Analysis of Diffusion MRI: Disentangling the Entangled Brain
Analysis of Diffusion MRI: Disentangling the Entangled Brain

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op 2 september 2015 om 15:00 uur

door

Jianfei YANG
Master of Science in Condensed Matter Physics
Shandong University, China
geboren te Linyi, China
Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. L. J. van Vliet
Prof. dr. ir. C. B.L.M. Majoie
Copromotor:
Dr. F. M. Vos

Samenstelling promotiecommissie:
Rector Magnificus voorzitter
Prof. dr. ir. L. J. van Vliet Delft University of Technology, promotor
Prof. dr. ir. C. B.L.M. Majoie Academic Medical Center, promotor
Dr. F. M. Vos Delft University of Technology, copromotor
Prof. dr. W. Niessen Delft University of Technology
Prof. dr. ir. B. P.F. Lelieveldt Delft University of Technology
Prof. dr. R. Driche INRIA Sophia-Antipolis Méditerranée, France
Prof. dr. J. Sijbers University of Antwerp, Belgium
Prof. dr. A. M. Vossepoel Erasmus Medical Center, reserved

The work in this dissertation was conducted at the Quantitative Imaging Group, Department of Imaging Physics, Faculty of Applied Science, Delft University of Technology.

The usage of the painting in the cover has been authorized by the its author dr. Greg Dunn.
Cover Design: Proefschriftmaken.nl || Uitgeverij BOXPress
Printed & Lay Out by: Proefschriftmaken.nl || Uitgeverij BOXPress
Published by: Uitgeverij BOXPress, ’s-Hertogenbosch
Copyright © 2015 by Jianfei Yang
A digital copy can be download from repository.tudelft.nl
## Contents

1  Introduction .................................................................................................................. 1

   1.1 Methods of brain imaging ......................................................................................... 4

   1.2 Diffusion tensor imaging ......................................................................................... 5

   1.3 Thesis objectives ..................................................................................................... 11

   1.4 Outline .................................................................................................................... 12

2  Reliable dual tensor modeling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI ................................................................. 15

   2.1 Introduction ............................................................................................................. 16

   2.2 Methods .................................................................................................................. 18

      2.2.1 Dual tensor diffusion model .............................................................................. 19

      2.2.2 Maximum Likelihood Estimation of a constrained DTM .............................. 19

      2.2.3 Automatic Relevance Determination ........................................................... 21

      2.2.4 Jeffreys Prior ................................................................................................. 23

   2.3 Experiment and Results ......................................................................................... 24

      2.3.1 ARD versus MLE: Simulation Experiment ....................................................... 26

      2.3.2 ARD versus MLE: brain imaging ...................................................................... 29

      2.3.3 ARD versus MLE: Fractional Anisotropy along the CC ............................... 32

      2.3.4 ARD versus MLE: Dual-tensor FA and volume-fraction maps .................... 34

      2.3.5 TBSS based on dual tensor FA and volume-fraction maps ............................ 35

   2.4 Discussion and Conclusion ..................................................................................... 37

3  Estimation of diffusion properties in three-way fiber crossings without overfitting ......................................................................................................................... 45

   3.1 Introduction ............................................................................................................. 47

   3.2 Methods .................................................................................................................. 49

      3.2.1 Triple-tensor Model ......................................................................................... 50
3.2.2 Constraints ........................................................................................ 50
3.2.3 Priors on the parameters of the triple tensor model ....................... 52
3.2.4 Maximum a-posteriori (MAP) estimator ........................................... 52
3.2.5 Model Selection: ICOMP-TKLD ...................................................... 53
3.3 Experiment and Results ........................................................................... 56
3.3.1 Three-way Crossings: A Simulation Study ......................................... 56
3.3.2 Validation of Model Selection on Brain Data .................................... 63
3.4 Discussion and Conclusion ...................................................................... 69
4 A Structure-adaptive Noise Filter for Accurate Estimation of Diffusion Properties in Complex Fibre Structures ................................................................. 77
4.1 Introduction ............................................................................................ 79
4.2 Methods ................................................................................................... 81
4.2.1 Estimation of Local Structure Information ........................................ 82
4.2.2 Structure-Adaptive LMMSE .............................................................. 85
4.3 Validation on Synthetic ........................................................................... 88
4.3.1 Tensor Similarity ............................................................................... 88
4.3.2 Validation using phantom DWIs ....................................................... 89
4.4 Experiments and Results .......................................................................... 92
4.4.1 Data .................................................................................................. 92
4.4.2 Validation on experimental data ........................................................ 92
4.4.3 Application to a demo neuroimaging problem ................................... 98
4.5 Discussion and Conclusion .................................................................... 101
5 Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering .......................................... 109
5.1 Introduction .......................................................................................... 111
5.2 Methods ................................................................................................. 112
5.2.1 Determination of the Fiber Structure .............................................. 112
5.2.2 Diffusion Tensor Models................................................................. 113
1 Introduction
“The brain, the masterpiece of creation, is almost unknown to us.” -- Nicolaus Steno, 1669

Questions related to the brain’s physiology, such as what is the function of the brain, how does the brain work and control a person’s behaviour, have puzzled mankind for thousands of years. Effectively, ideas on the brain’s function evolved over time [1]:

(I.) The “Cranial Stuffing” concept. The earliest recorded reference to the brain was found in the hieroglyphic Edwin Smith Surgical Papyrus (see Figure 1.1) from the 17th century B.C. In ancient Egypt, the brain was widely regarded as a form of ‘cranial stuffing’. Accordingly, the brain was removed in the preparation of mummifications, while the heart was believed to be the seat of intelligence.

(II.) Cooling Agent. In the 4th century B.C. Aristotle considered the brain a secondary organ, which served to cool the heart. He defined a new concept, the ‘sensus communis’, to represent the origin of spirits. Supposedly, it was located in the heart. The concept is at the basis of our term ‘common sense’.

(III.) Seat of mental activities. The Roman physician Galen noticed the effect of brain injuries on mental activities. After this observation, the interpretation of the brain’s function entered a new period: the seat of mental activities. Until the 18th century, however, brain research was performed by gross anatomic dissection, due to the lacking of sophisticated equipment (see Figure 1.2). Therefore, knowledge on the brain’s physiology remained limited for long, as marvelled by Nicolaus Salernitanus “The brain… is, according to some, of hot complexity; according to others, cold; according to others, moist.”
Chapter 1. Introduction

Figure 1.1 Hieroglyphs representing the word “brain” (c. 1700 BC).

Figure 1.2 Base of the brain shown by Andreas Vesalius (1514-1564), showing the optic chiasma, cerebellum, olfactory bulbs, etc.
(figure source: http://en.wikipedia.org/wiki/Brain#/media/File:1543,AndreasVesalius%27sFabrica,BaseOfTheBrain.jpg)

After the 18th century A.D., the idea of the brain as the center of the nervous system became widely accepted. Recently, after centuries of investigating the brain by means of dissectioning, non-invasive imaging based on magnetic resonance imaging (MRI) (see Figure 1.3) enabled highly sophisticated, in vivo studies of the human brain.
Chapter 1. Introduction

1.1 Methods of brain imaging

The approaches to investigate the human brain can be classified into ex-vivo and in-vivo approaches.

Ex-vivo approaches

Initially, Andreas Vesalius (1514-1564) observed common structural characteristics in the human brain and the peripheral nervous system (see Figure 1.2). Brain dissection has been the most important and direct way to investigate structural features of the human brain ever since. Even now, it is regarded the sole way to obtain a ground truth on structural features of the brain. Since dissection falls outside the scope of this thesis, we simply acknowledge its significance for in-vivo research.

In-vivo approaches

The techniques for in-vivo investigation of the brain are generally referred to as neuroimaging. These approaches include various methods to directly or indirectly image the structure, function and chemistry of the nervous system [2].

The origin of neuroimaging work traces back to the pioneering work by neuroscientist Angelo Mosso in the 19th century. Mosso introduced a technique
Chapter 1. Introduction

called the “human circulation balance” which non-invasively measured the redistribution of blood during emotional and intellectual activities [3]. Presently, it is considered a crude precursor to more refined, currently used techniques.

In the 1970s, the introduction of computerized axial tomography based on X-ray imaging is a milestone in the history of neuroimaging. As such, cross-sectional images of the brain became available for diagnostic and research purposes. A little later, other cross-sectional modalities were invented such as single-photon emission computed tomography (SPECT) and positron emission tomography (PET), both depicting the distribution of radioactive isotopes in the body. Also, magnetic resonance imaging was introduced, which used magnetic fields at radiofrequencies to produce images of the brain. The method manipulates the proton spins to generate a variety of tissue contrasts.

Compared to the X-ray and radioactive imaging techniques, MRI attracted more and more attention, especially because it does not expose the subject to ionizing radiation and also due to the versatility of the method. Accordingly, MRI based methods were widely used to investigate structural as well as functional aspects of the brain. The scope of this thesis is on one specific MRI modality named diffusion-weighted imaging, which was found to depict microscopic details about the tissue architecture. Essentially, it measures aspects of the diffusion of water molecules.

1.2 Diffusion tensor imaging

All conventional MR imaging techniques (e.g. T1- or T2-weighted) use three-dimensional phase and frequency encoding to spatially resolve the generated signals (see Figure 1.4). This is performed by application of magnetic field gradients that impose spatially varying precession frequencies of the spins [4]. During data acquisition, the measured MR signals are stored in k-space, which contains the spatial frequency domain of the acquired image. In other words, the raw MRI data in k-space is the Fourier transform of the frequency-encoded MR data.
**Diffusion-weighted MR imaging** (DW-MRI or DWI) was originally proposed in the mid-1980s by D. LeBihan [5]. Essentially, two identical, but oppositely oriented gradient pulses are added (see Figure 1.5) to a conventional pulse sequence to sensitize the imaging to molecular movement. If there is no molecular movement in the gradient direction, the effects of the two balanced gradient pulses cancel out. In that case, there is no influence on the signal intensity, i.e. the signal intensity will be the same as without the gradients. Instead, if there is movement along the gradient orientation, the effect of the two gradients will not cancel out, leading to a lower signal intensity (see Figure 1.6). DW-MRI is applied in clinical practice, especially for detecting cerebral abnormalities in acute stroke [6] (see Figure 1.7).

![Pulse Sequence Timing Diagram](http://www.mr-tip.com/serv1.php?type=db1&dbs=Spin+Echo+Sequence)

**Figure 1.4** A spin echo pulse sequence, which is one of the most commonly used MRI sequences. The pulse sequence timing can be adjusted to give T1-weighted, proton or spin density, and T2-weighted images. RF: Radio Frequency pulse. (figure source: http://www.mr-tip.com/serv1.php?type=db1&dbs=Spin+Echo+Sequence)
Chapter 1. Introduction

Figure 1.5 A pulsed gradient spin echo sequence used for diffusion MR imaging. Two diffusion-encoding gradient pulses (GDiff) are added to the standard spin echo sequence, which will introduce a phase shift that is proportional to molecular displacement along the gradient orientation. GS, GP, and GF represent Slice selective, Phase encoding, and Frequency encoding gradients, respectively. Echo represents the signal received from the slice of interest. (figure source: http://radiopaedia.org/articles/spin-echo-sequences)

Figure 1.6 The anisotropic nature of diffusion in the brain. Transverse DW MR images due to diffusion gradients applied along the x (Gx, left), y (Gy, middle), and z (Gz, right) axes demonstrate diffusional anisotropy. The signal intensity decreases when white matter tracts have the same orientation as the DW gradient because water protons move preferentially along such tracts. Note that the corpus callosum (arrow in the left image) is hypointense when the gradient is applied in the x (right-to-left) direction; the frontal and posterior white matter (arrowhead in the middle image) are hypointense when the gradient is applied in the y (anterior-to-posterior) direction, and the corticospinal tract (arrow in the right image) are hypointense when the gradient is applied in the z (superior-to-inferior) direction. (figure source: [7])
Chapter 1. Introduction

Figure 1.7 MRI of a patient with acute stroke, 2 hours postictus. The T2-weighted image is normal. The fluid-attenuated inversion recovery image (FLAIR) shows hyperintense vessels in the region of the middle carotid artery (MCA) (arrow), consistent with slow arterial flow, but it shows no tissue abnormality. The DWI shows hyperintensity in the deep middle cerebral artery's territory consistent with cytotoxic edema in acute stroke (arrow), but there is no evidence of cortical ischemia. Perfusion-weighted imaging (PWI) (time-to-peak image) shows reduced perfusion in the full (cortical and subcortical) MCA region. (figure source: [6]).

**Diffusion Tensor Imaging** (DTI) classically fits a single rank-2 tensor model to a series of DWIs to summarize the principal directions and shape characteristics of the molecular motion [8]. It is generally presumed that the myelin sheet encapsulating the nerve cells hinders the diffusion perpendicular to the fiber orientation. Consequently, the diffusion in this direction is slower than the diffusion in the axial direction: the diffusion is anisotropic (see Figure 1.8a). In the absence of such fibrous structures, the diffusion is approximately isotropic (see Figure 1.8b). The anisotropic diffusion is characterized by the shape and orientation information of the estimated rank-2 tensor model (see Figure 1.9a). As such, the diffusion process and the orientation of fibre bundles can be reconstructed (see Figure 1.9b and c).
Figure 1.8. Anisotropic and isotropic diffusion: (a) when motion is constrained, as inside white–matter tracts (illustrated on the right), diffusion is anisotropic, meaning that there is more motion in one direction than in another (e.g. more movement along the tract than perpendicular to the tract); (b) water molecules in the brain constantly exhibit Brownian motion. When motion is unconstrained, as in the large fluid–filled spaces in the brain (i.e., the ventricles, as illustrated in the MR image), diffusion is isotropic, which means that the motion occurs equally and randomly in all directions. (figure source: [9]);

Figure 1.9 (a) Visualization of the shape of a rank-2 tensor model; (b) Fractional Anisotropic (FA) map showing a rotation-invariant feature of diffusion shape; (c) tractography reconstruction of neural connections from DTI. (figure source: (a), http://www.diffusion-imaging.com/2012/10/voxel-based-versus-track-based.html ; (b) and (c), http://neuroimaging.tau.ac.il/ya/research1.html)
DTI and DW-MRI have been successfully applied both in research and clinical practice. Still, several challenges remained:

(I.) A single rank-2 tensor cannot describe complex fiber structures. Classical DTI uses a single tensor to describe the diffusion. Such a single tensor is known to accurately characterize the diffusion in voxels containing a single-fiber or isotropic matter. However, it is well known that a large part of the brain contains two-way and three-way fiber crossings. A single ellipsoid is not appropriate to describe the diffusion shape of such structures. Typically, the classical single tensor will give ‘disc’ shape in crossings (see Figure 1.10). There is an on-going debate regarding how to effectively model the diffusion process in complex fiber structures.

(II.) The characterization of complex fiber structures is hampered by a low signal-to-noise ratio (SNR). To better represent the diffusion in two-way and three-way fiber crossings, models with more degrees of freedom (DOF), such as kank-2 multi-tensor models, rank-4 tensor models, CHARMED, DKI etc, are required. To map the diffusion process in voxels with complex fiber-structures, strong diffusion-weightings are needed to facilitate the parameter estimation. Unfortunately, this goes hand in hand with a lower signal-to-noise-ratio (SNR) of the MR measurements. Therefore, finding an effective way to cope with the low SNR of MR signals is needed.

(III.) The application of complex models to the DW signals in voxels encompassing a single fiber structure is ill-posed. The parameters of a complex diffusion model such as the dual-tensor model cannot be uniquely determined in voxels encompassing a single fiber structure. Finding an effective way to avoid this ill-posedness in the application of higher-order rank-2 tensor models is an important challenge.
1.3 Thesis objectives

This thesis addresses the following objectives to improve the mapping of diffusion processes in the brain and facilitate better quantification of (pathological) changes.

We aim to accurately and precisely estimate diffusion properties in the brain based on multi-tensor representations by carefully modelling the fiber structures. Particularly, we aim to improve the estimation of diffusion properties of complex structures, such as two-way or three-way crossings. Therefore, we target to introduce novel model selection approaches that data-adaptively estimate diffusion properties. The order of a multi-tensor model, i.e. the number of tensor compartments in the multi-tensor model, must accord with the number of fiber orientations in a voxels: a single-tensor model for single fiber voxels, dual- and triple-tensor models for voxels with two- and three-way crossings respectively.

A related objective is to improve the estimation of diffusion properties in complex fiber structures by filtering the data to enhance the signal to noise ratio. Therefore, we aim to improve a Linear Minimum Mean Square Error (LMMSE) estimator for DWI data and a bilateral filter for the tensor data.

All techniques target to enhance the sensitivity of statistical analysis for detection of deviations in brain structure.
Chapter 1. Introduction

1.4 Outline

The content of this thesis can be divided into two parts:

PART I: Structure-adaptive modelling of the diffusion process

Chapter 2 introduces a method for reliable dual-tensor modelling of simple and two-way crossing fiber geometries. A framework using Automatic Relevance Determination steered by Jeffreys prior is described, which enables the dual-tensor model to data-adaptively characterize the diffusion. Hence, a second tensor compartment only survives the estimation if it is supported by the data.

Chapter 3 explores data-adaptive estimation of single, two-way and three-way fiber crossings. The estimation of diffusion properties is made reliable by a new model-order selection technique for rank-2 tensor models based on the total Kullback-Leibner divergence. The novel model-order selection criterion balances two terms in the application of a more complex model: a lower data misfit term and a potential higher parameter dependence term.

PART II: DWI and DTI filtering steered by structural information

Chapter 4 describes a novel, compartment-specific and data-adaptive DWI noise-suppression filter to enhance the estimation of diffusion properties in complex fiber structures. In this chapter, we will introduce a linear minimum mean square error (LMMSE) estimator, which is steered by tentatively estimated structural information and applied to compartment-specific DWI contributions.

Chapter 5 goes into estimating diffusion properties in complex configurations based on data-adaptive tensor-field filtering. In this chapter, a ‘bilateral’ filter is introduced which is guided by tensor similarity, c.q. representing the similarity of the diffusion process.

Finally, in Chapter 6 the merits and limitations the proposed methods will be summarized and discussed.
Chapter 1. Introduction

References


2 Reliable dual tensor modeling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

This chapter studies a framework for reliable modeling of diffusion MRI using a data-adaptive prior.

Automated relevance determination (ARD) estimates the mean of the posterior distribution of a rank-2 dual tensor model employing Jeffreys prior. This adaptive prior is based on the Fisher information matrix and enables the assessment whether two tensors are mandatory to describe the data. The method is compared to Maximum Likelihood Estimation (MLE) of the dual tensor model and to FSL’s ball-and-stick approach.

Monte Carlo experiments demonstrated that ARD’s volume fractions correlated well with the ground truth for single and crossing fiber configurations. In single fiber configurations ARD automatically reduced the volume fraction of one compartment to (almost) zero. Thereby, the variance in fractional anisotropy (FA) of the main tensor component was reduced compared to MLE. ARD and MLE gave a comparable outcome in data simulating crossing fibers. On brain data, ARD yielded a smaller spread in FA along the corpus callosum compared to MLE. Tract-based spatial statistics demonstrated a higher sensitivity in detecting age-related white matter atrophy using ARD compared to MLE by either single or dual tensor modeling and the ball-and-stick approach.

The proposed framework offers accurate and precise estimation of diffusion properties in single and dual fiber regions.
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

2.1 Introduction

Diffusion-weighted magnetic resonance imaging (DW-MRI) can provide unique information about the integrity of white matter structures in the brain. Conventionally, the diffusion is described by a single rank-2 diffusion tensor [1], which is estimated from diffusion weighted images (DWIs). There is an ongoing debate on how to effectively characterize the diffusivities in voxels containing complex anatomical structures such as crossing fibers. In these voxels the diffusion profile is not adequately described by a single rank-2 tensor [2] [3].

Several sophisticated models of the diffusion in white matter have shown the potential to estimate more plausible anatomical properties of the tissue, for instance the ‘ball & stick’ model [4], the composite hindered and restricted model of diffusion (CHARMED) [5], and the neurite orientation dispersion and density imaging approach (NODDI) [6]. These references relate to our approach, but many more techniques were proposed. A comprehensive overview is presented in [7].

As a first, the ‘ball & stick’ approach models the diffusion by one isotropic compartment and an array of linear, 1D diffusion profiles. The method is widely applied and worked well for reconstructing the orientations of fiber bundles [8], even though these fiber bundles are not represented by full, 3D diffusion profiles. Alternatively, CHARMED explicitly models the slowly diffusing component arising from restricted, intra-axonal diffusion (a non-Gaussian process). It yielded unbiased estimation of the orientations of two or more fiber compartments with low angular uncertainty. However, the application of CHARMED is challenging due to measurement at very high b-values causing signal-to-noise and scanning time limitations.

More recently, a clinically feasible technique for in vivo neurite orientation dispersion and density imaging (NODDI) [6] was proposed. NODDI adopts a tissue model that distinguishes three types of microstructural environment: intra-cellular, extra-cellular, and cerebrospinal fluid (CSF) compartments. The signal of intra-cellular diffusion is described by zero radius cylinders [9] (like in the ball-and-stick approach); the extra-cellular part is modeled by anisotropic, Gaussian diffusion and the CSF compartment is modeled as isotropic Gaussian diffusion. Experiments showed that indices of derived
from NODDI such as the neurite orientation dispersion provided more specific markers of brain tissue microstructure than standard indices from classical, single-tensor DTI [6].

Furthermore, other methods [10] [11] [12] were proposed that also aim to recover specific tissue parameters from the diffusion signal, such as cell size and cell density. However, these models are not directly compartment-specific. Instead, a multiple-tensor model [13] describes the diffusion through a sum of fully 3D tensor compartments. The model is an intuitive, physical representation that is a natural extension of the classical single tensor. Also, it does not necessarily require extremely high b-values. Previously, we introduced an optimization framework that rendered a constrained dual tensor model (DTM) as well as a set of diffusion weighting parameters, such that both the diffusion shape and the diffusion orientations of crossing fibers could be accurately estimated [3].

Unfortunately, even a constrained DTM is prone to overfitting in areas containing a single fiber bundle, causing biased volume fractions and imprecise diffusivity estimates. As such, a new challenge arises: how to automatically adapt the model complexity to warrant an accurate characterization of the underlying diffusion processes?

Many model selection methods were introduced in the DWI field, for instance based on constrained spherical deconvolution (CSD) [14], the Bayesian information criterion (BIC) and the generalization-error [15]. A limitation of the CSD approach is that it requires tuning of a threshold to reject small contributions. Furthermore, BIC is only determined by non-estimated factors such as the number of parameters and the sample size and the generalization-error method is a non-local model selection technique. Importantly, all these approaches involve model selection techniques that make hard decisions to select an appropriate model.

