

UNIVERSITÄT HAMBURG

**Fakultät für Mathematik, Informatik und  
Naturwissenschaften  
Fachbereich Physik**

Microstructure imaging using MRI: diffusion MRI and  
biophysical models under the influence of noise

**Dissertation**

**zur Erlangung des Doktorgrades  
an der Fakultät für Mathematik, Informatik und Naturwissenschaften  
Fachbereich Physik  
der Universität Hamburg**

von  
**Jan Malte Oeschger**

Hamburg im Mai 2024



**Universität Hamburg**

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Gutachter der Dissertation:

Dr. Siawoosh Mohammadi  
Prof. Dr. Florian Grüner

Zusammensetzung der Prüfungskommission:

Dr. Siawoosh Mohammadi  
Prof. Dr. Florian Grüner  
Dr. Karsten Tabelow  
Prof. Dr. Sune Nørhøj Jespersen  
Prof. Dr. Michael Potthoff

Vorsitzender der Prüfungskommission:

Prof. Dr. Michael Potthoff

Datum der Disputation:

21.06.2024

Vorsitzender des Fach-Promotionsausschusses PHYSIK:

Prof. Dr. Markus Drescher

Leiter des Fachbereichs PHYSIK:

Prof. Dr. Wolfgang J. Parak

Dekan der Fakultät MIN:

Prof. Dr.-Ing. Norbert Ritter

## Danksagung

Mein großer Dank gilt Herrn Dr. Siawoosh Mohammadi für seine wertvolle Unterstützung, die inspirierenden Fachdiskussionen und das stetig konstruktive Feedback, das mir im Verlauf meiner Doktorarbeit immer geholfen hat. Seine aufmerksame Förderung und Hinweise haben meine Doktorarbeit maßgeblich unterstützt.

Ein besonderer Dank gilt ebenso Herrn Dr. Karsten Tabelow vom Weierstraß-Institut für Angewandte Analysis und Stochastik in Berlin. Seine großzügige Hilfsbereitschaft und offene Diskussionskultur haben mir während meiner Doktorarbeit an wichtigen Stellen sehr geholfen.

Sehr danken möchte ich auch Herrn Prof. Dr. Florian Grüner vom Institut für Experimentalphysik der Universität Hamburg, für die freundliche Unterstützung als weiterer Betreuer meiner Doktorarbeit.

Des Weiteren möchte ich dem gesamten "qMRI and in vivo histology" Team danken. Dr. Francisco Javier Fritz, Laurin Mordhorst, Nina Lüthi, Björn Fricke, Tobias Streubel und Behnam Ashtarayeh haben mich mit ihrem Fachwissen immer unterstützt und bereichert und eine positive Atmosphäre und ein positives Miteinander in der Gruppe geschaffen.

Darüber hinaus gilt mein Dank auch den weiteren Büronachbarinnen Ivana Tanasic, Antoniya Boyanova und Alena Rußmann für ihre Kollegialität und Unterstützung.

## Abstract

The microstructure of white matter in the human brain plays a central role in healthy human development and non-invasive imaging of it is crucial for understanding the neurological function as well as disorders. While *ex vivo* microscopy provides detailed insights into the microstructure of nervous tissue, *in vivo* imaging of nervous tissue and its changes remains a challenge. Diffusion magnetic resonance imaging (dMRI) is a non-invasive, *in vivo* imaging technique built around diffusion of water molecules in nervous tissue which is sensitive to the tissue microstructure. Hereby, dMRI offers a non-invasive alternative to *ex vivo* microscopy. Diffusion kurtosis imaging (DKI) is a physical framework used to interpret the measured dMRI signal. DKI shows promise for clinical use due to its ability to capture restricted diffusion which is typical for diffusion in nervous tissue due to the complex, cellular microstructure hindering it. Axisymmetric DKI, a modification of DKI, reduces complexity by introducing additional symmetry assumptions, theoretically making it more noise-robust and data efficient. Both DKI variants are diffusion models used to estimate physical diffusion properties of the investigated tissue which are "only" correlated to the tissue microstructure. Here, biophysical models go one step further and enhance the interpretability of the dMRI signal by connecting it to specific metrics of the actual biological tissue microstructure. For both DKI and biophysical models, noise in dMRI images poses a critical hurdle for accurate and precise parameter estimation. For example, noise in dMRI can lead to a bias in the parameter estimates, the so-called "Rician bias" which burdens parameter estimation for both DKI and biophysical models. This thesis investigates the effects of the Rician bias and how to mitigate them, axisymmetric DKI's inherent bias, caused by violation of its additional symmetry assumptions, as well as the performance and accuracy of a variety of currently used biophysical models under the influence of noise. Through published, peer-reviewed articles, insights into bias-free parameter estimation at low signal to noise ratios are presented. Furthermore, the severity of the axisymmetric DKI inherent bias is quantified and ways to deal with it are explored. Finally, bias propagation of DKI parameters used to compute the biophysical parameters are investigated and the accuracy of various biophysical models are evaluated against a biological gold standard. Through this, the role of the combination of axisymmetric DKI and Rician bias correction as an highly effective "all-rounder" for reducing the Rician bias in DKI parameter estimation is demonstrated. Furthermore, the potential of a Bayesian-enhanced machine learning based approach for biophysical parameter estimation named "Baydiff" for becoming an accurate tool used for microstructure imaging under the influence of noise is highlighted.

## Zusammenfassung

Die Mikrostruktur der weißen Gehirns substanz des Menschen spielt eine zentrale Rolle für die gesunde menschliche Entwicklung. Hierbei ist die nicht-invasive Bildgebung dieser Mikrostruktur entscheidend für das Verständnis der neurologischen Funktion sowie von neurologischen Störungen. Während in diesem Zusammenhang *ex vivo* Mikroskopie detaillierte

Einblicke in die Mikrostruktur des Nervengewebes bietet, ist die Bildgebung des Nervengewebes und seiner Veränderungen *in vivo*, also im lebenden Menschen, eine Herausforderung. Diffusionsgewichtete Magnetresonanztomographie (dMRI) ist eine nicht-invasive, *in vivo* Bildgebungstechnik, die die Diffusion von Wassermolekülen im Nervengewebe messbar macht und deshalb auch sensitiv für die Gewebemikrostruktur ist. dMRI stellt deshalb eine nicht-invasive Alternative zur *ex vivo* Mikroskopie dar. Die Diffusions-Kurtosis-Bildgebung (DKI) ist hier ein physikalisches Modell, das verwendet wird, um das gemessene dMRI-Signal zu interpretieren. DKI bietet vielversprechende, potenzielle Anwendungsmöglichkeiten in der klinischen Praxis, weil es imstande ist gehinderte Diffusion zu erfassen, die typisch für die Diffusion im Nervengewebe wegen seiner komplexen, zellulären Mikrostruktur ist. Axialsymmetrische DKI ist eine Modifikation vom standard DKI Modell, die die Komplexität reduziert, indem sie zusätzliche Symmetrieannahmen einführt. Theoretisch machen diese zusätzlichen Annahmen die axialsymmetrische DKI robuster gegen die Einflüsse von Rauschen und dateneffizienter. Beide DKI Varianten sind Diffusionsmodelle, die verwendet werden, um die physikalischen Diffusionseigenschaften vom Nervengewebe zu bestimmen, die „nur“ mit der Mikrostruktur korreliert sind. Biophysikalische Modelle gehen hier einen Schritt weiter und erweitern die Interpretierbarkeit des dMRI-Signals, indem sie es mit spezifischen Metriken der tatsächlichen biologischen Mikrostruktur in Verbindung bringen. Sowohl für DKI als auch biophysikalische Modelle stellt Rauschen in dMRI-Bildern eine kritische Hürde für die exakte und präzise Parameterbestimmung dar. Zum Beispiel kann Rauschen in dMRI-Bildern zu einer systematischen Über- oder Unterschätzung der bestimmten Parameter führen, dem sogenannten „Rician Bias“, der die Parameterabschätzung sowohl für DKI als auch biophysikalische Modelle erschwert. Diese Thesis untersucht die Auswirkungen des Rician Bias und wie man ihn reduzieren kann, den inhärenten Bias in der axialsymmetrischen DKI, der durch die Verletzung ihrer zusätzlichen Symmetrieannahmen verursacht wird, sowie die Performance einer Vielzahl von derzeit verwendeter biophysikalischer Modelle unter dem Einfluss von Rauschen. Anhand von veröffentlichten, peer-reviewten, wissenschaftlichen Artikeln werden Einblicke in die bias-freie Parameterbestimmung bei niedrigen Signal-Rausch-Verhältnissen präsentiert. Zusätzlich wird die Stärke des inhärenten Bias der axialsymmetrischen DKI quantifiziert und Möglichkeiten zum Umgang mit diesem Bias erforscht und präsentiert. Außerdem wird die Fortpflanzung des inhärenten Bias der axialsymmetrischen DKI über die DKI Parameter, die zur Berechnung der biophysikalischen Parameter verwendet werden, untersucht und die Genauigkeit verschiedener biophysikalischer Modelle in Relation zu einem biologischen Goldstandard evaluiert. Insgesamt zeigt diese Arbeit so, dass die Kombination aus axialsymmetrischer DKI und Rician Bias-Korrektur eine sehr effektive „all-rounder“ Kombination zur Reduzierung des Rician Bias für die praktische Parameterbestimmung ist. Darüber hinaus wird das Potenzial eines machine-learning basierten, biophysikalischen Modells namens „Baydiff“ herausgestellt, ein verlässlicher Algorithmus für die biophysikalische Mikrostruktur-Bildgebung unter dem Einfluss von Rauschen zu werden.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methods</b>	<b>5</b>
2.1	Diffusion MRI and diffusion tensor imaging frameworks . . . . .	5
2.2	Noise in diffusion MRI . . . . .	9
2.3	Simulation studies . . . . .	10
2.4	Biophysical models . . . . .	11
<b>3</b>	<b>Results</b>	<b>14</b>
3.1	Biases in diffusion-weighted MRI . . . . .	14
3.1.1	Influence of noise in diffusion kurtosis imaging: Rician bias and precision and the effects of axisymmetric DKI and Rician bias correction in various tissue types . . . . .	14
3.1.2	Axisymmetric DKI inherent bias: Variance versus model inherent bias in the axisymmetric DKI tensor metrics . . . . .	19
3.2	Biophysical models: Influence of noise and axisymmetric DKI inherent bias . .	22
3.2.1	Axisymmetric DKI inherent bias propagation into biophysical parameters	22
3.2.2	Estimation of biophysical parameters under the influence of noise . . .	24
3.3	Biophysical models: Assessing the accuracy under the influence of noise with a histology based gold standard . . . . .	28
3.3.1	Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy	28
3.3.2	Accuracy of MRI based axonal volume fraction estimates of various biophysical models under the influence of noise . . . . .	30
<b>4</b>	<b>Discussion</b>	<b>33</b>

4.1	Overcoming biases in standard DKI and axisymmetric DKI and possible implications for practical applications . . . . .	34
4.2	Biophysical modeling under uncertainty and how to navigate it . . . . .	35
<b>5</b>	<b>Conclusion</b>	<b>36</b>
<b>6</b>	<b>Appendix</b>	<b>37</b>
6.1	Paper 1: Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study . . . . .	37
6.2	Paper 2: Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates . . . . .	71
6.3	Paper 3: Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy . . .	102

# 1 Introduction

The white matter tissue of the brain is a central constituent of the human nervous system. Its microstructure plays a crucial role in healthy human development and neurological disorders<sup>1;2;3;4</sup>. Accurate insights into the microstructure of the white matter therefore is central for understanding the neuronal function and integrity of the human brain and the origin and progression of neurological disorders.

Today, the gold standard for estimation of microstructure parameters describing nervous tissue on the micrometer scale is using *ex vivo* microscopy techniques<sup>5</sup>. Here, an example for an important microstructure parameter would be the volume fraction of axons in relation to the other microscopic tissue compartment like for example myelin, astrocytes and others. However, since in this way only *ex vivo* tissue is investigated, this technique is not suitable for tracking tissue changes *in vivo*, which is crucial, e.g., for understanding the progression of diseases. A possible alternative to *ex vivo* microscopy are diffusion MRI<sup>6</sup> (dMRI) and biophysical diffusion models. dMRI is a non-invasive, *in vivo* imaging technique that is measuring the ensemble averaged, macroscopic diffusion of water molecules through nervous tissue caused by Brownian motion shaped by the tissue microstructure.

The dMRI signal can physically be described by diffusion tensor imaging (DTI)<sup>7</sup> which is an established imaging technique in clinical practice. A more complex approach which is able to capture more realistic diffusion processes in nervous tissue, like restricted and hindered diffusion, is diffusion kurtosis imaging (DKI)<sup>8;9</sup>. Since DKI is an expansion of DTI, it has the additional advantage that its measurement protocols only need to be adapted slightly. However, DKI is not yet established in clinical practice even though it offers several advantages over DTI, which is why the focus of this thesis is on DKI. Both DTI and DKI are physical frameworks describing the diffusion of water through nervous tissue and allow derivation of parameters quantifying its diffusion properties<sup>10</sup>. These parameters can be derived from the diffusion and kurtosis tensors estimated in DTI or DKI and can be thought of as "material properties" describing the diffusion characteristics of the tissue on a purely physical level. The DTI and DKI signal equations already possess certain symmetry assumptions like inversion symmetry, reducing the number of free parameters. Axisymmetric DKI<sup>11;12;13</sup> is a promising modification of DKI which introduces an additional symmetry assumption: axons of a fiber bundle are axisymmetrically aligned around an axis of symmetry, which could be the direction of the main fiber bundle. This additional symmetry assumption imposes that diffusion processes in the imaged MRI voxel can only be axisymmetric which reduces the parameter space describing the diffusion pattern even more. Hereby, axisymmetric DKI becomes more noise-robust and possibly less data demanding and consequently more acquisition-time efficient. However, if the underlying, real diffusion pattern in an MRI voxel is not axisymmetric, e.g., due to a complex nerve fiber structure, the assumptions made in axisymmetric DKI are violated, leading to a model



inherent bias in axisymmetric DKI. Within this context, an important subset of rotationally invariant DKI parameters capturing complementary diffusion information are the five axisymmetric DKI tensor metrics (AxTM), the parallel and perpendicular diffusivity ( $D_{\parallel}$  and  $D_{\perp}$ ) and kurtosis ( $W_{\parallel}$  and  $W_{\perp}$ ) and the mean of the kurtosis tensor ( $\bar{W}$ ). The AxTM are important, because they can be computed with both standard DKI and axisymmetric DKI. Hence they can be used for performance comparison of both DKI frameworks, e.g., in a simulation study which quantifies the degree to which results of axisymmetric DKI differ from standard DKI using the AxTM as target metrics.

Another way to interpret the dMRI signals is by using biophysical models<sup>14;15;16;17;18;19;20;21;22</sup> which relate the dMRI signal to specific, biological properties of the tissue microstructure. Biophysical models establish this relation via a model that identifies the main tissue components and their interactions and how they influence the measured MR signal. This, e.g., allows estimation of the axon water fraction which can be used to compute the volume fraction of axons also accessible with *ex vivo* microscopy. Other parameters are the tissue compartment specific diffusivities and dispersion properties of the axons. Through these parameters, biophysical models hold the potential to enable novel insights into the development of the human brain and the mechanisms behind neurodegenerative diseases<sup>23;24;25;26;27;28;29</sup>. This, however, requires that the estimated biophysical parameters actually describe genuine, microscopic tissue properties which needs to be shown, e.g., by comparison to an established gold standard method. Depending on the biophysical model, the biophysical parameters are either obtained directly from the dMRI signals or by using parameters that were first estimated with DTI or DKI, e.g., the AxTM. If the AxTM are estimated using axisymmetric DKI, the axisymmetric DKI inherent bias could propagate into the biophysical parameters computed from them which requires additional investigation and consideration in this case.

Noise in the acquired dMRI<sup>30</sup> signals can pose a severe obstacle for DKI or biophysical parameter estimation in practice because the measured dMRI signal is weak and contrast is created via a signal decay. Noise competes with the acquired MR signal and can distort dMRI parameter estimates to a point where even models that are highly accurate in principle cannot produce useful results due to a low precision (see Figure 1 for a graphical depiction of the difference between accuracy and precision). Noise in MRI can have different origins with different corresponding distributional properties. Here we focus on thermal noise in the form of electrical fluctuations in the electrical components of the MRI scanner modeled by a Gaussian distribution in k-space<sup>31</sup> where the raw signals are recorded. This thermal noise has to be distinguished from physiological noise originating from processes of the human body like breathing, the heartbeat, swallowing movements or bodily movements in general. Both DKI and biophysical parameters are often estimated from MR magnitude images. MR magnitude images are computed from the raw, complex MRI signals recorded in k-space. When computing magnitude images, the raw, Gaussian noise in k-space is transformed to so-called "Rician noise" which follows a Rician or,

more generally a non-central  $\chi$  distribution. Since the Rician distribution is lopsided, especially for low signal to noise ratios (SNR), this leads to the so-called "Rician Bias"<sup>32;30;33</sup> in the estimated DKI parameters if this change of noise distribution is not properly accounted for. The severity of the Rician bias in the DKI parameters is determined by the SNR<sup>34;35</sup> which, in addition to the MR sequence and device, also depends on the tissue microstructure. The dependency on the tissue microstructure is intimately linked with dMRI which creates contrast via a signal decay depending on whether water molecules are more or less restricted along a given diffusion gradient direction which is dictated by the tissue microstructure, e.g., the main fiber orientation. In addition to directly biasing the DKI parameters, a Rician bias in the DKI parameters can also propagate into the biophysical parameters where it might even be enhanced due to non-linear dependencies. Bias-free DKI parameter estimation therefore is one necessary prerequisite for accurate biophysical parameter estimation.

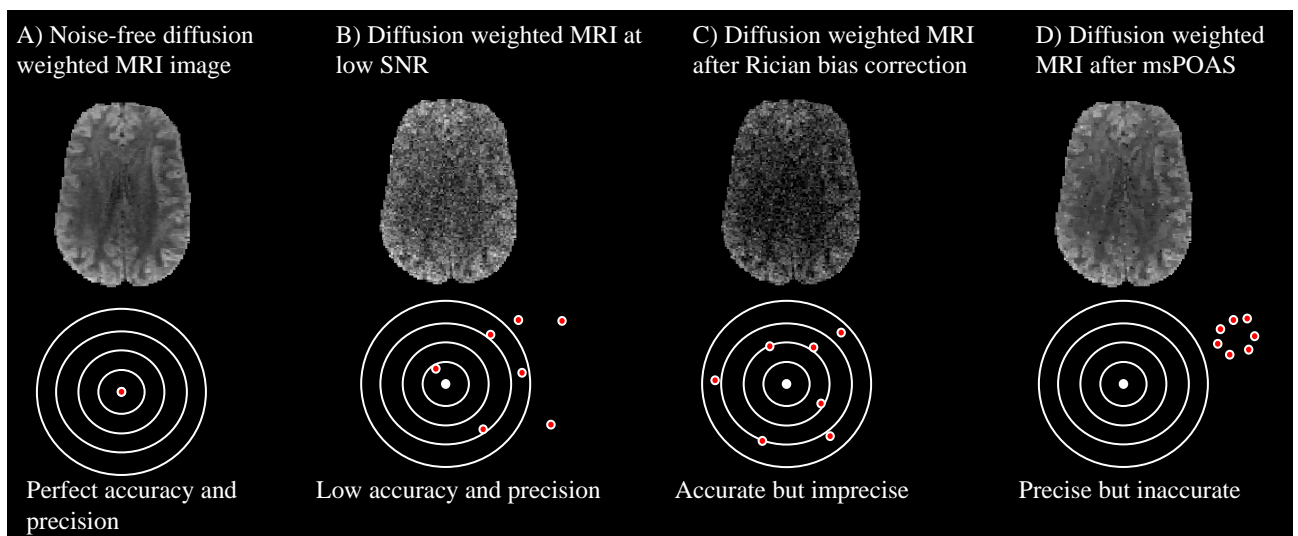


Figure 1: Graphic depicting the difference between accuracy and precision in diffusion weighted MRI. The bullseye and the red dots represent repeated measurements and estimation of a desired parameter. Shown are examples of: A) An axial view of a noise-free diffusion weighted MRI image as a reference, B) the same image shown in A) at SNR = 5 with low accuracy, i.e., a bias and low precision, C) the same image after Rician bias correction which improves accuracy by removing the bias but does not improve precision and D) the same image after using the adaptive denoising method "Multi-shell Position-Orientation Adaptive Smoothing" (msPOAS) which improves precision but does not remove the bias.

