carlo study of a two-compartment exchange model of diffusion. *NMR in Biomedicine*, 23(7):711–724.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355.

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., and Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion MRI data. *Frontiers in neuroinformatics*, 8:8.

George, R. and Griffin, J. W. (1994). The proximo-distal spread of axonal degeneration in the dorsal columns of the rat. *Journal of neurocytology*, 23(11):657–667.

Gibbons, E. K., Hodgson, K. K., Chaudhari, A. S., Richards, L. G., Majersik, J. J., Adluru, G., and DiBella, E. V. (2019). Simultaneous NODDI and GFA parameter map generation from subsampled q-space imaging using deep learning. *Magnetic resonance in medicine*, 81(4):2399–2411.

Gilani, N., Hildebrand, S., Schueth, A., and Roebroeck, A. (2019). Monte carlo simulation of diffusion MRI in geometries constructed from two-photon microscopy of cortical gray matter. *bioRxiv*, page 626945.

Ginsburger, K., Matuschke, F., Poupon, F., Mangin, J.-F., Axer, M., and Poupon, C. (2019a). Medusa: A gpu-based tool to create realistic phantoms of the brain microstructure using tiny spheres. *NeuroImage*, 193:10–24.

Ginsburger, K., Matuschke, F., Poupon, F., Mangin, J.-F., Markus, A., and Poupon, C. (2019b). Machine learning based estimation of axonal properties in the presence of beading. In *Proceedings of the 27th Annual Meeting of ISMRM (Montreal)*.

Ginsburger, K., Poupon, F., Beaujoin, J., Estournet, D., Matuschke, F., Mangin, J.-F., Axer, M., and Poupon, C. (2018). Improving the realism of white matter numerical phantoms: A step toward a better understanding of the influence of structural disorders in diffusion MRI. *Frontiers in Physics*, 6:12.

Girard, G., Daducci, A., Petit, L., Thiran, J.-p., Whittingstall, K., Deriche, R., Wassermann, D., and Descoteaux, M. (2017). AxTract: Toward microstructure informed tractography. *Human brain mapping*, 38(11):5485–5500.

Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256.

Golkov, V., Dosovitskiy, A., Sperl, J. I., Menzel, M. I., Czisch, M., Sämann, P., Brox, T., and Cremers, D. (2016). Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. *IEEE transactions on medical imaging*, 35(5):1344–1351.

Grebenkov, D. (2007). Residence times and other functionals of reflected Brownian motion. *Physical Review E*, 76(4):041139.

Grebenkov, D. S. (2008). Laplacian eigenfunctions in NMR. I. A numerical tool. *Concepts in Magnetic Resonance Part A*, 32(4):277–301.

Gross, B. and Kosfeld, R. (1969). Anwendung der spin-echo-methode der messung der selbstdiffusion. *Messtechnik*, 77:171–177.

Gudbjartsson, H. and Patz, S. (1995). The rician distribution of noisy MRI data. *Magnetic resonance in medicine*, 34(6):910–914.

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

Hall, M. G., Nedjati-Gilani, G., and Alexander, D. C. (2017). Realistic voxel sizes and reduced signal variation in Monte-Carlo simulation for diffusion MR data synthesis. *arXiv preprint arXiv:1701.03634*.

Harkins, K. D. and Does, M. D. (2016). Simulations on the influence of myelin water in diffusion-weighted imaging. *Physics in Medicine & Biology*, 61(13):4729.

Harms, R., Fritz, F., Tobisch, A., Goebel, R., and Roebroeck, A. (2017). Robust and fast nonlinear optimization of diffusion MRI microstructure models. *NeuroImage*.

Hill, I., Palombo, M., Santin, M., Branzoli, F., Philippe, A.-C., Wassermann, D., Aigrot, M.-S., Stankoff, B., Evercooren, A. B.-V., Felfi, M., et al. (2019). Machine learning based white matter models with permeability: An experimental study in cuprizone treated in-vivo mouse model of axonal demyelination. *arXiv preprint arXiv:1907.02324*.

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

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Horsfield, M. A. and Jones, D. K. (2002). Applications of diffusion-weighted and diffusion tensor MRI to white matter diseases–a review. *NMR in Biomedicine*, 15(7-8):570–577.