Automatic relevance determination (ARD) aims to eliminate the redundant parameters in a complex model, such that a simplified model yields a better description of the data [16]. Behrens [8] adopted ARD for assessing the appropriate number of fiber orientations in each voxel for fiber tracking. This method ensures that if there is no evidence for a second fiber orientation in the data, the volume fraction attributed to this fiber will automatically be forced to zero. ARD methods assume a
prior distribution for the model parameters. A Gaussian distribution is a straightforward choice for a prior [16]. Such a prior may involve hyper-parameters to tune its shape. Previous ARD approaches [8] [17] [18] involved marginalization (integration) over the hyper-parameters to get a prior for each parameter separately. Such a prior is likely to be suboptimal for individual voxels since potential correlations between parameters are ignored.

We present a new framework for data-adaptive estimation of the diffusion shape of simple and complex white matter structures. We consider the method data-adaptive as it takes properties of the data-acquisition into account such as the number of gradient directions, b-values used and the noise level. The method is based on ARD for a rank-2 dual tensor model and assesses whether two anisotropic tensors are ‘mandatory’ to model the acquired diffusion-weighted signals. Our ARD estimates the mean of the a posteriori distribution, i.e. the model parameters given the data, employing Jeffreys prior [19] [20]. This data-adaptive prior is based on the Fisher’s information matrix [21]. Previous work on ARD for diffusion weighted MRI primarily focused on the accurate reconstruction of fiber orientations based on the ball-and-stick model [8] [17]. This rank-1 tensor model is not appropriate for estimating the diffusion shape as reflected by a rank-2 tensor model. The proposed ARD method is particularly suited for application in comparative studies in which the goal is to assess subtle differences in diffusion shape between patients and matched controls.

2.2 Methods

The proposed ARD framework for estimation of the diffusion shape processes every voxel in the same way. It estimates the parameters of a constrained dual tensor model (DTM) by computing the mean of the posterior distribution sampled by a Markov Chain Monte Carlo (MCMC) approach. The algorithm is initialized by applying the constrained DTM to the measured diffusion-weighted signals using maximum likelihood estimation (MLE). The prior on the parameters in the MCMC sampling is given by the non-informative Jeffreys prior. This prior forces parameters, particularly the volume fraction, towards zero when there is little to no information in them. This will be verified experimentally in the Experiments and Results section.
2.2.1 Dual tensor diffusion model

We assume that the diffusion in every fiber bundle is mono-exponential and Gaussian. The diffusion-weighted signal in all voxels is initially modeled by a so-called dual tensor model (DTM) [3]. This model contains signal contributions of up to two fiber bundles and an isotropic component and is given by

\[
S_{\theta,j} = S_0 \left( \sum_{i \in \{1,2,\text{iso}\}} f_i \exp(-b_j g_j^T D_i g_j) \right),
\]  

(2.1)

where \(S_{\theta,j}\) is the diffusion-weighted signal given parameter vector \(\theta\) for diffusion weighting \(b_j\) in gradient direction \(g_j\) and \(S_0\) the signal without diffusion weighting. \(D_1\) and \(D_2\) are rank-2 tensors to model the anisotropic diffusion in each fiber, \(D_{\text{iso}}\) is the amount of isotropic diffusion (i.e., \(D_{\text{iso}} = 1_{\text{3x3}}\)), \(D_{\text{iso}}\) representing the scalar amount of isotropic diffusion), and \(f_i\) represents the volume fraction of component \(D_i\). Note that the DTM in Eq. (2.1) reduces to a single tensor model (STM) – reflecting a single fiber – if \(f_1 > 0\) and \(f_2 = 0\) or vice versa. The volume fraction parameters play an essential role in our ARD scheme.

2.2.2 Maximum Likelihood Estimation of a constrained DTM

The measured diffusion weighted image (DWI) \(\tilde{S}_{j,\sigma}\) with diffusion weighting \(b_j\) in direction \(g_j\) is corrupted by Rician noise of standard deviation \(\sigma\) [22]. Therefore, the probability density function (PDF) for \(\tilde{S}_{j,\sigma}\) is given by

\[
p(\tilde{S}_{j,\sigma} \mid \theta) = \frac{\tilde{S}_{j,\sigma}}{\sigma^2} \exp \left( \frac{\tilde{S}_{j,\sigma}^2 + S_{\theta,j}^2}{2\sigma^2} \right) I_0 \left( \frac{\tilde{S}_{j,\sigma} S_{\theta,j}}{\sigma^2} \right),
\]  

(2.2)

with \(I_0(\cdot)\) the zero-th order modified Bessel function of the first kind. The DWIs are statistically independent, so that the joint probability density function \(p(\tilde{S}_{\sigma} \mid \theta)\) of the signal profile \(\tilde{S}_{\sigma}\) is given by the product of the marginal distributions for the measured signals \(\tilde{S}_{j,\sigma}\) in each of the \(N_g\) diffusion weighted directions \(g_j\):
\[
\mathbf{p}(\mathbf{S}_\sigma | \mathbf{\theta}) = \prod_{j=1}^{N_s} \mathbf{p}(\mathbf{S}_{j,\sigma} | \mathbf{\theta})
\]  

(2.3)

Here, \( p(\mathbf{S}_\sigma | \mathbf{\theta}) \) is the likelihood function of \( \mathbf{\theta} \) given \( \mathbf{S}_\sigma \). The underlying parameter values can be inferred by maximizing this likelihood function [23]

\[
\hat{\mathbf{\theta}}_{\text{MLE}} = \arg \max_{\mathbf{\theta}} \left\{ p(\mathbf{S}_\sigma | \mathbf{\theta}) \right\}.
\]  

(2.4)

Maximum likelihood estimation (MLE) has a number of favorable statistical properties in the estimation of diffusion properties in crossing bundles [3]. First, under very general conditions, MLE asymptotically reaches the Cramér-Rao lower bound (CRLB). The CRLB is a theoretical lower bound on the variance of any unbiased estimator. Second, the MLE is consistent, which means that it asymptotically \( (N_s \to \infty) \) converges to the true value of the parameter in a statistically well-defined way [24].

The dual tensor model given in (2.1) should be parameterized such that its parameter values reside in a well-defined range. In previous work [3], we parameterized the tensor \( \mathbf{D}_i \) as follows:

\[
\mathbf{D}_i = \mathbf{R}_i^T \mathbf{E}_i \mathbf{R}_i,
\]

where \( \mathbf{E}_i = \text{diag}(\lambda_{i,1//}, \lambda_{i,11}, \lambda_{i,12}) \) is a diagonal matrix with the eigenvalues of the tensor \( \mathbf{D}_i \) on its diagonal. The non-negativity constraint that is imposed on the estimated diffusivity values is accomplished by employing an exponential mapping [3]. The matrices \( \mathbf{R}_{i=1,2} \) describe rotations around the \( x- \), \( y- \) and \( z- \) axes:

\[
\mathbf{R}_i(\alpha_{i-}) = \mathbf{R}_z(\alpha_i) \mathbf{R}_y(\alpha_2) \mathbf{R}_z(\alpha_3 \pm \alpha_4) .
\]

The first two rotations \( \mathbf{R}_z(\alpha_i) \mathbf{R}_y(\alpha_2) \) determine the orientation of the plane in which the first principal eigenvectors of both tensor reside. \( \mathbf{R}_z(\alpha_3 + \alpha_4) \) and \( \mathbf{R}_z(\alpha_3 - \alpha_4) \) denote the in-plane rotations of the first principal eigenvector of the two tensors. As such, the parameter vector to be estimated for a dual-tensor model MLE becomes

\[
\mathbf{\theta} = \{ f_1, f_2, f_{\text{iso}}, \lambda_{i,1//}, \lambda_{i,11}, \lambda_{i,12}, \lambda_{2,1//}, \lambda_{2,11}, \lambda_{2,12}, D_{\text{iso}}, \alpha_1, \alpha_2, \alpha_3, \alpha_4 \} \ldots (2.5)
\]

However, MLE does not necessarily yield useful estimates. A potential error in the estimated parameters is greatly influenced by the degrees of freedom (DOFs) and the covariance(s) between parameters. We demonstrated that restricting the DOFs by imposing constraints on the DTM greatly reduces the covariance between parameters.
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

The experiments in [3] showed that precise and accurate estimation can be achieved if we apply the following constraints:

\[ \lambda_{1,\parallel} = \lambda_{2,\parallel}, \]
\[ \lambda_{1,\perp} = \lambda_{1,\perp}, \]
\[ \lambda_{2,\perp} = \lambda_{2,\perp}, \]
\[ D_{iso} = C_{\text{free-water}}, \]
\[ f_1 + f_2 + f_{iso} = 1. \]

Eq. (2.6) imposes that the “unrestricted” diffusivity (i.e., free diffusivity) along the fibers, denoted by the first eigenvalues of \( D_1 \) and \( D_2 \), are equal. Eq. (2.7) states that the diffusion perpendicular to the fiber orientation is assumed to be axially symmetric, which models the average shape of axons. Eq. (2.8) defines that \( D_{iso} \) equals \( C_{\text{free-water}} = 3 \times 10^{-3} \text{mm}^2\text{s}^{-1} \), the diffusivity of free water at body temperature 37°C, and Eq. (2.9) states that the two anisotropic tensors plus the isotropic compartment fill the entire volume of each voxel.

Constraining the DTM cannot avoid the inherent risk of overfitting. This happens when a complex model is fitted to simple data, e.g., fitting multiple tensors to data of a single fiber bundle. Typically, this yields an increase of the variance and the covariance of the parameters, but also leads to biased diffusivity estimates.

2.2.3 Automatic Relevance Determination

Bayes factors offer an alternative to model selection by the classical likelihood test [25]. It computes the evidence for a model to be used in model selection. However, calculating the evidence for any model requires integration over all model parameters, weighted by the parameter priors. This is computationally unfeasible, especially with high-dimensional parameter spaces for which no analytical solution exists. ARD was introduced exactly to cope with such issues [8] [16] [17] [26] [27]. Compared to the Bayes factors approach, ARD does not fit competing potential models to the data and compares them on the basis of the residual after fitting. Instead, ARD always fits a complex model to the data and forces irrelevant parameters to zero, so that a complex model reduces to a simpler one.
Our ARD estimates the mean of the posterior distribution of the constrained DTM based on Bayes’ theorem [16]. The posterior distribution $p(\theta | \tilde{S}_\sigma)$ is

$$p(\theta | \tilde{S}_\sigma) = \frac{p(\tilde{S}_\sigma | \theta)p(\theta)}{p(\tilde{S}_\sigma)}, \tag{2.10}$$

where $p(\tilde{S}_\sigma | \theta)$ is the aforementioned likelihood function, $p(\theta)$ the prior probability of $\theta$ in the DTM, and $p(\tilde{S}_\sigma)$ the evidence for the DTM. As the evidence term in Eq. (2.10) is constant for any measured signal, the posterior probability distribution in ARD becomes

$$p(\theta | \tilde{S}_\sigma) \propto p(\tilde{S}_\sigma | \theta)p(\theta). \tag{2.11}$$

Our framework estimates the posterior distribution given the data, which is influenced by the likelihood function and the prior. We introduce a data-adaptive prior for DTM parameters based on Jeffreys theorem (see next subsection). It allows simplifying a complex model to a simple model by automatically forcing volume fractions, which are not supported by data to zero.

Our ARD employs a Markov Chain Monte Carlo (MCMC) technique with Metropolis-Hasting sampling of the posterior distribution $p(\theta | \tilde{S}_\sigma)$ [28] [29]. Our MCMC draws 5000 samples from the posterior distribution in the nine-dimensional parameter space. The algorithm is listed in Table 2.1. It is initialized by MLE of the constrained DTM. The final ARD estimate $\hat{\theta}_{ARD}$ is the mean of 3000 accepted samples after a burn-in period of 2000.

If the posterior estimates of both anisotropic fractions lie in a small interval around their MLE value, then this would indicate that Jeffreys prior did not significantly change the outcome and that fitting the initial dual tensor model was justified. Reversely, if the posterior estimate for one of the two anisotropic fractions does not significantly differ from zero, then its corresponding tensor compartment can be treated an unnecessary parameter. In such a case, the estimation essentially returns a ‘single-tensor’ model.
Table 2.1: Algorithm for estimating the mean of a multivariate posterior distribution with Jeffreys prior using a Markov Chain Monte Carlo method employing Metropolis-Hastings sampling. The ‘proposal’ distribution $Q(\theta' | \theta_t) = \theta_t + \Delta \cdot N(0,1)$ with $N(0,1)$ denoting a multivariate Gaussian distribution with zero mean and standard variance 1; and $\Delta$ the step size for parameter vector $\theta$. The step sizes for all parameters are: $10^{-5}$ (with unit mm$^2$s$^{-1}$) for the diffusivity parameters, $10^{-2}$ (with unit rad) for the angles, and $10^{-2}$ for the volume fractions.

Algorithm: Markov Chain Monte Carlo with Metropolis-Hastings sampling

For all voxels

$\theta_0 = \hat{\theta}_{MLE}$ // Initialize vector

For $t = 0$ to $N-1$

$\theta' = Q(\theta_t | \theta_t)$ // Draw candidate from ‘proposal’ distribution

$p(\theta') = \det(I(\theta'))^{1/2}$ // Calculate Jeffreys prior

$p(\theta', \tilde{S}) = p(\tilde{S} | \theta') p(\theta')$ // Calculate posterior probability

$\alpha = p(\theta | \tilde{S}) / p(\theta_t | \tilde{S})$ // Calculate the acceptance ratio

if $\alpha \geq 1$ then

$\theta_{t+1} = \theta'$ // Accept the candidate vector

else

$r = U(0,1)$ // Draw random variable $r$ between 0 and 1;

if $r \leq \alpha$

$\theta_{t+1} = \theta'$ // Accept the candidate vector

else

$\theta_{t+1} = \theta_t$; // Keep the previous vector

endif

endif

Endfor

$\theta_{ARD} = \frac{1}{N - N_{burn-in}} \sum_{t=N_{burn-in}}^{N-1} \theta_t$ // Compute mean of accepted samples after burn-in

Endfor

2.2.4 Jeffreys Prior

Methods to choose the prior for a Bayesian analysis can be divided into two groups: informative and non-informative priors [27] [30]. Informative priors use knowledge on the distributions of the parameters e.g. [27] [31], whereas non-informative priors rely on the data to steer the posterior [19] [20]. We aim to introduce a new, data-adaptive prior, which renders the ARD method essentially non-informative. Specifically, we adopt Jeffreys non-informative prior $p(\theta)$ which can be written as:

$$p(\theta) \propto \det(I(\theta))^{1/2},$$

(2.12)

where $I(\theta)$ denotes the Fisher information matrix given by
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

\[ I(\theta) = -\mathbb{E}_S \left\{ \frac{\partial^2 \ln(p(S|\theta, \sigma))}{\partial \theta \partial \theta^T} \right\}. \]  

(2.13)

The Fisher information matrix \( I(\theta) \) provides the amount of expected information about the parameter vector \( \theta \) in measurements. By definition, it is influenced by properties of the data-acquisition such as the number of data points and the noise. Jeffreys prior is in agreement with one’s intuition that if a parameter is necessary, it must be supported by the data. Poot [32] showed that the Fisher information matrix for Rice distributed measurements given by Eq. (2.13) can be efficiently computed. Jeffreys prior conveys support for the dual tensor representation in crossing fibers, as should be large. In that case, the prior only mildly affects the posterior distribution, which typically yields a peak near the initial dual tensor parameters obtained by MLE. Reversely, the determinant of Fisher’s information matrix is expected to be smaller in a single fiber region, where a dual-tensor model is overfitting to the data. Then, the prior becomes harsh, promoting a near-zero volume fraction in the posterior distribution.

2.3 Experiment and Results

All experiments below were carried out on a DELL laptop computer with an Intel i7-2620CPU @2.7GHz and 4.00 GB RAM running the Window-7 64-bit operating system. The method was implemented in MATLAB_R2014b. The average execution time on the brain image data was 6s/voxel. In the first part of this section we evaluate the performance of estimating the parameters of our constrained dual tensor model by ARD and by MLE on simulated data. Henceforth, these two approaches are simply abbreviated by ARD and MLE. We studied the differences between ARD and MLE as a function of the volume fraction for simulated crossing fibers under realistic conditions. Diffusion measurements were simulated by means of the model presented in Eq. (2.1). The parameters of crossing fibers are listed in Table 2.2 and are in agreement with the work of Pierpaoli [33] who reported diffusivities ranging from \( 0.25 \times 10^{-3} \) to \( 1.5 \times 10^{-3} \ mm^2/s \). The SNR
(defined by $S_0/\sigma$) was 25 \cite{3}. The gradient directions for the two b-values (1.0 \cdot 10^3 \text{ mm}^{-2}\text{s} and 3.0 \cdot 10^3 \text{ mm}^{-2}\text{s} ) were homogeneously distributed over the surface of a sphere as reported by \cite{3}. Two measurements at $b = 0 \text{ mm}\text{-2 s}$ were simulated. Furthermore, 92 gradient directions were adopted for each of two b-values (1.0 and 3.0 ), homogeneously distributed over the surface of a sphere. These settings are identical to dataset B (see below) and equal to those reported in \cite{3}.

\textbf{Table 2.2}: Model parameters for generating synthetic data. The units of the diffusion parameters $\lambda$ are $10^3 \text{ mm}^2\text{s}^{-1}$.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Quantity/Description</th>
<th>Value ($\theta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{1,1}$</td>
<td>largest eigenvalue of tensor-1 representing the diffusivity along fiber-1</td>
<td>1.480</td>
</tr>
<tr>
<td>$\lambda_{2,1}$</td>
<td>second largest eigenvalue of tensor-1 representing the largest diffusivity perpendicular to fiber-1</td>
<td>0.15</td>
</tr>
<tr>
<td>$\lambda_{2,2}$</td>
<td>smallest eigenvalue of tensor-1</td>
<td>0.12</td>
</tr>
<tr>
<td>$\lambda_{2,2}$</td>
<td>largest eigenvalue of tensor-2 representing the diffusivity along fiber-2</td>
<td>1.400</td>
</tr>
<tr>
<td>$\lambda_{2,1}$</td>
<td>second largest eigenvalue of tensor-2 representing the largest diffusivity perpendicular to fiber-2</td>
<td>0.4</td>
</tr>
<tr>
<td>$\lambda_{2,2}$</td>
<td>smallest eigenvalue of tensor-2</td>
<td>0.38</td>
</tr>
<tr>
<td>$\alpha_{1,2}$</td>
<td>angles defining the plane in which the fibers cross</td>
<td>Random</td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>mean orientation of the crossing fibers in the plane of the crossing</td>
<td>0.8 $\pi$</td>
</tr>
<tr>
<td>$f_1$</td>
<td>volume fraction of tensor-1</td>
<td>Variable</td>
</tr>
<tr>
<td>$f_{iso}$</td>
<td>volume fraction of isotropic component</td>
<td>0.1</td>
</tr>
<tr>
<td>$D_0$</td>
<td>diffusivity constant</td>
<td>3.0</td>
</tr>
<tr>
<td>$S_0$</td>
<td>signal without diffusion weighting</td>
<td>250</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>standard deviation of the noise</td>
<td>10</td>
</tr>
</tbody>
</table>

In the second part of this section we demonstrate the potential of the proposed framework for some neuroimaging applications. Therefore, we evaluated the performance on varying types of brain datasets to verify whether a reliable estimation could be achieved. Initially, we applied ARD and MLE to the genu of the corpus callosum (CC) representing a single fiber region enclosed at both ends by a crossing with the corticospinal tract (CST). Subsequently, we compared the volume-fraction weighted orientations obtained by ARD and MLE and test the reproducibility of ARD based on six randomly selected subjects. Finally, we show a neuroimaging application
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

of our proposed framework, i.e. automatic estimation of diffusion properties.

Three different datasets, acquired with different acquisition protocols, were adopted to explore the two methods. Dataset A concerned diffusion data from seven subjects of the Human Connectome Project (HCP) [34]. The relevant acquisition parameters of this dataset were: three b-values 1000, 2000 and 3000 s/mm$^2$, 288 gradient directions, TE/TR 89.5/5520 ms, voxel size 1.25×1.25×1.25 mm$^3$. Dataset B was acquired from a control subject (see also [3]). The acquisition parameters: two b-values 1000 and 3000 s/mm$^2$, 185 gradient directions, TE/TR 84/3800 ms, voxel size 1.7×1.7×2.2 mm$^3$. Dataset C consisted of data from 24 healthy controls from an ongoing DTI study into the effects of HIV on the brain [35]. The acquisition parameters were: two b-values 1000 and 2000 s/mm$^2$, 130 gradient directions, voxel size 2.0×2.0×2.0 mm$^3$. The SNR for each of the three datasets was found to be higher or equal than 20. This was determined by fitting a single tensor model to a selected region of the CC, after which we took the ratio between $S_{v,0}$ and the model residual as the estimated SNR.

2.3.1 ARD versus MLE: Simulation Experiment

The volume fraction of the constituting compartments of a dual tensor model is a crucial parameter for the modeling of simple and complex fiber geometries. Therefore, we evaluated the performance of the methods as the volume fractions of the two compartments were varied. To assess the robustness for variations in volume fraction, for each volume fraction and divergences of 90 degrees and 45 degrees we generated 100 realizations with the parameters given in Table 2.2 and $\alpha_4$. For each realization $\hat{\theta}_{ARD}$ and $\hat{\theta}_{MLE}$ were computed.
Figure 2.1 Results of dual tensor model estimation by ARD (red) and MLE (blue) on 100 noisy realizations (SNR 25) of two crossing fibers (divergence: 90 degrees in a-c and 45 degrees in d-f) as a function of volume fraction. The boxplots show: a) and d) volume fraction of the first compartment; b) and e) fractional anisotropy (FA) of the first component; c) and f) FA of the second compartment. The simulated model listed in Table 2 and acquisition parameters accords with Dataset B. The boxes display the median and 25th, respectively 75th percentiles of the data distribution; whiskers extend to 1.5 times the interquartile range; values outside these ranges are indicated as individual points.
Figure 2.1 shows boxplots depicting the results of dual tensor estimation by ARD (red) and MLE (blue). Since the estimation procedure assigns a random label to the first and second tensor, we need to assign the two estimated tensors to the corresponding ground truth compartment. The estimated tensors are sorted by tensor similarity based on the Frobenius norm: the tensor with the smallest Frobenius norm with respect to the ground truth of compartment 1 received the label 1. Figure 2.1(a-c) show the results for the 90° crossing. The estimation volume fraction by ARD nearly ideally correlates with the ground truth over the entire range of volume fractions, both for single fiber \( f_1 = 0 \lor f_1 = 0.9 \) and crossing fiber configurations \( f_1 \in (0, 0.9) \), see Figure 2.1(a). Clearly, MLE yields a random \( f_1 \) as the true volume fraction approaches that of a single fiber configuration.