Because of this, approaches for quantification of biases in different scenarios and ways of dealing with them are important prerequisites for a clinical application of DKI and biophysical models. Therefore, before dMRI frameworks and biophysical models can be used in clinical or scientific practice, at least two conditions have to be fulfilled: a) the noise susceptibility of the dMRI and biophysical parameters have to be investigated to establish suitable protocols for bias-free parameter estimation at sufficient high SNR and b) it has to be shown that the specific biophysical parameters describe genuine microstructural properties by comparison against an established gold standard. Here, approaches like Rician bias correction (RBC) (e.g., Figure 1),

that account for the correct noise distribution in the fitting process or the axisymmetric DKI framework with its smaller parameter space can mitigate the effects of the Rician bias. To shine light on issues a) and b), this thesis addresses the following concrete questions:

1. How does the Rician bias affect estimation of all five AxTM, can its influence be mitigated with axisymmetric DKI or Rician bias correction and is there a synergistic effect between these two?
2. Is axisymmetric DKI with its smaller parameter space more precise than standard DKI?
3. How severe is the axisymmetric DKI inherent bias, can it be seen under realistic, noisy conditions and can it be reduced by confining the analysis to voxels with lower levels of fiber complexity?
4. How does the axisymmetric DKI inherent bias in the AxTM affect biophysical parameters that are computed from them?
5. How well can the biophysical models estimate the biophysical parameters under the influence of noise?
6. How accurate are the most commonly used biophysical models in comparison with an established, histological, electron microscopy gold standard? How is this performance under the influence of noise?

To answer these points, this thesis presents published, peer-reviewed scientific articles as well as not yet published studies investigating the raised questions. First, in the methods section (Section 2), the main theories and methods required to understand the presented studies are outlined. Then, the studies are presented in the results section (Section 3). Section 3.1.1 addresses question 1. by exploring the Rician bias in both standard DKI and axisymmetric DKI in combination with a Rician bias correction algorithm in a noise simulation study. This same noise study was also used to address question 2. by evaluating the precision of axisymmetric DKI compared to standard DKI via the standard deviation of their parameter estimates. Section 3.1.2 presents a study where noise-free simulations of a healthy human brain are used to quantify the axisymmetric DKI inherent bias. Furthermore, voxel selection masks are used to see if confining the analysis to areas with lower levels of fiber complexity reduces the axisymmetric DKI inherent bias. Additionally, noisy data are investigated to see if the axisymmetric DKI inherent bias is typically visible under realistic conditions. Together, this section answers the points raised in question 3. Questions 1 through 3 provide detailed insights into the accuracy and precision of standard DKI and axisymmetric DKI and provide hands-on approaches for dealing with both the Rician bias and the axisymmetric DKI inherent bias. Then, biophysical models are investigated. Section 3.2.1 explores propagation of the axisymmetric DKI inherent bias in AxTM estimates into the biophysical parameters of a commonly used biophysical model

(WMTI-Watson), addressing question 4. Finally, accuracy and noise-susceptibility of a variety of biophysical models is investigated. Here, in Section 3.2.2, a multimodal mouse dataset consisting of dMRI data and electron microscopy (EM) data is introduced and used to perform a noise simulation to assess how well the investigated biophysical models perform under the influence of noise, addressing question 5. The EM data of the multimodal dataset were then used to assess the accuracy of the investigated biophysical models in Section 3.3.1 and Section 3.3.2, addressing question 6. The findings reported in the results section are then put into perspective of the bigger picture and their indications for dMRI and biophysical modeling is discussed in Section 4. In its entirety, this thesis provides insights into synergistic combinations of algorithms that enable bias-free parameter estimation for low SNRs, concrete, hands-on approaches for dealing with axisymmetric DKI's inherent bias as well as insights into the noise-robustness of a variety of commonly used biophysical models and their accuracy with respect to an actual biological gold standard. The published, peer-reviewed articles are appended in the Appendix at the end of this thesis.

## 2 Methods

### 2.1 Diffusion MRI and diffusion tensor imaging frameworks

The basis of the MRI signal is a macroscopic magnetization generated in the tissue by strong magnetic fields of a MR scanner systems. The strong magnetic field of the MR scanner interacts with the nuclear spin angular momentum  $\vec{I}$  of hydrogen atoms bound in water molecules (carrying the magnetic moment  $\vec{\mu} = \gamma \cdot \vec{I}$  where  $\gamma = \frac{42.58 \text{ MHz}}{2\pi \text{ T}}$  is the gyromagnetic ratio of the positively charged hydrogen nucleus) and exerts a force causing a spin precession at the so-called Larmor frequency  $\omega = \gamma |\vec{B}(\vec{r}, t)|$  where  $\vec{B}(\vec{r}, t) = \vec{B}_0 + \vec{B}_1(\vec{r}, t)$  is the local magnetic field and  $\vec{B}_0 = B_0 \cdot \vec{e}_z$  is the MR scanner's external magnetic field and  $\vec{B}_1(\vec{r}, t)$  is the magnetic field component of the RF-pulse. To first approximation, the Larmor frequency is constant across the tissue if only the external field of the MR scanner system is active and magnetic field susceptibilities are negligible.

This roughly aligns  $\approx \frac{1}{10^6}$  of the nuclear spins in the tissue of the human body resulting in the macroscopic magnetization. The macroscopic magnetization has two components, a longitudinal z-component  $M_z$  aligned along the external magnetic field and a transversal component  $M_{xy}$  perpendicular to the z-direction which, by itself, is zero. The longitudinal component is then deflected into the perpendicular plane via a oscillating radio frequency (RF) pulse at the Larmor frequency resulting in a non-zero transversal magnetization set by the flip angle  $\alpha$  of the RF pulse and a precession of the macroscopic magnetization which induces a change of magnetic flux in a nearby coil that can be measured as an induced voltage. Over time, the magnetization

vector is described by the Bloch equation that is valid if no diffusion is present:

$$\frac{\partial}{\partial t} \begin{pmatrix} M_x(\vec{r}, t) \\ M_y(\vec{r}, t) \\ M_z(\vec{r}, t) \end{pmatrix} = \gamma(\mathbf{M}(\vec{r}, t) \times \mathbf{B}(\vec{r}, t)) - \begin{pmatrix} \frac{1}{T_2} M_x(\vec{r}, t) \\ \frac{1}{T_2} M_y(\vec{r}, t) \\ \frac{1}{T_1} (M_z(\vec{r}, t) - M_{z,0}) \end{pmatrix} \quad (1)$$

Here  $M_{z,0}$  is the equilibrium magnetization z-component of  $M_0$ ,  $T_2$  is the transverse relaxation time describing the time it takes for the transverse component of the magnetization  $M_{xy}$  to decay back to  $\frac{1}{e}$  where the decay is caused by a dephasing of the aligned spins of the macroscopic magnetization due to local perturbations in the magnetic field caused by microscopic spin-spin interaction.  $T_1$  is the longitudinal relaxation time describing the time the longitudinal magnetization  $M_z$  needs to recover to its equilibrium state  $M_0$

The solutions of  $M_{xy}$  and  $M_z$  of the Bloch equation are:

$$M_z(t) = M_0(1 - e^{-\frac{t}{T_1}}) + M_z(0)e^{-\frac{t}{T_1}} \quad (2)$$

$$M_{xy}(\vec{r}, t) = M_{xy}(0) \cdot e^{-i\gamma|\vec{B}(\vec{r}, t)|t} e^{-\frac{t}{T_2}} \quad (3)$$

here  $M_z(0)$  and  $M_{xy}(0)$  are the initial states of the magnetizations (set by the RF pulse) and  $|\vec{B}(\vec{r}, t)|$  is the local magnetic field present in the tissue. The transverse component  $M_{xy}$  precesses at the Larmor frequency  $\omega = \gamma|\vec{B}(\vec{r}, t)|$  and induces a measurable voltage in a coil if there is a non-zero  $M_{xy}(0)$ . For acquiring MR images, this signal generation process has to be confined to a small three dimensional tissue volume called "voxel". This is achieved by employing additional linear, magnetic spatial encoding gradients  $\vec{g}$  that linearly modify the magnetic field across the sample and hence cause a spatially dependent Larmor frequency  $\omega(\vec{r}) = \gamma\vec{B}_0 \cdot \vec{e}_z + \gamma\vec{g}\vec{r}$ . MR images can be generated by changing the gradient strength in between measurements which leads to a spatially dependent on-resonance frequency of the RF pulse so that spin excitation can be confined (or ascribed) to an image voxel.

The differential equation governing diffusion-weighted MRI was written down in 1956 by H.C. Torrey who modified the Bloch equation (1) to account for magnetization transfer<sup>36</sup> in fluids. He did this using the diffusion coefficient  $D$  of a fluid and analyzing the diffusion current density of spins and the resulting Bloch-Torrey equation reads:

$$\frac{\partial}{\partial t} \begin{pmatrix} M_x(\vec{r}, t) \\ M_y(\vec{r}, t) \\ M_z(\vec{r}, t) \end{pmatrix} = \gamma(\mathbf{M}(\vec{r}, t) \times \mathbf{B}(\vec{r}, t)) + \begin{pmatrix} \nabla D \nabla (M_x - M_{x,0}) \\ \nabla D \nabla (M_y - M_{y,0}) \\ \nabla D \nabla (M_z - M_{z,0}) \end{pmatrix} - \begin{pmatrix} \frac{1}{T_2} M_x(\vec{r}, t) \\ \frac{1}{T_2} M_y(\vec{r}, t) \\ \frac{1}{T_1} (M_z(\vec{r}, t) - M_{z,0}) \end{pmatrix} \quad (4)$$

here  $M_{x,0}$ ,  $M_{y,0}$  and  $M_{z,0}$  are the magnetization components of the equilibrium magnetization  $M_0$ . The Bloch-Torrey equation (4) is the fundamental differential equation of diffusion tensor imaging because it describes the precession of the induced transversal magnetization and accounts for the effects of spin diffusion within a fluid.

The fundamental imaging principle of diffusion-weighted MRI is introduction of an additional linear magnetic "diffusion gradients"  $\vec{B}(\vec{r}) = \vec{B}_0 + \mathbf{G} \cdot \vec{r}$ , where  $\mathbf{G}$  is the gradient's field strength. The additional linear magnetic gradient  $\vec{g}(\vec{r}) = \mathbf{G} \cdot \vec{r}$ , again, introduces a Larmor frequency shift  $\Omega(\vec{r}) = \gamma(\mathbf{G} \cdot \vec{r})$  similar to the spatially encoding gradients which in this case lead to different precession frequencies  $\omega(\vec{r}) = \gamma\vec{B}_0 + \Omega(\vec{r}, t)$  across the tissue causing a dephasing and consequent attenuation of the measured signal relative to a non-diffusion weighted reference signal.

A well-known imaging sequence making use of this principle is the Pulsed Gradient Spin Echo (PGSE) sequence<sup>6</sup> also known as the Stejskal-Tanner sequence which is a modification of the Spin Echo sequence. The PGSE sequence uses additional linear magnetic diffusion gradients after the excitation pulse ( $90^\circ$ ) and the refocusing pulse ( $180^\circ$ ) of the Spin Echo sequence which sensitize the MRI signal to dephasing of the molecular spins, picked up if they travel along the direction of the diffusion gradient. The introduced frequency offset  $\Omega(\vec{r}, t)$  depends on the direction of the diffusion gradient  $\vec{g}$ . The measured signal described by equation (5) only attenuates (relative to a non-diffusion weighted reference signal) if the underlying tissue structure allows water molecules to diffuse into the direction of the diffusion gradient  $\vec{g}$ . For practical MRI image generation, diffusion-weighted images are usually acquired with the so-called "echo-planar-imaging-technique" (EPI) which, consists of a successive measurement of pulsed gradient spin echoes and allows to acquire image slices within a time of 50 to 100 ms<sup>37;38;39</sup>.

The fundamental signal equation for the diffusion-weighted signal measured in diffusion tensor imaging follows from the Bloch-Torrey equation (4) and reads<sup>40;41</sup>:

$$S(t, \vec{q}) = \int_V d\vec{r} \frac{d\vec{r}_0}{V} e^{-i\vec{q}(\vec{r}-\vec{r}_0)} \Psi(t, \vec{r}, \vec{r}_0) \quad (5)$$

a derivation of this equation can for example be found in<sup>40</sup>, here  $\Psi(t, \vec{r}, \vec{r}_0)$  is the probability density function of displacement of molecular spins from point  $\vec{r}_0$  to  $\vec{r}$  over time  $t$ <sup>40;41</sup> and  $\vec{q}(t) = -\int_0^t \vec{g}(\tau) d\tau$  is found via the magnetic gradient pulse  $\vec{g}(t)$ . The signal decay being measured is

the result of the diffusion of molecular spins within a voxel and consequent signal dephasing in the presence of a spatially dependent magnetic field  $\vec{B}(\vec{r}, t)$ .

The probability density function  $\Psi(t, \vec{r}, \vec{r}_0)$  describes the probability of molecules moving from point  $\vec{r}_0$  to  $\vec{r}$  and it is intimately linked to the underlying tissue micro structure. The diffusion-weighted signal is the result of considering all possible starting points  $\vec{r}_0$  and end points  $\vec{r}$  of spin displacements where the average spin dephasing depends on the tissue micro structure and results in a measurable signal decay.

Any arbitrary probability density function (PDF) can be expressed in terms of its cumulants<sup>40;42</sup> where only the Gaussian and delta PDF can be represented using a finite set of cumulants ( $c_1$  and  $c_2$  for the Gaussian PDF,  $c_1$  for the delta PDF) while any other PDF has an infinite set of nonzero cumulants<sup>42</sup>.

The PDF of spin displacement in formula (5) is complex but can be approximated using its first two cumulants. Using only the first cumulant is referred to as "Gaussian approximation" and yields the DTI framework equation:

$$S_{b, \vec{g}}(D) = S_0 e^{-b \vec{g}^T D \vec{g}} \quad (6)$$

Here,  $b$  is the so-called b-value, which tunes the strength of the diffusion weighting of the MR signal,  $S_0$  is the non-diffusion weighted signal and  $D$  is the so-called diffusion tensors. The  $b$  and the diffusion gradient  $\vec{g}$  can be set in the imaging sequence. In three dimensions, the symmetric diffusion tensor  $D$  consists six distinct directionally dependent diffusion constants  $D_{ij} = D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz}, D_{yz}$  (SI units  $\frac{m^2}{s}$ ).

A more complex approximation incorporates the second order cumulant, extending the DTI framework to media that highly restrict diffusion like nervous tissue, where diffusion is not Gaussian (or free) anymore. The dimensionless ratio  $\frac{c_4}{c_2^2}$  is called kurtosis and describes deviation from the Gaussian diffusion. Accounting for kurtosis leads to the diffusion kurtosis imaging framework:

$$S_{b, \vec{g}}(D, W) = S_0 \exp \left( -b \sum_{i,j=1}^3 g_i g_j D_{ij} + \frac{b^2}{6} \left( \sum_{i=1}^3 \frac{D_{ii}}{3} \right)^2 \sum_{i,j,k,l=1}^3 g_i g_j g_k g_l W_{ijkl} \right) \quad (7)$$

Here  $g_1, g_2, g_3$  are the entries of the diffusion gradient and  $\bar{D} = \frac{\sum_{i=1}^3 D_{ii}}{3}$  is the mean of the diffusion tensor (mean diffusivity). The DTI and DKI framework account for inversion symmetry reflected in the symmetry of the diffusion and kurtosis tensors. Assuming nerve fibers to be axially symmetrically organized around an axis of symmetry  $\vec{c}$  is referred to as "axially symmetric DKI", here called "axisymmetric DKI". These additional symmetry assumption drastically re-

duce the parameter space of DKI from 22 in "standard DKI" to 8 in axisymmetric DKI<sup>11</sup> where the signal equation can be written using the set of axisymmetric DKI tensor metrics, AxTM =  $\{\bar{W}, W_{\parallel}, W_{\perp}, D_{\parallel}, D_{\perp}\}$ ,  $S_0$  and the axis of symmetry  $\vec{c}$  as follows:

$$S_{b,\vec{g}}(\text{AxTM}) = S_0 \exp\left(-B_{ij}D_{ij} + \frac{1}{6}\bar{D}^2 B_{ij}B_{kl}W_{ijkl}\right) \quad (8)$$

where

$$B_{ij}D_{ij} = \text{Tr}(B)D_{\perp} + (D_{\parallel} - D_{\perp})\vec{c}^T B \vec{c} \quad (9)$$

and

$$B_{ij}B_{kl}W_{ijkl} = \frac{1}{2}(10W_{\perp} + 5W_{\parallel} - 15\bar{W})(\vec{c}^T B \vec{c})^2 \quad (10)$$

$$+ \frac{1}{2}(5\bar{W} - W_{\parallel} - 4W_{\perp})(\vec{c}^T B \vec{c} \text{Tr}(B)) \quad (11)$$

$$+ 2\vec{c}^T B B \vec{c} + \frac{W_{\perp}}{3}(\text{Tr}(B))^2 + 2\text{Tr}(B \otimes B)) \quad (12)$$

with

$$B = b \begin{pmatrix} g_x^2 & g_x g_y & g_x g_z \\ g_x g_y & g_y^2 & g_y g_z \\ g_x g_z & g_y g_z & g_z^2 \end{pmatrix}$$

Rotationally invariant, tissue characterizing DTI and DKI metrics can be either estimated directly in case of axisymmetric DKI where the axisymmetric DKI tensor metrics (AxTM)  $\bar{W}, W_{\parallel}, W_{\perp}, D_{\parallel}, D_{\perp}$  are estimated directly via fitting the signal equation or via the diffusion and kurtosis tensors metrics in DTI and DKI where the tensors eigenvalues and parameters are used to calculate the AxTM.

## 2.2 Noise in diffusion MRI

Parameter estimation using MRI is burdened by noise originating from many different sources but can be categorized as uncorrelated, random (white) noise and systematic noise. Uncorrelated white noise instances include the spontaneous movement of charge carriers in the electrical components (primarily the signal recording coils) of the MRI system due to thermic agitation, variations and imperfections in the radio frequency pulses, or fluctuations in the thermal radi-



ation emitted, for instance, by a patient undergoing imaging. On the other hand, systematic noise, such as physiological noise, arises from a systematic change in spin density within an imaged tissue volume due to factors like blood flow or breathing.

The noise contamination process in diffusion-weighted MRI can be modeled as following:

in the k-space where the raw MRI signals are recorded, the signal in each coil is contaminated with noise that can be described as an additive, white Gaussian noise process with a mean of zero and standard deviation  $\sigma$  as  $S_{cont.,k} = S_{0,k} + a + bi$  where  $\sigma$  is assumed to be stationary, so that it does not depend on the location and  $a, b \in \mathcal{N}(0, \sigma_k)$ . Images are depicted in spatial "x" space, so that  $S_{cont.,k}$  needs to be Fourier transformed. Since the Fourier transform of a Gaussian is still a Gaussian, the noise-contaminated signal in x-space after Fourier transformation is still Gaussian. The contaminated signal in x-space can be written as  $S_{cont.,x} = S_{0,x} + \alpha + \beta i$  with  $\alpha, \beta \in \mathcal{N}(0, \sigma_x)$ . Magnitude images are then created by computing the magnitude  $abs(S_{cont.,x})$ , which changes the Gaussian noise to Rician noise in case of one receiver coil  $L=1$  or more generally a noncentral  $\chi$  distribution in case of arbitrary number of receiver coils  $L$ . Not considering the proper noise distribution, e.g, when fitting the DTI, DKI or axisymmetric DKI signal equations directly to the diffusion-weighted signals, leads to the so-called "Rician bias" in the estimated parameters which becomes more severe, the lower the SNR is. There are several Rician bias correction approaches<sup>43;44;45;46</sup>, Rician bias correction in the context of Section 6.1 refers to an approach where the noise-free signal equations of DTI (Equation (6)), DKI (Equation (7)) or axisymmetric DKI (Equation (8)) are modified by calculating the corresponding expectation value of a Rician distribution and optimizing the framework parameters such that this Rician expectation value fits best to the measured signals instead of fitting the noise-free signal equations.

The Rician bias is a phenomenon that becomes visible if many repeated parameter estimates of identical measurements could be compared to a ground truth and would show itself as an SNR dependent bias. On the level of an individual measurement, the variation of parameter estimates introduced by noise for repeated measurements directly impacts parameter estimation in the form of variance. This variance has to be contrasted from the bias and methods supposed to deal with biased parameter estimates. While Rician bias correction is supposed to get rid of the systematic bias introduced by Rician noise, denoising methods like, e.g., msPOAS<sup>47</sup> reduce the variance in the acquired signals and thus the estimated parameters.

## 2.3 Simulation studies

In this thesis a number of simulation studies were performed and analyzed. Within this context, "simulation" refers to using standard DKI or axisymmetric DKI as a forward model for generation of noise-free signals  $S_{noise-free}$  which are then noise contaminated according to

$S_{\text{cont}} = |S_{\text{noise-free}} + \alpha + \beta i|$ , where  $\alpha, \beta \in \mathcal{N}(0, \sigma)$  were drawn from a zero mean Gaussian with standard deviation  $\sigma$ . The input into the simulation are the DKI tensors (standard DKI) or axisymmetric DKI tensor metrics (axisymmetric DKI), obtained from a previous fit on dMRI data. In this way the dMRI measurement under the influence of noise can be simulated for various tissues.

