Higher-Order Tensors and Differential Topology in Diffusion MRI Modeling and Visualization

Kumulative Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dipl.-Phys. Dipl.-Math. Michael Peter Ankele
aus Altbach

Bonn, November 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Thomas Schultz
2. Gutachter: Assist.-Prof. Dr. Andrea Fuster

Tag der mündlichen Prüfung: 2. April 2019
Erscheinungsjahr: 2019

Contents

Abstract

Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) is a noninvasive method for creating three-dimensional scans of the human brain. It originated mostly in the 1970s and started its use in clinical applications in the 1980s. Due to its low risk and relatively high image quality it proved to be an indispensable tool for studying medical conditions as well as for general scientific research. For example, it allows to map fiber bundles, the major neuronal pathways through the brain. But all evaluation of scanned data depends on mathematical signal models that describe the raw signal output and map it to biologically more meaningful values. And here we find the most potential for improvement.

In this thesis we first present a new multi-tensor kurtosis signal model for DW-MRI. That means it can detect multiple overlapping fiber bundles and map them to a set of tensors. Compared to other already widely used multi-tensor models, we also add higher order kurtosis terms to each fiber. This gives a more detailed quantification of fibers. These additional values can also be estimated by the Diffusion Kurtosis Imaging (DKI) method, but we show that these values are drastically affected by fiber crossings in DKI, whereas our model handles them as intrinsic properties of fiber bundles. This reduces the effects of fiber crossings and allows a more direct examination of fibers.

Next, we take a closer look at spherical deconvolution. It can be seen as a generalization of multi-fiber signal models to a continuous distribution of fiber directions. To this approach we introduce a novel mathematical constraint. We show, that state-of-the-art methods for estimating the fiber distribution become more robust and gain accuracy when enforcing our constraint. Additionally, in the context of our own deconvolution scheme, it is algebraically equivalent to enforcing that the signal can be decomposed into fibers. This means, tractography and other methods that depend on identifying a discrete set of fiber directions greatly benefit from our constraint.

Our third major contribution to DW-MRI deals with macroscopic structures of fiber bundle geometry. In recent years the question emerged, whether or not, crossing bundles form two-dimensional surfaces inside the brain. Although not completely obvious, there is a mathematical obstacle coming from differential topology, that prevents general tangential planes spanned by fiber directions at each point to be connected into consistent surfaces. Research into how well this constraint is fulfilled in our brain is hindered by the high precision and complexity needed by previous evaluation methods. This is why we present a drastically simpler method that negates the need for precisely finding fiber directions and instead only depends on the simple diffusion tensor method (DTI). We then use our new method to explore and improve streamsurface visualization.

Zusammenfassung

Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) ist ein nicht-invasives bildgebendes Verfahren, mit dem das menschliche Gehirn räumlich abgebildet werden kann. Es wurde hauptsächlich in den 1970er Jahren entwickelt und fand ab den 1980ern Einzug in klinische Anwendungen. Seine Sicherheit und relativ hohe Bildqualität machen es zu einem unverzichtbaren Werkzeug für medizinische Untersuchungen und wissenschaftliche Forschung im Allgemeinen. Es erlaubt zum Beispiel die Kartierung von Faserbündeln, den neuronalen Hauptverbindungen durch das Gehirn. Jede Auswertung von Aufnahmen ist allerdings auf mathematische Signalmodelle angewiesen, die das rohe aufgenommene Ausgangssignal beschreiben und mit biologisch bedeutsamen Eigenschaften in Verbindung bringen. Und hier bietet sich uns das größte Potential für Verbesserungen.

In dieser Arbeit stellen wir zunächst ein neues Multi-Tensor-Kurtosis-Signalmodell für DW-MRI vor. Das heißt, es kann mehrere überlappende Faserbündel erkennen und auf eine Reihe von Tensoren abbilden. Im Vergleich zu anderen, bereits weit verbreiteten Multi-Tensor-Modellen, fügen wir jeder Faser außerdem Kurtosis-Terme höherer Ordnung hinzu. Dies führt zu einer detaillierteren numerischen Beschreibung der Fasern. Diese zusätzlichen Werte können alternativ auch per Diffusion Kurtosis Imaging (DKI) geschätzt werden. Wie wir jedoch zeigen werden, werden diese Werte in DKI durch Faserkreuzungen drastisch beeinflusst, während sie in unserem Modell als intrinsische Eigenschaften der Faserbündel behandelt werden. Dies reduziert Verfälschungen durch Faserkreuzungen und ermöglicht eine unmittelbarere Untersuchung der Fasern selbst.

Danach widmen wir uns der sphärische Dekonvolution. Diese kann als eine Verallgemeinerung von Mehrfaser-Signalmodellen hin zu einer kontinuierlichen Verteilung von Faserrichtungen verstanden werden. Zu diesem Ansatz führen wir eine neuartige mathematische Nebenbedingung ein. Wie wir zeigen werden, werden state-of-the-art Methoden zur Schätzung der Faserverteilung robuster und gewinnen an Genauigkeit, wenn zusätzlich unsere Nebenbedingung gefordert wird. Darüber hinaus ist es im Kontext unseres eigenen Dekonvolutionsschemas algebraisch gleichbedeutend mit der Forderung, dass das Signal in einzelne Fasern zerlegt werden kann. Dies bedeutet, dass Traktographie und andere Methoden, die von der Identifizierung einer diskreten Menge von Faserrichtungen abhängen, von unserer neuen Bedingung stark profitieren.

Unser dritter, größerer Beitrag zur DW-MRI befasst sich mit makroskopischen Strukturen der Faserbündelgeometrie. In den letzten Jahren stellte sich die Frage, ob kreuzende Bündel im Gehirn zweidimensionale Oberflächen aufspannen. Mathematisch gibt es eine nicht ganz offensichtliche Integrationsbedingung aus der Differentialtopologie, die von Tangentialebenen erfüllt werden muss, damit sie zu konsistenten Flächen verbunden werden können. Für Ebenen, die an jedem Punkt von allgemeinen Faserrichtungen aufgespannt werden, ist zu erwarten, dass diese Bedingung nicht erfüllt wird. Die Unter-

suchung der Frage, ob und wie gut diese Integrationsbedingung jedoch in unserem Gehirn erfüllt ist, wird durch die hohe Präzision und Komplexität behindert, die bisherige Tests erfordern. Aus diesem Grund stellen wir eine drastisch einfachere Testmethode vor, die die genaue Bestimmung der Faserrichtungen überflüssig macht und nur auf dem einfachen Diffusion Tensor Imaging (DTI) Verfahren basiert. Unseren neuen Test verwenden wir dann, um die Visualisierung von Streamsurfaces zu untersuchen und weiterzuentwickeln.

Acknowledgments

I want to thank my colleagues and friends who helped during the work on this thesis, especially my supervisor Prof. Dr. Thomas Schultz, who brought me into this topic and patiently kept his door open for many long discussions. I'm also grateful for the support and distractions by my former office roommates Amin Abbasloo and Vitalis Wiens.

Dedicated to my father.

Part I.

Background

1. Introduction

1.1. Motivation

Diffusion Weighted Magnetic Resonance Imaging (DW-MRI) gives us a powerful tool to research the living human body and is one of the most impressive technical achievements of the 20th century, combining discoveries from quantum mechanics, like nuclear spin and superconductivity, with data processing capabilities of modern computers. Extraordinary engineering obstacles had to be overcome resulting in the surprising measurement accuracy and speed of this method.

DW-MRI allows to shed light into the mysteries of the human brain. The brain is the most complex and still the least well understood organ in the human body. Besides pure desire for understanding the inner workings of the mind, diseases and disorders of the brain often have devastating effects. Simple malfunctions can completely alter someone's personality, render them unable to move and communicate with the outside world, or cause death by ceasing vital body functions.

Other research methods brought invaluable information about specific parts of the brain. Most information into neuronal chemistry was found by experiments performed in vitro. Other methods, though resulting in valuable information, are more ethically questionable. The immense value of DW-MRI is its non-disruptive nature. Scans can be performed on a large group of subjects without any significant risks.

Like any other method, DW-MRI does not come without limitations, though. Currently, it is only feasible to perform scans of the whole brain at a resolution that is far too low to observe individual nerve cells. The smallest significant structure this allows to detect are fiber bundles, directed groups of nerve cells connecting various parts of the brain and forming its underlying structure. Many diseases affect different properties of these bundles, making them an enticing subject of research.

Increasing the precision with which these fiber bundles can be detected and measured will open up the doors for numerous medical applications and is the main focus of this thesis. We present several methods that yield higher angular accuracy as well as additional parameters that allow some insight into their microstructure. We also expand our studies into their geometry on a macroscopic level into previously hypothesized surfaces formed by intersecting fibers.

Mathematically we are guided on our quest by the powerful allies of tensor algebra and the calculus and topology of tensor fields. Applying these ideas results in maps of new scalar measures and more reliable tractography from scanning methods that are already commonly used.

Why tensors?

We use tensors mostly to represent directional information in our models. Tensors, and equivalently spherical harmonics, give a natural representation for functions on a sphere. Besides having many elegant mathematical properties that make them easy to handle, they come with an *order* or *degree*, which allows us to choose an angular resolution suitable for our problem.

Tensors also generalize the vector concept and automatically appear in calculations with vector fields involving derivatives.

For most of this work, tensors, like vectors and matrices, can be seen as a simple list of numbers without too much loss of understanding. Although especially for chapter 7, proper introduction into the darker magic of algebra will be necessary.

1.2. Related work

Basic DW-MRI has already been extended in several directions. One of the most basic models and a starting point for most of these extensions is Diffusion Tensor Imaging (DTI) (see section 5.3). It takes scans from a single shell of b-values as input and estimates a field of diffusion tensors with only 6 degrees of freedom for each point inside the brain. This method, albeit being fast and robust, only allows very limited insight into the underlying tissue structures from these 6 numbers.

Earlier generalizations include multi-diffusion-tensor approaches [92, 47, 82] and higher-order tensors like Diffusional Kurtosis Imaging (DKI) [36], requiring multi-shell scans. Although the goal of the former is to describe multiple fiber bundles that constitute the brain's white matter and the latter seeks to fix problems with non-exponential signal decay between shells, there are also deeper connections between these two directions (see [35]). Both seek to extract more fine-grained tissue information without making any assumptions about the expected signals.

Here we present a new model that combines both approaches into a linear combination of higher-order tensors. Since each fiber bundle is mapped to one tensor, certain scalar measures of these tensors can give intrinsic properties of the fibers.

Another development is spherical deconvolution. It further generalizes the multi-tensor approach and can directly detect fiber bundles [89] making it more suitable for fiber tractography. Its precision benefits from mathematical constraints like non-negativity [87] and it was shown that the angular resolution required for tractography is further enhanced by low-rank tensor approximations [70, 41, 99, 68]. Other improvements in this field include multi-shell, multi-tissue deconvolution [40], Diffusion Spectrum Imaging (DSI) tractography [98, 12] and blind source deconvolution [17].

We expand upon this with a new and stronger non-negativity constraint originating in abstract tensor algebra. This constraint will be proven equivalent to decomposability into fiber bundles.

The third major field we will discuss is streamsurface integrability. It may seem like a natural step from deconvolution and tractography along a single fiber direction to surfaces that are tangential to two dominant fiber directions [103]. Although these surfaces

can only exist when certain mathematical conditions are fulfilled, they are nevertheless widely used for visualization and spawned a fierce discussion whether or not they exist as a natural phenomenon in our brains [97, 14, 95]. To explore this question more rigorously, Tax et al. defined the Sheet Probability Index (SPI) [84, 83], a computationally expensive measure. Some authors have claimed that the question depends strongly on the particular fiber model used in deconvolution and the restricted angular precision.

Our contribution here is a far simpler method for estimating the SPI. Instead of requiring complex fiber detection on high precision scans, our method is based on simple DTI.

1.3. Publications

This dissertation is based upon the following four publications:

- [4]: "*Quantifying Microstructure in Fiber Crossings with Diffusional Kurtosis*" by Michael Ankele et al. in proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.
- [2]: "*Versatile, robust, and efficient tractography with constrained higher-order tensor fODFs*" by Michael Ankele et al. in International Journal of Computer Assisted Radiology and Surgery, 2017.
- [6]: "*A Sheet Probability Index from Diffusion Tensor Imaging*" by Michael Ankele et al. in Computational Diffusion MRI, 2018.
- [5]: "*DT-MRI Streamsurfaces Revisited*" by Michael Ankele et al. in IEEE Transactions on Visualization and Computer Graphics, 2019.

Additionally, we published [3]: "*Fast and Accurate Multi-Tissue Deconvolution Using SHORE and H-psd Tensors*" by Michael Ankele et al. in MICCAI, 2016. This got extended into [2].

1.4. Outline and Contributions

Chapters 2 and 3 present the mathematical background of tensor algebra, spherical harmonics and a small amount of differential topology that is required to discuss our main methods. Since also some rudimentary knowledge about structures expected to be found in the brain is necessary, chapter 4 gives a small introduction into brain anatomy. Chapter 5 then shortly explains, how DW-MRI works, how it can quantify water diffusion in fiber bundles and lists mathematical models that are commonly used to estimate certain tissue parameters from scans.

The main research of our published papers is presented in chapter 6-9. Chapter 6 starts with a new model for diffusional kurtosis, resulting in new tissue parameters sensitive to individual fiber bundle's microstructure. Compared to similar state-of-the-art models, our parameters are shown to be less affected by fiber crossings.

Chapter 7 improves on stability and accuracy of fiber direction estimation. We introduce a new algebraic constraint for deconvolution methods that can be implemented efficiently in widely available numerical libraries. The model is defined in terms of SHORE basis functions, allowing its usage beyond the classical multi-shell setup. We also incorporate multi-tissue capabilities to model white matter, grey matter and corticospinal fluid simultaneously.

Lastly, chapter 8 and 9 delve into the analysis of fiber geometry. They give additional insight into the possible existence of two-dimensional sheet structures discussed in recent years. For this task, we introduce a new, more efficient and stable method to evaluate the sheet integrability and show that this method does not require sophisticated fiber estimation but only uses the simple diffusion tensor.

2. Tensors

The main findings of this thesis will be presented in the language of tensors. This chapter provides a quick introduction into tensor calculations and the relationship to spherical harmonics and homogenous polynomials.

2.1. Tensors

Vectors

Even though the pragmatic definition of vectors as tuples of numbers $\mathbf{v} \in \mathbb{R}^n$ is enough to understand most of this work, giving a cleaner, more abstract definition will help easing the mind into the theory of tensor algebra.

Definition 2.1. *A (real)* ***vector space*** V *is a set with two operations, addition and scalar multiplication, so that*

- $(V, +)$ *is a commutative group*
- *scalar multiplication is distributive and respects multiplication in* $\mathbb{R}$

Since we are exclusively concerned with **finite dimensional** vector spaces, we can always find a **basis**, i.e. a set of n vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, so that any vector $\mathbf{v} \in V$ can be decomposed into

$$\mathbf{v} = \sum_{i=1}^{n} v_i \, \mathbf{e}_i \tag{2.1}$$

with the components $v_i \in \mathbb{R}$.

Definition 2.2. *A symmetric, non-degenerate bilinear map* $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ *is called a* ***scalar product****.*

In general $g = (\langle \mathbf{e}_i, \mathbf{e}_k \rangle)_{i,k=0\dots n}$ is a symmetric matrix and the scalar product can take the form of a matrix product $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i,k} v_i \, g_{i,k} \, w_k = v^t \cdot g \cdot w$. A basis $(\mathbf{e}_i)$ is called **orthonormal** in the special case that $\langle \mathbf{e}_i, \mathbf{e}_k \rangle = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$. The right-hand expression will show up frequently enough deserving its own symbol and is called the **Kronecker delta**

$$\delta_{i,k} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} . \tag{2.2}$$

.

Definition 2.3. *The **dual space** $V^\star$ of V is the space of linear functions $\mathbf{f} : V \to \mathbb{R}$.*

Given a basis $(\mathbf{e}_i)$ of V, the dual basis $(\mathbf{e}_i^\star)$ is defined by $\mathbf{e}_i^\star(\mathbf{e}_k) = \delta_{i,k}$. Especially is $\dim V^\star = \dim V$. For $\mathbf{f} \in V^\star$ and $\mathbf{v} \in V$, evaluation is simply

$$\mathbf{f}(\mathbf{v}) = \sum_i f_i \, \mathbf{e}_i^\star \left(\sum_k v_k \, \mathbf{e}_k \right) = \sum_{i,k} f_i \, v_k \, \delta_{i,k} = \sum_i f_i \, v_i \tag{2.3}$$

and coincides with the canonical scalar product of the component vectors. This allows to compute basis expansions as

$$v_i = \mathbf{e}_i^\star(\mathbf{v}), \quad f_i = \mathbf{f}(\mathbf{e}_i) \, . \tag{2.4}$$

A scalar product defines a natural isomorphism $\flat : V \to V^\star$ by $\mathbf{v}^\flat = \langle \mathbf{v}, \cdot \rangle$.

Tensors

The concept of tensors can be confusing, especially since textbooks from different fields rarely agree on a consistent definition. Historically, tensors were introduced to represent certain physical quantities, like material stress, that express linear relationships between other vectorial quantities. These new quantities can require more numerical measurements to be fully defined than the usual three that are be required for vectors. So tensors can be seen as a generalization of vectors.

Among algebraists it is popular to define the tensor product $V \otimes W$ of vector spaces V, W abstractly by its "universal property". As elegant as this approach might be, it is a rather hostile one. We will instead use a more practical working definition given by the realization as multilinear functions on a fixed vector space V:

Definition 2.4. *A **covariant tensor** $\boldsymbol{T} \in \mathcal{T}^d = \underbrace{V^\star \otimes \cdots \otimes V^\star}_{d}$ of order d is a multi-linear map $\boldsymbol{T} : \underbrace{V \times \cdots \times V}_{d} \to \mathbb{R}$.*

The $(\mathbf{e}_{i_1}^\star \otimes \cdots \otimes \mathbf{e}_{i_d}^\star)$ for $i_1, \dots, i_d \in \{1, \dots, n\}$ form a basis of $\mathcal{T}^d$. We conclude that the space $\mathcal{T}^d$ has dimension n^d. The basis expansion is

$$\boldsymbol{T} = \sum_{i_1, \dots, i_d = 1}^{n} T_{i_1, \dots, i_d} \, \mathbf{e}_{i_1}^\star \otimes \cdots \otimes \mathbf{e}_{i_d}^\star \tag{2.5}$$

with the coefficients $T_{i_1, \dots, i_d}$ forming a d-dimensional array.

Because of $\left(\mathbf{e}_{i_1}^\star \otimes \cdots \otimes \mathbf{e}_{i_d}^\star \right) (\mathbf{v}_1, \dots, \mathbf{v}_d) = \prod_{k=1}^d v_{k,i_k}$, function evaluation takes the simple form

$$\boldsymbol{T}(\mathbf{v}_1, \dots, \mathbf{v}_d) = \sum_{i_1, \dots, i_d = 1}^{n} T_{i_1, \dots, i_d} \, v_{1,i_1} \, \dots \, v_{d,i_d} \, . \tag{2.6}$$

Symmetric tensors

For this work a special class of tensors will be of particular interest:

Definition 2.5. *A tensor is (totally)* ***symmetric****, if*

$$\boldsymbol{T}(\mathbf{v}_1, \dots, \mathbf{v}_d) = \boldsymbol{T}(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(d)}) \tag{2.7}$$

for every permutation $\sigma \in S(d)$. For its components this means:

$$T_{i_1,\dots,i_d} = T_{i_{\sigma(1)},\dots,i_{\sigma(d)}} \tag{2.8}$$

The space of all symmetric tensors of order d will be denoted by $\mathsf{Sym}_{n,d}$.

A symmetric tensor $\boldsymbol{T}$ is already completely determined by its components $T_{i_1,\dots,i_d}$ with $i_1 \leq \dots \leq i_d$, the other components can be recovered via eq. (2.8). This also fixes the dimension of $\mathsf{Sym}_{n,d}$ to $\binom{n+d-1}{d}$. Our work focuses on low order symmetric tensors in $\mathbb{R}^3$, typical dimensions are:

order d	0	1	2	4	6	8
$\dim(\mathsf{Sym}_{3,d})$	1	3	6	15	28	45

2.2. Polynomials

Indexing

To capture the intuitive indexing scheme for multi-dimensional polynomials like

$$p(x, y, z) = a_{2,3,4}\, x^2 y^3 z^4 + a_{1,5,0}\, x y^5 + \dots \tag{2.9}$$

an efficient method is needed:

Definition 2.6. *A tuple $\boldsymbol{i} \in \{1, \dots, d\}^n$ is called a* ***multi-index****. We set $|\boldsymbol{i}| = \sum_k \boldsymbol{i}_k$ and $\mathbf{v}^{\boldsymbol{i}} = \prod_{k=1}^{n} (v_k)^{i_k}$ for an n-dimensional vector $\mathbf{v}$. The values*

$$\binom{\boldsymbol{i}}{} = \frac{|\boldsymbol{i}|!}{\prod \boldsymbol{i}_k!} \tag{2.10}$$

are called ***multinomial coefficients****.*

This allows the elegant notation

Definition 2.7. *A* ***polynomial*** *in n variables is an expression of the form*

$$p(\mathbf{x}) = p(\mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{\boldsymbol{i}} \binom{\boldsymbol{i}}{} a_{\boldsymbol{i}}\, \mathbf{x}^{\boldsymbol{i}} \tag{2.11}$$

with constants $a_{\boldsymbol{i}}$.

The arbitrary factor $\binom{\boldsymbol{i}}{}$ in this definition will be justified in the proof of theorem 2.1.

Definition 2.8. *A polynomial* p *is* ***homogeneous*** *or a* ***form*** *of degree* d *if*

$$p(\lambda \mathbf{x}) = \lambda^d \, p(\mathbf{x}), \quad \forall \lambda \in \mathbb{R} \,. \tag{2.12}$$

Forms vs. symmetric tensors

Theorem 2.1. *The space of homogeneous polynomials of degree* d *in* n *dimensions is isomorphic to the space of symmetric tensors* $\mathsf{Sym}_{n,d}$.

Proof. To show this, we need the natural homomorphism that maps a tensor $\boldsymbol{T}$ to a polynomial $\boldsymbol{T}(\mathbf{x})$ by setting $\boldsymbol{T}(\mathbf{x}) = \boldsymbol{T}(\mathbf{x}, \dots, \mathbf{x})$. We'll show that this is an isomorphism on $\mathsf{Sym}_{n,d}$. Let's make this map more explicit for a symmetric tensor

$$\boldsymbol{T}(\mathbf{x}) = \sum_{i_1,\dots,i_d} T_{i_1,\dots,i_d} \, x_{i_1} \dots x_{i_d} \,. \tag{2.13}$$

Since neither $T_{i_1,\dots,i_d}$ nor $x_{i_1} \dots x_{i_d}$ depend on the permutation of the indices i, we can use multi-indices $\boldsymbol{i}_k = \#\{i_\star = k\}$ instead, counting how many of the is share each possible value k. Some $\boldsymbol{i}$s will appear multiple times in the sum, coming from different permutations of a tuple i. These multiplicities are $\frac{d!}{\prod_k \#\{i_\star = k\}!}$ which is exactly the multinomial coefficient $\binom{\boldsymbol{i}}{}$. So we have the non-redundant form

$$\boldsymbol{T}(\mathbf{x}) = \sum_{|\boldsymbol{i}|=d} \binom{\boldsymbol{i}}{} T_{\boldsymbol{i}} \, \mathbf{x}^{\boldsymbol{i}} \,. \tag{2.14}$$

To show that this is an isomorphism: Assuming that the polynomial is $\boldsymbol{T}(\mathbf{x}) \equiv 0$, then all coefficients $T_{\boldsymbol{i}}$ have to be zero which means also all $T_{i_1,\dots,i_d}$ are zero, resulting in a vanishing tensor $\boldsymbol{T} = 0$. This shows that the map is injective. For surjectivity note, that every possible polynomial with coefficients $a_{\boldsymbol{i}}$ can be the image of a tensor with coefficients $T_{\boldsymbol{i}} = a_{\boldsymbol{i}}$. □

Remark. Symmetric tensors can also be recovered from polynomials by polarization

$$\boldsymbol{T}(\mathbf{v}_1, \dots, \mathbf{v}_d) = \frac{1}{d!} \frac{\partial}{\partial \lambda_1} \dots \frac{\partial}{\partial \lambda_d} \, \boldsymbol{T} \left(\sum_k \lambda_k \mathbf{v}_k \right) \,. \tag{2.15}$$

Harmonic forms

Some useful preparation for the next section:

Definition 2.9. *The **Laplace operator** is the second order differential operator*

$$\Delta = \sum_{k=1}^{n} \frac{\partial^2}{\partial {x_k}^2} \,. \tag{2.16}$$

*A function $f(\mathbf{x})$ is called **harmonic** if it solves **Laplace's equation***

$$\Delta f = 0 \,. \tag{2.17}$$

Theorem 2.2. *Harmonic forms are equivalent to trace-free symmetric tensors.*

To keep the number of indices low and readability high, we'll show this explicitly in 3 dimensions. The generalization should be obvious.

Proof. First apply a single term of the Laplace operator:

$$\frac{\partial^2}{\partial x^2} \mathbf{x}^{\boldsymbol{i}} = \frac{\partial}{\partial x} \boldsymbol{i}_x \, \mathbf{x}^{\boldsymbol{i}-(1,0,0)} = \boldsymbol{i}_x (\boldsymbol{i}_x - 1) \, \mathbf{x}^{\boldsymbol{i}-(2,0,0)} \tag{2.18}$$

And so:

$$\Delta p = \sum_{|\boldsymbol{i}|=d} \binom{\boldsymbol{i}}{} p_{\boldsymbol{i}} \left[\boldsymbol{i}_x (\boldsymbol{i}_x - 1) \, \mathbf{x}^{\boldsymbol{i}-(2,0,0)} + \boldsymbol{i}_y (\boldsymbol{i}_y - 1) \, \mathbf{x}^{\boldsymbol{i}-(0,2,0)} + \dots \right] = q \tag{2.19}$$

This is a new form of degree $d-2$ with the general coefficient $q_{\boldsymbol{j}}$

$$\begin{aligned} \binom{\boldsymbol{j}}{} q_{\boldsymbol{j}} = \big[& \binom{\boldsymbol{j}+(2,0,0)}{} p_{\boldsymbol{j}+(2,0,0)} (\boldsymbol{j}_x + 2)(\boldsymbol{j}_x + 1) \\ & + \binom{\boldsymbol{j}+(0,2,0)}{} p_{\boldsymbol{j}+(0,2,0)} (\boldsymbol{j}_y + 2)(\boldsymbol{j}_y + 1) + \dots \big] \end{aligned} \tag{2.20}$$

for $|\boldsymbol{j}| = d-2$. Observing that

$$\binom{\boldsymbol{j}+(2,0,0)}{} (\boldsymbol{j}_x + 2)(\boldsymbol{j}_x + 1) = \frac{d!}{(\boldsymbol{j}_x + 2)! \boldsymbol{j}_y! \boldsymbol{j}_z!} (\boldsymbol{j}_x + 2)(\boldsymbol{j}_x + 1) \tag{2.21}$$

$$= \binom{\boldsymbol{j}}{} d\,(d-1) \,, \tag{2.22}$$

making all factors equal, allows to factor them out:

$$\binom{\boldsymbol{j}}{} q_{\boldsymbol{j}} = d\,(d-1) \binom{\boldsymbol{j}}{} \left[p_{\boldsymbol{j}+(2,0,0)} + p_{\boldsymbol{j}+(0,2,0)} + p_{\boldsymbol{j}+(0,0,2)} \right] \tag{2.23}$$

The term $p_{\boldsymbol{j}+(2,0,0)} + p_{\boldsymbol{j}+(0,2,0)} + p_{\boldsymbol{j}+(0,0,2)}$ is exactly the trace of the tensor representing p, leading to the relation

$$\Delta p \sim d\,(d-1) \operatorname{tr} \boldsymbol{T}_p \tag{2.24}$$

□

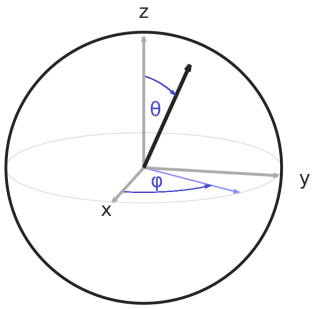
Remark. From this proof it can be seen that the space $\mathsf{Sym}^{tr=0}_{n,d}$ of trace-free forms of degree $d \geq 2$ has dimension $\dim \mathsf{Sym}_{n,d} - \dim \mathsf{Sym}_{n,d-2}$. Applying this recursively even gives us an identification

$$\mathsf{Sym}_{n,d} \cong \mathsf{Sym}^{tr=0}_{n,d} \oplus \mathsf{Sym}^{tr=0}_{n,d-2} \oplus \mathsf{Sym}^{tr=0}_{n,d-4} \oplus \dots \tag{2.25}$$

Especially in our case $\dim \mathsf{Sym}^{tr=0}_{3,d} = 2n+1$.

2.3. Spherical Harmonics

Functions on the sphere



$\mathbb{R}^3$ can be parametrized by **polar coordinates**

$$\mathbf{r}(r,\theta,\varphi) = r \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix} \quad \begin{matrix} r \geq 0 \\ \text{with } 0 \leq \theta \leq \pi \\ 0 \leq \varphi < 2\pi \end{matrix} . \tag{2.26}$$

For $r = 1$ we can cover the **sphere**

$$S^2 = \{\mathbf{r} \in \mathbb{R}^3 : \|\mathbf{r}\| = 1\} \tag{2.27}$$

by the two angular parameters θ, φ.

This section will examine complex functions $f : S^2 \to \mathbb{C}$ on the sphere with the **scalar product** defined by integration

$$\langle f, g \rangle_{S^2} = \int_{S^2} f \cdot \bar{g}\, dA = \int_0^{\pi} \int_0^{2\pi} f(\theta,\varphi)\, \bar{g}(\theta,\varphi)\, \sin\theta\, d\varphi\, d\theta\,. \tag{2.28}$$

We will also sometimes use the notation $f(\mathbf{u}) = f(\theta,\varphi)$ for a unit vector $\mathbf{u}(\theta,\varphi)$.

Construction

The complex spherical harmonics appear as solutions of Laplace's equation $\Delta f(\mathbf{x}) = 0$ in spherical coordinates when enforcing separability $f(\mathbf{x}) = R(r)\,\Theta(\theta)\,\Phi(\varphi)$. The rest of this segment delves into the details of their construction with an eye on their connection to harmonic forms. This can be skipped when only interested in their applications.

The Laplace operator transformed into spherical coordinates takes the form

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\, \frac{\partial}{\partial\theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\,. \tag{2.29}$$

Applying this to $0 = \Delta f = \Delta\left(R(r)\,\Theta(\theta)\,\Phi(\varphi)\right)$ and multiplying both sides by $\frac{r^2}{f}$ gives

$$0 = \frac{2r\,R'(r) + r^2 R''(r)}{R(r)} + \frac{\cos\theta\,\Theta'(\theta) + \sin\theta\,\Theta''(\theta)}{\Theta(\theta)\,\sin\theta} + \frac{\Phi''(\varphi)}{\sin^2\theta\,\Phi(\varphi)}\,. \tag{2.30}$$

The first term only depends on r while the rest is independent of it. This means it has to be constant. We can even make the ansatz $R(r) = R_l(r) = r^l$ for an integer $l \geq 0$ since we want to find harmonic forms. The first term then becomes

$$\frac{2r\,R'(r) + r^2 R''(r)}{R(r)} = \frac{2r\,l\,r^{l-1} + r^2 l\,(l-1) r^{l-2}}{r^l} = l\,(l+1)\,. \tag{2.31}$$

Similarly the third term has the only φ-dependence, which also forces $\frac{\Phi''}{\Phi}$ to be constant. The simple guess $\Phi(\varphi) = \Phi_m(\varphi) = e^{im\phi}$ with $m \in \mathbb{Z}$ gives

$$\frac{\Phi_m''}{\Phi_m} = \frac{-m^2 e^{im\varphi}}{e^{im\varphi}} = -m^2 \tag{2.32}$$

turning eq. (2.30) into a messy but ordinary differential equation for $\Theta_l^m(\theta)$:

$$0 = l\,(l+1) + \frac{\cos\theta\,\Theta'(\theta) + \sin\theta\,\Theta''(\theta)}{\Theta(\theta)\,\sin\theta} - \frac{m^2}{\sin^2\theta}\,. \tag{2.33}$$

This is the **general Legendre equation** for $\Theta_l^m(\theta) = P_l^m(\cos\theta)$. It is solved by the **associated Legendre polynomials** $P_l^m(\cos\theta)$ for $-l \leq m \leq l$. Technically, these are no polynomials, but form homogeneous polynomials of degree l when expressed in $\cos\theta$ and $\sin\theta$:

	$m=-2$	-1	0	1	2
$l=0$	–	–	1	–	–
1	–	$\frac{1}{2}\sin\theta$	$\cos\theta$	$-\sin\theta$	–
2	$\frac{1}{8}\sin^2\theta$	$\frac{1}{2}\cos\theta\,\sin\theta$	$\cos^2\theta - \frac{1}{2}\sin^2\theta$	$-3\cos\theta\,\sin\theta$	$3\sin^2\theta$
$\vdots$			$\vdots$		

Definition

Putting the angular parts together encourages the

Definition 2.10. *The **complex spherical harmonics** are the set of functions*

$$\mathcal{Y}_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\theta)\,e^{im\varphi} \tag{2.34}$$

for $(l,m) \in \mathbb{N}\times\mathbb{Z}$ with $-l \leq m \leq l$. The $P_l^m(\cos\theta)$ are the aforementioned associated Legendre polynomials.

The alternative choice of $\Phi(\varphi) = \begin{cases}\cos|m|\varphi & m \leq 0\\ \sin\,m\,\varphi & m > 0\end{cases}$ instead of $\Phi_m(\varphi) = e^{im\varphi}$ to solve eq. (2.32), leads to

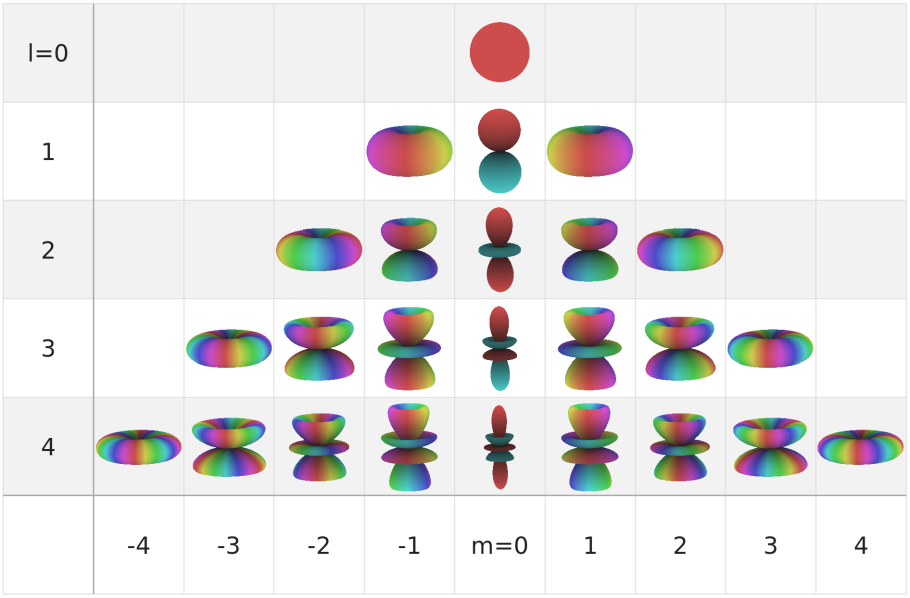


Figure 2.1.: The complex spherical harmonics $\mathcal{Y}_l^m$ of degree $l \leq 4$. Hue represents the complex phase and the radius is scaled by the absolute value.

Definition 2.11. *The **real** spherical harmonics (SH) are:*

$$Y_l^m = \begin{cases} \sqrt{2}\,\mathrm{Re}(\mathcal{Y}_l^{-m}) & m < 0 \\ \mathcal{Y}_l^0 & m = 0 \\ \sqrt{2}\,(-1)^{m+1}\,\mathrm{Im}(\mathcal{Y}_l^m) & m > 0 \end{cases} \tag{2.35}$$

The awkward constant factor in $\mathcal{Y}_l^m$ ensures **orthonormality**

$$\langle \mathcal{Y}_l^m, \mathcal{Y}_{l'}^{m'} \rangle_{S^2} = \langle Y_l^m, Y_{l'}^{m'} \rangle_{S^2} = \delta_{m,m'}\delta_{l,l'}\,. \tag{2.36}$$

This means the integral is one for $(m, l) = (m', l')$ and vanishes otherwise. A direct result of orthonormality is **linear independence**, serving as basis elements of certain function spaces on the sphere.

Their role in solving Laplace's equation makes them highly prevalent in the physics of classical field theories like electrodynamics and the heat equation. They also play an important role in pure math, since for fixed l the tuple $(\mathcal{Y}_l^m)$ spans irreducible representation spaces of the groups $\mathsf{SO}(3)$ and $\mathsf{SU}(2)$, which is again intimately connected to the theory of angular momentum and spin in quantum mechanics. A very brief glimpse into this connection will be given in section 5.1.

For this thesis their value lies mostly in forming a basis on S^2 and giving a natural way to approximate and smoothen functions and also to represent them on a computer as a list of coefficients. The spherical harmonics play a similar role to plane waves in

Fourier analysis and also allow to easily convolve functions as discussed further down.

Properties

Inversion symmetry:

$$Y_l^m(-\mathbf{u}) = (-1)^l Y_l^m(\mathbf{u}) \tag{2.37}$$

That means these functions are symmetric for even degree l and antisymmetric for odd degree.

They also form a basis of the space of (real) functions on the sphere. This allows to expand any smooth function $f : S^2 \to \mathbb{R}$ into

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m \cdot Y_l^m(\theta, \varphi) \tag{2.38}$$

with expansion coefficients

$$f_l^m = \langle f, Y_l^m \rangle_{S^2} . \tag{2.39}$$

This can be used to approximate functions up to degree l_{max}

$$f(\theta, \varphi) \approx \sum_{l=0}^{l_{\mathsf{max}}} \sum_{m=-l}^{l} f_l^m \cdot Y_l^m(\theta, \varphi) \tag{2.40}$$

Equivalence to tensors and polynomials

Let's call SH_l the space spanned by $(Y_l^m)_{m=-l\dots l}$ for a fixed $l \in \mathbb{N}$. As discussed before, the functions $r^l Y_l^m$ for any $m \in [-l, \dots, l]$ solve Laplace's equation $\Delta \left(r^l Y_l^m \right) = 0$. These functions are called **solid harmonics** and are homogeneous polynomials by construction. The radial term r^l tells us that the degree is l. Since their construction lead to $2l + 1$ independent functions, the whole space of harmonic forms is covered. By Thm. 2.2 these functions are equivalent to trace-free tensors, giving even a natural identification $\mathsf{SH}_l \cong \mathsf{Sym}_{3,l}^{tr=0}$. And so

$$\mathsf{Sym}_{3,l} \cong SH_l \oplus SH_{l-2} \oplus SH_{l-4} \oplus \dots \tag{2.41}$$

Convolution

Observation. The $Y_l^{m=0}(\theta, \varphi)$ are rotationally symmetric around the z-axis and independent of φ. Functions $f(\theta)$ with this symmetry are sometimes called **isotropic** or **zonal** and can be expanded into

$$f(\theta) = \sum_l f_l \, Y_l^0(\theta) \tag{2.42}$$

and will find their use as convolution kernels.

In general, convolution of a function can be seen as a smoothing operation, averaging the function with a kernel. The kernel needs to be shifted around in the domain of the original function to get the average everywhere. A clean definition of this process requires the kernel to be defined on the set of possible shifts. In our case this set is the group of rotations $\mathsf{SO}(3)$ that map the sphere onto itself.

Definition 2.12. *The **spherical convolution** of a function $f : S^2 \to \mathbb{C}$ with a kernel $k : \mathsf{SO}(3) \to \mathbb{C}$ is the weighted average over the group of rotations*

$$(k \star f)(\mathbf{u}) = \int_{\mathsf{SO}(3)} k(R)\, f(R^{-1}\mathbf{u})\, dR\,. \tag{2.43}$$

The integral uses the invariant Haar measure[a] with $\int_{\mathsf{SO}(3)} dR = 1$.

[a] In literature like [21] sometimes a measure with $\int_{\mathsf{SO}(3)} dR = 8\pi^2$ is used leading to a different constant in the convolution theorem.

Remark. A kernel $k : S^2 \to \mathbb{C}$ can be lifted from the sphere into $\mathsf{SO}(3)$ by $\tilde{k}(R) = k(R\mathbf{e}_z)$.

Theorem 2.3. *For an isotropic kernel $k(\theta)$ the convolution becomes multiplication of the coefficients*

$$(k \star f)_l^m = \frac{1}{\sqrt{4\pi(2l+1)}} k_l\, f_l^m\,. \tag{2.44}$$

2.4. Approximation

The spectral theorem states that any symmetric (real) $n \times n$-matrix M can be decomposed into a sum of n products

$$M = \sum_{i=1}^{n} \lambda_i\, \mathbf{v}_i \otimes \mathbf{v}_i\,. \tag{2.45}$$

For positive semi-definite matrices M, the coefficients λ_i will be non-negative. This can be used to construct rank-R approximations of M via

$$M \approx \sum_{i=1}^{R<n} \lambda_i\, \mathbf{v}_i \otimes \mathbf{v}_i\,, \tag{2.46}$$

including only the terms from dominant eigenvalues λ_i.

Symmetric tensors $\boldsymbol{T}$ of higher order can still be decomposed into a sum of simple tensor products, as seen in eq. (2.5). But in general, there is no decomposition of non-negative tensors $\boldsymbol{T}$ into

$$\boldsymbol{T} = \sum_i \lambda_i \mathbf{v}_i \otimes \mathbf{v}_i \otimes \mathbf{v}_i \otimes \dots\,. \tag{2.47}$$

with $\lambda_i \geq 0$. Also, finding a low-rank approximation for $\boldsymbol{T}$ is a very hard problem. Several heuristic methods can be used, all computationally expensive and accompanied by many drawbacks[27].

In chapter 7 we introduce a simple criterion for symmetric fourth order tensors in three dimensions to be decomposable into rank-1 terms like in eq. (2.47). Further information about the construction that is involved can be found in appendix C.

3. Differential Topology

We won't need the full power of abstract differential topology here. But some basic concepts, like derivatives of maps and vector fields, are unavoidable and demand a proper introduction. Especially, vector field integrability will prove crucial for chapter 8.

3.1. Diffeomorphisms

Definition 3.1. *A map $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is called a **diffeomorphism**, if it is smooth (i.e. C^∞ differentiable) and has an inverse map ϕ^{-1} that is also smooth.*

For any smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ we can calculate the **Jacobi matrix** containing the first order partial derivatives

$$D\phi = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \phi_m}{\partial x_1} & \cdots & \frac{\partial \phi_m}{\partial x_n} \end{pmatrix} . \tag{3.1}$$

If ϕ is a diffeomorphism then $D\phi$ is square and has full rank.

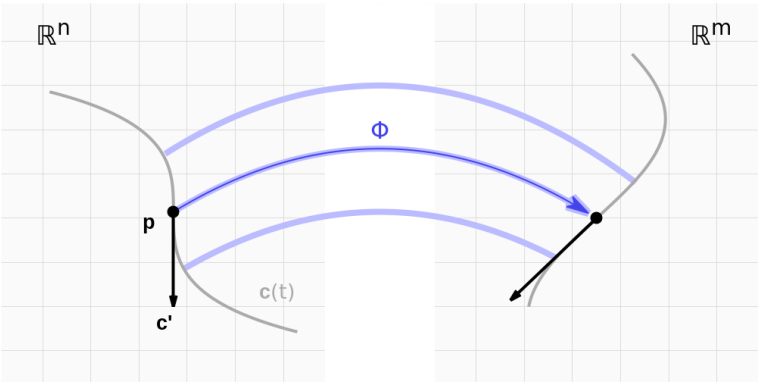


Figure 3.1.: $\phi : \mathbb{R}^n \to \mathbb{R}^m$ smoothly mapping points $\mathbf{p}$, curves $\mathbf{c}$ and tangential vectors $\mathbf{c}'$.