Hutchinson, E. B., Avram, A. V., Irfanoglu, M. O., Koay, C. G., Barnett, A. S., Komlosh, M. E., Özarslan, E., Schwerin, S. C., Juliano, S. L., and Pierpaoli, C. (2017). Analysis of the effects of noise, DWI sampling, and value of assumed parameters in diffusion MRI models. *Magnetic resonance in medicine*, 78(5):1767–1780.

Ianuş, A., Drobnjak, I., and Alexander, D. C. (2016). Model-based estimation of microscopic anisotropy using diffusion MRI: a simulation study. *NMR in Biomedicine*, 29(5):672–685.

Innocenti, G. M., Caminiti, R., Rouiller, E. M., Knott, G., Dyrby, T. B., Descoteaux, M., and Thiran, J.-P. (2018). Diversity of cortico-descending projections: histological and diffusion MRI characterization in the monkey. *Cerebral Cortex*, 29(2):788–801.

Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Jelescu, I. O. and Budde, M. D. (2017). Design and validation of diffusion mri models of white matter. *Frontiers in physics*, 5:61.

Jelescu, I. O., Veraart, J., Adisetiyo, V., Milla, S. S., Novikov, D. S., and Fieremans, E. (2015). One diffusion acquisition and different white matter models: how does microstructure change in human early development based on wmti and noddi? *Neuroimage*, 107:242–256.

Jelescu, I. O., Veraart, J., Fieremans, E., and Novikov, D. S. (2016). Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR in Biomedicine*, 29(1):33–47.

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

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

Joyashiki, E., Matsuya, Y., and Tohda, C. (2011). Sominone improves memory impairments and increases axonal density in alzheimer's disease model mice, 5XFAD. *International Journal of Neuroscience*, 121(4):181–190.

Kakkar, L. S., Bennett, O. F., Siow, B., Richardson, S., Ianuş, A., Quick, T., Atkinson, D., Phillips, J. B., and Drobnjak, I. (2017). Low frequency oscillating gradient spin-echo sequences improve sensitivity to axon diameter: An experimental study in viable nerve tissue. *NeuroImage*.

Kay, K. R., Smith, C., Wright, A. K., Serrano-Pozo, A., Pooler, A. M., Koffie, R., Bastin, M. E., Bak, T. H., Abrahams, S., Kopeikina, K. J., et al. (2013). Studying synapses in human brain with array tomography and electron microscopy. *Nature protocols*, 8(7):1366–1380.

Kellner, E., Dhital, B., Kiselev, V. G., and Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic resonance in medicine*, 76(5):1574–1581.

Kim, J. H., Loy, D. N., Liang, H.-F., Trinkaus, K., Schmidt, R. E., and Song, S.-K. (2007). Noninvasive diffusion tensor imaging of evolving white matter pathology in a mouse model of acute spinal cord injury. *Magnetic Resonance in Medicine*, 58(2):253–260.

Koay, C. G., Özarslan, E., and Basser, P. J. (2009). A signal transformational framework for breaking the noise floor and its applications in MRI. *Journal of Magnetic Resonance*, 197(2):108–119.

Kodiweera, C., Alexander, A. L., Harezlak, J., McAllister, T. W., and Wu, Y.-C. (2016). Age effects and sex differences in human brain white matter of young to middle-aged adults: a dti, noddi, and q-space study. *NeuroImage*, 128:180–192.

Kunz, N., da Silva, A. R., and Jelescu, I. O. (2018). Intra-and extra-axonal axial diffusivities in the white matter: Which one is faster? *Neuroimage*, 181:314–322.

Lam, W. W., Jbabdi, S., and Miller, K. L. (2015). A model for extra-axonal diffusion spectra with frequency-dependent restriction. *Magnetic resonance in medicine*, 73(6):2306–2320.

Lampinen, B., Szczepankiewicz, F., Mårtensson, J., van Westen, D., Sundgren, P. C., and Nilsson, M. (2017). Neurite density imaging versus imaging of microscopic anisotropy in diffusion MRI: a model comparison using spherical tensor encoding. *Neuroimage*, 147:517–531.

Larsson, E. G., Erdogmus, D., Yan, R., Principe, J. C., and Fitzsimmons, J. R. (2003). SNR-optimality of sum-of-squares reconstruction for phased-array magnetic resonance imaging. *Journal of Magnetic Resonance*, 163(1):121–123.

Lawson, C. L. and Hanson, R. J. (1995). *Solving least squares problems*, volume 15. Siam.