## 2.4 Biophysical models

The DTI and DKI frameworks can be considered mathematical representations<sup>48</sup> of the dMRI signal. Within this context, the DKI parameters can be thought of as describing the material properties that are probed by the water molecules diffusing through it. At their heart is a cumulant expansion of a probability density function that is used to mathematically approximate molecular movement that is sensitive to the tissue geometry. These frameworks have to be distinguished from biophysical tissue models which are based on theories identifying the relevant tissue structures and models of their interactions with each other and external forces of the MR scanner which can be used to derive "biophysical parameters". Through this, biophysical tissue models establish a specific correspondence between the underlying microstructural tissue structure, restricting molecular diffusion, and the measured dMRI signal. The investigated biophysical models and their parameters can either be estimated based on DKI parameters like the AxTM or the dMRI signals themselves and are thus accessible via an *in vivo* MR measurement. A bias in the AxTM fit results or the dMRI signals can therefore be expected to propagate into the biophysical parameters.

All biophysical models investigated in this thesis can be derived from the four compartment, *ex vivo* neurite orientation dispersion and density imaging (NODDI) signal model<sup>17</sup> whose "total signal" is comprised of an axonal, extracellular, isotropic and dot compartment:

$$\frac{S_{dMRI}}{S_0} = (1 - f_{dot})[(1 - f_{iso})(fS_{axonal} + (1 - f)S_{extra-axonal}) + f_{iso}S_{iso}] + f_{dot}S_{dot} \quad (13)$$

where  $f$ ,  $f_{iso}$  and  $f_{dot}$  are the axonal, isotropic and dot signal fractions and  $S$  denotes the respective signals. The signal fraction  $f$  is often used as an approximation to the axonal water fraction (AWF), see also Section 3.3.1 or<sup>23</sup>.

The signal originating from the axonal compartment,  $S_{axonal}$ , is modeled as zero-radius sticks where the intra-axonal diffusivity  $D_a$  quantifies the diffusion along the axons<sup>49;23</sup>. The signal originating from the extra-axonal compartment  $S_{extra-axonal}$  is modeled as a homogeneous medium that can be described as axisymmetric Gaussian diffusion with a parallel diffusivity

$D_{e,\parallel}$  along the axons and a perpendicular diffusivity  $D_{e,\perp}$  perpendicular to the axons. Finally, the isotropic compartment models diffusion as isotropic Gaussian which typically describes the cerebrospinal fluid (csf) with a diffusion constant  $D_{csf}$  ( $\approx 3 \frac{\mu m^2}{ms}$  at body temperature)<sup>16</sup>. The extra-axonal space is modeled as a homogeneous medium that can be described with an axisymmetric diffusion tensor with the parallel and perpendicular extra axonal diffusivities  $D_{e,\parallel}$  and  $D_{e,\perp}$ . Furthermore, axon orientation is modeled by a fiber orientation distribution function (fODF) that can be expanded using spherical harmonics and, in the well-known "WMTI-Watson" variant of the standard model, is assumed to be the axisymmetric Watson distribution parameterized by the axon alignment parameter  $\kappa$ . Here, the dot compartment is only present in fixed, *ex vivo* tissue and accounts for trapped water in small structures like cell bodies and its signal is measured mainly when the signal of the remaining compartments are suppressed (e.g. by acquiring highly diffusion-weighted shells along the main fiber orientation<sup>50</sup>).

The biophysical parameters of the standard model carry tissue specific meaning that can be associated with diseases or pathological processes like demyelination ( $D_{e,\perp}$ ), axonal or dendritic loss and oedema ( $f$ ) or beading and inflammation ( $f_{iso}$  and  $D_a$ ,  $D_{e,\parallel}$  and  $D_{e,\perp}$ )<sup>23</sup>. In this thesis, a variety of biophysical models were investigated: the white matter tract integrity ("WMTI")<sup>15</sup> model and the WMTI model with the additional assumption that the fiber orientation distribution function (fODF) can be modeled as the Watson distribution ("WMTI-Watson")<sup>19;20</sup>, the neurite orientation dispersion and density imaging ("NODDI")<sup>17</sup> model and its DTI based variant ("NODDI-DTI")<sup>18</sup>, standard model imaging ("SMI")<sup>22</sup> and the "Bayesian estimation of microstructural diffusion parameters" model ("Baydiff")<sup>21</sup>. Baydiff ("Bayesian Estimation of Microstructural Diffusion Parameters") uses Bayesian-enhanced, supervised machine learning for estimating the expectation values of the posterior probability distribution of the model parameters by minimizing a quadratic risk function<sup>51;21</sup> to obtain the biophysical parameter estimates. A huge advantage of this approach is that the model also implicitly learns the Rician noise distribution via the simulated training data<sup>21</sup>. The biophysical parameters of all of these models can be derived based on diffusion-weighted MRI images or are directly linked to the parameters of the diffusion tensor imaging (DTI) or diffusion kurtosis imaging (DKI) frameworks. Table 1 lists the investigated biophysical models, their input parameters, biophysical model parameters and model assumptions.

DWI-model	Input	Biophysical parameters $p_j$	Assumptions
SMI	dMRI data, noise map	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, f, p_2$	$f_{dot} = f_{iso} = 0$ fODF = arbitrary, represented via spherical harmonics decomposition
BAYDIFF <sup>21</sup>	dMRI data, noise map	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, f, f_{iso}, p_2$	$f_{dot} = 0$
WMTI-W <sup>+</sup>	$D_{\parallel}, D_{\perp}, \bar{W}, W_{\parallel}, W_{\perp}$	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, \kappa, f$	$f_{dot} = f_{iso} = 0,$ $D_a^{\parallel} > D_e^{\parallel},$ fODF = Watson distribution <sup>52</sup>
WMTI	All 21 DKI tensor metrics	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, f, p_2$	$f_{dot} = f_{iso} = 0,$ completely parallel fibers ( $\kappa \rightarrow \infty$ )
NODDI	dMRI data	<i>in vivo</i> : $\kappa, f, f_{iso}$ <i>ex vivo</i> : $\kappa, f, f_{iso}, f_{dot},$	$D_e^{\parallel} = D_a^{\parallel},$ $D_e^{\perp} = (1 - f)D_e^{\parallel},$ $D_{iso} = 2.0\mu\text{m}^2/\text{ms},$ $D_e^{\parallel} = 0.35\mu\text{m}^2/\text{ms}$ fODF = Watson distribution
NODDI-DTI	$FA, MD$	$\kappa, f$	$f_{dot} = f_{iso} = 0,$ $D_e^{\parallel} = D_a^{\parallel},$ $D_e^{\perp} = (1 - f)D_e^{\parallel},$ $D_e^{\parallel} = 0.35\mu\text{m}^2/\text{ms}$ fODF = Watson distribution

Table 1: Table 2 from<sup>53</sup> (adapted to fit the terminology of this thesis and with extra information about the model's fODF) showing a summary of the biophysical model ("DWI model"), the required input parameters ("Input"), the biophysical model parameters ("Biophysical parameters  $p_j$  and the biophysical model assumptions ("Assumptions")). Apart from the already introduced parameters, there are the fractional anisotropy (FA), mean diffusivity (MD), a noise map estimating the noise per voxel (e.g., measured as the standard deviation per voxel for repeated, identical measurements, see<sup>54</sup>), signal fraction of the axonal compartment  $f$ , signal fraction of the isotropic compartment  $f_{iso}$ , signal fraction of the dot compartment  $f_{dot}$  and  $p_2$  is one of the expansion coefficients of the fiber orientation distribution function that describes axon alignment, the expansion coefficients are also referred to as "anisotropy metrics", see, e.g.,<sup>22</sup>.

## 3 Results

### 3.1 Biases in diffusion-weighted MRI

There are two main sources of bias in DKI and axisymmetric DKI that were investigated in this work. First the noise driven Rician bias in standard DKI and axisymmetric DKI and second the model inherent bias of axisymmetric DKI. These biases are not only relevant for DKI or axisymmetric DKI and their parameter estimates but also for biophysical models that are based on them. Therefore, a precise understanding of the behavior of these biases is crucial. In the following, two publications investigating these biases are presented:

- First publication: "Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study"<sup>55</sup>. This publication deals with the question whether axisymmetric DKI is able to reduce the Rician bias and the variance in the parameter estimates. Furthermore, the combination with Rician bias correction and axisymmetric DKI was investigated and the effectiveness was tested on all five axisymmetric DKI tensor metrics.
- Second publication: "Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates"<sup>56</sup>. This publication deals with the question whether the axisymmetric DKI framework comes with an inherent bias related to its additional symmetry assumptions. This inherent bias was quantified in simulations of the white matter of a healthy brain and propagation into the biophysical parameters of the WMTI-Watson model was investigated.

#### 3.1.1 Influence of noise in diffusion kurtosis imaging: Rician bias and precision and the effects of axisymmetric DKI and Rician bias correction in various tissue types

**Subsection introduction:** Noise in MRI introduces a random variation in the measured signals leading to a variance in parameter estimates as well as a Rician bias if MR magnitude images are obtained. Since the use of MR magnitude images is widespread in dMRI, the Rician bias is often encountered here. The influence of noise is inversely proportional to the signal-to-noise ratio (SNR), i.e., the lower the SNR the more severe the Rician bias<sup>45</sup> and the lower the precision. Particularly the Rician bias therefore depends on a variety of factors that influence the SNR like, e.g., image resolution or the diffusion-weighting strength which is typically higher in DKI than in DTI. Additionally, the SNR depends on the underlying tissue microstructure, since it defines the diffusion characteristics in the tissue and thus the strength of the signal decay of the diffusion-weighted signals. In this section the paper "Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study"<sup>55</sup> is summarized to answer the questions how the Rician bias affects the different AxTM, if its influence can effectively be

mitigated by axisymmetric DKI and a Rician bias correction and if there is a synergistic effect between these two (question 1. from the introduction). Furthermore, the question whether axisymmetric DKI is more precise than standard DKI is answered, addressing question 2. from the introduction.

**Subsection methods:** Two approaches for dealing with the Rician bias and their interaction with each other were investigated. First, a Rician bias correction (RBC) algorithm that is modifying the noise-free standard DKI or axisymmetric DKI signal model prediction<sup>57;55</sup> using the Rician expectation value was investigated. Second, the performance of axisymmetric DKI with its reduced parameter space, which should make it more noise-robust, was investigated. The work was using a simulation study (see Section 2.3) based on synthetic and *in vivo* based white matter for densely sampled SNRs= [1,2,3...200], simulated for 2500 noise samples per voxel. The synthetic white matter was generated with varying degrees of fiber alignments from low alignment ("LA") to high alignment ("HA") using axisymmetric DKI as a forward model. The *in vivo* based white matter simulation was based on a dMRI measurement of the brain of a healthy human participant and standard DKI was used as a forward model to generate signals for 12 selected voxels from that dataset. The ground truth against which the parameter estimates of the various combinations of algorithms were compared were "model based", i.e., they were the standard DKI fit results obtained from noise-free signals in case of the *in vivo* based white matter simulation or simply the five AxTM input parameters used for signal generation in case of the synthetic white matter simulation. The target metric for measuring the Rician bias was the absolute value of the mean percentage error ("A-MPE"):

$$\text{A-MPE} = 100 \cdot \frac{|\text{GT} - \overline{\text{FitResults}}(\text{SNR})|}{\text{GT}} \quad (14)$$

here "GT" refers to the ground truth parameters and " $\overline{\text{FitResults}}$ " refers to the average of the fit results obtained from fitting 2500 noise samples per SNR. The investigated DKI parameters were the five axisymmetric DKI tensor metrics  $\overline{W}$ ,  $W_{\parallel}$ ,  $W_{\perp}$ ,  $D_{\parallel}$  and  $D_{\perp}$  computed with standard DKI or axisymmetric DKI with and without Rician bias correction. The leading question in this analysis was "at what SNR is the A-MPE smaller than 5%?", i.e., at what SNR is the accuracy of the average estimator within 5% of its ground truth. Here, the 5% threshold was considered a reasonable trade-off between accuracy and SNR requirement. Being below this threshold is also referred to as being "bias-free".

To measure the precision of both standard DKI and axisymmetric DKI with and without Rician bias correction, the standard deviation in reference to the ground truth ("R-STD") was calculated according to  $\text{R-STD} = 100 \cdot \frac{\text{std}(\text{Distribution}_{\text{Estimator}})}{\text{GroundTruth}}$ . Here  $\text{std}(\text{Distribution}_{\text{Estimator}})$  is the standard deviation over the distribution of the 2500 fit results per SNR and AxTM. The leading question here was "at what SNR is the R-STD smaller than 5%?", i.e., at what SNR is the precision of a certain method within 5% of its ground truth?

**Subsection results:** Figure 2 (Figure 6 from<sup>55</sup>) shows the SNR at which  $A\text{-MPE} \leq 5\%$ . Here, a lower SNR is favorable because it means that the parameter can be accurately estimated for lower SNRs. Figure 2 demonstrated that RBC itself (hatched barplots) worked particularly well for the parallel diffusion ( $D_{\parallel}$ ) and kurtosis ( $W_{\parallel}$ ) and mean of the kurtosis tensor ( $\bar{W}$ ). However, RBC needed at least some level of fiber alignment to be effective, since it did not improve parameter estimation in the "LA" voxel mimicking highly isotropic tissue with very low fiber alignment. Furthermore, Figure 2 demonstrated that the axisymmetric DKI framework was superior to standard DKI for estimating the perpendicular diffusion ( $D_{\perp}$ ) and kurtosis ( $W_{\perp}$ ) and parameters in the "LA" dataset. An important result was that the combination of axisymmetric DKI and Rician bias correction was the overall best approach for bias-free parameter estimation *in vivo*, reaching the  $A\text{-MPE} \leq 5\%$  threshold for SNRs of 15 ( $W_{\perp}$ ) and lower, with an SNR as low as 4 required for  $D_{\perp}$ <sup>55</sup>.

Figure 3 (Figure S2 from the Supporting Information of<sup>55</sup>) shows the SNR at which  $R\text{-STD} \leq 5\%$ . Again, a lower SNR would be favorable, since this would indicate a higher precision at a lower SNR. Figure 3 demonstrates that there is no improvement in precision associated with using axisymmetric DKI compared to standard DKI. This is documented by the fact that the  $R\text{-STD} \leq 5\%$  threshold was not reduced to lower SNRs when using axisymmetric DKI (see red barplots of Figure 3). In fact, axisymmetric DKI almost always required the same or higher SNRs for reaching the SNR threshold. Figure 3 furthermore demonstrated that there is no positive effect on the precision associated with Rician bias correction.

**Subsection discussion:** This study demonstrated that the combination of axisymmetric DKI and RBC was a highly effective "all-rounder" for reducing the Rician bias in all five AxTM since both algorithms synergistically compensate the weaknesses of each other. Interestingly, axisymmetric DKI did not have a higher precision compared to standard DKI despite its smaller parameter space. Furthermore, and as expected, RBC did not improve precision of parameter estimates since it is a method targeting the removal of a bias as opposed to improving precision which would be achieved by de-noising algorithms. The whole, published paper<sup>55</sup> and the Supporting Information can be found in the Appendix in Section 6.1.

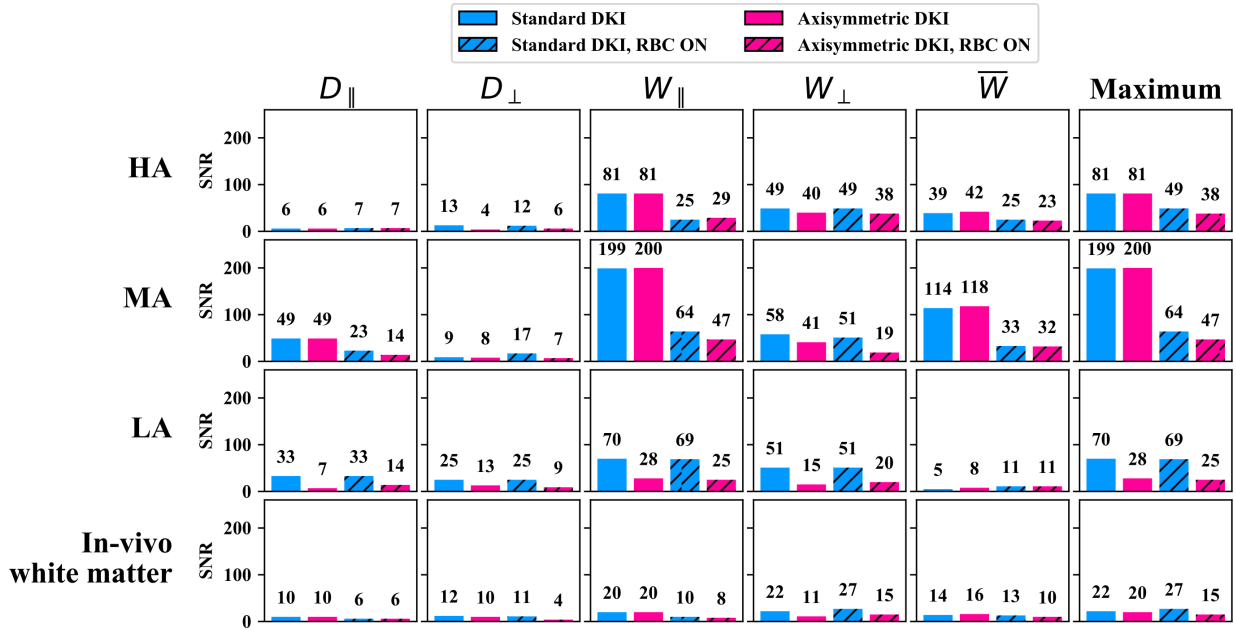


Figure 2: Shown is Figure 6 from<sup>55</sup>, showing the signal-to-noise ratio (SNR) above which the absolute value of the mean percentage error ("A-MPE") is  $< 5\%$  for the synthetic dataset with high, medium and low fiber alignment ("HA", "MA", "LA") and the "in-vivo white matter" dataset. For the "in-vivo white matter" dataset, the A-MPE was averaged across the 12 simulated *in vivo* voxels and the SNR above which this average A-MPE  $< 5\%$  is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve A-MPE  $< 5\%$  across all five AxTM.



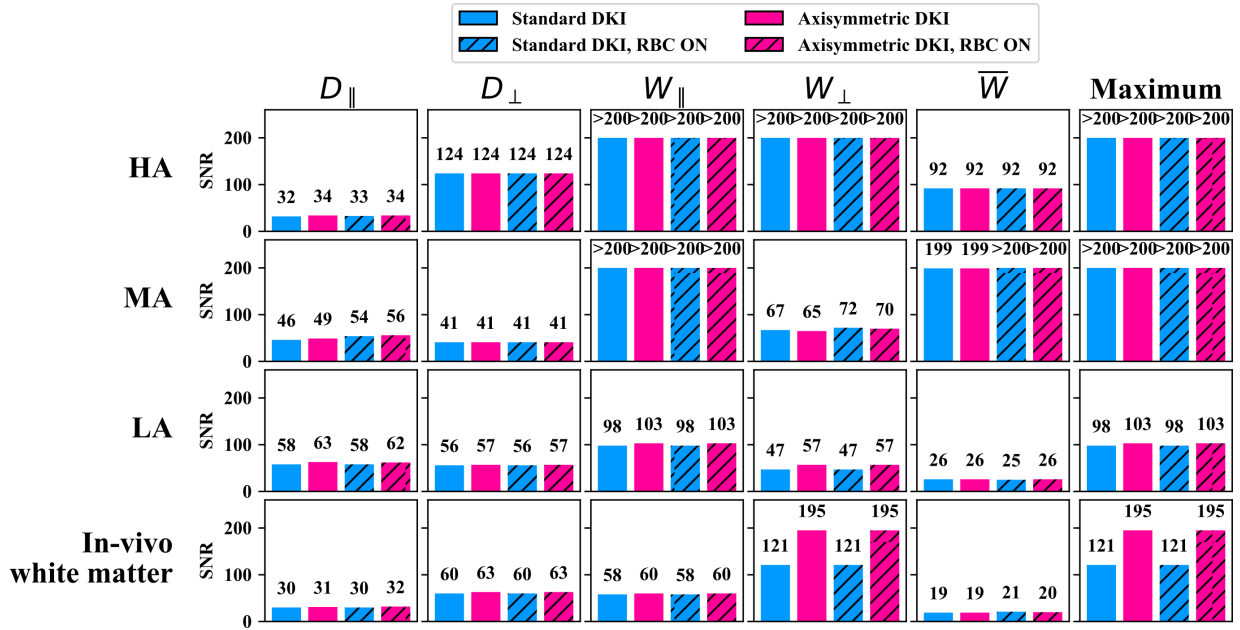


Figure 3: Shown is Figure S2 from<sup>55</sup>, showing the signal-to-noise ratio (SNR) above which the the standard deviation in reference to the ground truth ("R-STD") is  $< 5\%$  for the synthetic dataset with high, medium and low fiber alignment ("HA", "MA", "LA") and the "in-vivo white matter" dataset. For the "in-vivo white matter" dataset, the R-STD was averaged across the 12 simulated *in vivo* voxels and the SNR above which this average R-STD  $< 5\%$  is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve A-MPE  $< 5\%$  across all five AxTM.