Now consider a smooth vector field $\mathbf{v}(\mathbf{p}) = \mathbf{v_p}$ on $\mathbb{R}^n$. A smooth map $\phi : \mathbb{R}^n \to \mathbb{R}^m$ can be used to push the vectors $\mathbf{v}$ forward into $\mathbb{R}^m$ by setting

$$(\phi_\star \mathbf{v})_{\phi(\mathbf{p})} = D\phi_{\mathbf{p}} \cdot \mathbf{v_p} , \tag{3.2}$$

which means we take the vector at position $\mathbf{p}$, transform it by $D\phi_{\mathbf{p}}$ and attach it at position $\phi(\mathbf{p})$. This defines a new vector field $\phi_\star \mathbf{v}$ if ϕ is a diffeomorphism; otherwise a point in the target might be covered multiple times or not at all. The transformation by $D\phi$ also ensures that vectors tangential to a curve again yield vectors tangential to the image of the curve.

Similarly, a dual vector field $\boldsymbol{\omega}_{\mathbf{p}}$ on $\mathbb{R}^m$ can be pulled back along ϕ via

$$(\phi^\star \boldsymbol{\omega})_{\mathbf{p}} = \boldsymbol{\omega}_{\phi(\mathbf{p})} \circ D\phi_{\mathbf{p}} \,. \tag{3.3}$$

3.2. Integral curves

In our every-day lives, wind and water currents are well known examples of phenomena described by vector fields. And from these we carry an intuitive understanding of movement along these currents and global paths of transport created by local vectors. Mathematically, this intuitive picture gives rise to:

Definition 3.2. *The **flow** of a vector field* $\mathbf{v}$ *is a map* $\Phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ *with*

- $\Phi_0 = \mathsf{id}$
- $\Phi_t(\Phi_s(\mathbf{p})) = \Phi_{s+t}(\mathbf{p})$
- $\left(\frac{d}{dt}\Phi_t(\mathbf{p})\right)_{|t=0} = \mathbf{v}_{\mathbf{p}}$

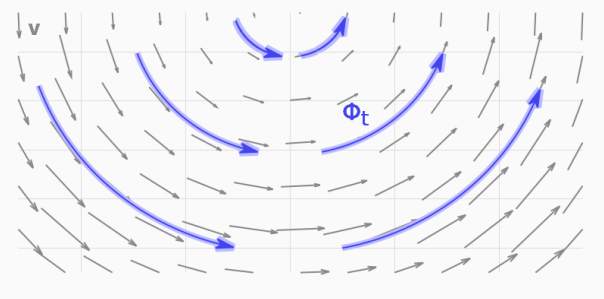


Figure 3.2.: A vector field $\mathbf{v}$ and its flow Φ_t for a specific value of t.

The path passing through a single point $\mathbf{p}_0$ at time $t = 0$ is $\mathbf{c}(t) = \Phi_t(\mathbf{p}_0)$ which is a curve tangential to $\mathbf{v}$ at every image point and usually called an **integral curve** of $\mathbf{v}$.

By the Picard–Lindelöf theorem the flow is guaranteed to exist uniquely at least in small regions for vector fields $\mathbf{v}$ that are Lipschitz continuous, i.e. if there exists a constant $K > 0$ so that for any two points $\mathbf{p}$, $\mathbf{q}$ the following condition holds:

$$\|\mathbf{v}_{\mathbf{p}} - \mathbf{v}_{\mathbf{q}}\| \leq K\|\mathbf{p} - \mathbf{q}\| \tag{3.4}$$

On compact domains the flow will exist even globally. Also the continuity condition can be dropped when dealing with vector fields defined on a discrete grid, since it can

always be interpreted as discretisation of a continuous field and interpolated smoothly. So, from this work's perspective there is no constraint on the existence of flows and integral curves.

In section 8.2 we will find a simplified introduction into Frobenius theory which generalizes this concept of integral curves tangential to a single vector field to higher dimensional surfaces tangential to multiple vector fields. The main message will be, that these do not exist for arbitrary vector fields and additional conditions have to be met to form these surfaces.

3.3. Derivatives of Vector Fields

In $\mathbb{R}^n$ there is a simple way to define the directional derivative of a vector field. For the derivative of the field $\mathbf{w}$ in the direction of $\mathbf{v}$ we can set

$$(\nabla_{\mathbf{v}}\mathbf{w})_{\mathbf{p}} = \left(\frac{d}{dt}\mathbf{w}_{(p+t\mathbf{v}_{\mathbf{p}})}\right)_{|t=0}, \tag{3.5}$$

which can be expressed via the Jacobi matrix of $\mathbf{w}(\mathbf{p})$:

$$(\nabla_{\mathbf{v}}\mathbf{w})_{\mathbf{p}} = (D\mathbf{w})_{\mathbf{p}} \cdot \mathbf{v}_{\mathbf{p}}\,. \tag{3.6}$$

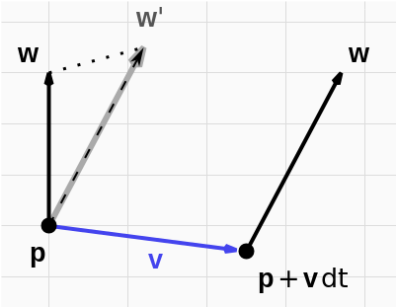


Figure 3.3.: The covariant derivative $\nabla_{\mathbf{v}}\mathbf{w}$ compares $\mathbf{w}_{\mathbf{p}}$ with the parallel transport of $\mathbf{w}_{\mathbf{p}+\mathbf{v}dt}$ (dashed).

The operator ∇ (or rather a version generalized to curved manifolds) is called the **covariant derivative** because the object $\nabla\mathbf{w} = D\mathbf{w}$ is a linear function in the parameter $\mathbf{v}$. In general ∇ can be defined for tensor fields $\boldsymbol{T}$ resulting in a tensor field $\nabla\boldsymbol{T}$ with the covariant order increased by one.

Differential geometry and topology provide an assortment of alternative and non-equivalent constructions for vector field derivatives. This is unavoidable because our simple definition will fail in a setup of abstract manifolds. It depends on the ability to compare vectors at different locations $\mathbf{p}$ and $\mathbf{p} + t\mathbf{v}$ which is not possible when each location is equipped with its own, distinct tangential vector space. The alternatives arise from different methods of comparing these vector spaces.

Another type of derivative is called the **Lie derivative** $\mathcal{L}_{\mathbf{v}}\mathbf{w}$ and is defined as

$$(\mathcal{L}_{\mathbf{v}}\mathbf{w})_{\mathbf{p}} = \lim_{h\to 0} \frac{1}{h}(\mathbf{w}_{\mathbf{p}} - (\Phi_{h\star}\mathbf{w})_{\mathbf{p}}) \tag{3.7}$$

with Φ the flow of $\mathbf{v}$. This means we use Φ to transport $\mathbf{w}$ along the integral curve of $\mathbf{v}$. This circumvents the problem of distinct vector spaces by only comparing vectors after moving them to the same location $\mathbf{p}$.

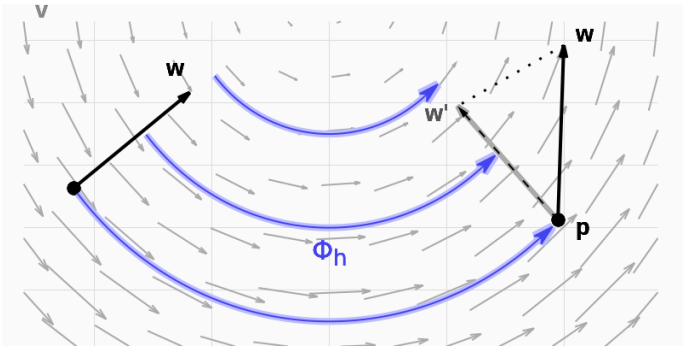


Figure 3.4.: The Lie derivative $\mathcal{L}_{\mathbf{v}}\mathbf{w}$ compares $\mathbf{w}_{\mathbf{p}}$ with the image (dashed) of $\mathbf{w}$ along the flow of $\mathbf{v}$.

The major difference between $\nabla_{\mathbf{v}}\mathbf{w}$ and $\mathcal{L}_{\mathbf{v}}\mathbf{w}$ is that $\nabla_{\mathbf{v}}\mathbf{w}$ has tensor behaviour in $\mathbf{v}$, i.e. it only depends on the value of $\mathbf{v}$ at one single point. $\mathcal{L}_{\mathbf{v}}\mathbf{w}$ on the other hand also depends on the partial derivatives of $\mathbf{v}$. A surprising result is, that even $\mathcal{L}_{\mathbf{v}}\mathbf{w} = -\mathcal{L}_{\mathbf{w}}\mathbf{v}$.

3.4. Lie Bracket

The Lie bracket of two vector fields $\mathbf{v}$ and $\mathbf{w}$ is often defined as $[\mathbf{v},\mathbf{w}] = \mathbf{vw} - \mathbf{wv}$ using the inconspicuous notation identifying a vector $\mathbf{v} = \sum_i v_i \mathbf{e}_i$ with the first order linear differential operator $\sum_i v_i \frac{\partial}{\partial x_i}$ acting on scalar functions $f(\mathbf{p})$. Making this definition more explicit will prove insightful:

$$[\mathbf{v},\mathbf{w}]f = \left(\sum_i v_i \frac{\partial}{\partial x_i}\right)\left(\sum_k w_k \frac{\partial}{\partial x_k}\right) f - \left(\sum_i w_i \frac{\partial}{\partial x_i}\right)\left(\sum_k v_k \frac{\partial}{\partial x_k}\right) f \tag{3.8}$$

$$= \sum_{i,k}\left(v_i \frac{\partial w_k}{\partial x_i}\frac{\partial f}{\partial x_k} + v_i w_k \frac{\partial^2 f}{\partial x_i \partial x_k} - w_i \frac{\partial v_k}{\partial x_i}\frac{\partial f}{\partial x_k} - w_i v_k \frac{\partial^2 f}{\partial x_i \partial x_k}\right) \tag{3.9}$$

$$= \sum_{i,k}\left(v_i \frac{\partial w_k}{\partial x_i} - w_i \frac{\partial v_k}{\partial x_i}\right)\frac{\partial f}{\partial x_k} \tag{3.10}$$

Surprisingly, even though $[\mathbf{v},\mathbf{w}]$ was defined as a second order differential operator, both second order terms cancel each other out and only a first order operator remains. Since we identify first order operators with vectors, we end up with the vector equation

$$[\mathbf{v},\mathbf{w}] = \nabla_{\mathbf{v}}\mathbf{w} - \nabla_{\mathbf{w}}\mathbf{v}\,. \tag{3.11}$$

Another far-reaching classical result is the equivalence with the Lie derivative:

$$[\mathbf{v}, \mathbf{w}] = \mathcal{L}_{\mathbf{v}}\mathbf{w}\,. \tag{3.12}$$

This connection between the Lie bracket and vector flows will be further explored in section 8.2.

4. Brain Anatomy

Before we can discuss the various signal models in the next chapter, a brief introduction of basic brain anatomy will be helpful. The rest of this work will require an understanding of the different tissue types of grey and white matter, as well as the corticospinal fluid. A main topic will also be the shape of fiber bundles in the white matter.

A very good reference is [28]. The images of figure 4.2–4.4 below were adapted from the 20th edition of this book, which was released into the public domain and is available online.

4.1. Neurons

The average human brain comprises about 80 to 90 billion neurons and 100 to 500 billion glia cells. Neurons are connected, forming a complex network that can transmit signals. This allows for simple signal processing, controlling basic body functions, as well as higher mental tasks, like memory, decision making, personality and consciousness.

The glia cells are often described as the glue, holding the neurons together, supporting and maintaining the infrastructure of the brain. Although, recent research also finds more evidence for their active role in the neurotransmission process.

Neurons have a cell body, called *soma*, containing the nucleus. Thin, tentacle-like structures emerge from the soma. These are called *axons* and *dendrites*. Axons emit signals from the soma outwards, while dendrites receive signals. A typical neuron has one axon and multiple dendrites, both types usually branching out into complex tree structures (see fig. 4.1). Most connections between neurons are formed between the axons and dendrites. These touch in so-called *synapses*.

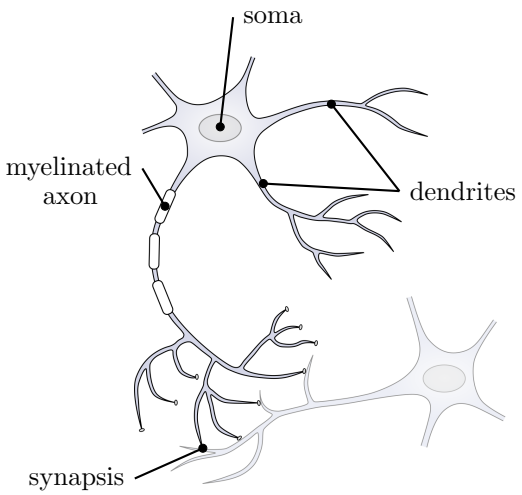


Figure 4.1.: Simplified neuron structure and basic synaptic connections.

Signals are transported in the form of electric potential. Non-excited neurons maintain a negative potential of ca. $-70\,\mathrm{mV}$ through exchange of ions with the surrounding medium. When excited, the potential shortly becomes more positive before returning to $-70\,\mathrm{mV}$. This pulse travels along the axon and triggers the release of *neurotransmitters* in the synapses. These chemicals are

able to cross the gap to connected neurons where they are absorbed and cause an increase or decrease in its dendrites' potential. These changes received from multiple synapses add up. If the overall potential rises above a certain threshold, the neuron also starts the excitation process.

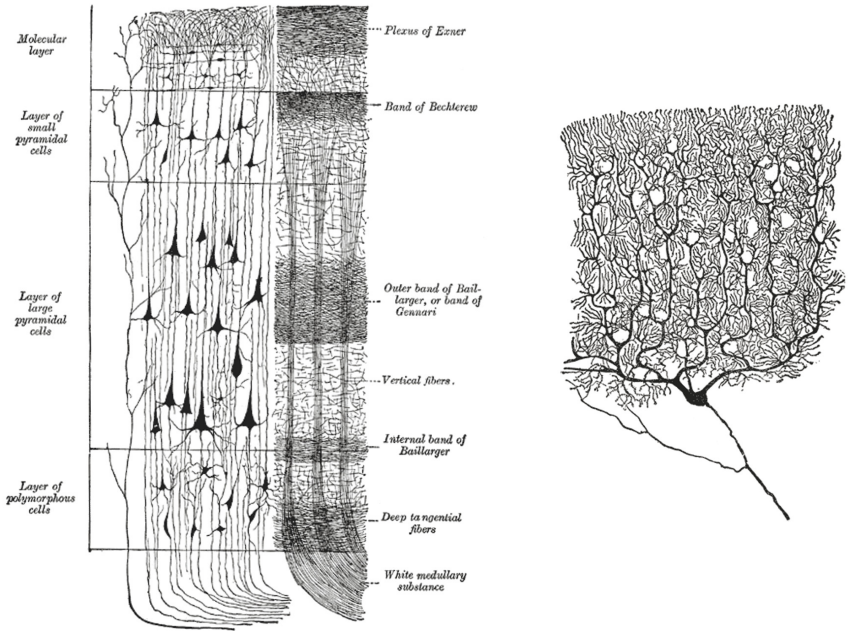


Figure 4.2.: Examples of more complicated neuron structures. *Left*: Different layers in the grey matter. *Right*: Single neuron in the cerebellum with branching dendrites.

Axons typically have a diameter between 1 and 15 μm. Their length can vary greatly. Most axons only connect neighboring neurons and are therefore similarly short as the dendrites. Some axons on the other hand can attain extreme lengths, connecting different parts of the brain and even running along the spinal cord and reaching the toes.

Longer axons are often coated in *myelin*, a lipid-rich material from specialized glia cells. Tissue with a high concentration of myelinated axons is called **White Matter** (WM) because of its distinctive white color. White matter will be the main interest of this thesis. It contains **fiber bundles**, groups of long axons connecting different regions of the brain.

4.2. Anatomy

Overview

The brain can roughly be divided into three parts - the hind-, mid- and forebrain. For humans, the forebrain is by far the largest part and is mostly composed of the two heavily wrinkled cerebral hemispheres. The hindbrain contains the very fractal looking

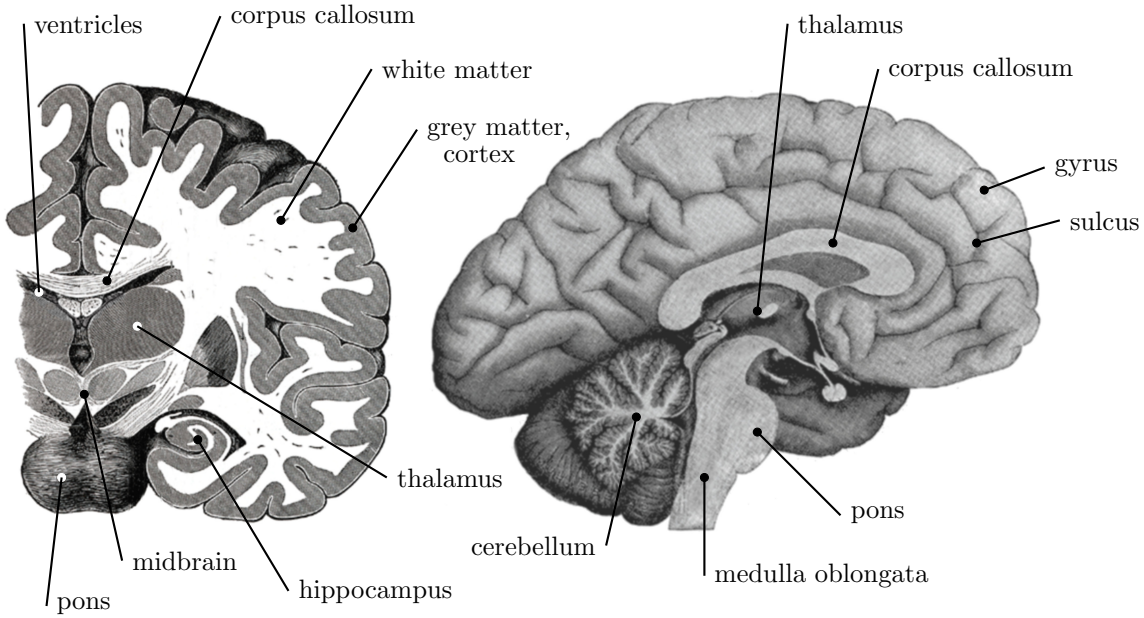


Figure 4.3.: *Left*: Coronal cut through the brain showing white and grey matter. *Right*: Sagittal cut with the corpus callosum in the center.

cerebellum as well as the *pons* and *medulla oblongata*. The later two together with the midbrain form the brain stem which connects to the spinal cord.

Beneath the cerebral hemispheres, the forebrain also contains smaller structures. Most prominent is the *corpus callosum*, which acts as a bridge between both hemispheres, as well as the *thalamus* and *hypothalamus*. Also long cavities can be found in each hemisphere and the central plane between them. These are the *ventricles* which are filled with **Corticospinal Fluid**[1] (CSF).

Cerebral hemispheres

The largest structures in the brain are the two cerebral hemispheres. The outer layer is the *cerebral cortex* and is formed by **Grey Matter** (GM) - tissue with a high density of short-ranged connections with a high degree of branching. The cerebral cortex has a very distinctive, folded shape - nature's solution to the problem of fitting a large surface area of ca. $2\,500\,\mathrm{cm}^2$ into the small volume of the skull cavity. The protruding ridges are called *gyri*, the grooves are the *sulci*.

Beneath a ca. $2.5\,\mathrm{mm}$ thick layer of grey matter comes a larger volume of white matter. Myelinated axon fibers transport signals between different parts of the cortex. The majority of connections are confined to each hemisphere. But there are also fibers bridging the fissure between the hemispheres forming the thin, $10\,\mathrm{cm}$ wide corpus callosum

[1] Also called *cerebrospinal fluid*.

and other, smaller commissures.

Deep inside the white matter, one can find *nuclei*, smaller structures of grey matter. The thalamus and *basal ganglia* belong to this category.

Fibers

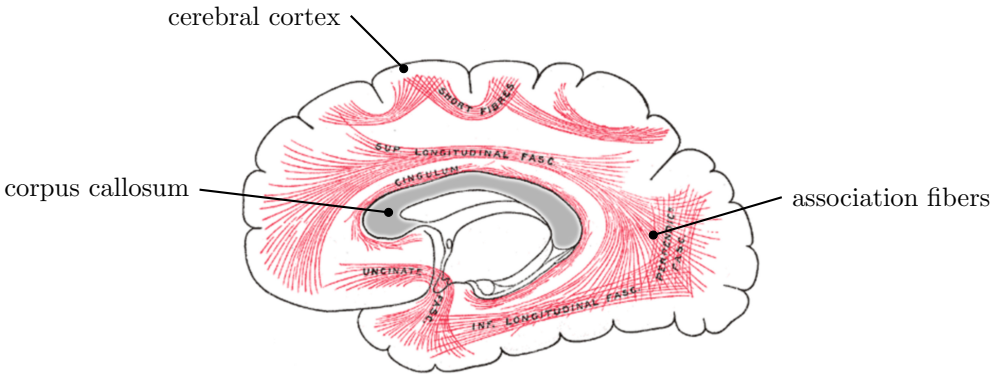


Figure 4.4.: Sagittal cut through the forebrain showing association fibers (red) in the white matter.

Fiber bundles in the white matter fall into one of three categories:

- **Association fibers** (fig. 4.4) stay within one hemisphere. Their length varies greatly as they can either connect neighboring sulci or distant parts of the hemisphere. They run mostly tangential to the surface
- **Commissural fibers** connect the two hemispheres. They run through the corpus callosum and the smaller commissures through the center of the brain and spread out radially in both hemispheres to the surface of the cerebral cortex.
- **Projection fibers** form connections between the cerebral cortex and other parts of the brain, like the brain stem. A larger number runs through the thalamus and the basal ganglia, radiating to the surface.

The next chapter will introduce the diffusion weighted MRI method to measure water diffusivity in the brain. Even though it is not feasible to scan the whole brain at a resolution that can resolve single axons on a μm-scale, the dominant direction of fiber bundles is clearly visible in a typical scan of 1 mm resolution.

Different mathematical tools will be discussed throughout this thesis to estimate fiber bundle directions from a scan. The goal will be to increase the angular accuracy, especially in areas where two or three fiber bundles cross. Chapter 6 introduces additional scalar measures that are influenced by each individual fiber bundle's shape.

5. Diffusion Weighted MRI

The method of **Diffusion-Weighted Magnetic Resonance Imaging** (DW-MRI or DWI or dMRI) has proven to be an invaluable tool in brain tissue research. It allows acquiring three-dimensional images of living subjects without significant risk. Even though a compromise between image resolution and acquisition time is necessary, clinical practice and research of structures within brain has made tremendous progress due to accessibility of information.

DW-MRI is based on the physical effect of Nuclear Magnetic Resonance (NMR) which was pioneered around 1940 by Isidor Rabi, Felix Bloch and Edward Purcell. During the 1970s Paul Lauterbur and Peter Mansfield improved on existing imaging techniques (MRI), laying the foundations for first clinical applications in the 1980s.

5.1. MRI

MRI uses interactions between nuclear magnetic moments of particles with external magnetic fields to create three-dimensional images. Specifically, the hydrogen nuclei of water molecules inside the human brain tissue will be of interest. These are made up of a single proton carrying a property called spin-vector $\mathbf{s}$ of length $\|\mathbf{s}\| = \frac{\sqrt{3}}{2}\hbar$ generating a magnetic moment $\boldsymbol{\mu} = \gamma\mathbf{s}$ with the reduced Planck's constant $\hbar \approx 1.0546 \cdot 10^{-34} Js$ and the gyromagnetic moment $\gamma \approx 267.5 \cdot 10^6 \frac{\text{rad}}{sT}$ of protons.

Since this work is not directly concerned with the detailed process of image acquisition, we will only give a small introduction into the theory behind it. More exhaustive explanations and be found in the well-written books [52] and [29].

For the rest of this chapter, spin will be described by classical mechanics. The terms *spin* and *magnetic moment* will even be used interchangeably. For the more curious readers, appendix A presents a brief introduction into the quantum mechanical treatment of spin and its connections to spherical harmonics from section 2.3.

Spins interacting with magnetic fields

Without an external magnetic field, the spins or magnetic moments of water molecules inside a volume of material are oriented randomly cancelling each other out. The net magnetization of a volume V is therefore $\mathbf{M} = \frac{1}{V}\sum \boldsymbol{\mu} = 0$.

In an external field $\mathbf{B}$, each magnetic moment $\boldsymbol{\mu}$ will experience a torque $\boldsymbol{\mu} \times \mathbf{B}$ and rotate along the axis of $\mathbf{B}$ with **Larmor frequency** $\omega = \gamma B$. For protons in a typical field of $1T$, this corresponds to a radio frequency $f = \frac{\omega}{2\pi} = 42.58\text{MHz}$.

When applying a constant field $\mathbf{B}$ over a longer time, we will observe a second effect. Each moment's orientation defines an energy level $E = -\langle \boldsymbol{\mu}, \mathbf{B} \rangle$. Due to thermal motion

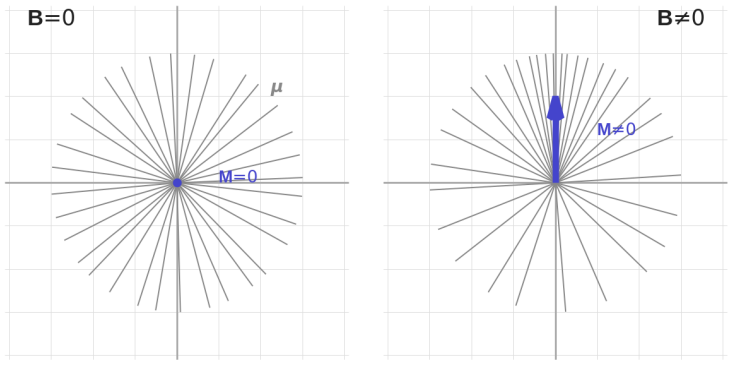


Figure 5.1.: Thermal equilibrium spin alignment altered by an external field $\mathbf{B}$ (*right*) creating net magnetization $\mathbf{M}$.

at room temperature, not all the spins will be aligned with $\mathbf{B}_0$, minimizing the energy. The system will instead end up in a thermal equilibrium with only a very small preference of spins along $\mathbf{B}$. This results in a small but measurable net magnetization $\mathbf{M}_0 \neq 0$ in the direction of $\mathbf{B}$.

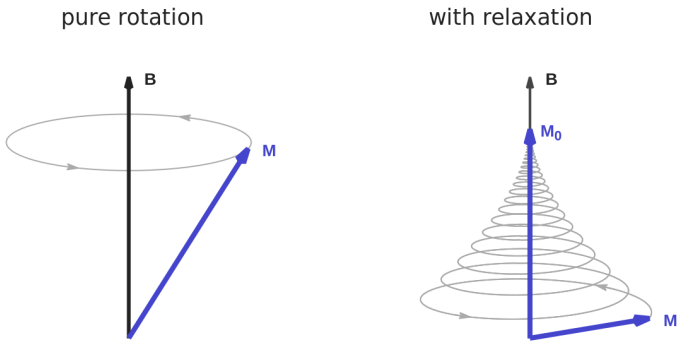


Figure 5.2.: *Left*: magnetization $\mathbf{M}$ will precess around an applied $\mathbf{B}$ field. *Right*: additional relaxation effects as described by the Bloch equation (eq. (5.1)).

The relaxation process into thermal equilibrium is rather complex and is caused by several effects. Phenomenologically, all this complexity is modelled by the simple Bloch equation

$$\dot{\mathbf{M}} = \mathbf{M} \times \gamma \mathbf{B} - \frac{\mathbf{M}_\perp}{T_2} - \frac{\mathbf{M}_\parallel - \mathbf{M}_0}{T_1} . \tag{5.1}$$

It describes the relaxation of the longitudinal and transversal components of $\mathbf{M}$ as exponential decay with the time parameters T_1 and T_2 respectively. $\mathbf{M}_0$ is the equilibrium magnetization.

Longitudinal relaxation is mostly governed by thermal fluctuations causing energy exchange with the surrounding material. That is why T_1 is called the **spin-lattice** time

constant. Common values range from 100 to 1500ms and can depend on $\mathbf{B}$.

Transversal relaxation is also affected by the same processes. Additionally, field inhomogeneities in $\mathbf{B}$ as well as susceptibility differences in the surrounding material can locally change the Larmor frequency. Since $\mathbf{M}_\perp \neq 0$ requires that the $\boldsymbol{\mu}$ exhibit a small preference for an angular phase due to exterior excitation, this phase preference will vanish quickly, if the $\boldsymbol{\mu}$ rotate with different frequencies. We will come back to this effect when discussing spin echos. T_2 is called the **spin-spin** time constant and ranges between 20 and 300ms with typical values of 100ms in brain tissue. We can always expect $T_2 \leq T_1$.

Radio pulses

Now let us observe the effects of a short radio signal $\mathbf{B}_{HF}$ added to the constant background field $\mathbf{B}_0$. To simplify explanations, we will align $\mathbf{B}_0$ along the z-axis. The radio signal is supposed to rotate in the x-y-plane and be of the form

$$\mathbf{B}_{HF}(t) = B_{HF}(t)\left[\cos(\omega_{HF}t)\,\mathbf{e}_x + \sin(\omega_{HF}t)\,\mathbf{e}_y\right]. \tag{5.2}$$

For the short duration of the pulse, the relaxation terms in eq. (5.1) can be ignored. Also, the situation simplifies when going from the static laboratory coordinate frame into the a rotating frame defined by the Larmor frequency ω_0 of $\mathbf{B}_0$. In this frame, the effects of $\mathbf{B}_0$ are trivial and the radio signal $\mathbf{B}_{HF}$ effectively rotates with frequency $\omega_{HF} - \omega_0$.

If the radio frequency ω_{HF} matches the Larmor frequency ω_0, then the radio field's direction is constant in this frame and the magnetization $\mathbf{M}$ experiences rotation around this fixed axis in the rotating x-y-plane. Intensity and duration of the radio pulse can easily be adjusted to cause a rotation by any given angle, for example pushing $\mathbf{M}$ by 90° into the x-y-plane.

After such a pulse, the excited transversal magnetization will keep rotating with ω_0 in the static frame until its decay due to T_2-relaxation. Since a rotating magnetic moment emits radio waves, we can **detect a signal** with a receiver coil tuned to ω_0.

Note, that if the frequency ω_{HF} does not match ω_0, its effects in the rotating frame will be more complicated and parts of the rotation will cancel each other out, resulting in smaller $\mathbf{M}_\perp$ and lower signal output.

Gradients

So far, the signal detectable after a 90° radio pulse has no spatial information. If $m(x, y, z)$ is the signal density of the material, then the overall received signal is the pure integral of all material in the scanner volume

$$S = \int_V m(x, y, z)\, dV\,. \tag{5.3}$$

For imaging purposes, this is bad. There are several common techniques to solve this problem by adding gradients to the field $\mathbf{B}_0$. The intricate details of coil design necessary for this task will be completely glossed over. For an introduction, see [32].

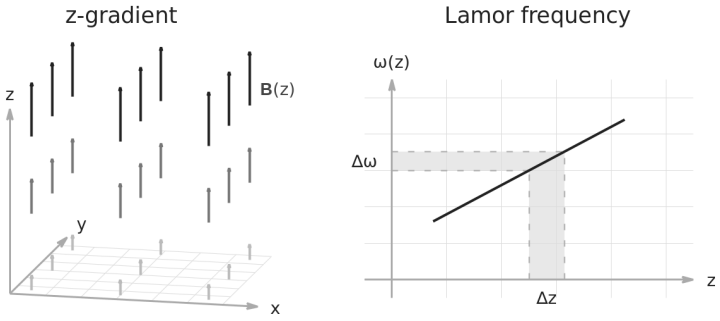


Figure 5.3.: Slice selection: A field gradient in the z-direction giving the Larmor frequency a z-dependency. A radio pulse will only be in resonance with a small slice around z.

A simple gradient in the z-direction accompanying the radio pulse for example results in a background field $\mathbf{B}_0(z) = (B_0 + G_z z)\mathbf{e}_z$ and also makes the Larmor frequency $\omega_0(z) = \gamma(B_0 + G_z z)$ dependent on the location as seen in fig. 5.3. This allows to **select a thin slice** around a specific z-value by tuning the radio pulse ω_{HF} to it. Material outside this slice is far less affected by the pulse.

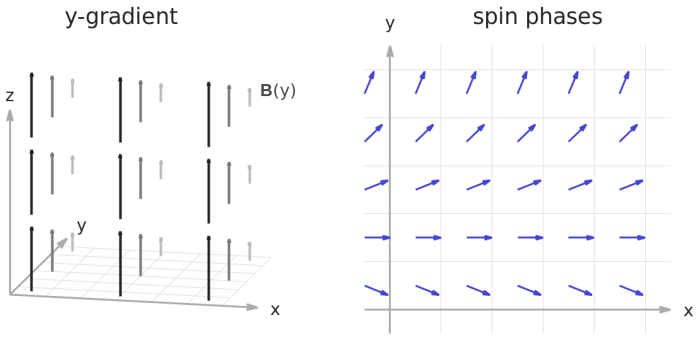


Figure 5.4.: Phase encoding: A y-gradient of $\mathbf{B}$ changes the rotation frequency $\omega(y)$ and gives phase differences along the y-axis after the gradient pulse.

A second gradient G_x can be applied during the signal readout, giving every x-slice a different output frequency (fig. 5.4). This is called **frequency encoding**. A simple Fourier transform of the received signal allows to separate the signals from each x-slice.

Directly after the initial 90° radio pulse, $\mathbf{M}$ rotates in the x-y-plane with constant frequency ω_0. Applying another gradient G_y for a short duration causes the magnetic moments to rotate with different speeds depending on their y-location. Afterwards, their phase will still show this difference. This is called **phase encoding**. In complex notation, it has the effect of multiplying each location's output signal $m(y)$ by a complex phase factor before summing to produce the total signal. We recognize that the new

signal

$$\int_{y_0}^{y_1} m(y)\,dy \quad \rightarrow \quad \int_{y_0}^{y_1} m(y)\,e^{i\,\gamma G_y y}\,dy \tag{5.4}$$

is just the Fourier transform of $m(y)$ for a fixed "frequency" γG_y. Repeating the whole process for enough different values of G_y allows to reconstruct $m(y)$.

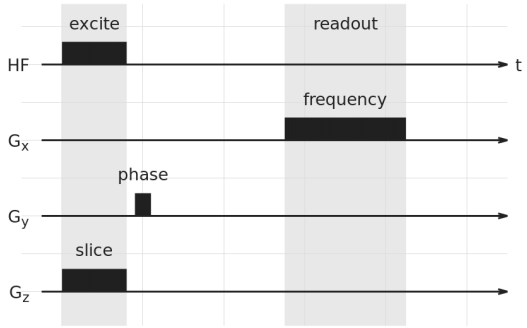


Figure 5.5.: Simplified MRI sequence: The radio pulse with a simultaneous gradient G_z excites spins within a selected z-slice. Afterwards a short G_y pulse shifts rotation phases to encode the y coordinate. During readout G_x is applied for frequency encoding.

By combining these three gradient techniques as shown in fig. 5.5, we can produce a three-dimensional image. $N_y N_z$ measurements are necessary for a grid of $N_x \times N_y \times N_z$ voxels.

Spin echo

As already mentioned, one main contributing factor to transversal relaxation are inhomogeneities in $\mathbf{B}_0$ and adding gradients enhances this problem even further. Immediately after the initial $90°$ pulse, the moments $\boldsymbol{\mu}$ point into the same direction in the x-y-plane, their angular phase in the plane is focused, but differences in rotation speed ω_0 will drag the individual phases apart, leading to loss of focus.

This process can be reversed by a second pulse, flipping by $180°$ and thereby inverting the phases. Performing this Δt after the $90°$ flip will refocus the phases after another Δt (see [30]). Common terminology is that T_2 denotes the irreversible part of transversal relaxation and T_2' the reversible part due to inhomogeneities. In the Bloch equation eq. (5.1) the effective relaxation time $T_2^\star$ with

$$\frac{1}{T_2^\star} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{5.5}$$

is used.

Similar results can be achieved by inverting the gradient field after Δt. This creates a **gradient echo** if the gradients are the dominant source of inhomogeneities.

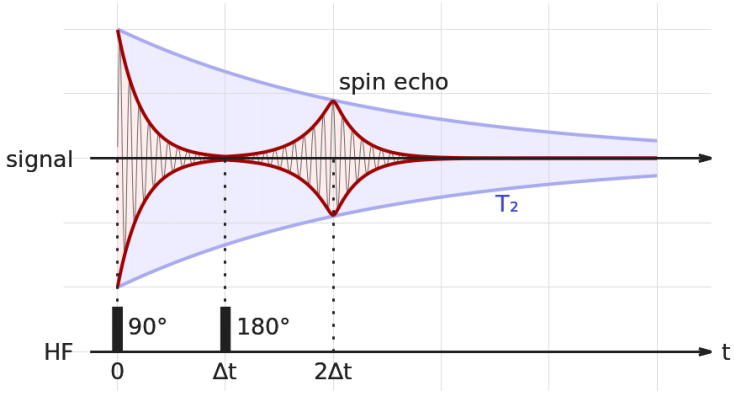


Figure 5.6.: A spin echo is created at $t = 2\Delta t$ after an initial $90°$ pulse at $t = 0$ and a refocusing $180°$ flip at $t = \Delta t$. The blue curve shows the irreversible T_2-relaxation.

5.2. Diffusion

Brownian motion

Diffusion is the transport process of particles in a medium from regions with higher concentration to lower concentration, not due to the material's overall momentum, but rather relative to it. A phenomenological description of this process is Fick's first law

$$\mathbf{J} = -D\,\nabla\rho \tag{5.6}$$

and the resulting diffusion equation for $D = \text{const}$

$$\dot{\rho} = D\,\Delta\rho\,. \tag{5.7}$$

$\rho(\mathbf{r}, t)$ is the particle concentration, $\mathbf{J}$ the flux due to diffusion, and the material's diffusivity D describes how freely the particles can move through the medium.

It is usually explained microscopically by Einstein's model[22] of Brownian motion. In this theory the particles follow random trajectories due to collisions with other particles. The resulting motion is far too complex to resolve for individual particles, but some macroscopic conclusions are possible. Each particle will travel a certain distance $\delta\mathbf{r}$ during a time interval τ. Dealing with a very large number of particles, this gives a probability distribution $P(\delta\mathbf{r})$ of displacements. P is also called the **propagator**, since it describes the probability of particles at position $\mathbf{r}$ to move to position $\mathbf{r} + \delta\mathbf{r}$ during τ.

Using kinetic gas theory, Einstein was able to recover the diffusion equation as a low order approximation of his model and predict concrete values for the diffusivity D. Useful information can be extracted from an analog of his calculations. We start by formulating the propagator's definition as an integral over displacements and performing a Taylor

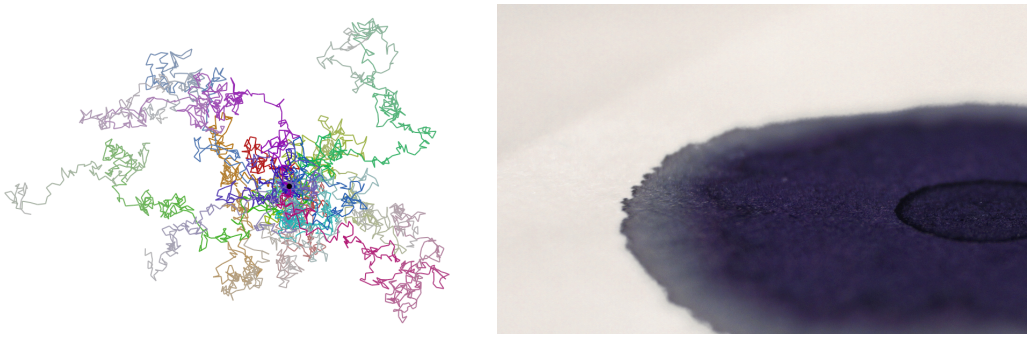


Figure 5.7.: Diffusion as a result of random particle motion in a medium.

expansion of ρ in $\mathbf{r}$:

$$\rho(\mathbf{r}, t+\tau) = \int_{\mathbb{R}^3} \rho(\mathbf{r}-\delta\mathbf{r}, t)\, P(\delta\mathbf{r})\, d\delta\mathbf{r} \tag{5.8}$$

$$= \int_{\mathbb{R}^3} \left(\rho(\mathbf{r}, t) - \sum_i \frac{\partial \rho}{\partial x_i}\delta x_i + \frac{1}{2}\sum_{i,k} \frac{\partial^2 \rho}{\partial x_i \partial x_k}\, \delta x_i\, \delta x_k + \ldots \right) P(\delta\mathbf{r})\, d\delta\mathbf{r} \tag{5.9}$$

$$= \rho(\mathbf{r}, t)\cdot 1 + 0 + \sum_{i,k} \frac{\partial^2 \rho}{\partial x_i \partial x_k} \underbrace{\left(\int_{\mathbb{R}^3} \frac{\delta x_i\, \delta x_k}{2} P(\delta\mathbf{r})\, d\delta\mathbf{r} \right)}_{=\tau D_{i,k}} + 0 + \ldots \tag{5.10}$$

The last step used normalization of P and its assumed symmetry $P(-\delta\mathbf{r}) = P(\delta\mathbf{r})$. Comparing this result to the Taylor expansion of ρ in t gives a generalization of the familiar diffusion equation:

$$\dot{\rho} = \sum_{i,k} D_{i,k} \frac{\partial^2 \rho}{\partial x_i \partial x_k} + \text{higher terms} \tag{5.11}$$

$D_{i,k}$ are the components of the **diffusion tensor** which is elegantly predicted by this theory to be the particle average of square displacements:

$$\boldsymbol{D} = \frac{1}{2\tau} \left\langle \delta\mathbf{r} \otimes \delta\mathbf{r} \right\rangle_P \tag{5.12}$$

The simple diffusion equation eq. (5.7) is recovered for isotropic tensors $\boldsymbol{D} = D\mathbf{1}$. On the other hand, eq. (5.11) suggests, that even the anisotropic tensor $\boldsymbol{D}$ is just an approximation to the general diffusion behavior. Even worse. P was assumed independent of the location $\mathbf{r}$ inside the observed volume. Hard boundaries that locally limit the movement of particles will create indisputable $\mathbf{r}$-dependence.

For $D = \text{const}$ the diffusion equation simplifies to eq. (5.7) and can be solved analytically. Particles concentrated at a single point will dissolve into a **Gaussian distribution** around that point after τ. This Gaussian behavior with its exponential decay will be

studied extensively in chapter 6 where we discuss ways to quantify deviations from this simple assumption.

But first, diffusion needs to be measured.

Effect on MRI

Ignoring the difficulties of measurement, let's assume we could record the ratio between the signal S at $t = 0$ immediately after the $90°$ pulse and the echo $S(t = 2\Delta)$ in a simple spin-echo experiment. This ratio is governed by the exponential relaxation with T_2. If a gradient field is present during this time interval, the echo will decay even further because some of the water molecules carrying a magnetic moment $\boldsymbol{\mu}$ will move due to diffusion. Movement parallel to the gradient changes the Larmor frequency and reduces the focusing effect of the $180°$ pulse[85].

Stejskal and Tanner calculated[80] this effect on the spin echo intensity $S = S(t = 2\Delta)$ for a sequence of two short gradient pulses of duration δ before and after the $180°$ pulse:

$$S(\mathbf{g}, b) = S_{b=0}\, e^{-bD} \tag{5.13}$$

with the diffusivity D in the gradient direction and by setting

$$b = \gamma^2 G^2 \delta^2 \left(\Delta_G - \frac{\delta}{3} \right) . \tag{5.14}$$

Δ_G is the time between gradient pulses of strength G.

As described before, diffusion processes even in a small volume of tissue can be complex. Different media with different diffusion properties are present and interact with the tissue's microstructure. To account for all these effects, eq. (5.13) is still used for complex materials while D is interpreted as the **Apparent Diffusion Coefficient** (ADC).

Eq. (5.13) forms the basis of **Diffusion Weighted MRI** (DW-MRI). It allows to extract information about tissue diffusion properties and will be crucial to all further discussions. Also, it will shield us from the gritty engineering details of the MRI process.

Q-space and shells

The main objective is to measure the function $D(\mathbf{g}, b)$, or rather the ratio $S(\mathbf{g}, b)/S_{b=0}$ between diffusion-weighted and non-diffusion-weighted spin echo signals and find good models for it. Since $\mathbf{g}$ is a unit vector and $b \geq 0$, these functions can be parametrized by "b-vectors" $\mathbf{b} = b\,\mathbf{g} \in \mathbb{R}^3$. It is also common to use the alternative scaling

$$\mathbf{q} = q\,\mathbf{g}, \qquad \text{with } q = \gamma G \delta\,. \tag{5.15}$$

The space of these vectors is called **q-space**.