Lee, C.-Y., Bennett, K. M., and Debbins, J. P. (2013). Sensitivities of statistical distribution model and diffusion kurtosis model in varying microstructural environments: a monte carlo study. *Journal of Magnetic Resonance*, 230:19–26.

Liewald, D., Miller, R., Logothetis, N., Wagner, H.-J., and Schüz, A. (2014). Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque. *Biological cybernetics*, 108(5):541–557.

Lin, M., He, H., Tong, Q., Ding, Q., Yan, X., Feiweier, T., and Zhong, J. (2018). Effect of myelin water exchange on DTI-derived parameters in diffusion MRI: Elucidation of TE dependence. *Magnetic resonance in medicine*, 79(3):1650–1660.

Liu, M., Gross, D. W., Wheatley, B. M., Concha, L., and Beaulieu, C. (2013). The acute phase of Wallerian degeneration: longitudinal diffusion tensor imaging of the fornix following temporal lobe surgery. *Neuroimage*, 74:128–139.

Lovas, G., Szilágyi, N., Majtényi, K., Palkovits, M., and Komoly, S. (2000). Axonal changes in chronic demyelinated cervical spinal cord plaques. *Brain*, 123(2):308–317.

Maxwell, W. L., Bartlett, E., and Morgan, H. (2015). Wallerian degeneration in the optic nerve stretch-injury model of traumatic brain injury: a stereological analysis. *Journal of neurotrauma*, 32(11):780–790.

McCall, D. W., Douglass, D. C., and Anderson, E. W. (1963). Self-diffusion studies by means of nuclear magnetic resonance spin-echo techniques. *Berichte der Bunsengesellschaft für physikalische Chemie*, 67(3):336–340.

McHugh, D. J., Zhou, F., Cristinacce, P. L. H., Naish, J. H., and Parker, G. J. (2015). Ground truth for diffusion MRI in cancer: a model-based investigation of a novel tissue-mimetic material. In *International Conference on Information Processing in Medical Imaging*, pages 179–190. Springer.

McMahan, H. B. and Streeter, M. (2010). Adaptive bound optimization for online convex optimization. *arXiv preprint arXiv:1002.4908*.

Mercredi, M. and Martin, M. (2018). Toward faster inference of micron-scale axon diameters using Monte Carlo simulations. *Magnetic Resonance Materials in Physics, Biology and Medicine*, pages 1–20.

Merlet, S., Caruyer, E., Ghosh, A., and Deriche, R. (2013). A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features. *Medical Image Analysis*, 17(7):830–843.

Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N.-J., Dougherty, R. F., Perry, M. L., Parvizi, J., Hua, L. H., Butts-Pauly, K., et al. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nature medicine*, 19(12):1667.

Nashmi, R. and Fehlings, M. (2001). Changes in axonal physiology and morphology after chronic compressive injury of the rat thoracic spinal cord. *Neuroscience*, 104(1):235–251.

Nedjati-Gilani, G. L., Schneider, T., Hall, M. G., Cawley, N., Hill, I., Ciccarelli, O., Drobnjak, I., Wheeler-Kingshott, C. A. G., and Alexander, D. C. (2017). Machine learning based compartment models with permeability for white matter microstructure imaging. *NeuroImage*, 150:119–135.

Nedjati-Gilani, G. L., Schneider, T., Hall, M. G., Wheeler-Kingshott, C. A., and Alexander, D. C. (2014). Machine learning based compartment models with permeability for white matter microstructure imaging. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 257–264. Springer.

Nilsson, M., Alerstam, E., Wirestam, R., Sta, F., Brockstedt, S., Lätt, J., et al. (2010). Evaluating the accuracy and precision of a two-compartment kärger model using Monte Carlo simulations. *Journal of Magnetic Resonance*, 206(1):59–67.

Nilsson, M., Lasi**c**, S., Drobnjak, I., Topgaard, D., and Westin, C.-F. (2017). Resolution limit of cylinder diameter estimation by diffusion MRI: The impact of gradient waveform and orientation dispersion. *NMR in Biomedicine*, 30(7):e3711.

Nilsson, M., Lätt, J., Nordh, E., Wirestam, R., Ståhlberg, F., and Brockstedt, S. (2009). On the effects of a varied diffusion time in vivo: is the diffusion in white matter restricted? *Magnetic resonance imaging*, 27(2):176–187.