### 3.1.2 Axisymmetric DKI inherent bias: Variance versus model inherent bias in the axisymmetric DKI tensor metrics

**Subsection introduction:** The additional, axisymmetry assumptions (see Section 2.1) of axisymmetric DKI could be violated, e.g., in voxels with complex fiber structures like fiber crossings. This violation would become visible as an apparent difference between standard DKI and axisymmetric DKI parameter estimates from the same data. This deviation is referred to as "axisymmetric DKI inherent bias" in this work. This inherent bias likely depends on the severity of the violation of the symmetry assumptions of the fiber structure in an image voxel. One of the main objectives of the study "Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates"<sup>56</sup> was to quantify the severity of the axisymmetric DKI inherent bias in a noise-free simulation of the white matter of a healthy human brain. Here, the inherent bias was quantified using the five AxTM. This quantification is important, since a substantial inherent bias would counteract axisymmetric DKI's benefits regarding the reduction of the Rician bias reported in Section 3.1.1. The study was also investigating if the relative number of affected voxels with a substantial inherent bias could be reduced by using voxel selection masks, filtering out white matter voxels with high fiber complexity. Finally, it was investigated whether the axisymmetric DKI inherent bias is visible under realistic, noisy conditions or if it is hidden by the effects of noise. All together, the addressed points answer question 3. from the introduction.

**Subsection methods:** To quantify the axisymmetric DKI inherent bias, noise-free simulations (see Section 2.3) of healthy, *in vivo* white matter were performed using standard DKI as a forward model for signal generation. Here, standard DKI as a forward model was chosen, because this explicitly allows for the inclusion of complex fiber structures in the simulation. Since no noise was added to the simulation data, the observed, apparent difference between AxTM estimates obtained with standard DKI and axisymmetric DKI could only be due to the different DKI models. The observed difference between both DKI variants was also referred to as "baseline difference". This baseline difference was computed according to the voxel-wise absolute percentage error (A-PE)  $A-PE = 100 \cdot \frac{\Theta_{standardDKI} - \Theta_{axisymmetricDKI}}{\Theta_{standardDKI}}$ , where  $\Theta$  is an AxTM and the subscript indicated whether that AxTM was estimated using standard DKI or axisymmetric DKI. Here, the ground truth reference were the noise-free standard DKI estimates of the AxTM. If the A-PE was greater than 5%, the corresponding voxel was classified as a "substantially differing voxel" (SDV). To see if the relative number of SDV could effectively be reduced by confining the analysis to voxels with lower fiber complexity, two voxel selection masks were introduced. Here, the first mask was filtering out voxels using a fractional anisotropy (FA) threshold derived from the FA of a unidirectional phantom<sup>58</sup>. The FA is a parameter obtainable with DKI or DTI quantifying the degree of anisotropy of diffusion, where a higher FA indicates more anisotropic diffusion associated with less complex, single fiber bundles. The second voxel selection mask was based on the "Westin indices"<sup>56;59;60</sup>.

Finally, to investigate if the axisymmetric DKI inherent bias is visible under realistic, noisy conditions, a simulation study Section 2.3 with noisy diffusion-weighted MRI signals of five well-known fiber tracts at SNR= 39 was performed. In addition to a standard dMRI protocol using the "full" standard number of acquired diffusion gradients, the fast "199" protocol<sup>61;11</sup> was simulated. The fast "199" protocol was specifically designed to leverage axisymmetric DKI's reduced data demand using the theoretical minimum number of diffusion gradient directions, which is a total of 18 diffusion weighted images and one non diffusion weighted image. Furthermore, in the evaluation of this study, the adaptive denosing algorithm "msPOAS"<sup>62;47;63</sup> was used.

**Subsection results:** The main result with respect to quantifying the axisymmetric DKI inherent bias in the whole white matter was that the diffusion metrics ( $D_{\parallel}$  and  $D_{\perp}$ ) and the mean of the kurtosis tensor ( $\overline{W}$ ) could be estimated with a very low number (only  $\approx 1\%$  to  $2\%$ ) of SDV, see plot A.) in Figure 5. However, for the parallel and perpendicular kurtosis  $W_{\parallel}$  and  $W_{\perp}$ , the number of SDV was high with 22% and 51%, respectively. However, here an important finding is that the relative number of SDV and the median inherent axisymmetric DKI bias could effectively be reduced by confining the analysis to voxels with lower fiber complexity via the voxel selection masks. Particularly the mask based on the "Westin indices"<sup>59;60;56</sup> was able to reduce the relative number of SDV for  $W_{\parallel}$  from 22% to 3% while it had only small effects in case of  $W_{\perp}$ .

With respect to the visibility of the inherent bias under realistic, noisy conditions, it was shown that the differences between the two DKI frameworks is very likely smaller than the effects of noise if a standard MRI acquisition protocol is used. This finding is demonstrated by Figure 4 (Figure 6 of<sup>56</sup>), which shows the median inherent bias obtained in major white matter fiber tracts for a typical SNR= 39 found in dMRI data<sup>64</sup>. Here, the results for standard DKI and axisymmetric DKI overlap almost exactly for almost every AxTM if a standard dMRI protocol is used. In this case, the only exception is  $W_{\perp}$  where a small but noticeable difference between both DKI variants was observed. However, if the fast "199" protocol was used, a substantial difference to both standard DKI and axisymmetric DKI estimates from the standard dMRI protocol was observed.

**Subsection discussion:** The axisymmetric DKI inherent bias was mostly confined to the parallel and particularly the perpendicular kurtosis, where  $\approx 50\%$  of voxels were found to be "substantially differing". However, at SNR= 39, the only difference between standard DKI and axisymmetric DKI when using a standard dMRI protocol was observed for  $W_{\perp}$ , demonstrating that under realistic, noisy conditions, the axisymmetric DKI inherent bias does not play a critical role for most AxTM and is hidden in the noise-induced Rician bias. The fact that only  $W_{\perp}$  differed noticeably under the influence of noise nicely matches the finding that  $W_{\perp}$  has also been the parameter with the most number of SDV and also had the highest median inherent bias across the whole white matter. It was also demonstrated that the inherent bias could be

reduced when using voxel selection masks confining the analysis to regions with less complex fiber structure, supporting the hypothesis that fiber complexity is one root cause of the observed differences.

Results of the fast "199 protocol" were more biased than the results of the standard dMRI protocol. This is likely due to a stronger effect of noise since in the fast "199" protocol significantly fewer diffusion gradient measurements are obtained, making parameter estimation more noise susceptible because fewer data points are fitted. Here it is important to note that the axisymmetric DKI inherent bias should be independent of the number of diffusion gradient directions, if the minimum number of gradients is met, so that the worse performance of the fast "199" protocol can only be attributed to a heightened noise susceptibility. The results in Figure 4 were obtained despite using Multi-shell Position-Orientation Adaptive Smoothing (msPOAS)<sup>62;54</sup> as a pre-processing which reduces the variance of the estimated parameters, making a potential axisymmetric DKI inherent bias easier to spot in the fit results.

As a summary, the axisymmetric DKI inherent bias will very likely not be visible under realistic, noisy conditions for the diffusion parameters and the mean of the kurtosis tensor while it might become visible for  $W_{\perp}$  but only for high SNRs and if a standard MRI protocol is used for imaging.

Another subject of the work was the propagation of the axisymmetric DKI inherent bias into the biophysical parameters of the WMTI-Watson model, this part of the work is presented in this thesis in Section 3.2.

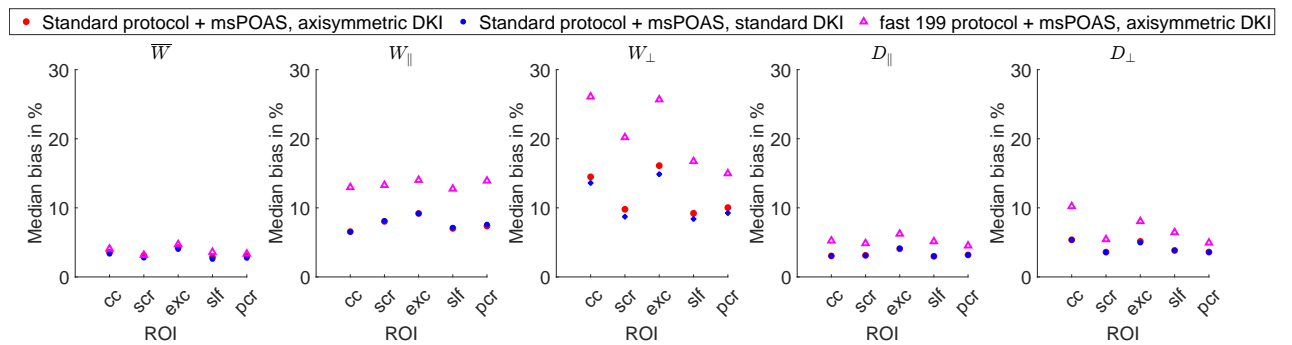


Figure 4: Shown is Figure 6 from<sup>56</sup>, showing the median total bias (axisymmetric DKI inherent bias and Rician bias) observed for noisy human brain data at an SNR = 39, smoothed with msPOAS in the five white matter fiber tracts corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf) and posterior corona radiata (pcr). The median bias was computed for a simulation of the standard protocol that was fitted with standard DKI (red data points) and axisymmetric DKI (blue data points) and for a simulation of the fast "199" protocol that was fitted with axisymmetric DKI (pink data points).

## 3.2 Biophysical models: Influence of noise and axisymmetric DKI inherent bias

Biophysical models potentially are highly valuable tools for non-invasive microstructure imaging. However, it has to be shown that they actually describe genuine, microscopic tissue properties by comparison with an established gold standard. Furthermore, even accurate biophysical models could be unuseful in practice if they are highly noise sensitive resulting in a low precision. Additionally, propagation of the axisymmetric DKI inherent bias into the biophysical parameters needs to be investigated, if axisymmetric DKI is used as a basis for estimation of the biophysical parameters. To address these question, in the following, a peer-reviewed study investigating the propagation of the axisymmetric DKI inherent bias into the parameters of the WMTI-Watson model is presented<sup>56</sup>. Furthermore, another peer-reviewed study investigating the accuracy of a variety of biophysical models<sup>53</sup> is shown. These findings of the peer-reviewed studies are also complemented by not yet published noise simulation studies that investigated how well biophysical models can estimate the biophysical parameters under the influence of noise, as well as their accuracy under the influence of noise.

### 3.2.1 Axisymmetric DKI inherent bias propagation into biophysical parameters

**Subsection introduction:** Bias-free axisymmetric DKI tensor metric (AxTM) estimates are a necessary prerequisite for accurate estimation of the biophysical parameters of the WMTI-Watson model. One part of the work<sup>56</sup> presented in 3.1.2 was dealing with the inherent bias in AxTM estimates, estimated with axisymmetric DKI. In this section, propagation of the inherent bias in the AxTM estimates into the biophysical parameters of the WMTI-Watson model is investigated, answering question 4. from the introduction.

**Subsection methods:** This study is part of the work "Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates"<sup>56</sup>. The methods regarding the performed simulation study were already presented in 3.1.2. Here, the AxTM estimates of the noise-free simulation obtained with axisymmetric DKI were used to compute the biophysical parameters of the WMTI-Watson model. These biophysical parameter estimates were then compared with the corresponding standard DKI based biophysical parameters estimates using the A-PE described in 3.1.2.

**Subsection results:** The study showed that there is a significant bias enhancement in biophysical parameter estimates of the WMTI-Watson model when the AxTM estimated with axisymmetric DKI are used. The number of SDV in the biophysical parameters were bigger compared to the AxTM from which they were computed with roughly one third ( $D_a$  and  $\kappa$ ) to half (AWF,  $D_{e,\perp}$ ,  $D_{e,\parallel}$ ) of the voxels being classified as SDV, as shown in Figure 5 B.1) (Figure 3 in the

original publication<sup>56</sup>). Furthermore, the median inherent bias in the biophysical parameters (11% to 18%) was also higher compared to the AxTM (7% to 11%). Again, the voxel selection via the "Westin indices" managed to reduce the number of SDV down to 10% for  $D_a$  and  $\kappa$  and roughly one third for AWF,  $D_{e,\perp}$  and  $D_{e,\parallel}$ .

**Subsection discussion:** All five AxTM are needed to compute the biophysical parameters of the WMTI-Watson model and the relationship is non-linear. This likely explains the bias enhancement observed in the biophysical parameters, since here biases in the AxTM are "stacked" upon each other. Another finding was that the inherent bias in the biophysical parameters could significantly be reduced by using voxel selection masks that were filtering out voxels with complex fiber structures, despite this voxel selection not being able to reduce the inherent bias of  $W_{\perp}$  in the AxTM, as reported in Section 3.1.2. However, the findings still urge caution when planning to use axisymmetric DKI for biophysical parameter estimation because even after applying a voxel selection mask, the number of SDV was still between  $\approx 10\%$  to  $30\%$ .

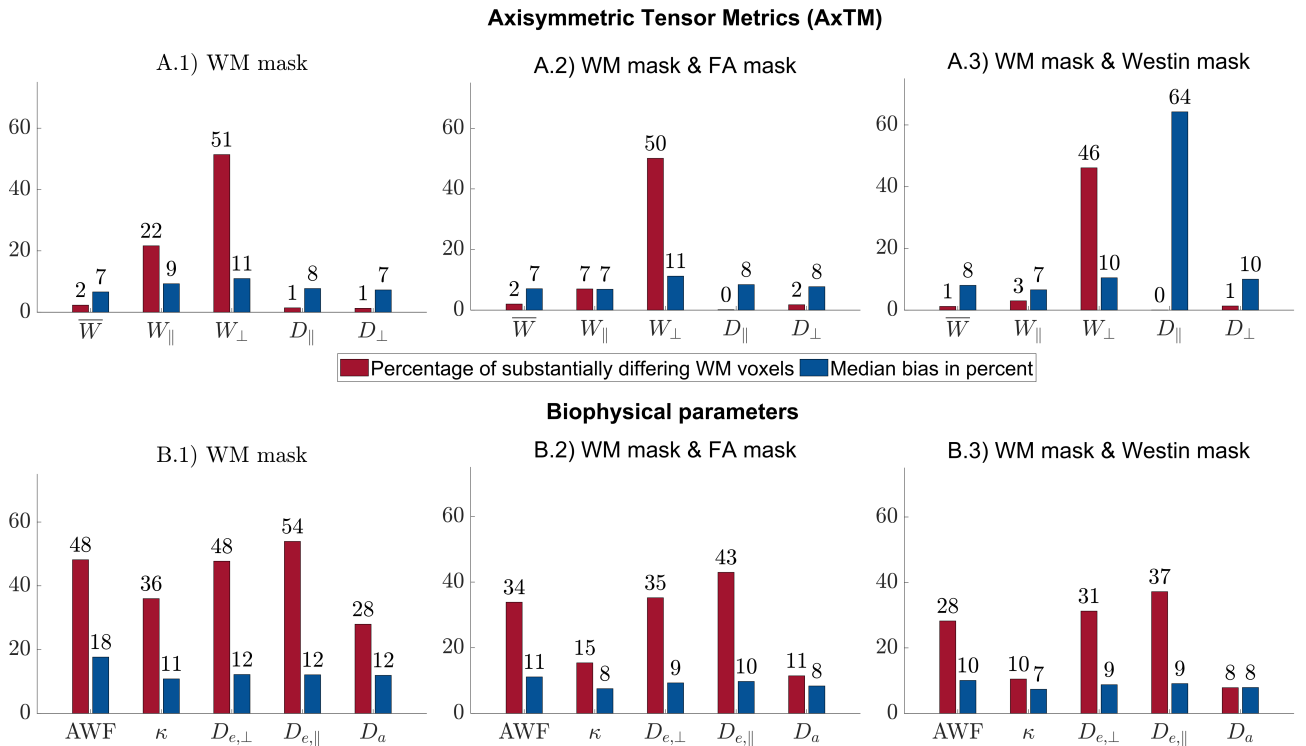


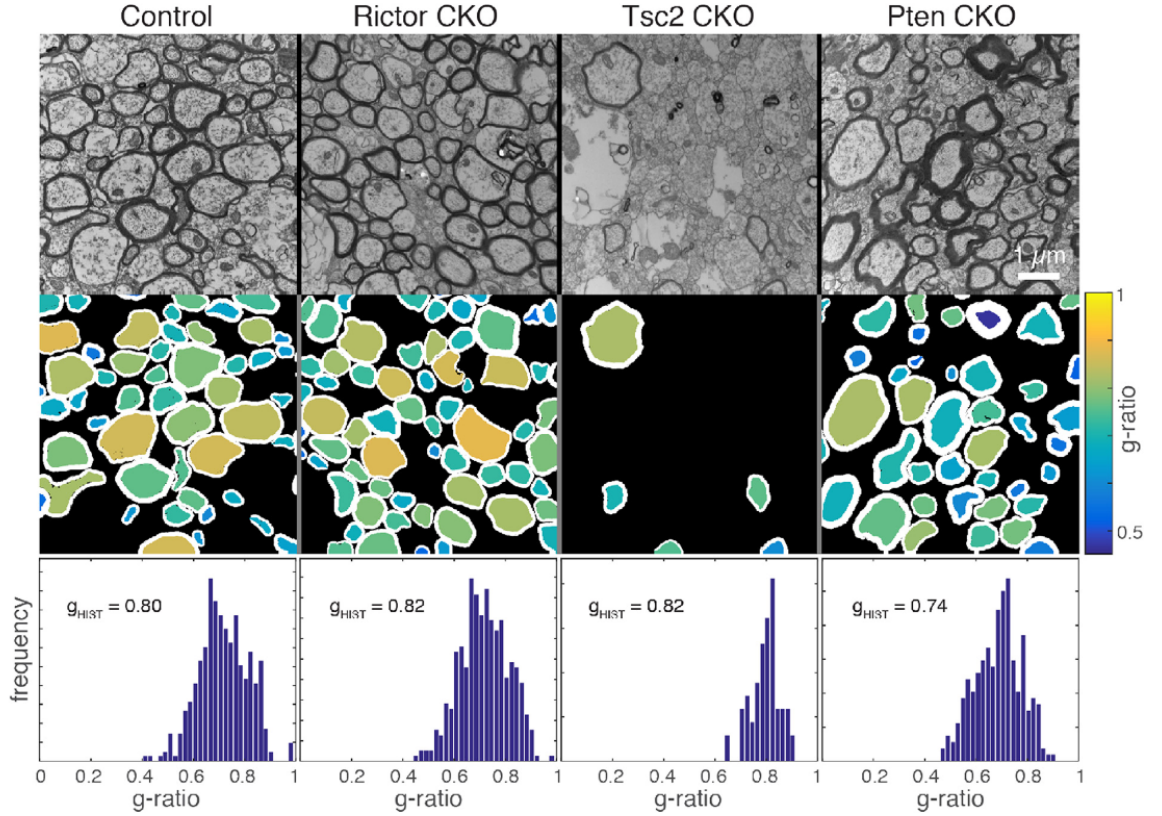
Figure 5: Shown is Figure 3 from<sup>56</sup>, showing barplots summarizing the number of substantially differing voxels (SDV) of the axisymmetric DKI tensor Metrics (AxTM) (top) and biophysical parameters (bottom), estimated in a noise-free simulation of health white matter. Shown are the number of SDV (red barplots) and the median difference ("bias") in those voxels (blue barplots). The subplot's titles indicate the subset of voxels that were analyzed, i.e., "WM mask" = white matter voxels were analyzed, "WM mask & FA mask" = white matter voxels in the FA mask were analyzed and "WM mask & Westin mask" = white matter voxels in Westin mask were analyzed. Number of voxels in the white matter mask: 101521, number of voxels in the WM mask & FA mask analysis: 28741, number of voxels in the WM mask & Westin mask analysis: 20527.

### 3.2.2 Estimation of biophysical parameters under the influence of noise

**Subsection introduction:** In this section, a study that is not yet peer-reviewed and published was carried out, investigating the performance of the biophysical models under the influence of noise. Particularly, estimation of all five biophysical parameters  $D_{e,\parallel}$ ,  $D_{e,\perp}$ ,  $D_a$ ,  $\kappa$  and AWF was investigated. Here, it is important to note that the ground truth with which the parameter estimates were compared were model based, i.e., the noise-free parameter estimates of each biophysical parameter were taken as a ground truth. This approach neglects any question regarding the actual accuracy of the biophysical models, e.g., with respect to an established EM gold standard (this comparison is done in Section 3.3.2). This study addresses question 5. raised in the introduction.