Ideally, one would want to measure $S(\mathbf{g}, b)$ on the whole q-space but of course this is not possible. So there are several commonly used acquisition schemes that sample the signal at few discrete points while still allowing to get enough useful information. These schemes range from the bare minimum of 7 measurements for finding the diffusion tensor

to sampling q-space on a large Cartesian grid in **Diffusion Spectrum Imaging** (DSI). The gap between these extremes is filled by **High Angular Resolution Diffusion Imaging** (HARDI)[92].

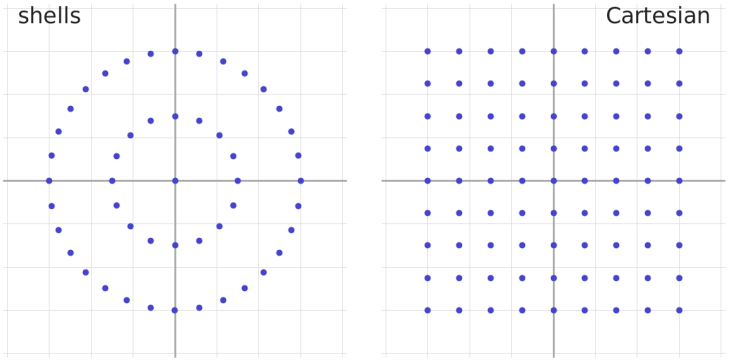


Figure 5.8.: Q-space sampling schemes (2d for illustration); *Left*: multi-shell with two shells, *right*: on a Cartesian grid.

Although DSI acquisition requires a large number of measurements making it highly time-consuming, it offers a big advantage. The signal in q-space is a good approximation of the Fourier transformed propagator P. This means, DSI allows to reconstruct the propagator via inverse Fourier transform instead of just giving simple diffusivities.

One very popular category of HARDI schemes uses **shells**. Besides measurements for $b = 0$, a small number of discrete b-values is selected and for each the associated sphere in q-space is sampled. Single- and multi-shell schemes are differentiated because these might require different methods for reconstruction.

5.3. Signal Models

To gain any information about the scanned materials, the acquired signal needs to be interpreted. For some simple structures it is possible to calculate how this signal would look, this can be fitted to the measurement to extract certain parameters of the material. Of course creating these mathematical models involves a great deal of simplification and heuristics.

On the other hand, models do not have to stem from real structures and can be rather abstract. In most cases the goal is to distinguish and quantify different structures and one is not directly interested in their underlying details.

Diffusion tensor

The simple second order diffusion tensor from eq. (5.12) yields

$$S(b, \mathbf{g}) = S_0 \, e^{-b \, \boldsymbol{D}(\mathbf{g})} \tag{5.16}$$

with $\boldsymbol{D}(\mathbf{g}) = \sum_{i,k} D_{ik} \, g_i g_k$.

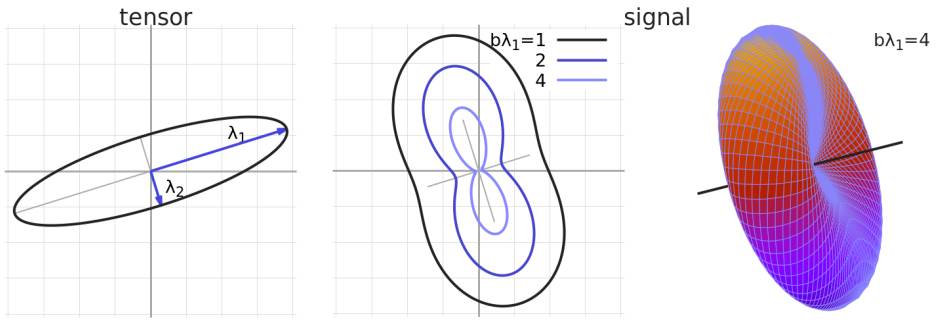


Figure 5.9.: *Left*: Ellipsoid of a diffusion tensor $\boldsymbol{D}$ with its eigenvalues λ_1, λ_2, *center, right*: A radial plot of the associated signal for different b-values.

We will denote the eigenvalues of $\boldsymbol{D}$ by $\lambda_1 > \lambda_2 > \lambda_3$. From the physical considerations in the previous section, $\boldsymbol{D}$ can be assumed to be positive definite: $\lambda_{1,2,3} > 0$. Other simple scalar measures derived from the eigenvalues include the **Mean Diffusivity** (MD)

$$\overline{D} = \frac{1}{3}\operatorname{tr}\boldsymbol{D} = \frac{1}{3}(\lambda_1 + \lambda_2 + \lambda_3) \tag{5.17}$$

and **Fractional Anisotropy** (FA)

$$\mathrm{FA} = \sqrt{\frac{(\lambda_2 - \lambda_1)^2 + (\lambda_3 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{5.18}$$

as well as the linear, planar and spherical **Westin measures** (fig. 5.10)

$$c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5.19}$$

$$c_p = 2\frac{\lambda_2 - \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5.20}$$

$$c_s = 3\frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}\,. \tag{5.21}$$

FA measures the deviation from an isotropic tensor. High anisotropy is often associated with the presence of a dominant fiber direction. The Westin measures additionally characterize the type of anisotropy - linear or planar.

Ball-and-stick model

A single diffusion tensor only describes the ideal Gaussian diffusion of a single type of uniform material and can not accurately describe crossing fibers. A better approach would be a multi-tensor model with a linear combination of DTI signals:

$$S(b, \mathbf{g}) = S_0 \sum_i f_i\, e^{-b\,\boldsymbol{D}_i(\mathbf{g})} \tag{5.22}$$

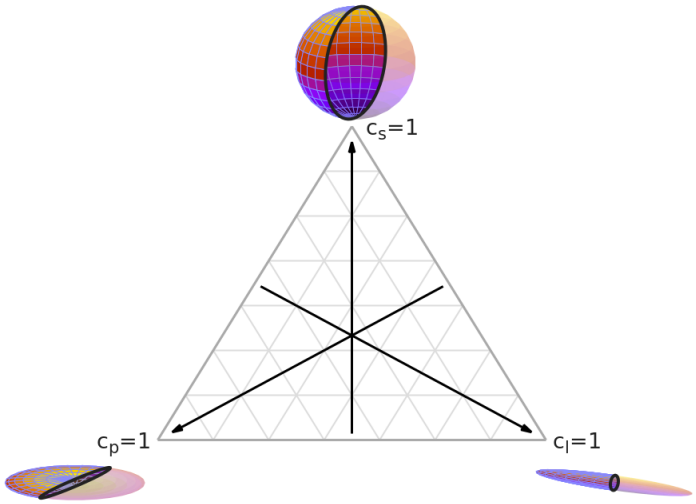


Figure 5.10.: Tensor anisotropy: spherical, linear and planar tensors.

Unfortunately this model has redundant parameters making it unsuitable for fitting.

To get rid of the redundancy, the tensors can be restricted. One tensor is assumed isotropic ("ball" $\boldsymbol{D} \propto \mathbf{1}$) while the others are linear ("stick" $\boldsymbol{D} \propto \mathbf{v} \otimes \mathbf{v}$) modeling the fibers, sharing a common diffusivity parameter:

$$S(b, \mathbf{g}) = S_0 \left(f_{\text{ball}}\, e^{-b\,d} + \sum_{\text{sticks } i} f_i\, e^{-b\,d\langle \mathbf{g}, \mathbf{v}_i \rangle^2} \right) \tag{5.23}$$

with $f_\star \geq 0$ and $\sum f_\star = 1$.

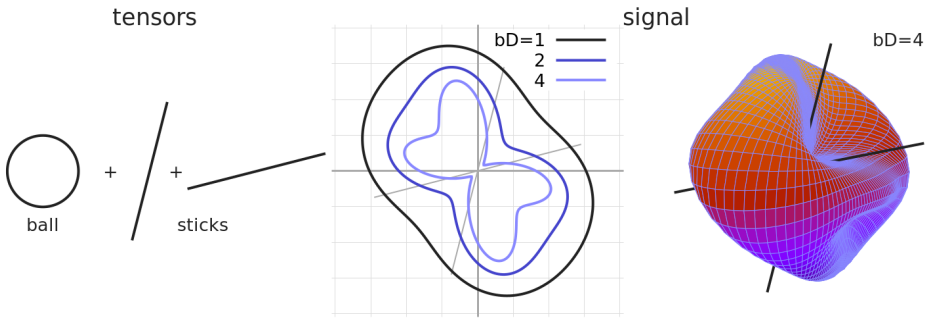


Figure 5.11.: *Left*: Ball-and-stick model mixing one ball and two sticks and the generated signal (*center*, *right*).

The sticks can be generalized to rotationally symmetric tensors by adding a radial component $-b(d_\parallel \langle \mathbf{g}, \mathbf{v}_i \rangle^2 + d_\perp (1 - \langle \mathbf{g}, \mathbf{v}_i \rangle^2))$.

There are also several alternatives for the isotropic part of the signal besides the ball compartment $S/S_0 \propto e^{-b\,d}$. Most commonly used is the **dot compartment** $S/S_0 \propto 1$ representing completely constrained particles.

Kurtosis tensor

Another generalization of the diffusion tensor model is the usage of higher order terms, especially the inclusion of a fourth order tensor as proposed by Jensen et. al[36]:

$$\ln \frac{S(b,\mathbf{g})}{S_0} = -b\,\boldsymbol{D}(\mathbf{g}) + \frac{b^2}{6} D^2_{\text{mean}} \boldsymbol{W}(\mathbf{g}) \tag{5.24}$$

with the kurtosis tensor $\boldsymbol{W}(\mathbf{g}) = \sum_{i,k,l,m} W_{iklm}\, g_i g_k g_l g_m$.

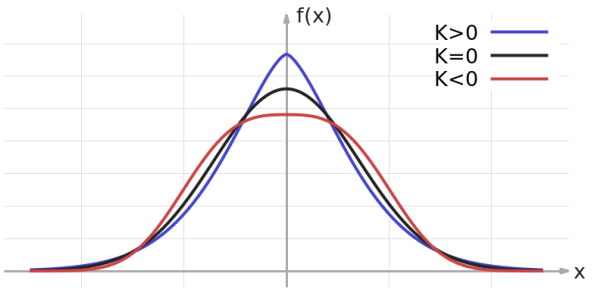


Figure 5.12.: Distribution functions with equal means and variances but different kurtosis values.

In statistics (excess) kurtosis quantifies the deviation of a probability distribution from Gaussianity. In case of diffusion with displacement $\delta\mathbf{r}$ and distribution $P(\delta\mathbf{r})$ its value in the direction $\mathbf{n}$ is

$$K(\mathbf{n}) = \frac{\left\langle \langle \mathbf{n}, \delta\mathbf{r}\rangle^4 \right\rangle_P}{\left\langle \langle \mathbf{n}, \delta\mathbf{r}\rangle^2 \right\rangle_P^2} - 3\,. \tag{5.25}$$

For the kurtosis tensor model eq. (5.24) this evaluates to

$$K(\mathbf{g}) = \frac{D^2_{\text{mean}}}{\boldsymbol{D}(\mathbf{g})^2}\, \boldsymbol{W}(\mathbf{g})\,. \tag{5.26}$$

Spherical deconvolution

While the ball-and-stick model describes a finite linear combination of fiber signals oriented along the directions $\mathbf{v}_i$, it loses accuracy if the fibers are not perfectly focused inside the measuring volume. This can be solved by the generalization to a continuous integral over an infinite number of fibers. The finite set of weights f_i for the fiber directions is replaced by a continuous density function $f(\mathbf{v})$ on the sphere called the

Fiber Orientation Distribution Function (ODF or fODF). We demand that $f \geq 0$, $\int_{S^2} f \, dA = 1$ and $f(-\mathbf{v}) = f(\mathbf{v})$.

For each shell of fixed b-value the signal can also be seen as a discretized function on a sphere $S : S^2 \to \mathbb{R}$. Each fiber is then described by a rotated copy of a single kernel function $K : S^2 \to \mathbb{R}$ on this shell with symmetry around the z-axis. The integral of rotated copies of K weighted by f is the convolution[75]

$$S = f \star K \,. \tag{5.27}$$

Because of Thm. 2.3 the convolution turns into a simple product when expressing these functions in the spherical harmonics basis. This also allows to easily invert the relation to find the fODF f for a given signal S and kernel K. This process is known as deconvolution. In practice it is more beneficial to express this as a linear least squares problem and impose additional constraints for f.

It is common to use a SH basis of degree 4 or 8 in calculations. Lower order results in loss of angular resolution while higher degrees are in most cases too susceptible to noise.

Note that degree n spherical harmonics are equivalent to order n tensors. But although the higher order tensor models define each tensor for all b-values, here all functions are only evaluated on a single shell. To describe a multi-shell signal, each shell requires its own kernel K_b. Section 7.2.1 describes a generalization of this model via the SHORE basis avoiding the need for a shell structure altogether.

Also, as argued in [70], naive deconvolution inverting eq. 5.27 can lead to artifacts and negativity. The fODF estimated from the signal of a single discrete fiber would be a δ-function projected onto the space of low degree SHs. This function exhibits high frequency ringing around the peak. Deconvolution becomes more stable when altering the kernel so that it maps the single fiber signal to a rank-1 tensor of order n when using the SH basis of degree n.

Part II.

Contributions

6. Quantifying Microstructure in Fiber Crossings with Diffusional Kurtosis

Abstract *Diffusional Kurtosis Imaging (DKI) is able to capture non-Gaussian diffusion and has become a popular complement to the more traditional Diffusion Tensor Imaging (DTI). In this paper, we demonstrate how strongly the presence of fiber crossings and the exact crossing angle affect measures from diffusional kurtosis, limiting their interpretability as indicators of tissue microstructure. We alleviate this limitation by modeling fiber crossings with a mixture of cylindrically symmetric kurtosis models. Based on results on simulated and on real-world data, we conclude that explicitly including crossing geometry in kurtosis models leads to parameters that are more specific to other aspects of tissue microstructure, such as scale and homogeneity.*

Comment *This chapter corresponds to the paper[4]: "Quantifying Microstructure in Fiber Crossings with Diffusional Kurtosis" by Michael Ankele et al. in proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015. For more details about the new kurtosis signal model and its relationship to DKI see section B.1. Also, a more efficient formula for mean kurtosis MK is given in eq.* (B.35) *in the appendix.*

6.1. Introduction

Diffusional Kurtosis Imaging (DKI) is a natural and popular extension of Diffusion Tensor Imaging (DTI) that accounts for the empirically observed non-Gaussianity of diffusion in biological tissue. Measures of diffusional kurtosis are known to be affected by factors such as the scale and homogeneity of obstacles to the molecular motion [35], and therefore provide useful information on tissue microstructure, complementing the information captured in the diffusion tensor.

Many studies of white matter are motivated by an interest in structural parameters, such as nerve fiber density or myelination. They use diffusion MRI because it provides quantities that are affected by such factors, and that are easy and safe to obtain *in vivo*. A known limitation of diffusion tensor imaging is the fact that measures such as fractional anisotropy are sensitive, but not specific to those parameters of interest: The effect of confounding factors, such as the presence of orientational dispersion or fiber crossings, can be substantial.

In Section 6.3 of this paper, we discuss an analogous limitation in DKI: We show that common measures of diffusional kurtosis are not specific to microstructural parameters of individual fibers, but are heavily affected by the presence, and the exact angle, of fiber

crossings. This motivates development of a novel computational method in Section 6.4, in which the impact of those nuisance parameters is greatly reduced. Its building block is a cylindrically symmetric kurtosis model. In Section 6.5, we present results on simulated data, confirming that our newly derived kurtosis measures are affected far less by the crossing angle than results of traditional DKI. We also show parameter maps that demonstrate the effectiveness of our method on real data.

6.2. Related Work

Several recent works have aimed to reduce the effects of fiber crossings and orientational dispersion on quantitative markers from diffusion MRI. NODDI disentangles the effects of neurite orientation dispersion and density, but does not model fiber crossings [102]. Spherical deconvolution can be used to quantify fiber properties in a way that is robust to fiber crossings, either by analyzing the fiber orientation distribution function after using a fixed deconvolution kernel [62, 19], or by calibrating the kernel itself [69]. Finally, estimation of per-compartment diffusion parameters can be integrated into crossing-fiber tractography [48, 63].

Our work is most closely related to a series of approaches that have fitted multiple diffusion tensors [92, 47, 82]. However, none of them model diffusional kurtosis. We demonstrate how fiber crossings affect kurtosis measures and propose novel kurtosis measures whose sensitivity to crossings is greatly reduced.

6.3. How Fiber Crossings Affect Diffusional Kurtosis

It is well-known that fiber crossings strongly affect measures derived from the diffusion tensor model, such as Fractional Anisotropy (FA). It is unsurprising that the same is true for measures of diffusional kurtosis. Our first goal is to systematically demonstrate the exact extent of this dependence.

We have synthesized crossings with varying crossing angles between 0 and 90 degrees, and created plots of how diffusional kurtosis depends on it. Signal synthesis was performed in a data-driven manner from a subject from the Human Connectome Project (288 DWIs on shells at $b \approx \{5, 1000, 2000, 3000\}\,\mathrm{s/mm^2}$). It is based on 300 voxels thought to contain a single dominant fiber compartment, given as the voxels with highest FA within a white matter mask. The DKI model was fit to the data, and model parameters were analytically rotated by the desired crossing angle. Diffusion-weighted signals were computed from the original and rotated model, and averaged. This simulates two fiber compartments that cross at a known angle, with no significant exchange within the diffusion time. Since we use the full DKI model, it does not impose cylindrical symmetry.

We performed a constrained least squares fit of the diffusional kurtosis model to the simulated data [81]. Fig. 6.1 plots mean and one standard deviation of the resulting mean, axial, and radial kurtosis, over the 300 voxels used to simulate the crossings. The plots in the top row show that the dependence on the crossing angle is substantial:

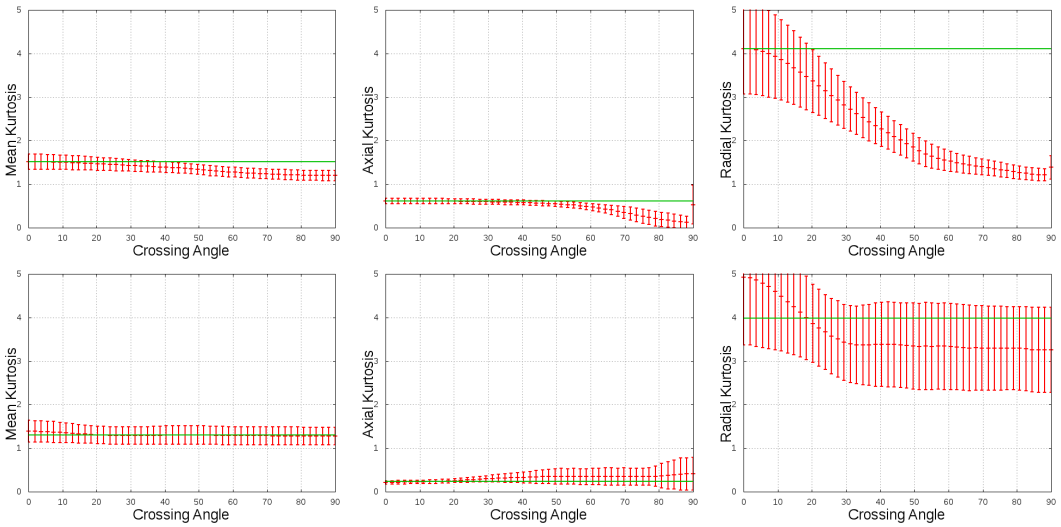


Figure 6.1.: *Top:* The angle at which fibers cross has a severe impact on the parameters of DKI. *Bottom:* The proposed model reduces effects of crossing geometry, leading to biomarkers that more specifically quantify microstructure properties.

Compared to the baseline (value for a single fiber compartment, indicated by a green line), the crossing changes radial kurtosis by a factor of up to 4.23, and axial kurtosis up to 5.65. A very similar dependence is observed for axial and radial diffusivities (not shown). Studies that aim to use kurtosis to specifically quantify microstructure properties, without confounding effects from crossings, should be aware of this problem. We will now introduce a method to compute novel kurtosis measures, shown in the lower row of Fig. 6.1, that are less susceptible to the undesired impact of crossing geometry.

6.4. A Mixture of Kurtosis Models

The general strategy of our method is to fit a mixture of kurtosis models to fiber crossings. In effect, this adds kurtosis to previous methods that have modeled fiber crossings using multiple diffusion tensors.

6.4.1. A Cylindrically Symmetric Kurtosis Model

The full kurtosis model has six parameters for the diffusion tensor, plus 15 for the kurtosis tensor. This seems prohibitive for fitting a mixture. We thus constrain the kurtosis models that will represent the individual fiber compartments in our mixture to be cylindrically symmetric around the principal diffusion direction.

The same symmetry is frequently assumed in multi-tensor models [92, 47]. It reduces the 21 parameters in the full kurtosis model to only 7: Two angles that parameterize a unit vector $\mathbf{v}$ indicating the fiber direction, axial and radial diffusivities ($\lambda_{\parallel}$, $\lambda_{\perp}$), as

well as three kurtosis-related parameters: In addition to $\kappa_\parallel$ and $\kappa_\perp$, which are related to axial and radial kurtosis, the fact that kurtosis is a fourth-order quantity introduced a third parameter $\kappa_\diamond$. The resulting signal equation as a function of gradient direction $\mathbf{g}$ and b value is

$$\begin{aligned}\ln \frac{S_{cyl}(\mathbf{g}, b; \mathbf{v})}{S_0} = &-b\left[\lambda_\perp + (\lambda_\parallel - \lambda_\perp)\langle \mathbf{v}, \mathbf{g}\rangle^2\right] + \\ &\frac{b^2}{6}\left[\kappa_\perp + (\kappa_\diamond - 2\kappa_\perp)\langle \mathbf{v}, \mathbf{g}\rangle^2 + (\kappa_\parallel - \kappa_\diamond + \kappa_\perp)\langle \mathbf{v}, \mathbf{g}\rangle^4\right]\end{aligned} \tag{6.1}$$

To simplify Eq. (6.1), κ_* absorb the square of the mean diffusivity $\bar{\lambda} = (\lambda_\parallel + 2\lambda_\perp)/3$ that usually occurs as a factor in diffusional kurtosis. This means that axial and radial kurtosis K_* can be computed from our parameters as

$$K_\parallel = \frac{\kappa_\parallel}{\lambda_\parallel^2} \qquad \text{and} \qquad K_\perp = \frac{\kappa_\perp}{\lambda_\perp^2}. \tag{6.2}$$

Assuming that the fiber is oriented along the z axis, our model parameters translate to a standard kurtosis tensor via

$$W_{xxxx} = W_{yyyy} = 3W_{xxyy} = \frac{\kappa_\perp}{\bar{\lambda}^2},\ W_{xxzz} = W_{yyzz} = \frac{\kappa_\diamond}{6\,\bar{\lambda}^2},\ W_{zzzz} = \frac{\kappa_\parallel}{\bar{\lambda}^2}\,. \tag{6.3}$$

This allows computation of mean kurtosis MK using the equations given in [81].

We have used the Bayesian Information Criterion (BIC) to compare our cylindrically symmetric model with two variants, one with a ball compartment, the other one with a dot compartment [56], and with the full kurtosis model. Ranking them with respect to their BIC preferred "symmetric+dot" in 96.3% of all cases, "symmetric" in 3.0%, the full kurtosis model in 0.7%, and "symmetric+ball" in 0%. Therefore, we include a dot compartment in all our experiments.

6.4.2. Strategy for Fitting the Final Mixture

Our final signal equation results from using Eq. (6.1) to model each of k crossing fiber compartments and adding the dot compartment:

$$S(\mathbf{g}, b) = S_0\left[f_{dot} + \sum_{i=1}^{k} f_i S_{cyl}(\mathbf{g}, b; \mathbf{v}_i)\right] \tag{6.4}$$

Volume fractions f_* are constrained to be non-negative, and to add to one. We ensure numerical stability in evaluating Eq. (6.2), and force $\mathbf{v}_i$ to align with a principal diffusion direction, by constraining $\lambda_\perp \in [0.01\lambda_\parallel, \lambda_\parallel]$ and $\lambda_\parallel > \epsilon$. We also impose the same constraints on our kurtosis parameters as Tabesh et al., $3/(b_{max}\lambda) \geq K \geq 0$ [81]. As in the widely used ball-and-stick model [9], the diffusion and kurtosis parameters of all compartments are coupled. Trying to obtain stable estimates without this constraint is a topic for future work.

SNR	Kurtosis Tensor Model			Our Kurtosis Mixture Model		
	MK	$K_\parallel$	$K_\perp$	MK	$K_\parallel$	$K_\perp$
∞	-0.15 ± 0.17	-0.09 ± 0.22	-1.69 ± 1.60	-0.00 ± 0.07	0.02 ± 0.07	-0.20 ± 1.01
40	-0.15 ± 0.28	-0.09 ± 0.21	-1.75 ± 1.76	0.01 ± 0.31	0.02 ± 0.08	-0.14 ± 1.62
30	-0.15 ± 0.28	0.17 ± 0.38	-1.95 ± 1.36	0.16 ± 0.40	0.20 ± 0.50	-0.36 ± 1.48
20	-0.04 ± 0.45	0.35 ± 0.51	-1.89 ± 1.35	0.42 ± 0.84	0.52 ± 1.00	-0.27 ± 1.95
10	0.49 ± 1.18	0.83 ± 0.96	-1.43 ± 2.02	1.79 ± 5.05	1.61 ± 2.52	1.88 ± 14.90

Table 6.1.: Statistics on the difference between kurtosis estimates in simulated crossings and the single fiber voxels from which they were generated quantify the extent to which we reduce the impact of crossings. At low SNR, neither model gives useful results.

Even though Eq. (6.4) is relatively straightforward conceptually, fitting it to a given set of measurements amounts to a difficult non-convex optimization problem. We have developed the following strategy for solving it: A suitable initialization is obtained from a diffusion tensor fit, by setting $\lambda_\parallel$ to the largest eigenvalue and $\lambda_\perp$ to the mean of the two smaller ones. Fiber volume fractions f_i and directions $\mathbf{v}_i$ are initialized by discretizing an orientation distribution function from spherical deconvolution, as proposed in [75]. The kurtosis parameters and f_{dot} are initialized to zero.

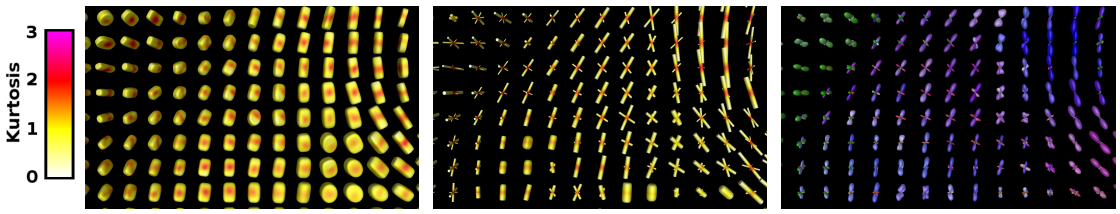
The fitting itself is performed with constrained Levenberg-Marquardt optimization. We found that it can be accelerated greatly by re-parametrizing diffusivity and kurtosis parameters. The actual parameters visible to the optimizer are $\ln(\lambda_\parallel)$, $\lambda_\perp/\lambda_\parallel$, $1000\kappa_\parallel/\lambda_\parallel$, $1000\kappa_\perp/\lambda_\perp$, $1000\kappa_\diamond/\lambda_\parallel$. Moreover, we observed that convergence benefits from splitting the parameters into two blocks, and alternating between their optimization. The first block contains the volume fractions and directions, which we parameterize using elevation and azimuth angles. The second block optimizes diffusion and kurtosis. Despite these optimization, processing a slice of 174×145 voxels on 6 cores of a 3.4GHz *i7* takes 6 minutes.

6.5. Results

6.5.1. Simulated Data

As an initial validation of our model and fitting procedure, we applied it to the simulated data that was described in Section 6.3. The results are shown in the bottom row of Fig. 6.1. They confirm that our crossing model succeeds in absorbing a significant part of the variation in kurtosis measures which is otherwise caused by crossing geometry. The results remain much closer to the baseline, which is indicated by the green line. Note that differences between the baselines in both rows are due to the presence of the dot compartment.

For a quantitative summary, we have taken the difference of kurtosis parameters estimated in the crossing by the two models, and a baseline, computed by the same method from the single-fiber voxel that was used to simulate the fiber crossing. Table 6.1 reports



(a) DKI (b) Our Model (c) CSD

Figure 6.2.: Our mixture of cylindrically symmetric kurtosis compartments (b) results in principal fiber directions that agree well with constrained spherical deconvolution (c). In contrast to the traditional kurtosis model (a), it leads to measures that disentangle the effects of microstructure and crossing geometry.

the mean and standard deviation of this difference over all 300 voxels and all crossing angles. It confirms that our model greatly reduces the impact of crossings, in particular in case of radial kurtosis. The relatively low standard deviations indicate that fitting works reliably. Table 6.1 also shows the results of adding Rician noise to the simulation, indicating that our fitting starts to degrade around $\mathsf{SNR} \approx 20$. At this point, even values from the full model start to exhibit a noticeable bias.

6.5.2. Real Data

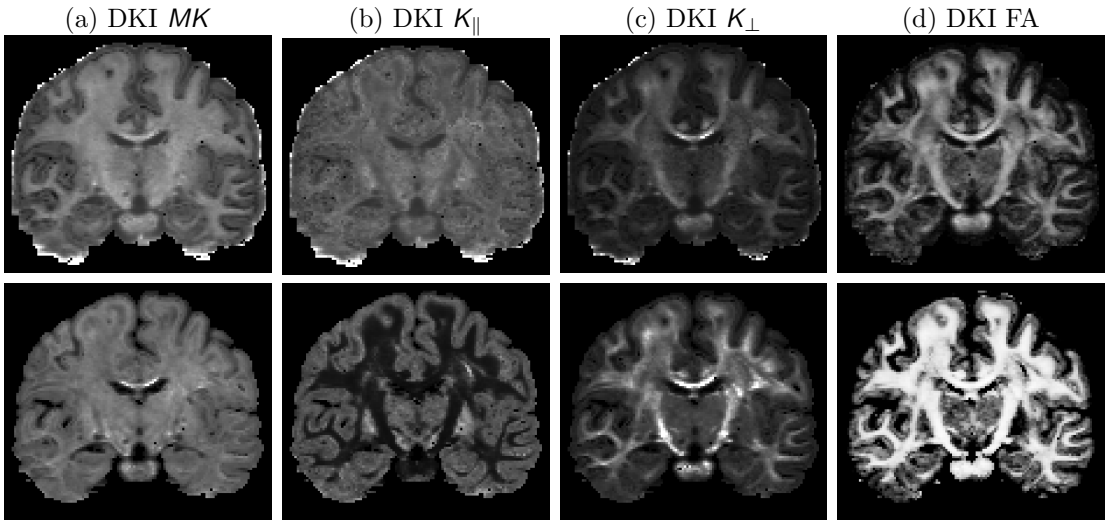
In addition to the quantitative validation on simulated data, we have verified that our model produces plausible results on real human brain scans by fitting it to data from the human connectome project. In each voxel, the BIC has been used to select between models with a single, two, or three cylindrically symmetric kurtosis compartments.

A detail of the result on a coronal slice, in the region where fibers from the corpus callosum, corticospinal tract, and superior longitudinal fasciculus cross, is visualized in Fig. 6.2 using superquadric glyphs [23] for the diffusion tensor part of the kurtosis model. Glyphs have been scaled with the volume fraction of the respective compartment, and color coded with directional kurtosis. In contrast to the traditional kurtosis model in (a), directions of crossing fibers are immediately apparent from our result (b).

A comparison to the widely used constrained spherical deconvolution model [87], which we fitted to the subset of measurements with $b \approx 3000\,\mathsf{s/mm^2}$, is shown in Fig. 6.2 (c). The agreement of principal fiber directions and relative volume fractions confirms that the individual kurtosis compartments in our model successfully capture the dominant fiber populations in real crossings.

However, our main interest is in the kurtosis measures themselves, which are mapped in the bottom row of Fig. 6.3 and compared to the corresponding ones from standard kurtosis imaging in the top row. MK and $K_{\parallel}$ are mapped with range $[0, 2]$; $K_{\perp}$ is mapped with range $[0, 5]$; FA is shown with range $[0, 1]$.

Within the white matter, our model measures a much lower $K_{\parallel}$ than the classical DKI model, close to that of free diffusion. In gray matter, our $K_{\parallel}$ remains high, providing a



(a) DKI MK (b) DKI $K_{\parallel}$ (c) DKI $K_{\perp}$ (d) DKI FA

(e) Our MK (f) Our $K_{\parallel}$ (g) Our $K_{\perp}$ (h) Our FA

Figure 6.3.: Differences in MK and $K_{\parallel}$ between our model and standard DKI appear to be due to including a dot compartment, while those in $K_{\perp}$ and FA are more strongly affected by the reduced impact of fiber crossings.

clear contrast between the two tissue types. This correlates with the volume fraction of the dot compartment; after factoring it out, MK is nearly uniform over the brain tissue (Fig. 6.3 (e)).

$K_{\perp}$ and FA have been computed from the diffusion tensors of both models. There is a clear visual similarity between structures in Fig. 6.3 (c) and (d), which is much reduced in the corresponding Fig. (g) and (h): While the FA from our model remains high throughout the white matter (in agreement with the results in [69]), confirming the reduced impact of fiber crossings, our radial kurtosis still shows substantial variation, which reflects more subtle aspects of tissue architecture. We believe that the similarity between Fig. 6.3 (c) and (d) is caused by the fact that, in standard DKI, FA and $K_{\perp}$ are both reduced in regions of fiber crossings, and that factoring out the effect of crossings emphasizes the information specific to diffusional kurtosis.

6.6. Conclusion

In this work, we have demonstrated how strongly measures from diffusional kurtosis are affected by fiber crossings, which limits their interpretability as indicators of tissue microstructure. To alleviate this, we have explicitly accounted for crossings by adding a cylindrically symmetric kurtosis term to the popular multi-tensor model. Results on simulated data confirm that the resulting model remains tractable, and successfully disentangles the effects of crossings and per-compartment tissue parameters.

In real data, the maps from our model differ significantly from standard diffusional

kurtosis imaging; we believe that they more specifically indicate factors such as scale and homogeneity of tissue microstructure. As a next step, we plan to use additional simulations and a systematic comparison to other MR-derived quantities to gain more insight into the exact interpretation of these maps. We also plan to use spatial regularization to achieve stable fitting on noisy data.

7. Versatile, Robust, and Efficient Tractography With Constrained Higher-Order Tensor fODFs

Abstract

Purpose *Develop a multi-fiber tractography method that produces fast and robust results based on input data from a wide range of diffusion MRI protocols, including high-angular resolution diffusion imaging (HARDI), multi-shell imaging, and clinical diffusion spectrum imaging (DSI).*

Methods *In a unified deconvolution framework for different types of diffusion MRI protocols, we represent fiber orientation distribution functions (fODFs) as higher-order tensors, which permits use of a novel positive definiteness constraint (H-psd) that makes estimation from noisy input more robust. The resulting directions are used for deterministic fiber tracking with branching.*

Results *We quantify accuracy on simulated data, as well as condition numbers and computation times on clinical data. We qualitatively investigate the benefits when processing suboptimal data, and show direct comparisons to several state-of-the-art techniques.*

Conclusion *The proposed method works faster than state-of-the-art approaches, achieves higher angular resolution on simulated data with known ground truth, and plausible results on clinical data. In addition to working with the same data as previous methods for multi-tissue deconvolution, it also supports DSI data.*

Comment *This chapter corresponds to the paper[2]: "Versatile, robust, and efficient tractography with constrained higher-order tensor fODFs" by Michael Ankele et al. in International Journal of Computer Assisted Radiology and Surgery (2017). The paper heavily uses a matrix H_T associated to tensors or polynomials. See appendix C for more information about this matrix and further analysis of its spectrum.*

7.1. Introduction

Tractography algorithms reconstruct the trajectories of major fiber bundles based on data from diffusion MRI (dMRI), and are now firmly established in neurosurgical planning [15] and studies of brain white matter [97]. It has long been known that tractography based on diffusion tensor imaging [51, 8] is unable to deal with the large number of voxels in which multiple fiber populations cross or fan out. Constrained spherical deconvolution [87] is a widely used alternative that successfully reconstructs fiber crossings from high

angular resolution diffusion imaging (HARDI) data.

For each voxel, spherical deconvolution computes a fiber orientation distribution function (fODF). While local fODF maxima are often taken as indicators of main fiber directions [87], previous work from ourselves and other groups [70, 41, 99, 68] suggests that mathematically representing fODFs as higher-order tensors and performing a low-rank approximation reproduces fiber directions with increased accuracy, especially when fibers cross at small angles.

Our present work significantly extends this higher-order tensor based approach by applying a constraint that was first introduced in our recent conference publication [3]. We name this constraint H-psd, since it requires positive semidefiniteness (psd) of a matrix H that is related to the Hankel form of the higher-order tensor. It can be imposed on fODFs to make their estimation more robust, especially when only relatively little or noisy dMRI data is available. Its benefit is analogous to the addition of a non-negativity constraint [87] to the original spherical deconvolution approach [89], which has greatly increased its practical utility. However, our H-psd constraint is stronger than non-negativity, and its exact mathematical motivation and implementation are substantially different. This is the first part of our technical contribution.

Spherical deconvolution has recently been extended to work with multi-shell data, i.e., dMRI data which has been acquired at multiple levels of diffusion weighting [40]. In regions of partial voluming between white and gray matter or corticospinal fluid, this makes it possible to isolate the part of the dMRI signal corresponding to white matter, which makes tractography in those regions more robust. As a second contribution, we present a unified framework that allows us to apply this idea also to diffusion spectrum imaging (DSI). The traditional algorithms used to perform tractography based on DSI data [98, 12] are different from those commonly used on HARDI and multi-shell data, and our experimental results provide a direct comparison.

7.2. Methods

Our proposed method supports single-tissue, as well as multi-tissue deconvolution. Section 7.2.1 describes our unified framework that permits multi-tissue deconvolution even in cases where measurements include multiple levels of diffusion weighting, but are not organized on shells. Our H-psd constraint, which makes deconvolution numerically more robust, is explained in Section 7.2.2, with mathematical details given in an appendix. Finally, Section 7.2.3 describes the fiber tracking algorithm used in our experiments.

7.2.1. A Unified Deconvolution Framework

Single tissue deconvolution is based on the assumptions that, up to rotations that account for differences in orientation, all fibers within a voxel give rise to the same dMRI signal, and that signals from differently oriented fibers add up linearly. Under these conditions, in voxels that contain only white matter, the measured dMRI signal can be expressed as a convolution integral on the sphere. In particular, a fiber orientation distribution function (fODF) that captures the fraction of fibers in each direction is convolved with a

kernel that reflects the common dMRI response from a single, coherently oriented fiber compartment [89]. fODFs are antipodally symmetric and do not indicate the direction of signal transmission along the bundle; in accordance with the literature, we use the terms "direction" and "orientation" interchangeably.

Multi-tissue deconvolution involves multiple tissue response functions [40, 17]. It has mostly been used to separate out signal contributions from gray matter (GM) and corticospinal fluid (CSF), which have been modeled as isotropic, and therefore only add respective volume fraction parameters, rather than additional fODFs. These tissue types are distinguished based on differences in how their dMRI signal is attenuated at different levels of diffusion weighting, which are commonly quantified as b values. Previous approaches [40, 17] employ separate response functions for each tissue and each distinct b value. This is well-suited for multi-shell data, in which the full space of orientations is sampled for each of a small number of different b values.

When using Diffusion Spectrum Imaging [96], or some recently proposed dMRI protocols that distribute samples freely in q-space [46, 57], many different b values are available, but with few or even only a single orientation each. In this case, modeling independent spherical functions for each b value results in an unreasonably large number of model parameters. An immediate consequence is that the previously proposed method for response function estimation [40] cannot be used on such data.

Therefore, our unified framework for multi-tissue deconvolution instead builds on a continuous model of functions $F(\mathbf{q} = q\mathbf{u})$ in q-space, using the SHORE basis functions [50]

$$\phi_{lnm}(\mathbf{q}) = \left[\frac{2(n-l)!}{\zeta^{3/2}\Gamma(n+3/2)}\right]^{1/2} \left(\frac{q^2}{\zeta}\right)^{l/2} \exp\left(\frac{-q^2}{2\zeta}\right) L_{n-l}^{l+1/2}\left(\frac{q^2}{\zeta}\right) Y_l^m(\mathbf{u}) \tag{7.1}$$

with the associated Laguerre polynomials L_n^α, the real spherical harmonics Y_l^m, and a radial scaling factor $\zeta = 700$. We use maximum radial and angular orders four, which leads to a fixed number of 6 parameters to describe a cylindrically symmetric white matter response function, while per-shell modeling as in [40], even when reducing the angular order to four, would still require $1 + 3 \times B$ parameters, where B is the number of unique non-zero b values, e.g., $B = 16$ in the DSI data used in our experiments.

Let $K(\mathbf{q}) = \sum_{ln} K_{ln}\, \phi_{ln0}(\mathbf{q})$ be the white matter single-fiber response, with $m = 0$ due to cylindrical symmetry. The signal from an fODF f is then modeled by a convolution on the sphere, $S(\mathbf{q}) \approx K \star_{\mathbb{S}^2} f$ [16]. For a given K and signal vector $S_i = S(\mathbf{q}_i)$, finding the spherical harmonics coefficients f via deconvolution becomes a linear least squares problem:

$$\mathsf{argmin}_f \, \|Mf - S\|^2 \tag{7.2}$$

with convolution matrix

$$M_{(i)(lm)} = \sum_n \frac{1}{\alpha_l} K_{ln}\, \phi_{lnm}(\mathbf{q}_i). \tag{7.3}$$

Tournier et al. [87] constrain the optimization in Equation (7.2) by requiring f to be non-negative and define the α_l in Equation (7.3) from the truncated spherical harmonics

transform of the unit delta function. Our method improves numerical stability by making these two choices differently, as will be explained in Section 7.2.2.

Multi-tissue support is added by concatenating individual tissue matrices

$$M = [M_{\mathsf{CSF}}, M_{\mathsf{GM}}, M_{\mathsf{WM}}], \quad f = \begin{bmatrix} f_{\mathsf{CSF}} \\ f_{\mathsf{GM}} \\ f_{\mathsf{WM}} \end{bmatrix}. \tag{7.4}$$

Since CSF and GM are isotropic, M_{CSF} and M_{GM} are single-column matrices. Given single-shell data, multiple tissues cannot be estimated. In this case, we simply replace the SHORE with the spherical harmonics basis, which leads to omission of the radial parameter n from Equation (7.3).

Tissue response functions are estimated in close analogy to a previous method for multi-tissue deconvolution [40]: Initial masks for WM, GM, and CSF are created from an intensity-based tissue segmentation of a coregistered T_1 image [104], thresholded at 95% for each tissue type. These masks are refined based on Fractional Anisotropy (FA) from a diffusion tensor fit, by restricting them to $\mathsf{FA} > 0.7$ for core white matter, and $\mathsf{FA} < 0.2$ for GM and CSF. SHORE coefficients of the three response functions are obtained by fitting and averaging within the respective masks. Due to the stringent FA threshold, the core white matter mask is thought to contain voxels with a single dominant fiber in each. They were aligned by rotating the principal eigenvector of the diffusion tensor to the z axis, which was achieved by rotating the B matrix before the SHORE fit.

7.2.2. Constrained Higher-Order Tensor fODFs

In a previous work [70], we proposed to describe fODFs f by fully symmetric fourth order tensors:

$$f(\mathbf{v}) = T(\mathbf{v}) = \sum_{i,j,k,l=1}^{3} T_{ijkl}\, v_i v_j v_k v_l, \quad \mathbf{v} \in \mathbb{S}^2 \tag{7.5}$$

Such fODF tensors T are obtained by deconvolution in the Spherical Harmonics basis, as in Equation (7.2), and subsequently changing to the monomial basis, which can be done using a nonsingular linear transformation [53]. In the higher-order tensor framework, it is natural to represent a single fiber contribution with volume fraction λ in unit direction $\mathbf{u}$ as a rank-one tensor $\lambda \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u} \otimes \mathbf{u}$, which can be achieved by setting the α_l in Equation (7.3) to the spherical harmonics coefficients of a unit rank-1 tensor. In the resulting representation, k principal fiber directions for deterministic tractography should be extracted from an fODF by performing a symmetric rank-k approximation, rather than using the k dominant peaks of the fODF function [70, 41].