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

Nilsson, M., van Westen, D., Ståhlberg, F., Sundgren, P. C., and Lätt, J. (2013). The role of tissue microstructure and water exchange in biophysical modelling of diffusion in white matter. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 26(4):345–370.

Ning, L., Özarslan, E., Westin, C.-F., and Rathi, Y. (2017). Precise inference and characterization of structural organization (PICASO) of tissue from molecular diffusion. *NeuroImage*, 146:452–473.

Novikov, D. S., Kiselev, V. G., and Jespersen, S. N. (2018a). On modeling. *Magnetic resonance in medicine*, 79(6):3172–3193.

Novikov, D. S., Veraart, J., Jelescu, I. O., and Fieremans, E. (2018b). Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *NeuroImage*, 174:518–538.

Özarslan, E., Koay, C., and Basser, P. (2008). Simple harmonic oscillator based estimation and reconstruction for one-dimensional q-space MR. In *Proc. Intl. Soc. Mag. Reson. Med*, volume 16, page 35.

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

Palombo, M., Alexander, D. C., and Zhang, H. (2019). A generative model of realistic brain cells with application to numerical simulation of the diffusion-weighted mr signal. *NeuroImage*, 188:391–402.

Palombo, M., Ligneul, C., Hernandez-Garzon, E., and Valette, J. (2017). Can we detect the effect of spines and leaflets on the diffusion of brain intracellular metabolites? *NeuroImage*.

Palombo, M., Ligneul, C., Najac, C., Le Douce, J., Flament, J., Escartin, C., Hantraye, P., Brouillet, E., Bonvento, G., and Valette, J. (2016). New paradigm to assess brain cell morphology by diffusion-weighted MR spectroscopy in vivo. *Proceedings of the National Academy of Sciences*, 113(24):6671–6676.

Panagiotaki, E., Hall, M. G., Zhang, H., Siow, B., Lythgoe, M. F., and Alexander, D. C. (2010). High-fidelity meshes from tissue samples for diffusion MRI simulations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 404–411. Springer.

Payne, S. C., Bartlett, C. A., Harvey, A. R., Dunlop, S. A., and Fitzgerald, M. (2011). Chronic swelling and abnormal myelination during secondary degeneration after partial injury to a central nervous system tract. *Journal of neurotrauma*, 28(6):1077–1088.

Peled, S., Cory, D. G., Raymond, S. A., Kirschner, D. A., and Jolesz, F. A. (1999). Water diffusion, T2, and compartmentation in frog sciatic nerve. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 42(5):911–918.

Pirkl1, C. M., Lipp, I., Buonincontri, G., Molina-Romero1, M., Sekuboyina1, A., Waldmannstetter, D., Dannenberg, J., Tomassini, V., Tosetti, M., Jones, D. K., Menzel, M. I., Menze, B. H., and Gómez, P. A. (2019). Deep learning-enabled diffusion tensor MR fingerprinting. In *Proceedings of the 27th Annual Meeting of ISMRM (Montreal)*.

Pizzolato, M., Fick, R., Boutelier, T., and Deriche, R. (2016). Noise floor removal via phase correction of complex diffusion-weighted images: Influence on DTI and q-space metrics. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 21–34. Springer.

Pizzolato, M., Wassermann, D., Boutelier, T., and Deriche, R. (2015). Exploiting the phase in diffusion MRI for microstructure recovery: towards axonal tortuosity via asymmetric diffusion processes. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 109–116. Springer.

Pohmann, R., Shajan, G., and Balla, D. (2011). Contrast at high field: relaxation times, magnetization transfer and phase in the rat brain at 16.4 T. *Magnetic resonance in medicine*, 66(6):1572–1581.

Price, W. S. (2009). *NMR Studies of Translational Motion: Principles and Applications*. Cambridge Molecular Science. Cambridge University Press.

Raffelt, D., Tournier, J.-D., Rose, S., Ridgway, G. R., Henderson, R., Crozier, S., Salvado, O., and Connelly, A. (2012). Apparent fibre density: a novel measure for the analysis of diffusion-weighted magnetic resonance images. *Neuroimage*, 59(4):3976–3994.

Ramirez-Manzanares, A., Rivera, M., Vemuri, B. C., Carney, P., and Mareci, T. (2007). Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. *IEEE transactions on medical imaging*, 26(8):1091–1102.