**Subsection methods:** To analyze the noise-susceptibility of the biophysical models and their underlying algorithms ranging from established, semi-analytical models like WMTI-Watson, NODDI and NODDI-DTI to machine learning based approaches like Baydiff and SMI, a noise simulation was performed. Noisy dMRI data were simulated using standard DKI as a forward model for SNRs= [5, 15, 30, 52, 100] for 100 noise samples, see Section 2.3. For this simulation study, data of a multimodal dMRI dataset of 15 mice published in<sup>53;65;66</sup> was used. The multimodal dataset consisted of 15 mice separated into two groups 1 and 2 for which there were dMRI data acquired with a 15.2 Tesla Bruker BioSpec scanner at  $150\mu\text{m}$  isotropic resolution<sup>65</sup> as well as EM histology data of both axonal volume fraction and myelin volume fraction for four regions of interest (ROIs). In this study, only the dMRI data in the ROIs were used for the simulation while the EM data were used later in the study presented in Section 3.3.2. The two groups of mice were defined by their level of myelination: group 1 consisted of controls with normal levels of myelination as well as mice with moderately higher ("Pten CKO") or lower ("Rictor CKO") levels of myelination while the second group 2 consisted of mice that were heavily hypomyelinated ("Tsc2 CKO"), i.e., they had a severe deficiency of myelin, see Figure 6. Group 2 was therefore specifically designed to investigate the effects of unmyelinated axons. Grouping of mice was motivated by similarities in their axonal volume fractions estimated with EM (see als Figure 2 of<sup>53</sup> in Section 6.3).

The simulated dMRI data were either fitted directly to estimate the biophysical parameters of SMI, BAYDIFF and NODDI or the DKI tensor, the axisymmetric DKI tensor metrics as well as FA and mean diffusivity (MD) were estimated based on standard DKI and then used to estimate the biophysical parameters of WMTI-Watson, WMTI and NODDI-DTI. Performance of the biophysical models was compared by computing the root mean square error between the biophysical model's fit results and the noise-free fit results of each biophysical model for every voxel in the simulated ROIs. The root mean square error was then normalized to the noise-free results to get the normalized root mean square error (NRMSE). Then, the NRMSE and its standard deviation over the ROI voxels and mice in groups 1 and 2 were plotted as a function



**Fig. 1.** (top) Representative electron microscopy from the MCC of four different mice, one of each type. (middle) The same images after segmentation, with each axon's fill-color defined by its g-ratio. (bottom) Histograms of g-ratio values of all MCC axons measured from the individual mice corresponding to the histology above. For reference, the  $g_{\text{HIST}}$  value is shown for each histogram, and the corresponding algebraic mean g-ratios are 0.72, 0.72, 0.80, and 0.69. For ease of comparison, each histogram has been scaled to have the same peak amplitude.

Figure 6: Figure 1 from<sup>65</sup> showing electron microscopy from the mid-body of the corpus callosum (MCC) of the four different mouse models "Control", "Rictor CKO", "Tsc3 CKO" and "Pten CKO" simulated in this study (top), the corresponding g-ratios (see, e.g., Section 3.3.1) for each axon after segmentation (middle) and histograms of the g-ratio of all MCC axons corresponding to the top image.

of SNR in Figure 7, where the axon water fraction (AWF), the parallel and perpendicular extra axonal diffusivity ( $D_{e,\parallel}$  and  $D_{e,\perp}$ ), the intra axonal diffusivity  $D_a$  and the axon alignment were evaluated.

The biophysical models estimated two different types of parameters describing axon alignment. First, the alignment parameter  $\kappa$  describing axon alignment in biophysical models using the Watson distribution and second,  $p_2$ , an anisotropy invariant used to describe the fiber orientation dispersion function<sup>19</sup>, see Table 1. While  $p_2$  lies between 0 (extremely low fiber alignment) and 1 (extremely highly aligned fibers),  $\kappa$  lies between 0 (extremely low fiber alignment) and  $\text{inf}$  (extremely highly aligned fibers). To make the NRMSE comparable in this case,  $\kappa$  can be used to compute the orientation dispersion index (ODI):  $\text{ODI} = \frac{2}{\pi} \arctan(\frac{1}{\kappa})$ . The ODI is bound between 0 (extremely highly aligned fibers) and 1 (extremely low fiber alignment). Furthermore, for the models that estimate  $p_2$ , first  $\kappa$  and then the ODI can be computed according to Equation (13) from<sup>20</sup> when assuming that the fiber orientation distribution function (fODF) is the Watson distribution:  $p_2 = \frac{1}{4}(\frac{3}{\sqrt{\kappa F(\sqrt{\kappa})}} - 2 - \frac{3}{\kappa})$  where F is Dawson's function. For WMTI



which is based on the assumption of extremely highly aligned fibers ( $\kappa = \text{inf}$ ), estimation of a fiber alignment is still possible using the three eigenvalues of the intra axonal diffusion tensor<sup>67</sup>. The quantity computed in this case is  $\langle \cos(\Theta) \rangle$  where  $\Theta$  is the angle of the fODF which can be used to calculate  $p_2$  according to  $\langle \cos(\Theta) \rangle = \frac{2p_2+1}{3}$ . However, WMTI is expected to become invalid in case of low fiber alignment. Figure 7 shows the results obtained for AWF,  $D_{e,\parallel}$ ,  $D_{e,\perp}$ ,  $D_a$  and the ODI.

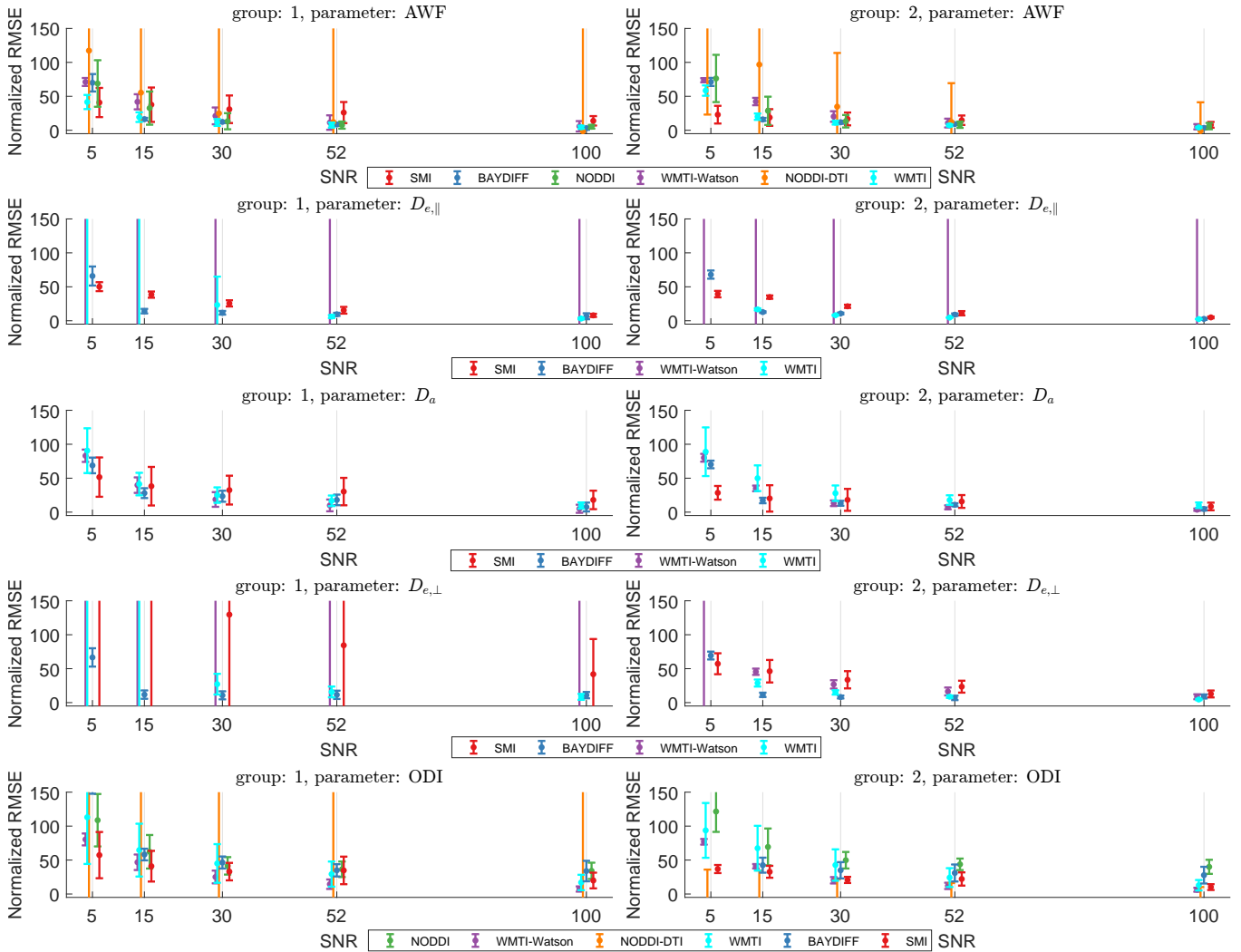


Figure 7: Results of biophysical model parameter estimation for simulation data based on the two groups of mice "1" and "2". The data points of each model were slightly shifted on the x-axis to make comparison easier. Shown is the average normalized root mean square error (NRMSE) and standard deviation across regions of interest (ROI) and the mice of each group for every simulated SNR. The groups of mice contained 12 (group 1) or 3 (group 2) mice

### Subsection results:

**Summary of most important points:** Baydiff was the overall best performing model across all parameters and SNRs and also had the smallest standard deviation across ROIs and mice. Interestingly, problems with negative and thus physically not meaningful  $D_{e,\parallel}$  values, estimated by WMTI-Watson and observed in group 1, were not present in group 2.

**AWF:** Baydiff and WMTI were the most precise biophysical models for AWF estimation in both group 1 and group 2 across all SNRs. SMI was least influenced by noise, i.e., its change across SNRs was smaller than other biophysical models, but also did not profit as much from higher SNRs. Interestingly, the standard deviation across the NRMSEs was reduced in Group 2 compared to Group 1. NODDI-DTI was heavily affected by outliers, meaning that in some voxels there were very high NRMSEs. For a very high SNR of 100, all models except NODDI-DTI performed similarly, while NODDI-DTI, still at this SNR, was heavily affected by outliers.

$D_{e,\parallel}$ : The machine learning based algorithms Baydiff and SMI had a significantly smaller standard deviation than WMTI and WMTI-Watson for small SNRs and Baydiff overall performed the best. Interestingly again, standard deviation over the NRMSE was reduced in group 2 and furthermore, negative, physically not meaningful WMTI-Watson based estimates of  $D_{e,\parallel}$  (caused by negative noise-free results) in group 1 became physically plausible, i.e., positive in group 2.

$D_a$ : Baydiff and WMTI-Watson were the two best performing models in both groups. Again, there was an improvement of standard deviation in Group 2.

$D_{e,\perp}$ : Baydiff was the best performing biophysical model across all SNRs and significantly outperformed the other models for SNRs 15 to 52 where, especially for low SNRs, SMI and WMTI-Watson had an extremely high standard deviation caused by outliers (Figure 7). This sensitivity to noise caused by outliers, again, was significantly reduced in Group 2.

**ODI:** WMTI-Watson and SMI were the two best performing models over all SNRs. Baydiff struggled for the very low SNR 5 where it had NRMSE values of  $\approx 200$  (group 1) and  $\approx 300$  (group 2). Furthermore, NODDI-DTI produced physically not meaningful, negative values indicating that the fit failed.

**Subsection discussion:** In this study, "Baydiff" a biophysical model whose parameters are estimated with a machine learning based algorithm was found to be the best performing model in terms of noise-robustness in both groups of mice. Also, Baydiff was shown to have less variance in the normalized root mean square error (NRMSE) across different ROIs and mice within both groups, indicating more consistent results.

### 3.3 Biophysical models: Assessing the accuracy under the influence of noise with a histology based gold standard

#### 3.3.1 Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy

**Subsection introduction:** Biophysical models are increasingly being used to estimate the axonal water fraction (AWF) which is a key biomarker for estimation of the axonal volume fraction (AVF) noninvasively. The axonal volume fraction in turn can for example be used for estimation of the g-ratio<sup>68;69</sup> which is a key biomarker for the neuronal conduction velocity. However, accurate transformation of the AWF to the AVF requires thorough calibration with an established gold standard like, e.g., electron microscopy (EM). Accurate calibration in turn requires tackling two hurdles. First, it is important to account for the differential sensitivity of dMRI and EM to unmyelinated axons. Here, dMRI is sensitive to unmyelinated axons while EM typically is not. Neglecting the fraction of unmyelinated axons during calibration introduces a bias in the AVF calibration parameters reducing the accuracy of biophysical models based AVF estimation. Second, differences in compartmental relaxation times are often neglected in dMRI models which, again, leads to a possible bias in the AVF calibration parameters.

The study<sup>53</sup> set out to investigate the effects of introducing an offset and a scaling when calibrating biophysical model based AVF estimates using an EM based gold standard. The multimodal mouse dataset used in this study was already introduced in Section 3.2.2 and consisted of two groups of mice 1 and 2 containing mouse models with different levels of myelination. Particularly, group 2, consisting of severely hypomyelinated, i.e., demyelinated mice was a prime example designed for investigation of the effects and contributions of unmyelinated axons to the calibration process. Amongst other things, the study allowed for a biological interpretation of the calibration parameters. This study addressed the part of question 6. raised in the introduction regarding the accuracy of commonly used biophysical models.

**Subsection methods:** A variety of biophysical models, see Table 1, were used to estimate the AWF and then calibrated using an EM based AVF gold standard. Performance of the combination of calibration parameters was assessed using the Bayesian information criterion (BIC). To this end  $\Delta\text{BIC}$  was calculated as the difference between the calibrated (baseline) BIC and the BIC obtained from a calibrated biophysical model. Accuracy of the calibrated MRI based AVF estimation was quantified using the Bland-Altman metrics bias ( $\delta$ ) and error according to:  $\delta = \text{AVF}_{\text{MRI}} - \text{AVF}_{\text{EM}}$  and error as 1.96 times the standard deviation of  $\delta$ . In order to calculate these metrics, the MRI based axonal volume fraction ("AVF<sub>MRI</sub>") was computed according to  $\text{AVF}_{\text{MRI}} = (1 - \text{MVF}_{\text{EM}}) \cdot s \cdot \text{AWF} - U$  where  $\text{MVF}_{\text{EM}}$  is the EM based myelin volume fraction (MVF) and AWF was estimated using different biophysical models, see Table 1. The offset  $U$  and scaling  $s$  were estimated by minimizing the residual sum of squares between  $\text{AVF}_{\text{MRI}}$  and

the EM histology based AVF counterpart  $AVF_{EM}$ .

**Subsection results:** The study<sup>53</sup> demonstrated that a linear calibration with an offset and a scaling can improve comparability of AVF estimated using dMRI based AWF estimates and the EM based AVF gold standard. Figure 8 (Figure 3 in the original publication<sup>53</sup>) shows the contribution of the calibration parameters to the improvement of comparability between the EM based AVF gold standard and DWI based AVF estimates based on different biophysical models using  $\Delta BIC$ . Here, a lower  $\Delta BIC$  indicates less information loss and is desirable because the observed data are better explained by the given biophysical model. Figure 8 shows that the combination of U2, the offset of the severely hypomyelinated mice of group 2, together with a scaling  $s$  had the lowest  $\Delta BIC$  for most models, except for NODDI-DTI where only using a scaling  $s$  and no offset had the lowest BIC. Furthermore, the Baydiff model had the lowest  $\Delta BIC$ . It was also found that, biophysical models like WMTI, Baydiff, WMTI-Watson and SMI that are fitting the compartmental diffusivities were performing better than NODDI and NODDI-DTI that are using fixed diffusivities (see Table 1). In terms of the Bland-Altman metrics bias and error, WMTI was the best performing model, closely followed by Baydiff and SMI.

**Subsection discussion:** Within the context of this work the AWF plays an important role since it can be used to calculate the axon volume fraction. The axon water fraction (AWF) can be approximated via biophysical models with the signal fraction  $f$  of the intra axonal compartment. The AWF and signal fraction  $f$  often are assumed to be equivalent under the assumption that the small contribution of the myelin water and differences in the compartmental  $T_2$ <sup>23</sup> can be neglected. Here one important finding was that the compartmental  $T_2$  differences could be accounted for by a scaling in the calibration process. Furthermore, the offset can be interpreted as accounting for the volume fraction of unmyelinated axons. Hence, both calibration parameters carry biologically interpretable meaning. The whole, peer-reviewed paper<sup>53</sup> can be found in the Appendix of this thesis in Section 6.3.

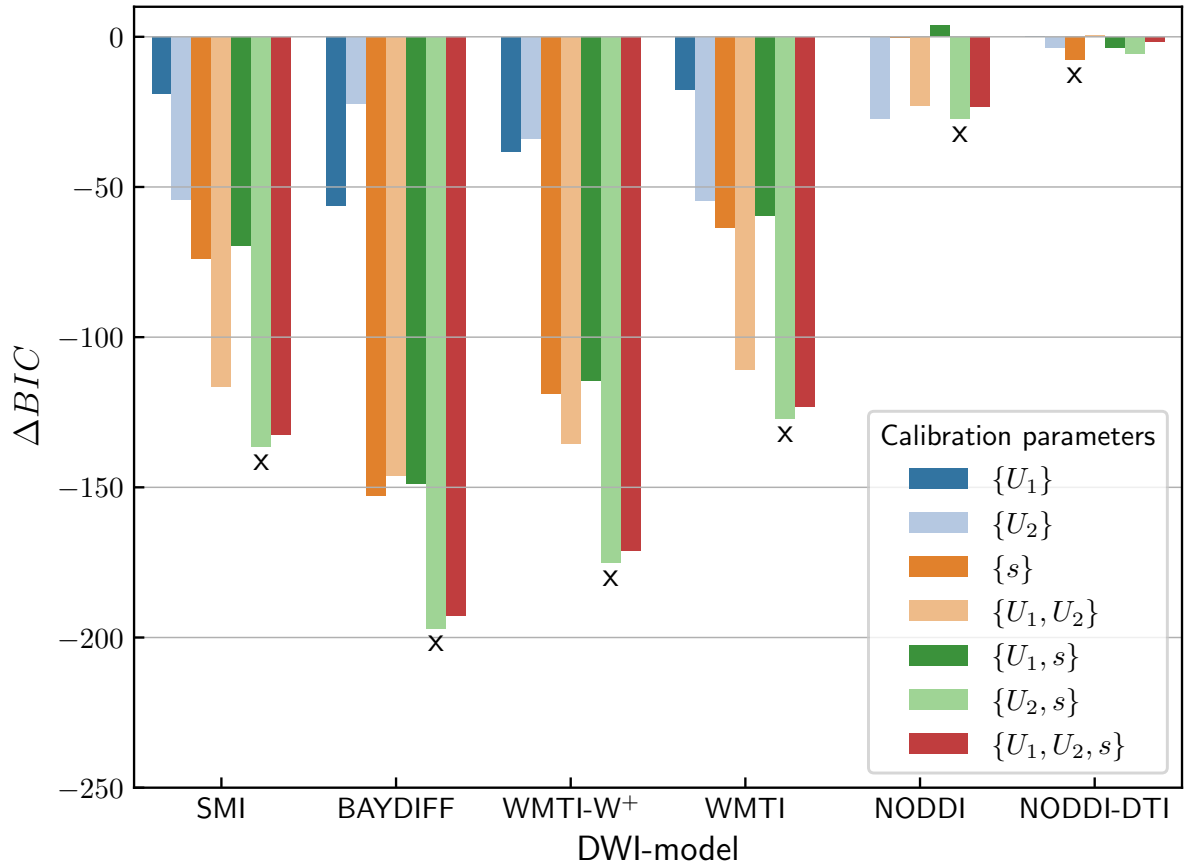


Figure 8: Shown is Figure 3 from<sup>53</sup>, showing the contribution of calibration parameters to DWI model improvement where 1 and 2 are the two offsets adapted to two groups of mice with different levels of myelination and  $s$  is the scaling. Shown are the differences  $\Delta BIC$  with respect to the parameter combination with the smallest BIC in each DWI model. A lower  $\Delta BIC$  value indicates better model performance. Values for the uncalibrated case (i.e.,  $AWF \equiv AVF$ ) served as baseline, that is, for this case  $\Delta BIC=0$ . An  $x$  indicates the calibration parameter combination with the largest evidence of improvement with respect to the baseline without further calibration for each DWI model, respectively.

### 3.3.2 Accuracy of MRI based axonal volume fraction estimates of various biophysical models under the influence of noise

**Subsection introduction:** The study reported in Section 3.2.2 was modified to assess how well the biophysical models reproduce an actual electron microscopy (EM) based, established AVF gold standard under the influence of noise. Hence, this study addresses the noise part of question 6 from the introduction but also further investigated the accuracy of the biophysical models. Here, the methods and results of this not yet published and peer-reviewed or published study are presented and discussed.