A key technical contribution of our current work is to formulate a constraint which we call H-psd. It is the suitable higher-order tensor counterpart of the nonnegativity constraint on f that is commonly imposed in Equation (7.2) [87]. It can be enforced in an exact, simple, and efficient manner using standard optimization packages, and it is theoretically well-founded in the sense that it can be shown to be equivalent to the condition

that a valid fODF should represent a mixture of non-negative fiber compartments. It is more rigorous than non-negativity in the sense that any H-psd fODF is non-negative, but the reverse may not be true: Intuitively, some non-negative fODFs are so sharp that they cannot possibly arise from a non-negative mixture of fiber compartments, which have a minimum width in the higher-order tensor framework.

Since the exact mathematical derivation of the H-psd constraint will only be relevant to part of the intended audience of our work, we present it in an appendix. The main result, which also explains the name H-psd, is the fact that the matrix H in Equation (7.18), which is composed of the coefficients of T, has to be positive semidefinite (psd).

In practice, this constraint can be imposed on Equation (7.2) using the quadratic cone program (QCP)

$$\mathsf{argmin}_f \, \frac{1}{2}\langle f, Pf\rangle + \langle q, f\rangle \quad \mathsf{subject\,to} \quad (Gf)\,\mathsf{psd} \tag{7.6}$$

with f, M and S as in Equation (7.2), $P = M^T M$, $q = -M^T S$, and a matrix G that first maps f from spherical harmonics to the monomial basis, and then to its H matrix from Equation (7.18), $Gf = H_f$. In multi-tissue deconvolution, additional non-negativity constraints are enforced for f_{GM} and f_{CSF}.

The QCP is solved using the routine CONEQP in the publicly available software package CVXOPT, which requires the above-mentioned vectors and matrices as its only input. Details on its implementation are given in [93], and are beyond the scope of our paper.

7.2.3. Deterministic Tractography with Branching

Similar to previous work [98, 70, 20, 88], deterministic multi-fiber tractography has been implemented with Euler integration. In each of the local fiber directions at a given seed point, a streamline is extended bi-directionally with an integration stepsize that was set to $0.5\,\mathsf{mm}$ in our experiments. In each integration step, a local fODF is interpolated trilinearly, and a set of fiber directions is extracted from it. In case of multiple options, the fiber direction that leads to lowest tract curvature is selected.

Streamline integration stops when no fiber is found within a given turning angle (exact values specified below), after a maximum number of integration steps (400), or when the white matter volume fraction was below 0.5. This volume fraction is estimated by the multi-tissue deconvolution itself or, in case of single-shell data, using a coregistered T_1 tissue segmentation.

In order to capture the effect of fiber dispersion, we implement branching. In particular, when multiple directions are permissible, the tract branches out, i.e., a new streamline is seeded in the direction leading to the second lowest curvature. In order to avoid excessive branching, streamlines that result from branching are not allowed to branch again, and the original ones can only branch again after some number of integration steps ($4\,\mathsf{mm}$).

In our framework, main fiber directions are extracted using low-rank approximation, which requires deciding on a suitable number k of local fiber compartments. As discussed more formally in the appendix, k amounts to the numerical rank of the fODF tensor, which coincides with the rank of the same matrix H that underlies the H-psd constraint.

Therefore, we estimate k as the number of eigenvalues of H above some threshold θ. Since we are not aware of any regions where more than three fiber tracts cross, we impose a maximum of $k \leq 3$. We found that the exact values of θ that lead to plausible results depended on the acquisition scheme and regularization. We set it by visually inspecting the spatial maps of the number of fiber compartments resulting from different choices, and report values for the individual experiments below. Rank-one terms with scaling factor $\lambda_i < 0.15$ were discarded.

For comparison, we also created results using methods that take local fODF maxima as estimates of main fiber directions. In this case, we first roughly localize local maxima on a discretization of the sphere (321 unique directions), and subsequently refine their exact positions using gradient ascent. Peaks with magnitude below 0.1 were discarded.

7.3. Results

We report results of our method on two clinical dMRI datasets, and on data that we synthesized from one of them, as described in Section 7.3.1. Motion correction was performed in a pre-process using the tools available in FSL [1].

The first dataset, *clin-2-sh,* is a two-shell dataset from a healthy volunteer, measured on a 3T Skyra (Siemens, Germany) with $96 \times 96 \times 50$ voxels of $2\,\mathrm{mm}$ isotropic size, $\mathrm{TE/TR} = 89/9100\,\mathrm{ms}$. Three images were taken at $b = 0$, 30 DWIs at $b = 700$, and 64 at $b = 2000$ (all b values in $\mathrm{s/mm^2}$). Some experiments only use the data from the inner or outer shell (*clin-700* and *clin-2000,* respectively). We note that traditional single-shell deconvolution does not make use of the $b = 0$ images.

The second dataset, *clin-dsi,* was kindly provided by Katrin Sakreida and Georg Neuloh (RWTH Aachen University Hospital). It was acquired on a 3T Prisma (Siemens, Germany) with $136 \times 136 \times 84$ voxels of $1.5\,\mathrm{mm}$ isotropic size, $\mathrm{TE/TR} = 69/11600\,\mathrm{ms}$. One $b = 0$ image and 128 DWIs with b values up to $b = 3000$ were taken on a Cartesian grid.

7.3.1. Simulation Experiment

In order to quantify the accuracy of our method, we generate dMRI data for which volume fractions and orientations of crossing fiber compartments are known. Crossings are simulated based on resampling the clinical data *clin-2-sh.* This process avoids potentially oversimplifying assumptions in mathematical models of water diffusion and noise.

Data was generated based on the same voxels used to estimate response functions for the three tissue types, as explained in Section 7.2.1. The respective white matter voxels are thought to contain a single fiber compartment, whose orientation was estimated via the diffusion tensor model.

For each simulated fiber crossing, we averaged the signals from two randomly chosen single fiber voxels, after applying a random rotation to them, and one voxel from either the gray matter or CSF mask. The respective volume fractions were selected randomly, and normalized to sum to one. In the presence of noise, we cannot expect any algorithm to reconstruct the orientation of fibers with a very small volume fraction. Therefore,

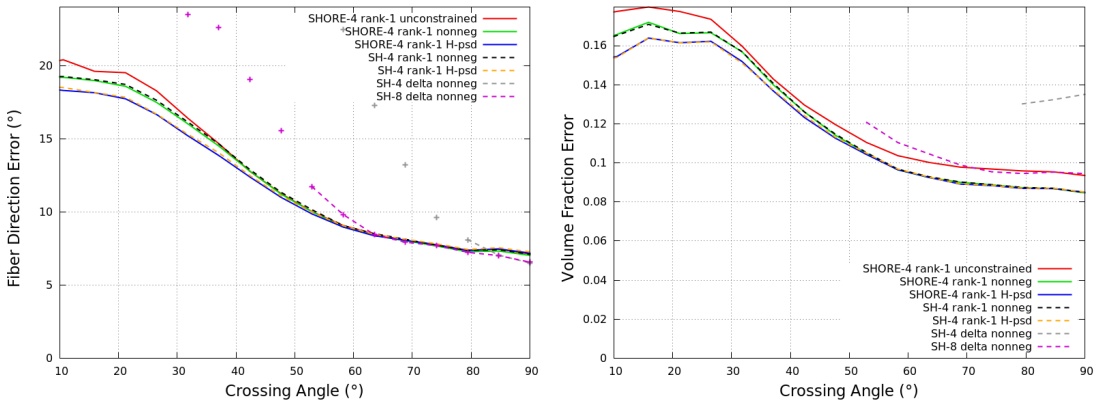


Figure 7.1.: Average errors of fiber directions and individual fiber volume fractions reconstructed from two-fiber crossings that have been simulated from clinical two-shell data. The proposed method is SHORE-4 rank-1 H-psd, the current state-of-the-art is SH-8 delta nonneg. Results from several hybrid methods are shown to assess the impact of individual factors.

we discarded samples in which the weaker fiber contributed less than 20% to the white matter, or in which both fibers combined accounted for less than 50% of the voxel.

We used different deconvolution approaches to reconstruct the fiber directions and volume fractions from the resulting data, and compared the results to the known values, which were recorded during the simulation. Figure 7.1 plots the average angular error in the reconstructed fiber directions (top) and the average absolute errors in their estimated volume fractions (bottom), each as a function of the ground truth crossing angle between the two fibers.

The scaling parameters λ_i of the low-rank approximation were directly taken as estimates of the corresponding fiber volume fraction, as motivated in Section 7.2.2. When extracting fiber directions from ODF peaks (in "SH-4 delta nonneg" and "SH-8 delta nonneg"), volume fraction estimates were obtained by dividing the magnitude of the corresponding peak by the peak magnitude observed when applying constrained deconvolution of the same order to the single fiber response.

The main comparison is between our proposed method ("SHORE-4 rank-1 H-psd") and the previous state-of-the-art [40] ("SH-8 delta nonneg"). We note that, for angles smaller than approximately 55°, the latter method no longer reliably resolved the crossing, with two significant fODF peaks being detected in less than 50% of the cases. We continue to show fiber direction errors based on the remaining few cases as isolated markers.

In order to investigate the effects of individual factors that make up the difference between our method and the previous state-of-the-art, we also performed experiments with several hybrid methods, including one that combines lower order with peak finding ("SH-4 delta nonneg"), ones that use traditional per-shell signal modeling instead of the SHORE basis ("SH" variants of "SHORE"), or replace our H-psd with a traditional

	clin-700	clin-2000	clin-2-sh	clin-dsi
Proposed Method	1m49s	1m26s	2m10s	4m45s
Order-8 CSD [87]	3m08s	3m19s	N/A	N/A
MSMT Deconvolution [40]	N/A	N/A	4m10s	N/A

Table 7.1.: Computation times for deconvolution of whole-brain data using different approaches, all implemented in the same optimization framework. This does not include tractography, whose computational cost is also reduced by using low-rank approximation.

nonnegativity constraint in 300 discrete directions [87].

7.3.2. Single-Shell Deconvolution

Similar to Tournier et al. [87], who demonstrate that their non-negativity constraint allows them to produce plausible results even with relatively low b value and few gradient directions, we test our H-psd constraint by applying it to *clin-700* (30 DWIs at $b = 700$). Figure 7.2 visualizes the resulting fODFs and a deterministic streamline tractography in the brainstem region.

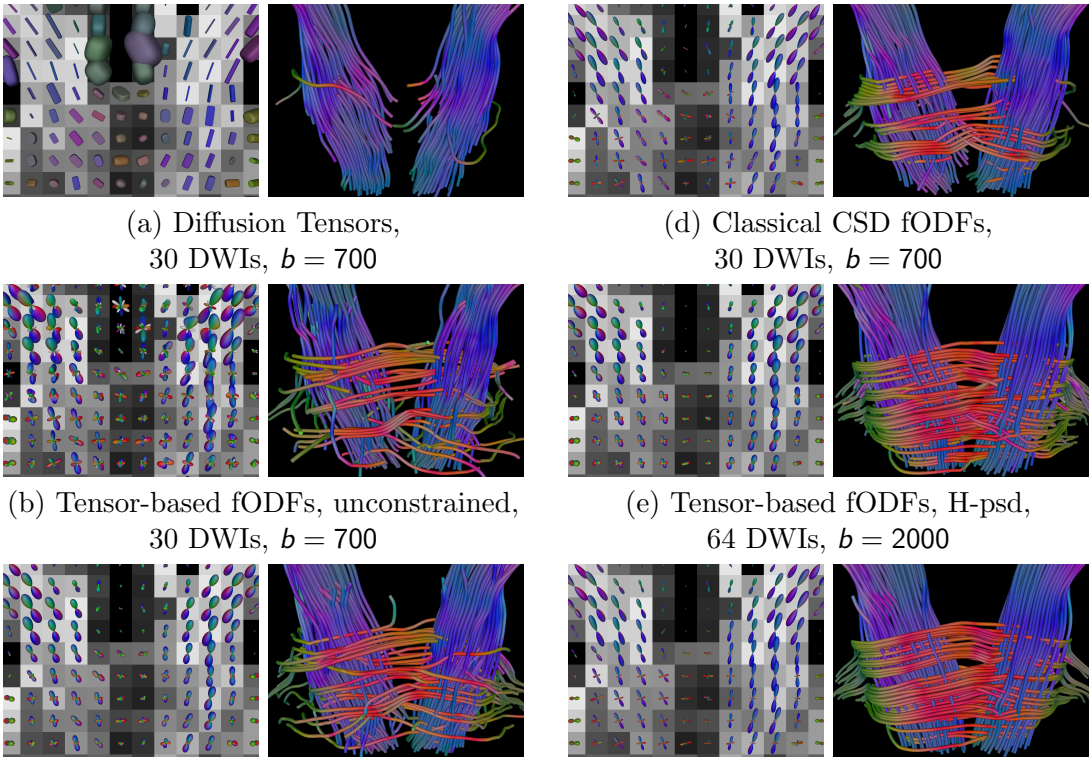
Unlike the standard diffusion tensor model, which is shown for comparison, combining spherical deconvolution with low-rank approximation allows us to reconstruct pontine crossing tracts even with this limited b value and number of gradient directions (maximum turning angle $40°$; $\theta = 0.35$ for unconstrained, $\theta = 0.1$ for H-psd). Negative values of the fODFs are shown in white in Figure 7.2 (b). They are removed, and fODFs in adjacent voxels become more similar, when enforcing the H-psd constraint.

For comparison, Subfigure (e) shows results based on *clin-2000* (64 DWIs at $b = 2000$; $\theta = 0.1$), which is commonly considered a more suitable input for spherical deconvolution [87]. Subfigures (d) and (f) provide a direct comparison with classical constrained spherical deconvolution (CSD).

Table 7.1 compares the times for whole-brain deconvolution, measured on a workstation with a six-core CPU at 3.4 GHz, and using the same optimization package (CVXOPT) for all methods. Our method takes only about half as long as CSD, mostly due to our use of a lower model order. In Figure 7.2, this is reflected in the increased smoothness of our fODFs. Despite this, applying low-rank approximation allows us to obtain very similar tractography results.

In addition to the reduced computational effort for whole-brain deconvolution, tractography is also less costly when using order-4 low-rank approximation, compared to order-8 peak finding: In our implementation, one integration step with the former took 16 ms on average, while one step with the latter took around 125 ms.

While the H-psd constraint improves the results of fourth order tensor based deconvolution, Figure 7.2 (b) shows that we achieve useful results even without it, even on the *clin-700* data. This is in contrast to standard order-8 deconvolution, whose results



(a) Diffusion Tensors, 30 DWIs, $b = 700$

(d) Classical CSD fODFs, 30 DWIs, $b = 700$

(b) Tensor-based fODFs, unconstrained, 30 DWIs, $b = 700$

(e) Tensor-based fODFs, H-psd, 64 DWIs, $b = 2000$

(c) Tensor-based fODFs, H-psd, 30 DWIs, $b = 700$

(f) Classical CSD fODFs, 64 DWIs, $b = 2000$

Figure 7.2.: Unlike diffusion tensors (a), fODFs allow us to resolve the pontine crossing tracts. In (c), our H-psd constraint reduces the impact of noise and removes negative fODF lobes, shown in white in the left part of (b). Even though it is possible to reconstruct fiber crossings even from few gradient directions at low b values, cleaner results are obtained when HARDI data is available, as shown in (e) and (f). We achieve similar tractography as classical CSD, at a reduced computational cost.

	clin-700	clin-2000	clin-2-sh	clin-dsi
Proposed Method	189	7.86	458	536
Order-8 CSD [87]	∞	3.22×10^6	N/A	N/A
MSMT Deconvolution [40]	N/A	N/A	1.02×10^7	N/A
Order-4 CSD / MSMT	1.18×10^4	478	1.47×10^3	N/A

Table 7.2.: The condition numbers of the matrices defining the quadratic optimization problem in our method are much lower than in competing ones. This indicates its intrinsic numerical stability, and explains why constraints are less crucial than in traditional deconvolution.

without enforcing non-negativity were dominated by noise so severely, even in the *clin-2000* data, that it did not make sense to present them. This is explained by inspecting the condition number of the matrix P that defines the respective quadratic optimization, listed in Table 7.2. Estimating order-8 fODFs from *clin-700* is an underdetermined problem, represented symbolically in Table 7.2 by an infinite condition number. As we saw in Figure 7.1, order-4 variants of standard deconvolution do not provide a useful angular resolution. We still include them in the final row of Table 7.2 to illustrate that the improved conditioning of our method is only partially explained by the reduced model order. Using a rank-1 representation of single fibers, as it was introduced in Section 7.2.2, also improves conditioning considerably.

In theory, our H-psd constraint forces fODFs to be nonnegative everywhere, while standard CSD [87] only enforces nonnegativity in 300 discrete directions. We quantified to which extent our numerical implementations of both constraints satisfy nonnegativity in practice. This was done by evaluating ODFs along a dense set of 1281 unique directions, and taking the minimum of all results. Computed over the whole brain estimates in *clin-700,* a minimum of -1.38×10^{-7} indicated a negligible numerical inaccuracy in enforcing the H-psd constraint. Even though we used the capabilities of CVXOPT to enforce nonnegativity in 300 directions as hard constraints, as opposed to implementing soft constraints as proposed in [87], the minimum ODF value observed for classical CSD was much more substantial, -0.164.

7.3.3. Multi-Shell Deconvolution

Using all data from *clin-2-sh,* we compared results from state-of-the-art multi-shell multi-tissue deconvolution [40], which involves order-8 fODFs based on the truncated delta peak and a non-negativity constraint, to our proposed method, which uses a fourth-order tensor representation of fODFs with our H-psd constraint.

We found the tissue volume fraction maps from both methods, shown in Figure 7.3, to be quite similar. Averaged over a brain mask, the mean absolute difference was 0.005 in CSF, 0.026 in gray matter, and 0.025 in white matter. As in Figure 7.2, our fODFs are less sharp compared to the ones from the existing approach [40]. The bottom row of

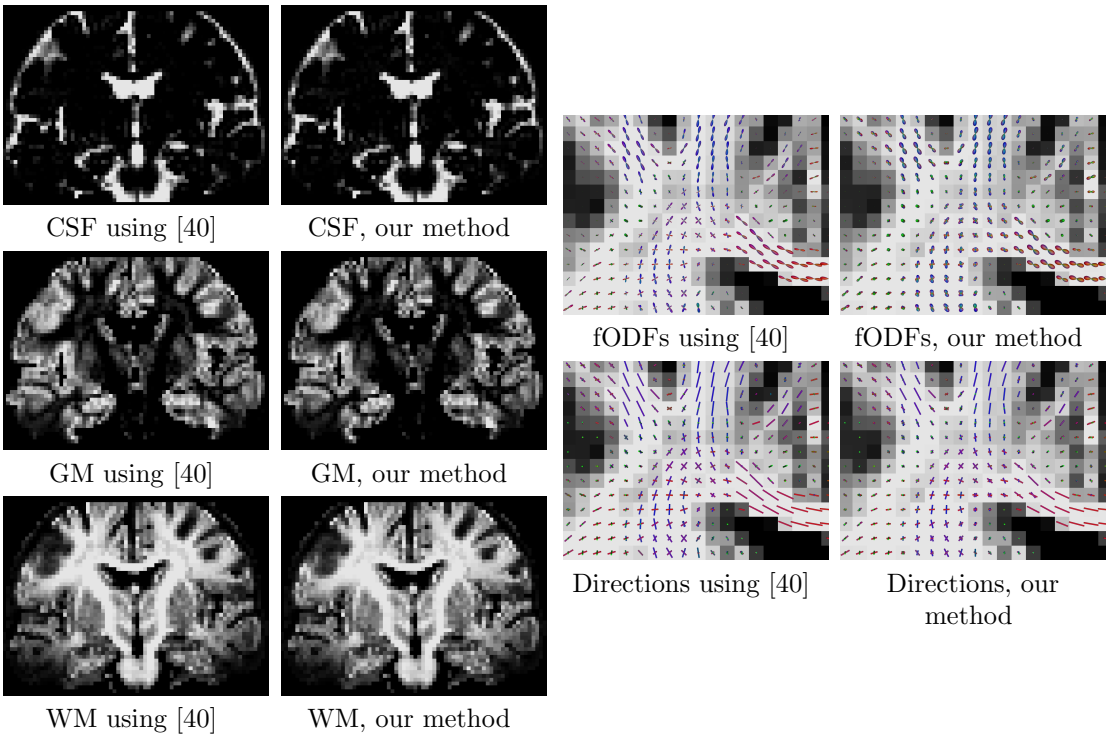


Figure 7.3.: On two-shell clinical data, very similar tissue volume fraction maps are obtained using a state-of-the-art approach (left column) and ours (right column). Our fODFs (fourth row) are smoother, which accounts for the improved speed and conditioning of our method. The final row (right column) shows that low-rank approximation allows us to reliably resolve crossing fibers directions despite the smoothness of our fODFs.

Figure 7.3 shows that low-rank approximation allows us to resolve two- and three-fiber crossings despite the smoothness of our fODFs. Averaged over the white matter, the angular deviation, weighted by volume fractions, was $8.24°$ when comparing fiber directions from order-8 fODF peaks to the results of fourth-order low-rank approximation.

Similar to previous work on crossing fiber tractography [70, 49], we compared the ability of different fiber tracking algorithms to reliably reconstruct transcallosal fibers from seeds in the corpus callosum near the mid-sagittal plane. The left column of Figure 7.4 shows results from our own implementation of deterministic tractography with branching (maximum turning angle $45°$, $\theta = 0.02$). It finds many lateral projections when using fiber estimates from fourth-order low-rank approximation (top), but only very few based on the peaks in traditional order-8 fODFs (bottom).

For comparison, we re-processed the same data with the software package MRtrix3 [88], which includes a reference implementation of the original multi-shell multi-tissue deconvolution approach [40]. In order to make the results of tracking methods with and without branching more comparable, we varied the number of tracts per seed voxel to achieve a similar final number of displayed fibers (around 4 400) in all cases. We used the recommended default step sizes with MRtrix, which depend on the algorithm.

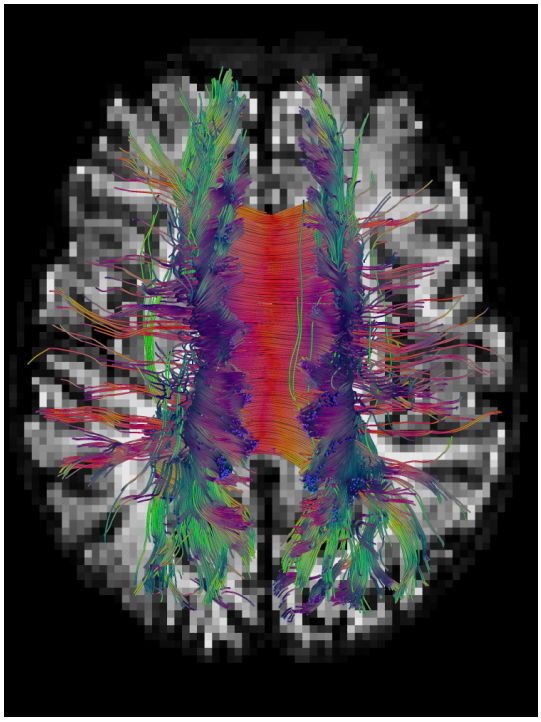
Since the deterministic tracking in MRtrix does not implement branching, it finds even fewer lateral projections (bottom right). However, it includes a probabilistic tractography approach [86] that is based on sampling the fODFs rather than finding their maxima, and extracts lateral fibers similar to the ones found by our deterministic technique (top row), albeit with the increased wiggling that is characteristic of probabilistic tracking.

7.3.4. Clinical DSI

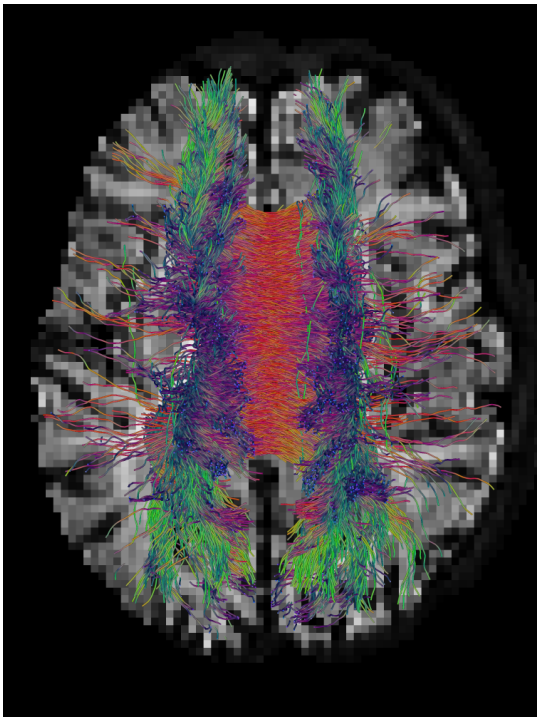
After applying our variant of multi-tissue deconvolution to clinical DSI data *(clin-dsi),* we evaluated the extent to which the resulting fODFs allow us to reconstruct the corticospinal tract from seeds in the internal capsule. A similar task was recently used by others to evaluate crossing-fiber tractography [15]. The results are shown in the top row of Figure 7.5, and are compared to results from standard DSI reconstruction [98], as implemented in the open source software package dipy [26].

In contrast to the fiber ODFs (fODFs) from our method, traditional DSI reconstruction results in diffusion ODFs (dODFs) that reflect the fraction of diffusion that happens in each direction, as opposed to the fraction of fiber bundles. As can be seen in the center row of Figure 7.5, dODFs are generally less sharp then fODFs. Canalez-Rodríguez et al. [12] have used deconvolution to compute sharper ODFs from DSI data. The results from their method, again using the publicly available implementation from dipy, is shown in the right-hand column of Figure 7.5. We note that in contrast to our method and related ones [40], their deconvolution approach avoids assumptions on the response of single fiber compartments, and therefore does not produce fODFs.

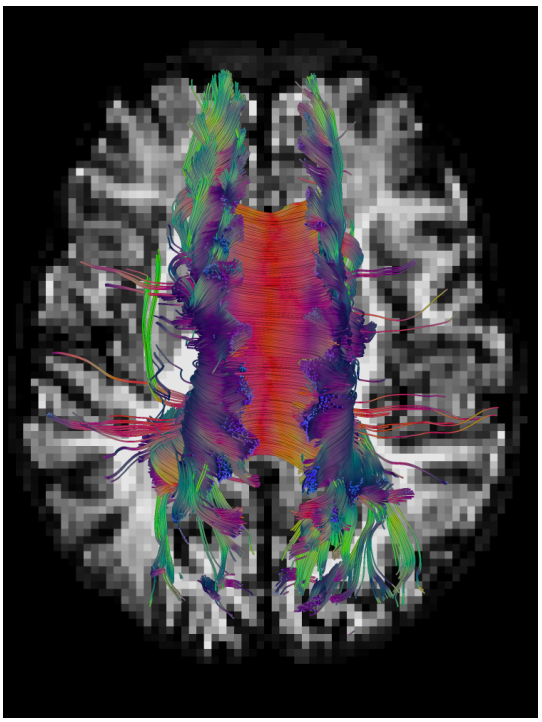
All tractography results in Figure 7.5 have been obtained with the algorithm from Section 7.2.3 (maximum turning angle $50°$). It is using fiber estimates from low-rank approximation ($\theta = 0.02$) in case of our fODFs, and local maxima in case of dODFs, as



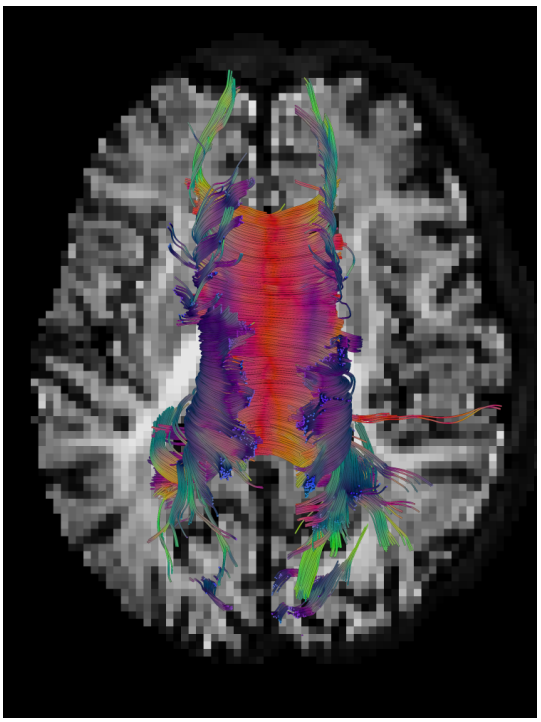
Deterministic tracking with low-rank approximation (order 4)



Probabilistic tracking (order 8)



Deterministic tracking with peak finding and branching (order 8)



Deterministic tracking with peak finding, reference implementation (order 8)

Figure 7.4.: Our deterministic tractography based on low-rank approximation successfully reconstructs transcallosal fibers from seeds near the mid-sagittal plane. Using previously available tools, we only managed to reconstruct them with probabilistic tracking. As shown on the bottom row, deterministic tracking based on peak finding reconstructs much fewer of them, both with our own implementation (left) and a publicly available one (right).

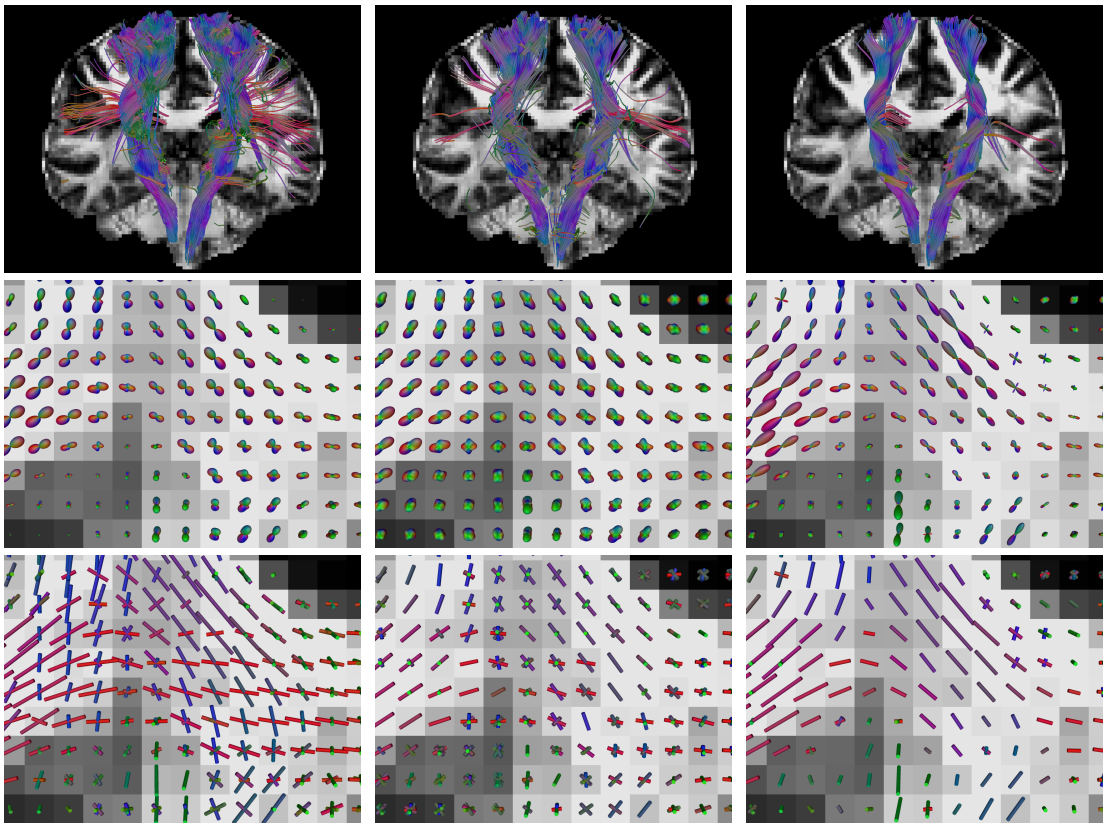


Figure 7.5.: Our method can also be applied to clinical DSI data, and reconstructs more lateral projections than standard DSI tools when tracking the corticospinal tract from seeds in the internal capsule. The reason for this becomes apparent when comparing fODFs from our method to dODFs from DSI (middle row), and the directions resulting from low-rank approximation or peak finding, respectively (last row).

is customary in standard DSI tractography [98]. Corresponding directions are visualized in the final row of Figure 7.5. In all cases, streamlines that cross the mid-sagittal plane have been removed in a post-process.

7.4. Discussion

Our proposed method differs from the previous state-of-the-art in several aspects, which we will discuss separately.

7.4.1. Unified Deconvolution Framework

The main benefit of our unified deconvolution framework is the fact that it allows us to perform multi-tissue deconvolution also on dMRI data that include multiple b values, but is not organized on shells. The comparison in Figure 7.5 confirms that this extension of multi-tissue deconvolution produces plausible results on clinical DSI data, and compares favorably to standard DSI processing in its ability to reconstruct also the lateral projections of the cortico-spinal tract from seeds in the internal capsule.

In Figure 7.5, a previously proposed method for DSI deconvolution led to an implausibly low number of secondary fiber directions in a known fiber crossing region, and decreased our ability to reconstruct branching fibers. This might indicate that the characteristics of the *clin-dsi* data, which include fewer measurements and lower maximum b value than the data used by Canalez-Rodríguez et al. [12], might be less suitable for that approach. Due to the long computational times of its DIPY implementation, we only tried it with default parameters.

The simulation results in Figure 7.1 indicate that, in case the input data is organized on shells, it makes very little difference whether we model the response functions per-shell, as in [40, 17], or with our unified framework. Despite the fact that the angular part of the SHORE basis in Equation (7.1) consists of spherical harmonics, i.e., the same basis that is otherwise used for per-shell modeling, this result is not trivial: Our approach reduces the number of parameters needed to represent the response function by assuming that radial and angular parts factorize, which is not enforced by the more flexible per-shell modeling. The fact that both methods produced almost identical results suggests that this assumption is not violated to an extent that it would impair the deconvolution.

7.4.2. H-psd Constraint

A theoretical benefit of our H-psd constraint is the fact that it is easy to enforce exactly, while non-negativity is either approximated by enforcing it on a discrete set of points on the sphere [87], or requires a costly Riemannian gradient descent [16]. It was confirmed visually and quantitatively in Section 7.3.2 that enforcing H-psd removes nonnegative fODF lobes, and makes the fODFs and resulting tractography more regular.

Even though this constraint is beneficial, Figure 7.1 suggests that it results in an increased accuracy mostly at small crossing angles. Clearly, constraining deconvolution based on fourth-order tensors is not as crucial as enforcing non-negativity in order-8

CSD. This can be understood from the greatly reduced condition numbers, which were presented in Table 7.2, and which are made possible not only by the reduction in model order, but also by representing single fibers as rank-one tensors rather than truncated delta peaks.

A direction that we would like to follow in our future work is the use of our H-psd constraint to prevent overfitting of misspecified deconvolution models, e.g., ones that assume a response function that has been derived from healthy tissue, and may not match the true response in regions that suffer from demyelination [69].

7.4.3. Model Order and Low-Rank Approximation

An obvious factor that contributes to the improved conditioning and speed of our method is the fact that we use a lower order representation of fODFs, i.e., order-4 with just 15 degrees of freedom, compared to the more common order-8 models with 45 degrees of freedom. It is well-known that lower orders lead to better numerical behavior, while higher orders are commonly selected to increase angular resolution, defined as the minimum angle at which two crossing fiber populations can still be reliably distinguished [87].

At this point, the use of low-rank approximations to extract fiber directions from fODFs is the key ingredient of our method. It can be seen from Figure 7.1 that low-rank approximation of a fourth-order fODF leads to an even higher angular resolution than extracting peaks of an order-8 fODF, and that reducing the order in a peak finding approach would lead to unacceptably poor resolution. Even though we demonstrated this benefit of low-rank approximations before [70], we previously compared to filtered spherical deconvolution [89], while our current work shows that the advantage is maintained compared to state-of-the-art constrained deconvolution [87].

It was shown in Fig. 7.1 that fiber volume fraction estimates from low-rank approximation are more accurate than ones based on ODF peak magnitudes. The reason for this is the same as for the increased angular resolution: Unlike simple maximum finding, low-rank approximation amounts to an optimization that correctly accounts for the interference of ODF peaks from mixing fibers [70]. We note that, in any case, volume fraction estimates derived from deconvolution will be affected by potential mismatches between the deconvolution kernel and the actual tissue characteristics [62].

As shown in [70], it is possible to generalize the tensor-based approach to orders higher than four. According to Equation (7.26) in the appendix, the corresponding H-psd constraints are stricter than requiring non-negativity of the fODF, but no longer ensure that the fODFs will be a mixture of non-negative fiber contributions. Preliminary experiments, whose results are not shown, indicate that the additional benefits of taking the tensor-based approach to higher orders appear to be marginal, while the increase in computational effort is significant. Future work might investigate this in more detail.

7.4.4. Multi-Fiber Tractography

Several approaches to multi-fiber tractography are based on the idea that estimating multiple fiber directions in each voxel independently is not reliable enough, and should be regularized by using information from spatial neighborhoods. Examples include unscented Kalman filters [49, 15] or global optimization frameworks [64]. Other authors have argued that a larger number of plausible tracts can be reconstructed by moving from deterministic to probabilistic tractography [88].

Our results have been achieved by improving the per-voxel estimates used within a deterministic streamline-based tractography algorithm whose benefits include its simplicity and speed. In Figure 7.4, the improved angular resolution of our technique allowed us to reconstruct deterministic tracts that otherwise would have required probabilistic tracking. A comprehensive comparison of deterministic branching and probabilistic tracking as strategies for capturing the full extent of dispersing bundles will require a separate study. Similarly, it is an obvious question for future work to which extent results can be further improved by integrating our refined per-voxel estimates into more complex tractography algorithms.

Of course, many of the well-known limitations that are shared by all dMRI-based fiber tractography approaches also apply to our proposed method, and results should be interpreted with due care [43]. This includes the possibility of false positives and false negatives, i.e., missing or spurious fibers. While theoretically justified constraints such as our H-psd can contribute to reducing them, and to generating tractography that is more helpful for neurosurgery and scientific investigation, it is unclear whether this fundamental issue can be fully resolved, given the gap in spatial scales between individual axons and MR image resolution.

7.5. Conclusion

We have introduced a multi-fiber tractography method that is based on a unified deconvolution framework for HARDI, multi-shell, and diffusion spectrum imaging data, and makes use of fourth-order tensor based deconvolution with a novel and theoretically justified H-psd constraint. Results indicate that our method is faster and numerically more robust than previous alternatives, achieves higher angular resolution on simulated data, and compares favorably to existing techniques in qualitative comparisons on clinical data. Despite the promising results, more extensive validation should be performed before our method is used in clinical practice.

Appendix: Mathematical Derivation of H-psd Constraint

This appendix presents the formal definition of our H-psd constraint (Definition 4), based on a matrix representation $\boldsymbol{H}$ of higher-order tensor fODFs (Definition 3). According to Corollary 1, $\boldsymbol{H}$ is positive semidefinite if and only if the corresponding fODF can be decomposed into a non-negative weighted sum of rank-1 terms, which correspond to

single fiber compartments in our framework. This provides the theoretical justification of our H-psd constraint, whose practical effectiveness is demonstrated in the main part of the paper.

We represent fODFs as forms or symmetric tensors p of even degree $d = 4$ in $n = 3$ dimensions. In tensor notation:

$$p(\mathbf{x}) = p(\underbrace{\mathbf{x}, \dots, \mathbf{x}}_{d}) = \sum_{i \in \{1,\dots,n\}^d} p_{i_1,\dots,i_d}\, x_{i_1} \dots x_{i_d}, \quad \mathbf{x} \in \mathbb{R}^n . \tag{7.7}$$

We will call the set of these forms $F_{n,d}$. Symmetry allows us to use a different, non-redundant indexing scheme

$$p(\mathbf{x}) = \sum_{i \in I(n,d)} \binom{d}{i} p_i\, x^i \tag{7.8}$$

with multi-indices $i \in I(n,d) = \{i \in {\mathbb{N}_0}^n | \sum_k i_k = d\}$, multinomial coefficients $\binom{d}{i} = \frac{d!}{\prod i_k!}$ and monomial terms $x^i = \prod_{k=1}^n (x_k)^{i_k}$.

Our constraint stems from the relations of three subsets of $F_{n,d}$:

$$P_{n,d} = \{p \in F_{n,d} : p(\mathbf{x}) \geq 0, \forall\, \mathbf{x} \in \mathbb{R}^n\} \qquad \text{“positive semidefinite”} \tag{7.9}$$

$$\Sigma_{n,d} = \{p \in F_{n,d} : p(\mathbf{x}) = \sum_k h_k(\mathbf{x})^2\} \qquad \text{“sums of squares”} \tag{7.10}$$

$$Q_{n,d} = \{p \in F_{n,d} : p(\mathbf{x}) = \sum_k \langle \mathbf{a}_k, \mathbf{x}\rangle^d\} \qquad \text{“sums of d-th powers”} \tag{7.11}$$

Here, the $h_k(\mathbf{x})$ denote forms of degree $d/2$, $\mathbf{a}_k$ denote the individual vectors that define the rank-1 terms of a non-negative decomposition.

As shown in [65], these three subsets obey

$$Q_{n,d} \subset \Sigma_{n,d} \subset P_{n,d} . \tag{7.12}$$

These sets cannot be vector spaces, since if $p \neq 0$ is positive, $-p$ is not. So the concept of vector spaces has to be weakened:

Definition 7.1. *A **convex cone** is a subset C of $F_{n,d}$ that obeys:*

- $p, q \in C \quad \Rightarrow \quad p + q \in C$
- $p \in C, \lambda \geq 0 \quad \Rightarrow \quad \lambda p \in C$

$P_{n,d}$, $\Sigma_{n,d}$ and $Q_{n,d}$ are closed convex cones.

The definition of our constraint and its properties depend on the choice of a scalar product for forms.

Definition 7.2. *A **scalar product** on $F_{n,d}$ can be defined as*

$$[p, q] = \sum_i \binom{d}{i} p_i \, q_i \,. \tag{7.13}$$

For $Q_{n,d}$ this scalar product has a particularly simple form:

Lemma 7.1. *For $p = \sum_k \langle a_k, \cdot \rangle^d \in Q_{n,d}$ and $q \in F_{n,d}$ we have*

$$[q, p] = \sum_k q(a_k) \,. \tag{7.14}$$

Proof. By the multinomial theorem:

$$\langle \mathbf{a}, \mathbf{x} \rangle^d = \left(\sum_{i=1}^n a_i \, x_i \right)^d = \sum_{i_1 + \cdots + i_n = d} \binom{d}{i} a^i \, x^i \tag{7.15}$$

And so

$$\begin{aligned} [q, p] &= [q, \sum_k \langle \mathbf{a}_k, \cdot \rangle^d] = \sum_k [q, \langle \mathbf{a}_k, \cdot \rangle^d] \\ &= \sum_k \sum_i \binom{d}{i} q_i \cdot (a_k)^i = \sum_k q(\mathbf{a}_k) \,. \end{aligned}$$

□ □

Also note that $[\cdot, \cdot]$ corresponds to the usual scalar product for tensors and $[p, p]$ is the square of the Frobenius norm $\|p\|_F$.

In order to derive a matrix representation, we want to reduce a form of even degree $d = 2s$ to $d' = 2$. For this we need

$$L(\mathbf{x}, t) = \sum_i x^i \, t_i \tag{7.16}$$

with a vector of variables t indexed by $i \in I(n, s)$. For fixed t, this is a $F_{n,s}$ form in $\mathbf{x}$. For fixed $\mathbf{x}$, this is a linear form in t.

Definition 7.3. *For $p \in F_{n,2s}$ the **Hankel form** is the quadratic form*

$$H_p(t) = [p, L^2(\cdot, t)] = \sum_{i,j} p_{i+j} \, t_i \, t_j \; \in F_{|I(n,s)|,2} \,. \tag{7.17}$$

We can put this into matrix form as $H_p(t) = t^T H_p \, t$. For $p \in F_{3,4}$ the matrix is

$$H_p = \begin{pmatrix} p_{xxxx} & p_{xxxy} & p_{xxxz} & p_{xxyy} & p_{xxyz} & p_{xxzz} \\ p_{xxxy} & p_{xxyy} & p_{xxyz} & p_{xyyy} & p_{xyyz} & p_{xyzz} \\ p_{xxxz} & p_{xxyz} & p_{xxzz} & p_{xyyz} & p_{xyzz} & p_{xzzz} \\ p_{xxyy} & p_{xyyy} & p_{xyyz} & p_{yyyy} & p_{yyyz} & p_{yyzz} \\ p_{xxyz} & p_{xyyz} & p_{xyzz} & p_{yyyz} & p_{yyzz} & p_{yzzz} \\ p_{xxzz} & p_{xyzz} & p_{xzzz} & p_{yyzz} & p_{yzzz} & p_{zzzz} \end{pmatrix} . \tag{7.18}$$

Definition 7.4. *A form p with positive semidefinite H_p will be called **H-psd**.*

$$H_{n,d} = \{p \in F_{n,d} : H_p(t) \geq 0 \; \forall \, t\} \tag{7.19}$$

The method we propose enforces the H-psd constraint on fODFs during deconvolution.