Regan, D. G. and Kuchel, P. W. (2000). Mean residence time of molecules diffusing in a cell bounded by a semi-permeable membrane: Monte carlo simulations and an expression relating membrane transition probability to permeability. *European Biophysics Journal*, 29(3):221–227.

Reisert, M., Skibbe, H., and Kiselev, V. G. (2014). The diffusion dictionary in the human brain is short: Rotation invariant learning of basis functions. In *Computational Diffusion MRI and Brain Connectivity*, pages 47–55. Springer.

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

Rensonnet, G., Jacobs, D., Macq, B., and Taquet, M. (2016). Fast and accurate simulations of diffusion-weighted MRI signals for the evaluation of acquisition sequences. In *SPIE Medical Imaging*, pages 97843L–97843L. International Society for Optics and Photonics.

Rensonnet, G., Scherrer, B., Girard, G., Jankovski, A., Warfield, S. K., Macq, B., Thiran, J.-P., and Taquet, M. (2019). Towards microstructure fingerprinting: Estimation of tissue properties from a dictionary of monte carlo diffusion mri simulations. *NeuroImage*, 184:964–980.

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

Romascano, D., Canales-Rodríguez, E., Rafael-Patino, J., Pizzolato, M., Rensonnet, G., Barakovic, M., Girard, G., Daducci, A., Dyrby, T., and Thiran, J.-P. (2019). HOTmix: characterizing hindered diffusion using a mixture of generalized higher order tensors. In *Proceedings of the 27th Annual Meeting of ISMRM (Montreal)*.

Rooney, W. D., Johnson, G., Li, X., Cohen, E. R., Kim, S.-G., Ugurbil, K., and Springer Jr, C. S. (2007). Magnetic field and tissue dependencies of human brain longitudinal 1H2O relaxation in vivo. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 57(2):308–318.

Sasaki, S. and Maruyama, S. (1992). Increase in diameter of the axonal initial segment is an early change in amyotrophic lateral sclerosis. *Journal of the neurological sciences*, 110(1-2):114–120.

Sato, K., Kerever, A., Kamagata, K., Tsuruta, K., Irie, R., Tagawa, K., Okazawa, H., Arikawa-Hirasawa, E., Nitta, N., Aoki, I., et al. (2017). Understanding microstructure of the brain by comparison of neurite orientation dispersion and density imaging (noddi) with transparent mouse brain. *Acta radiologica open*, 6(4):2058460117703816.

Scherrer, B., Schwartzman, A., Taquet, M., Sahin, M., Prabhu, S. P., and Warfield, S. K. (2016). Characterizing brain tissue by assessment of the distribution of anisotropic microstructural environments in diffusion-compartment imaging (DIAMOND). *Magnetic resonance in medicine*, 76(3):963–977.

Scherrer, B., Taquet, M., Schwartzman, A., St-Onge, E., Rensonnet, G., Prabhu, S. P., and Warfield, S. K. (2017). Decoupling axial and radial tissue heterogeneity in diffusion compartment imaging. In *International Conference on Information Processing in Medical Imaging*, pages 440–452. Springer.

Scherrer, B. and Warfield, S. K. (2012). Parametric representation of multiple white matter fascicles from cube and sphere diffusion MRI. *PLoS one*, 7(11):e48232.

Schilling, K., Gao, Y., Janve, V., Stepniewska, I., Landman, B. A., and Anderson, A. W. (2017). Can increased spatial resolution solve the crossing fiber problem for diffusion MRI? *NMR in Biomedicine*, 30(12):e3787.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61:85–117.

Schneider, T., Brownlee, W., Zhang, H., Ciccarelli, O., Miller, D. H., and Wheeler-Kingshott, C. G. (2017). Sensitivity of multi-shell noddi to multiple sclerosis white matter changes: a pilot study. *Functional neurology*, 32(2):97.

Sepehrband, F., Alexander, D. C., Kurniawan, N. D., Reutens, D. C., and Yang, Z. (2016). Towards higher sensitivity and stability of axon diameter estimation with diffusion-weighted MRI. *NMR in Biomedicine*, 29(3):293–308.

Setsompop, K., Kimmlingen, R., Eberlein, E., Witzel, T., Cohen-Adad, J., McNab, J. A., Keil, B., Tisdall, M. D., Hoecht, P., Dietz, P., et al. (2013). Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *Neuroimage*, 80:220–233.