Similarly, Figure 2.1(b) shows the estimated \( FA_1 \) against the true \( f_1 \) and Figure 2.1(c) shows the estimated \( FA_2 \) against the true \( f_2 \). Figure 2.1(b) shows that the median \( FA_1 \) estimation is almost identical for both estimation methods in crossing fiber configurations, irrespective of the volume fraction. Figure 2.1(c) shows that the estimated \( FA_2 \) converges to the true value as the true \( f_2 \) increases. The \( FA \) estimation in single fibers \( f_1 = 0.9 \lor f_2 = 0.9 \) appears equally unbiased for both methods, but a considerably larger spread is encountered with MLE than with ARD. Note that the estimated \( FA \)'s with \( f_1 = 0 \lor f_2 = 0 \) essentially represent degenerate measurements and are therefore not shown in the graphs. In the absence of a ground truth this can be detected with ARD as the corresponding volume fraction is automatically forced to (near) zero (see Figure 2.1 (a)). MLE does not offer such a mechanism, which might lead to a fictitious fiber compartment.

Figure 2.1(d-f) show the results for the 45° crossing. This figure shows similar trends as obtained for the 90° crossing in Figure 2.1(a-c), albeit with a larger spread.

The performance of the two methods in single fiber regions is further corroborated in Figure 2.2. It shows the dual tensor model estimation by ARD (red) and MLE (blue) on a single fiber with a small isotropic compartment while only the single fiber’s \( FA \) is varied.
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

Figure 2.2 Results of dual tensor model estimation by ARD (red) and MLE (blue) on 100 noisy realizations (SNR 25) of a single fiber and an isotropic compartment as a function of FA. a) Volume fraction of the first compartment; b) Fractional Anisotropy (FA) of the first component. The simulated model are listed in Table 2 and acquisition parameters accords with Dataset B.

Figure 2.2(a) shows that the estimated $f_1$ with ARD improves with increasing FA and approximates the ground truth. MLE essentially yields a random estimate of $f_1$ irrespective of the actual FA. Figure 2.2(b) confirms that the estimation of $FA_1$ remains largely unbiased with both methods. Clearly, the spread in the estimated $FA_1$ with ML is much larger than with ARD.

Summarizing, the graphs demonstrate two things: 1) ARD facilitates accurate estimation of volume fractions especially with increasingly unbalanced real volume contributions, in which case MLE grossly fails; 2) the estimation of FA by ARD shows a much narrower distribution than by MLE, especially in single fibers and for lower actual FA’s.

2.3.2 ARD versus MLE: brain imaging

To demonstrate the performance of ARD and MLE on brain data, one subject was randomly selected from each of the three aforementioned datasets (A, B, and C). Figure 2.3 shows approximately corresponding coronal-views of regions of interest where the Corpus Callosum (CC) crosses the Corticospinal Tract (CST). Specifically,
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on
Jeffreys prior for diffusion MRI

it contains the ARD (left) and MLE (middle) and FSL’s ARD [8] (ARD-FSL, right)
estimates in this slice. From top to bottom are shown data from respectively datasets A,
B, and C. A single fiber region, i.e. the central part of CC, is indicated by the yellow
ellipse and the region where CC and CST cross by the green circle. For ARD and
MLE, the line segments visualize the orientations of the largest eigenvectors of the
underlying tensors and the length of each line segment is scaled by the volume fraction.
For FSL’s ARD1 the line segments visualize stick orientations, also scaled with the
estimated volume fractions. For FSL’s ARD the line segments visualize stick
orientations, also scaled with the estimated volume fractions. Therefore, we applied the
command “bedpostX” in FSL with set parameters Fibres=2, weight=1, Burn in = 1000,
and switching to multi-shell model and rician noise (leaving model noise floor off) for
a comparable outcome.

Clearly, ARD forces the volume fraction of one tensor compartment nearly to zero in
the single fiber region of the CC. Evidently, MLE yields an erratic outcome regarding
both volume fraction and fiber orientation in the same region. Furthermore, notice
that FSL-ARD often returns two fiber orientations in this region. In crossings like the
region where CC crosses with CST, ARD, MLE and FSL-ARD yield a similar
outcome regarding fiber orientations. These trends can be observed for each of the
three datasets.

The proposed framework was subsequently applied to the same ROI in three other,
arbitrarily selected HCP subjects. The outcome, depicted in Figure 2.4 confirms the
previous findings. Furthermore, it demonstrates that despite the inter-subject
anatomical variations, our ARD is stable and yields consistent results across different
individuals.

---

1 applied by means of the command “bedpostX” in FSL
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

Figure 2.3 Results of our ARD (a, d, g), MLE (b, e, h) and FSL’s ARD using the ball & stick model (c, f, i) in a region where the corpus callosum (CC) crosses the corticospinal tract (CST) in three randomly selected subjects; (a,b,c) Dataset A, i.e. HCP dataset; (d, e, f) Dataset B; (g, h, i) Dataset C. The length of a line segment is proportional to the corresponding volume fractions; the colors indicate the orientation of fibers: medio-lateral (red), anterior-posterior (green), and superior-inferior (blue).
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

2.3.3 ARD versus MLE: Fractional Anisotropy along the CC

Figure 2.5 shows the estimated FA along the Genu of the Corpus Callosum (GCC) in one subject from dataset B. The GCC is indicated by the yellow trajectory superimposed on the red-colored structure of the inset. Notice that the center of the GCC is a single fiber bundle but there is a crossing with the CST at its lateral sides. The FA along the tract was estimated by the proposed ARD framework (red) and MLE (green). We had to select the tensor compartment that corresponds to the GCC since the labels assigned by MLE and ARD are random. To solve this, the FA belonging to CC was selected based on “front evolution” [36]. In front evolution, the
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

estimated tensor of one compartment is randomly chosen as the reference tensor. Then, the tensor of a neighborhood voxel with the smallest Frobenius norm to the reference tensor receives the same label. After processing all neighbors of the current front as such, these neighbors become the new reference tensors for the next iteration. The green (MLE) and red (ARD) areas indicate the uncertainty in the estimated value as quantified by the square root of the CRLB.

Confirming the above findings, estimation by ARD yields a rather constant FA with small variance. In contrast, MLE yields FA values with a larger variance, particularly in the central part of GCC. We attribute this to overfitting.

Figure 2.5 Estimated FA by ARD (red) and MLE (green) as a function of position along the genu of the Corpus Callosum (GCC trajectory is indicated by the yellow arrow on the inset). The estimated uncertainty (indicated by the colored background) was calculated by +/- the square root of the Cramer-Rao lower bound. The trajectory along the GCC consists of a single fiber region in the middle (30-70) and enclosed by crossing fibers region (with CST) on both ends. The figure displays the FA of the estimated tensor compartments assigned to the GCC by front evolution for both ARD and MLE.
2.3.4 ARD versus MLE: Dual-tensor FA and volume-fraction maps

FA as well as volume-fraction maps have been used to detect white matter changes [36] [37]. Here, we will display FA and volume fraction maps generated by ARD and MLE and point out the differences.

Figure 2.6 Color-coded output displaying the FA (red channel) and the corresponding volume fraction (green channel) for ARD and MLE. The tensor compartments were classified into first and second by front evolution. FA of first tensor and its corresponding volume fraction by ARD (a) and by MLE (b); FA of the second tensor and its corresponding volume fraction by ARD (c) and by MLE (d).
Specifically, Figure 2.6 shows color-coded RGB maps encoding FA in the red channel and the corresponding volume fraction in the green channel. Left images show the outcome by ARD, right images by MLE; top images reflect the first tensor, bottom images denote the second tensor. The ordering of the tensors was performed by front evolution [36].

Regarding the ARD outcomes, one can observe that in Figure 2.6(a) grey matter is displayed in green reflecting large $f_1$ and small $FA_1$; in Figure 2.6(c) grey matter is dark, representing both small $f_2$ and small $FA_2$. Furthermore, in Figure 2.6(a) single fiber regions, particularly the central part of CC, are yellowish due to simultaneously large $f_1$ and $FA_1$; in Figure 2.6(c) the corresponding regions are reddish because of a small $f_2$. In both Figure 2.6(a) and (c) crossing fiber regions are yellowish reflecting large volume fractions and FA’s.

In contrast, MLE does not specifically force the volume fraction of one tensor to zero in single fiber and gray matter regions. Therefore, all white matter regions in Figure 2.6(b) and (d) are yellowish reflecting large volume fractions and FA’s. At the same time gray matter regions in Figure 2.6(b) and (d) are greenish due to the small FA for a substantial volume fraction.

In general, Figure 2.6 confirms that MLE estimation is not able to automatically cope with single fiber regions whereas ARD suppresses the volume fraction of one tensor in such areas. Practically, FA2 should be ignored when $f_2$ is very small, since in such cases it has no meaning.

### 2.3.5 TBSS based on dual tensor FA and volume-fraction maps

For a statistical analysis of the FA and volume fractions with age, we included the healthy controls from dataset C [35]. The subjects aged between 45 and 50 (12 subjects, mean-age: 46.2, standard deviation 1.49) and those aged between 65 and 75 (12 subjects, mean-age: 68.2, standard deviation 3.72) were selected from the full control group. All data was registered to the MNI152 standard space using FNIRT [38]. We used the version of FNIRT that is implemented in FSL (version 5.0.7). Subsequently, differences between the two age groups were analyzed by means of the classical TBSS technique, i.e. single tensor analysis, [37] as well as the extended TBSS method for the dual tensor models [36]. Compared to the classical TBSS method, the
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

Extended TBSS technique employs ‘front evolution’ to avoid swapping of the two anisotropic components between the different images. Extended TBSS was used to analyze differences in the dual tensor volume fractions (comparable to [36]) as well as differences in the dual tensor FA maps between the two age groups. Importantly, our approach facilitates such a separate analysis of tensor volume and shape. Notice that the volume fraction used in [36] is a different variable, representing the stick strength in a ball-and-stick model.

Figure 2.7 shows a typical axial slice containing multiple areas with crossing fibers and regions with just a single fiber. Figure 2.7(a) shows the classical, single tensor TBSS analysis, Figure 2.7(b,d) the extended TBSS analysis of the dual tensor FA maps, and Figure 2.7(c,e) the extended TBSS analysis of the dual tensor volume fraction maps. The dual tensor estimations in Figure 2.7(b,c) were obtained by ARD, those in Figure 2.7(d,e) by means of MLE. The red-yellow colored regions in Figure 2.7(a-e) identify regions where the differences are significant.

It can be observed that in single fiber regions (indicated by the green ellipse) the extended TBSS analysis based on MLE (Figure 2.7(d,e)) yields smaller regions with significant differences compared to the classical approach (Figure 2.7(a)). We attribute this to the large variability in such regions that we already observed with MLE e.g. in Figure 3.2. Instead, the ARD method (Figure 2.7(b,c)) finds similar or slightly more significant differences compare to the classical method especially when the changes in volume fraction and FA are considered together. This signifies that the superfluous parameters are effectively eliminated by ARD in single fiber regions. Furthermore, we found the expected similarities regarding detected differences in crossing regions (e.g. the red circles) between the MLE and ARD . This indicates that Jeffreys prior does not affect the MLE outcome in such regions. The extended TBSS analyses based on ARD expands regions with significant differences compared to the classical approach. All significant differences are negative, i.e. reduced FA and volume fraction with increasing age. This outcome confirms the finding that significant age-related white matter atrophy was found in the Corpus Callosum [28].
Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI

2.4 Discussion and Conclusion

We developed a new framework for estimating the parameters of a constrained dual tensor model in diffusion-weighted MRI. It automatically determines to what extend the diffusion in a voxel should be modeled by one or two rank-2 tensors. In essence, the complexity of the model is implicitly inferred with ARD in a Bayesian probabilistic manner. An initial guess of the diffusion model is obtained by fitting a dual tensor model to the data with MLE. Subsequently, a new automated relevance determination method assesses whether two tensors are ‘mandatory’ to model the data. If this is not the case, the volume fraction of the superfluous tensor automatically reduces to (nearly) zero.

The proposed framework extends previous work by Behrens et al and Jbabdi et al [36] [37]. A crucial difference is that we employ a rank-2 tensor model, whereas the previous works concerned ball-and-stick models. As such, we aim to recover the full diffusion shape. Furthermore, an important novelty of our work is that Jeffreys non-informative prior is introduced into ARD, which is an alternative for previously used informative priors. It facilitates accurate and precise estimation of the volume fractions as well as the diffusion properties with a DTM in single fiber and crossing fiber regions. Jeffreys prior is based on the Fisher’s information matrix which accounts for properties of the data acquisition, such as diffusion weighting b-value, the gradient directions and the effective SNR. Therefore we call this method data-adaptive.
We demonstrated that both in single fibers as well as in fiber crossings the configuration inferred by our method corresponds to the expected neuro-anatomy [40]. The proposed framework has been compared with direct MLE of the same dual tensor model. Several differences between the proposed framework and MLE were observed. In regions that were considered to contain just a single fiber, MLE typically inferred a large volume fraction for both tensor components (see Figure 2.3). Here, the proposed framework yielded a single tensor representation by diminishing the volume fraction of the second tensor component. Furthermore, the FA estimated by MLE showed much more variation than the FA estimated by ARD in such a region (Figure 2.5). In regions that were considered to contain crossing fibers, the results of MLE and ARD were similar.

There are a few limitations of our method. Firstly, we assume a mono-exponential decay along the eigenvectors of the three compartments up to $b=3000$ s/mm$^2$. Measuring at higher $b$-values will certainly introduce sensitivity to different compartments such as the myelin sheet [41] with the associated restricted and hindered diffusion [5]. In the latter case, the Gaussian diffusion assumption is no longer valid. However, investigating such non-Gaussian diffusion is beyond the scope of this work. As such, we follow [3] and [42]. Secondly, recent studies reported the presence of a three-way crossing of fiber bundles [14] [43]. In our framework a dual-tensor model is employed to characterize voxels encompassing crossing fibers. The reason for limiting the number of anisotropic components to two is the limited SNR in our HARDI data. Notice that whereas [14] and [43] only recover the fiber orientation, we aim to reconstruct the full diffusion shape, which requires a higher SNR. In [3] we showed that estimating a dual rank-2 tensor model already requires HARDI at two $b$-values, data of sufficient SNR, and some model restrictions. The latter is needed to ensure stability as the number of model parameters may approach or even surpass the number of degrees of freedom present in the data. Therefore, fitting a triple rank-2 tensor model to voxels with a three-way-crossing will be even more challenging [13]. Developing methods for estimation of the diffusion properties in three-way-crossing fiber bundles will remain an important challenge for future research.
Diffusion imaging may reveal several aspects to white matter integrity: (I) locations of alterations; (II) which fiber tract is affected; (III) the exact change in diffusion. Previously, many solutions were already proposed for the first two aspects. Our work focused on all three aspects. Particularly, our framework may aid a more accurate characterization of the diffusion shape.

References


Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI


Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI


Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI


Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI


Chapter 2. Reliable dual tensor modelling of simple and complex fiber geometries based on Jeffreys prior for diffusion MRI
3 Estimation of diffusion properties in three-way fiber crossings without overfitting

Diffusion-weighted magnetic resonance imaging permits assessment of the structural integrity of the brain’s white matter. This requires unbiased and precise quantification of diffusion properties. We aim to estimate such properties in simple and complex fiber geometries up to three-way fiber crossings using rank-2 tensor model selection.

A maximum a-posteriori (MAP) estimator is employed to determine the parameters of a constrained triple tensor model. A prior is imposed on the parameters to avoid the degeneracy of the model estimation. This prior maximizes the divergence between the three tensor’s principal orientations. A new model selection approach quantifies the extent to which the candidate models are appropriate, i.e. a single-, dual- or triple-tensor model. The model selection precludes overfitting to the data. It is based on the goodness of fit and information complexity measured by the total Kullback-Leibler divergence (ICOMP-TKLD). The proposed framework is compared to maximum likelihood estimation on phantom data of three-way fiber crossings. It is also compared to the ball-and-stick approach from the FMRIB Software Library (FSL) on experimental data.

The spread in the estimated parameters reduces significantly due to the prior. The fractional anisotropy (FA) could be precisely estimated with MAP down to an angle of approximately 40° between the three fibers. Furthermore, volume fractions between 0.2 and 0.8 could be reliably estimated. The configurations inferred by our method corresponded to the anticipated neuro-anatomy both in single fibers and in three-way fiber crossings. The main difference with FSL was in single fiber regions. Here, ICOMP-TKLD predominantly inferred a single fiber configuration, as preferred, whereas FSL mostly selected dual or triple order ball-and-stick models.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

The prior of our MAP estimator enhances the precision of the parameter estimation, without introducing a bias. Additionally, our model selection effectively balances the trade-off between the goodness of fit and information complexity. The proposed framework can enhance the sensitivity of statistical analysis of diffusion tensor MRI.
3.1 Introduction

Diffusion-weighted magnetic resonance imaging (DW-MRI) is a unique tool for assessing the integrity of white matter (WM) in vivo [1] [2] [3]. Essentially, it provides a way to measure the 3D diffusion profile of water in the brain. Classically, a single rank-2 diffusion tensor, estimated from a series of DW images, is used to model the local diffusion of water molecules: hence the name Diffusion Tensor Imaging (DTI) [4] [5]. Typically, properties are derived from such a tensor by invariants like the fractional anisotropy (FA) [6].

However, it is widely known that a single rank-2 tensor as applied in classical DTI cannot accurately characterize complex fiber structures such as crossings and bifurcations [4]. There is an ongoing debate on how to best characterize the orientation and diffusion properties in these configurations [7] [8]. This is relevant since accurate orientation and diffusion estimation has led to promising applications in brain mapping [9] and detection of pathologies [10]. Many approaches such as the ball-and-stick model (essentially a rank-1 tensor model) [7], spherical deconvolution [11] [12], diffusion spectrum imaging [13], the composite hindered and restricted model of diffusion (CHARMED) [14] [15] and diffusion kurtosis imaging (DKI) [16] enable a more sophisticated characterization of the diffusion in order to achieve a better reconstruction, particularly of the principal fiber orientations. A potential advantage of the multi-tensor model is in that it asserts a Gaussian diffusion partitioned over multiple compartments. As such, an intuitive physical interpretation can be attributed to the parameters of the dual tensor model (as with the conventional, single tensor model). Furthermore, notice also that the full diffusion shape is captured by such tensors and not only the diffusion along single directions (as in the ball-and-stick model).

Recently, the existence of three-way crossings has been reported [17]. Although the full diffusion shape of two-way fiber crossings has been characterized by a single rank-4 tensor model [18] [19] and a dual rank-2 tensor model [8] [20], there is no literature doing the same for voxels with triple-way crossing fibers as revealed in [17] [11].
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

Essentially, two issues hamper estimating the full diffusion shape of three-way fiber crossings by means of a high-rank tensor model:

I. Unbiased estimation of a triple tensor model requires high-quality DW-MRI data and proper constraints on the parameters to be estimated, i.e. the degrees of freedom (DOF) embedded in the data should support the large number of parameters to be estimated;

II. Fitting dual or triple tensor models to voxels comprising a simpler fiber structure will inevitably cause unreliable estimates as a result of overfitting. Therefore, a proper model should be selected based on the available data from a voxel.

Several model selection methods were introduced in the field of DW-MRI. Automatic relevance determination (ARD) aims to eliminate the redundant parameters in a complex model, such that a simplified model yields a better description of the data [21]. Behrens [22] adopted ARD for assessing the appropriate number of fiber orientations in each voxel for fiber tracking. ARD methods assume a prior distribution for the model parameters. A Gaussian distribution is a straightforward choice for a prior [21]. Previous ARD approaches [22] [23] [24] involved marginalization (integration) over the hyper-parameters to get a prior for each parameter separately. Such a prior is likely to be suboptimal for individual voxels since potential correlations between parameters are ignored. Alternatively, model selection techniques related to constrained spherical deconvolution (CSD) [25], the Bayesian information criterior (BIC) [26], and the generalization-error [27] were used. A potential drawback of model selection in CSD is that the model selection criterion is not implicitly defined, so that it requires tuning of a threshold. A limitation of the BIC is that it is not determined by the data itself, but through non-estimated factors such as the number of parameters and the sample size. Finally, a restriction of the generalization-error method [27] is that it is a non-local model selection technique. To our knowledge,
model selection methods have not been studied especially for rank-2 triple tensor models.

To address the aforementioned challenges, we need (1) to extend the range of rank-2 tensor models to triple-tensor estimation and simultaneously achieve a high accuracy and precision by incorporating suitable priors and imposing appropriate constraints; and (2) to select the right model for a given set of diffusion measurements per voxel.

This chapter introduces a maximum a-posteriori estimator to characterize diffusion profiles particularly in voxels comprising three-way crossing fiber bundles. The proposed model selection method selects from the single-, dual- and triple-tensor models the most suitable representation given the data. This data-adaptive model selection framework, which compares the competing models by a measure of information complexity based on the total Kullback-Leibler divergence, dubbed ICOMP-TKLD (see below), is an explicit model selection approach. The whole framework will be validated on DW-MRI data from the Human Connectome Project (HCP) [28] [29] and data from an ongoing study into the effects of HIV on the brain [30]. Moreover, the performance of the proposed framework will be compared to a state-of-the-art approach [7] [31].

A. Significance

The aim of the proposed framework is to achieve unbiased estimation of diffusion properties in white matter structures throughout the brain. In particular, our work targets to improve neuroimaging research by providing the estimation of diffusion properties along fiber bundles, which cross multiple other structures along theirs tract. Specifically, the properties extracted from the fitted diffusion model may facilitate a more sensitive detection of WM changes, e.g. related to aging or disease processes.

3.2 Methods

We will first present the triple-tensor model and its constrained parameterization for estimating diffusion profiles in three-way fiber crossings. Second, we will introduce a
prior, which is required to solve this ill-posed inverse problem. Third, we describe the posterior probability function that is optimized to obtain an unbiased fit of the triple-tensor model. Fourth, we introduce the structure-adaptive model selection method that balances the goodness of fit and a new measure for information complexity named ICOMP-TKLD.

3.2.1 Triple-tensor Model

We assume that the diffusion in a single fiber bundle is mono-exponential and can be described by an anisotropic Gaussian profile. Accordingly, the measured DW signal \( S_j \) in voxels comprising a three-way fiber crossing is modeled by a weighted sum of three anisotropic, Gaussian basis functions:

\[
S_j = S_0 \left( \sum_{i=1,2,3} f_i \exp(-b_j g_j^T D_i g_j) \right). \tag{3.1}
\]

In (3.1) \( S_j \) is the signal predicted by the triple-tensor model; \( S_0 \) is the signal without diffusion weighting; \( b_j \) is the amount of diffusion weighting in gradient direction \( g_j \); \( b_j \) is selected from a vector \( b \) whose length \( n_b \) equals the number of unique \( b \)-values; \( D_1, D_2 \) and \( D_3 \) are three positive definite rank-2 tensors (i.e. \( 3 \times 3 \) matrices), which independently describe the diffusion profile of a fiber; \( f_i \) quantifies the volume fraction of each fiber bundle, with the constraint: \( f_1 + f_2 + f_3 = 1 \). For the single and dual tensor model we limit ourselves to one respectively two anisotropic compartments.