In the rest of this section we will discuss some properties of the set $H_{n,d}$ that are relevant to our method. The main tool will be duality:

Definition 7.5. *The **dual cone** of a convex cone C is the set*

$$C^\star = \{p \in F_{n,d} : [p, q] \geq 0, \; \forall \, q \in C\} . \tag{7.20}$$

If C is a closed convex cone then (see Reznick [65])

$$C^{\star\star} = C . \tag{7.21}$$

Theorem 7.1. *$P_{n,d}$ and $Q_{n,d}$ are dual to each other:*

$$Q_{n,d}{}^\star = P_{n,d}, \qquad P_{n,d}{}^\star = Q_{n,d} \tag{7.22}$$

Proof.

$$\begin{aligned} p \in Q_{n,d}{}^\star &\iff 0 \leq [q, p] = [\sum_k \langle \mathbf{a}_k, \cdot \rangle^d, p] = \sum_k p(\mathbf{a}_k), \quad \forall \, q \in Q_{n,d} \\ &\iff 0 \leq p(\mathbf{a}), \quad \forall \, \mathbf{a} \in \mathbb{R}^n \\ &\iff p \in P_{n,d} \end{aligned}$$

The second equation is a consequence of $Q_{n,d}{}^{\star\star} = Q_{n,d}$. □ □

In the special case of $(n, d) = (3, 4)$ the H-psd constraint is equivalent to decomposability into rank-1 terms. This can be shown with the following two theorems:

Theorem 7.2. *(Hilbert)*

$$P_{n,d} = \Sigma_{n,d} \tag{7.23}$$

if and only if $n = 2$ or $d = 2$ or $(n, d) = (3, 4)$.

Theorem 7.3.

$$\Sigma_{n,d}{}^{\star} = H_{n,d} \tag{7.24}$$

Proof. Observe that $t \mapsto L(x,t)$ is a bijection between vectors $t \in \mathbb{R}^{|I(n,s)|}$ and forms in $F_{n,s}$. So:

$$\begin{aligned} p \in \Sigma_{n,d}{}^{\star} &\iff 0 \leq [p,q] = [p, \sum_k h_k^2] \quad \forall\, q \in \Sigma_{n,d} \\ &\iff 0 \leq [p,h^2] \quad \forall\, h \in F_{n,s} \\ &\iff 0 \leq [p, L(\cdot,t)^2] = H_p(t) \quad \forall\, t \in \mathbb{R}^{|I(n,s)|} \end{aligned}$$

□ □

A direct consequence for $(n,d) = (3,4)$ is:

Corollary 1. *$p \in F_{3,4}$ is a sum of fourth powers iff it is H-psd, since*

$$Q_{3,4} = P_{3,4}{}^{\star} = \Sigma_{3,4}{}^{\star} = H_{3,4}\,. \tag{7.25}$$

For higher degrees the relation is weaker:

$$Q_{3,n} \subsetneq H_{3,n} \subset P_{3,n} \tag{7.26}$$

Another property of the H matrix is that it can be used to estimate the number of fibers in an fODF.

Definition 7.6. *The **rank** of $p \in Q_{n,d}$ is the smallest integer $\mathsf{rank}(p) = r$ for which $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r \in \mathbb{R}^n$ can be found with*

$$p = \sum_{k=1}^{r} \langle \boldsymbol{a}_k, \cdot \rangle^d\,. \tag{7.27}$$

For the cases in Hilbert's theorem 7.2, the ranks of p and H_p are equal as shown in theorem 4.6 in [65]. In general $\mathsf{rank}(p) \geq \mathsf{rank}(H_p)$.

Conflict of Interests

This work was supported by the DFG under grant SCHU 3040/1-1. LHL is supported by AFOSR FA9550-13-1-0133, DARPA D15AP00109, NSF IIS 1546413, DMS 1209136, and DMS 1057064. The authors declare that they have no conflict of interest.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

The study was approved by the local ethics committee. All volunteers gave their written informed consent for participation in the study.

8. A Sheet Probability Index from Diffusion Tensor Imaging

Abstract *A sheet probability index (SPI) has recently been derived from high angular resolution diffusion MRI to quantify the hypothesis that white matter tracts are organized in parallel sheets of interwoven paths. In this work, we derive the DTI-SPI, a variant of the SPI that can be computed from the widely available, simple, and fast diffusion tensor imaging, by considering the normal component of the Lie bracket of the major and medium eigenvector fields. We observe that, despite the fact that DTI does not allow us to infer crossing fiber orientations, the DTI-SPI has a meaningful interpretation in terms of sheet structure if the largest pair of eigenvectors spans the same plane as the two dominant fibers. We report empirical results that support this assumption. We also show a direct comparison to the previously proposed SPI on data from the human connectome project, and demonstrate that major features in maps of our DTI-SPI remain recognizable in standard clinical DTI data.*

Comment *This chapter corresponds to the paper[6]: "A Sheet Probability Index from Diffusion Tensor Imaging" by Michael Ankele et al. in Computational Diffusion MRI, 2018. The paper introduces a simple method for numerical estimation of the Lie bracket of certain vector fields. For additional experiments evaluating this method on synthetic tensor fields see chapter 9.*

8.1. Introduction

Do white matter fiber tracts cross in such a way that they form two-dimensional sheets, similar to pages in a book? This hypothesis was first suggested a few years ago by Wedeen et al. [97]. Based on inspecting streamlines from diffusion spectrum imaging tractography, which they performed on humans and on four nonhuman primates, and supported by additional numerical experiments, they claimed that, independent of species, fiber orientation, and curvature, crossing pathways formed well-defined two-dimensional sheets. They point out that it is highly unlikely a priori that two independent families of 3D curves should form such surfaces, and claim that their finding therefore reveals a previously unknown organizational principle of cortico-cortical pathways.

The sheet structure hypothesis has been a subject of controversial debate. In a technical comment, Catani et al. [14] focus on the fact that its original formulation contained frequent references to the "near-orthogonality" of the claimed grid structure. Based on results from constrained spherical deconvolution, they demonstrate that fibers cross at a wide range of angles, with no specific preference for orthogonality. As part of their

response, Wedeen et al. [95] emphasize that they consider integrability in terms of the Frobenius theorem to be at the core of their hypothesis, and that integrability does not depend on orthogonality.

Clearly, more formal and quantitative experiments are required to settle this dispute. In particular, visual inspection is an error-prone way of assessing sheet structure, since even nonintegrable curves might appear to form a surface from certain viewpoints. Therefore, Tax et al. [84, 83] recently proposed a sheet probability index (SPI) as an objective measure of the local presence of sheet structure in crossing fiber tracts.

Given that crossing tracts are at the core of the sheet structure hypothesis, it is unsurprising that all previous works on this topic have been based on crossing fiber models, and some have even emphasized the differences between specific alternatives [14, 24]. In this paper, we argue that, even though it is widely known that the diffusion tensor model is unable to resolve the orientations of crossing fibers, it is still possible to derive an informative measure of sheet structure from it. Our main argument is based on

1. the mathematical observation that, in order to quantify the extent to which crossing fibers form a sheet, it is enough to know the plane spanned by them, we do not require their exact orientations within that plane (cf. Section 8.2);

2. experimental evidence in support of a previously formulated hypothesis stating that, in voxels that contain two main fibers, the diffusion tensor provides an accurate estimate of the plane spanned by them (cf. Section 8.3).

Based on this line of reasoning, we derive the DTI-SPI, a variant of the SPI that can be computed from the widely available diffusion tensor imaging (Section 8.4). We also describe an algorithm to compute the DTI-SPI, visually compare it to the established SPI, and demonstrate structures in its maps that can be reproduced across subjects (Section 8.5).

8.2. Theoretical Part of Our Argument

To formalize the hypothesis that two given tracts form a two-dimensional sheet, let $\mathbf{v}$ and $\mathbf{w}$ denote vector fields that locally indicate the directions of the tracts, and assume that they are linearly independent. We call $\mathbf{v}$ and $\mathbf{w}$ integrable if there exists a surface whose tangential plane at every point $\mathbf{p}$ is given by $\Delta_p = \mathsf{span}(\mathbf{v_p}, \mathbf{w_p})$. Similar to previous work [84, 83], we count non-integrability of $\mathbf{v}$ and $\mathbf{w}$ as evidence against the sheet structure hypothesis.

Interestingly, standard references on differential geometry [79, Chapter 6] formulate Frobenius integrability theory not in terms of vector fields, but rather in terms of a field of tangential planes Δ. In this context, such a field is called a two-dimensional distribution. This clarifies that integrability depends only on the planes spanned by the crossing tracts at each location, their orientations within that plane do not play any role for integrability. Even though Wedeen et al. [95] previously clarified that orthogonality is not a requirement for integrability, the mathematical fact that the fiber orientations are

indeed irrelevant for integrability as long as they remain in the same plane is not widely appreciated, as can be seen from recent claims such as "Without an accurate estimation of the intravoxel fiber distribution, any statements about sheet structure, however they are formulated, will be relatively meaningless." [24]

The exact condition of integrability is in terms of the Lie bracket $[\Delta, \Delta]$

$$[\Delta, \Delta] \subset \Delta \tag{8.1}$$

The definition of the Lie bracket is well-known in differential geometry, and has been explained in an accessible manner in related work [84, 83]. We repeat it in Appendix 8.7 to make our paper self-contained. In terms of the vector fields $\mathbf{v}$ and $\mathbf{w}$, integrability requires that

$$[\mathbf{v}, \mathbf{w}] \in \Delta \quad \forall \mathbf{v}, \mathbf{w} \in \Delta\,. \tag{8.2}$$

As in the previously defined sheet probability index [84, 83], we use the normal component of the Lie bracket

$$[\mathbf{v}, \mathbf{w}]^N = \left\langle [\mathbf{v}, \mathbf{w}], \frac{\mathbf{v} \times \mathbf{w}}{\|\mathbf{v} \times \mathbf{w}\|} \right\rangle \tag{8.3}$$

as a measure of nonintegrability, since $\mathbf{v}, \mathbf{w}$ describe a surface iff $[\mathbf{v}, \mathbf{w}]^N = 0$.

8.3. Empirical Part of Our Argument

In the previous section, we argued that it does not matter which specific pair of vector fields we use to test integrability: As long as two pairs span the same planes at all points of the domain, one pair will be integrable if and only if the other pair is integrable. We will now argue that this allows us to formulate a meaningful measure of sheet structure from the diffusion tensor model, even though we cannot derive the exact orientations of crossing fibers from it.

The key of our argument is that, in voxels in which two or more fiber compartments cross, the dominant two fibers will lie in the plane spanned by the major and medium eigenvectors of the diffusion tensor. We consider this assumption to be quite natural, and we note that it has been used by others to constrain two-fiber tractography [58]; the resulting method has been validated with the help of functional MRI [61].

We provide additional evidence for this assumption by demonstrating a close agreement, in eight subjects from the human connectome project (HCP) [78], between the planes spanned by the major and medium DTI eigenvectors, and the planes spanned by the two dominant fibers as estimated by a crossing fiber model. In particular, we use constrained higher order tensor-based multi-tissue deconvolution [3]. It represents fiber orientation distribution functions (fODFs) as fourth-order tensors, and finds the two dominant fiber directions via a rank-two approximation [70]. This method has recently been shown to produce similar crossing-fiber tractography as standard constrained spherical deconvolution [40], while enabling faster and numerically more stable computation [2].

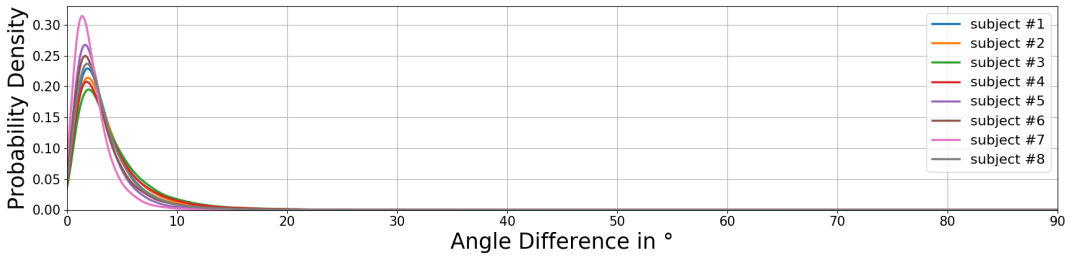


Figure 8.1.: The distribution of angles between the minor eigenvector from DTI and the normal vector from a crossing fiber model indicates a strong agreeement of the two.

For all eight subjects, Figure 8.1 shows the distribution of angles between the minor eigenvector and the vector orthogonal to the two dominant fiber directions from the crossing fiber model. It only makes sense to compute this angle when both vectors are uniquely defined, at least up to their signs, which we discard after taking their dot product. This requires that the crossing fiber model found at least two fibers, and that the smallest eigenvalue has multiplicity one. We enforce the latter condition by computing diffusion tensor planarity c_p, as defined in terms of sorted eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ by Westin et al. [100],

$$c_p := \frac{\lambda_2 - \lambda_3}{\lambda_1}, \tag{8.4}$$

and restricting our computation to a mask in which $c_p > 0.2$.

As shown in Figure 8.1, our assumption that the two vectors should coincide is met in all subjects, up to an average deviation between $2.3° - 4.1°$. Since this is below the expected reconstruction error of crossing fiber directions ($5° - 10°$ according to [3], depending on the crossing angle) and since the plot does not indicate any notable outliers, we conclude that our assumption is met with reasonably high accuracy.

Of course, this assumption is limited to the regions where two main fiber directions are dominant. Both above-mentioned conditions together were met in $54\% - 64\%$ of all white matter, which is a quite similar range as the fraction of two-fiber voxels estimated in a previous study by Jeurissen et al. [39]. As shown in Figures 8.2, 8.3 and 8.4, our DTI-based measure of sheet structure can be computed even outside of this mask. While it is certainly not obvious how it should be interpreted in regions with only one tract, it might retain a meaningful interpretation in terms of the two strongest fibers in some regions where a third fiber is close to perpendicular to them, and has a sufficiently low volume fraction. We hypothesize that, in these cases, the third fiber will not have a strong effect on the eigenvectors of the diffusion tensor. However, we leave a more detailed investigation as a potential topic for future work.

In any case, it is clearly beyond the possibilities of DTI to study crossing sheet structures, which might arise where three significant bundles cross, and for which Tax et al. [84] have demonstrated one example based on crossing fiber tractography.

8.4. Quantifying Sheet Structure With DTI

8.4.1. Brief Review of the Existing Sheet Probability Index

Tax et al. [84] define their sheet probability index (SPI) as the local probability that the normal component of the Lie bracket is within a symmetric interval $[-\lambda, \lambda]$, where λ is a tunable parameter. This probabilistic formulation accounts for the fact that we cannot expect the normal component of the Lie bracket to be exactly zero, since it is estimated from noisy measurements.

They proposed two different algorithms to compute the SPI in practice [84, 83]. Since diffusion MRI measurements are so time consuming that it is not feasible to repeat them many times, both implementations are based on a bootstrapping technique [18, 38] that estimates the measurement noise from model fitting residuals in a single set of measurements.

In each bootstrapping iteration and each voxel, the first algorithm [84], named *flows-and-limits,* approximates the Lie bracket of any pair of local fiber directions using crossing fiber tractography. The approximation is obtained by integrating several quadrilaterals, similar to those in Figure 8.5 right, for each fiber pair, and consolidating the results in a least squares fit.

The second algorithm [83], named *the coordinate approach,* first derives a multi-vector field, in which each point is assigned up to three local fODF peak directions, and then clusters it into three separate fields that isolate different tracts. The definition of the Lie bracket in terms of vector field derivatives (cf. Eq. (8.6)) is then evaluated using normalized convolution, which accounts for missing peaks and sufficiently large spatial neighborhoods.

In both cases, computational effort is reduced by taking relatively few bootstrap realizations (50 in [84], 20 in [83]), and fitting a Gaussian to the results. The final estimate of the SPI is computed based on that Gaussian; no SPI is computed in voxels where the samples fail a statistical test for Gaussianity.

8.4.2. The Normal Part of the Lie Bracket in DTI

According to our argument, a meaningful variant of the sheet probability index can be derived from the diffusion tensor model, based on the normal part of the Lie bracket of the major and medium eigenvector fields. We will denote this normal part as ζ. Since the three eigenvector fields, which we will sort in descending order of associated eigenvalues and refer to as $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, are orthonormal, Eq. (8.3) simplifies to

$$\zeta := [\mathbf{X}, \mathbf{Y}]^N = \langle [\mathbf{X}, \mathbf{Y}], \mathbf{Z} \rangle \,. \tag{8.5}$$

A simple and efficient algorithm for practical computation of ζ, based on evaluating a few closed-form equations, can be derived by expressing it in terms of partial derivatives of the tensor field D, and the eigenvectors and -values of D. It is given in Appendix 8.8. Alternatively, ζ can be expressed exclusively in terms of the minor eigenvector field, as demonstrated in Appendix 8.9.

8.4.3. A DTI-Based Sheet Probability Index

Our DTI-SPI is defined from ζ, i.e., the normal part of the Lie bracket in the diffusion tensor field, in analogy to the standard SPI: It equals the probability of ζ falling into the interval $[-\lambda, \lambda]$. We also use the same bootstrapping approach to estimate the distribution of tensor fields given the measurement noise. However, in each iteration, the equations from Appendix 8.8 allow us to compute ζ in closed form, without having to perform any tractography, clustering, or convolution over large neighborhoods. Rather, we compute derivatives using a computationally efficient B-spline approximation [54], whose continuous nature also makes it easy to compute DTI-SPI maps at a resolution that exceeds the one of the original data (in our results, factor 4 along both axes).

Due to the greatly reduced computational cost of evaluating the Lie bracket, we can afford not just higher resolution, but also a much larger number of iterations, which makes it possible to estimate the DTI-SPI nonparametrically, i.e., by simply counting the fraction of bootstrap samples in which $\zeta \in [-\lambda, \lambda]$. Shown results use 1000 iterations, but we found that maps had visually converged already at 100. Based on inspecting histograms of ζ at different locations, it did appear to approximately follow a Gaussian distribution in most cases. However, not making an explicit assumption about this simplifies the algorithm, and allows us to compute the DTI-SPI in all voxels.

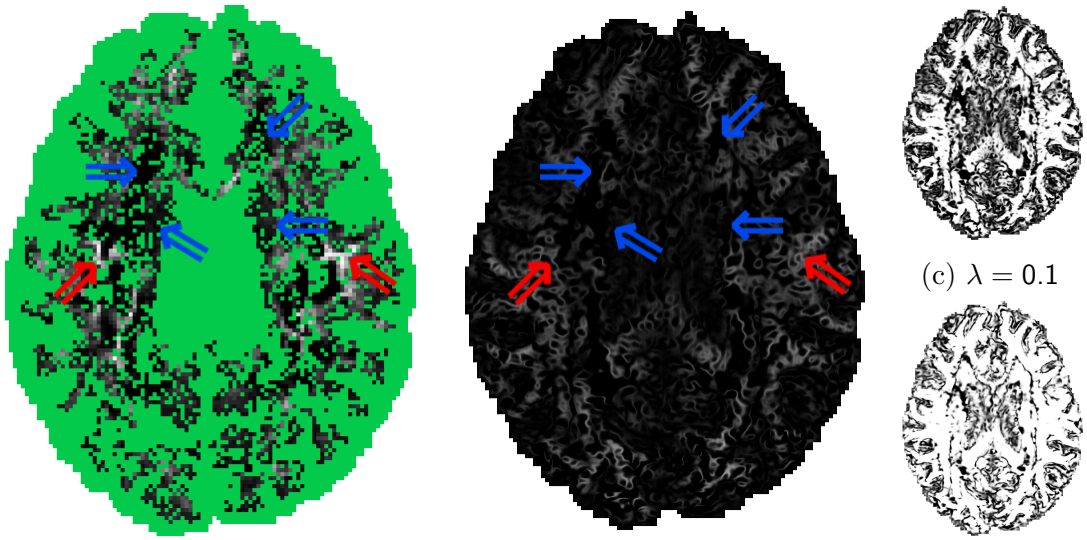
8.5. Results and Discussion

8.5.1. Impact of λ and Comparison to Traditional SPI

Depending on its λ parameter, the SPI might either give the impression that sheet structure is rare or very common in the brain. This is illustrated for our DTI-SPI in Fig. 8.2 (b)–(d), but also applies to its original definition. In the light of previous claims that "no brain pathways were observed without sheet structure" [97], one notable observation is that there are certain white matter structures that are persistently classified as nonintegrable, even with rather large λ.

Tax et al. compute their SPI for each pair of peaks. According to Section 8.3, we expect our DTI-SPI to be most comparable to their SPI when it is computed from the two dominant fiber directions. The authors of [83] kindly provided such a map to us, shown in Fig. 8.2 (a). It has been computed using the coordinate approach on the same HCP subject (101006) as our own maps. Locations at which the SPI has not been computed, due to a lack of a second fODF peak or a violation of Gaussianity, are shown in green.

We found the similarity to be highest when comparing to the DTI-SPI with $\lambda = 0.01$ ($\lambda = 0.008$ is used in [83]). Several regions of high or low SPI in Fig. 8.2 (a) can be recognized in (b), marked by red and blue arrows, respectively. However, the agreement is not perfect, and the fragmented nature of Fig. 8.2 (a) makes a more detailed comparison difficult. Unlike the traditional SPI, our DTI-SPI can be evaluated everywhere in the brain. As noted in Section 8.3, interpretation in regions with more than two tracts requires additional investigation, and might account for some of the differences observed



(a) SPI from [83] (b) DTI-SPI with $\lambda = 0.01$ (c) $\lambda = 0.1$ (d) $\lambda = 0.2$

Figure 8.2.: Our DTI-SPI with $\lambda = 0.01$ (b) reproduces several salient regions of high (red) or low (blue arrows) values in the previous SPI [83], computed from the two largest fODF peaks (a). The effect of different settings of λ on our DTI-SPI is shown in (b)–(d).

in Fig. 8.2.

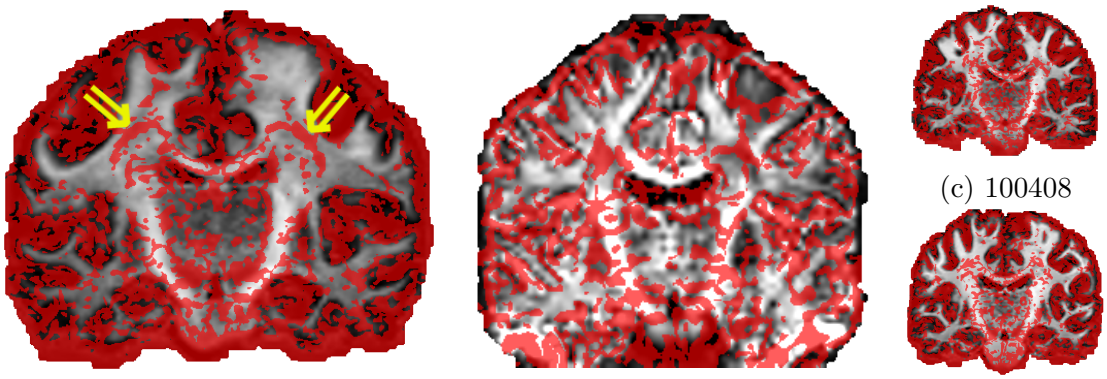
Most of the brighter regions in Fig. 8.2 (b) exhibit texture-like patterns, at the scale of the voxel resolution. They indicate that ζ fluctuates around zero in those regions. The exact configuration of these patterns probably carries no anatomical significance, and they get merged into uniform regions with increasing λ.

8.5.2. Comparison of Subjects and Acquisition Schemes

For a more detailed investigation of those white matter regions in which our DTI-SPI indicates nonintegrability, we set $\lambda = 0.1$, and thresholded DTI-SPI < 0.5. In Figures 8.3 and 8.4, the results are overlaid (in red) on white matter segmentations that provide anatomical context.

We compare results on three HCP subjects (90 DWIs with $b = 1000$, 18 references with $b = 5$, spatial resolution $(1.25\,\mathrm{mm})^3$) to another subject that has been scanned with the default DTI protocol on a clinical 3T Philips Achieva (32 DWIs with $b = 1000$, one reference with $b = 0$, $(2\,\mathrm{mm})^3$). All data remains in the respective scanner space, but we selected anatomically corresponding slices.

The most salient structure in Fig. 8.3 is a region of low DTI-SPI that extends over part of the region where projection fibers and the corpus callosum merge, as well as the boundary towards the superior longitudinal fasciculus (SLF). It is marked by arrows in (a), but is present also in the other subjects, and remains recognizable even in data



(a) HCP 101006 (b) Clinical (c) 100408 (d) 101107

Figure 8.3.: Our DTI-SPI indicates nonintegrability in part of the region where the corpus callosum merges with projection fibers, and at the interface towards the SLF. This is highlighted by yellow arrows in (a), but visible also in other HCP subjects (c/d), as well as in a subject that has been scanned with a clinical DTI protocol (b).

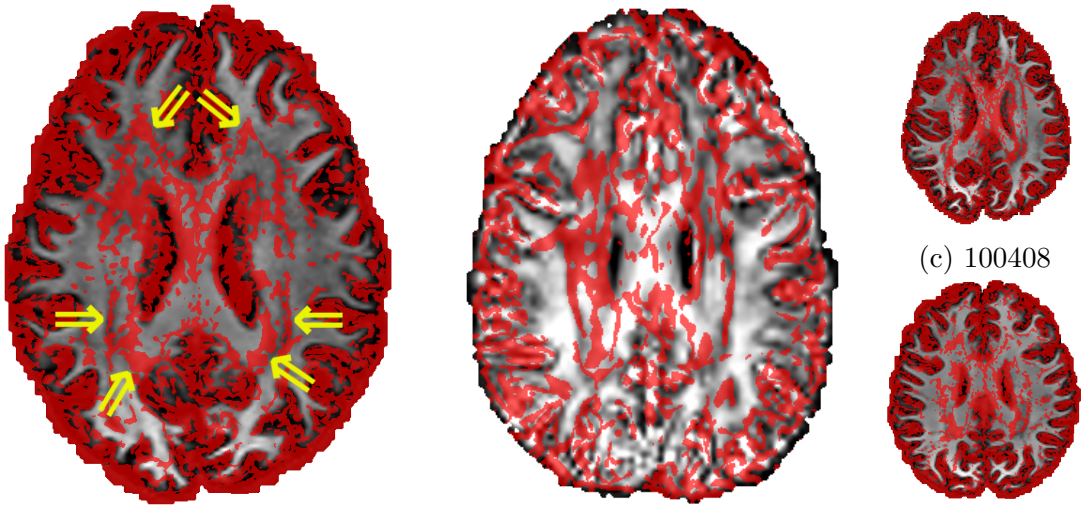
that was acquired under clinical conditions. We inspected results from tensor-based spherical deconvolution [3] on HCP subject 101006 to confirm the presence of two dominant fibers in a large part of these regions, so interpreting our DTI-SPI as an indication of nonintegrability should be justified.

The axial slices in Fig. 8.4 confirm that the low values of DTI-SPI in regions of merging tracts, as well as at the interface between corticospinal tract and SLF, are not limited to the particular coronal slice in Fig. 8.3, but have a significant spatial extent in all subjects.

8.6. Conclusion

Based on a mathematical and an empirical observation, we argue that the diffusion tensor model can be used to study the sheet structure hypothesis despite its inability to indicate crossing fiber directions. Consequently, we have derived a novel sheet probability index from diffusion tensor imaging (DTI-SPI), and presented an algorithm for computing it.

A theoretical insight from our work is that, in a substantial part of the brain, the assumptions required to infer sheet structure from diffusion MRI are weaker than it has been thought previously [83]: In the presence of two dominant tracts, we do not have to assume that fODF peaks directly indicate their directions; it is enough if we are able to deduce the planes in which the crossings happen. Practical benefits of our DTI-SPI are the facts that it is simple and fast to compute, and that it imposes more modest requirements on the data acquisition. The price we have to pay for this is the inability to investigate crossing sheet structures, which have been shown to occur in voxels with more than two tracts.



(a) HCP 101006 (b) Clinical (c) 100408 (d) 101107

Figure 8.4.: Extended regions of low DTI-SPI that occur at the interfaces of tracts, or where fibers from different tracts merge, can be observed also in these axial slices. They are again reproduced in all subjects.

A preliminary comparison indicates several similarities to the standard SPI from the two major fibers [83]. A more detailed anatomical interpretation and comparison to the traditional SPI are planned for our future work. We also found salient features in our maps that were reproducible across subjects and acquisition protocols. This suggests a potential application of ζ as a biomarker that is simple to evaluate, and complementary to the commonly used DTI-based measures, such as Fractional Anisotropy or Mean Diffusivity. In the future, we plan to apply statistical hypothesis testing to investigate if age, gender, or disease may have an effect on sheet structure as quantified by ζ.

8.7. Definition and Intuition of the Lie Bracket

Our test of integrability makes use of the Lie bracket $[\mathbf{X}, \mathbf{Y}]$ of differentiable vector fields $\mathbf{X}$ and $\mathbf{Y}$, which is formally defined as

$$[\mathbf{X}, \mathbf{Y}] = \nabla_{\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}\mathbf{X} \tag{8.6}$$

$$= \sum_{i,j} (X^i \partial_i)(Y^j \mathbf{e}_j) - (Y^i \partial_i)(X^j \mathbf{e}_j) \tag{8.7}$$

$$= \sum_{i,j} (X^i \, \partial_i Y^j - Y^i \, \partial_i X^j)\mathbf{e}_j, \tag{8.8}$$

with $\mathbf{e}_j$ indicating the canonical basis vectors.

One interpretation of this bracket comes from the flow Φ of a vector field $\mathbf{X}$ [79, Chapter 5], which maps each point $\mathbf{p}$ to its integral curve $\Phi_t(\mathbf{p})$, i.e. $\Phi_0(\mathbf{p}) = \mathbf{p}$ and

Figure 8.5.: *Left:* The flow Φ of a vector field $\mathbf{X}$ and a specific integral curve. *Right:* A loop of vector field integral curves failing to close due to non-vanishing Lie brackets.

everywhere tangential to $\mathbf{X}$, see Figure 8.5 (left). Then, given two vector fields $\mathbf{X}$ and $\mathbf{Y}$ with corresponding flow Φ and Ψ and a step size h, we can start at point $\mathbf{p}$ and first move along Φ_h, then from the new point along Ψ_h, then backwards along Φ_{-h} and finally backwards along Ψ_{-h} (right subfigure).

This will form a closed loop if Φ and Ψ commute, that is if

$$[\mathbf{X}, \mathbf{Y}] = 0\,. \tag{8.9}$$

The difference between $\mathbf{p}$ and the end point $\mathbf{q}$ can be given in terms of the bracket:

$$\mathbf{q} - \mathbf{p} = h^2[\mathbf{X}, \mathbf{Y}] + O(h^3) \tag{8.10}$$

8.8. Practical Computation of the Lie Bracket

The diffusion tensor D can be rotated into its eigenbasis $\mathbf{v}_i$:

$$RDR^T = \Sigma = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3) \tag{8.11}$$

with

$$R = \begin{pmatrix} \mathbf{v}_1{}^T \\ \mathbf{v}_2{}^T \\ \mathbf{v}_3{}^T \end{pmatrix}, \qquad R^T = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} & \mathbf{Z} \end{pmatrix}. \tag{8.12}$$

Define the matrices

$$M_{i,jk} = \langle \partial_i \mathbf{v}_j, \mathbf{v}_k \rangle\,. \tag{8.13}$$

These are antisymmetric for every i since

$$0 = \partial_i \underbrace{\langle \mathbf{v}_j, \mathbf{v}_k \rangle}_{0 \text{ or } 1} = \langle \partial_i \mathbf{v}_j, \mathbf{v}_k \rangle + \langle \mathbf{v}_j, \partial_i \mathbf{v}_k \rangle \tag{8.14}$$

so we can write them as

$$M_i = (\partial_i R) R^T = \begin{pmatrix} 0 & c_i & -b_i \\ -c_i & 0 & a_i \\ b_i & -a_i & 0 \end{pmatrix} \tag{8.15}$$

with

$$a_i = \langle \partial_i \mathbf{Y}, \mathbf{Z} \rangle \qquad b_i = \langle \partial_i \mathbf{Z}, \mathbf{X} \rangle \qquad c_i = \langle \partial_i \mathbf{X}, \mathbf{Y} \rangle \tag{8.16}$$

$$\mathbf{a} = \nabla \mathbf{Y} \cdot \mathbf{Z} \qquad \mathbf{b} = \nabla \mathbf{Z} \cdot \mathbf{X} \qquad \mathbf{c} = \nabla \mathbf{X} \cdot \mathbf{Y} \,. \tag{8.17}$$

Rotating the partial derivatives $\partial_i D$ of the tensor field then yields

$$\begin{aligned} R(\partial_i D)R^T &= \partial_i(RDR^T) - (\partial_i R)DR^T - RD(\partial_i R^T) && (8.18)\\ &= \partial_i \Sigma - (\partial_i R)R^T\Sigma - \Sigma R(\partial_i R^T) && (8.19)\\ &= \partial_i \Sigma - (\partial_i R)R^T \cdot \Sigma + \Sigma \cdot (\partial_i R)R^T && (8.20)\\ &= \partial_i \Sigma - M_i\Sigma + \Sigma M_i && (8.21)\\ &= \begin{pmatrix} \partial_i \lambda_x & c_i(\lambda_1 - \lambda_2) & b_i(\lambda_3 - \lambda_1) \\ c_i(\lambda_1 - \lambda_2) & \partial_i \lambda_2 & a_i(\lambda_2 - \lambda_3) \\ b_i(\lambda_3 - \lambda_1) & a_i(\lambda_2 - \lambda_3) & \partial_i \lambda_3 \end{pmatrix} . && (8.22) \end{aligned}$$

This leads to the following closed-form equation for ζ:

$$\begin{aligned} \langle [\mathbf{X}, \mathbf{Y}], \mathbf{Z} \rangle &= \langle \nabla_{\mathbf{X}} \mathbf{Y}, \mathbf{Z} \rangle - \langle \nabla_{\mathbf{Y}} \mathbf{X}, \mathbf{Z} \rangle && (8.23)\\ &= \langle \nabla_{\mathbf{X}} \mathbf{Y}, \mathbf{Z} \rangle + \langle \mathbf{X}, \nabla_{\mathbf{Y}} \mathbf{Z} \rangle && (8.24)\\ &= \mathbf{X} \cdot \nabla \mathbf{Y} \cdot \mathbf{Z} + \mathbf{Y} \cdot \nabla \mathbf{Z} \cdot \mathbf{X} && (8.25)\\ &= \langle \mathbf{X}, \mathbf{a} \rangle + \langle \mathbf{Y}, \mathbf{b} \rangle && (8.26) \end{aligned}$$

with

$$a^i = \frac{1}{\lambda_2 - \lambda_3} \left[R(\partial_i D)R^T \right]_{2,3} \tag{8.27}$$

$$b^i = \frac{1}{\lambda_3 - \lambda_1} \left[R(\partial_i D)R^T \right]_{1,3} \tag{8.28}$$

where the square bracket notation $[\cdot]_{i,j}$ denotes taking the (i,j)th component of a matrix.

8.9. Expression in Terms of Minor Eigenvector Field

Since the eigenvectors are orthonormal, it is obvious that a surface is everywhere tangential to the major and medium eigenvectors if and only if its surface normal is everywhere parallel to the minor eigenvector. This can be illustrated by the following derivation of

an alternative expression of ζ, that only involves the minor eigenvector field $\mathbf{Z}$:

$$\begin{aligned}
\zeta &= \langle \nabla_{\mathbf{X}} \mathbf{Y}, \mathbf{Z} \rangle - \langle \nabla_{\mathbf{Y}} \mathbf{X}, \mathbf{Z} \rangle && (8.29)\\
&= -\langle \mathbf{Y}, \nabla_{\mathbf{X}} \mathbf{Z} \rangle + \langle \mathbf{X}, \nabla_{\mathbf{Y}} \mathbf{Z} \rangle && (8.30)\\
&= -\sum_{j,k} Y^k X^j \, \partial_j Z^k + \sum_{j,k} X^k Y^j \, \partial_j Z^k && (8.31)\\
&= \sum_{j,k,l,m} X^l Y^m (-\delta_{lj}\delta_{mk} + \delta_{lk}\delta_{mj}) \partial_j Z^k && (8.32)\\
&= \sum_{i,j,k,l,m} X^l Y^m \, \epsilon_{ilm} \, \epsilon_{ikj} \, \partial_j Z^k && (8.33)\\
&= \sum_{i,j,k} Z^i \, \epsilon_{ikj} \, \partial_j Z^k && (8.34)\\
&= -\langle \mathbf{Z}, \operatorname{rot} \mathbf{Z} \rangle && (8.35)
\end{aligned}$$

with the Kronecker delta

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{8.36}$$

and the totally antisymmetric Levi-Civita symbol

$$\begin{aligned}
\epsilon_{123} = \epsilon_{231} = \epsilon_{312} &= 1 && (8.37)\\
\epsilon_{321} = \epsilon_{132} = \epsilon_{213} &= -1 && (8.38)\\
\epsilon_{\text{other}} &= 0\,. && (8.39)
\end{aligned}$$

9. DT-MRI Streamsurfaces Revisited

Abstract *DT-MRI streamsurfaces, defined as surfaces that are everywhere tangential to the major and medium eigenvector fields, have been proposed as a tool for visualizing regions of predominantly planar behavior in diffusion tensor MRI. Even though it has long been known that their construction assumes that the involved eigenvector fields satisfy an integrability condition, it has never been tested systematically whether this condition is met in real-world data. We introduce a suitable and efficiently computable test to the visualization literature, demonstrate that it can be used to distinguish integrable from nonintegrable configurations in simulations, and apply it to whole-brain datasets of 15 healthy subjects. We conclude that streamsurface integrability is approximately satisfied in a substantial part of the brain, but not everywhere, including some regions of planarity. As a consequence, algorithms for streamsurface extraction should explicitly test local integrability. Finally, we propose a novel patch-based approch to streamsurface visualization that reduces visual artifacts, and is shown to more fully sample the extent of streamsurfaces.*

Comment *This chapter corresponds to the paper[5]: "DT-MRI Streamsurfaces Revisited" by Michael Ankele et al. in IEEE Transactions on Visualization and Computer Graphics (2019).*

9.1. Introduction

Streamlines and streamsurfaces are among the first and most fundamental visualization techniques for diffusion tensor MRI [103, 73]. While streamlines are defined as curves that are everywhere tangential to the major eigenvector field, and are used in "linear" regions where the major eigenvalue is clearly dominant, streamsurfaces have been proposed for "planar" regions, in which the major and medium eigenvalues are similar in magnitude, but larger than the minor eigenvalue. They have been defined as surfaces that are everywhere tangential to the major and medium eigenvector fields [103].

Despite the close analogy between streamlines and streamsurfaces, there is an important difference between them: Any generic vector field can be integrated to form streamlines, but a pair of vector fields defines a consistent surface, even locally, only if it meets a specific integrability condition, which is formulated in the Frobenius theorem, one of the fundamental theorems of differential geometry and topology [79].

Even though this fact has been pointed out already when streamsurfaces were first introduced to DT-MRI visualization [103], its consequences for visualization have remained controversial. Some authors have reported that they were able to compute streamsurfaces without visible artifacts [103], and conclude that, therefore, the integra-

bility condition is likely met in practice. A recent textbook [59] introduces streamsurfaces without even mentioning the caveat of integrability. Other authors have concluded that streamsurfaces are unsafe, based on a counter example in which the integration result strongly depended on arbitrary choices during their extraction, which they attributed to the fact that the eigenvector fields did not define a surface locally [72].

The key contribution of our work is to investigate the following questions, which we consider to be critical for DT-MRI streamsurfaces as a visualization tool:

Q1 How can we test whether the integrability condition, which is a crucial assumption of DT-MRI streamsurfaces, is satisfied at a given point in a given DT-MRI dataset? The relevance of such a test was already mentioned in the original paper on streamsurfaces, which stated: "We have yet to test if the condition is satisfied by the first two eigenvector fields because calculating the Lie bracket is nontrivial." [103]

Q2 Is the failure case of DT-MRI streamsurfaces that was reported in [72] an isolated incident, or do diffusion tensor fields from different subjects and acquisition protocols contain significant regions in which streamsurfaces cannot be used for visualization?

Q3 DT-MRI streamsurfaces have been proposed primarily to visualize regions of planar diffusion. Is the integrability condition reliably met in regions of planarity?

Q4 Do insights on streamsurface integrability suggest changes to existing algorithms for streamsurface extraction and rendering?

Section 9.3 answers **Q1** by introducing a measure of non-integrability that we denote as ζ. Section 9.4 investigates **Q2** and **Q3** by computing ζ on an ensemble of 15 healthy subjects and correlating it with anatomical structures and regions of planarity. Finally, Section 9.5 discusses practical consequences for DT-MRI visualization and suggests an alternative, patch-based algorithm for streamsurface rendering (**Q4**), before Section 9.6 concludes our paper.

9.2. Related Work

In the visualization literature, the word streamsurface is used to refer to two different mathematical concepts. Our paper is concerned with the definition that was first provided in Zhang et al.'s work on the visualization of diffusion tensor MRI [103]. Here, streamsurfaces are defined as surfaces that are everwhere tangential to two eigenvector fields, and an algorithm is introduced that grows them from a single seed point. Sondershaus and Gumhold prefer to call these surfaces diffusion surfaces [77] to reduce the risk of confusing them with the second definition of streamsurfaces, which is discussed below. Vilanova et al. [94] propose an algorithm that switches between creating streamlines and streamsurfaces in a data-adaptive manner.

Our work relies on the Lie bracket of the first two eigenvector fields. This goes back to well-known results from differential topology [79] and its relevance for DT-MRI streamsurfaces was pointed out already by Zhang et al. [103]. However, this test has never been

performed in practice, nor have its implications for DT-MRI visualization been studied. Our work is inspired by papers outside of visualization, which used the Lie bracket to define a Sheet Probability Index [84, 83] in the context of the sheet structure hypothesis, which states that fiber bundles in the brain form parallel sheets of interwoven paths [97]. These works did not compute the Lie bracket for the major and medium eigenvector fields, but for different pairs of vector fields that result from a multi-fiber model and may be non-orthogonal. In our own recent work, we argue that, even though DT-MRI cannot resolve fiber crossings, it can be used to derive an alternative Sheet Probability Index [6]. Unlike this prior work, our present paper investigates DT-MRI streamsurfaces as a visualization technique. Section 9.5.1 will put our present main findings into context of those from [6].

In vector field visualization, the word streamsurface refers to surfaces that are generated by advecting a given seeding curve along a vector field [33]. In the context of stress tensors from geomechanics, this concept has also been applied to the visualization of eigenvector fields [37]. The problem of integrability, which is at the core of our paper, does not affect streamsurfaces in this second sense, due to their completely different definition. However, as we will discuss in more detail in Section 9.5.2, our work is related to attempts of finding surfaces that are as perpendicular as possible to 3D steady vector fields [76].

9.3. Testing DT-MRI Streamsurface Integrability

This section addresses **Q1** by introducing the normal part of the Lie bracket, which we denote ζ, to the visualization literature. It is relevant for DT-MRI streamsurface visualization because it can serve as a measure of surface nonintegrability.

We start with examples that illustrate the problem of DT-MRI streamsurface integrability (Section 9.3.1). To make our paper more self-contained, Section 9.3.2 introduces the involved mathematical concepts, before Section 9.3.3 goes into detail on the detection and practical computation of ζ. Section 9.3.4 establishes a novel theoretical result on the sign of the Lie bracket of eigenvector fields, which is relevant to our experiments later on.

9.3.1. The Problem of DT-MRI Streamsurface Integrability

When Zhang et al. introduced DT-MRI streamsurfaces, they described them as "the approximation of the surface that extends along both the major eigenvector field and medium eigenvector field. At any point on a streamsurface, the major and medium eigenvectors lie in the tangent plane to the surface" [103]. Most subsequent publications simply defined streamsurfaces as the surfaces that result from the previously proposed algorithm for their extraction [77, 94], or assumed that they are intuitively understood [59].

In order to discuss the problem of DT-MRI streamsurface integrability, we require a more formal mathematical definition. Based on the explanation above, the following definition seems natural:

Definition 1: Given a smooth symmetric 3D tensor field $\mathcal{D} : \Omega \to \mathbb{R}^{3\times3}_{\mathsf{Sym}}$ defined on $\Omega \subset \mathbb{R}^3$, a DT-MRI streamsurface $\mathcal{S}$ seeded at point $\mathbf{p} \in \Omega$ is a two-manifold that is embedded in Ω, includes $\mathbf{p}$, and is locally tangent to the major and medium eigenvector fields of $\mathcal{D}$ at all points on $\mathcal{S}$.