Shintaku, M., Hirano, A., and Llena, J. (1988). Increased diameter of demyelinated axons in chronic multiple sclerosis of the spinal cord. *Neuropathology and applied neurobiology*, 14(6):505–510.

Slawski, M. and Hein, M. (2011). Sparse recovery by thresholded non-negative least squares. In *Advances in Neural Information Processing Systems*, pages 1926–1934.

Smith, S. A., Edden, R. A., Farrell, J. A., Barker, P. B., and Van Zijl, P. (2008). Measurement of T1 and T2 in the cervical spinal cord at 3 tesla. *Magnetic resonance in medicine*, 60(1):213–219.

Song, S.-K., Sun, S.-W., Ju, W.-K., Lin, S.-J., Cross, A. H., and Neufeld, A. H. (2003). Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage*, 20(3):1714–1722.

Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., et al. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage*, 80:125–143.

Spijkerman, J. M., Petersen, E. T., Hendrikse, J., Luijten, P., and Zwanenburg, J. J. (2018). T2 mapping of cerebrospinal fluid: 3T versus 7T. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 31(3):415–424.

Stanisz, G. J., Odrobina, E. E., Pun, J., Escaravage, M., Graham, S. J., Bronskill, M. J., and Henkelman, R. M. (2005). T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magnetic resonance in medicine*, 54(3):507–512.

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

Stikov, N., Campbell, J. S., Stroh, T., Lavelée, M., Frey, S., Novek, J., Nuara, S., Ho, M.-K., Bedell, B. J., Dougherty, R. F., et al. (2015). In vivo histology of the myelin g-ratio with magnetic resonance imaging. *Neuroimage*, 118:397–405.

Sun, S.-W., Liang, H.-F., Cross, A. H., and Song, S.-K. (2008). Evolving Wallerian degeneration after transient retinal ischemia in mice characterized by diffusion tensor imaging. *Neuroimage*, 40(1):1–10.

Szafer, A., Zhong, J., Anderson, A. W., and Gore, J. C. (1995). Diffusion-weighted imaging in tissues: theoretical models. *NMR in biomedicine*, 8(7):289–296.

Tanner, J. E. and Stejskal, E. O. (1968). Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. *The Journal of Chemical Physics*, 49(4):1768–1777.

Taquet, M. (2013). *Multi-Fascicle Models of the Brain Microstructure for Population Studies: Acquisition, Estimation, Registration and Statistical Analysis*. PhD thesis, UCL, Louvain-la-Neuve, Belgium.

Taquet, M., Scherrer, B., Boumal, N., Peters, J. M., Macq, B., and Warfield, S. K. (2015). Improved fidelity of brain microstructure mapping from single-shell diffusion MRI. *Medical image analysis*, 26(1):268–286.

Timmers, I., Roebroeck, A., Bastiani, M., Jansma, B., Rubio-Gozalbo, E., and Zhang, H. (2016). Assessing microstructural substrates of white matter abnormalities: A comparative study using dti and noddi. *PloS one*, 11(12):e0167884.

Ting, Y.-L. and Bendel, P. (1992). Thin-section MR imaging of rat brain at 4.7 T. *Journal of Magnetic Resonance Imaging*, 2(4):393–399.

Topgaard, D. (2017). Multidimensional diffusion MRI. *Journal of Magnetic Resonance*, 275:98–113.

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

Tournier, J., Calamante, F., Gadian, D. G., and Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185.

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582.

van de Ven, R. C., Hogers, B., van den Maagdenberg, A. M., de Groot, H. J., Ferrari, M. D., Frants, R. R., Poelmann, R. E., van der Weerd, L., and Kiihne, S. R. (2007). T1 relaxation in in vivo mouse brain at ultra-high field. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(2):390–395.

Van Nguyen, D., Li, J.-R., Grebenkov, D., and Le Bihan, D. (2014). A finite elements method to solve the bloch–torrey equation applied to diffusion magnetic resonance imaging. *Journal of Computational Physics*, 263:283–302.

Vangelderen, P., DesPres, D., Vanzijl, P., and Moonen, C. (1994). Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle. *Journal of Magnetic Resonance, Series B*, 103(3):255–260.

Vellmer, S., Edelhoff, D., Suter, D., and Maximov, I. I. (2017). Anisotropic diffusion phantoms based on microcapillaries. *Journal of Magnetic Resonance*, 279:1–10.