3.2.2 Constraints

In this chapter, the tensor \( D_i \) is parameterized by a diagonal eigenvalue-matrix \( E_i \) and eigenvector-matrix \( V_i : D_i = V_i^T E_i V_i \). The diffusion perpendicular to the fiber orientation is assumed to be isotropic. Accordingly, we model the diffusivities by an axial and a radial diffusivity, denoted by \( \lambda_{||} \) and \( \lambda_{\perp} \) respectively. Now \( E_i \) yields:

\[
E_i = \text{diag}\left( \lambda_{||}, \lambda_{\perp}, \lambda_{\perp} \right) \text{ with } i \in \{1,2,3\}. \tag{3.2}
\]

As tensor \( D_i \) must be symmetric positive-definite, non-negativity of the diffusivities is enforced by adopting exponential mappings \( \exp(\cdot) \) as in [8]; additionally, to impose that diffusion along a fiber is faster than perpendicular to it, i.e. \( \lambda_{||} > \lambda_{\perp} \), we further constrain the diffusivity parameters by
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

\[ \lambda_{ii} = c_{ii} \lambda_{\text{free}}, \quad (3.3) \]

with
\[ \lambda_{i\perp} = c_{i\perp} \lambda_{ii}, \quad (3.4) \]
where \( \lambda_{\text{free}} \) represents the approximated diffusivity of free water at body temperature and the coefficients \( c_{ii} \) and \( c_{i\perp} \) are positive fractions ranging between 0 and 1.

Additionally, we expect that the variance of the estimated parameters is significantly reduced by constraining the axial diffusivity of all tensor compartments to be equal, as was previously reported for the dual-tensor case [8], i.e. \( \lambda_{ii} = \lambda_{2ii} = \lambda_{3ii} \) so that \( c_{11} = c_{12} = c_{13} \). The benefit of this constraint will be demonstrated in the experiment section.

It is assumed that the orientation of each fiber is aligned with the first principal eigenvector of the corresponding tensor in model (3.1), i.e. \( \mathbf{v}_{1,i} \) in the matrix \( \mathbf{V} = [\mathbf{v}_{1,i}, \mathbf{v}_{2,i}, \mathbf{v}_{3,i}] \). We use spherical coordinates to represent \( \mathbf{v}_{1,i} \) uniquely by its zenith angle \( \theta_{1,i} \) and azimuth angle \( \phi_{1,i} \):
\[ \mathbf{v}_{1,i} = [\cos(\phi_{1,i}) \sin(\theta_{1,i}) \sin(\phi_{1,i}) \sin(\theta_{1,i}) \cos(\theta_{1,i})]^T. \quad (3.5) \]
The other two eigenvectors, i.e. \( \mathbf{v}_{2,i} \) and \( \mathbf{v}_{3,i} \), are defined in the plane perpendicular to \( \mathbf{v}_{1,i} \). Volume fractions of tensor compartments are defined in the range \( 0 < f_i < 1 \) by means of an error function \( \text{erf}(\cdot) : \text{erf}(x) = (2 / \sqrt{\pi}) \int_0^x \exp(-t^2) \, dt \). This mapping of \( f_i, i \in \{1, 2\} \) is achieved by
\[ f_{(1,2)} = \frac{1}{2} \left( \text{erf}(f_{(1,2)}) + 1 \right). \quad (3.6) \]

In summary, the \( i^{th} \)-fiber in a three-way crossing will be characterized the parameters \( \{\lambda_{ii}, \lambda_{i\perp}, \theta_{1,i}, \phi_{1,i}, f_i\} \). Employing the aforementioned parameterization, the complete DW-signal is described by a 16-dimensional parameter vector \( \Theta \):
\[ \Theta = \left\{ S_0, c_{11}, c_{21}, c_{12}, c_{22}, c_{13}, c_{23}, \theta_1, \theta_2, \theta_3, \phi_1, \phi_2, \phi_3, f_1, f_2, f_3 \right\}. \quad (3.7) \]

However, with the constraint that \( c_{11} = c_{12} = c_{13} \) (i.e. \( \lambda_{1ii} = \lambda_{2ii} = \lambda_{3ii} \)) and \( f_s = 1 - f_1 - f_2 \), 13 DOFs need to be estimated.
3.2.3 Priors on the parameters of the triple tensor model

The number of parameters to be estimated in the triple-tensor model is large, and correlations between some of these parameters lead to high variances in the estimated parameters. Therefore, imposing a prior on a subset of the parameters is needed to facilitate precise estimation. In [8], we demonstrated that the performance of estimation got worse as the angle between two tensors decreased below 45 degrees. Accordingly, the prior applied by us is given by

\[
p(v_{1,1}, v_{1,2}, v_{1,3}) = \prod_{i,j=1, i>j}^{3} \exp \left( - \frac{1}{2} \left( v_{1,i} \cdot v_{1,j} \right)^2 \right), \quad (3.8)
\]

in which \( v_{1,i} \cdot v_{1,j} \) represents the inner product between the principal eigenvectors of \( D_i \) and \( D_j \). The prior (3.8) essentially states that a relatively large divergence between any pair of fibers is more likely than a configuration in which they make a small angle. This is reasonable, since complex models with a small divergence between any pair of fibers will not survive the model selection step anyway (see below). No restrictions will be imposed on the other parameters, i.e. these priors are given by uniform distributions: \( p(S_0) = p(\lambda_{i||}) = p(\lambda_{i\perp}) = U(0, \infty) \), \( f_j = U(0,1) \). The complete multivariate prior for parameter vector \( \Theta \) is given by

\[
p(\Theta) = p(v_{1,1}, v_{1,2}, v_{1,3}) p(f_j) p(\Delta \lambda_{i||}) p(\Delta \lambda_{i\perp}) p(S_0). \quad (3.9)
\]

3.2.4 Maximum a-posteriori (MAP) estimator

The probability density function (PDF) of a measured DWI signal \( \tilde{S}_j \) is a non-central \( \chi \)-distribution [32]:

\[
p(\tilde{S}_j | S_j, \sigma, n) = \frac{S_j}{\sigma^2} \left( \frac{\tilde{S}_j}{S_j} \right)^{n-1} \exp \left( - \frac{\tilde{S}_j^2 + S_j^2}{2\sigma^2} \right) I_{n-1} \left( \frac{\tilde{S}_j S_j}{\sigma^2} \right), \quad (3.10)
\]

where \( n \) determines the rank of the \( \chi \)-distribution. The value of \( n \) depends on the DWI reconstruction protocol: for ‘GRAPPA’ [33], \( n \) is assigned by the number of receiver coils involved, whereas for ‘SENSE’ [34], , i.e. the non-central \( \chi \)-distribution degenerates to the Rician distribution. \( I_{n-1}(\cdot) \) denotes the \((n-1)^{th}\) order modified Bessel function of the first kind.
Likelihood function. The DWIs are independent; therefore the joint PDF \( p(\tilde{S} | \Theta) \) of the signal profile \( \tilde{S} \) is given by the product of the marginal distributions, i.e. the likelihood function becomes

\[
p(\tilde{S} | \Theta) = \prod_{j=1}^{N_s} p(\tilde{S}_j | S_j, \sigma). \tag{3.11}
\]

Posterior distribution: Combining the likelihood function (3.11) and the prior (3.9) in Bayes' rule yields the posterior distribution of parameter vector \( \Theta \)

\[
p(\Theta | \tilde{S}) \propto p(\tilde{S} | \Theta) p(\Theta). \tag{3.12}
\]

The parameter values can be estimated by maximizing (3.12), yielding the maximum a posteriori (MAP) estimate

\[
\hat{\Theta} = \arg \max_{\Theta} (\ln(p(\Theta | \tilde{S}))). \tag{3.13}
\]

We use Levenberg-Marquardt optimization to solve (3.13) [35]. Therefore, the triple-tensor model is initialized multiple times based on an initial single-tensor fit:

\[
\lambda_{\parallel} = |\lambda_{s} + \lambda_{i} \cdot N(0,1)| \quad \text{(i.e. } C_{i} = \lambda_{\text{free}} / \lambda_{\parallel}) \quad \text{and} \quad \lambda_{\perp} = |\lambda_{2s} + \lambda_{2i} \cdot N(0,1)| \quad \text{(i.e. } C_{2i} = \lambda_{\parallel} / \lambda_{\perp});
\]

furthermore, the zenith and azimuth angles are initialized randomly and \( f_i = 1/3 \). Ultimately, we select the parameters corresponding to the maximum a-posteriori probability of all these fits.

Unfortunately, simply applying the triple-tensor model to all voxels yields massive over-fitting, since only a small fraction of them comprises triple-way fiber crossings. Therefore, we fit single-, dual-, and triple-tensor models followed by model selection based on the total Kullback-Leibler divergence (ICOMP-TKLD) to select the ‘best’ model for each voxel.

3.2.5 Model Selection: ICOMP-TKLD

Model selection methods evaluate the appropriateness of competing models. Typically, these methods assess each model by balancing a measure for goodness-of-fit (e.g. residual energy) with a measure for model complexity. First, we briefly review the standardly used ICOMP criterion [36], which is based on the conventional Kullback-Leibler divergence (KLD), and show its shortcomings in evaluating nested rank-2 tensor models. Second, we present our ICOMP-TKLD technique, which employs the
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

**total Kullback-Leibler divergence (TKLD)** [37] to assess the information complexity of each model, and show that this solves the problems associated with t KLD.

**Information complexity (ICOMP).** ICOMP quantifies a model’s complexity not simply by the number of parameter in the model such as the Akaike information criterion (AIC) [38] (see Appendix for details), but as the degree of interdependence among the parameters [36]. Therefore, it employs the Kullback-Leibler divergence [39] between the joint PDF and the product of the marginal PDF’s of the parameters to measure the model complexity. For a model in which the parameters are totally independent, proper scaling of the parameters can be achieved such that the complexity, i.e. the KLD, is zero. This accords with one’s intuition that a nested model may be expanded as long as parameters are not ‘superfluous’ with respect to the data. ICOMP quantifies an overall criterion for a model in two terms: the first denoting the residual energy that remains after ML or MAP estimation and the second expressing the complexity:

\[
ICOMP(\widehat{\Theta}_i) = -2\ln(L(\widehat{\Theta}_i, | \widehat{S})) + 2C_i(I^{-1}(\widehat{\Theta}_i)),
\]

where

\[
C_i(I^{-1}) = \max_T (T'C_0(I^{-1})T) = \frac{k_i}{2} \log\left(\frac{tr(I^{-1})}{k_i}\right) - \frac{1}{2} \log\left(||I^{-1}||\right),
\]

with \(C_0(I^{-1})\) being the KLD defined as

\[
C_0(I^{-1}) = \frac{1}{2} \sum_{j=1}^{k_i} \log(\varepsilon_j^2) - \frac{1}{2} \log\left(||I^{-1}||\right),
\]

and the Fisher information matrix given by

\[
I = -\mathbb{E}_S \left\{ \frac{\partial^2 \ln(p(S | \widehat{\Theta}_i, \sigma))}{\partial \widehat{\Theta}_i \partial \widehat{\Theta}_i^T} \right\}.
\]

Here, \(\widehat{\Theta}_i (i = 1|2|3)\) denotes the estimated parameter vector of respectively the single-, dual- and triple-tensor models; \(I^{-1}\) represents the inverse of the Fisher information matrix. Furthermore, \(\varepsilon_j^2\) denotes the \(j\)-th diagonal element of \(I^{-1}\) and \(k_i\) the number of parameters to be estimated in model \(i\). However, the KLD \(C_0(I^{-1})\) is a coordinate dependent measure, so that mere changes in the parameterization (i.e. coordinate
system) of a model already yield a different KLD. To overcome this issue, eq. (3.15) calculates $C_i(I^{-1})$ by maximizing the KLD over all possible orthogonal transforms (denoted by $T$). As such, the measure of complexity becomes coordinate independent. The closed-form expression for $C_i(\cdot)$ is the theoretically maximum measure of complexity for a given covariance matrix $I^{-1}$. Finally, $ICOMP$ (3.14) is evaluated for all models and the model with the smallest value is selected as the appropriate model for the given data.

Unfortunately, the maximum KLD is not necessarily a fair measure to compare competing models, because the maximum KLD for the different diffusion-tensor models will generally occur for different orthogonal transformations, i.e. different “choices” of the coordinate system. This is a disadvantage for complex models, whose larger number of parameters increases the chance of finding an unfavorable orthogonal transformation.

$ICOMP$-$TKLD$. The concept of total Kullback-Leibler divergence (TKLD) was proposed in [37] as a rotation-invariant divergence for a completely different application than ours. It measures the orthogonal distance [37] between two distributions. Compared to the original KLD, TKLD can be viewed as a weighted KLD:

$$C_{tot}(I^{-1}) = \frac{C_0(I^{-1})}{\sqrt{1 + \|\nabla f(q)\|^2}},$$

(3.18)

where the numerator $C_0(I^{-1})$ is the KLD (defined in (3.16)) between two PDFs and the denominator is the weighting of the KLD in which $f(\cdot)$ is a convex function defined as $f(q) = \int q \log q$ and $q = N(\hat{\Theta}, I^{-1})$ is the approximate joint posterior PDF of $\Theta_i$. Vemuri [37] has derived that $\|\nabla f(q)\|^2 = \int (1 + \log q)^2 q$.

In (3.15), the theoretical maximum value of KLD was employed to obtain a measure of complexity that was invariant to a rotation of the coordinate system. As such the KLD of the more complicated model is potentially penalized more, so that its $ICOMP$ measure will be non-comparable to that of the simpler model. Replacing the theoretical maximum KLD, i.e. $C_i(I^{-1})$ in (3.14) by TKLD solves this:
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

\[ ICOMP_{TKLD}(\hat{\Theta}_i) = -2 \ln(L(\hat{\Theta}_i | \hat{S})) + 2 C_{tot}(\Gamma^{-1}(\hat{\Theta}_i)), \]  
(3.19)

with \( C_{tot}(\Gamma^{-1}(\hat{\Theta}_i)) \) as defined in eq.(3.18).

3.3 Experiment and Results

The proposed MAP estimate and the ICOMP-TKLD model selection were tested on simulated data and on experimental data. The performance of ICOMP-TKLD was compared with a state-of-the-art model-order selection technique for complex fiber geometries (bedpostx in toolbox FSL). Specifically, we applied our framework to estimate the FA on both simple and complex fiber geometries along the corpus callosum. Finally, we applied our framework to a neuroimaging case study involving TBSS analysis.

3.3.1 Three-way Crossings: A Simulation Study

We simulated data of three-way crossings by means of (3.1) with varying divergence between the fibers to determine the angular resolution of the model estimation. Therefore, we considered fiber-1 to cross a plane spanned by fiber-2 and fiber-3; the angle between the orientation of fiber-1 and this plane was set to different angles: 90°, 60°; furthermore, the divergence between fiber-2 and fiber-3 varied from 0° to 90°. For each configuration, 500 noisy realizations (SNR = 25) were synthesized using Rician noise. The diffusion parameters were set as defined in Table 3.1. As such, the FA’s of the fibers were: \( FA_1 = 0.7, FA_2 = 0.8, FA_3 = 0.9 \). Furthermore, a scan protocol was emulated comprising 92 gradient directions for each of the three \( b \)-values (1000, 2000 and 3000 \( mm^2/s \)), homogeneously distributed over the surface of a sphere as in [8]. The parameters of the triple-tensor model were estimated by maximizing the posterior probability (3.12) (MAP) and the likelihood function (3.11) (MLE). We did not apply model selection at this stage. The estimated tensors were sorted by orientation similarity. Basically, label 1 was assigned to the estimated tensor with the smallest divergence to the ground truth of tensor-1; subsequently, label 2 was given to the remaining tensor with the smallest divergence with ground truth of tensor-2 and the remaining tensor was assigned label 3.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

Table 3.1: Diffusion model parameters for generating synthetic data of three-way crossings. The unit of the diffusion parameters are $10^{-3} \text{mm}^2/\text{s}$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{1,\parallel}$</td>
<td>Axial diffusivity of fiber-1</td>
<td>1.5</td>
</tr>
<tr>
<td>$\lambda_{1,\perp}$</td>
<td>Radial diffusivity of fiber-1</td>
<td>0.35</td>
</tr>
<tr>
<td>$\lambda_{2,\parallel}$</td>
<td>Axial diffusivity of fiber-2</td>
<td>1.45</td>
</tr>
<tr>
<td>$\lambda_{2,\perp}$</td>
<td>Radial diffusivity of fiber-2</td>
<td>0.25</td>
</tr>
<tr>
<td>$\lambda_{3,\parallel}$</td>
<td>Axial diffusivity of fiber-3</td>
<td>1.4</td>
</tr>
<tr>
<td>$\lambda_{3,\perp}$</td>
<td>Radial diffusivity of fiber-3</td>
<td>0.15</td>
</tr>
<tr>
<td>$f_1$</td>
<td>Volume fraction of fiber-1</td>
<td>0.4</td>
</tr>
<tr>
<td>$f_2$</td>
<td>Volume fraction of fiber-2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Estimated fractional anisotropy in three-way fiber crossings. Figure 3.1 shows line plots of the estimated FA (vertically) as a function of the angular divergence between fiber-2 and fiber-3 (horizontally) based on the proposed MAP approach (red) and MLE [8] (blue). The top, middle and bottom graphs depict FA of the first, second and third tensor, respectively. The left, middle and right graphs represent different angles that tensor-1 makes with the plane spanned by tensors 2 and 3: 90°, 60°. It can be observed that the spread in the estimations by MLE is generally much larger than those of MAP. Furthermore, it may be noticed that the FA of the second tensor is precisely estimated down to an angle of approximately 40° with the third tensor. Below this angle there is large spread in both the MLE and MAP estimates. The small spread that can be observed in the FA estimation of the third tensor is due to a biased estimation and labeling errors that may occur for small angular divergences between the fibers. In previous work on MLE of the dual tensor model applied to two-way fiber crossings [8] we also noticed a biased, high FA value for small angular divergence. Below we will show that the dual and triple tensor models tends not to survive the model selection for these configurations.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

Figure 3.1 Statistics of compartment-specific FA estimation in three-way fiber crossings. Each graph depicts FA (from top to bottom FA1, FA2 and FA3) as a function of the angular divergence between fiber-2 and fiber-3, estimated by respectively MAP (red) and MLE (blue). The shaded region around each line indicates the interval bounded by the 25 and 75 percentiles of the estimated values. The left and right graphs represent different angles that fiber-1 makes with respect to the plane spanned by fibers 2 and 3: 90° respectively 60°. For each configuration 500 noisy realizations were synthesized with SNR=25.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

Figure 3.2 (a, b) Statistics of angular divergence between fiber-2 and fiber-3 as a function of true fiber divergence estimated by MAP (red) and MLE (blue). The shaded region around each line indicates the interval bounded by the 25 and 75 percentiles of the estimated values. Each plot presents a different angle that fiber-1 makes with respect to the plane spanned by fiber-2 and fiber-3: 90° respectively. (c, d) Scatterplots of estimated orientations by means of MAP estimation from 500 noise realizations. Again, the plots reflect different angles of tensor-1 with respect to the plane spanned by fibers 2 and 3: 90°, 60°; the angle between fiber-2 and fiber-3 was kept fixed at 90°.

Estimated fiber orientation in three-way fiber crossings. Figure 3.2 (a-b) present line plots of the estimated angular divergence between fiber-2 and fiber-3 based on the proposed MAP approach (red) and MLE (blue). This orientation divergence was calculated by \( \langle \text{fiber-1}, \text{fiber-2} \rangle = \cos \left( \frac{\langle \hat{v}_{12} \cdot \hat{v}_{13} \rangle}{\| \hat{v}_{12} \| \| \hat{v}_{13} \|} \right) \), in which \( \hat{v}_{12} \) and \( \hat{v}_{13} \) are the largest principal eigenvectors of estimated tensor-2 and tensor-3 (after sorting). Clearly, the estimation of the angular divergence of both MAP and MLE demonstrates a large spread below 40°. However, above 40° the spread of the MAP estimation is small and the median estimated divergence shows a nice linear relation with the imposed
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

divergence. Alternatively, the spread of MLE is somewhat larger in this range. Figure 3.2(a-b) demonstrate that the MAP estimation of the angular divergence is precise down to 40°, just like the MAP estimation of FA. Figure 3.2(c-d) show scatterplots of projected orientations estimated by means of the MAP approach for 500 noisy realizations of two configurations in which tensor-1 makes an angle of either 90° or 60° with the plane spanned by the first principal eigenvectors of tensor-2 and tensor-3 while the angle between tensor-2 and tensor-3 remains fixated at 90°. Essentially, these figures confirm the small spread that is also observed in the corresponding line plots of Figure 3.2(a-b) (see the 90° result in each plot).

Estimated volume fraction in three-way fiber crossings. To assess the influence of compartment size, we simulated three-way crossings by means of (3.1) while varying the volume fractions of the underlying fibers. For simplicity the three fibers were taken perpendicular to each other. Furthermore, the volume fraction of fiber-1 (i.e. \( f_1 \)) was varied from 0 to 1 and the volume fractions of fiber-2 and fiber-3 were taken equal, i.e. \( f_2 = f_3 = (1 - f_1)/2 \). Thus, \( f_1 = 0 \) corresponds to a two-way crossing and \( f_1 = 1 \) represents a single-fiber case. The other parameters were the same as in the previous experiment: SNR=25, 92 gradient directions for each of the three \( b \)-values, diffusivities c.f. Table 3.1. Notice that 500 noisy realizations were generated for each configuration.

Figure 3.3 shows the estimated volume fraction (Figure 3.3(a, b, c)) and FA (Figure 3.3(d, e, f)) based on MAP estimation (red) and MLE (blue). The top, middle and bottom images represent the parameters estimated for tensors-1, -2 and -3, respectively. Again MLE generally yields a larger spread in its estimations than MAP. The FA of tensor-1 is precisely estimated for volume fractions larger than 0.2 (top right). The FA estimation of tensor-2 and tensor-3 are precise for volume fractions of tensor-1 up to 0.7 (middle/bottom right). Observe that this corresponds to volume fractions of 0.15 for each of the other tensors. The volume fraction of the tensor-1 is precise also at small volume fractions, but becomes biased below 0.1 (top left). On the other side it becomes no longer precise for volume fractions of 0.7 and higher. Notice that with increasing volume fraction of fiber-1 (and consequently decreasing volume
fractions of the other two fibers) the configuration tends to resemble a single fiber structure, but that the uncertainty and sometimes the bias of the estimated parameters increase.