**Subsection methods:** The noisy AVF estimates of Section 3.2.2 obtained using a variety of

biophysical models were used to estimate the MRI based axonal volume fraction ("AVF<sub>MRI</sub>") for SNRs= [5, 15, 30, 52, 100]. To compute the MRI based AVF, AWF estimates of each biophysical model were averaged in each ROI (see Section 3.2.2). These average AWF markers were then used to compute AVF<sub>MRI</sub> using the best calibration parameters  $s$  (scaling) and  $U2$  (offset for group 2) from the work reported in Section 3.3.1. The values of these best calibration parameters can be found in<sup>53</sup> Table 4. The AVF was computed according to  $AVF_{MRI} = (1 - MVF_{EM}) \cdot s \cdot AWF - U2$ <sup>53</sup>, for all biophysical models except for NODDI-DTI that did not use an offset  $U$ , i.e.,  $U2 = 2$ . Here,  $MVF_{EM}$  is the EM based myelin volume fraction. Then, the NRMSE was calculated using the EM histology AVF estimates as a gold standard, see Section 3.3.1 and<sup>53;65;66</sup>. This gold standard was both used as a reference for calculation of the RMSE and for normalization to get the NRMSE. The results of this study, i.e., the SNR dependent NRMSE of the AVF estimates of each biophysical model, are plotted in Figure 9. In addition to the noisy NRMSE results, the average noise-free NRMSE across mouse specimen and ROIs were plotted in Figure 9 to be able to see the noise-free accuracy of AVF<sub>MRI</sub>.

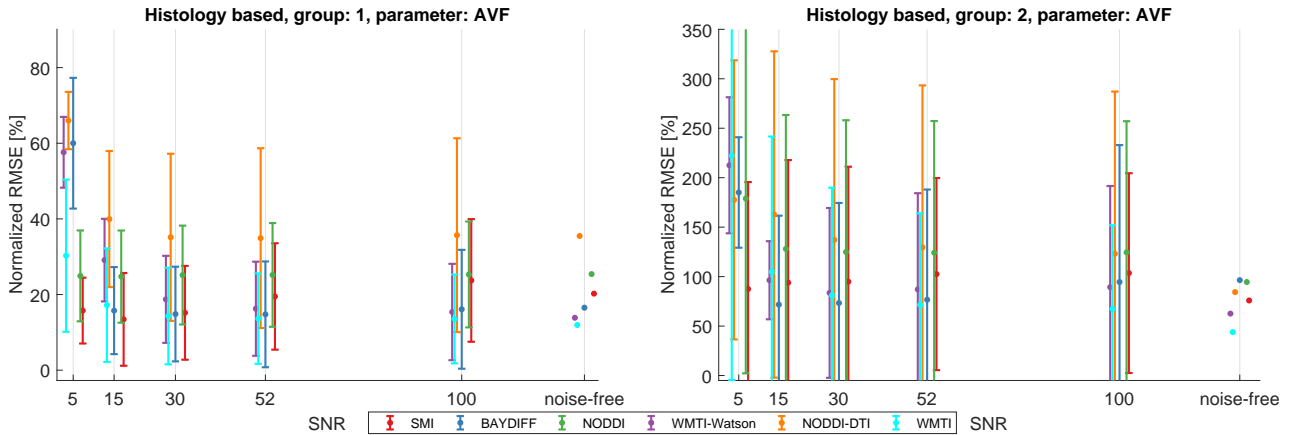


Figure 9: Normalized root mean square error (NRMSE) of MRI based AVF estimates computed for the two mouse groups 1 and 2. Shown is the average normalized root mean square error (NRMSE) and standard deviation across regions of interest (ROI) and the mice of each group for every simulated SNR. The groups of mice contained 12 (group 1) or 3 (group 2) mice. Data points of each model were slightly shifted on the x-axis to make comparison easier. Note the different scaling of the y-axis of group 1 (left) and 2 (right).

**Subsection results:** A big difference between the two mouse groups (1 and 2) was observed. In group 1, WMTI ( $\approx 12\%$ ) and WMTI-Watson ( $\approx 14\%$ ) were the best performing models in terms of noise-free NRMSE, closely followed by Baydiff with an NRMSE of  $\approx 16.5\%$ . Similar to the model based noise simulation, SMI was least influenced by noise even for low SNRs and its mean value was very similar across all simulated SNRs. NODDI-DTI performed the worst in terms of average, noise-free NRMSE in group 1. However, both the spread and average of the noise-free NRMSE across all biophysical models were significantly smaller in group 1 (average  $\approx 20.5 \pm 8.7\%$ ) compared to group 2 (average  $\approx 76.2 \pm 20.2\%$ ). The noise simulation results in mouse group 2 were characterized by higher average NRMSE across all SNRs and roughly 10 times bigger standard deviations across ROIs and mice even though group 2 contained only

three mice, i.e., it was only a quarter of the size of group 1, which contained 12 mice. In terms of noise-free NRMSE in group 2, both WMTI and WMTI-Watson were still the best performing models, while here Baydiff was the worst performing model.

### **Subsection discussion:**

When evaluating the accuracy of the biophysical models, a clear distinction between both investigated groups of mice was observed. In group 1, WMTI (NRMSE =  $\approx 12\%$ ) and WMTI-Watson (NRMSE =  $\approx 13.9\%$ ) were the best performing models, i.e., they were most accurate with respect to the actual EM based gold standard in the noise-free case, closely followed by Baydiff (NRMSE =  $\approx 16.5\%$ ). Interestingly, in group 1, the results of all investigated biophysical models only changed slightly for  $\text{SNR} \geq 30$ , i.e., they were already close to the noise-free results at  $\text{SNR}=30$ , which is encouraging, since this is close to typical SNRs found in dMRI data<sup>64</sup>. Even the variance across NRMSE estimates did not change significantly anymore after this SNR, indicating consistently good results across ROIs and mice. The finding that WMTI was the most accurate biophysical model is consistent with the findings reported in<sup>53</sup>, see Section 3.3.1. However, it is important to note that the focus of the study were areas that are known to host many highly aligned fiber bundles like the corpus callosum and the fornix which are well suited for WMTI's assumption of straight fibers. WMTI might therefore not perform as well when imaging the whole brain including brain areas where fibers are not highly aligned.

In group 2 the NRMSE was significantly worse for all investigated biophysical models (WMTI-Watson NRMSE =  $\approx 44\%$ , WMTI-Watson NRMSE =  $\approx 63\%$ , Baydiff NRMSE =  $\approx 96\%$ ). Additionally, there was a high variance across ROIs and mice. Judging from the noise-free results, WMTI-Watson and WMTI were still the best performing models while Baydiff was the worst performing model in this group. One reason for the generally worse performance of all biophysical models in group 2 could be the violation of a fundamental assumption of the biophysical models which is that there is no exchange happening between intra and extra axonal compartments (see Section 2.4). The myelin sheath is the main substance that separates the intra and extra axonal compartments from one another and hinders exchange. Therefore, in case of severe demyelination, exchange between compartments is likely to occur which violates fundamental assumptions of biophysical models possibly rendering them invalid.

Taking into account the performance across SNRs and biophysical parameters reported in Section 3.2.2, Baydiff<sup>21</sup> was the best performing model. Baydiff also ranked amongst the best in terms of accuracy measured with respect to the EM based histological gold standard in healthy to moderately hyper or hypo myelinated tissue. Here, it is very encouraging that the algorithm was used on *ex vivo* mouse data by adjusting the ranges for biophysically plausible parameters used for training and was still very accurate when compared to an actual EM histological gold standard, i.e., the Baydiff model generalizes well. Furthermore, its good performance even for low SNRs is likely due to implicitly learning the Rician distribution of the noisy training

data. However, given the findings regarding group 2, Baydiff (and all other biophysical models) struggled to perform as well in terms of accuracy as in group 1, indicating that it might not be useful for imaging of pathologies in its current state and would need to be retrained on specialized data.

## 4 Discussion

Standard DKI and axisymmetric DKI are promising dMRI frameworks that hold the potential for clinical applications, especially if they are used together with biophysical models for computation of specific, biophysical parameters. However, noise and the axisymmetric DKI inherent bias pose a potential hurdle for their application. This thesis provided insights into biases and precision of standard DKI and axisymmetric DKI as well as the accuracy and performance of biophysical models under the influence of noise. Through this, possibly, contributions to moving one step closer to a clinical application of these imaging techniques could be made.

Turning to research questions 1. regarding the Rician bias raised in the introduction, an important finding is that the combination of axisymmetric DKI and Rician bias correction synergistically mitigate effects of the Rician bias and improve each others weaknesses. While the investigated Rician bias correction method was particularly efficient for estimation of the parallel parameters and the mean of the kurtosis tensor, axisymmetric DKI was superior to standard DKI for estimation of the perpendicular parameters. Together, the combination of axisymmetric DKI and Rician bias correction enabled bias-free parameter estimation for all AxTM for an SNR as low as 15. Regarding the precision of axisymmetric DKI (question 2. from the introduction), it was found that interestingly, axisymmetric DKI itself was not more precise than standard DKI, despite its smaller parameter space. Question 3. from the introduction was dealing with the axisymmetric DKI inherent bias. Here, it is encouraging that it was demonstrated that the axisymmetric DKI inherent bias was mostly confined to the parallel and particularly the perpendicular kurtosis. Furthermore, under noisy conditions, this inherent bias is not visible since the effects of noise typically dominate parameter estimation under real life circumstances. Additionally, it was demonstrated that confining the analysis to voxels with a lower level of fiber complexity achieved a significant reduction of the relative number of parallel kurtosis voxels with a substantial inherent bias. The next question (4.) was dealing with effects of the axisymmetric DKI inherent bias in the AxTM on the biophysical parameters of the WMTI-Watson model. Here, the axisymmetric DKI inherent bias noticeably propagated into all five biophysical parameters of the WMTI-Watson model, where more voxels were affected by this bias compared to the AXTM. This finding urges caution when planing to use axisymmetric DKI for estimation of the biophysical parameters. The final questions addressed in this thesis were dealing with the performance of the biophysical models under the influence of noise (ques-



tion 5.) and the accuracy (question 6.). Regarding the noise-susceptibility, the machine learning based Baydiff algorithm was the overall best performing model, closely followed by SMI which is also a machine learning based biophysical model. Amongst the "classical", (semi-) analytical biophysical models, WMTI performed best under the influence of noise. This finding is furthermore supported by WMTI, WMTI-Watson and Baydiff being the most accurate biophysical models in healthy tissue when comparing it to an actual, electron microscopy based gold standard. When taking into account both the accuracy and the precision under the influence of noise (question 6.) (see Figure 1), Baydiff was the overall best performing biophysical model.

#### **4.1 Overcoming biases in standard DKI and axisymmetric DKI and possible implications for practical applications**

Real life measurements using MR scanners are always affected by noise stemming from the electrical components of the MR scanner system. This noise interferes and competes with the desired MR signal, posing a significant challenge to signal detection and analysis. Noise affects the measured signals and thereby shows up as variance in the estimated parameters from these signals lowering the precision of an estimator. In MRI images, noise additionally leads to a systematic over or underestimation ("Rician bias") of parameters if magnitude signals are fitted, which is often standard operating procedure in diffusion MRI and clinical practice. Therefore, quantification of possible biases in standard DKI and axisymmetric DKI, as well as assessment of their performance under the influence of noise is a necessary prerequisite for a possible clinical application of these imaging techniques.

Here, this thesis provided insights into approaches on how to mitigate the effects of both the noise induced Rician bias as well as the axisymmetric DKI inherent bias. The Rician bias in all five axisymmetric DKI tensor metrics (AxTM) could effectively be reduced by using a combination of axisymmetric DKI with Rician bias correction. This combination of algorithms synergistically worked together and enabled bias-free AxTM estimation for SNRs of 15 down to as low as 4, depending on the investigated parameter. This is encouraging, since SNRs of DKI measurements are typically higher (e.g., 39<sup>64</sup>), producing an "SNR gain" that could be used to favorably adjust other imaging parameters that typically go along with an SNR reduction. Here, examples would be an increase of image resolution, an increase of contrast through higher diffusion weighting or reduction of scan time.

However, for this combination to be useful, it had to be shown that these positive findings are not counteracted by a significant axisymmetric DKI inherent bias caused by a violation of its additional symmetry assumptions, e.g., in tissues with complex fiber structures. Here, it was demonstrated that the axisymmetric DKI inherent bias was confined mostly to  $W_{\parallel}$  and  $W_{\perp}$ . Furthermore, the relative number of affected voxels could efficiently be reduced by confining

the analysis to brain regions with lower fiber complexity, supporting the hypothesis that fiber complexity is in fact one cause of the observed inherent bias in axisymmetric DKI. However, it was also shown that under realistic, noisy conditions, the Rician bias typically is the dominating bias in parameter estimation. Therefore, the axisymmetric DKI inherent bias will be invisible in practical applications in most cases. Here, the only exception is  $W_{\perp}$  for which the inherent bias was visible even under the influence of noise. Furthermore, the inherent bias in  $W_{\perp}$  could not significantly be reduced by confining the analysis to brain regions with lower fiber complexity, urging caution with respect to this parameter.

## 4.2 Biophysical modeling under uncertainty and how to navigate it

Biophysical parameters of the standard model offer valuable insights for tracking the progress of neuro-degenerative diseases, e.g., going along with axonal and dendritic loss (AWF<sup>23;24;25;26;27</sup>) or demyelination ( $D_{e,\perp}$ <sup>23;28;24;29</sup>). However, noise can distort parameter estimates to a point where even a highly accurate biophysical model would not be useful anymore because the precision would be too low. Furthermore, biophysical parameters of models like WMTI, WMTI-Watson and NODDI-DTI are based on the AxTM. If these AxTM are estimated using axisymmetric DKI, the axisymmetric DKI inherent bias in the AxTM estimates can be enhanced in the biophysical parameters. Therefore, understanding bias propagation into the biophysical models, as well as their behavior under the influence of noise is crucial for a more widespread application.

Overall, it was found that Baydiff was the best performing biophysical model, taking into account performance under the influence of noise as well as accuracy in healthy or moderately hyper or hypo myelinated mouse tissue. Here it is encouraging that the accuracy was tested on *ex vivo* mouse data, because it showed that Baydiff, a machine-learning based algorithm, generalizes well since only its training boundaries were adjusted to the *ex vivo* case. However, accuracy of the Baydiff model deteriorated significantly in the mouse group consisting of severely demyelinated mice, indicating the need for specialized retraining on data containing these pathologies. Purely judging by the accuracy, the (semi) analytical biophysical models WMTI and WMTI-Watson performed slightly better in terms of normalized root mean square error (NRMSE), however, these models were more sensitive to the effects of noise. Furthermore, WMTI comes with the assumption of straight fibers in nervous tissues, possibly limiting its applicability.

Additionally, WMTI-Watson and WMTI are computed from the AxTM. Here, a significant bias enhancement, affecting all biophysical parameters of the WMTI-Watson model was demonstrated when using axisymmetric DKI to compute them. Similar to the applicability of WMTI discussed above, brain regions with highly aligned fibers are more likely to conform with ax-

isymmetric DKI's additional symmetry assumptions. Confining the analysis to areas with lower fiber complexity is therefore less likely to result in a significant axisymmetric DKI inherent bias which was demonstrated in Figure 5. Here, an introduction of a fiber selection mask could reduce the relative number of voxels with an axisymmetric DKI inherent bias in the biophysical parameters of the WMTI-Watson model significantly, offering a hands-on approach for dealing with the axisymmetric DKI inherent bias. Still, the findings urge caution when planning to use axisymmetric DKI or possibly low SNR data with a high Rician bias as a basis for estimation of the biophysical parameters. Particularly the extra axonal diffusivities were shown to be highly noise sensitive when estimated with WMTI-Watson (as reported in Figure 7) and severely affected by the axisymmetric DKI inherent bias.

## 5 Conclusion

Non-invasive imaging of the microstructure of nervous tissue is crucial for understanding neurological disorders. Here, DKI and biophysical models are a promising, non-invasive *in vivo* alternative to *ex vivo* microscopy for microstructure imaging. However, noise poses a critical hurdle for parameter estimation of both DKI and biophysical models.

It was demonstrated that the combination of axisymmetric DKI with Rician bias correction is a highly effective combination of algorithms for reducing the SNR requirements for Rician bias-free estimation of all five axisymmetric DKI tensor metrics. In practice, this "SNR gain" could be used to favorably optimize MRI acquisition parameters like resolution or scan time that typically lower the SNR. With respect to axisymmetric DKI, it was shown that the axisymmetric DKI inherent bias is confined to the parallel and particularly the perpendicular kurtosis parameters. Furthermore, voxel selection masks designed to filter out voxels with complex fiber structures could reduce the number of substantially differing voxels, suggesting that fiber complexity is one cause of the axisymmetric DKI inherent bias. Additionally, the effects of noise likely overshadow the axisymmetric DKI inherent bias under realistic conditions. Finally, Baydiff was shown to be the overall best performing biophysical model taking into account precision of the biophysical parameter estimates and accuracy of the axonal volume fraction measured with an actual electron microscopy gold standard. All together, these findings provide insights into possible biases in DKI and axisymmetric DKI encountered in practice while providing hands-on ways of mitigating them. Furthermore, the findings highlight the potential for Baydiff to become a reliable tool in dMRI based microstructure imaging.

## **6 Appendix**

### **6.1 Paper 1: Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study**

The scientific article "Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study" was published in "Magnetic Resonance in Medicine" (Wiley) in 2022.

## RESEARCH ARTICLE

# Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study

Jan Malte Oeschger<sup>1</sup>  | Karsten Tabelow<sup>2</sup>  | Siawoosh Mohammadi<sup>1,3</sup> 

<sup>1</sup>Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>2</sup>Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

<sup>3</sup>Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

## Correspondence

Jan Malte Oeschger, Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany.  
Email: [j.oeschger@uke.de](mailto:j.oeschger@uke.de)

## Funding information

Bundesministerium für Bildung und Forschung, Grant/Award Numbers: 01EW1711A, 01EW1711B; Deutsche Forschungsgemeinschaft, Grant/Award Numbers: MO 2397/4-1, MO 2397/5-1, MO 2397/5-2

Correction added after online publication 12 December 2022. Due to a production error, the following changes were not incorporated before publication: The delta symbol in the paragraph preceding equation 3 is restored to  $\delta xy$ ; one of the 'g vector' symbols of Figure 4 has been corrected; the author's name in the 'How to cite this article' information has been corrected.

**Purpose:** To compare the estimation accuracy of axisymmetric diffusion kurtosis imaging (DKI) and standard DKI in combination with Rician bias correction (RBC).

**Methods:** Axisymmetric DKI is more robust against noise-induced variation in the measured signal than standard DKI because of its reduced parameter space. However, its susceptibility to Rician noise bias at low signal-to-noise ratios (SNR) is unknown. Here, we investigate two main questions: first, does RBC improve estimation accuracy of axisymmetric DKI?; second, is estimation accuracy of axisymmetric DKI increased compared to standard DKI? Estimation accuracy was investigated on the five axisymmetric DKI tensor metrics (AxTM): the parallel and perpendicular diffusivity and kurtosis and mean of the kurtosis tensor, using a noise simulation study based on synthetic data of tissues with varying fiber alignment and in-vivo data focusing on white matter.

**Results:** RBC mainly increased accuracy for the parallel AxTM in tissues with highly to moderately aligned fibers. For the perpendicular AxTM, axisymmetric DKI without RBC performed slightly better than with RBC. However, the combination of axisymmetric DKI with RBC was the overall best performing algorithm across all five AxTM in white matter and axisymmetric DKI itself substantially improved accuracy in axisymmetric tissues with low fiber alignment.

**Conclusion:** Combining axisymmetric DKI with RBC facilitates accurate DKI parameter estimation at unprecedented low SNRs ( $\approx 15$ ) in white matter, possibly making it a valuable tool for neuroscience and clinical research studies where scan time is a limited resource. The tools used here are available in the open-source ACID toolbox for SPM.

## KEYWORDS

axisymmetric DKI, microscopic fiber alignment, noise, Rician bias correction, simulation

## 1 | INTRODUCTION

Diffusion weighted MRI is an in-vivo imaging modality used in neuroscience and clinical research. It is sensitive to changes in nervous tissues that, for example, go along with neurodegenerative diseases like epilepsy and multiple sclerosis.<sup>1,2</sup> Diffusion MRI measures the net diffusion of nuclear spins of hydrogen nuclei in water molecules that are omnipresent in nervous tissue.

Diffusion of water molecules within the microstructural tissue landscape can be arbitrarily complex. A data-efficient method that captures both standard Gaussian diffusion and more complex restricted diffusion processes (e.g., due to diffusion of water trapped in the cell-body of axons), is the recently introduced axisymmetric diffusion kurtosis imaging (DKI) framework.<sup>3,4</sup> Its data-efficiency stems from requiring only eight parameters due to assuming axisymmetrically distributed axons, instead of 22 parameters like standard DKI.<sup>5</sup> This is likely a reasonable assumption in major white matter fiber bundles.<sup>3</sup> Furthermore, this is expected to make axisymmetric DKI less susceptible to the noise induced variation of the acquired diffusion MRI signals.