Veraart, J., Novikov, D. S., and Fieremans, E. (2018). TE dependent diffusion imaging (TEdDI) distinguishes between compartmental T2 relaxation times. *NeuroImage*, 182:360–369.

Visser, F., Zwanenburg, J. J., Hoogduin, J. M., and Luijten, P. R. (2010). High-resolution magnetization-prepared 3D-FLAIR imaging at 7.0 Tesla. *Magnetic resonance in medicine*, 64(1):194–202.

Wachowicz, K. and Snyder, R. E. (2002). Assignment of the T2 components of amphibian peripheral nerve to their microanatomical compartments. *Magnetic Resonance in Medicine*, 47(2):239–245.

Waller, A. (1850). Experiments on the section of the glossopharyngeal and hypoglossal nerves of the frog, and observations of the alterations produced thereby in the structure of their primitive fibres. *Philosophical Transactions of the Royal Society of London*, 140:423–429.

Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., and Westin, C.-F. (2016). The white matter query language: a novel approach for describing human white matter anatomy. *Brain Structure and Function*, 221(9):4705–4721.

Waudby, C. A. and Christodoulou, J. (2011). Gpu accelerated monte carlo simulation of pulsed-field gradient nmr experiments. *Journal of Magnetic Resonance*, 211(1):67–73.

Wedeen, V. J., Hagmann, P., Tseng, W.-Y. I., Reese, T. G., and Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic resonance in medicine*, 54(6):1377–1386.

Whitaker, S. (1967). Diffusion and dispersion in porous media. *AIChE Journal*, 13(3):420–427.

Winston, G. P., Micallef, C., Symms, M. R., Alexander, D. C., Duncan, J. S., and Zhang, H. (2014). Advanced diffusion imaging sequences could aid assessing patients with focal cortical dysplasia and epilepsy. *Epilepsy research*, 108(2):336–339.

Xing, D., Papadakis, N. G., Huang, C. L.-H., Lee, V. M., Carpenter, T. A., and Hall, L. D. (1997). Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. *Magnetic resonance imaging*, 15(7):771–784.

Xing, H., Lin, F., Wu, Q., and Gong, Q. (2013). Investigation of different boundary treatment methods in monte-carlo simulations of diffusion nmr. *Magnetic resonance in medicine*, 70(4):1167–1172.

Xu, J., Li, H., Harkins, K. D., Jiang, X., Xie, J., Kang, H., Does, M. D., and Gore, J. C. (2014). Mapping mean axon diameter and axonal volume fraction by MRI using temporal diffusion spectroscopy. *NeuroImage*, 103:10–19.

Ye, C. (2017). Tissue microstructure estimation using a deep network inspired by a dictionary-based framework. *Medical image analysis*, 42:288–299.

Zaimi, A., Duval, T., Gasecka, A., Côté, D., Stikov, N., and Cohen-Adad, J. (2016). AxonSeg: open source software for axon and myelin segmentation and morphometric analysis. *Frontiers in neuroinformatics*, 10:37.

Zhang, H., Dyrby, T. B., and Alexander, D. C. (2011). Axon diameter mapping in crossing fibers with diffusion MRI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 82–89. Springer.

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., and Alexander, D. C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61(4):1000–1016.

Zhang, J., Jones, M., DeBoy, C. A., Reich, D. S., Farrell, J. A., Hoffman, P. N., Griffin, J. W., Sheikh, K. A., Miller, M. I., Mori, S., et al. (2009). Diffusion tensor magnetic resonance imaging of Wallerian degeneration in rat spinal cord after dorsal root axotomy. *Journal of Neuroscience*, 29(10):3160–3171.

Zhang, Q., Su, P., Liao, Y., Guo, R., Qi, H., Hu, Z., Lu, H., and Chen, H. (2019). Deep learning based MR fingerprinting ASL reconstruction (DeepMARS). In *Proceedings of the 27th Annual Meeting of ISMRM (Montreal)*.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57.

Zucchelli, M., Brusini, L., Méndez, C. A., Daducci, A., Granziera, C., and Menegaz, G. (2016). What lies beneath? diffusion EAP-based study of brain tissue microstructure. *Medical image analysis*, 32:145–156.