Figure 3.3 Statistics of estimated volume fractions (a, b, c) and compartment-specific FAs (d, e, f) as obtained by MAP (red) and MLE (blue) as a function of the volume fraction of fiber-1 in three-way fiber crossings under 90° angles and for \( f_1 = f_3 = (1 - f_2) / 2 \). The shaded region around each line indicates the interval bounded by the 25 and 75 percentiles of the estimated values. For each configuration 500 noisy realization were synthesized with SNR=25.

Summarizing, MAP clearly outperforms MLE in estimating the parameters of the
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

triple-tensor model. However, the reliability of triple-tensor estimation by MAP reduces with decreasing divergence and volume fraction, i.e. where the three-way crossing degenerates into a two-way crossing or a single fiber. In these voxels, a simpler representation by the dual-tensor or the single-tensor model is more appropriate.

Model-order selection of three-way fiber crossings. We applied the proposed ICOMP-TKLD model selection to the data from the previous experiments. As such, Figure 3.4 collates the performance of the proposed ICOMP-TKLD model selection approach. Figure 3.4(a) shows the percentage of cases the triple-tensor model was selected as a function of the angle between fiber-2 and fiber-3 on the data from Figure 3.1. Likewise, Figure 3.4(b) depicts the same percentage as a function of the volume fraction of fiber-1 on the data from Figure 3.3. Figure 3.4(a) demonstrates that the success rate of triple-tensor model selection decreases significantly as the divergence between the two fibers is smaller than 40°, which corresponds with the lower precision of the estimated parameters in Figure 3.1 and Figure 3.2. Furthermore, Figure 3.4(b) shows that the triple-tensor model is particularly preferred when the volume-fraction ratio’s of the three tensors are in the same range, i.e. $f_i \in [0.1, 0.5]$. Finally, in all cases MAP outperforms MLE, which demonstrates the benefit of the presented prior (3.9).

Figure 3.4 (a) Estimated probability of selecting the triple-tensor model selection as a function of the angle between fiber-2 and fiber-3 on the data from Figure 3.1 by MAP (red) and MLE (blue). The different line styles reflect different orientations of fiber-1 with respect to the plane spanned by fibers 2 and 3. (b) Estimated probability of selecting the triple-tensor model selection as a function of the volume fraction of fiber-1 on the data from Figure 3.3 by MAP (red) and MLE (blue).
3.3.2 Validation of Model Selection on Brain Data

The proposed model selection framework for single-, dual-, and triple-tensor models was applied to DTI data of the Human Connectome Project (HCP) [28] [29] and DTI data of an ongoing study into the effects of HIV on the brain [3]. The HCP data is an open-access dataset intended for characterization of brain connectivity and function and their relationship to behavior [17] [29] [28]. The relevant acquisition parameters of the HCP dataset were: three $b$-values 1000, 2000 and 3000 s/mm$^2$, 288 gradient directions, TE/TR 89.5/5520 ms, voxel size 1.25×1.25×1.25 mm$^3$. Similarly, the relevant acquisition parameters for the HIV data were: two $b$-values, i.e. 1000 and 2000 s/mm$^2$, 130 gradient directions, voxel size 2.0×2.0×2.0 mm$^3$.

Comparison of model-order selection. Figure 3.5(a,d) show images of a randomly selected subject from the HCP dataset and the HIV dataset, respectively. The left sub-images in Figure 3.5(a,d) reflect the principal directions of single tensor fits to the data: red denotes left-right, blue top-bottom, and green front-back. The right sub-images display the FA derived from single tensor fits and the overlay depicts the probabilistic tractography outcome of FSL-PROBTRACKX (FSLV6.64). Notice that the ROI’s contain among others a single fiber region (the center of the corpus callosum) and a triple tensor region (as reported previously [40] [41]). Figure 3.5(b,e) depicts the model selection and fiber-orientation estimation results from FSL (by means of the command ‘bedpostx’ [7] [17] [31]) on these data. The number of line segments quantifies the number of fiber bundles in a voxel, while the color of each line segment indicates the orientation of the fiber (red for left to right, green for front to back, and blue for bottom to top). Much in the same way Figure 3.5(c,f) demonstrate the model selection and MAP estimation by the proposed framework.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

Figure 3.5 (a,d) DTI data of subjects selected from the HCP dataset (a) and the HIV dataset (d). The left sub-images of (a,d) depict the principal orientations of single tensor fits to the data: red corresponds to left-right, blue to top-bottom, and green to front-back. The right sub-images of (a,d) display the FA derived from single tensor fits and the overlay depicts the probabilistic tractography outcome. Images (b,c,e,f) focus on the regions indicated in (a,b), which supposedly contain a three-way crossings. (b,e) Model selection and parameters estimation by FSL and (c,f) by the proposed ICOMP-TKLD framework. The number of line-segments indicates number of inferred compartments. The color of line segments represents the estimated orientations as in (a,d).

The consistency of the model selection was explored by registering 10 subjects from the two datasets using FNIRT registration [42]. Subsequently, we determined the
average number of fibers in each voxel of the region of interest. Figure 3.6 shows this number in a false color representation. The main difference between the results of FSL and that of the proposed framework is in the central part of the CC. Here, ICOMP-TKLD predominantly infers a single fiber configuration as preferred, whereas FSL mostly selects dual or triple ball-and-stick models. The proposed framework gives comparable inference to FSL in the region supposedly containing crossing fibers. Thus, Figure 3.5(c, f) and Figure 3.6 essentially convey that the proposed method yields improved adaptation to the underlying structures compared to the FSL technique particularly in ‘simpler’ structures.

Figure 3.6 The average number of fibers inferred in spatially aligned ROIs of 10 subjects from the HCP dataset (a,b) and HIV dataset (c,d). Note that the ROI supposedly contains a three fiber crossing, as in Fig.5; subjects were registered using fnirt [42].

---

Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting
Influence of model selection on fractional anisotropy of Corpus Callosum. Figure 3.7 shows the estimated FA along the Genu of the Corpus Callosum (GCC) in one subject from the HCP dataset. The graph corresponds to the voxels in the yellow box (see inset) highlighted by the red-colored tractography outcome. The FA along the tract was estimated by: (A) the proposed single-, dual- and triple-tensor model selection framework (red); (B) a simplified version of this framework that focused on single- or dual-tensor models only, thus not considering three-way crossings (blue); (C) a conventional single tensor fit (green). The former, multi tensor methods require selection of the tensor compartment that corresponds to the GCC since the labels assigned to the compartments are random. To solve this, the FA belonging to CC was selected based on “front evolution” [46]. In front evolution, the estimated tensor of one compartment is randomly chosen as the reference tensor. Then, the tensor of a neighborhood voxel with the smallest Frobenius norm to the reference tensor receives
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

the same label. After processing all neighbors of the current front, these neighbors become the new reference tensors for the next iteration. The solid lines in the figure denote the median value of the estimated FA value perpendicular to the tract. Furthermore, the shaded areas comprise the 25 and 75 percentiles of the local distributions of estimated FA values.

The figure illustrates how estimation of the FA along a tract is influenced by the model choice. We hypothesize that the large fluctuations of the single tensor approach is caused by space-variant partial volume effects, c.q. the mixing with neighboring structures. Clearly, only the estimation by the proposed framework (red) is without large fluctuations. This outcome accords with one’s intuition that the FA should vary smoothly along a tract. Essentially, the stable estimation signifies that the confounding influence of partial volume effects are effectively dealt with by our framework.

A case study into the influence of handedness using TBSS. The application and potential benefit brought by our framework will be demonstrated by a small case study into the influence of handedness on brain structure using TBSS. The handedness of subjects from the HCP dataset is quantified on a scale from -100 to 100 (left-handed to right-handed) by the handedness inventory of Schachtar et al [43]. We included the 12 most left-handed and the 12 most right-handed female subjects from the HCP dataset. The handedness of the selected subjects ranged from -100 to -30 in the left-handed group and from 80 to 95 for the right-handed group.

All data was registered to the MNI152 standard space using the version of FNIRT that is implemented in FSL (version 5.0.7) [44]. Subsequently, differences between the left and right brain were analyzed in relation to the handedness. Therefore, we used the TBSS symmetry test. We performed this test with the classical TBSS technique [45], i.e. based on single tensor data, as well as the extended TBSS [46] method for the one/dual/triple tensor models. Compared to the classical TBSS method, the extended TBSS technique [46] employs ‘front evolution’ to avoid swapping of the tensor components, i.e. so that corresponding tensor components are properly compared.

Figure 3.8 shows the outcome of testing whether one or more of the left brain’s FAs or the left brain’s volume fractions is significantly greater than the corresponding
parameter in the right brain. The top images focus on the left-handed subjects, the bottom images on the right-handed subjects. The first column shows the classical TBSS outcome (based on a single tensor); the second column visualizes the extended TBSS outcome for the volume fractions of FSL’s ball-and-stick model; the third and fourth column display the extended TBSS outcome for FAs respectively volume fractions estimated by the proposed three-tensor approach. Notice that in such a comparison TBSS imposes symmetry of the data. Therefore significant differences are indicated only on the right brain’s skeleton.

Figure 3.8 FSL-TBSS outcome testing whether the brain’s FA or volume fraction in the left hemisphere is significantly larger than in the right hemisphere. Results of a group of predominantly left-handed subjects (top row) and predominantly right-handed subjects (bottom row). The first column shows the classical TBSS outcome for FA; the second column visualizes the outcome of the extended TBSS framework for the volume fraction of FSL’s ball-and-stick model; similarly, the third and fourth columns show the extended TBSS outcome for FA respectively volume fraction generated with the proposed triple-tensor framework. The green lines denote the white matter’s skeleton; the red points indicate clusters of significant difference.

It can be clearly seen that the classical TBSS analysis and the TBSS analysis of FSL’s volume fraction find little differences in both groups. Instead, the TBSS analysis of the triple-tensor FA and volume fraction yields substantially more significant differences. Furthermore, it can be observed that the classical DTI differences are identifiable in either the triple-tensor’s FA or volume fraction. Finally, slightly more differences were
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

found in the right-handed group than in the group of left-handers. This outcome essentially confirms previous results (see e.g. [47]).

To our opinion, the outcome signifies how the proposed method can enhance the sensitivity of statistical analysis of experimental brain data. We believe that the enhanced sensitivity is due to a more accurate and precise estimation of FA and volume fraction in fiber crossings.

3.4 Discussion and Conclusion

We developed a framework for adaptive estimation of diffusion properties particularly for three-way fiber crossings. A MAP estimator involving a mild prior was used to estimate the parameters of a triple-tensor model. The prior precluded the degeneracy of the model estimation. A new model selection technique quantified the extent to which candidate models were appropriate, i.e. single-, dual- or triple-tensor model. As such, overfitting was circumvented. The model selection employed the total Kullback-Leibler divergence [37] to assess the information complexity [36] (ICOMP-TKLD), thereby avoiding problems related to the choice of coordinate system. It balanced the trade-off between the goodness of fit and the complexity brought by the order of the diffusion model. The model selection used the Cramer-Rao lowest bound of variance, so that it implicitly adapts to the data acquisition protocol.

The proposed technique extends previous work, which aimed to reconstruct fiber orientations in three-way crossings [11] [17]. A crucial difference is that we employ a rank-2 tensor model, whereas the previous works concerned non-parametric [11] or ball-and-stick models [17]. In other words, we recover the full diffusion shape in different types of structures.

The proposed MAP estimation was compared to MLE using phantom data of three-way fiber crossings. The experiments showed that the prior helps to improve the precision of the triple-tensor model, without introducing a bias in the estimated diffusivities. The spread in the parameters from MLE was generally much larger than those from MAP estimation. The FA of the third tensor could be precisely estimated with MAP down to an angle of approximately 40° with the second tensor.
Furthermore, the volume fractions and FAs could be accurately and precisely estimated if the volume fraction of the first tensor was between 0.2 and 0.7. At the latter boundary (0.7), the volume fractions of the remaining two tensors were merely 0.15. The proposed framework was also compared with the ball-and-stick modeling approach from FSL using experimental brain data, acquired with two different protocols. The configurations inferred by our method corresponded to the anticipated neuro-anatomy, both in the single fiber and in the triple-crossing regions. The main difference with FSL was observed in a single fiber region. Here, ICOMP-TKLD predominantly inferred a single fiber configuration, as preferred, whereas FSL mostly selected dual or triple ball-and-stick models. The different performance is due to ICOMP-TKLD’s model selection. FSL’s model selection method is done in a Bayesian way, merely by adding a prior to the parameter estimation. Finally, a TBSS experiment showed the enhanced sensitivity of statistical analysis using the proposed method.

There are a few limitations of our approach. Firstly, we assume a mono-exponential decay along the eigenvectors of the three compartments up to $b=3000 \text{ s/mm}^2$. Measuring at higher $b$-values will certainly introduce sensitivity to compartments such as the myelin sheet [48] with severely restricted and hindered diffusion [49]. In the latter case, the Gaussian diffusion assumption is no longer valid. We considered investigating non-Gaussian diffusion beyond the scope of our current work.

Secondly, our framework focused on single, dual- and triple-tensor models to characterize the full diffusion profile. There might exist four-way crossings or even more complicated fiber structures, although we did not find papers reporting about such configurations. In [8] we showed that estimating a dual rank-2 tensor model already requires HARDI at two $b$-values, data of sufficient SNR, and some model restrictions. The latter is needed to ensure stability as the number of model parameters may approach or even surpass the number of degrees of freedom present in the data. Presently, the modeling was extended to triple-way crossings by employing several sophisticated concepts. Fitting a quadruple rank-2 tensor model to voxels with a four-way-crossing will be even more challenging. Developing methods for estimation of the diffusion properties in four-way-crossing fiber bundles will remain an important challenge for future research.
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting

References


Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting


72
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting


Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting


Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting


75
Chapter 3. Estimation of diffusion properties in three-way fiber crossings without overfitting


4 A Structure-adaptive Noise Filter for Accurate Estimation of Diffusion Properties in Complex Fibre Structures

Fitting of complex diffusion models to diffusion-weighted magnetic resonance images (DW-MRI) is often hampered by a low signal to noise ratio (SNR) or a limited number of diffusion-weighted images (DWIs). A standard solution is to suppress the noise by filtering the DWIs. However, it remains challenging to adapt the filter kernel such that the noise is suppressed, while the signal’s diffusion properties are preserved.

We propose a structure-adaptive, compartment-specific version of the Linear Minimum Mean Square Error (LMMSE) filter to achieve this. Initially, we estimate structural information from the data by means of a severely constrained dual tensor model (DTM) using automated relevance determination (ARD). The resulting DTM is used to decompose the DWIs into component-specific contributions. Subsequently, each compartment-specific contribution is filtered by a space-variant, structure-adaptive kernel. The filter coefficients reflect the tensor similarity between the neighborhood tensors and the central tensor. Thereafter, the regularized DWIs are reconstructed from the filtered DWI contributions. Finally, the regularized DWIs are used as input to estimate a mildly constrained dual tensor model using ARD.

The results demonstrate that the method preserves the edges of fiber compartments, while denoising the DWIs. Therefore, it is expected that subsequent dual compartment diffusion estimation will achieve a better accuracy and precision. The benefit of the proposed framework is shown through a case study into the relation between white matter atrophy and aging. It is demonstrated that the noise on the
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

DWIs limits the regions with significant differences between two age groups, whereas subtle differences are identified after DWI filtering. Discarding up to fifty percent of the DWIs followed by our denoising framework still yields results comparable to those obtained on the original set of non-denoised DWIs. Consequently, our method may enhance the sensitivity of DWI studies or achieve identical sensitivity using a more limited scan protocols.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

4.1 Introduction

Diffusion-weighted Magnetic Resonance Imaging (DW-MRI) can provide unique information about the integrity of white matter structures in vivo [1]. The diffusion profile in a voxel is often described by a positive-definitive tensor of rank-2, which is estimated from a series of diffusion-weighted images (DWIs). However, a single diffusion tensor is insufficient to characterize the diffusivities in crossing fiber configurations. Hence, for those configurations, more complex models such as the dual tensor model (DTM) [2] [3] and constrained spherical deconvolution [4] have been proposed. Unfortunately, accurate estimation of diffusion properties is often hindered by a low signal to noise ratio (SNR), especially when complex diffusion models with many degrees of freedom are used [5] [2]. Images with stronger diffusion weighting, introduced to facilitate the disentanglement of diffusion profiles from fascicles in a fiber crossing, have lower MR signal magnitude and consequently a lower SNR. Increasing the number of acquired DWIs can compensate the detrimental effect of noise on the parameter estimation. However, increasing the number of DWIs is restricted by scan time constraints, especially in clinical practice. Even advances in multi-slice and multi-band protocols [6], which have been proposed to accelerate the acquisition, do not yet solve this problem. Moreover, many existing datasets will have an insufficient number of DWIs to permit reliable estimation of the diffusion properties in voxels encompassing complex fiber geometries.

Several approaches have been introduced to increase the SNR of the diffusion data. These techniques can be roughly divided into three categories: (a) methods that perform denoising on the DWIs such as the linear minimum mean square error (LMMSE) filter [5] [7], non-local means (NLM) and unbiased NLM [8] [9] [10], Markov random fields [11] [12], and anisotropic Wiener filtering [13]; (b) methods that estimate diffusion tensors from the unfiltered DWI after which denoising is performed in tensor space, e.g. by a Riemannian approach [14], regularization of the principal eigenvectors [15], compartment-specific structure-adaptive tensor filtering [16] or the spectral decomposition of tensors [17]; (c) methods that try to simultaneously smooth and estimate model parameters subject to anatomical
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

constraints [18] [19]. Among these techniques, denoising of DWIs based on the LMMSE scheme has attracted much attention. This is because – unlike other schemes – it has a closed-form expression, making it computationally efficient. In addition, it is easy to use and to implement [7] [20] [21].

However, the aforementioned techniques, except the LMMSE filtering schemes, ignore inherent structural differences between voxels. LMMSE avoids blurring across boundaries between anisotropic and isotropic regions by effectively turning off the filter near such boundaries. As a side effect, no noise suppression takes place. To limit this negative side effect, we previously proposed to steer the kernel of the LMMSE filter along fiber tracts [22]. However, this method did not consider crossing fiber structures and therefore the signals belonging to different fiber compartments in fiber crossings did not benefit from LMMSE-based noise suppression. To extend the method to complex fiber geometries might prove highly relevant since a large percentage of the voxels appears to contain crossing or branching fiber bundles [23] [24] [25] [26].

In this chapter, we propose a new filtering framework for DWIs in which a closed-form LMMSE filter is constructed that adapts to complex fiber structures. The framework contains five steps: (1) parameter estimation of a severely constrained dual tensor model by using automated relevance determination (ARD); (2) decomposition of each DWI into compartment-specific DWIs; (3) smoothing the compartment-specific DWIs with a structure-adaptive LMMSE filter steered by tensor similarity; (4) reconstruction of the regularized DWI from its filtered compartment-specific constituents; (5) parameter estimation of a mildly constrained dual tensor model by using ARD on the regularized DWIs.

It will be shown that the proposed framework facilitates the application of complex diffusion models to limited scan-time datasets from neuroimaging practice, i.e. datasets with a limited number of DWIs and/or a low SNR. Therefore, it may shed a new light on the diffusion properties derived from such datasets.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

Incorporation of structural information is the key to feature-preserving denoising of DWIs. A flowchart of the presented framework is given in Figure 4.1. We start in Section 4.2.1 with a detailed description of how we model the diffusion structure by a multi-tensor model (4.2.1.1). Then we briefly introduce maximum likelihood estimation (MLE) for initial parameter estimation (4.2.1.2) and automated relevance determination (ARD) for estimating a mean posterior probability (4.2.1.3). Finally, the compartment-specific structure-adaptive LMMSE filter will be presented in Section 4.2.2.

Figure 4.1 Flowchart of the framework for component-specific structure-adaptive denoising of diffusion-weighted images (DWIs). ARD stands for automatic relevance determination, DTM for dual tensor model.

4.2 Methods
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

4.2.1 Estimation of Local Structure Information

4.2.1.1 Multi Tensor Model

Conventionally, the diffusion of water in the brain is described by a rank-2 diffusion tensor, which is estimated from a large number of DWIs [1]. There is an ongoing debate on how to effectively characterize the diffusivities in complex fiber configurations, where the diffusion profile is not adequately described by a single rank-2 tensor [2] [3]. Previously, we introduced an optimization framework based on a constrained dual tensor model that enabled accurate estimation of both the diffusion shape and the diffusion orientations in crossing fiber configurations [2]. The model is given by

\[
S_{\theta,j} = S_0 \left( \sum_{i=1,2,iso} f_i \exp(-b_j g_j^T D_i g_j) \right),
\]

(4.20)

where \(S_{\theta,j}\) is the expected ‘true’ signal value given the model parameter vector \(\theta\), and \(j\) indicates the DWI corresponding to diffusion weighting \(b_j\) in gradient direction \(g_j\); \(S_0\) is the signal without diffusion weighting; \(D_1\) and \(D_2\) are two anisotropic rank-2 diffusion tensors and \(D_{iso}\) is an isotropic diffusion tensor (i.e., \(D_{iso} \cdot I_{3x3}\)); \(f_i \geq 0\) \((i=1,2,iso)\) represent volume fractions, and it is assumed that \(f_1 + f_2 + f_{iso} = 1\). In (4.20), if \(f_2\) and \(f_{iso}\) are set to be zero, then the dual tensor model degenerates to the classical single tensor model, which is often employed to model voxels encompassing single fiber data.

4.2.1.2 Maximum Likelihood Estimation

For a voxel with 3D spatial coordinates \(\upsilon\), the measured diffusion-weighted signal \(S_{\upsilon,j}\) in direction \(g_j\) is Rician distributed [27]. Therefore, the probability density function (PDF) for \(S_{\upsilon,j}\) with a chosen diffusion weighting \(b_j\) is given by

\[
p(\tilde{S}_{\upsilon,j} \mid S_{\theta,j}, \sigma) = \tilde{S}_{\upsilon,j} \exp\left(-\frac{\tilde{S}_{\upsilon,j}^2 + S_{\theta,j}^2}{2\sigma^2}\right) I_0\left(\frac{\tilde{S}_{\upsilon,j} S_{\theta,j}}{\sigma^2}\right),
\]

(4.21)

in which \(\sigma\) denotes the standard deviation of the noise and \(I_0\) is the modified Bessel function of zero-th order. The DWIs can be assumed to be independent, which means
that the joint probability density function \( p(\tilde{S}_v | \theta, \sigma) \) of the signal profile \( \tilde{S}_v \) is given by the product of the marginal distributions for the measured signals \( \tilde{S}_{v,j} \) in each of the \( N_g \) diffusion weighting directions

\[
p(\tilde{S}_v | \theta, \sigma) = \prod_{j=1}^{N_g} p(\tilde{S}_{v,j} | S_{\theta,j}, \sigma),
\]

where \( p(\tilde{S}_v | \theta, \sigma) \) is the likelihood function which will henceforth be denoted by \( L(\theta, \sigma | \tilde{S}_v) \).