Noise in MRI images introduces a random variation into the measured diffusion signals and a bias for the estimated DKI parameters when the signal-to-noise-ratio<sup>6</sup> (SNR) is low. This bias is known as the “Rician noise bias”<sup>7-9</sup> and becomes more severe, the lower the SNR is. Diffusion MRI is prone to a low SNR because it generates image contrast from additional spin dephasing associated with water mobility leading to a signal attenuation. DKI is even more susceptible to the Rician noise bias compared to conventional diffusion tensor imaging, since estimating the DKI parameters requires multiple diffusion shells including higher diffusion weighting, lowering the SNR. This increases the demand for effective Rician bias correction (RBC) schemes<sup>10,11</sup> in DKI. Currently, it is unclear whether fitting the axisymmetric DKI framework with its reduced parameter space is better suited for parameter estimation from noisy diffusion MRI data than standard DKI and if its susceptibility to Rician noise bias is reduced.

The effect of the Rician noise bias on the fractional anisotropy (FA), mean diffusivity (MD), mean kurtosis (MK), diffusion tensor elements and diffusion kurtosis tensor elements was shown to be mitigated by using RBC in standard DKI.<sup>10,12-14</sup> Of these parameters only the mean kurtosis provides similar contrast as the mean of the kurtosis tensor  $\overline{W}$  and thus can be considered part of the axisymmetric DKI tensor metrics (AxTM). The AxTM are the parallel and perpendicular diffusivities ( $D_{\parallel}$  and  $D_{\perp}$ ), the parallel and perpendicular kurtosis and mean of the kurtosis tensor ( $W_{\parallel}$ ,  $W_{\perp}$ , and  $\overline{W}$ ). Here, parallel and perpendicular are in reference to the axis of symmetry. The

**TABLE 1** AxTM and standard diffusion kurtosis imaging (DKI) tensor metrics with which they are calculated

Axisymmetric DKI tensor metric (AxTM)	Corresponding standard DKI tensor metrics
$D_{\parallel}$	1: $\lambda_1$
$D_{\perp}$	2: $\lambda_2, \lambda_3$
$W_{\parallel}$	1: $W_{1111}$
$W_{\perp}$	3: $W_{2222}, W_{3333}, W_{2233}$
$\overline{W}$	6: $W_{1111}, W_{2222}, W_{3333}, W_{1122}, W_{1133}, W_{2233}$

*Notes:* The numbers show how many standard DKI tensor metrics are needed to compute the AxTM.  $\lambda$  refers to the eigenvalues of the diffusion tensor,  $W$  refers to the components of the kurtosis tensor.

axisymmetric DKI framework contains three additional parameters, the two angles of the unit vector pointing along the axis of symmetry, and the nondiffusion-weighted signal ( $b = 0$ ).

The AxTM can be estimated based on standard DKI and computed as aggregates from its 22 tensor metrics or directly with axisymmetric DKI (see Table 1). The AxTM are of particular interest for neuroscience and clinical research<sup>15-17</sup> because they are invariant against coordinate transformations and describe free and restricted diffusion within nervous tissue. Furthermore, the AxTM can be directly related to the tissue microstructure<sup>18,19</sup> via the axon water fraction, axon dispersion, and three diffusivities associated with the intra- and extra-axonal space.

It was shown empirically,<sup>13</sup> that RBC will impact the estimation of the parallel and perpendicular AxTM differently. It was speculated that the parallel and perpendicular AxTM are associated with different levels of water mobility and consequently different levels of SNR. Furthermore, another open question is the influence of fiber alignment on the effectiveness of RBC. It can be expected that the degree of fiber alignment within a white matter MRI voxel affects water mobility and thereby also the effectiveness of RBC.

In this work two main questions are investigated: First, we investigate whether RBC also increases the estimation accuracy of axisymmetric DKI. Second, we investigate whether the estimation accuracy is improved by using axisymmetric DKI as compared to standard DKI. Moreover, we investigate whether the performance of RBC depends on tissue fibre alignment and investigate differences in effectiveness for the parallel and perpendicular AxTM. To study these questions, we simulated two classes of datasets: the “synthetic dataset” is based on three sets of synthetic AxTM describing tissues with varying degrees of fiber alignment which allows us to assess AxTM estimation accuracy as a function of fiber alignment; the “in-vivo white matter dataset”

and “in-vivo gray matter dataset” are based on in vivo measurements of white matter tissue fiber tracts with a high to moderate fiber alignment (in-vivo white matter dataset) or typical gray matter areas (in-vivo gray matter dataset) which allows us to study AxTM estimation accuracy under realistic, in-vivo conditions. In both studies, axisymmetric DKI and standard DKI (with and without RBC) were used to obtain estimates of the five AxTM that could then be compared to the ground truth.

## 2 | METHODS

### 2.1 | Standard DKI signal representation

For a given diffusion weighting  $b$  and diffusion gradient  $\vec{g} = (g_1, g_2, g_3)^T$ , the noise-free DKI signal  $\tilde{S}_{b,\vec{g}}$  can be represented as in:<sup>5,20</sup>

$$\tilde{S}_{b,\vec{g}}(\tilde{S}_0, D, W) = \tilde{S}_0 \exp \left[ -bD + \frac{b^2}{6} \left( \frac{\text{Tr}(D)}{3} \right)^2 W \right] \quad (1a)$$

$$D = \sum_{i,j=1}^3 g_i g_j D_{ij} \quad (1b)$$

$$W = \sum_{i,j,k,l=1}^3 g_i g_j g_k g_l W_{ijkl} \quad (1c)$$

where  $D_{ij}$  are the diffusion tensor entries,  $W_{ijkl}$  are the kurtosis tensor entries and  $\tilde{S}_0$  is the non-diffusion-weighted signal ( $b = 0 \frac{\text{s}}{\text{mm}^2}$ ).

From the tensors  $D$  and  $W$ , the AxTM can be directly computed:  $D_{\parallel} = \lambda_1$  where  $\lambda_1$  is the first eigenvalue of the diffusion tensor  $D$ ,  $D_{\perp} = (\lambda_2 + \lambda_3)/2$ .

$W_{\parallel}$  and  $W_{\perp}$  can be computed from the fitted  $W$  tensor according to formulas 11 and 12 from:<sup>3</sup>  $W_{\parallel} = W(v_1) = W_{1111}$ , where  $v_1$  is the first eigenvector of the corresponding diffusion tensor and  $W_{\perp} = 3/8(W_{2222} + W_{3333} + 2W_{2233})$ .  $\overline{W}$  can be computed according to equation 10 from:<sup>21</sup>  $\overline{W} = 1/5(W_{1111} + W_{2222} + W_{3333} + 2W_{1122} + 2W_{1133} + 2W_{2233})$  (in:<sup>21</sup> 1 = x, 2 = y, 3 = z).

### 2.2 | Axisymmetric DKI

Axisymmetric DKI<sup>3</sup> assumes symmetric diffusion around an axis of symmetry  $\vec{c}$  inside an imaging voxel. Mathematically, this assumption leads to axisymmetric diffusion and kurtosis tensors with a drastically reduced number of independent tensor parameters compared to standard DKI (from 15 to 3 parameters for the kurtosis tensor and from 6

to 2 parameters for the diffusion tensor). In addition to the five AxTM, axisymmetric DKI contains two parameters for the axis of symmetry (inclination and azimuth). The symmetry assumptions are likely a reasonable approximation to diffusion in major white matter fiber bundles<sup>3</sup> due to their structural organization, which is why the focus of this study is white matter.

With the axis of symmetry  $\vec{c}$  parameterized by the inclination  $\theta$  and azimuth  $\phi$ :  $\vec{c} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^T$ , the diffusion and kurtosis tensors can be determined according to:<sup>3</sup>

$$D = D_{\parallel} \mathbf{I} + (D_{\parallel} - D_{\perp}) \vec{c} \vec{c}^T, \quad (2)$$

and

$$W = \frac{1}{2}(10W_{\perp} + 5W_{\parallel} - 15\overline{W})\mathbf{P} + W_{\perp}\mathbf{\Lambda} + \frac{3}{2}(5\overline{W} - W_{\parallel} - 4W_{\perp})\mathbf{Q},$$

where  $\Omega = \{D_{\parallel}, D_{\perp}, W_{\parallel}, W_{\perp}, \overline{W}, \tilde{S}_0, \theta, \phi\}$  are the eight framework's parameters ( $\tilde{S}_0$  is the nondiffusion-weighted signal) and  $\mathbf{I}$  is the three-dimensional identity matrix. The tensors  $\mathbf{P}$ ,  $\mathbf{\Lambda}$ , and  $\mathbf{Q}$  can be computed with the Kronecker delta  $\delta_{xy}$  and the components of the axis of symmetry  $c_x$  ( $x, y \in 1, 2, 3$ ) as:  $\mathbf{P}_{ijkl} = c_i c_j c_k c_l$ ,  $\mathbf{Q}_{ijkl} = (1/6)(c_i c_j \delta_{kl} + c_i c_k \delta_{jl} + c_i c_l \delta_{jk} + c_j c_k \delta_{il} + c_j c_l \delta_{ik} + c_k c_l \delta_{ij})$  and  $\mathbf{\Lambda}_{ijkl} = (1/3)(\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$ .<sup>3</sup> The associated noise-free signal  $\tilde{S}_{b,\vec{g}}(\Omega)$  can then be computed based upon the axisymmetric tensors:<sup>22</sup>

$$\tilde{S}_{b,\vec{g}}(\Omega) = \tilde{S}_0 \exp(-B_{ij} D_{ij} + \frac{1}{6} \overline{D}^2 B_{ij} B_{kl} W_{ijkl}), \quad (3)$$

where

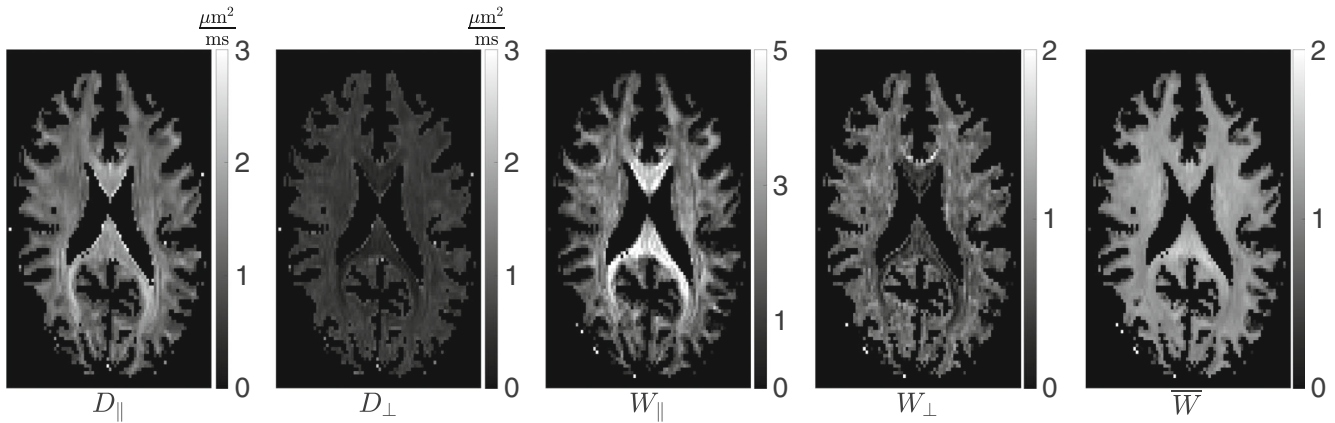
$$B_{ij} D_{ij} = \text{Tr}(B) D_{\perp} + (D_{\parallel} - D_{\perp}) \vec{c}^T B \vec{c}, \quad (4)$$

and

$$B_{ij} B_{kl} W_{ijkl} = \frac{1}{2}(10W_{\perp} + 5W_{\parallel} - 15\overline{W})(\vec{c}^T B \vec{c})^2 + \frac{1}{2}(5\overline{W} - W_{\parallel} - 4W_{\perp})(\vec{c}^T B \vec{c} \text{Tr}(B)) + 2\vec{c}^T B B \vec{c} + \frac{W_{\perp}}{3}(\text{Tr}(B))^2 + 2\text{Tr}(B \otimes B),$$

$$\text{with } B = b \begin{pmatrix} g_x^2 & g_x g_y & g_x g_z \\ g_x g_y & g_y^2 & g_y g_z \\ g_x g_z & g_y g_z & g_z^2 \end{pmatrix}.$$

The AxTM can be computed from the standard DKI tensor metrics assuming axial-symmetry, see Section 2.1. Table 1 shows the AxTM and the standard DKI tensor metrics needed to compute them. Figure 1 shows the five AxTM obtained with the axisymmetric DKI fit without RBC, available in the open source ACID toolbox for SPM.



**FIGURE 1** Axisymmetric diffusion kurtosis imaging (DKI) tensor metric (AxTM) results in white matter. The AxTM are the parallel and perpendicular diffusivity and kurtosis and the mean of the kurtosis tensor. The shown maps were obtained with the axisymmetric DKI fit available in the open source ACID toolbox for SPM that was used in this work. The AxTM were estimated from the in-vivo measurement used for the in-vivo white matter (and gray matter) dataset (Figure 3).

### 2.3 | Parameter estimation and the Rician noise bias

A detailed version of this section can be found in Section S1.1 of Appendix S1.

Often standard DKI or axisymmetric DKI parameters are estimated using the acquired magnitude signals  $S_{b,\bar{g}}$  and Equation (1a) or (3) with the least-squares approach<sup>23-25</sup> (found fit results are denoted with a hat, i.e.,  $\hat{S}_0, \hat{D}, \hat{W}$ ):

$$(\hat{S}_0, \hat{D}, \hat{W}) = \operatorname{argmin}_{\tilde{S}_0, D, W} \sum_i (S_{b,\bar{g}_i} - \tilde{S}_{b,\bar{g}_i}(\tilde{S}_0, D, W))^2. \quad (5)$$

However, this can be biased by the Rician noise bias. The severity of the Rician noise bias in the estimated parameters depends on the SNR:<sup>26</sup> the lower the SNR, the larger the bias. For RBC, we rely on an approach outlined in<sup>26</sup> that uses the expectation value  $\mathbf{E}(S_{b,\bar{g}})$  of the noisy composite magnitude dMRI signal. The probability density function of  $S_{b,\bar{g}}$  is a noncentral  $\chi$  distribution whose expectation value  $\mathbf{E}(S_{b,\bar{g}})$  is given by:<sup>26</sup>

$$\begin{aligned} \mathbf{E}(S_{b,\bar{g}}) &= \mathbf{E}(\tilde{S}_{b,\bar{g}}(\tilde{S}_0, D, W), \sigma) \\ &= \sigma \sqrt{\frac{\pi}{2}} \cdot \mathbf{L}_{1/2}^{(L-1)} \left( \frac{\tilde{S}_{b,\bar{g}}(\tilde{S}_0, D, W)^2}{2\sigma^2} \right), \end{aligned} \quad (6)$$

where  $\mathbf{L}_{1/2}^{(L-1)}(x) = \Gamma(L + 1/2) / (\Gamma(3/2)\Gamma(L)) \mathbf{M}(-1/2, L, x)$  is the generalized Laguerre polynomial which can be expressed using a confluent hypergeometric function  $\mathbf{M}$ , the Gamma function  $\Gamma$  and the number of receiver coils  $L$ . Following,<sup>26</sup> we implemented a time-efficient Gauss Newton fitting algorithm<sup>27</sup> (see Section S1.2 of Appendix S1) that, unlike Equation (5), accounts for Rician noise

in magnitude dMRI data by solving the optimization problem:

$$(\hat{S}_0, \hat{D}, \hat{W}) = \operatorname{argmin}_{\tilde{S}_0, D, W} \sum_i (S_{b,\bar{g}_i} - \mathbf{E}(\tilde{S}_{b,\bar{g}_i}(\tilde{S}_0, D, W), \sigma))^2. \quad (7)$$

Estimating parameters this way is referred to as “quasi-likelihood” estimation and is denoted as “RBC ON” in this paper.

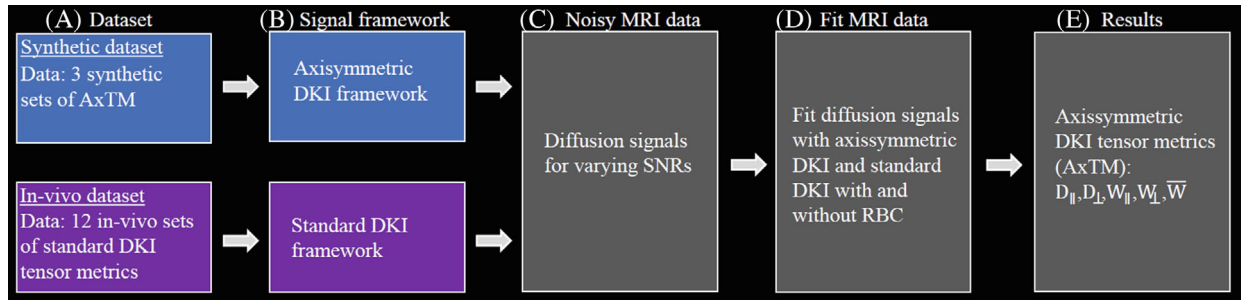
Rician bias corrected, standard DKI or axisymmetric DKI parameter estimation can be done by using Equation (1a) or Equation (3) to compute the noise-free signal predictions  $\tilde{S}_{b,\bar{g}}$ , then using Equation (6) to compute  $\mathbf{E}(\tilde{S}_{b,\bar{g}}(\tilde{S}_0, D, W), \sigma)$  and finally minimize Equation (7) to estimate the framework parameters  $(\hat{S}_0, \hat{D}, \hat{W})$  for standard DKI or  $\Omega$  for axisymmetric DKI.

### 2.4 | Simulation study: datasets and overview

A detailed version of this section can be found in Section S1.3 of Appendix S1.

We assessed estimation accuracy of the five AxTM as a function of the SNR in a simulation study with two classes of datasets, see Figure 2. The first class, the “synthetic dataset,” consisted of three synthetic voxels with varying fiber alignment (defined in Reference 28). The other class of datasets was based on an in-vivo measurement and consisted of either 12 major white matter fiber tract voxels (“in-vivo white matter dataset”) or 12 voxels from typical gray matter areas (“in-vivo gray matter dataset”). For all datasets, magnitude diffusion MRI data were simulated for varying SNRs and fitted





**FIGURE 2** Scheme of the simulation study. Simulations were performed in 5 steps: (A) choice of datasets, (B) signal framework used for simulation, (C) diffusion signal simulation and contamination with noise, (D) parameter estimation, and (E) results. Note that both simulation studies only differed in (A) and (B) but were identical in the following procedures. (A) the synthetic dataset consisted of  $3 \times 5$  sets of axisymmetric diffusion kurtosis imaging (DKI) tensor metric (AxTM) (Table S1) while the in-vivo dataset (either white matter or gray matter) consisted of  $12 \times 22$  standard DKI tensor metrics (Table S2 lists the white matter data while Table S4 lists the gray matter data). (B) The DKI signal framework used for diffusion signal simulation. (C) Diffusion signal data were contaminated with 2500 Rician noise samples for each signal-to-noise ratio,  $\text{SNR} = [1, 2, 3 \dots 200]$ . (D) Simulated diffusion data were fitted with axisymmetric DKI and standard DKI with and without Rician bias correction in both simulation studies for each of the 2500 noise samples. (E) The axisymmetric DKI tensor metrics (AxTM):  $D_{\parallel}, D_{\perp}, W_{\parallel}, W_{\perp}$  and  $\bar{W}$  were calculated for standard DKI (for axisymmetric DKI they were directly estimated), averaged across the 2500 noise samples per SNR and finally compared to the ground truth.

with standard DKI and axisymmetric DKI, with and without RBC (as described in Section 2.3) to obtain estimates of the five AxTM. Accuracy of the obtained AxTM estimates were evaluated as the absolute value of the mean percentage error (A-MPE):

$$\text{A-MPE} = 100 \cdot \frac{|\text{GT} - \overline{\text{FitResults}(\text{SNR})}|}{\text{GT}}. \quad (8)$$

Here GT refers to the ground truth and  $\overline{\text{FitResults}}$  refers to the average of the fit results over the noise samples. We evaluated the accuracy of the AxTM estimates for each estimation method by looking for the SNR after which the A-MPE was smaller 5%, which was considered an acceptable error in a trade-off between estimation accuracy and SNR requirement. As a summary to compare each method, we looked at the maximum SNR needed across the five AxTM for which A-MPE consistently  $< 5\%$  for all AxTM (“Maximum” column in Figure 6).