An obvious problem with this definition arises at tensor degeneracies: When the smaller two eigenvalues are equal, $\lambda_2 = \lambda_3$, the corresponding eigenvectors can be rotated freely within a two-dimensional subspace. Thus, the tangential plane specified in Definition 1 is no longer unique. Since streamsurfaces were introduced to visualize regions of planarity, this problem can be avoided by restricting Ω to a subdomain on which diffusion tensor planarity, as defined by Westin et al. [100]

$$c_p := \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{9.1}$$

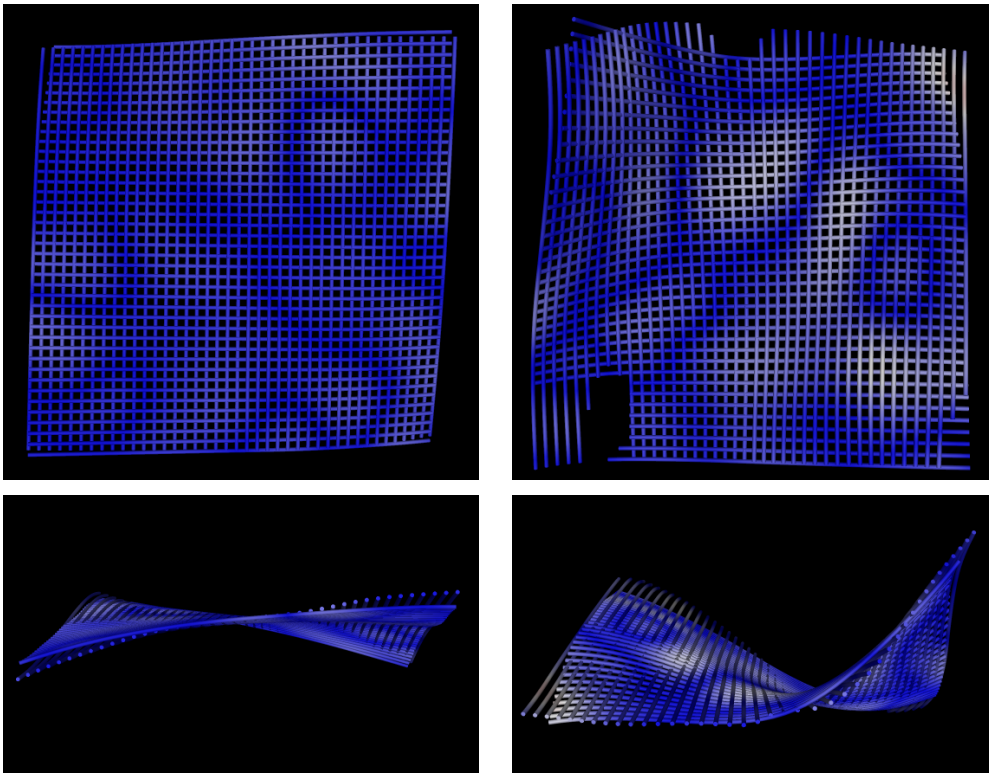
is greater than a suitable threshold.

However, even in regions of high c_p, the problem of surface integrability might remain. It refers to the fact that a given pair of vector fields may not permit a surface whose tangential plane includes both vectors everywhere. This is true even if the pairs of vectors are everywhere orthogonal, as in the case of eigenvector fields derived from the same symmetric tensor field.

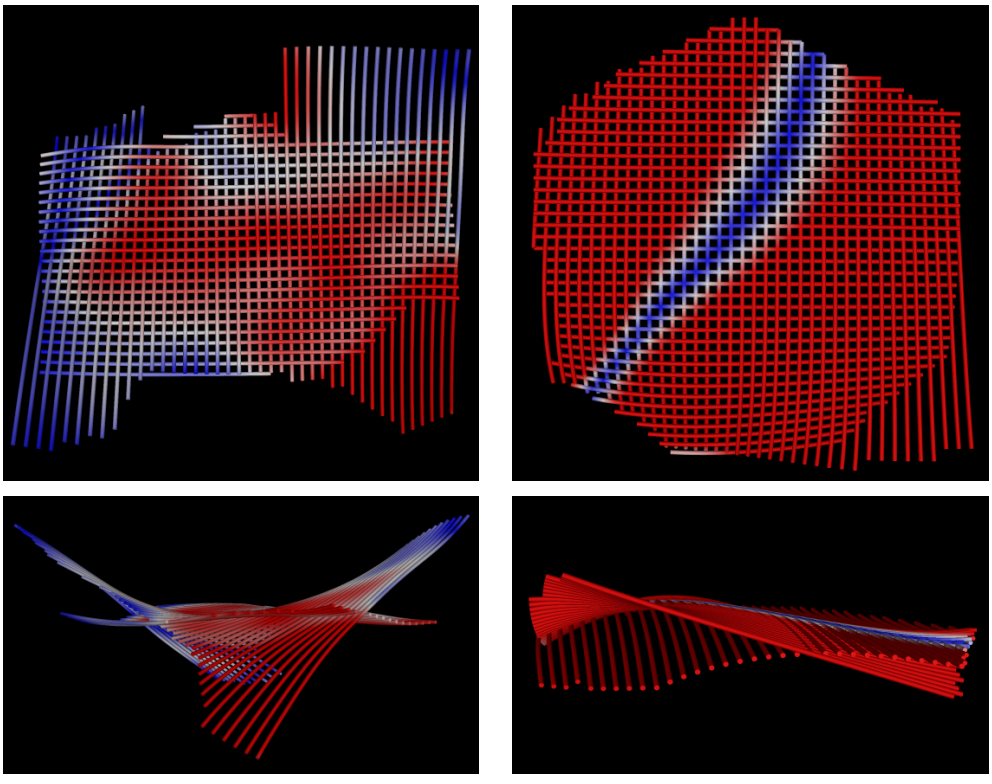
We present an initial experiment that introduces and illustrates this problem along with several implications that will be discussed more formally and in more detail later on. In particular, we illustrate the problem of surface integrability by integrating two sets of curves whose tangents remain within the plane spanned by the two eigenvectors, and which will form a grid on $1\,\mathsf{cm}^2$ large streamsurface patches if the underlying vector fields permit such a surface. Even though this exact patch size is arbitrary, we chose its order of magnitude in the spirit of a necessary condition for the practical utility of DT-MRI streamsurfaces based on typical voxel sizes in DT-MRI data ($\approx 2\,\mathsf{mm}$ in clinical data) and our opinion that a clear failure to integrate streamsurfaces across even just a few neighboring voxels should discourage their use.

The corresponding curves are shown in Figure 9.1, and are created by the following algorithm:

1. Determine the major and medium eigenvector directions $\mathbf{X}$ and $\mathbf{Y}$ at a given seed point $\mathbf{p}$.
2. Starting at $\mathbf{p}$, use the numerical method whose details are given in Step 4 to integrate seed curves that are as straight as possible, while remaining tangential to the plane spanned by the local major and medium eigenvectors. Extend these curves for $5\,\mathsf{mm}$ in steps of $0.25\,\mathsf{mm}$, starting in positive and negative $\mathbf{X}$ and $\mathbf{Y}$ directions.
3. At each vertex of each seed curve found in Step 2, determine a vector $\mathbf{v}$ that is within the plane spanned by the local major and medium eigenvectors, but orthogonal to the tangent of the curve. Use the numerical method from Step 4 to integrate curves that start at each vertex of the seed curves in positive and negative $\mathbf{v}$ direction, again for $5\,\mathsf{mm}$ and in steps of $0.25\,\mathsf{mm}$.



(a) CR (b) Sulcus



(c) CC x CST (d) Random Field

Figure 9.1.: The problem of streamsurface integrability can be illustrated by tracing curves that will form a grid if a surface can be integrated from the major and medium eigenvector fields, as it is approximately true in (a) and (b). Otherwise, horizontal and vertical curves will diverge, as it can be seen especially in the lateral views (bottom row) of (c) and (d). Colors indicate $|\zeta|$, our proposed measure of nonintegrability, from 0 (integrable, blue) over 0.1 (white) to 0.2 or greater (red).

4. Numerical integration is performed using second-order Runge-Kutta [60]. The tensor field is interpolated with a cubic B-spline kernel [54], and the propagation direction is given by the most recent tangent of the curve, projected to the plane spanned by the major and medium eigenvectors, and renormalized. To avoid artifacts in regions of singularities ($\lambda_2 = \lambda_3$), integration is terminated if $c_p < 0.2$.

Step 4 of this algorithm is analogous to the integration in standard DT-MRI streamsurface algorithms [103, 94]. Note that the resulting curves are *not* streamlines of the major and medium eigenvector fields. Rather, they are defined in such a way that they will intersect each other in a grid if major and medium eigenvector fields permit a streamsurface according to Definition 1. In Figure 9.1 (a) and (b), this is (approximately) the case. The bottom rows of the subfigures show the same patches as the top rows, but from the side, to make it easier to see whether the curves actually intersect, or bend away from each other, as it happens in Figure 9.1 (c) and (d).

In generic tensor fields, we cannot expect to find streamsurfaces according to Definition 1 since, in the absence of further constraints, it is extremely unlikely that a pair of eigenvector fields will permit a surface that is tangential to both eigenvectors at all points. This fact will be formally discussed in Section 9.3.4, and is illustrated in Figure 9.1 (d), which shows results from a random tensor field. It was generated by drawing eigenvalues in range $[0, 1]$ uniformly at random, and a random orthonormal set of eigenvectors, at each vertex of a provisional voxel grid. A final smooth tensor field was created by upsampling the result by factor 10 with a cubic B-spline kernel [54]. In this case, our set of curves clearly does not form a common surface, and the shown result is representative of different random seeds.

Despite this, at least some seed positions in a real-world DT-MRI dataset lead to (approximate) streamsurface patches, such as those shown in Figure 9.1 (a) and (b). Small deviations are still present, but should be seen in the context of Section 9.4.1, which will show that we have to admit small deviations between the tangent planes of $\mathcal{S}$ and the eigenvector fields to account for measurement noise even in cases where the underlying physical phenomenon (e.g., brain anatomy) might constrain the tensor field in such a way that streamsurfaces in the sense of Definition 1 emerge.

Subfigure (a) is from a crossing fiber region in the Corona Radiata (CR); Figure 9.9 presents a crossing fiber visualization in that same region. In this example, the fact that the curves approximate a grid might be attributed to the observation that they are relatively straight overall. The patch in Subfigure (b), which has been extracted from a seed point near a sulcus in the temporal lobe, exhibits more substantial curvature, which is reflected equally in both sets of curves, so that again only moderate deviations occur between them. However, even in DT-MRI data, emergence of streamsurface patches is not universal: The curves in Subfigure (c) are seeded in a region in which the Corpus Callosum (CC) and Corticospinal Tract (CST) cross. They very clearly bend away from each other, despite the fact that our algorithm stops integration when $c_p < 0.2$.

We emphasize that the diverging behavior in Subfigures 9.1 (c) and (d) is dominated by the fact that the underlying vector fields do not form a consistent surface, not by numerical errors. It cannot be fixed by using higher-order Runge-Kutta or smaller

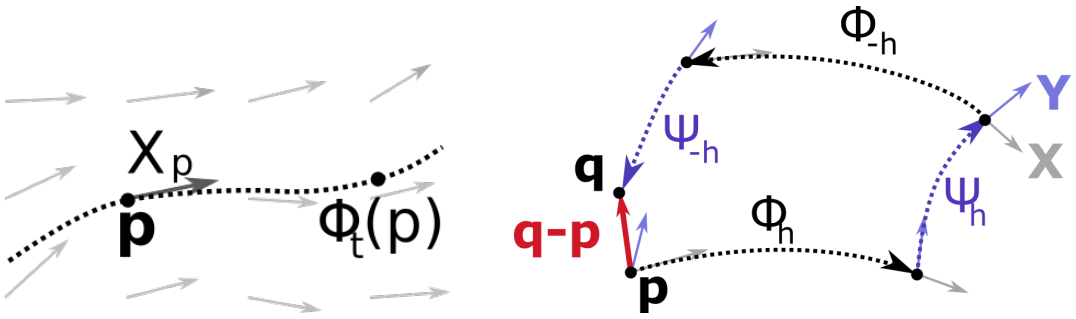


Figure 9.2.: *Left:* An integral curve in a vector field $\mathbf{X}$ (dashed) and the corresponding flow Φ. *Right:* A non-vanishing Lie bracket of two vector fields indicates that their flows do not commute, i.e., a loop of integral curves fails to close.

integration steps.

9.3.2. Introduction to the Lie Bracket

Our test for whether or not the major and medium eigenvector fields can locally be integrated into a streamsurface relies on concepts from differential geometry and topology, where the condition under which a surface can be constructed from any pair of differentiable vector fields $\mathbf{X}$ and $\mathbf{Y}$ has been established by Frobenius integrability theory [79]. It uses their Lie bracket $[\mathbf{X}, \mathbf{Y}]$, which is defined as

$$[\mathbf{X}, \mathbf{Y}] = \nabla_{\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}\mathbf{X} \tag{9.2}$$

$$= \sum_{i,j} (X^i \partial_i)(Y^j \mathbf{e}_j) - (Y^i \partial_i)(X^j \mathbf{e}_j) \tag{9.3}$$

$$= \sum_{i,j} (X^i \, \partial_i Y^j - Y^i \, \partial_i X^j)\mathbf{e}_j, \tag{9.4}$$

with canonical basis vectors $\mathbf{e}_j$.

An intuitive interpretation of the bracket can be given in terms of the flow map Φ, which is a well-established concept in the visualization literature [25, 66], and illustrated in Figure 9.2 (left): Given a vector field X, the corresponding flow Φ maps each point $\mathbf{p}$ to its integral curve $\Phi_t(\mathbf{p})$, which is everywhere tangential to $\mathbf{X}$ and starts at $\Phi_0(\mathbf{p}) = \mathbf{p}$.

Now, given two vector fields $\mathbf{X}$ and $\mathbf{Y}$ with corresponding flows Φ and Ψ, consider the attempt of forming a loop of integral curves, as indicated in Figure 9.2 (right): Starting at point $\mathbf{p}$, we first move along Φ_h for a given step size h, then follow Ψ_h from the new point, then backwards along Φ_{-h}, and finally backwards along Ψ_{-h}.

This will result in a closed loop if Φ and Ψ commute, which is indicated by a vanishing Lie bracket:

$$[\mathbf{X}, \mathbf{Y}] = 0 \tag{9.5}$$

In fact, the difference between starting point $\mathbf{p}$ and end point $\mathbf{q}$ can be approximated

based on the bracket as [79]:

$$\mathbf{q} - \mathbf{p} = h^2[\mathbf{X}, \mathbf{Y}] + O(h^3) \tag{9.6}$$

9.3.3. The Normal Part of the Lie Bracket as a Measure of Nonintegrability

Assuming that $\mathbf{X}, \mathbf{Y}$ are linearly independent, we want to know if a surface exists with $\Delta_p = \mathsf{span}(\mathbf{X_p}, \mathbf{Y_p})$ as its tangential plane at every point $\mathbf{p}$. According to Frobenius integrability theory [79, Chapter 6] such a surface exists iff

$$[\mathbf{X}, \mathbf{Y}] \in \mathsf{span}(\mathbf{X}, \mathbf{Y})\,. \tag{9.7}$$

This makes the normal component of the Lie bracket

$$[\mathbf{X}, \mathbf{Y}]^N = \left\langle [\mathbf{X}, \mathbf{Y}], \frac{\mathbf{X} \times \mathbf{Y}}{\|\mathbf{X} \times \mathbf{Y}\|} \right\rangle \tag{9.8}$$

a natural measure of nonintegrability, since $\mathbf{X}, \mathbf{Y}$ describe a surface iff $[\mathbf{X}, \mathbf{Y}]^N = 0$.

In the special case that $\mathbf{X} = \mathbf{v}_1, \mathbf{Y} = \mathbf{v}_2, \mathbf{Z} = \mathbf{v}_3$ are the three orthonormal eigenvector fields of a symmetric second order tensor field $\mathcal{D}$, we have

$$[\mathbf{X}, \mathbf{Y}]^N = \langle [\mathbf{X}, \mathbf{Y}], \mathbf{Z} \rangle =: \zeta\,. \tag{9.9}$$

In the remainder of this paper, we will use ζ as our measure of nonintegrability. An equation for its practical computation is derived in Appendix 9.7. We note that ζ only depends on the change of orientation of the spanned planes and is measured in units of length^{-1}. Equation (9.6) allows us to give an intuitive meaning to its value. For $\zeta = 0.1\,\mathrm{mm}^{-1}$, trying to follow a path around a square of area $1\,\mathrm{mm}^2$ leads $\approx 0.1\,\mathrm{mm}$ into the normal direction. Because Equation (9.6) is quadratic in h, an area of $1\,\mathrm{cm}^2$ already leads $\approx 1\,\mathrm{cm}$ outwards.

9.3.4. Interpreting the Sign of ζ

In this section, we argue that the sign of ζ is uniquely defined, and has a clear interpretation: According to Equation (9.9), it indicates whether the normal part of the Lie bracket is parallel or anti-parallel to the third eigenvector $\mathbf{Z}$. Since $\mathbf{Z}$ itself lacks an intrinsic orientation, i.e., $\mathbf{Z}$ and $-\mathbf{Z}$ equally satisfy the eigenvector equation $D\mathbf{Z} = \lambda_3\mathbf{Z}$, it might at first seem that the sign of ζ is arbitrary also. However, we will show that, as long as we require the three orthogonal eigenvector fields to form a right-handed system, inverting the sign of $\mathbf{Z}$ necessarily inverts the sign of the Lie bracket. Thus, the overall sign of ζ is preserved, since it reflects the *relative* orientation of these two vectors. We state this as

Theorem 1: Assuming that the eigenvector fields are defined so that they form a right-handed system, ζ has a well-defined sign, and is continuous almost everywhere.

The proof is based on the observation that, even though picking the eigenvector fields to be right-handed does not orient $\mathbf{Z}$, it makes it impossible to invert the sign of $\mathbf{Z}$ in

isolation, since that would make the system left-handed. Rather, we will have to invert the sign of exactly one other eigenvector field also.

Now, we observe that Equation (9.28) reformulates ζ as a sum of two products, each of which involves each eigenvector field, or its derivative, exactly once. This means that signs introduced when flipping pairs of eigenvectors will cancel out during computation of ζ.

Moreover, Appendix 9.7 computes ζ as a continuous function of the derivatives of a smooth tensor field, as well as its sorted eigenvalues and eigenvectors. Eigenvalues and eigenvectors are locally continuous functions of the tensor field almost everywhere, except at the singularities that arise when at least two eigenvalues are equal [44]. It is known that these singularities generically occur along curves in 3D symmetric tensor fields [105]. Away from these singularities, ζ is a continuous signed quantity. ■

We emphasize that our theorem does not follow directly from Theorem 1 in [55], which states that, on any simply-connected component on which tensors are sufficiently planar, $\mathbf{Z}$ can be converted into a continuous vector field. The key issue is that this conversion is not unique: It still leaves us with the choice of inverting the sign everywhere on the component. Our theorem states that this remaining ambiguity does not affect ζ. In Section 9.4.3, this allows us to compare ζ (as a signed quantity) across different subjects without having to worry about how to consistently orient eigenvectors in different datasets.

Note that the fact that the three eigenvector fields are everywhere orthogonal simplifies computation of ζ, but does not restrict its sign or value. This implies that, in a generic tensor field, we have to expect that ζ will fluctuate freely. This in turn indicates that generic tensor fields do not contain streamsurfaces in the sense of Definition 1, as it has been illustrated with the random tensor field in Section 9.3.1.

Theorem 1 also has an important consequence for correctly interpreting spatial maps of $|\zeta|$, which will be shown in Section 9.4.2: Even when the eigenvector fields are nonintegrable everywhere, we have to expect that zero isosurfaces will show up in such maps, since ζ is signed and continuous almost everywhere. These zero isosurfaces do *not* correspond to integrable surfaces in the eigenvector fields, since the eigenvectors will generally not be tangential to them. Rather, streamsurface integrability in a given region should be indicated by the fact that ζ is zero throughout that region, not just along isolated surfaces.

9.4. Systematic Study of Streamsurface Integrability

After introducing ζ as a nonintegrability measure for DT-MRI streamsurfaces, we will now use it to study the applicability of streamsurface visualizations in real-world data much more systematically than in previous works, which only highlighted individual cases where trying to integrate streamsurfaces seemed to work well [103], or where it resulted in visible artifacts [72].

Section 9.4.1 confirms the practical applicability of ζ via a quantitative evaluation in simulated data. We then study question **Q2** based on whole-brain maps of ζ in a

group of subjects. Section 9.4.2 shows qualitative results on a representative subject, Section 9.4.3 reports quantitative results from a statistical group analysis. Proceeding to question **Q3**, Section 9.4.4 clarifies how tensor planarity in the sense of Equation (9.1) relates to streamsurface integrability. Finally, Section 9.4.5 shows that our main findings are replicated on data from a different scanner and measurement setup.

9.4.1. Quantitative Evaluation on Simulated Data

We established that no streamsurfaces can be expected in generic tensor fields. However, it has been observed previously that constraints from the underlying neuroanatomy make diffusion tensor fields behave differently than purely random tensor fields [71]. This might explain why Figure 9.1 (a) and (b) still suggest that meaningful streamsurfaces exist in real-world DT-MRI data. However, even if we assume an exact value of $\zeta = 0$ based on the underlying structures, we have to expect that measurement noise will induce small deviations from zero. In order to establish a threshold $|\zeta| < \Theta$ that characterizes regions in which we consider nonintegrability to be low enough to still use streamsurfaces for visualization, we have to assess the expected effect of measurement noise on ζ.

To this end, we constructed perfectly planar tensors ($\lambda_1 = \lambda_2 \gg \lambda_3$) from two orthonormal sets of three eigenvector fields $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, so that the tangential planes of the intended streamsurfaces would be spanned by $\mathbf{X}$ and $\mathbf{Y}$. Based on these tensors and measurement settings from a clinical protocol, we simulated DW-MRI data sets (30 DWIs). After adding varying levels of noise, diffusion tensors were fitted, and we compared the values of ζ computed from the equations in Appendix 9.7 to their analytical ground truth.

Integrable Case: Concentric Spheres

When choosing the normal field to be the normalized radial field

$$\mathbf{Z}(\mathbf{p}) = \frac{\mathbf{p}}{\|\mathbf{p}\|}\,, \tag{9.10}$$

the $\mathbf{X}$-$\mathbf{Y}$-planes are tangential to concentric spheres around the origin. This configuration is obviously integrable and we expect $\zeta \equiv 0$. Figure 9.3 shows average empirical values of $|\zeta|$ as a function of the radius (left), as well as, for a fixed radius, mean and standard deviation of ζ for different levels of added noise (right).

Based on these plots, computed values of ζ appear unbiased (i.e., close to the ground truth value of zero on average). Unsurprisingly, variance increases with the level of measurement noise in the data. For small radii, we observe errors in the computed values of ζ even in the noise-free case. This can be explained by the onset of discretization artifacts as the streamsurface curvature increases.

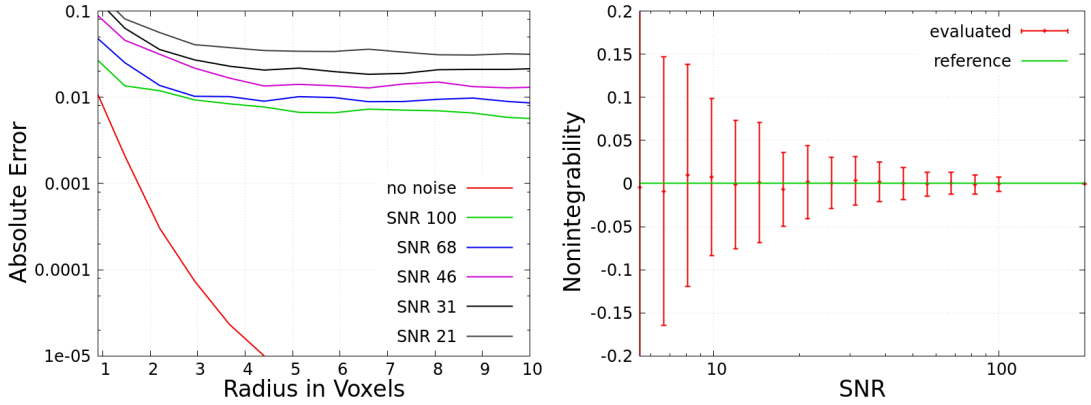


Figure 9.3.: Empirical values of ζ on the simulated "concentric spheres" data with different levels of noise. *Left*: The mean absolute error $|\zeta|$ plotted over the radius of a 3D Cartesian grid. *Right*: For each noise level, the mean and standard deviation of ζ at a fixed radius $r = 4.8$.

Nonintegrable Case: Linear Rotation Field

A nonintegrable configuration can be constructed by rotating the standard frame $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ by an angle $\theta(x) = \theta' x$ around the x-axis:

$$\mathbf{X} = \mathbf{e}_x \tag{9.11}$$
$$\mathbf{Y} = \cos\theta\,\mathbf{e}_y + \sin\theta\,\mathbf{e}_z \tag{9.12}$$
$$\mathbf{Z} = \cos\theta\,\mathbf{e}_z - \sin\theta\,\mathbf{e}_y\,. \tag{9.13}$$

In this case, the analytical ground truth is $\zeta = \theta'$, as derived in Appendix 9.8. The results of estimating ζ from fields with varying rotation speed θ' and different levels of added noise can be seen in Figure 9.4. Again, empirical values are reasonably centered on the ground truth, with a variance that depends on the level of measurement noise, as well as on errors from spatial discretization, which increase with rotation speed θ'.

Given the observed standard deviations at realistic noise levels (e.g., SNR=20), we can expect that empirical values of ζ will allow us to reliably distinguish between the integrable case shown at the bottom of Figure 9.3, and the nonintegrable one at the bottom of Figure 9.4. However, Figure 9.3 also illustrates that, even in cases where neuroanatomy would induce streamsurfaces in high-quality data, high levels of measurement noise might make it impossible to observe them.

In the remainder of our paper, we will use $|\zeta| < 0.1$ as a threshold for streamsurface integrability. This is motivated in two ways: First, it is clear from Figure 9.3 that, even if the anatomical structures that are reflected in a diffusion tensor field induce integrable surfaces, we have to expect deviations on this order of magnitude because of the measurement noise. Second, we visually explored a large number of patches with the algorithm from Section 9.3.1 and used different values of ζ as an additional termination

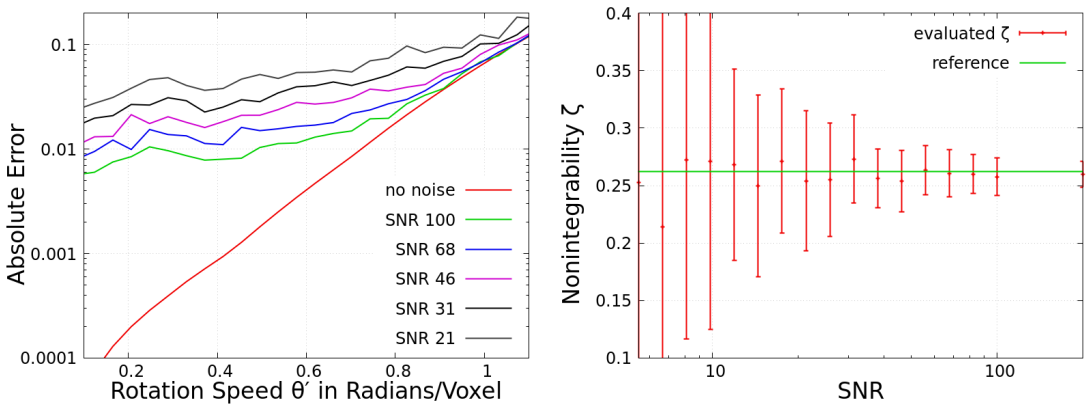


Figure 9.4.: Empirical values of ζ on the simulated "linear rotation field" data. *Left*: The average absolute deviation from the ground truth value ($|\zeta - \theta'|$) as a function of the rotation speed θ'. *Right*: For each noise level, the mean and standard deviation of ζ at a fixed speed $\theta' = 0.26$, with the ground truth value $\zeta = \theta'$ indicated in green.

criterion. We found that $|\zeta| < 0.1$ limited deviations on a $1\,\mathrm{cm}^2$ patch to a level where the curves still formed visually plausible surfaces.

Based on its definition, ζ is measured in units of $[\mathrm{length}]^{-1}$, so thresholds on $|\zeta|$ should only depend on the level of noise in the data, not on the spatial resolution of the sampling grid, as long as derivatives are computed with respect to the proper physical units of length.

9.4.2. Spatial Extent of Nonintegrability

To investigate where in the brain integrable or non-integrable behavior can be found, we color mapped ζ on whole-brain DT-MRI datasets from 15 healthy subjects, which were taken from the publicly available Human Connectome Project (HCP) database [78]. The diffusion MRI acquisition scheme of the HCP includes 90 diffusion-weighted images each at b values of $b \approx \{1000, 2000, 3000\}\,\mathrm{s/mm}^2$, as well as 18 images with minimal diffusion weighting, $b \approx 5$. The spatial resolution is $(1.25\,\mathrm{mm})^3$ isotropic. Since the diffusion tensor model [7] has been designed for b values up to $b \approx 1000$, we only used the corresponding part of the data to estimate the tensors.

ζ is a nonlinear function of the tensor field and its derivatives. Similar to ridges and valleys in tensor invariants [45, 91], it contains much higher spatial frequencies than the underlying tensor field itself. For this reason, we sampled it on a grid that, along each axis, was four times as fine as the original voxel grid. As before, upsampling used cubic B-splines.

Figure 9.5 presents results from a representative subject. To provide anatomical context, the left column shows a standard XYZ-RGB color coding of the principal eigenvector direction, indicating left-right as red, anterior-posterior as green, and superior-inferior

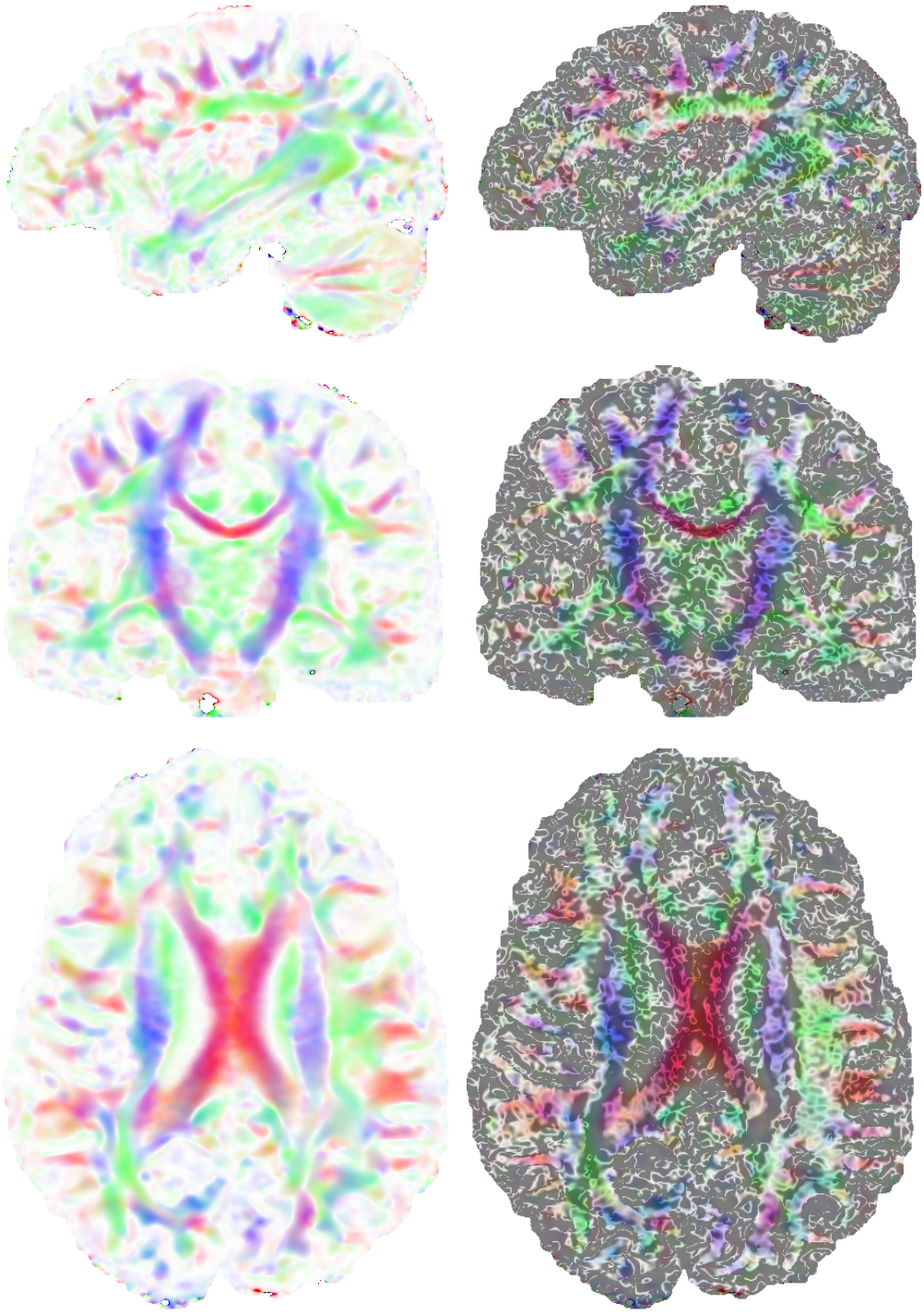


Figure 9.5.: In the left column, standard XYZ-RGB color maps of the principal eigenvector are shown for three orthogonal slices of a DT-MRI dataset. In the right column, the absolute value $|\zeta|$ of our nonintegrability measure is overlaid in black (transparent if $|\zeta| = 0$, $\alpha = 0.5$ for $|\zeta| \geq 0.1$). Even though $|\zeta|$ masks out several regions where we cannot expect to integrate meaningful streamsurfaces, substantial regions remain in which streamsurfaces can be used for visualization.

as blue. In the right column, $|\zeta|$ has been overlaid in black, with linearly increasing opacity so that $|\zeta| = 0$ is shown transparent, $|\zeta| \geq 0.1$ with opacity $\alpha = 0.5$. The resulting maps show several extended regions where $|\zeta|$ is consistently low and streamsurface integration may reasonably be attempted. However, there are also many regions that indicate high values of $|\zeta|$ or exhibit a fine-grained texture, indicating high amplitude oscillations around zero. As explained in Section 9.3.4, this indicates that trying to integrate streamsurfaces in those regions would lead to strong artifacts, as illustrated in Figure 9.1 (c) and (d).

9.4.3. Between-Subject Variability

When visually comparing all 15 subjects, we observed similar overall patterns in maps of ζ. In particular, we noticed extended anatomical regions in which $|\zeta|$ appeared to be high in all subjects. This motivated a more formal statistical analysis, whose results are presented in Figure 9.6.

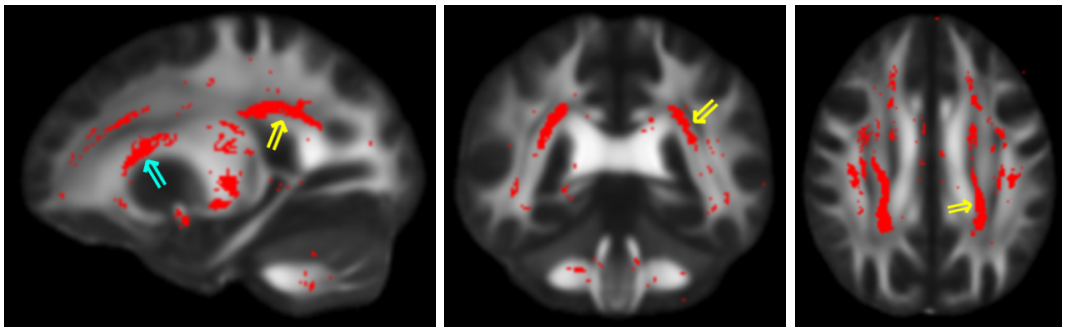
We brought the maps from different subjects into anatomical correspondence using the established and publicly available nonlinear registration algorithm FNIRT, which is part of FSL [34]. In particular, we registered brain extracted Fractional Anisotropy (FA) maps from all subjects to the FMRIB58_FA template provided with FSL, and applied the same transformations to the maps of ζ, which warps them to a common reference space with 1 mm isotropic voxel size.

In this common space, we performed a voxel-wise one-sample t-test of the null hypothesis that the local mean of ζ across subjects is zero. This null hypothesis is justified by observing that ζ is intrinsically signed (cf. Theorem 1) and there is no reason why noise should perturb it more frequently or more strongly in one direction than the other. If the mean across subjects deviates from zero despite this, the t-test judges the significance of this deviation by relating it to the number of included subjects, and the observed variability between them. We accounted for the multiple comparisons (one per voxel) by imposing a false discovery rate (FDR) of 5% [10].

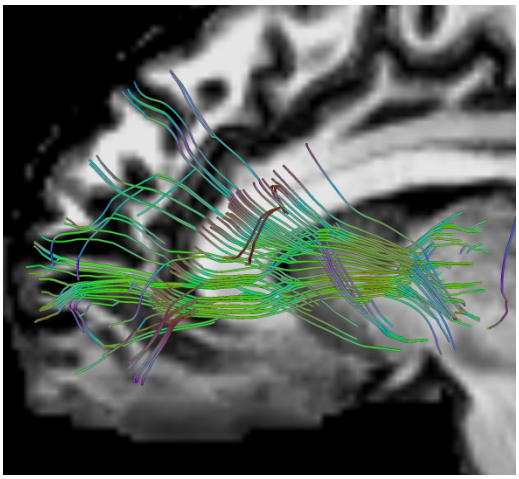
Red regions in Figure 9.6 (a) indicate locations where the null hypothesis was rejected, overlaid on the FA template. We identified several spatial clusters, which arose symmetrically in both hemispheres. Two of the clusters are marked with arrows, an anterior one in light blue, a posterior one in yellow. To identify the tracts that these clusters belong to, we warped the clusters back into the image space of one particular subject, and used them as seeds for fiber tractography.

Since a large fraction of white matter voxels contain more than a single dominant fiber orientation, and diffusion tensor tractography is known to be unable to accommodate this fact [90], tractography is based on a state-of-the-art deconvolution approach that makes use of all available diffusion measurements in the HCP data [2]. Results are shown in Subfigures (b) and (c). They suggest that the anterior region is part of the anterior thalamic radiation; in the posterior region, fibers from the corpus callosum and the corticospinal tract mix.

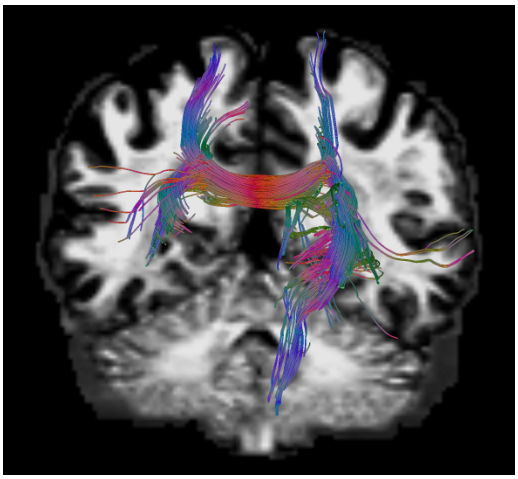
Concerning **Q2**, we conclude from this experiment that the previously observed failure case of DT-MRI streamsurface integration [72] is *not* a rare and isolated problem.



(a) Results of the statistical hypothesis test



(b) Tracts, light blue region



(c) Tracts, yellow region

Figure 9.6.: We coregistered maps of ζ from 15 healthy subjects and applied a one-sample t test to identify regions in which statistically significant nonintegrability occured in this population. Results with false discovery rate 5% are shown in red in (a). In one of the subjects, we performed tractography starting in the regions that are highlighted with light blue and yellow arrows. Results are shown in (b) and (c), respectively.

Rather, our results clearly identify anatomical regions in which streamsurface integrability is violated in a statistically significant sense, in a whole group of subjects.

9.4.4. Correlation with Diffusion Tensor Planarity

We now turn to question **Q3**. Given that streamsurfaces have been proposed for regions of sufficiently large diffusion tensor planarity c_p, we should clarify whether regions of nonintegrability, as they were shown in the previous subsections, may be restricted to regions of low c_p, which would make the issue less relevant for visualization.

Figure 9.7 visualizes the relationship between c_p and $|\zeta|$ with a bivariate color map that is adapted to the thresholds used throughout our paper. Fully saturated green indicates $c_p \geq 0.2$, i.e., regions where we consider planarity large enough for streamsurface integration. Fully saturated magenta stands for $|\zeta| \geq 0.1$, i.e., regions where nonintegrability is so large that no meaningful streamsurfaces exist. In fact, green and magenta dominate the maps, indicating that nonintegrability mostly affects regions of low tensor planarity.

We also confirmed this quantitatively, based on the joint histogram in the bottom right corner of Figure 9.7. It shows values from the full volume, and the full range of c_p and $|\zeta|$, whose infinite range has been compressed using the $\arctan$. The Pearson correlation between $\arctan|\zeta|$ and c_p is $\rho = -0.41$.

However, despite this negative overall correlation, there are non-negligible regions that show up in black in the slice images, indicating significant nonintegrability even though c_p is above our threshold. In these regions, we may be tempted to integrate streamsurfaces even though they are not meaningful. This settles **Q3**: The integrability condition is *not* reliably met in regions of planarity. We believe that this insight needs to be accounted for in streamsurface visualizations, and will suggest specific consequences in Section 9.5.

9.4.5. Impact of Data Acquisition Protocol

All experiments above used data from the Human Connectome Project, which has higher quality than the data on which DT-MRI streamsurfaces have been reported in the literature previously. In particular, it has been acquired with a highly modified Siemens 3T Skyra scanner that was specifically optimized for diffusion MRI, and provides higher spatial and angular resolution than typical clinical data [78]. We verified that our main findings still hold when using data that has been acquired on a widely used clinical MR scanner (3T Philips Achieva) with a standard DT-MRI protocol: 32 diffusion-weighted directions with $b = 1000$, one reference scan with $b = 0$, spatial resolution $(2\,\mathrm{mm})^3$ isotropic.

Results on this clinical data are shown in Figure 9.8. Structures in the brain map (left) have a coarser appearance than in Figure 9.7, reflecting the reduced spatial resolution. Despite this, our main findings carry over: Beside regions that are either non-planar and nonintegrable (magenta) or planar and integrable (green), we still observe non-negligible regions in which streamsurfaces are ill-defined despite high c_p (black). The

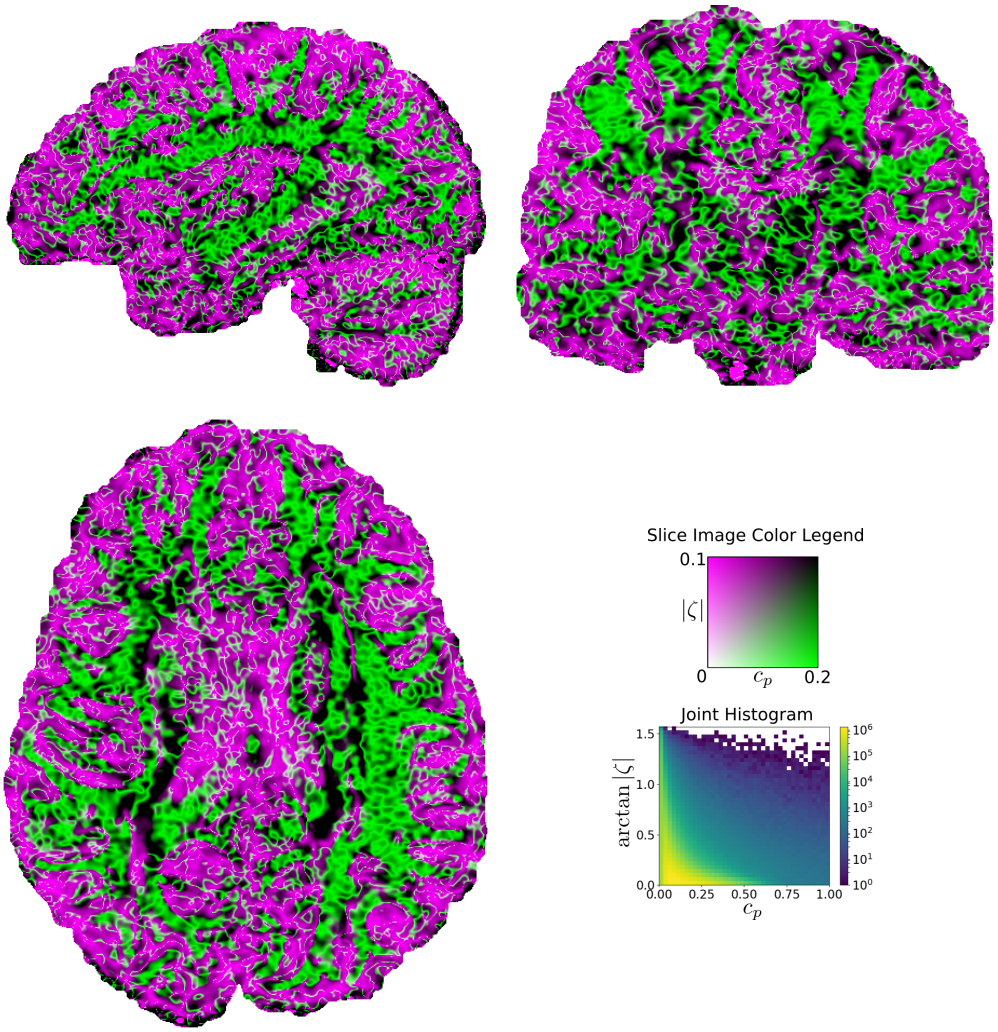


Figure 9.7.: Comparing nonintegrability $|\zeta|$ and diffusion tensor planarity c_p with a bivariate color map reveals that most regions exhibit either high planarity (green) or nonintegrability (magenta). However, there clearly exist regions, shown in black, where the major and medium eigenvector fields are nonintegrable despite high c_p. A joint histogram of both measures, computed over the full volume, confirms these observations.