In previous work [2], we parameterized the tensor \( D_i \) by a diagonal matrix \( E_i = \text{diag}(\lambda_{i//}, \lambda_{i\perp}, \lambda_{i\perp}) \) encoding the eigenvalues and a rotation matrix \( R_i \), \( D_i = R_i^T E_i R_i \). The matrices \( R_{i=1,2} \) describe the orientation of the largest, i.e. parallel, diffusivity \( \lambda_{i//} \) by rotations around the \( x- \), \( y- \) and \( z- \) axes: \( R_i(\alpha_{i-\pm}) = R_x(\alpha_i)R_y(\alpha_i)R_z(\alpha_i \pm \alpha_i) \). The first two rotations \( R_x(\alpha_i)R_y(\alpha_i) \) determine the plane in which the principal eigenvectors of the tensors reside; \( R_z(\alpha_i \pm \alpha_i) \) represents a rotation in the plane of the fiber crossing where the angle of the principal eigenvectors of the two tensors \( i=1,2 \) is given by the mean orientation \( \alpha_3 \) with deviations \( \pm \alpha_4 \). Inspection of the diagonal matrices shows that axially symmetric diffusion is assumed corresponding to the on average axial symmetry of the axons comprising white matter fibers.

The parameter values can be inferred by maximizing the constrained likelihood function [28], also referred to as maximum likelihood estimation (MLE):

\[
(\hat{\theta}, \hat{\sigma}) = \arg\max_{\theta, \sigma} (\ln L(\theta, \sigma | \tilde{S}_v)) \quad \text{s.t.} \quad C,
\]

with parameter vector

\[
\theta = \{ s_0, \lambda_{i//}, \lambda_{i\perp}, \lambda_{2//}, \lambda_{2\perp}, \alpha_1, \alpha_2, \alpha_3, \alpha_4, f_1, f_2, f_{iso}, D_{iso} \}
\]

and constraint

\[
C = \{ f_1 + f_2 + f_{iso} = 1 \}.
\]

This axially symmetric dual tensor model can be further constrained by assuming an equal diffusivity along the two fiber bundles, i.e. \( \lambda_{2//} = \lambda_{i//} \). As such a difference in diffusion shape between the two tensors is represented by a difference in perpendicular
diffusivity ($\lambda_{1\perp} \neq \lambda_{2\perp}$). Another restriction assumes that the isotropic diffusivity equals that of free water at body temperature of 37°C, $D_{iso} = 3 \cdot 10^{-3} \text{mm}^2\text{s}^{-1}$. These constraints only mildly affect the diffusion shape but substantially reduce the sensitivity of the estimates to noise [2]. Hence, the mildly constrained dual tensor model uses constraints

\[ C_m = \{ f_1 + f_2 + f_{iso} = 1, \lambda_{1/1} = \lambda_{2/1}, D_{iso} = 3 \cdot 10^{-3} \text{mm}^2\text{s}^{-1} \} \quad (4.26) \]

and the estimation (4.23) yields parameter vector $\hat{\theta}_{mc}$. This model enables accurate estimation of diffusion properties of crossing white matter structures. However, due to the large number of degrees of freedom (DOFs) of model (4.24) and (4.26) its applicability is restricted by, among others, the SNR and the number of DWIs. The primary objective of this chapter is to enhance the applicability of the mildly constrained model (4.24) in DW-MRI data by structure-adaptive denoising.

To achieve such denoising, an initial estimate of the local structure information is required. Since clinical datasets may have a (very) low SNR or are composed of a small number of DWIs, unbiased estimation of all DOFs in model (4.24) and (4.25) cannot always be achieved. To cope with this situation prior to filtering, we further constrain the model by omitting all shape differences between the two diffusion tensors, i.e., $\lambda_{2/1} = \lambda_{1/1}$ and $\lambda_{1\perp} = \lambda_{2\perp}$ and by setting $f_{iso}$ to a constant. This set of constraints

\[ C_s = \{ f_{iso} = 0.05, f_1 + f_2 + f_{iso} = 1, \lambda_{1/1} = \lambda_{2/1}, \lambda_{1\perp} = \lambda_{2\perp}, D_{iso} = 3 \cdot 10^{-3} \text{mm}^2\text{s}^{-1} \} \quad (4.27) \]

is used in our severely constrained dual tensor model of which the parameter estimates are denoted $\hat{\theta}_{sc}$. In the experiments below, we will evaluate the SNR and the number of DWIs required to support the application of the severely constrained dual tensor model. If even fitting a severely constrained dual tensor model is not supported by the acquired data, more constraints such as a fixed divergence between the crossing fibers (i.e., $\alpha_4$) might be needed. Initially, we estimate the parameters of the severely constrained model by computing the expectation of the posterior probability using ARD (c.f. 4.2.1.3). This delivers us a coarse description of the underlying fiber structure per voxel. Based on this structural information the data is filtered, after which the mildly constrained model (4.24) is fitted to the filtered data.
4.2.1.3 Automatic relevance determination (ARD)

Constrained dual tensor models such as the ones introduced in the previous subsection, are still prone to ‘overfitting’ in areas containing a single fiber bundle, where a single tensor representation would be appropriate. Therefore, we devised an ARD method for rank-2 tensor models, which ensures that the volume fraction of a second fiber compartment is automatically forced to be close to zero if there is no evidence for a second fiber present in the data. Our ARD estimates the mean posterior probability of the parameters, subject to the appropriate set of constraints, based on Bayes’ theorem

\[ p(\theta | \tilde{S}_v) \propto p(\tilde{S}_v | \theta) p(\theta), \]  

(4.28)

where \( p(\tilde{S}_v | \theta) \) is the aforementioned likelihood function and \( p(\theta) \) a prior probability for which we used Jeffreys prior

\[ p(\theta) \propto \det(I(\theta))^{1/2}, \]  

(4.29)

where \( I(\theta) \) is Fisher’s information matrix given by

\[ I(\theta) = -E_s \left\{ \frac{\partial^2 \ln(p(S | \theta, \sigma))}{\partial \theta \partial \theta^T} \right\}. \]

(4.30)

Our ARD employs a Markov Chain Monte Carlo (MCMC) technique with Metropolis-Hasting sampling of the posterior distribution. It is initialized by the MLE of the corresponding dual tensor model. Jeffreys prior conveys support for the dual tensor representation in crossing fibers, since in that case \( \det(I(\theta)) \) will be large. Consequently, the posterior distribution peaks near the initial dual tensor parameters obtained by MLE. Reversely, \( \det(I(\theta)) \) will be small in single fiber regions since there is very little information about the diffusion tensor of the second compartment. Then, the prior, which is proportional to \( \det(I(\theta)) \), becomes harsh, thereby promoting a near-zero volume fraction for one tensor component in the posterior distribution.

4.2.2 Structure-Adaptive LMMSE

A closed-form LMMSE framework to estimate the underlying ‘noise-free’ diffusion weighted signal values was proposed in [5] [29] to reduce the non-central \( \chi \)-distributed noise in MR magnitude images.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

\[
\hat{S}_{v,j} = \langle S_{v,j}^2 \rangle - 2N_{\text{coil}} \sigma_n^2 + \left( 1 - \frac{4\sigma_n^2 \left( \langle S_{v,j}^2 \rangle - N_{\text{coil}} \sigma_n^2 \right)}{\langle \hat{S}_{v,j}^4 \rangle - \langle \hat{S}_{v,j}^2 \rangle^2} \right) \left( \hat{S}_{v,j}^2 - \langle S_{v,j} \rangle \right),
\]

(4.31)

in which \( \hat{S}_{\theta,j} \) are the estimated, squared noise-free signals corresponding to gradient direction \( g_j \) in a voxel \( v \) and \( N_{\text{coil}} \) denotes the number of acquisition coils. If the DWI reconstruction scheme is based on a ‘SENSE’ protocol [30], then the noise is Rician distributed which is the first order case of the non-central \( \chi \)-distribution. In case of Rician noise, \( N_{\text{coil}} \) must be set to one. Alternatively, for datasets acquired based on ‘GRAPPA’ reconstruction [31], \( N_{\text{coil}} \) is set to the number of coils. Since all the datasets in the current study are reconstructed using the ‘SENSE’ protocol, we set \( N_{\text{coil}} = 1 \) in our framework. Furthermore, in (4.31), \( \langle \cdot \rangle \) represents the estimator of a sample’s expectation value, which is defined as:

\[
\langle X_v \rangle = \frac{1}{|\Omega_v|} \sum_{\rho \in \Omega_v} X_{\rho},
\]

(4.32)

with \( \Omega_v \) being a set of voxels in a 3D neighborhood around voxel \( v \) (including voxel \( v \)); \( |\Omega_v| \) denotes the size of the neighborhood in \( \Omega_v \) and \( X \) refers to the quantity of interest. Finally, \( \sigma_n^2 \) represents the local noise variance. For Rician distributed noise, \( \sigma_n^2 \) can be estimated by [32] [5]:

\[
\hat{\sigma}_n^2 = \text{mode} \left\{ \frac{1}{|\Omega_v| - 1} \sum_{\rho \in \Omega_v} S_{\rho}^2 - \langle S_{\rho} \rangle^2 \right\}
\]

(4.33)

One may observe that within the estimation window \( \Omega_v \) all voxels are weighted equally (see (4.32)). In case the window overlaps two (or more) populations of voxels belonging to different fiber compartments, the filtering is effectively turned off. Our structure adaptive LMMSE framework solves this problem in two steps: 1) decomposition of the original DWI into compartment-specific DWIs based on the estimated coarse structural information; 2) application of a structure-adaptive closed-form LMMSE filter (4.31) to each compartment-specific DWI, employing a kernel whose weights are not equal for all neighborhood voxels, but based on the tensor
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

similarity with the central voxel.

4.2.2.1 Compartment-specific DWI contribution

In equation (4.20), the contribution of the $i^{th}$ compartment in voxel $v$ to a DWI along gradient $g_j$ is modeled based on the severely constrained dual tensor model as

$$S_{v,j,i} = \hat{S}_{v,0} f_{v,i} \exp(-b_j g_j^T \hat{D}_v g_j). \quad (4.34)$$

where $v$ represents the 3D spatial coordinate of a voxel. Notice that in a single fiber situation $f_{v,2} = 0$, in which case the contribution $S_{v,j,2} = 0$. The residual part is calculated by $R_{v,j} = \hat{S}_{v,j} - S_{v,j,1} - S_{v,j,2}$, which contains the contribution of the isotropic part, potential model errors as well as noise.

4.2.2.2 Structure-adaptive LMMSE and DWI reconstruction

We assert that $C_{v,j,i}$ is an inherent property of the $i^{th}$ fiber compartment in a voxel $v$. However, due to noise, we can expect that even the tensor components estimated from the same fiber will not be exactly equal, i.e. $C_{v,j,i}$ will locally vary. To reduce the noise, we propose to apply a structure-adaptive LMMSE filter to $C_{v,j,i}$ such that voxels sharing a higher structural similarity to the voxel $v$ will be assigned a higher weight. We opt to measure this structural similarity by a tensor similarity measure (TS):

$$d_{pv}^{ki} = \exp\left(-\frac{|D_{p,k} - D_{v,j}|^2}{\beta |D_{p,k}||D_{v,j}|}\right) \quad (4.35)$$

where $|A|$ denotes the Frobenius norm of $A$: $|A| = \sqrt{\sum_{m,n} a_{mn}^2}$ with $a_{mn}$ representing one element of matrix $A$; $D_{p,k}$ is the $k^{th}$ tensor in voxel $p$; $\beta$ a scaling parameter. Subsequently, the expected contribution by the $i^{th}$ fiber in voxel $v$ to the compartment-specific DWI is calculated by a weighted sum

$$\hat{S}_{v,j,i} = \left\langle S_{v,j,i} \right\rangle = \frac{\sum_{p \in \Omega_v} \sum_k d_{pv}^{ki} \cdot S_{p,j,k}}{\sum_{p \in \Omega_v} \sum_k d_{pv}^{ki}}. \quad (4.36)$$

The compartment-specific contributions are substituted into the LMMSE framework via equations (4.31) and (4.33). Furthermore, a spatially isotropic Gaussian filter
denoises the model residual $R_{v,j}$. Finally, the regularized DWIs are synthesized based on the filtered constituents as follows

$$\hat{S}_{v,j} = \sum_{i=1,2} \hat{S}_{v,j,i} + \hat{R}_{v,j}$$

(4.37)

Subsequently, the regularized DWIs $\hat{S}_{v,j}$ are used for estimation of the diffusion properties of the mildly constrained dual tensor model using ARD.

### 4.3 Validation on Synthetic

In Section 4.3.1, we evaluate the selectivity of the proposed tensor similarity measure (4.16) and determine the value for the scaling parameter $\beta$. Subsequently, in Section 4.3.2, we illustrate the efficiency of our filtering approach on simulated DWIs.

#### 4.3.1 Tensor Similarity

The estimated diffusivity profile and orientation of the principal diffusivity typically vary along a fiber due to noise on the acquired DWIs. To prevent filtering across fiber boundaries, we restrict the filter kernel both in terms of spatial proximity and diffusion shape. The latter is accomplished by a measure of tensor similarity as computed in (4.35), which restricts the variations in the permitted diffusion shape between neighboring voxels within one compartment. The selectivity of the proposed tensor similarity was validated on synthetic data of fiber crossings in which either the angle between the fibers or the perpendicular diffusivity of one compartment were varied. The parameters of the synthetic configurations are given in Table 4.1. These parameters were previously used by us [2] and are in agreement with the work of Pierpaoli [33] who reported diffusivities ranging from $3 \cdot 10^{-3}$ to $1.5 \cdot 10^{-3} \text{mm}^2\text{s}$.

Figure 4.2 depicts the TS as a function of the angle between the fibers for three different axial diffusivities. We empirically selected the scaling parameter $\beta = 0.1$ to balance a high similarity for tensor fluctuations due to natural diffusivity variations and noise with a low similarity for tensors of a crossing or splitting fiber or as a result of fluctuations due to pathology. Likewise, Figure 4.2 (b) shows the tensor similarity as a function of the perpendicular diffusivity of the second fiber compartment for three
different angles between the fibers. For comparison, the tensors’ inner product (TIP), c.f. [22], is also included as an alternative measure of tensor similarity.

### Table 4.1: Diffusivity parameters of synthetic dual tensor configurations ($\lambda_i$ with unit of $10^{-3} \text{mm}^2\text{s}^{-1}$)

<table>
<thead>
<tr>
<th>Eigenvalues</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\lambda_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tensor-1</td>
<td>1.4</td>
<td>0.6</td>
<td>0.3</td>
</tr>
<tr>
<td>(fixed)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tensor-2</td>
<td>[0.6, 1.0, 1.4] (Figure 4.2(a)); 1.4 (Figure 4.2(b))</td>
<td>0.6 (Figure 4.2(a)); [0.05 – 1.4] (Figure 4.2(b))</td>
<td>0.3 (Figure 4.2(a)); $\lambda_1 = \lambda_2$, (Figure 4.2(b))</td>
</tr>
<tr>
<td>(variable)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.2 Tensor Similarity (TS, solid lines) with scale parameter $\beta = 0.1$ and Tensor Inner Product (TIP, dashed lines) as a function of the angular divergence between two fibers (a) and as a function of the perpendicular diffusivity of the second fiber (b). (a) The three colors reflect different axial diffusivities of the second fiber; the first fiber is modeled by $\lambda_1 = 1.4, \lambda_2 = 0.35, \lambda_3 = 0.3$ ($10^{-3} \text{mm}^2\text{s}^{-1}$); the perpendicular diffusivities of both fibers are equal $\lambda_1 = 0.3, \lambda_2 = 0.3$ ($10^{-3} \text{mm}^2\text{s}^{-1}$). (b) The diffusivities of the first fiber as in (a); the axial diffusivities of both fibers are equal; the perpendicular diffusivities of the second fiber both take the value indicated horizontally.

Figure 4.2(a) shows that our similarity measure decreases rapidly with increasing angle between the fibers, whereas TIP yields a much flatter profile. Particularly, TIP does not approach zero, even when the fibers are orthogonal. Similarly, Figure 4.2(b) demonstrates, for multiple angular divergences, a rapid decreasing similarity when the difference in radial diffusivity increases. Again, TIP yields a much flatter profile.

### 4.3.2 Validation using phantom DWIs

Phantom DWIs were created such that the top part contained crossing fiber structures and the bottom part a single fiber structure (see Figure 4.3(a) for an example). The simulated DWIs were created via (4.20), using $b = [1000, 2000] \text{mm}^2\text{s}$ and 64 gradient
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

directions per \( b \)-value. The simulated DWIs of the phantom were corrupted by Rician noise with SNR \( \left( \frac{S_0}{\sigma} \right) = 15 \) [22]. In the top part the angle between the tensors was set to 90 degrees; the volume fractions were \( f_1 = 0.4, f_2 = 0.5, f_{iso} = 0.1 \) and the diffusivities as given in Table 4.2. In the bottom part the volume fractions were \( f_1 = 0.9, f_2 = 0.0, f_{iso} = 0.1 \). Notice that 64 gradient directions is approximately three times lower than the required number for precise estimation of a mildly constrained dual tensor model using MLE [2].

Table 4.2: Model parameters for generating synthetic data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{1,</td>
<td></td>
<td>} )</td>
</tr>
<tr>
<td>( \lambda_{1,\perp} )</td>
<td>Radial diffusivity of fiber 1</td>
<td>0.15</td>
</tr>
<tr>
<td>( \lambda_{2,</td>
<td></td>
<td>} )</td>
</tr>
<tr>
<td>( \lambda_{2,\perp} )</td>
<td>Radial diffusivity of fiber 2</td>
<td>0.4</td>
</tr>
<tr>
<td>( \alpha_{1,2} )</td>
<td>Angles defining the plane</td>
<td>random</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>Mean orientation in the plane</td>
<td>0.8\pi</td>
</tr>
<tr>
<td>( \alpha_4 )</td>
<td>Angle between the crossing fibers</td>
<td>0.5\pi</td>
</tr>
<tr>
<td>( f_1 )</td>
<td>Volume fraction of fiber 1</td>
<td>variable</td>
</tr>
<tr>
<td>( f_{iso} )</td>
<td>Volume fraction of isotropic part</td>
<td>0.1</td>
</tr>
<tr>
<td>( f_{iso} )</td>
<td>Diffusivity constant</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Figure 4.3(b) and (c) represent the DWIs after denoising by the original LMMSE filter and the proposed filtering framework, respectively. Furthermore, Figure 4.3(d)-(f) visualize the outcomes of ARD with the mildly constrained DTM to the noisy input data (d), the LMMSE filtered data (e), and data filtered by the proposed method (f). Figure 4.3(d)-(f) visualize the diffusion ellipsoids that represent the estimated tensor components in which the size of the ellipsoid is proportional to the corresponding volume fraction and the color represents the orientation of the principal eigenvector: red for left to right, green for front to back, blue for top to bottom.
Figure 4.3 A simulated DWI of a transition between a crossing fiber region (top) and a single fiber region (bottom) of SNR = 15 before denoising (a), after DWI filtering by the LMMSE (b) and after filtering by the proposed framework (c). (d,e,f) Representations of the mildly constrained dual tensor model obtained by ARD applied to the corresponding data in (a,b,c). The ellipsoids represent both the shape and orientation of the local diffusion profile.

Clearly, the voxels in the crossing region (top part of the image) reflect large variations in model type (one versus two tensor), tensor shapes and orientations, especially without denoising (Figure 4.3(d)). This is due to the low SNR, which makes that Jeffreys prior becomes rather harsh after which ARD forces one of the anisotropic volume fractions to zero, thereby promoting a single tensor model. In other words, there is no support for the more complex DTM. After filtering by LMMSE, the dual tensor model was chosen more consistently. The LMMSE filtering reduces the number of voxels for which only a single compartment was selected away from the boundary, but does not substantially reduce this near the boundary, as expected. Figure 4.3(f) demonstrates improved model selection near the boundary without sacrificing performance away from the boundary.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

### 4.4 Experiments and Results

#### 4.4.1 Data

The experimental data consists of DW-MRI from 26 healthy controls that participated in a study into the effects of HIV on the brain [34]. The images were acquired on a Philips Intera 3.0T MRI scanner (Philips Intera, Philips Healthcare, Best, The Netherlands) by means of a spin echo EPI sequence. An echo time of TE=92 ms and a repetition time of TR=7725 ms were used. The diffusion weighting was along 64 three-fold tessellated icosahedral gradient directions, with two \( b \)-values; \( b = [1.0 \text{, } 2.0] \times 10^3 \text{ mm}^2/\text{s} \). Two non-diffusion-weighted images were obtained, the average of which was used in the calculations as the \( S_0 \) image. Fifty-five axial slices of the whole brain region were acquired with an imaging matrix of \( 112 \times 110 \) voxels. Slices of \( 112 \times 112 \) voxels were reconstructed with a voxel-size of \( 2.0 \times 2.0 \times 2.0 \text{ mm}^3 \). Deformations induced by eddy currents were corrected using an affine registration in the phase encoding direction [35]. In addition, the subjects’ head motion during acquisition was corrected by a rigid registration between the DWIs and the \( S_0 \) images. Subsequently, the dataset was filtered by the original closed-form LMMSE [5] and the proposed structure-adaptive LMMSE, respectively. Our DWIs were acquired using a ‘SENSE’ protocol, so that equation (4.31) reduces to the form corresponding to Rician-distributed noise, i.e. \( N_{\text{coil}} = 1 \). The acquired data does not comply with the requirements for accurate and precise estimation of parameters of the mildly constrained DTM since only 128 DWIs were acquired instead of the required number of ~180 as reported by Caan et al. [2].

#### 4.4.2 Validation on experimental data

The aim of the experiments in this section is twofold: to demonstrate the performance of the individual steps and to validate the entire denoising and parameter estimation framework on experimental data.

ARD of the mildly constrained DTM vs ARD of the severely constrained DTM applied to unfiltered data. Figure 4.4 shows that direct application of the mildly constrained DTM by ARD fails. This can be concluded by inspection of the shape and orientation of neighboring diffusion tensors in the selected regions of interest. ROI-1 contains the central part of the corpus callosum (CC): a region dominated by a single
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

fiber in each voxel. It shows that the severely constrained model shows considerably less variation than the mildly constrained model. ROI-2 shows the crossing of the CC with the corticospinal tract. Inspection of ROI-2 reveals that the mildly constrained model fails to find two tensors in a large number of the voxels, whereas the severely constrained model yields well-pronounced tensors of similar shape and orientation across the entire crossing region.