Gaëtan Rensonnet

Born in Brussels on Feb. 16, 1992

+41786317586, +32475335818
gaetan.rensonnet@epfl.ch
gaetan.rensonnet@uclouvain.be

Education

2015-2019	Ph.D. in Applied Sciences	Ecole polytechnique fédérale de Lausanne (EPFL), Switzerland & Université catholique de Louvain, Belgium
	In vivo diffusion MRI of the brain microstructure from Monte Carlo signals	
2016	Research stay (3 months)	Harvard Medical School (HMS), Boston Children's Hospital
2015	Summer internship Sparse medical image registration	Massachusetts Institute of Technology, USA
2013-2015	Master Applied Math. Engineering Summa cum laude	Université catholique de Louvain, Belgium
	Master thesis : Generation of a dictionary of brain diffusion MRI signals	
2014	International exchange program Regular classes for one semester	Universidad Técnica Federico Santa María, Chile
2010-2013	Bachelor Civil Engineering Summa cum laude	Université catholique de Louvain, Belgium
2009-2010	High School exchange Complete senior year	New Buffalo High School, MI, USA

Awards and Fellowships

2019	Three-Minute Thesis contest (Ma Thèse en 180 Secondes) Belgian Edition finalist
2017-2018	WBI.World Excellence Scholarship for a 12-month stay at EPFL, Switzerland
2017-2019	International Society for Magnetic Resonance in Medicine (ISMRM) Educational stipend (2017, 2018 & 2019)
2016	WBI.World Excellence Scholarship for a 3-month stay at HMS, Boston, MA
2015-2019	Fellow of the Belgian National Fund for Scientific Research (F.RS.-FNRS)
2015	AILouvain Best Master Thesis Presentation Award
2015	IEEE UCLouvain Student Chapter Runner-Up for Best Master Thesis Award

Programming

Languages: Python, C/C++, Matlab, Java, Linux Bash
Machine/Deep Learning: attended the 2016 Medical Imaging Summer School (MISS 2016), passed the EPFL Master Class "EE559 – Deep Learning", proficient in PyTorch
Others : CMake, cluster computing (SLURM), SSH (Unix & Windows), LaTeX, Office, …

Languages

French: native English: C2 Spanish: C1 Dutch: C1

Publications (journal papers)

- Taquet, M., Jankovski, A., Rensonnet, G., Jacobs, D., des Rieux, A., Macq, B., Warfield, S.K., Scherrer, B., 2019. Extra-axonal restricted diffusion as an in vivo marker of reactive microglia. Scientific Reports (in press).
- Schilling, K.G., […], Rensonnet, G., […] & Landman, B. A., 2019. Limits to anatomical accuracy of diffusion tractography using modern approaches. Neuroimage, 185, pp.1-11.
- Rensonnet, G., Scherrer, B., Girard, G., Jankovski, A., Warfield, S.K., Macq, B., Thiran, J.P. and Taquet, M., 2018. Towards microstructure fingerprinting: Estimation of tissue properties from a dictionary of Monte Carlo diffusion MRI simulations. NeuroImage, 184, pp.964-980.
- Canales-Rodríguez, E.J., Legarreta, J.H., Pizzolato, M., Rensonnet, G., Girard, G., Patiño, J.R., Barakovic, M., Romascano, D., Alemán-Gomez, Y., Radua, J. and Pomarol-Clotet, E., 2018. Sparse wars: A survey and comparative study of spherical deconvolution algorithms for diffusion MRI. NeuroImage, 184, pp.140-160.

- Rensonnet, G., Scherrer, B., Warfield, S.K., Macq, B. and Taquet, M., 2017. Assessing the validity of the approximation of diffusion-weighted-MRI signals from crossing fascicles by sums of signals from single fascicles. Magnetic Resonance in Medicine.

[redacted] Links

MT180 contest interview (in French): https://www.youtube.com/watch?v=XVDDNPECXUg
MT180 final (in French): https://mt180.be/candidat/gaetan-rensonnet/?wpv-annee=2019
LinkedIn: https://www.linkedin.com/in/ga%C3%ABtan-rensonnet-695911b4/
Google Scholar: https://scholar.google.be/citations?user=m_I_9cAAAAAJ&hl=en
Research Gate: https://www.researchgate.net/profile/Gaetan_Rensonnet
Open-source code: https://ch.mathworks.com/matlabcentral/fileexchange/66645-nrrd-nhdr-reader-and-writer
(Computer codes developed during the PhD thesis will be publicly shared upon graduation, which is due in September 2019).

Ce document a été imprimé au Centre d'impression EPFL,
imprimerie climatiquement neutre, certifiée myClimate.