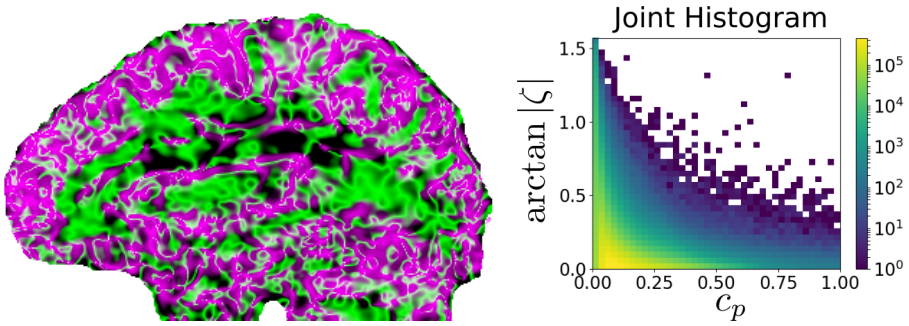


Figure 9.8.: Nonintegrability ζ can also be computed from data that has been acquired with a standard clinical DT-MRI protocol. Also in this case, bivariate color mapping and a joint histogram clearly indicate regions of nonintegrability despite high c_p, shown in black in the sagittal map on the left (please refer to Figure 9.7 for a full legend)

joint distribution of $\arctan|\zeta|$ and c_p (right) exhibits a similar pattern as in Figure 9.7.

9.5. Consequences for DT-MRI Visualization

Adressing **Q4**, we will now propose a novel, patch-based approach to DT-MRI streamsurface visualization. However, we believe that another issue is at least as important as algorithmic aspects, and should be discussed first: When and how should DT-MRI streamsurfaces be used, and how might they be interpreted?

9.5.1. Streamsurfaces in Visualization: When, How, and Why

Three important insights from our analysis can be summarized as:

I1 There are brain regions in which the extent of nonintegrability is small enough that it might be attributed to the effects of measurement noise. In these regions, major and medium eigenvector fields define small streamsurface patches with a visually satisfactory accuracy (cf. Sections 9.3.1, 9.4.1 and 9.4.3).

I2 There are other, non-negligible regions where, despite high values of diffusion planarity c_p, the major and medium eigenvector fields do not define meaningful streamsurfaces, not even on relatively small ($1\,\mathrm{cm}^2$) patches (cf. Sections 9.3.1 and 9.4.4).

I3 Even small amounts of nonintegrability will accumulate rapidly over larger distances (cf. quadratic dependence in Equation (9.6)).

From these, we draw three main conclusions on when and how to use streamsurfaces for DT-MRI visualization:

C1 When they are used to visualize small patches in certain brain regions, we consider streamsurfaces to be a meaningful visualization technique despite the failure case reported in [72].

C2 Algorithms for integrating streamsurfaces should not only check that c_p is sufficiently high, as it has been done previously [103, 94], but also that nonintegrability $|\zeta|$ is sufficiently low. Appendix 9.7 can be used for practical computation of ζ, and Section 9.4.1 as well as Figure 9.1 support its applicability.

C3 When interpreting visualizations of larger streamsurfaces, users have to account for the potential consequences of accumulating nonintegrability.

We emphasize that **C3** is more fundamental than the problem of accumulating integration errors that also affect the widely used streamline visualizations [42]. At least, we know that white matter contains curve-like anatomical structures (fiber tracts), and the question is merely whether streamlines correctly follow them. When showing streamsurfaces, *we are not even certain that surface-like anatomical structures exist that they could correspond to.*

What might be a valid interpretation of DT-MRI streamsurfaces in terms of brain anatomy? Zhang et al. point out that planar diffusion anisotropy "could result from a surface structure, a boundary between different materials, or a crossing of multiple linear features" [103]. This is true, but applies to the microscale, which is probed by the diffusion. It does not necessarily imply that surface structures exist at the macroscopic scale, at which DT-MRI streamsurfaces are extracted. Considering this, and the nonintegrability of generic tensor fields (Section 9.3.4), Zhang et al.'s observation that they could construct streamsurfaces without visually obvious errors is quite remarkable, and could have deeper implications than they might have realized.

We believe that there is a close link between this early observation in the visualization literature and the sheet structure hypothesis, which has started to attract significant interest outside of visualization about a decade later [97, 14, 95, 84, 83, 24, 6]: It claims that fiber tracts cross in such a way that they form surfaces. We note that the original publication by Wedeen et al. [97] additionally called the crossings "near-orthogonal", which has been challenged in a technical comment by Catani et al. [14]. In their response, Wedeen et al. [95] emphasize the aspect of integrability, and the fact that it is independent from orthogonality. Due to the well-known limitations of the DT-MRI model, it cannot be used to conclude anything about crossing angles; our discussion will therefore focus entirely on the aspect of integrability.

In examples such as Figure 9.9, we observed that, when seeding crossing fiber tractography [2] on streamsurfaces, the resulting streamlines tend to remain close to the surface. We emphasize that the streamlines in Figure 9.9 are not derived from the diffusion tensor, and are not constrained to cross at any particular angle. The fact that they approximately form a surface locally supports the sheet structure hypothesis. The fact that this surface can be approximated as a DT-MRI streamsurface supports our own conjecture, which states that *there are regions in the brain where DT-MRI streamsurfaces represent the sheet-like crossings postulated by Wedeen et al.* This assumes a close alignment between the planes in which the two dominant fiber tracts cross (possibly

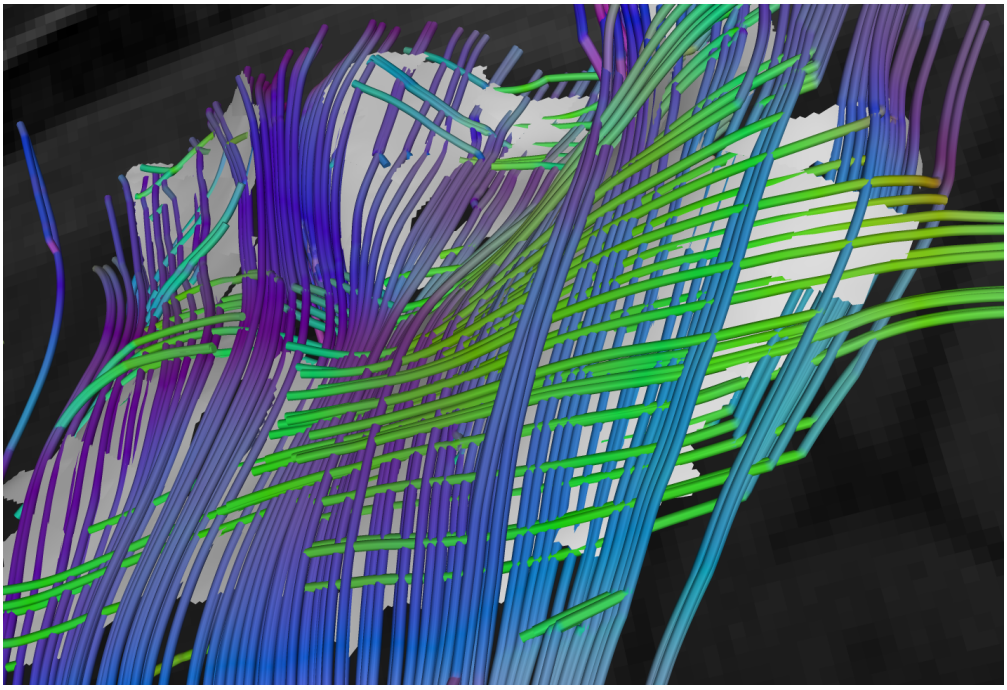


Figure 9.9.: We observed that seeding crossing fiber tractography on DT-MRI streamsurfaces often results in streamlines that remain close to the surface, which is shown here in gray. This specific example shows part of the corona radiata.

non-orthogonally) and the planes spanned by the major and medium DT-MRI eigenvectors. In a recent paper [6], we provide quantitative evidence for such an alignment. We believe that, even though the results in [6] and Figure 9.9 leave room for further investigation, our conjecture is relevant to visualization because it suggests a potential anatomical interpretation of DT-MRI streamsurfaces.

9.5.2. Patch-Based Rendering of Streamsurfaces

The most recent algorithm for streamsurface extraction is the one by Vilanova et al. [94]. It assumes a fixed topology (disk with holes), with a prespecified regular connectivity of potential vertices whose positions are found by integrating them in an order that forms a spiral starting at the seed, as shown in Figure 9.10.

Figure 9.11 (a) shows a result from this algorithm, with a red sphere indicating the seed point. Nonintegrability of the involved eigenvector fields causes vertices that are connected by the prespecified topology to end up far away from each other, which introduces spurious connections (marked in cyan). Moreover, cusps in the surface (marked with arrows) arise where the positions of adjacent vertices have been determined along different integration paths.

Following **C2**, we modified this algorithm to use $|\zeta| \geq 0.1$ as an additional termination criterion. The results are shown in Figure 9.11 (b). As expected, the new termination

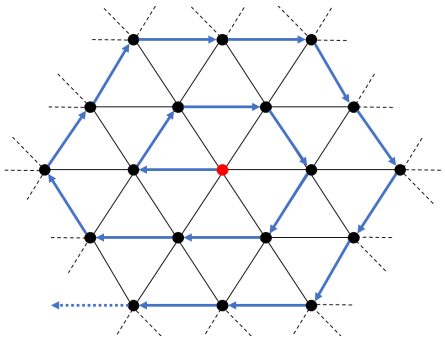


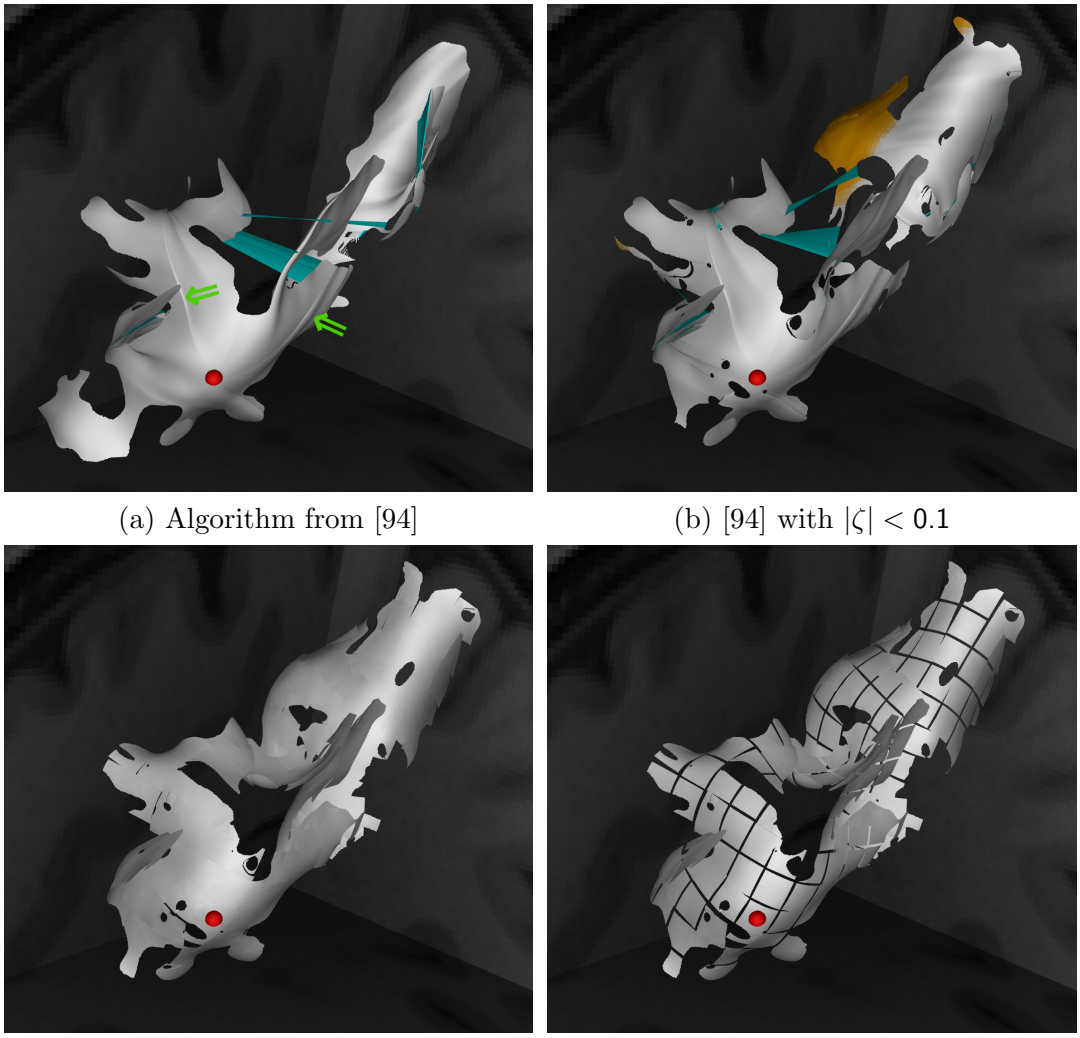
Figure 9.10.: A fixed streamsurface topology is assumed by the algorithm in [94]. The red vertex indicates the seed, blue arrows the order in which vertices are placed. Neighboring vertices might take very different integration paths, which leads to the artifacts illustrated in Figure 9.11.

criterion eliminated parts of the surface. Somewhat paradoxically, it also added new parts elsewhere, highlighted in (b) by coloring vertices orange if they are far away from the surface in (a). This is a side-effect of the spurious connections: Since the algorithm assumes that the respective vertices are close to each other, it does not attempt to grow the surface any further in these locations. Thus, the additional termination criterion may free up some of the available vertices to sample parts of the streamsurface that were previously not accessible to this algorithm. Cusps and some spurious connections remain in (b), which is unsurprising given the quickly accumulating effects of nonintegrability (**I3**). Therefore, it is necessary, but not sufficient to use a termination criterion based on $|\zeta|$.

As a consequence of **C1**, which asserts the value of small streamsurface patches as a visualization technique, and **C3**, which casts doubt on the interpretability of larger streamsurfaces, we propose a novel, patch-based approach to streamsurface visualization, whose results are shown in Figure 9.11 (c) and (d). It is based on rendering a set of small ($4\,\mathrm{mm} \times 4\,\mathrm{mm}$) streamsurface patches. They can be extracted with the modified algorithm from (b), which works well when constrained to sufficiently small surface patches in regions of sufficiently low $|\zeta|$.

Since manually seeding a large number of patches would be tedious and unlikely to yield a visually informative overall arrangement, our method automatically places additional seeds for neighboring patches based on an original one, which can be specified by the user. This process uses the same grid-like integration that was described in detail in Section 9.3.1. From each original and new seed point, curves are extended in all four directions. Integration terminates when $c_p < 0.2$, $|\zeta| \geq 0.1$, or at the boundary of the domain. The endpoints of these curves are only added to the seed pool if they are sufficiently far away from all existing seeds, which guarantees that this procedure terminates even for complex overall surfaces, including ones that might loop back on themselves.

In Figure 9.11 (c), patch sizes are slightly larger than the maximum integration distance used for seed placement, which makes them overlap and, when rendered together,



(a) Algorithm from [94] (b) [94] with $|\zeta| < 0.1$

(c) Patch-based with overlaps (d) Patch-based with gaps

Figure 9.11.: When using established algorithms for streamsurface integration (a), non-integrability of the eigenvector fields can lead to visual artifacts such as spurious connections (cyan) and cusps (arrows). Since these effects accumulate over larger distances, they are not fully avoided by using non-integrability $|\zeta|$ as a termination criterion (b). We therefore propose to render an automatically seeded set of small patches, which can either be made to overlap (c), or to leave gaps as a visual reminder that large-scale streamsurfaces are not well-defined mathematically, and their anatomical interpretation remains controversial (d).

gives the visual impression of a single surface. In regions of stronger nonintegrability, or at high zoom levels, small gaps between the patches can still be seen. They help avoid the accumulation of visual artifacts that otherwise occur when trying to integrate a not fully integrable pair of vector fields over larger distances.

Compared to the traditional approach [94], our patch-based technique removes the spurious connections and the related inability to fully sample the surface. It also greatly reduces surface cusps related to nonintegrability. A remaining limitation is the fact that the global shape of the (apparent) surface is still not well-defined, since our automated seeding scheme is still affected by accumulating nonintegrability, so that changing the integration order could alter the placement of patches. As a visual reminder of the challenges in interpreting larger streamsurfaces, which are intrinsic to the problem and not specific to any particular algorithm (**C3**), one might deliberately leave gaps between the patches by making them slightly smaller, as in Figure 9.11 (d).

In vector field visualization, Schulze et al. [76] have formally defined surfaces that are globally "as perpendicular as possible" to a given vector field. This could also be applied to DT-MRI streamsurfaces, by optimizing surfaces so that they become as perpendicular as possible to the minor eigenvector field. This would reduce the visual effects of accumulating nonintegrability. However, as a tool that we hope will contribute to the ongoing scientific debate about whether or not fiber bundles form surface-like anatomical structures (Section 9.5.1), we prefer a visualization that is "visually honest" in the sense that it will not hide the effects of nonintegrability if it is present in the data.

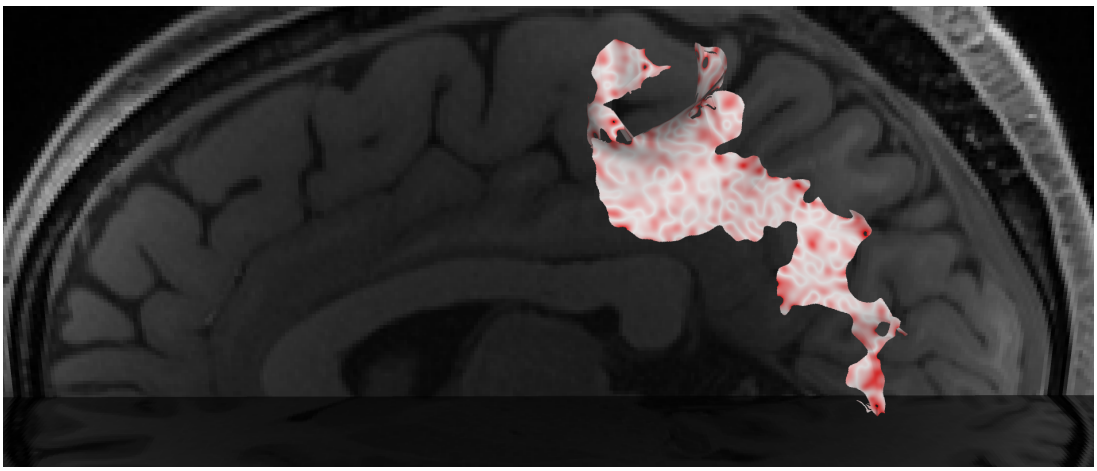
Finally, Figure 9.12 (b) presents a particularly large patch-based streamsurface that our proposed algorithm has grown from a single seed point, and that we have not been able to reproduce using the traditional algorithm (a). It extends over the whole length of the left hemisphere, illustrating both the ability of our approach to visualize structures at this scale and the striking fact that the brain indeed contains such large and connected regions in which diffusion tensor eigenvectors remain locally near-integrable.

9.6. Conclusion

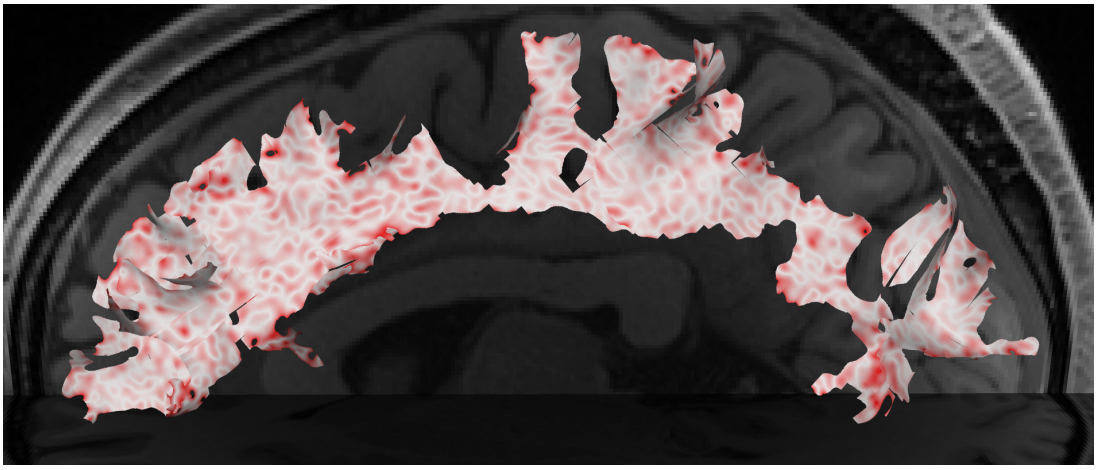
Streamsurfaces have been introduced early on to DT-MRI visualization, but are known and used less widely than streamlines or glyphs. Discovering a potential link to the recently proposed hypothesis that crossing fiber tracts in humans and nonhuman primate species form sheet-like structures [97] motivated us to revisit them as a visualization technique.

We focused on the issue of streamsurface integrability, which we studied with a suitable combination of mathematical reasoning, simulations, and an empirical investigation into an ensemble of real-world datasets, using both visual and statistical tools. This has led us to new conceptual insights and conclusions concerning the role of streamsurfaces in DT-MRI visualization (Section 9.5.1), as well as to a novel approach for streamsurface extraction (Section 9.5.2).

Our work proposes a measure ζ that quantifies the expected error up to which the major and medium eigenvector fields can locally be integrated into a surface. It can be



(a) [94] with $|\zeta| < 0.1$



(b) Patch-based with overlaps

Figure 9.12.: A large-scale patch-based streamsurface visualization (b), colored so that $|\zeta| = 0$ is white, $|\zeta| = 0.1$ fully saturated red. In comparison, the traditional algorithm (a) was unable to fully grow the surface.

used as a termination criterion in streamsurface integration, and it can be color coded to convey the local degree of nonintegrability. An important open issue concerns quantification of the *accumulated* uncertainty on streamsurfaces. For streamlines, there exists a wealth of techniques for probabilistic tractography, along with corresponding visualization methods [74]. Probabilistic construction of streamsurfaces is left as a challenge for future work.

Acknowledgments

This work was supported by the DFG under grant SCHU 3040/1-1. Data were provided by Tobias Schmidt-Wilcke (Ruhr-University Bochum, Germany) and by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

9.7. Computing the Lie Bracket from Tensor Field Derivatives

To apply ζ in practice, we want to calculate it from the eigenvectors and the gradient of the tensor field $\mathcal{D}$. The local diffusion tensor D can be rotated into its eigenbasis $\mathbf{v}_i$:

$$RDR^T = \Sigma = \mathsf{diag}(\lambda_1, \lambda_2, \lambda_3) \tag{9.14}$$

with

$$R^T = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} & \mathbf{Z} \end{pmatrix} . \tag{9.15}$$

As building blocks for computing the Lie bracket, we define matrices

$$M_{i,jk} = \langle \partial_i \mathbf{v}_j, \mathbf{v}_k \rangle \tag{9.16}$$

which are antisymmetric for every i since

$$0 = \partial_i \underbrace{\langle \mathbf{v}_j, \mathbf{v}_k \rangle}_{0 \text{ or } 1} = \langle \partial_i \mathbf{v}_j, \mathbf{v}_k \rangle + \langle \mathbf{v}_j, \partial_i \mathbf{v}_k \rangle . \tag{9.17}$$

Thus, we can write them as

$$M_i = (\partial_i R) R^T = \begin{pmatrix} 0 & c_i & -b_i \\ -c_i & 0 & a_i \\ b_i & -a_i & 0 \end{pmatrix} \tag{9.18}$$

with

$$a_i = \langle \partial_i \mathbf{Y}, \mathbf{Z} \rangle \qquad b_i = \langle \partial_i \mathbf{Z}, \mathbf{X} \rangle \qquad c_i = \langle \partial_i \mathbf{X}, \mathbf{Y} \rangle \tag{9.19}$$

$$\mathbf{a} = \nabla \mathbf{Y} \cdot \mathbf{Z} \qquad \mathbf{b} = \nabla \mathbf{Z} \cdot \mathbf{X} \qquad \mathbf{c} = \nabla \mathbf{X} \cdot \mathbf{Y} . \tag{9.20}$$

The rotated tensor gradient then becomes

$$
\begin{aligned}
R(\partial_i D)R^T &= \partial_i(RDR^T) - (\partial_i R)DR^T - RD(\partial_i R^T) && (9.21)\\
&= \partial_i\Sigma - (\partial_i R)R^T\Sigma - \Sigma R(\partial_i R^T) && (9.22)\\
&= \partial_i\Sigma - (\partial_i R)R^T \cdot \Sigma + \Sigma \cdot (\partial_i R)R^T && (9.23)\\
&= \partial_i\Sigma - M_i\Sigma + \Sigma M_i && (9.24)\\
&= \begin{pmatrix} \partial_i\lambda_x & c_i(\lambda_1-\lambda_2) & b_i(\lambda_3-\lambda_1) \\ c_i(\lambda_1-\lambda_2) & \partial_i\lambda_2 & a_i(\lambda_2-\lambda_3) \\ b_i(\lambda_3-\lambda_1) & a_i(\lambda_2-\lambda_3) & \partial_i\lambda_3 \end{pmatrix}. && (9.25)
\end{aligned}
$$

This leads us to the desired closed-form equation for ζ:

$$
\begin{aligned}
\langle[\mathbf{X},\mathbf{Y}],\mathbf{Z}\rangle &= \langle\nabla_{\mathbf{X}}\mathbf{Y},\mathbf{Z}\rangle - \langle\nabla_{\mathbf{Y}}\mathbf{X},\mathbf{Z}\rangle && (9.26)\\
&= \langle\nabla_{\mathbf{X}}\mathbf{Y},\mathbf{Z}\rangle + \langle\mathbf{X},\nabla_{\mathbf{Y}}\mathbf{Z}\rangle && (9.27)\\
&= \mathbf{X}\cdot\nabla\mathbf{Y}\cdot\mathbf{Z} + \mathbf{Y}\cdot\nabla\mathbf{Z}\cdot\mathbf{X} && (9.28)\\
&= \langle\mathbf{X},\mathbf{a}\rangle + \langle\mathbf{Y},\mathbf{b}\rangle && (9.29)
\end{aligned}
$$

with

$$
a^i = \frac{1}{\lambda_2-\lambda_3}\left[R(\partial_i D)R^T\right]_{2,3} \qquad (9.30)
$$

$$
b^i = \frac{1}{\lambda_3-\lambda_1}\left[R(\partial_i D)R^T\right]_{1,3} \qquad (9.31)
$$

where square brackets $[\cdot]_{i,j}$ denote the (i,j)th component of a matrix.

9.8. Analytical Lie Bracket of Linear Rotation Field

To calculate ζ for the linear rotation field from Section 9.4.1, we first need to evaluate the Lie bracket. To do so, we note that the bracket can be written as $[\mathbf{v},\mathbf{w}] = \mathbf{vw} - \mathbf{wv}$ using the inconspicuous notation identifying a vector $\mathbf{v} = \sum_i v_i\,\mathbf{e}_i$ with the first order linear differential operator $\sum_i v_i \frac{\partial}{\partial x_i}$ acting on scalar functions $f(\mathbf{p})$. Making this definition

more explicit will prove insightful:

$$\begin{aligned}
[\mathbf{v}, \mathbf{w}]f = & \left(\sum_i v_i \frac{\partial}{\partial x_i}\right)\left(\sum_k w_k \frac{\partial}{\partial x_k}\right) f \\
& - \left(\sum_i w_i \frac{\partial}{\partial x_i}\right)\left(\sum_k v_k \frac{\partial}{\partial x_k}\right) f \\
= & \sum_{i,k} \left(v_i \frac{\partial w_k}{\partial x_i}\frac{\partial f}{\partial x_k} + v_i w_k \frac{\partial^2 f}{\partial x_i \partial x_k} \right. \\
& \left. - w_i \frac{\partial v_k}{\partial x_i}\frac{\partial f}{\partial x_k} - w_i v_k \frac{\partial^2 f}{\partial x_i \partial x_k} \right) \\
= & \sum_{i,k} \left(v_i \frac{\partial w_k}{\partial x_i} - w_i \frac{\partial v_k}{\partial x_i} \right) \frac{\partial f}{\partial x_k}
\end{aligned} \tag{9.32}$$

Surprisingly, even though $[\mathbf{v}, \mathbf{w}]$ was defined as a second order differential operator, both second order terms cancel out and only a first order operator remains. Since we identify first order operators with vectors, Equation (9.32) is equivalent to the definition that was previously given in Equation (9.2).

With this notation, we proceed to compute $[\mathbf{X}, \mathbf{Y}]$ for $\mathbf{X}, \mathbf{Y}$ as defined in Equations (9.11) and (9.12):

$$\begin{aligned}
[\mathbf{X}, \mathbf{Y}]f = & \mathbf{X}\mathbf{Y}f - \mathbf{Y}\mathbf{X}f \\
= & \frac{\partial}{\partial x}\left(\cos\theta \frac{\partial}{\partial y} f + \sin\theta \frac{\partial}{\partial z} f \right) \\
& - \cos\theta \frac{\partial}{\partial y}\frac{\partial}{\partial x} f - \sin\theta \frac{\partial}{\partial z}\frac{\partial}{\partial x} f \\
= & -\theta' \sin\theta \frac{\partial}{\partial y} f + \theta' \cos\theta \frac{\partial}{\partial z} f
\end{aligned} \tag{9.33}$$

Thus, the value of ζ is given as

$$\langle [\mathbf{X}, \mathbf{Y}], \mathbf{Z} \rangle = \theta' \sin^2\theta + \theta' \cos^2\theta = \theta' \tag{9.34}$$

10. Conclusion

10.1. Contributions

Kurtosis

The first achievement of our work on this topic was demonstrating that fiber crossings affect the diffusion and kurtosis parameters found by the DKI model. To circumvent this problem we introduced a new per-fiber model, generalizing the ball-and-stick model by adding an axially symmetric kurtosis tensor term. This proved to reduce the effects of crossing angles on scalar kurtosis measures for simulated fiber crossings.

Our model includes a dot compartment to represent isotropic portions of the signal. This term was chosen because it greatly outperformed other alternatives like the ball compartment in BIC model selection experiments.

On real data our model predicts fiber directions that agree with CSD. In contrast, DKI does not directly give any fiber directions. As for the kurtosis measures, $K_{\parallel}$ is strongly reduced in white matter compared to DKI, giving contrast between WM and GM in our model. Our model also reduces the effects of fiber crossings on FA and $K_{\perp}$, as well as reducing the correlation between these two measures. We hope this allows to infer more subtle tissue properties from these maps.

Deconvolution

We introduced a multi-tissue deconvolution method using the SHORE basis and the H-psd constraint, a new positivity constraint for fODF tensors. This method allows the use fourth order tensors, increasing numerical stability and reducing time consumption compared to the widely used order 8 approach while maintaining the high angular resolution, as demonstrated on simulated fiber crossings.

Using the SHORE basis instead of per-shell spherical harmonics not only reduces the complexity of computations but also makes the method applicable to DSI and other data sets that are not organized into shells. Experiments on clinical DSI data sets found more fiber directions than state-of-the-art deconvolution and cleaner results than standard DSI methods without deconvolution.

The method also showed benefits for tractography. We implemented a deterministic fiber tracking scheme with branching ability. In experiments this scheme was able to find a higher number of transcallosal fibers and lateral projections, comparable only to probabilistic approaches, but without the drawback of noisy or jiggly fibers.

Integrability

Our first contribution to this subject was the formulation and empirical proof of the hypothesis that the two dominant fiber lie in the same plane as major and medium eigenvectors of the diffusion tensor. This allowed us to define our DTI-SPI measure, a much faster and simpler alternative to SPI, which only depends on the diffusion tensor and does not require any fiber estimation. Its accuracy and reliability was demonstrated on synthetic data sets with varying noise levels.

Evaluating DTI-SPI on 15 subjects revealed persistently nonintegrable WM structures. Particular areas of low surface integrability were found where projection fibers and corpus callosum merge. Even experiments with lower quality clinical data recovered the same features.

On the visualization side DTI-based ζ was found to improve surface integration when used as an additional termination criterion. This led us to state general rules for surface integration (sec. 9.5.1) and the implementation of a new patch-based integration method.

10.2. Future work

Our symmetric kurtosis model left several questions unanswered. Decoupling the parameters of individual fibers seemed to break the optimizer's convergence in quick experiments. It should be tested if there is a deeper reason behind the divergences or if a better implementation could fix the problem. Also a closer examination of the dot compartment would be helpful.

Another problem that deserves further analysis is overfitting of diseased tissue with response functions from healthy patients. Here we expect better prevention with our new deconvolution method. Also, since our constraint does not generally imply decomposability in tensor orders higher than four, a challenge would be to find a tighter constraint for these cases. On the other hand the H-matrix already proved useful in higher orders for estimating the local number of fibers (see sec. C.1) and its ability to clean up fODFs needs more investigation. We also expect that both the kurtosis model and deconvolution might benefit from spatial regularization.

We also consider the DTI-based ζ to be a highly promising candidate as a new biomarker due to its reliability and low computational cost. Extensive evaluation on a large set of subjects will be required to map typical ranges of values for healthy subjects and correlate deviations to diseases.

Part III.

Appendix

A. Appendix - Spin

The following chapter is meant as additional backstory, giving more details about the behaviour of spin and quantum mechanics, while also illustrating the deeper connection between spin and spherical harmonics mentioned in section 2.3.

As already mentioned at the beginning of section 5.1, particles have a quantum mechanical property called *spin* $\mathbf{s}$ - a special type of angular momentum. Its absolute value

$$\|\mathbf{s}\| = \sqrt{j\,(j+1)}\hbar \tag{A.1}$$

can be expressed with a spin-number $j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \dots\}$ and is a constant (at least for single electrons). For each particle's type this value can be derived from its equation of motion. Electrons and protons have spin-number $j = \frac{1}{2}$.

Spin $\frac{1}{2}$

First a quick disclaimer about quantum mechanics[67] that needs to be recognized but not necessarily completely digested:

Definition A.1. *In **quantum mechanics** the current state of a system is abstractly described by a complex vector Ψ with $\|\Psi\| = 1$.*

*To get any specific information about the system (for example a particle's position or speed) one needs to perform a **measurement**. The rules are:*

- *For each type of measurement an associated Hermitian matrix M can be found.*
- *The result of the measurement can only be one of the eigenvalues $\{\lambda_i\}$ of M.*
- *Nature picks one of the eigenvalues λ_i randomly with probability $|\langle \Psi, \mathbf{m}_i \rangle|^2$ with the eigenvectors $\{\mathbf{m}_i\}$.*
- *Immediately after the measurement the system changes its state into the corresponding eigenvector $\Psi' = \mathbf{m}_i$.*

The last rule is the famous "observation changes a system" statement.

Remark. Even though results are random, probabilities for certain values can be higher than for others, if the state is already close to an eigenvector. And directly after a measurement, when the state is exactly equal to an eigenvector, its probability even becomes 1.

In the case of a spin-$\frac{1}{2}$ particle the possible spin/magnetization states can be described by two-dimensional complex vectors $\Psi \in \mathbb{C}^2$. The matrices associated to measuring the components of the magnetic moment are proportional to the Pauli matrices:

$$\mu_x = \frac{\gamma\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \mu_y = \frac{\gamma\hbar}{2}\begin{pmatrix}0 & i\\ -1 & 0\end{pmatrix}, \quad \mu_z = \frac{\gamma\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{A.2}$$

As dictated by the weird rules of def. A.1, measuring μ_z can only give one of two results $\mu_z = \pm\frac{1}{2}\gamma\hbar$ and move the system into the respective eigenvector $\Psi \propto \begin{pmatrix}1\\0\end{pmatrix}$ or $\begin{pmatrix}0\\1\end{pmatrix}$. Repeating this measurement would give the same result with absolute certainty and does not change the state. But performing either μ_x or μ_y afterwards would give completely random results and change the state again because the eigenvectors of the different matrices are not compatible.

For our MRI setup, one could argue that the component μ_z is being measured constantly because it interacts with the background $\mathbf{B}$ field, turning it into physical information of the particle, although "quantized" - meaning, it can only take on one of two values. On the other hand, the state space is not large enough to also carry information about the other components.

This interpretation only works for single particles. When measuring information about multiple particles simultaneously, the exact behaviour will be far more complex but can be approximately described by classical mechanics again when averaging over a large enough number of particles as stated by the Ehrenfest theorem[31].

General spin

Spin and angular momentum have their roots in rotational symmetry with a deep connection to representation theory: A physical system can be transformed by a rotation matrix $R \in \mathsf{SO}(3)$. No changes will be measurable if the system is rotationally symmetric. Let $\Psi \in \mathbb{C}^n$ then be the original state and Ψ' the state after rotation. The mapping $D(R)\Psi = \Psi'$ is assumed to be linear and should respect

$$D(R_1) \cdot D(R_2) = D(R_1 \cdot R_2)\,. \tag{A.3}$$

The theory of pure representation theory tells us, that the $n\times n$ matrices $D(R)$ are built up from a fundamental set of pre-defined matrices $\mathcal{D}^j(R)$ for $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$. These fundamental matrices $\mathcal{D}^j(R)$ have dimension $(2j+1)\times(2j+1)$ and can be found explicitly from properties of the group[1] $\mathsf{SO}(3)$ without any information about the physical system (see next section). A particular version of these matrices was constructed by Eugene Wigner[101] and is often called Wigner D-matrices.

[1] To be more precise, the group $SU(2)$ is used instead of $SO(3)$. Also see next section.

Another observation is required to define a particle's spin: Rotations can be created from "infinitesimal" rotations. I.e. there are three *generator* matrices

$$r_x = \frac{d}{d\varphi} R(\mathbf{e}_x, \varphi)_{|\varphi=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad r_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad r_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{A.4}$$

so that a general rotation of angle φ around axis $\mathbf{n}$ is given by

$$R(\mathbf{n}, \varphi) = e^{i\varphi\,(r_x n_x + r_y n_y + r_z n_z)} = e^{i\varphi\,\mathbf{r}\cdot\mathbf{n}}\,. \tag{A.5}$$

Also generators $L^j_{x,y,z}$ for the representations $\mathcal{D}^j$ can be found, meaning $\mathcal{D}^j(R(\mathbf{n}, \varphi)) = e^{i\varphi\,\mathbf{L}^j\cdot\mathbf{n}}$. The Pauli matrices in eq. (A.2) are the special case of $L^{j=\frac{1}{2}}$. Also, for any j the Ls fulfill

$$\left(\mathbf{L}^j\right)^2 = \left(L^j_x\right)^2 + \left(L^j_y\right)^2 + \left(L^j_z\right)^2 = j\,(j+1)\,\mathbf{1}\,, \tag{A.6}$$

which explains the relation eq. (A.1) between absolute spin value and spin-number j:

Applying additional physical knowledge, it turns out, the matrices $\hbar \cdot L^j_{x,y,z}$ are exactly the measuring matrices for angular momentum and spin. And as the total system rotation matrix $D(R)$ can be split into fundamental parts $\mathcal{D}^j$, the state vector Ψ can be split into parts that have constant total angular momentum $\sqrt{j\,(j+1)}\hbar$ because of eq. (A.6).

Again, the $2j+1$ eigenvalues of L^j_z can be measured, but the eigenvectors are not compatible with $L^j_{x,y}$ leading to random measuring results for the x- and y-components.

Spinors

The theory of spin and representations has another surprising link to the tensor algebra via the dark magic of spinors[13].

The components of a vector $\mathbf{v} = (x, y, z)^T$ can be encoded in the Hermitian matrix

$$V = \sum_i v_i\,\sigma_i = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \tag{A.7}$$

with the aforementioned Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & i \\ -1 & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A.8}$$

The length of $\mathbf{v}$ is then encoded in the determinant of V because

$$\det V = -z^2 - x^2 - y^2 = -\|\mathbf{v}\|^2\,. \tag{A.9}$$

With a similar procedure one can also find a 2×2 analogon of rotation matrices. For any $R = R(\mathbf{n}, \theta) \in SO(3)$ define the unitary matrix

$$\mathcal{R} = e^{i\theta/2 \sum_i n_i \sigma_i} = \sin\frac{\theta}{2}\,\mathbf{1} - i\cos\frac{\theta}{2}\,M(\mathbf{n})\,. \tag{A.10}$$

With this the matrix V' associated to the rotated vector $\mathbf{v}' = R\mathbf{v}$ obeys the transformation law

$$V' = \mathcal{R}^\star\,V\,\mathcal{R}\,. \tag{A.11}$$

Definition A.2. *The complex two-dimensional vectors* $\xi = \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} \in \mathbb{C}^2$ *that are transformed by* $\xi' = \mathcal{R}\xi$ *are called* ***spinors*** *or spin vectors.*

Using the previous vocabulary, spinors with the matrices $\mathcal{R}$ form the two dimensional representation of $\mathcal{R} = \mathcal{D}^{j=1/2}$. Also the two dimensional state vectors of spin-$\frac{1}{2}$ particles are spinors.

Remark. Eq. (A.10) defines a mapping from $SO(3)$ into $SU(2)$. Actually both $\pm\mathcal{R}$ correspond to the same R since both perform the same vector rotation according to eq. (A.11).

As with regular vectors, one can build tensors $\xi \otimes \rho$ from spinors. Symmetric spin tensors $\tau_{abc\ldots} = \xi_a \xi_b \xi_c \ldots$ are particularly simple, since products of n copies of ξ only have $n+1$ independent components $(\xi_1^n, \xi_1^{n-1}\xi_2, \ldots, \xi_2^n)$. These symmetric products form exactly the higher spin representations $\mathcal{D}^j$ for $2j = n$.

Spin tensors form another equivalent alternative to tensor algebra. Calculations done with regular tensors or spherical harmonics can also be done in the world of spin tensors. Spinors have the additional property, that a regular vector is represented as a second order tensor in the spinor world. This mysteriously turns spinors into the *square roots of vectors*.

Spherical harmonics

Spinors and spherical harmonics, besides both being connected to the tensor algebra, also share a more direct link as demonstrated by an elegant calculation in [11]:

The sum $\sum_{m=-l}^{l} \xi_1^{l-m}\xi_2^{l+m}\, \mathcal{Y}_l^m(\mathbf{u})$ for $l \in \mathbb{N}$ and a spinor ξ is rotationally invariant[2]. Since $\xi_1^{l-m}\xi_2^{l+m}$ are components of a symmetric spin tensor of order l which transforms with the unitary matrices $\mathcal{D}^l$, the spherical harmonics have to transform inversely:

$$\mathcal{Y}_l^m(R\mathbf{u}) = \sum_{m'} \overline{\mathcal{D}}^l_{m,m'}(R)\, \mathcal{Y}_l^{m'}(\mathbf{u}) \tag{A.12}$$

The spherical harmonics can even directly be found in the matrix elements of the $\mathcal{D}^j$:

$$\overline{\mathcal{D}}^l_{m,0}(\alpha, \beta, 0) = \sqrt{\frac{4\pi}{2l+1}}\, \mathcal{Y}_l^m(\beta, \alpha) \tag{A.13}$$

when using Euler angles α, β, γ to parametrize rotations R.