Figure 4.4 A typical result of dual tensor modeling by ARD using severe constraints (b, c) and mild constraints (d, e) in the ROIs shown on the single tensor FA map in (a). ROI-1 is a region of interest around the corpus callosum; ROI-2 is a region where the corpus callosum (left-to-right) crosses the corticospinal tract (top-to-bottom). The color (as in Figure 4.3) indicates the fiber orientation, red for left to right, green for front to back, blue for top to bottom.
Figure 4.5 (a,b) Tensor fields sorted by applying front evolution to the estimated dual tensor model from the crossing in ROI-2 of Figure 4.4(a); (c-f) Tensor similarity $d_{ij}$ between the 'boxed' tensor in field $i$ and the neighboring tensors in field $j$. (c) $d_{11}$; (d) $d_{12}$; (e) $d_{21}$; (f) $d_{22}$. 
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

Tensor similarity for generating structure-adaptive filter kernels. The two tensors inferred in the heterogeneous region from Figure 4.4(b) (ROI-2) were sorted by means of ‘front evolution’ [36] [37]. This method was originally introduced to ensure the continuity of tensor fields in dual-tensor analyses [16]. Figure 4.5(a) and (b) show the sorting outcome. Note that the tensors with approximately the same orientation indicated by the colors of the ellipsoids are assigned to the same group. Furthermore, Figure 4.5(c) and (e) show the tensor similarity of the boxed tensor in Figure 4.5(a) to the neighborhood tensors from Figure 4.5(a), i.e. the same group, respectively the neighborhood tensors from Figure 4.5(b), i.e. the other group. Similarly, Figure 4.5(d) and (f) show the tensor similarity of the central boxed tensor in Figure 4.5(b) to the neighborhood tensors from Figure 4.5(b), respectively Figure 4.5(a). The figures illustrate that tensors associated with the same shape and orientation are assigned larger weight in (4.17), simply because they have a larger tensor similarity to the boxed tensor. As such, the smoothing kernel selects tensors belonging to the same tract as the selected tensor, thereby steering the filtering along tracts.

DWI decomposition, structure-adaptive denoising and DWI reconstruction. Figure 4.6 illustrates three steps of the proposed framework: decomposition of a DWI into compartment-specific DWIs, structure-adaptive filtering, and DWI reconstruction from the filtered components. Figure 4.6(a) shows an original DWI (top left) and the reconstructed DWI (bottom left) by taking the weighted sum of the filtered components as in (4.18). Additionally, the component specific contributions to the DWI are depicted prior to (top images) and after filtering (bottom images). Furthermore, Figure 4.6(b) shows two DWIs filtered by means of the classical LMMSE approach [5] and the proposed framework. Clearly, the DWIs filtered by both frameworks (middle and right columns) are smoother than the original DWIs, especially in the grey matter regions. However, note how the fiber structures are much better preserved in the DWI filtered by means of the proposed framework.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

Figure 4.6 (a) Illustration of the decomposition of a typical DWI into compartment-specific DWIs (top row, from left to right) and the synthesis of the filtered compartment-specific DWIs into the reconstructed DWI (bottom row, right to left). (b) Illustration of DWIs before and after filtering: before filtering (left), after classical LMMSE (middle), after the proposed structure-adaptive framework (right). The two depicted DWIs are typical examples of two coronal slices of the brain using different gradient orientations.
Influence of DWI denoising on the complexity after ARD. Figure 4.7 shows the component-specific FA in an axial slice estimated from the original data (left), after filtering by the classical LMMSE filter (middle) and after filtering by the proposed framework (right). The top images depict the FA of tensor component 1 and the bottom images show the FA of tensor component 2. We applied a threshold of 0.05 to the volume fractions to discard negligible compartments as in [23].

![Figure 4.7 Component-specific FA maps (FA1 and FA2) obtained by ARD of the mildly constrained dual tensor model, after sorting the tensors using front evolution. Results without DWI filtering (left), after the classical LMMSE filter (middle) and after the proposed structure-adaptive framework (right).](image)

The FA maps estimated from the original data (left) are noisier due to larger variations in the parameter estimation and fluctuations in model selection. An instance of the latter problem is indicated by the yellow circle. In this crossing region, ARD predominantly selects a single tensor model because the noise decreases the support for the model with a higher complexity. The FA map of the second component is black.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

on this location, because there is no second tensor. The classical LMMSE filter improves the SNR of the dataset (Figure 4.7, middle). However, the filter still does not facilitate full support for the complex model (see the yellow circle), which is signified by the large number of voxels without a second tensor. However, the FA maps generated after filtering by the proposed method (Figure 4.7, right), demonstrate both improved continuity of FA estimation and clearly more voxels with a second tensor in the crossing structure.

4.4.3 Application to a demo neuroimaging problem

The usefulness of our method will be assessed with a case study into the relation between white matter atrophy and age. The 26 control subjects from the HIV study were equally divided into two age groups. The younger group consisted of 13 subjects, mean age: 56.5, standard deviation: 1.2. The older group also consisted of 13 subjects, mean age: 68.9, standard deviation: 3.7.

ARD of the mildly restricted tensor model was applied to the original data, after filtering by the classical LMMSE method and after filtering by the proposed framework. Subsequently, the FA was calculated from the estimated tensor(s). All FA maps were registered to the FMRIB58_FA standard-space using FNIRT (FSL version 5.0.7.) Thereafter, the differences between the two age groups were analyzed by means of the extended TBSS approach [37], which enhances the classical single tensor TBSS to voxels comprising crossing fibers. Figure 4.8 shows the outcome. The red-yellow marked regions in Figure 4.8 identify regions where significant differences were found. Observe that both the classical LMMSE and our structure-adaptive filtering framework yield increased significant differences compared to the approach without filtering. What is more, our structure-adaptive LMMSE framework (bottom row of Figure 4.8) increases both the size and the number of regions with significant differences compared to the classical LMMSE filter, which we attribute to a more accurate and precise model estimation.

Figure 4.9 shows the outcome of the same experiment after discarding respectively 30% and 50% of the DWIs prior to any processing. First, 30% of the DWIs were discarded, so that only 85 out of 128 gradient directions were left: the DWI corresponding to
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

every third gradient was deleted. The top row of images in Figure 4.9 (a) demonstrates that without filtering the significant regions are further restricted, whereas the outcome after filtering by our framework is less affected (the bottom row of Figure 4.9 (a)). Secondly, half of all DWIs were discarded in a similar way, leaving only 64 out of 128 gradient directions. Clearly, the unfiltered results in top row of Figure 4.9 (b) demonstrate a further reduction of significantly different regions. Alternatively, the filtered results in the bottom row Figure 4.9 (b) still contain many regions that were also found to be significant in the original data set. Notice also that again the classical LMMSE filter gives slightly less significant differences than our structure adaptive version. An even further reduction of the number of DWIs shows a steady decline of the regions with significant differences, and marks the limit of the proposed method.

Figure 4.8 Extended TBSS analysis of differences in component-specific FA maps between two age groups of healthy controls. FA maps obtained by ARD of the mildly constrained dual tensor model applied to: unfiltered DWIs (top row), after filtering by the classical LMMSE (middle row) and after filtering by our structure-adaptive framework (bottom row). The red-yellow marked regions identify regions where the differences in FA maps are significant between the two age groups.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

Figure 4.9 Extended TBSS analysis of differences in component-specific FA maps, obtained from incomplete data, between two age groups of healthy controls. FA maps obtained by ARD of the mildly constrained dual tensor model applied to 30%-reduced DWIs (a): unfiltered DWIs (top row), after filtering with the classical LMMSE (middle row) and after filtering by our structure-adaptive framework (bottom row). (b) as in (a), but now applied to a 50%-reduced set of DWIs. The red-yellow marked regions identify regions where the differences in FA maps are significant between the two age groups.
4.5 Discussion and Conclusion

We proposed a novel, compartment-specific structure-adaptive LMMSE filter to suppress the noise in DWIs at complex fiber geometries while maintaining the underlying diffusion properties as well as the boundaries of the compartments. Denoising of the DWIs enhances the applicability of a mildly constrained dual tensor diffusion model in data sets with a low SNR or with a limited number of DWIs as frequently occurs in neuroimaging problems. To obtain an initial guess of the structural information, we fit a severely constrained tensor model to the data by ARD using Jeffreys prior. This structural information is used to decompose the DWIs into compartment-specific contributions and an isotropic (residual) fraction. Subsequently, each of the compartment-specific contributions is filtered by a space-variant, structure-adaptive kernel. The filter coefficients reflect the tensor similarity between the neighborhood tensors and the central tensor. After filtering, regularized DWIs are reconstructed from the filtered DWI contributions. Finally, these regularized DWIs serve as input for an improved estimation of the full diffusion shape using a mildly constrained dual tensor model.

The tensor similarity measure aims to select the tensors of neighboring voxels, which belong to the same fiber. The measure was evaluated on synthetic data, which demonstrated to which extent it supported variations in diffusion shape and orientation. Synthetic data was also used to show the edge-preserving, regularizing effect of the filter on the final estimates of tensor shape and orientation.

Experimental data were employed to first demonstrate that the severely constrained dual-tensor could be reliably estimated from these data (Figure 4.4 (a)). Subsequently, the effectiveness of the similarity measure was illustrated (Figure 4.5). Thereafter, both the filtering of the DWI contributions and the reconstructed DWIs (Figure 4.7) showed that our framework simultaneously facilitates preservation of structure information and improved SNR.

The proposed framework was also applied to a neuroimaging case study in which differences between two age groups were investigated. It was demonstrated that the noise in the original DWIs limited the size as well as the number of regions with significant changes related to aging. Reversely, it was possible to detect more subtle
alterations of the diffusion shape after filtering. We demonstrated that with our framework we still could find comparable results after discarding thirty or even fifty percent of all data. This may enable the use of limited scan protocols.

There are also a few limitations in our framework. Firstly, we assume a mono-exponential decay along the eigenvectors of the three compartments up to \( b=2000 \) mm\(^{-2}\)s\(^{-1}\). Measuring at higher \( b \)-values will certainly introduce sensitivity to different compartments such as the myelin sheet [38], giving rise to restricted and hindered diffusion [39]. In the latter case, the Gaussian diffusion assumption is no longer valid, which is beyond the scope of this work. As such, we follow [2] and [40]. Secondly, recent studies reported the presence of three-way fiber crossings [41]. In our framework a dual-tensor model is employed to characterize voxels encompassing crossing fibers. The reason behind this restricted choice is the limited SNR in conventional HARDI data with different \( b \)-values. Notice that we aim to reconstruct the full multi-compartment diffusion shape, whereas [41] only tries to recover the fiber orientations. In [2] we showed that estimating a dual rank-2 tensor model already requires HARDI at two \( b \)-values, data of sufficient SNR, and some model restrictions. The latter are needed to ensure stability as the number of model parameters may approach or even surpass the number of degrees of freedom present in the data. Therefore, fitting a triple rank-2 tensor model to the DWIs of voxels encompassing a three-way-crossing will definitely lead to unreliable estimation, even though scan times have been reduced by parallel imaging techniques [6]. It is our empirical finding that triple tensor models still cannot be reliably estimated. At the same time, developing methods to enhance the applicability of triple tensor model will certainly be an important challenge for future research.

The proposed framework can easily be applied to other approaches in which DWIs are modeled as a weighted sum of compartmental contributions. What is more, it may enhance the sensitivity of clinical DTI studies.
Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures

References


Chapter 4. *A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures*


Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures


Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures


Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures


Chapter 4. A structure-adaptive noise filter for accurate estimation of diffusion properties in complex fiber structures
5 Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

Conventionally, a single rank-2 tensor is used to assess the white matter integrity in diffusion imaging of the human brain. However, a single tensor fails to describe the diffusion in fiber crossings. Although a dual tensor model is able to do so, the low signal-to-noise ratio hampers reliable parameter estimation as the number of parameters is doubled.

We present a framework for structure-adaptive tensor field filtering to enhance the statistical analysis in complex fiber structures. In our framework, a tensor model will be fitted based on an automated relevance determination method. Particularly, a single tensor model is applied to voxels in which the data seems to represent a single fiber and a dual-tensor model to voxels appearing to contain crossing fibers. To improve the estimation of the model parameters we propose a structure-adaptive tensor filter that is applied to tensors belonging to the same fiber compartment only.

It is demonstrated that the structure-adaptive tensor-field filter improves the continuity and regularity of the estimated tensor field. It outperforms an existing denoising approach called LMMSE, which is applied to the diffusion-weighted images. Track-based spatial statistics analysis of fiber-specific FA maps show that the method sustains the detection of more subtle changes in white matter tracts than the classical single-tensor-based analysis.
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

Conclusion: Thus, the filter enhances the applicability of the dual-tensor model in diffusion imaging research. Specifically, the reliable estimation of two tensor diffusion properties facilitates fiber-specific extraction of diffusion features.
5.1 Introduction

Diffusion-weighted Magnetic Resonance Imaging (DW-MRI) provides unique information on the integrity of white matter (WM) structures [1]. In diffusion tensor imaging (DTI), a single rank-2 tensor model is fitted to the diffusion-weighted images (DWIs), after which features describing the diffusion are extracted from the fitted tensor. However, in voxels with crossing fibers, a single tensor model cannot adequately describe the diffusion process. Previously, we introduced a framework for estimating diffusion properties based on a constrained dual tensor model to solve this problem [2]. However, to avoid overfitting in single fiber bundles, a structure-adaptive approach is needed. In other words, voxels containing a single fiber bundle must be characterized by a single tensor, whereas voxels encompassing crossing fibers should be modeled with two tensors. Such an approach will yield unbiased diffusivity estimates in single fibers and in fiber crossings. However, the inherent noise in DW-MRIs hinders a stable fit of especially a dual tensor model.

In this chapter, a new framework will be presented which enhances the applicability of complex models in a structure-adaptive way. As demonstrated in Figure 5.1, the structural geometry of voxels is tentatively inferred by means of Automated Relevance Determination (see Chapter 2 of this thesis). As such, the diffusion is implicitly modeled by a single or a dual diffusion tensor model. Subsequently, a structure-adaptive tensor filter is proposed to improve the estimated diffusion properties. Particularly, the filter kernel is restricted to tensors belonging to the same compartment. This is highly relevant to complex fiber configurations. We will compare our structure-adaptive tensor filter to denoising in DWI space by the linear minimum mean square error (LMMSE) estimator [3]. Finally, fiber-specific features are extracted, which are expected to improve the sensitivity of statistical analysis such as tract-based spatial statistics (TBSS) [4].
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

Figure 5.1 Flowchart of the framework for structure-adaptive estimation of diffusion properties. ARD and MLE are abbreviations of Automatic Relevance Determination and Maximum Likelihood Estimation respectively. STM and DTM denote respectively the single tensor model and the dual tensor model.

5.2 Methods

5.2.1 Determination of the Fiber Structure

Automatic relevance determination (ARD) as described in Chapter 2 is employed to infer the local geometry by judging the properness of severely constrained dual-tensor model. ARD estimates the posterior probability for a dual-tensor model in every voxel. If the data supports a compartment, its estimated volume fraction will be non-zero. Therefore, if the data originates from a single fiber, the estimated volume fraction of all-but-one of the compartment will become approximately zero (smaller than the employed threshold 0.05). Usually, at most two tensors yield a non-zero volume fraction.
In the following section, the rank-2 diffusion model that is used in our framework is introduced. Subsequently, the structure-adaptive filtering approach is proposed to avoid unstable estimation of the diffusion properties.

### 5.2.2 Diffusion Tensor Models

We assume that the diffusion in fiber tracts is mono-exponential and anisotropic Gaussian. For a voxel with 3D spatial coordinates $v$, the underlying diffusion-weighted signal is initially modeled to contain a contribution of up to two fiber bundles and an isotropic component [2]

$$ S_{v,\theta,j} = S_{v,0} \left( \sum_{i=1,2,iso} f_i \exp(-b_j g_j^T D_{v,i} g_j) \right), $$

(5.38)

where $S_{v,\theta,j}$ is the modeled signal value given the model parameter vector $\theta$ and $S_{v,0}$ the signal without diffusion weighting. $b_j$ denotes the amount of diffusion weighting in gradient direction $g_j$; $D_{v,1}$ and $D_{v,2}$ are two rank-2 diffusion tensors, $D_{v,iso}$ represents the amount of isotropic diffusion (i.e., $D_{v,iso} = I_{3x3}$) and $f_i$ the volume fraction corresponding to $D_{v,i}$, which satisfy the relation $f_1 + f_2 + f_{iso} = 1$. The single tensor model is obtained by setting $f_2 = 0$ in (5.38). Besides the constraints of volume fraction, in the dual-tensor model we also set the largest eigenvalue of $D_{v,1}$ and $D_{v,2}$ to be equal, i.e. the axial diffusivity in two crossed fiber bundles are equal, which is referred as mildly constrained dual-tensor model in Chapter 4. Also as discussed in Chapter 4, for the severely constrained dual-tensor model, we further constrain the shape of $D_{v,1}$ and $D_{v,2}$ to be equal, i.e. both radial and axial diffusivities are equal. The parameters of the mildly constrained diffusion tensor models are estimated by maximum likelihood estimation (MLE) [2].

### 5.2.3 Structure-adaptive tensor filtering

Denoising in DWI space, such as the LMMSE approach [3], is often used to improve the performance of the tensor estimation. However, bluntly denoising without knowing the local fiber structure smoothen the DWI-space and changes the data such that the subsequent tensor estimation becomes biased. Alternatively, we propose a structure-adaptive fitting of tensor models to the original dataset followed by denoising of the estimated (multi-valued) tensor field. This approach may effectively preserve the
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

structural properties of the data. Currently, there is no established approach for tensor field filtering or regularization in a complex tensor field, i.e. fields with multiple tensors modeling the diffusion. We will introduce such a filtering method, which uses a measure for tensor similarity.

5.2.3.1 Tensor Similarity

Tensors characterizing nearby voxels of the same fiber bundle are assumed to have a similar shape as well as orientation. Reversely, tensors describing different fiber bundles will usually display a different fiber orientation. Furthermore, tensors at and across the borders of fiber bundles usually display a distinctively different diffusion shape. Therefore, we propose to quantify shape and orientation differences in a single measure for tensor similarity:

\[
d_{p\rho}^{ki} = \exp \left( -\frac{\left| D_{p,k}^{sc} - D_{\rho,i}^{sc} \right|^2}{\alpha |A|} \right) \quad (k=1|2; i=1|2),
\]

where \( d_{p\rho}^{ki} \) represents the calculated similarity between the \( i^{th} \) severely-constrained tensor in voxel \( \rho \) (i.e., \( D_{\rho,i}^{sc} \)) and the \( k^{th} \) severely-constrained tensor in voxel \( \rho \) (i.e., \( D_{\rho,k}^{sc} \)), which are estimated from ARD; \( |A| \) denotes the Frobenius norm of \( A \) :

\[
|A| = \left( \sum_{m,n} a_{mn}^2 \right)^{1/2} \quad \text{with} \quad a_{mn} \text{ being an element of } A; \alpha \text{ is a scaling parameter.}
\]

The scaling parameter allows us to define what variation is expected due to stochastic variations (noise) and what is seen as a systematic difference (another fiber bundle). We extensively validated the proposed tensor similarity approach in synthetic data (see Figure 5.2), which was simulated in agreement with the work of Pierpaoli [10]. From these experiments we concluded that the value of \( d_{p\rho}^{ki} \) decreases rapidly when either the orientation difference or shape divergence increases. Therefore, we employ the measure of equation (5.39) as structure-adaptive weight as it restricts the footprint of the tensor field filter to tensors of the same population. In the experiments section, the scaling parameter \( \alpha \) is set 0.1. Figure 5.2 shows how the tensor similarity varies with \( \alpha \), the angle between tensors and the first eigenvalue of one of the tensors.
5.2.3.2 Structure-adaptive filtering in Tensor Space

The objective is to denoise the multi-valued tensor field while preserving the underlying true geometries. In our structure-adaptive framework, the diffusion in one voxel might be modeled by one or by two tensors. In the latter case each tensor belongs to one individual fiber. In order to restrict the averaging to tensors of the same population, we apply a weighted tensor averaging, which is inspired by the well-established bilateral filter. Note that the product of the aforementioned tensor similarity and a spatial proximity kernel gives the weights:

\[
\hat{D}_{v,i} = \frac{\sum_{p\in\Omega_v} \sum_{k} k_{pu} d_{pu}^{ki} \cdot D_{p,k}}{\sum_{p\in\Omega_v} \sum_{k} k_{pu} d_{pu}^{ki}}, \quad (5.40)
\]

with

\[
k_{pu} = \exp \left( -\frac{|\rho - \psi|^2}{2\sigma_k^2} \right), \quad (5.41)
\]

where \(\Omega_v\) represents the neighborhood of voxel \(v\), \(d_{pu}^{ki}\) is defined in (5.39), and \(k_{pu}\) a Gaussian kernel with standard deviation \(\sigma_k\). In such a way, the tensors are essentially regularized based on the structure-adaptive information within a restricted
Chapter 5. Estimating diffusion properties in complex fiber configurations based on
structure-adaptive multi-valued tensor-field filtering

neighborhood. In the results section, the structure-adaptive tensor field filtering will be
compared to LMMSE denoising of the DWIs.

5.3 Experiment and Results

5.3.1 Data
The tensor filtering method was validated on data of 46 control subjects taken from a
study into white matter abnormalities in males with suppressed HIV-infection [6]. The
acquisition parameter settings were: TE/TR=92/7725 ms; FOV 128 × 128 (RO ×
PE); matrix 112 × 110 (RO × PE); slice thickness 2.0 mm, 55 slices; voxel size 2.0 ×
2.0 × 2.0 mm³; b-values: 1000 and 2000 s/mm².

5.3.2 Visual comparison of three modeling strategies
Three approaches were envisioned for estimating the diffusion properties based on the
rank-2 tensor models: (I) direct application the ARD method (i.e. no filtering); (II)
denoising in DWI-space by LMMSE [3] followed by ARD; (II) the proposed
approach consisting of ARD followed by structure-adaptive tensor averaging. Figure
5.3(a) shows ellipsoids representing the estimated diffusion as such: without filtering
(left), using the LMMSE filtering approach (middle), and based on the proposed
structure-adaptive tensor-field filtering (right). Figure 5.3(b) and (c) are zoomed-in
versions of ROI-1 and ROI-2 from Figure 5.3(a) showing details of the outcomes.
In the left images of Figure 5.3(a)-(c), the large variation in both shape and orientation
of diffusion ellipsoids indicates that the rank-2 tensor models were not reliably
estimated. The middle images of Figure 5.3(a)-(c) show that the continuity of
diffusion ellipsoids is improved at the expense of a loss of anisotropy of the tensors.
The right images in Figure 5.3(a)-(c) show further enhanced continuity/regularity
particularly of fiber-specific ellipsoids.
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

Figure 5.3 (a) Visualization of estimated rank-2 tensors using ellipsoids. The colors of the ellipsoids indicate the fiber orientation (principal axis of the rank-2 tensor): red for left-to-right, green for front-to-back, and blue for top-to-bottom. Results were generated directly from the original dataset (left, i.e. without filtering), data filtered in DWI-space by the LMMSE method (middle) and filtered tensor-field by the proposed framework (right); (b) and (c) show details of respectively ROI-1 and ROI-2; the ordering of the three figures is the same order as in (a). ROI-1 contains a single fiber – the Corpus Callosum (red ellipsoids). ROI-2 contains a crossing region between Corpus Callosum (red ellipsoids) and Corticospinal Tract (blue ellipsoids).

5.3.3 TBSS analysis based on classical and structure-adaptive DTI features

The application of our framework to a neuroimaging problem is demonstrated by a fiber-specific TBSS analysis of the control subjects from the HIV study [6] (see section 5.3.1) based on fractional anisotropy (FA) and volume fraction. Furthermore, the results of a conventional TBSS analysis based on the classical, single-tensor approach are included to compare with.
FA as well as volume fraction maps are widely used to detect WM changes related to aging [5] [7]. A first step to analyze such an effect was to register the data to the MNI152 standard space using the FNIRT method in FSL [8] [9] (FSL toolbox version-5.0.7). Subsequently, age-related effects in WM were analyzed by means of the classical TBSS technique, i.e. the single-tensor analysis, as well as by an extended TBSS method for dual component DTI analysis. The dual components were generated by means of the proposed framework. Compared to the classical TBSS method, the extended TBSS technique [5] employed ‘front evolution’ to avoid swapping of two anisotropic components. Subsequently, the TBSS analysis was performed based on the FA as well as the volume fraction of each component.

Figure 5.4 shows the result of the classical TBSS analysis for the single-tensor FA map. The blue colors represent regions where the FA has a significant negative correlation with age. Similar to Figure 5.4, Figure 5.5(a) shows the extended TBSS analysis on the two-tensor FA maps. In these images the blue regions represent regions where either one or both tensors correlate negatively with age. Furthermore, Figure 5.5(b) contains the TBSS outcome on the volume fraction of the isotropic part from the two-tensor model. Here, volume fractions having a significant, positive correlation with age are depicted in red. Notice that this reflects a negative correlation of $f_1 + f_2$ with age,
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

since $f_{iso} = 1 - f_1 - f_2$. We did not observe any significant positive correlation of FA with age or a negative correlation of the volume fraction with age.

![Figure 5.5](image)

Figure 5.5 TBSS analysis of features from a dual tensor representation generated by the proposed framework: analysis of FA (a) and $f_{iso}$ (b). In (a), regions where the FA of either tensor has a significant negative correlation with age are indicated in blue; in (b), red denotes regions with a significant, positive correlation of $f_{iso}$ with age.

Clearly, TBSS analysis of the dual tensor data extends the regions in which significant correlations are encountered with the classical analysis. Furthermore, it facilitates a more sophisticated characterization of which feature changes with age: FA or volume fraction. Additionally, the two tensor maps facilitate to distinct which component (tract) changes with age, especially in regions with crossing fiber structures.
Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering

5.4 Discussion and Conclusion

The proposed structure-adaptive tensor-field filter offers regularization of multi-valued tensor fields. The method preserves the structure of the data in voxels comprising of complex fiber configurations such as crossing fibers. The framework enhances the applicability of a dual tensor model, which facilitates structure-adaptive estimation of diffusion properties. Compared to the conventional DWI space filtering, the proposed framework enhances the continuity of the estimations. The application of the framework in TBSS analyses enables sensitive fiber-specific diffusion statistics in complex white matter structures. Extended TBSS of the dual tensor FA maps show more regions with a significant negative correlation by either one or both FA maps compared to classical TBSS based on a single tensor’s FA map. TBSS based on the volume fraction reveals that the fraction of the isotropic component’s volume fraction correlates positively with age, which might be an indication of white matter atrophy.

References


Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering


Chapter 5. Estimating diffusion properties in complex fiber configurations based on structure-adaptive multi-valued tensor-field filtering
6 Conclusion

Structural hindering of diffusion yields geometrically-constrained diffusion, which gives rise to a distinct three-dimensional diffusion profile. The 3D diffusion profile of water in the human brain can be probed in vivo by diffusion-weighted MRI. Since it is assumed that pathologies change the structural influence on diffusion and its 3D profile, DW-MRI provides the means to indirectly assess the integrity of white matter in a non-invasive way.

Conventionally, the diffusion is modelled by a single tensor. Features such as the Fractional Anisotropy and the Mean Diffusivity calculated from this diffusion tensor are often used to characterize the white matter. However, a single tensor is not appropriate for the characterization in two-way and three-way crossings. The goal of this thesis was to improve the estimation of higher-order diffusion models, i.e. more complex diffusion shapes, to neuroimaging data acquired with different acquisition protocols, without causing overfitting in voxels encompassing simpler geometries.

In Part-I of this thesis, new model selection approaches were proposed to accurately and precisely estimate diffusion properties of white matter based on multi-tensor representations.

In Chapter 2, we developed a new framework using automated relevance determination (ARD) for estimating the parameters of a constrained dual-tensor model. Basically, the proposed ARD enabled data-adaptive estimation of diffusion properties to overcome the well-known overfitting problem. This occurs when fitting a ‘complex’ model to ‘simple’ data. Maximum Likelihood Estimation (MLE) of the constrained dual-tensor model yields the initial values of the parameters at the start of the ARD procedure. Subsequently, ARD automatically estimated to what extent the diffusion in a voxel should be modelled by rank-2 dual-tensor model. ARD estimates the a-posteriori probability using Jeffreys prior. Jeffreys prior uses Fisher’s information
matrix to compute the prior probability for any combination of parameter values. The probability represents the degree to which a configuration of fibers under a certain angle and with certain diffusivities can be reliably estimated given the acquisition protocol and signal-to-noise ratio. In a single fiber voxel, the dual-tensor model was made to degenerate to a ‘single-tensor’ model by forcing the volume fraction of one tensor compartment to zero. The complexity of the model was implicitly inferred with ARD in a Bayesian probabilistic manner.

The proposed framework offered accurate and precise estimation of diffusion properties in single and dual fiber regions. Tract-based spatial statistics demonstrated a higher sensitivity in detecting age-related white matter atrophy using the ARD approach compared to MLE of either a single or a dual tensor model.

A limitation of the ARD framework is that the employed mildly constrained dual-tensor model primarily aims to characterize two-way crossings (and simpler configurations). Recent studies reported evidence for the existence of a three-way crossing of fiber bundles. It is not trivial to extend our ARD method to three-way crossings by employing a three-tensor model. The increased complexity of such a model will even make the ARD method prone to overfitting. Furthermore, a drawback of the ARD approach is the high computational complexity, which results in very long processing time.

In Chapter 3, we introduced a maximum a-posteriori (MAP) estimator to characterize the diffusion in three-way crossings based on a triple-tensor model. The MAP estimator involved a mild prior to prevent that the problem was ill-posed. Additionally, a new model selection technique quantified the extent to which candidate models were appropriate, i.e. single-, dual- or triple-tensor model. The model selection technique combined the total Kullback-Leibler divergence and information complexity (ICOMP-TKLD). ICOMP-TKLD balanced the trade-off between the goodness of fit and the complexity brought by this diffusion model. The model selection used Fisher’s information matrix, so that it implicitly adapted to the data by incorporating knowledge about the acquisition protocol and signal-to-noise ratio.
Chapter 6. Conclusion

The aforementioned prior of our MAP estimator was shown to enhance the precision of the parameter estimation, without introducing a bias. The proposed framework enhanced the sensitivity of statistical analysis of differences between left and right-handed subjects on publically accessible data from the Human Connectome Project.

The estimation of complex diffusion models from diffusion-weighted magnetic resonance images (DW-MRI) is hampered by a low signal to noise ratio (SNR). PART-II of this thesis studied new methods to filter the raw Diffusion Weighted Images and the Diffusion Tensor data to improve the SNR and in turn the estimation of the diffusion properties.

In Chapter 4, we proposed a structure-adaptive, compartment-specific version of the Linear Minimum Mean Square Error (LMMSE) estimator to suppress the noise in DWIs. Initially, we fitted a strictly constrained dual tensor model (DTM) to the DWI data using ARD. The resulting DTM was used to decompose the DWIs into compartment-specific contributions to the DWI signal. Subsequently, these DWI contributions were adaptively filtered, in which neighborhood data was weighed based on a tensor similarity measure. Thereafter, the regularized DWIs were reconstructed from the filtered DWI contributions. Finally, the regularized DWIs were used as input to estimate a mildly constrained dual tensor model.

The results demonstrated that the method preserved the edges of fiber compartments, while denoising the data. The benefit of the proposed framework was shown through a case study into the relation between white matter atrophy and aging. It was found that the noise on the DWIs limited the regions with significant differences between two age groups. Instead, subtle differences could be identified after DWI filtering. Discarding up to fifty percent of the DWIs still yielded comparable results by our framework. Consequently, our method may permit the use of ‘simpler’ scan protocols and even enhance the sensitivity in clinical DTI studies.

In Chapter 5, we proposed a structure-adaptive, bilateral tensor-field filter, which was steered both by spatial distance and the similarity between diffusion tensors. The
method restricted the tensor field denoising to tensors of the same population, even in complex fiber structures, such as crossing and touching fiber bundles. The filter facilitated reliable, fiber-specific estimation of diffusion properties, such as FA. TBSS analysis on the correlation between white matter changes and age in HIV-infected patients demonstrated that this enables more sensitive detection of changes in white matter properties.

Future work

Several aspects from the methods developed in this thesis could be extended.

1. **Real-time registration.** The computational time is one of the issues that was not elaborately considered in this thesis. The processing time of ARD involving 5000 iterations is approximately 0.35 s/voxel and the processing time of the ICOMP-TKLD framework that merely requires fitting of multi models is about 0.15 s/voxel on a conventional computer. Real-time performance is an issue, particularly for practical application. Clearly, parallel processing on multiple GPU’s could an obvious way to achieve this. Alternatively, the processing time might be further reduced by careful initialization of the estimations close to the global minimum.

2. **Modelling improvement.** The number of degrees of freedom (DOF) of the proposed models is large. Therefore, data of high quality is required. We introduced methods to improve the SNR of the underlying data both in DWI space (Chapter 4) and in tensor space (Chapter 5). However, as pointed out in [1], such approaches have one drawback: the errors in the filtering will influence the estimation. Therefore, in [1], simultaneous smoothing and estimation were employed. An important topic of our future work will be to truly integrate the filtering and the model estimation, particularly by using a minimum total-variation framework which essentially trades off the fitness of the multi-tensor model and the neighbourhood compatibility. Also, the modelling might be enhanced by adding data-driven priors or adding more physical constraints.
Chapter 6. Conclusion

Contributions to the neuroimaging community

We developed several methods for improved estimation of diffusion properties in white matter structures. This paves the way to more accurate assessment in clinical studies:

1. Detection of subtle microstructural changes. We validated our methods on small sized datasets to show the enhanced sensitivity for the detection of differences in brain structure. We now aim to apply our framework to study microstructural white matter changes in males with aviremic HIV-infection. Much in the same way our framework is applicable to other clinical DTI studies.

2. Multi-tensor modeling for reconstruction of brain connectivity. In general, brain connectivity refers to a pattern of anatomical links ("anatomical connectivity"), statistical dependencies ("functional connectivity") or causal interactions ("effective connectivity") between distinct parts of the nervous system. The connectivity pattern is formed by structural links such as synapses or fiber pathways, or it represents statistical or causal relationships measured as cross-correlations, coherence, or information flow [2]. DTI has enabled non-invasive investigation and characterization of the white matter architecture of the brain [3] [4]. Moreover, this thesis showed that multi-tensor models provide an effective way to characterize complex fiber structures. Recent studies [5] showed that compared to non-parametric approaches (such as spherical deconvolution [6]) the multi-tensor model facilitates higher accuracy in streamline tractography. Therefore, combining the multi-tensor model and the latest tractography methods might enable new ways to study brain connectivity.

3. Comparison with competing techniques Many competing techniques were proposed that relate to the topics of this thesis. Particularly, several sophisticated representations provided a more accurate description of the diffusion process than the conventional single tensor, e.g. the composite hindered and restricted model of diffusion (CHARMED) [7] [8] high-order tensor [9] and diffusion kurtosis imaging DKI [10] [11]. The CHARMED and DKI approaches both modelled non-Gaussian aspects of diffusion, whereas the multi-tensor model asserts a Gaussian diffusion partitioned over
multiple compartments. As such, an intuitive physical meaning is attributed to the parameters of the dual tensor model (as with the conventional, single tensor model). Also, many model selection methods were introduced in the DWI field, for instance based on constrained spherical deconvolution [6], the Bayesian information criterion (BIC) [12], and the generalization-error [13]. An important difference of our ICOMP-TKLD [14] with CSD is that the model selection criterion is implicitly defined, so that it does not require tuning of a threshold. Further, ICOMP-TKLD measures the model complexity using an estimated probability density function (PDF) of the model parameters. Instead, BIC is only determined by non-estimated factors such as the number of parameters and the sample size. Also, ICOMP-TKLD is essentially a local model selection method since the goodness of fit and model complexity are calculated per voxel. Alternatively, the generalization-error method [13] is a non-local model selection technique.

The introduced techniques were already compared with state-of-art approaches such as AIC, the Likelihood-ratio, and FSL’s-ARD involving a ball and –stick representation. The results showed that the inference from the proposed techniques in several white matter tracts more closely resembled the expected anatomy. Simultaneously, it must be conceded that the benefit of our ICOMP-TKLD and ARD to the detection of disease processes needs to be further investigated.

Diffusion imaging can reveal several aspects related to white matter integrity: (I) locations of alterations; (II) which fiber tract is affected; (III) the exact change in diffusion. Previously, many solutions were already proposed for the first two aspects. Our work focused on all three aspects. Particularly, we improved the accuracy and the precision of the estimation of diffusion properties by carefully modelling complex fiber structures based on multi-tensor representations. What is more, we enhanced the estimation by filtering the data to improve the signal to noise ratio. We anticipate that many diffusion MRI studies can benefit from our work.
References


Chapter 6. Conclusion


Summary

The white matter of the brain contains all the connections between different parts of the grey matter. Many diseases especially affect the brain’s white matter. For instance, the white matter tracts are destroyed in neurodegenerative diseases, such as Alzheimer’s disease. Accordingly, there is a large interest in features of the white matter to understand the pathophysiological mechanisms underlying these diseases. Diffusion MRI enables non-invasive characterization of the white matter architecture by measuring features of local diffusion processes. Importantly, the diffusion is larger along white matter tracts rather than perpendicular to them due to structural hindrance of the myelin sheets that surround nerve cells.

Diffusion MRI measures the diffusion in a large number of directions to yield so-called diffusion weighted images (DWI’s). Conventionally, the diffusion is modelled from the DWI’s by a single ellipsoidal shape, mathematically termed a tensor. Such a tensor reflects the principal directions of diffusion and the associated diffusion lengths. Typically, measures such as the imbalance in diffusion, i.e. the so-called fractional anisotropy, and the mean diffusivity, are calculated from the diffusion tensor to characterize the white matter. However, a single tensor is not appropriate for two-way and three-way crossings of tracts.

The goal of this thesis was to improve the modelling of more complex diffusion shapes. Therefore, a multi-tensor is used. However, a pitfall with such a complicated model is that it is prone to ‘overfitting’, for instance as a dual tensor model is fit to single tract data. In that case, the estimated diffusion features will be inaccurate and/or imprecise. The methods in this thesis aim to adapt the model to the underlying structures: a single tensor for single fiber data and dual tensors for fiber crossings. While doing so, the techniques must cope with different measurement circumstances, for instance a varying signal to noise ratio.

Initially, two different frameworks are proposed to structure adaptively determine diffusion parameters. The first framework (Chapter 2) involves a so-called automated relevance determination (ARD) approach to estimate the parameters of a dual tensor model. The dual-tensor model automatically adapts to single fibers by reducing one of the volume fractions to a near zero value in case there is no support for a second tensor in the data. It is demonstrated that the ARD approach gives a higher sensitivity in detecting age-related white
matter atrophy than standard techniques. A limitation of the ARD framework is that the employed dual-tensor model primarily aims to characterize two-way crossings and simpler configurations. Recent studies reported evidence for the existence of a three-way crossing of fiber bundles. The second framework (Chapter 3) relies on a maximum a-posteriori (MAP) estimator to characterize the diffusion in three-way crossings based on a triple-tensor model. A new model selection technique quantifies the extent to which candidate models are appropriate, i.e. single-, dual- or triple-tensor model. The MAP estimator combined with the model selector is shown to enhance the precision of the parameter estimation, without decreasing the accuracy. The proposed framework improved the sensitivity of statistical analysis of differences between left and right-handed subjects from a large a publically accessible dataset.

Unfortunately, the estimation of complex diffusion models is often hampered by a low signal to noise ratio (SNR) of the underlying DWI data. Therefore, two methods are studied to filter the data in order to improve the SNR and in turn enhance the estimation of the diffusion properties. At first (Chapter 4), we propose a structure-adaptive technique to suppress the noise in the underlying DWI data. Initially, the DWI data is decomposed into compartment-specific contributions. Subsequently, these contributions are filtered, after which noise suppressed data is reconstructed from the filtered contributions. Finally, the noise-suppressed DWI data is used to estimate the parameters of a complicated dual tensor model. The results demonstrate that noise limits the regions with significant differences between two age groups. Instead, subtle differences are identified after filtering. Secondly, a method is introduced to estimate a multi-tensor model from the unfiltered diffusion data after which denoising is performed by filtering the tensors (Chapter 5). The method restricts the denoising to tensors of the same population, even in complex fiber structures, such as crossings and touching fiber bundles. An analysis on the correlation between white matter atrophy and age demonstrates that this enables more sensitive detection of changes in white matter properties.
In conclusion, this thesis improves the accuracy and the precision of the estimation of diffusion properties by carefully modelling complex fiber structures based on multi-tensor representations. What is more, it enhances the estimation by filtering the data to suppress the noise on the data. We anticipate that many diffusion MRI studies can benefit from this work.

Jianfei Yang
Acknowledgement

On Jan. 4th 2011 in a solid-state laser lab, China, I received the message from Frans that I got the PhD position on medical imaging, for which I was filled with excitement and curiosity. New life began on Jun. 28th, 2011, which was exciting but also difficult. I still remember I was even too shy to say hello to Lucas when we met him in front of Frans’ office, also my educational background on optics and condensed matter physics didn’t help directly facing the challenges in Diffusion Tensor Imaging. Without the help of my PhD promoters and colleagues, I could not accomplish this thesis. Working in QI must be one of the best experiences in my life, people here are friendly and open minded. Four years is not a short time, although it flies fast. In retrospect of this experience, I would like to express my appreciation to many people:

Lucas, I want to say I am so lucky and honoured under the your guidance. In the past four years, you have spent so much your valuable time on guiding me to view problems and help me on my research, from which I believe I will be lifelong benefitted. In Confucianism, we define the essence of a teacher -- A teacher, therefore proselytizes, instructs, resolve doubts (original: 师者，所以传道受业解惑也). I truly think you have done this! In life, I learn a lot from your open-mind viewing things; I greatly appreciate your help and consideration to complete my PhD career.

Frans, thanks for your trust which opened a door for me to enter the medical image processing field. In the past four years, we have had hundreds of progress meetings, through which I believe I have obtained full training. I would like to express my great appreciation on your patience on modifying my manuscripts. In life, you are also enthusiastic about helping others, from helping me kill the virus to explaining the art in museum.

Charles, it was always nice to meet you in AMC and thanks for all the suggestions you made on my research. Thanks for your diffusion MRI data. Also, thanks for modifying proposition 4 in this dissertation. Matthan and Dirk, without your help I could not have accomplished my research. From discussing the mathematical formulas, debugging MATLAB scripts to modifying the manuscripts, your suggestions have rooted in all my frameworks, thanks for all the help!
I would like to thanks my colleagues, a group of great people without whom my PhD career will not be complete: Zhang, so many common experiences and the daily-based ‘summit’, thanks for all the sharing and help, good luck buddy; Qiaole, you are always pleased to help and nice to offer delicious food, the life in Netherlands becomes more colourful when you are around, Tiago thanks for the interesting activities and games at your home; my dear roommates: Robert, Farzad and Vincent, working become more alive when you are/were in the office, thanks for the topics you brought and chatting; Ganesh, thanks for a lot of help in the beginning period of my PhD career; Kadir, thanks for a lot of sharing and understanding; Joor, it is a nice experience to work at the same research topic with you, thanks for your suggestions in our progress meeting and the travel guidance when we were in New York; Robiel, talking to you is always full of fun, with you the topics in coffee room becomes more interesting; Jeroen van Schie, thanks for the help when we assisted the bachelor courses, where I sent a lot of intractable students to you; Hui Tang, thanks for a lot of frank sharing and help; Tian Zhang, it is nice to have you in the group so I get a helper in a lot of issues; Thanks to Alex, Milos, Mojtaba, Nadya, Lena, Jelle, Jeroen and Anna, David, Juan, Tom, Babak for all the relaxing topics in the coffee/tea break (although a lot of time I am a silent listener). I would also like to thanks our supporting staff: Ronald, Annelies, Mandy and Wim, with your help, a lot of tough works becomes easier.

In the four years’ life in the Netherlands, I luckily meet a lot of Chinese friends: Yan Ren and Ting Hao, Changlin and Hongling, Wenbo and Zhang, thanks for all kinds of activities, spending Chinese festival with you are great memories; Yimin, many thanks for the guidance on life and career, I greatly appreciate your kind recommendation of many opportunities; Chunyang and Ling Zhang, the travel to Austria and Germany was a great success, thanks for nice guidance and arrangement; It is great that I can still have several old friends around in Netherlands: Shuqi, not necessary to say many words but you know this is wonderful! Qiang Wang, thanks for the regular visiting and those long conversations.

I want to thanks my family: to my dear mother and father, thanks for giving me your best without any expectation of return; to my parents in-law, thanks for your trust; to my brother Jian Yang, thanks for taking care of everything when I am in the Netherlands; Last but not least, to my lovely wife, Shaoying Wang, I am grateful and lucky to have you!
Curriculum Vitae

Jianfei Yang, was born in January 1985, Linyi (in Shandong Province), China. In June 2008, Jianfei obtained his Bachelor degree on optics from Shandong University, China. From September 2008 to June 2011, Jianfei did his master in State Key Lab of Crystal Materials, Shandong University, China. In his master period, Jianfei was majored on condensed matter physics and did research on solid-state laser. From July 2011 to September 2015, Jianfei did his PhD projects in Quantitative Imaging Group, TU Delft, The Netherlands. His PhD project was on analysis of diffusion magnetic resonance images, specialised on diffusion tensor imaging.