[2] The scaling of $\mathcal{Y}_l^m$ in [11] might differ from the convention in this thesis.

B. Appendix - Kurtosis

B.1. Cylindrically Symmetric Kurtosis Tensor

The general kurtosis tensor is

$$\begin{aligned}\boldsymbol{W}(\mathbf{v}) &= \sum_{i,k,l,m} W_{iklm}\, v_i v_k v_l v_m && \text{(B.1)}\\ &= W_{xxxx}x^4 + W_{yyyy}y^4 + W_{zzzz}z^4 \\ &\quad + 4\left(W_{xxxy}x^3y + W_{xyyy}xy^3 + W_{xxxz}x^3z + W_{xzzz}xz^3 + W_{yyyz}y^3z + W_{yzzz}yz^3\right) \\ &\quad + 6\left(W_{xxyy}x^2y^2 + W_{xxzz}x^2z^2 + W_{yyzz}y^2z^2\right) \\ &\quad + 12\left(W_{xxyz}x^2yz + W_{xyyz}xy^2z + W_{xyzz}xyz^2\right) && \text{(B.2)}\end{aligned}$$

Projection

To project $\boldsymbol{W}$ onto the subspace of tensors symmetric around the z-axis, we can average it over rotations around the z-axis:

$$\begin{aligned}2\pi\overline{\boldsymbol{W}}(\mathbf{v}) &= \int_0^{2\pi} \boldsymbol{W}(R_{z,\varphi}\mathbf{v})\, d\varphi && \text{(B.3)}\\ &= \int_0^{2\pi} \boldsymbol{W}(x\cos\varphi + y\sin\varphi, -x\sin\varphi + y\cos\varphi, z)\, d\varphi && \text{(B.4)}\\ &= \int_0^{2\pi} \left[(x\cos\varphi + y\sin\varphi)^4 W_{xxxx} + (-x\sin\varphi + y\cos\varphi)^4 W_{yyyy} + \ldots\right] d\varphi && \text{(B.5)}\\ &= \frac{3\pi}{4}(x^2+y^2)^2 W_{xxxx} + \frac{3\pi}{4}(x^2+y^2)^2 W_{yyyy} + 2\pi z^4 W_{zzzz} \\ &\quad + 4\,(0+0+0+0+0+0) \\ &\quad + 6\left[\frac{\pi}{4}(x^2+y^2)^2 W_{xxyy} + \pi(x^2+y^2)z^2 W_{xxzz} + \pi(x^2+y^2)z^2 W_{yyzz}\right] \\ &\quad + 12\,(0+0+0) && \text{(B.6)}\end{aligned}$$

resulting in

$$\begin{aligned}\overline{\boldsymbol{W}}(\mathbf{v}) &= (x^2+y^2)^2 \underbrace{\frac{3}{8}(W_{xxxx} + 2W_{xxyy} + W_{yyyy})}_{:=W_\perp} + z^4 \underbrace{W_{zzzz}}_{:=W_\parallel} && \text{(B.7)}\\ &\quad + (x^2+y^2)z^2 \underbrace{3(W_{xxzz} + W_{yyzz})}_{:=W_\diamond}\,. && \text{(B.8)}\end{aligned}$$

Components

Identifying this expression with the basis decomposition of the tensor $\overline{\boldsymbol{W}}$ gives us

$$\overline{W}_{xxxx} = \overline{W}_{yyyy} = W_\perp \tag{B.9}$$
$$3\overline{W}_{xxyy} = W_\perp \tag{B.10}$$
$$\overline{W}_{zzzz} = W_\parallel \tag{B.11}$$
$$6\overline{W}_{xxzz} = 6\overline{W}_{yyzz} = W_\diamond\,. \tag{B.12}$$

All the other components are zero.

Signal parameters

The signal of the kurtosis tensor model contains the term $\overline{D}^2\,\boldsymbol{W}(\mathbf{g})$. Choosing our signal parameters to be

$$\kappa_\parallel = \overline{D}^2\,W_\parallel, \quad \kappa_\perp = \overline{D}^2\,W_\perp, \quad \kappa_\diamond = \overline{D}^2\,W_\diamond\,, \tag{B.13}$$

the signal term becomes

$$\overline{D}^2\,\boldsymbol{W}(\mathbf{g}) = \kappa_\parallel\,a^4 + \kappa_\diamond\,a^2(1-a^2) + \kappa_\perp\,(1-a^2)^2 \tag{B.14}$$

with the scalar product $a = \langle \mathbf{g}, \boldsymbol{\mu} \rangle$ between the gradient $\mathbf{g}$ and the symmetry axis $\boldsymbol{\mu}$.

Because of the relationship

$$K(\mathbf{g}) = \frac{\overline{D}^2}{\boldsymbol{D}(\mathbf{g})^2}\boldsymbol{W}(\mathbf{g}) \tag{B.15}$$

between kurtosis K and kurtosis tensor $\boldsymbol{W}$, radial and axial kurtosis simply become

$$K(\boldsymbol{\mu}) = K_\parallel = \frac{\kappa_\parallel}{D_\parallel^2}, \quad K_\perp = \frac{\kappa_\perp}{D_\perp^2}\,. \tag{B.16}$$

The parameter $\kappa_\diamond$ is needed to calculate $K(\mathbf{g})$ in any other direction.

B.2. Mean Kurtosis for Cylindrically Symmetric Fibers

In this case the kurtosis is

$$K(\theta) = \frac{\overline{D}^2}{\boldsymbol{D}(\theta)^2}\boldsymbol{W}(\theta) = \frac{\kappa_\parallel \cos^4\theta + \kappa_\diamond \cos^2\theta \sin^2\theta + \kappa_\perp \sin^4\theta}{(D_\parallel \cos^2\theta + D_\perp \sin^2\theta)^2}\,. \tag{B.17}$$

Mean Kurtosis is the average over the unit sphere:

$$\overline{K} = \frac{\int_{S^2} K\,dA}{\int_{S^2} 1\,dA = 4\pi} \tag{B.18}$$
$$= \frac{1}{4\pi}\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \frac{\kappa_\parallel \cos^4\theta + \kappa_\diamond \cos^2\theta \sin^2\theta + \kappa_\perp \sin^4\theta}{(D_\parallel \cos^2\theta + D_\perp \sin^2\theta)^2} \sin\theta \tag{B.19}$$
$$= \frac{1}{2}\int_0^{\pi} d\theta \frac{\kappa_\parallel \cos^4\theta + \kappa_\diamond \cos^2\theta \sin^2\theta + \kappa_\perp \sin^4\theta}{(D_\parallel \cos^2\theta + D_\perp \sin^2\theta)^2} \sin\theta \tag{B.20}$$

Complex variables

We can turn this real integral into a complex integral over the upper half of the unit circle $S_+ = \{z : |z| = 1 \wedge \mathrm{Im}\, z \geq 0\}$, substituting $z = e^{i\theta} \quad \Rightarrow \quad \frac{dz}{d\theta} = ie^{i\theta} = iz$. Analytical continuation on $\mathbb{C}$ gives us:

$$\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) = \frac{1}{2}(z + \frac{1}{z}) \tag{B.21}$$

$$\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta}) = \frac{1}{2i}(z - \frac{1}{z})\,, \tag{B.22}$$

resulting in

$$\overline{K} = \frac{1}{2}\int_{S_+} dz \frac{\kappa_\parallel \frac{1}{16}(z+\frac{1}{z})^4 - \kappa_\diamond \frac{1}{16}(z+\frac{1}{z})^2(z-\frac{1}{z})^2 + \kappa_\perp \frac{1}{16}(z-\frac{1}{z})^4}{(D_\parallel \frac{1}{4}(z+\frac{1}{z})^2 - D_\perp \frac{1}{4}(z-\frac{1}{z})^2)^2} \frac{1}{2i}(z-\frac{1}{z})\frac{1}{iz} \tag{B.23}$$

$$= -\frac{1}{4}\int dz \frac{(\kappa_\parallel(z^2+1)^4 - \kappa_\diamond(z^4-1)^2 + \kappa_\perp(z^2-1)^4)(z^2-1)}{z^2(D_\parallel(z^2+1)^2 - D_\perp(z^2-1)^2)^2} \tag{B.24}$$

Divisor

$$N(z) = z^2\left(D_\parallel(z^2+1)^2 - D_\perp(z^2-1)^2\right)^2 \tag{B.25}$$

$$= z^2\left(z^4(D_\parallel - D_\perp) + 2z^2(D_\parallel + D_\perp) + (D_\parallel - D_\perp)\right)^2 \tag{B.26}$$

We need the roots of $N(z)$ and substitute $y = z^2$:

$$y_\pm = \frac{-(D_\parallel + D_\perp) \pm 2\sqrt{D_\parallel D_\perp}}{D_\parallel - D_\perp} \tag{B.27}$$

giving

$$\frac{N}{(D_\parallel - D_\perp)^2} = y\,(y-y_+)^2(y-y_-)^2 = z^2\,(z-z_+)^2(z+z_+)^2(z-z_-)^2(z+z_-)^2 \tag{B.28}$$

with $z_\pm = \sqrt{y_\pm}$. Because of $D_\parallel > D_\perp > 0$ we know that

$$y_+ \in (-1,0) \qquad y_+ \in (-\infty,-1) \qquad y_- y_+ = 1 \qquad z_+ \in (0,i) \qquad z_- \in (i, i\infty)\,. \tag{B.29}$$

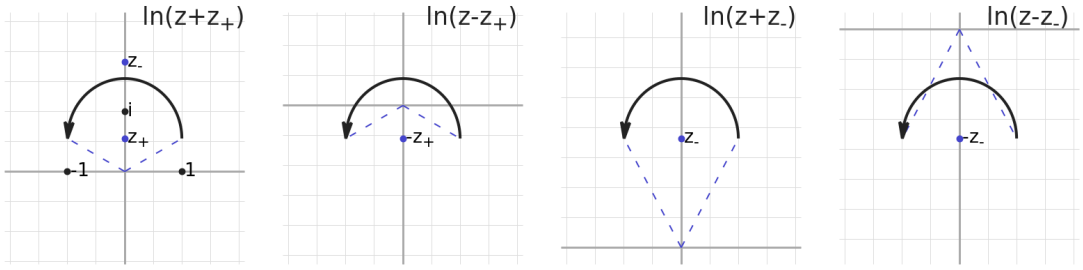
Logarithms

The formula now looks like this:

$$\overline{K} = \kappa_\parallel \int \frac{p_\parallel(z)}{N(z)}\,dz + \kappa_\diamond \int \frac{p_\diamond(z)}{N(z)}\,dz + \kappa_\perp \int \frac{p_\perp(z)}{N(z)}\,dz\,. \tag{B.30}$$

Each integrand is a rational function and will contain logarithm terms $\ln(z \pm z_\pm)$ after integration. To evaluate these integrals along the path S_+, one needs to find a branch of $\ln$ that consistently covers the whole path and then take the difference at the endpoints.

We can use the knowledge from eq. (B.29) about the position of the poles and draw the following image:



This gives the following integrals:

$$[\ln(z+z_+)]_{S_+} = i\pi - 2i\arctan|z_+| \tag{B.31}$$
$$[\ln(z-z_+)]_{S_+} = i\pi + 2i\arctan|z_+| \tag{B.32}$$
$$[\ln(z+z_-)]_{S_+} = i\pi - 2i\arctan|z_-| \tag{B.33}$$
$$[\ln(z-z_-)]_{S_+} = -(i\pi - 2i\arctan|z_-|) \tag{B.34}$$

Result

Putting it all together we get the following formula for the mean kurtosis in the radially symmetric model:

$$\overline{K} = \frac{C_\parallel\kappa_\parallel + C_\diamond\kappa_\diamond + C_\perp\kappa_\perp}{4(D_\parallel - D_\perp)^2} \tag{B.35}$$

with

$$C_\parallel = 3(y_- + 1)A + 4\frac{y_-^2+1}{(y_- - 1)^2} + 2 \tag{B.36}$$
$$C_\diamond = \left(-3(y_- + 1) + \frac{4}{y_+ + 1}\right)A - 6 \tag{B.37}$$
$$C_\perp = \left(3(y_- + 1) - \frac{8}{y_+ + 1} - 16\frac{y_+}{(y_+ + 1)^3}\right)A + 6 - 8\frac{y_-}{(y_- + 1)^2} \tag{B.38}$$
$$A = 2|z_+|\arctan|z_+| \tag{B.39}$$
$$y_- = \frac{-(D_\parallel + D_\perp) - 2\sqrt{D_\parallel D_\perp}}{D_\parallel - D_\perp} \tag{B.40}$$
$$y_+ = 1/y_- \qquad |z_+| = \sqrt{-y_+} \tag{B.41}$$

In the special case that $D_\parallel = D_\perp =: D$, the integral simplifies to

$$\overline{K} = \frac{3\kappa_\parallel + 2\kappa_\diamond + 8\kappa_\perp}{15D^2}\,. \tag{B.42}$$

C. Appendix - H-Psd

C.1. Spectrum and fiber estimation

Idea

At the end of section 7.5, we mentioned, that for fourth order tensors the rank of its H-matrix is equivalent to the rank of the tensor. In the context of our deconvolution model, this is equivalent to the number of fibers described by the fODF tensor. The connection between tensor rank and H-matrix was only shown in [65] for a special class of tensors that includes order four but might not be exact for higher orders. Also because of noise and anatomical complications H can always be expected to have full rank, albeit some eigenvalues will be close to zero.

This section will explore the feasibility of a simple machine learning approach to estimate the number of fibers from the spectrum of H.

Experiment

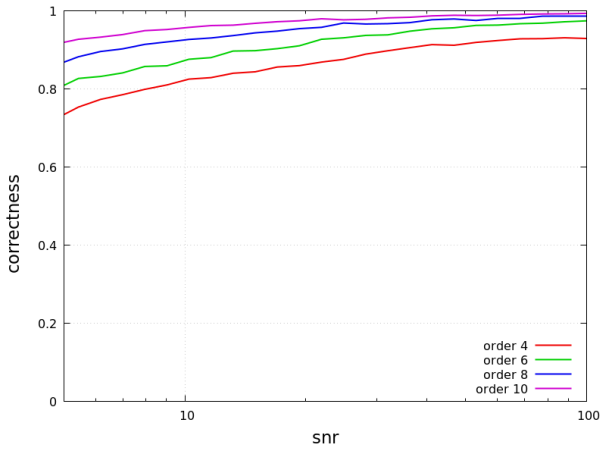


Figure C.1.: Evaluation of the machine learning method to estimate fiber numbers from the spectrum of H for varying levels of noise.

For training we created tensors of order $4, 6, 8, 10$ with $0 - 3$ fibers. Tensors with k fibers were built from a random linear combination of k powers of random direction vectors. 25 different levels of noise ranging from snr 100 to 5 were then added to an intermediate DWMRI signal, produced via convolution and deconvolved to estimate fiber tensors again. Each class of order/fibers/snr included 2000 tensors.

As the learning method we used random forests from the python package `sklearn`. The largest five eigenvalues of H were used as features. For evaluation the same classes and numbers of tensors were created.

The result can be seen in figure C.1. For low noise, orders $6, 8, 10$ can correctly label fibers in at least 95% of the cases, with higher orders performing better, but order 4 still being in the 90% range. Orders $6, 8, 10$ only drop below 90% for SNR ≤ 15.

Repeating the experiment and allowing up to four fibers does not change the accuracy for low noise. At SNR 10, tensor orders 4 and 6 drop by 10 percentage points compared to figure C.1 while order 10 does not change.

C.2. Invariant spectrum

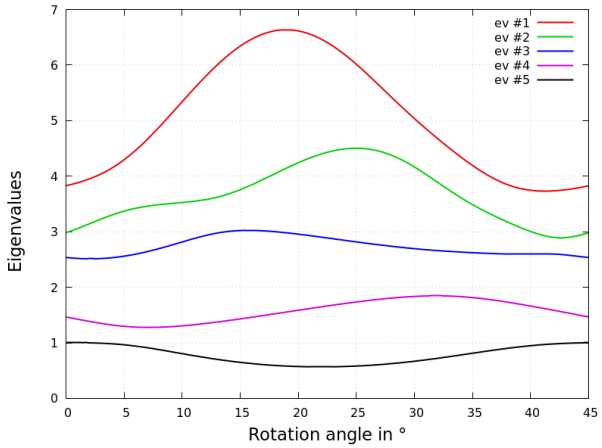


Figure C.2.: The spectrum of H_p for a random form $p(\mathbf{x})$ of degree 6 continously rotated around a fixed axis.

Observation. The spectrum of the matrix H_p is **not** rotationally invariant. The eigenvalues as well as their ratios change when rotating $p(\mathbf{x})$, as can be seen in figure C.2.

The problem stems from the function $L(\mathbf{x}, t)$ as defined in equation (7.16) and its interactions with the space of $t \in \mathbb{R}^{\binom{n+s-1}{s}}$. L can be seen as an isomorphic mapping from parameters $t \in \mathbb{R}^{\binom{n+s-1}{s}}$ into the space of forms $\mathsf{Sym}_{n,s}$. Both spaces have a natural scalar product - $\mathbb{R}^{\binom{n+s-1}{s}}$ has the canonical scalar product while $\mathsf{Sym}_{n,s}$ has the product of forms - but L does not preserve these scalar products. This leads to undesired effects because the definition of eigenvalues for this type of matrix implicitly assumes the canonical scalar product.

New version

To fix this problem we can find an orthonormal basis of $\mathsf{Sym}_{n,s} = \mathbb{R}^{\binom{n+s-1}{s}}$ and define a unitary version of $L(\mathbf{x}, t)$:

$$L^{\mathsf{u}}(\mathbf{x}, t) = \sum_{\boldsymbol{i}} \mathbf{x}^{\boldsymbol{i}} \left(\boldsymbol{i}\right)^{1/2} t_{\boldsymbol{i}} \tag{C.1}$$

The new mapping preserves the scalar product

$$\langle L^{\mathsf{u}}(\cdot, t), L^{\mathsf{u}}(\cdot, v)\rangle = \sum_{\boldsymbol{i}} \left(\boldsymbol{i}\right) \frac{t_{\boldsymbol{i}}}{\left(\boldsymbol{i}\right)^{1/2}} \frac{v_{\boldsymbol{i}}}{\left(\boldsymbol{i}\right)^{1/2}} = \langle t, v\rangle\,. \tag{C.2}$$

Using L^{u}, a modified H-matrix can be defined:

Definition C.1. *The **"covariant" H-form** of $p(\mathbf{x})$ is*

$$H_p^{\mathsf{cov}}(t) = \langle p, L^{\mathsf{u}}(\cdot, t)^2\rangle = \sum_{\boldsymbol{i},\boldsymbol{j}} p_{\boldsymbol{i}+\boldsymbol{j}} \left(\boldsymbol{i}\right)^{1/2} \left(\boldsymbol{j}\right)^{1/2} t_{\boldsymbol{i}}\, t_{\boldsymbol{j}}\,. \tag{C.3}$$

Its matrix can be found via

$$H_{p\,\boldsymbol{i},\boldsymbol{j}}^{\mathsf{cov}} = H_{p\,\boldsymbol{i},\boldsymbol{j}} \left(\boldsymbol{i}\right)^{1/2} \left(\boldsymbol{j}\right)^{1/2}. \tag{C.4}$$

Properties

To make life easier, we need two results about multinomial coefficients. The first one parallels the rule of Pascal's triangle $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ for multi-indices:

Lemma C.1.

$$\left(\boldsymbol{i}\right) = \sum_{r=x,y,z,\ldots} \left(\boldsymbol{i} - \mathbf{1}_r\right) \tag{C.5}$$

Proof.

$$\sum_{r=x,y,z,\ldots} \left(\boldsymbol{i} - \mathbf{1}_r\right) = \frac{(|\boldsymbol{i}| - 1)!}{(i_x - 1)!\, i_y!\, i_z!\, \ldots} + \frac{(|\boldsymbol{i}| - 1)!}{i_x!\, (i_y - 1)!\, i_z!\, \ldots} + \ldots \tag{C.6}$$

$$= \frac{(|\boldsymbol{i}| - 1)!}{i_x!\, i_y!\, i_z!\, \ldots} \underbrace{\left(i_x + i_y + i_z + \ldots\right)}_{|\boldsymbol{i}|} = \frac{|\boldsymbol{i}|!}{i_x!\, i_y!\, i_z!\, \ldots} = \left(\boldsymbol{i}\right) \tag{C.7}$$

□

Lemma C.2. *For* $\boldsymbol{k}$ *with* $|\boldsymbol{k}| = d$ *and* $0 \leq s \leq d$

$$\sum_{\substack{\boldsymbol{i}+\boldsymbol{j}=\boldsymbol{k} \\ |\boldsymbol{i}|=s}} \binom{\boldsymbol{i}}{} \binom{\boldsymbol{j}}{} = \binom{\boldsymbol{k}}{} \tag{C.8}$$

Proof. For $s = 0$ the lemma states $1 \cdot \binom{\boldsymbol{k}}{} = \binom{\boldsymbol{k}}{}$ which is trivially true. The rest follows by induction over s using lemma C.1 in the induction step:

$$\sum_{|\boldsymbol{i}|=s} \binom{\boldsymbol{i}}{} \binom{\boldsymbol{k}-\boldsymbol{i}}{} = \sum_{r=x,y,z,\ldots} \sum_{|\boldsymbol{i}|=s} \binom{\boldsymbol{i}-\mathbf{1}_r}{} \binom{\boldsymbol{k}-\boldsymbol{i}}{} \tag{C.9}$$

$$= \sum_{r=x,y,z,\ldots} \sum_{|\boldsymbol{i}'|=s-1} \binom{\boldsymbol{i}'}{} \binom{\boldsymbol{k}-\boldsymbol{i}'-\mathbf{1}_r}{} \tag{C.10}$$

$$= \sum_{|\boldsymbol{i}'|=s-1} \binom{\boldsymbol{i}'}{} \binom{\boldsymbol{k}-\boldsymbol{i}'}{} \tag{C.11}$$

(*for the purpose of clearity, multinomial coefficients* $\binom{\boldsymbol{i}}{}$ *are assumed to vanish for negative components, allowing us to ignore complications with the range of sums*) □

Now to the two main properties of the new H-matrix:

Theorem C.1. $\boldsymbol{T}_p$ *and* H_p^{cov} *have the same Frobenius norm:*

$$\|H_p^{\mathrm{cov}}\|_F = \|\boldsymbol{T}_p\|_F \tag{C.12}$$

Proof.

$$\|H_p^{\mathrm{cov}}\|_F^2 = \sum_{\boldsymbol{i},\boldsymbol{j}} \left(H_{p\,\boldsymbol{i},\boldsymbol{j}}^{\mathrm{cov}}\right)^2 = \sum_{\boldsymbol{i},\boldsymbol{j}} p_{\boldsymbol{i}+\boldsymbol{j}}^2 \binom{\boldsymbol{i}}{} \binom{\boldsymbol{j}}{} \tag{C.13}$$

$$= \sum_{\boldsymbol{k}} p_{\boldsymbol{k}}^2 \sum_{\boldsymbol{i}+\boldsymbol{j}=\boldsymbol{k}} \binom{\boldsymbol{i}}{} \binom{\boldsymbol{j}}{} \tag{C.14}$$

$$= \sum_{\boldsymbol{k}} p_{\boldsymbol{k}}^2 \binom{\boldsymbol{k}}{} = \langle p, p \rangle = \|\boldsymbol{T}_p\|_F^2 \tag{C.15}$$

$$\tag{C.16}$$

□

Theorem C.2. *The spectrum of* H_p^{cov} *is invariant under rotations* $p(\mathbf{x}) \mapsto p(R\mathbf{x})$.

Proof. For any rotation $R \in SO(3)$ there exists a matrix $\overline{R}$ so that

$$L^{\mathsf{u}}(\mathbf{x}, t) = L^{\mathsf{u}}(R\mathbf{x}, \overline{R}t)\,. \tag{C.17}$$

A similar matrix could also be found for $L(\mathbf{x}, t)$, but by construction of $L^{\mathsf{u}}(\mathbf{x}, t)$ we get the nice property $\overline{R} \in SO(\binom{n+s-1}{s})$. This leads to the simple transformation behavior:

$$\begin{aligned}
\langle t, H_p^{\mathsf{cov}} v\rangle &= \langle p, L^{\mathsf{u}}(\cdot, t) L^{\mathsf{u}}(\cdot, v)\rangle_{\mathbf{x}} && \text{(C.18)}\\
&= \langle p, L^{\mathsf{u}}(\cdot, t) L^{\mathsf{u}}(\cdot, v)\rangle_{R\mathbf{x}} && \text{(C.19)}\\
&= \langle p \circ R, L^{\mathsf{u}}(\cdot, \overline{R}^T t) L^{\mathsf{u}}(\cdot, \overline{R}^T v)\rangle_{x} && \text{(C.20)}\\
&= \langle \overline{R}^T t, H_{p\circ R}^{\mathsf{cov}}\, \overline{R}^T v\rangle && \text{(C.21)}
\end{aligned}$$

or in short

$$H_p^{\mathsf{cov}} = \overline{R} \cdot H_{p\circ R}^{\mathsf{cov}} \cdot \overline{R}^T\,. \tag{C.22}$$

Let $\lambda \in \mathbb{R}$ be an eigenvalue of H_p^{cov}. Then there is a v so that

$$\lambda\langle t, v\rangle = \langle t, H_p^{\mathsf{cov}} v\rangle \quad \forall t\,. \tag{C.23}$$

The transformation formula allows to rewrite this as

$$\begin{aligned}
\lambda\langle \overline{R}^T t, \overline{R}^T v\rangle = \lambda\langle t, v\rangle &= \langle t, H_p^{\mathsf{cov}} v\rangle && \text{(C.24)}\\
&= \langle \overline{R}^T t, H_{p\circ R}^{\mathsf{cov}}\, \overline{R}^T v\rangle \quad \forall t\,. && \text{(C.25)}
\end{aligned}$$

This means, λ is also an eigenvalue of the H^{cov}-matrix of $p(R\mathbf{x})$. □

Bibliography

[1] Andersson, J.L.R., Sotiropoulos, S.N.: An integrated approach to correction for off-resonance effects and subject movement in diffusion mr imaging. NeuroImage 125, 1063–1078 (2016)

[2] Ankele, M., Lim, L.H., Groeschel, S., Schultz, T.: Versatile, robust, and efficient tractography with constrained higher-order tensor fODFs. Int'l J. of Computer Assisted Radiology and Surgery 12(8), 1257–1270 (2017)

[3] Ankele, M., Lim, L.H., Gröschel, S., Schultz, T.: Fast and accurate multi-tissue deconvolution using SHORE and H-psd tensors. In: Proc. Medical Image Analysis and Computer-Aided Intervention (MICCAI) Part III. LNCS, vol. 9902, pp. 502–510 (2016)

[4] Ankele, M., Schultz, T.: Quantifying microstructure in fiber crossings with diffusional kurtosis. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, Springer (2015)

[5] Ankele, M., Schultz, T.: Dt-mri streamsurfaces revisited. IEEE Transactions on Visualization and Computer Graphics 25(1) (2019)

[6] Ankele, M., Schultz, T.: A sheet probability index from diffusion tensor imaging. In: Kaden, E., Grussu, F., Ning, L., Tax, C.M.W., Veraart, J. (eds.) Computational Diffusion MRI, pp. 141–154. Springer (2018)

[7] Basser, P.J., Jones, D.K.: Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review. NMR in Biomedicine 15(7–8), 456–467 (2002)

[8] Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. Magnetic Resonance in Medicine 44, 625–632 (2000)

[9] Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34, 144–155 (2007)

[10] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57(1), 289–300 (1995)

[11] Brinkman, H.: Applications of Spinor Invariants in Atomic Physics. North-Holland Publishing Company, 1st edn. (1956)

[12] Canales-Rodríguez, E.J., Iturria-Medina, Y., Alemán-Gómez, Y., Melie-García, L.: Deconvolution in diffusion spectrum imaging. NeuroImage 50, 136–149 (2010)

[13] Élie Cartan: The Theory of Spinors. Dover Publications, 1st edn. (1981)

[14] Catani, M., Bodi, I., Dell'Acqua, F.: Comment on "the geometric structure of the brain fiber pathways". Science 337(6102), 1605–1605 (2012)

[15] Chen, Z., Tie, Y., Olubiyi, O., Zhang, F., Mehrtash, A., Rigolo, L., Kahali, P., Norton, I., Pasternak, O., Rathi, Y., Golby, A.J., O'Donnell, L.J.: Corticospinal tract modeling for neurosurgical planning by tracking through regions of peritumoral edema and crossing fibers using two-tensor unscented kalman filter tractography. International Journal of Computer Assisted Radiology and Surgery 11(8), 1475–1486 (2016)

[16] Cheng, J., Deriche, R., Jiang, T., Shen, D., Yap, P.T.: Non-negative spherical deconvolution (NNSD) for estimation of fiber orientation distribution function in single-/multi-shell diffusion MRI. NeuroImage 101, 750–764 (2014)

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

[18] Chung, S., Lu, Y., Henry, R.G.: Comparison of bootstrap approaches for estimation of uncertainties of DTI parameters. NeuroImage 33(2), 531–541 (2006)

[19] Dell'Acqua, F., Simmons, A., Williams, S.C.R., Catani, M.: Can spherical deconvolution provide more information than fiber orientations? hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. Human Brain Mapping 34(10), 2464–83 (2013)

[20] Descoteaux, M., Deriche, R., Knösche, T.R., Anwander, A.: Deterministic and probabilistic tractography based on complex fibre orientation distributions. IEEE Transactions on Medical Imaging 28(2), 269–286 (2009)

[21] Driscoll, J.R., Healy, D.M.: Computing fourier transforms and convolutions on the 2-sphere. Advances in Applied Mathematics 15(2), 202–250 (1994)

[22] Einstein, A.: Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. Annalen der Physik 322(8), 549–560 (1905)

[23] Ennis, D.B., Kindlmann, G., Rodriguez, I., Helm, P.A., McVeigh, E.R.: Visualization of tensor fields using superquadric glyphs. Magnetic Resonance in Medicine 53(1), 169–176 (2005)

[24] Galinsky, V.L., Frank, L.R.: The lamellar structure of the brain fiber pathways. Neural Computation 28(11), 2533–2556 (2016)

[25] Garth, C., Gerhardt, F., Tricoche, X., Hagen, H.: Efficient computation and visualization of coherent structures in fluid flow applications. IEEE Trans. on Visualization and Computer Graphics 13(6), 1464–1471 (2007)

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

[27] Grasedyck, L., Kressner, D., Tobler, C.: A literature survey of low-rank tensor approximation techniques (2013)

[28] Gray, H., Carter, H.V., Standring, S.: Gray's Anatomy - The Anatomical Basis of Clinical Practice. Churchill Livingstone, Elsevier, 40th edn. (2008)

[29] Haacke, B., Vankatesan: Magnetic Resonance Imaging: Physical Principles and Sequence Design. Wiley (1999)

[30] Hahn, E.L.: Spin echoes. Physical Review 80(4), 580–594 (1950)

[31] Hanson, L.G.: Is quantum mechanics necessary for understanding magnetic resonance? Concepts in Magnetic Resonance Part A 32A, 329–340 (2008)

[32] Hidalgo-Tobon, S.: Theory of gradient coil design methods for magnetic resonance imaging. Concepts in Magnetic Resonance Part A 36A(4), 223–242 (2010)

[33] Hultquist, J.: Constructing stream surfaces in steady 3d vector fields. In: Proc. IEEE Visualization. pp. 171–178 (1992)

[34] Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.: FSL. NeuroImage 62(2), 782–790 (2012)

[35] Jensen, J.H., Helpern, J.A.: MRI quantification of non-gaussian water diffusion by kurtosis analysis. NMR in Biomedicine 23(7), 698–710 (2010)

[36] Jensen, J.H., Helpern, J.A., Ramani, A., Lu, H., Kaczynski, K.: Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magnetic Resonance in Medicine 53, 1432–1440 (2005)

[37] Jeremić, B., Scheuermann, G., Frey, J., Yang, Z., Hamann, B., Joy, K.I., Hagen, H.: Tensor visualizations in computational geomechanics. International Journal for Numerical and Analytical Methods in Geomechanics 26, 925–944 (2002)

[38] Jeurissen, B., Leemans, A., Jones, D.K., Tournier, J.D., Sijbers, J.: Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. Human Brain Mapping 32, 461–479 (2011)

[39] Jeurissen, B., Leemans, A., Tournier, J.D., Jones, D.K., Sijbers, J.: Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. Human Brain Mapping 34(11), 2747–66 (2013)

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

[41] Jiao, F., Gur, Y., Johnson, C.R., Joshi, S.: Detection of crossing white matter fibers with high-order tensors and rank-k decompositions. In: Székely, G., Hahn, H.K. (eds.) IPMI. LNCS, vol. 6801, pp. 538–549 (2011)

[42] Jones, D.K.: Challenges and limitations of quantifying brain connectivity in vivo with diffusion mri. Future Medicine 2(3), 341–355 (2010)

[43] Jones, D.K., Knösche, T.R., Turner, R.: White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. NeuroImage 73, 239–254 (2013)

[44] Kato, T.: Perturbation theory for linear operators, Die Grundlehren der mathematischen Wissenschaften, vol. 132. Springer, 2nd edn. (1976)

[45] Kindlmann, G., Tricoche, X., Westin, C.F.: Delineating white matter structure in diffusion tensor MRI with anisotropy creases. Medical Image Analysis 11(5), 492–502 (2007)

[46] Knutsson, H., Westin, C.F.: Tensor metrics and charged containers for 3d q-space sample distribution. In: Mori et al., K. (ed.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI) Part I. LNCS, vol. 8149, pp. 679–686. Springer (2013)

[47] Kreher, B., Schneider, J., Mader, I., Martin, E., Hennig, J., Il'yasov, K.: Multitensor approach for analysis and tracking of complex fiber configurations. Magnetic Resonance in Medicine 54, 1216–1225 (2005)

[48] Malcolm, J.G., Shenton, M.E., Rathi, Y.: Filtered multitensor tractography. IEEE Trans. on Medical Imaging 29(9), 1664–75 (2010)

[49] Malcolm, J.G., Michailovich, O., Bouix, S., Westin, C.F., Shenton, M.E., Rathi, Y.: A filtered approach to neural tractography using the Watson directional function. Medical Image Analysis 14, 58–69 (2010)

[50] Merlet, S.L., Deriche, R.: Continuous diffusion signal, EAP and ODF estimation via compressive sensing in diffusion MRI. Medical Image Analysis 17, 556–572 (2013)

[51] Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C.M.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Annals of Neurology 45(2), 265–269 (1999)

[52] Nishimura, D.G.: Principals of Magnetic Resonance Imaging. Stanford University (2010)

[53] Özarslan, E., Mareci, T.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. Magnetic Resonance in Medicine 50, 955–965 (2003)

[54] Pajevic, S., Aldroubi, A., Basser, P.J.: A continuous tensor field approximation of discrete DT-MRI data for extracting microstructural and architectural features of tissue. Journal of Magnetic Resonance 154, 85–100 (2002)

[55] Palacios, J., Yeh, H., Wang, W., Zhang, Y., Laramee, R.S., Sharma, R., Schultz, T., Zhang, E.: Feature surfaces in symmetric tensor fields based on eigenvalue manifold. IEEE Trans. on Visualization and Computer Graphics 22(3), 1248–1260 (2016)

[56] Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C.: Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. NeuroImage 59, 2241–54 (2012)

[57] Paquette, M., Merlet, S., Gilbert, G., Deriche, R., Descoteaux, M.: Comparison of sampling strategies and sparsifying transforms to improve compressed sensing diffusion spectrum imaging. Magnetic Resonance in Medicine 73(1), 401–416 (2015)

[58] Peled, S., Friman, O., Jolesz, F., Westin, C.F.: Geometrically constrained two-tensor model for crossing tracts in DWI. Magnetic Resonance Imaging 24(9), 1263–1270 (2006)

[59] Preim, B., Botha, C.: Visual Computing for Medicine: Theory, Algorithms, and Applications. Morgan Kaufmann (2014)

[60] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C++: The Art of Scientific Computing, Second Edition. Cambridge Univ. Press (2002)

[61] Qazi, A.A., Radmanesh, A., O'Donnell, L., Kindlmann, G., Peled, S., Westin, C.F., Golby, A.J.: Resolving crossings in the corticospinal tract by two-tensor streamline tractography: Method and clinical assessment using fMRI. NeuroImage 47(2), T98–T106 (2009)

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

[63] Reisert, M., Kiselev, V.G., Dihtal, B., Kellner, E., Novikov, D.S.: MesoFT: unifying diffusion modelling and fiber tracking. In: Proc. Medical Image Computing and Computer-Assisted Intervention. LNCS, vol. 8675, pp. 201–208. Springer (2014)

[64] Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., Kiselev, V.: Global fiber reconstruction becomes practical. NeuroImage 54(2), 955–962 (2011)

[65] Reznick, B.A.: Sums of even powers of linear forms. Memoirs of the American Mathematical Society 96(463), 1–155 (1992)

[66] Sadlo, F., Peikert, R.: Efficient visualization of lagrangian coherent structures by filtered AMR ridge extraction. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE Visualization) 13(6), 1456–1463 (2007)

[67] Sakurai, J.: Modern Quantum Mechanics. Pearson (1993)

[68] Schultz, T., Fuster, A., Ghosh, A., Deriche, R., Florack, L., Lim, L.H.: Higher-order tensors in diffusion imaging. In: Westin, C.F., Vilanova, A., Burgeth, B. (eds.) Visualization and Processing of Tensors and Higher Order Descriptors for Multi-Valued Data, pp. 129–161. Springer (2014)

[69] Schultz, T., Groeschel, S.: Auto-calibrating spherical deconvolution based on ODF sparsity. In: Mori et al., K. (ed.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI) Part I. LNCS, vol. 8149, pp. 663–670. Springer (2013)

[70] Schultz, T., Seidel, H.P.: Estimating crossing fibers: A tensor decomposition approach. IEEE Trans. on Visualization and Computer Graphics 14(6), 1635–1642 (2008)

[71] Schultz, T., Theisel, H., Seidel, H.P.: Topological visualization of brain diffusion MRI data. IEEE Trans. on Visualization and Computer Graphics 13(6), 1496–1503 (2007)

[72] Schultz, T., Theisel, H., Seidel, H.P.: Crease surfaces: From theory to extraction and application to diffusion tensor MRI. IEEE Trans. on Visualization and Computer Graphics 16(1), 109–119 (2010)

[73] Schultz, T., Vilanova, A.: Diffusion MRI visualization. NMR in Biomedicine (2018), early view. DOI: 10.1002/nbm.3902

[74] Schultz, T., Vilanova, A., Brecheisen, R., Kindlmann, G.: Fuzzy fibers: Uncertainty in dMRI tractography. In: Hansen, C., Chen, M., Johnson, C., Kaufman, A., Hagen, H. (eds.) Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization, pp. 79–92. Springer (2014)

[75] Schultz, T., Westin, C.F., Kindlmann, G.: Multi-diffusion-tensor fitting via spherical deconvolution: A unifying framework. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI). LNCS, vol. 6361, pp. 673–680. Springer (2010)

[76] Schulze, M., Rössl, C., Germer, T., Theisel, H.: As-perpendicular-as-possible surfaces for flow visualization. In: Proc. IEEE Pacific Visualization Symposium. pp. 153–160 (2012)

[77] Sondershaus, R., Gumhold, S.: Meshing of diffusion surfaces for point-based tensor field visualization. In: Proc. 12th International Meshing Roundtable (IMR). pp. 177–188 (2003)

[78] Sotiropoulos, S.N., Jbabdi, S., Xu, J., Andersson, J.L., Moeller, S., Auerbach, E.J., Glasser, M.F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D.A., Yacoub, E., Lenglet, C., Essen, D.C.V., Ugurbil, K., Behrens, T.E.: Advances in diffusion {MRI} acquisition and processing in the human connectome project. NeuroImage 80, 125–143 (2013)

[79] Spivak, M.: A Comprehensive Introduction to Differential Geometry, vol. 1. Publish or Perish, 3rd edn. (1999)

[80] Stejskal, E., Tanner, J.: Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. Journal of Chemical Physics 42, 288–292 (1965)

[81] Tabesh, A., Jensen, J.H., Ardekani, B.A., Helpern, J.A.: Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. Magnetic Resonance in Medicine 65, 823–836 (2011)

[82] Taquet, M., Scherrer, B., Boumal, N., Macq, B., Warfield, S.K.: Estimation of a multi-fascicle model from single b-value data with a population-informed prior. In: Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI) Part I. LNCS, vol. 8149, pp. 695–702. Springer (2013)

[83] Tax, C.M.W., Westin, C.F., Dela Haije, T., Fuster, A., Viergever, M.A., Calabrese, E., Florack, L., Leemans, A.: Quantifying the brain's sheet structure with normalized convolution. Medical Image Analysis 39, 162–177 (2017)

[84] Tax, C.M., Dela Haije, T., Fuster, A., Westin, C.F., Viergever, M.A., Florack, L., Leemans, A.: Sheet probability index (SPI): Characterizing the geometrical organization of the white matter with diffusion MRI. NeuroImage 142, 260–279 (2016)

[85] Torrey, H.C.: Bloch equation with diffusion terms. Physical Review 104(3), 563–565 (1956)

[86] Tournier, J.D., Calamante, F., A.Connelly: Improved probabilistic streamlines tractography by 2nd order integration over fiber orientation distributions. In: Proc. International Society of Magnetic Resonance in Medicine (ISMRM). p. 1670 (2010)

[87] Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. NeuroImage 35, 1459–1472 (2007)

[88] Tournier, J.D., Calamante, F., Connelly, A.: MRtrix: Diffusion tractography in crossing fiber regions. International Journal of Imaging Systems and Technology 22(1), 53–66 (2012)

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

[90] Tournier, J.D., Mori, S., Leemans, A.: Diffusion tensor imaging and beyond. Magnetic Resonance in Medicine 65(6), 1532–1556 (2011)

[91] Tricoche, X., Kindlmann, G., Westin, C.F.: Invariant crease lines for topological and structural analysis of tensor fields. IEEE Transactions on Visualization and Computer Graphics (Proc. IEEE Visualization) 14(6), 1627–1634 (2008)

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

[93] Vandenberghe, L.: The CVXOPT linear and quadratic cone program solvers. Tech. rep., UCLA Electrical Engineering Department (2010), `http://www.seas.ucla.edu/~vandenbe/publications/coneprog.pdf`

[94] Vilanova, A., Berenschot, G., van Pul, C.: DTI visualization with stream surfaces and evenly-spaced volume seeding. In: Deussen, O., Hansen, C., Keim, D.A., Saupe, D. (eds.) Proc. Joint Eurographics – IEEE TCVG Symposium on Visualization (VisSym). pp. 173–182 (2004)

[95] Wedeen, V.J., Rosene, D.L., Wang, R., Dai, G., Mortazavi, F., Hagmann, P., Kaas, J.H., Tseng, W.Y.I.: Response to comment on "the geometric structure of the brain fiber pathways". Science 337(6102), 1605–1605 (2012)

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

[97] Wedeen, V.J., Rosene, D.L., Wang, R., Dai, G., Mortazavi, F., Hagmann, P., Kaas, J.H., Tseng, W.Y.I.: The geometric structure of the brain fiber pathways. Science 335(6076), 1628–1634 (2012)

[98] Wedeen, V., Wang, R., Schmahmann, J., Benner, T., Tseng, W., Dai, G., Pandya, D., Hagmann, P., D'Arceuil, H., de Crespigny, A.: Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. NeuroImage 41(4), 1267–1277 (2008)

[99] Weldeselassie, Y.T., Barmpoutis, A., Atkins, M.S.: Symmetric positive semi-definite cartesian tensor fiber orientation distributions (CT-FOD). Medical Image Analysis 16(6), 1121–1129 (2012)

[100] Westin, C.F., Maier, S., Mamata, H., Nabavi, A., Jolesz, F., Kikinis, R.: Processing and visualization for diffusion tensor MRI. Medical Image Analysis 6, 93–108 (2002)

[101] Wigner, E.P.: Group Theory: And its Application to the Quantum Mechanics of Atomic Spectra. Academic Press, 1st edn. (1959)

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

[103] Zhang, S., Demiralp, C., Laidlaw, D.H.: Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. IEEE Transactions on Visualization and Computer Graphics 9(4), 454–462 (2003)

[104] Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans. on Medical Imaging 20(1), 45–57 (2001)

[105] Zheng, X., Parlett, B., Pang, A.: Topological lines in 3D tensor fields and discriminant hessian factorization. IEEE Transactions on Visualization and Computer Graphics 11(4), 395–407 (2005)