### 2.4.1 | Datasets

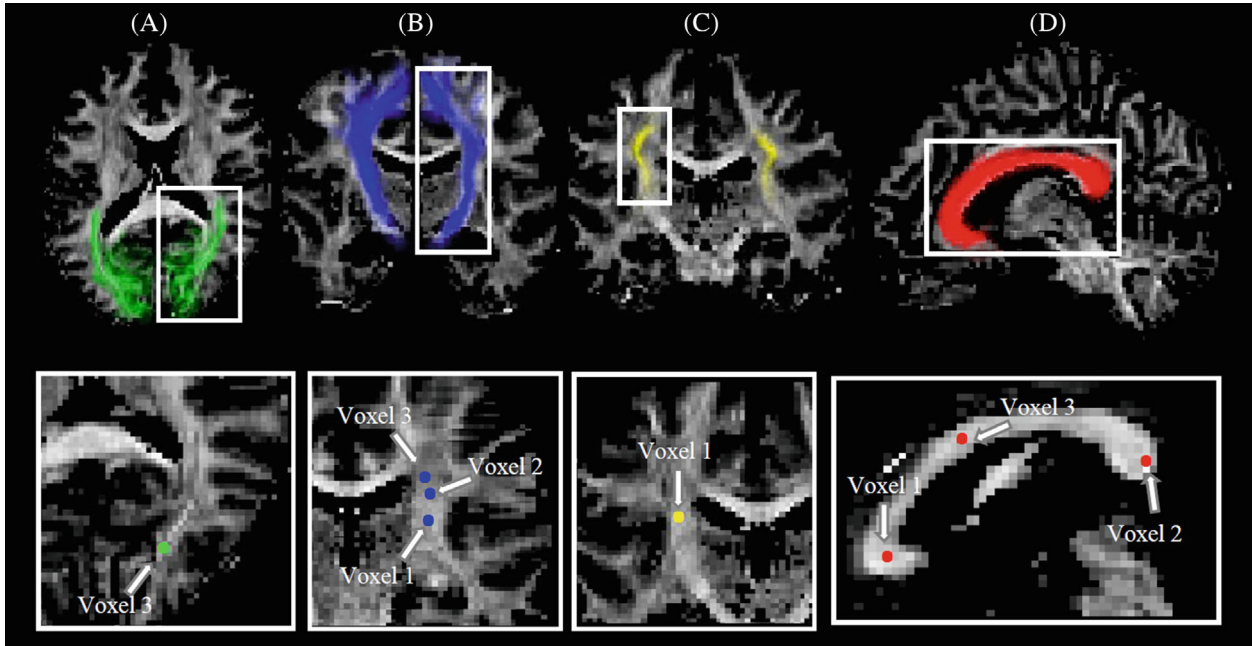
The synthetic dataset consisted of three synthetic sets of AxTM (from Reference 28) describing three voxels with varying fiber alignment, one with fibers with low alignment (“LA”,  $\text{FA} = 0.067$ ), one with fibers with moderate alignment (“MA”,  $\text{FA} = 0.24$ ) and one with highly aligned fibers (“HA”,  $\text{FA} = 0.86$ ). The AxTM of the three synthetic voxels are summarized in Table S1. Figure 4 shows two areas of typical brain regions in a map of the mean of the kurtosis tensor  $\bar{W}$  where LA and HA voxels

can be found and the corresponding idealized fiber stick model.

The in-vivo white matter dataset consists of 12 voxels extracted from four major white matter tracts (three voxels from each of the four fiber tracts, see Figure 3) from an in-vivo brain measurement ( $\text{SNR} = 23.4$ ) of a healthy volunteer, details on the DWI acquisition sequence can be found in Section S1.3 of Appendix S1. The 12 voxels were extracted from the in-vivo measurement by fitting the standard DKI framework in 12 white matter voxels of the acquired in-vivo DWI magnitude images to get the corresponding 22 standard DKI tensor metrics, the derived data are therefore referred to as “in-vivo white matter”. Three voxels each with HA to MA (defined through the FA threshold  $\text{FA} \geq 0.4$ <sup>29</sup>) were extracted, see Figure 3. The selected voxels differ from the synthetic voxels in that here only HA and MA voxels were selected. The extracted white matter standard DKI tensor metrics (Table S2) and corresponding AxTM (Table S3) are documented in the Appendix S1.

The in-vivo gray matter dataset was produced according to the same procedure used for the in-vivo white matter dataset, only that the voxels were selected from typical gray matter areas. The extracted gray matter standard DKI tensor metrics (Table S4) and corresponding AxTM (Table S5) are documented in the Appendix S1. Since white matter is the focus of this article, details and results of the in-vivo gray matter dataset can be found in Section S1.5 of Appendix S1.

The in-vivo dMRI data used for this study were acquired with the help of a human research participant.



**FIGURE 3** Selection of voxels of the in-vivo white matter dataset: The four white matter fiber pathways within which voxels were selected and used as a basis for the in-vivo white matter dataset. Top: (A) Optic radiation (or), (B) cortico spinal tract (ct), (C) superior longitudinal fasciculus (slf) and (D) callosum body (cb) in a fractional anisotropy (FA) map of a healthy human brain. The fiber pathways were identified with the coregistered Jülich fiber atlas.<sup>30</sup> Bottom: Voxels in the fiber pathways used for the in-vivo white matter dataset. In each fiber pathway, three voxels were chosen for the in-vivo white matter dataset (for slf an or only one is shown here because the three chosen voxels were not in the same slice).

The participant provided written informed consent and was compensated for its participation. The local ethics committees at University Medical Center Hamburg-Eppendorf approved the study (PV5141).

#### 2.4.2 | Signal framework used for simulation

The three synthetic voxels of AxTM were simulated with the axisymmetric DKI framework to first obtain noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$ . The 12 in-vivo white matter and gray matter voxels were simulated with the standard DKI framework to first obtain noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$ .

#### 2.4.3 | Contamination with noise

For both the synthetic and the in-vivo dataset (white matter or gray matter), the noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$  were contaminated with noise for SNRs [1, 2, 3...200] and magnitude signals  $S_{\text{cont}}$  were computed. The noisy magnitude signals were computed according to  $S_{\text{cont}} = |\tilde{S}_{\text{noise-free}} + \alpha + \beta i|$ , where  $\alpha, \beta \in \mathcal{N}(0, \sigma)$  are drawn from a zero mean Gaussian with SD  $\sigma$ , yielding different  $\text{SNR} = \sqrt{2}S_0/\sigma$  (for one receiver coil) for a given  $S_0 = 1$ .

#### 2.4.4 | Estimating the five AxTM

Both, the simulated signals  $S_{\text{cont}}$  from the synthetic and the in-vivo dataset were fitted with axisymmetric DKI and standard DKI, with and without RBC (Section 2.3) to obtain estimates of the AxTM whose accuracy could then be investigated as a function of SNR.

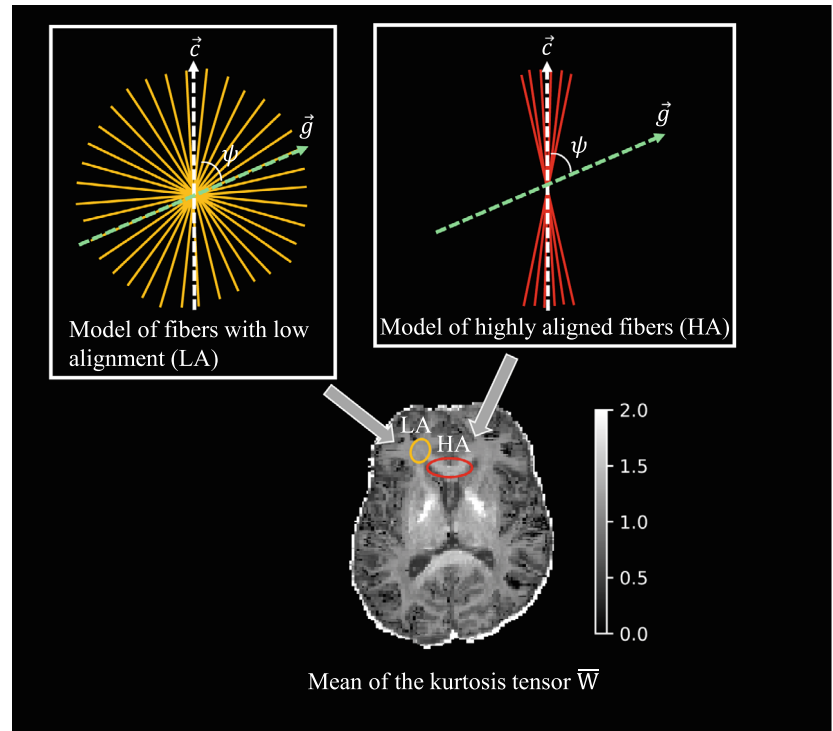
### 2.5 | Diffusion signal profiles influenced by fiber alignment

To further elucidate differences between tissues with different levels of fiber alignment, angular signal profiles under the influence of noise were studied for the three voxels of the synthetic dataset. Noise-free and noise-contaminated signals have been simulated ( $\text{SNR} = 20$ ). The simulated signal's mean and SD could then be plotted as a function of angle  $\psi$  (in degree) between diffusion gradient  $\vec{g}$  and axis of symmetry  $\vec{z}$ . For a graphical representation of angle  $\psi$ , see Figure 4.

## 3 | RESULTS

First, the results of the diffusion signal profiles in voxels with different levels of fiber alignments (Section 3.1) are

**FIGURE 4** Model of fiber alignment in characteristic areas of the brain. Bottom: in-vivo map of the mean of the kurtosis tensor  $\overline{W}$  with a typical area where fibers with low alignment (LA) are found and a typical area where highly aligned fibers (HA) are found. Top: the corresponding golden and red sticks depict an idealized model of the underlying fiber arrangement, the white dashed line indicates the axis of symmetry  $\vec{c}$ , the green dashed line indicates the diffusion gradient direction  $\vec{g}$ ,  $\psi$  is the angle between  $\vec{g}$  and  $\vec{c}$ .



shown because these not only explain the results obtained in different tissues but also help to understand the difference between estimating the parallel or the perpendicular AxTM. After that, our main findings are stated and the corresponding results are reported (Section 3.2).

### 3.1 | Diffusion signal profiles influenced by fiber alignment

Each of the simulated voxels of the synthetic dataset shows a characteristic,  $\psi$  dependent shape, see Figure 5. For smaller angles  $\psi$  between  $\psi = 10^\circ$  and  $\psi = 0^\circ$ , the simulated signals of the HA voxel are strongly diffusion weighted and are close or below the noise floor (SNR = 1), see Figure 5A. In the simulated voxel of MA, the noise floor is already reached for angles  $\psi \approx 50^\circ$ , see Figure 5B. In the simulated LA voxel the noise floor is never reached and the simulation shows a seemingly constant,  $\psi$  independent signal form, see Figure 5C. In summary, the signal in HA to MA decays along the direction of symmetry, whereas there is almost no decay in LA.

### 3.2 | Results of simulation study

Figure 6 shows the SNRs required to accurately estimate the AxTM (A-MPE < 5%, Equation 8) in the synthetic and in-vivo white matter dataset. Shown are

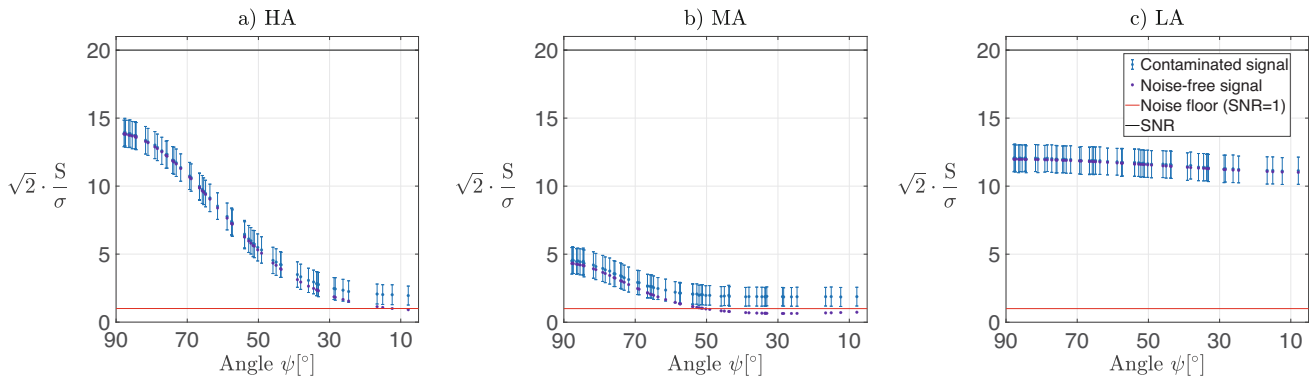
the results for axisymmetric DKI and standard DKI with (hatched) or without RBC.

#### 3.2.1 | RBC is most effective in highly aligned fibers and parallel diffusion case

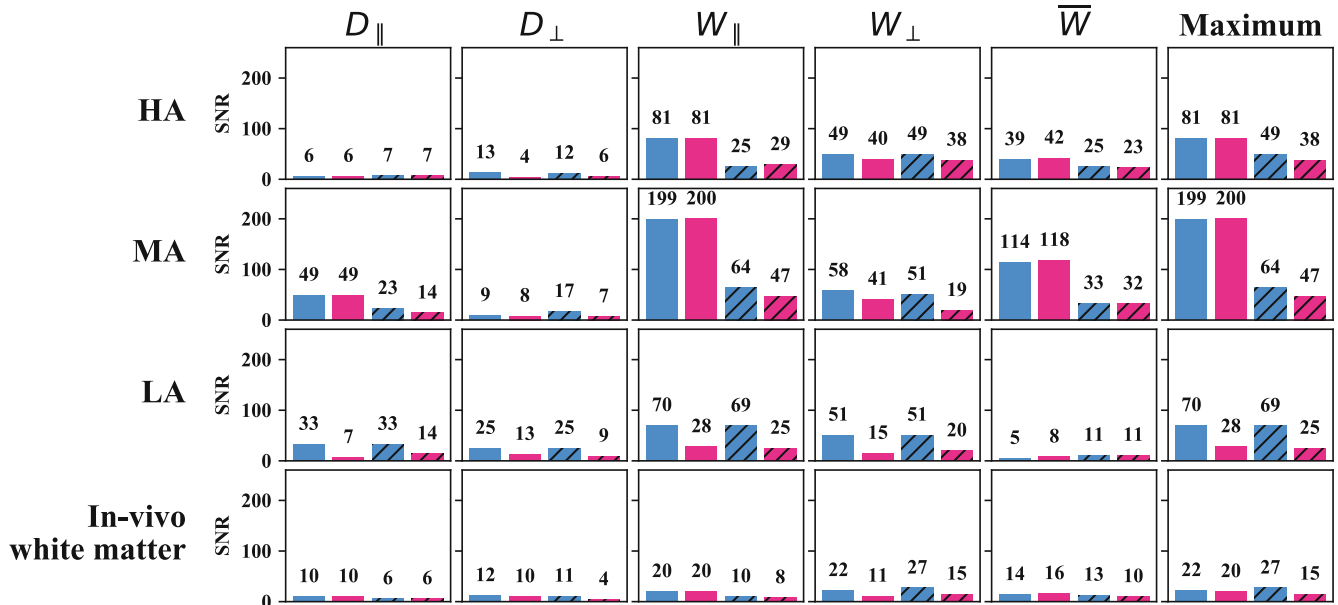
RBC was most effective for the parallel parameters  $D_{\parallel}$  and  $W_{\parallel}$  in HA to MA (for both the synthetic and in-vivo white matter dataset), see Figure 6. For example, achieving A-MPE < 5% for  $W_{\parallel}$  in the synthetic HA voxel could be reduced from SNR = 81 to SNR = 25 (standard DKI) or SNR = 29 (axisymmetric DKI). In the in-vivo white matter voxels, the SNR requirements for achieving A-MPE < 5% for  $W_{\parallel}$  could be reduced from 20 to 10 (standard DKI) or to 8 (axisymmetric DKI). Estimation of  $\overline{W}$  was significantly improved by RBC in the synthetic HA and MA voxels but only slightly in the in-vivo white matter datasets.

#### 3.2.2 | Superiority of axisymmetric DKI in axisymmetric fibers with low alignment where RBC is ineffective

In the synthetic LA voxel, estimation of  $D_{\parallel}$ ,  $D_{\perp}$ ,  $W_{\parallel}$ , and  $W_{\perp}$  was substantially improved by using the axisymmetric DKI framework instead of standard DKI. For example, it only required an SNR = 15 (axisymmetric DKI) instead of SNR = 51 (standard DKI) to achieve A-MPE < 5% for



**FIGURE 5** Simulated signal decay along the axis of symmetry at signal-to-noise ratio (SNR) =  $\sqrt{2}S_0/\sigma = 20$ . Signal decay is shown for the synthetic dataset consisting of A) highly aligned fibers (HA), B) fibers with moderate alignment (MA) and C) fibers with low alignment (LA) as a function of angle  $\psi$  between the diffusion gradient  $\vec{g}$  and the axis of symmetry  $\vec{c}$ , see Section 2.5. The “Contaminated signal” shows the mean and SD over 2500 noise-samples. Plotted is the quantity  $\sqrt{2} \cdot S/\sigma$ , which is the SNR computed for the diffusion weighted signal  $S$  (instead of  $S_0$ ) at a given angle  $\psi$  (for a graphical representation of angle  $\psi$  see Figure 4). The red horizontal line indicates the noise floor where SNR = 1. All three plots were calculated for the highest b-shell ( $b = 2500 \text{ s/mm}^2$ ).



**FIGURE 6** Signal-to-noise ratio (SNR) above which the absolute value of the mean percentage error (A-MPE, Equation 8) < 5% for the synthetic dataset with high, medium and low fiber alignment (“HA,” “MA,” “LA”) and the in-vivo white matter dataset. For the in-vivo white matter dataset, the A-MPE was averaged across the 12 simulated voxels and the SNR above which this average A-MPE < 5% is shown. The SD across the in-vivo white matter dataset voxels is not shown here because the values were between  $\approx 0.5$  and  $\approx 6$  with an average of  $\approx 2.5$  and thus too small to display. The number above the barplots indicates the barplot’s height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. “Maximum” shows the maximum SNR needed to achieve A-MPE < 5% across all five AxTM.

$W_{\perp}$ . For  $\bar{W}$ , axisymmetric DKI performed slightly worse than standard DKI (SNR = 8 instead of SNR = 5). Interestingly, RBC did not influence the fitting results much in this fiber alignment configuration but even worsened them in some cases (e.g., axisymmetric DKI results for  $D_{\parallel}$ ), see Figure 6.

### 3.2.3 | Axisymmetric DKI improves estimation of perpendicular parameters

Estimation of  $D_{\perp}$  and  $W_{\perp}$  could also be improved (A-MPE < 5% reached for lower SNRs) by using the axisymmetric DKI framework in HA and MA of the synthetic

and in-vivo white matter dataset (e.g., for HA,  $D_{\perp}$ : SNR = 4 instead of SNR = 13;  $W_{\perp}$ : SNR = 40 instead of SNR = 49).

### 3.2.4 | RBC can also worsen accuracy

Interestingly, there were also few scenarios in which RBC increased the SNR requirements. As described above this was observed for  $D_{\parallel}$ ,  $W_{\perp}$ , and  $\overline{W}$  in the LA voxel and, outside the LA voxel, predominantly for the perpendicular parameters, for example, from SNR = 11 to SNR = 15 for the axisymmetric DKI fit of  $W_{\perp}$  in the in-vivo white matter dataset.

## 4 | DISCUSSION

### 4.1 | Summary of main findings

Overall, we found, that the combination of axisymmetric DKI with RBC was the best option for estimating all five AxTM in the synthetic and in-vivo white matter datasets, see “Maximum” column of Figure 6. This combination achieved an absolute value of the mean percentage error (A-MPE, Equation 8) < 5% in our simulated in-vivo white matter data if the SNR  $\geq 15$ , making this combination a possibly valuable tool in neuroscience and clinical research studies focusing on white matter. Specifically, we found that RBC is highly effective for increasing estimation accuracy of the AxTM associated with diffusion parallel to the main fiber orientation, that is, parallel diffusivity and kurtosis, in in-vivo white matter. In contrast, RBC fails in improving estimation accuracy in parameters perpendicular to the main fiber orientation, that is, perpendicular diffusivity and kurtosis, or if fiber alignment is too low. For the latter scenarios, axisymmetric DKI is more effective than standard DKI in in-vivo white matter or the synthetic LA voxel.

### 4.2 | Rician noise bias and its correction: Effectiveness for different DKI parameters and levels of fiber alignment within white matter

The effectiveness of RBC correlated with the severity of the Rician noise bias, it varied between individual AxTM and depended on the level of fiber alignment. The severity of the Rician noise bias is inversely proportional to the SNR which in turn depends on a variety of parameters. One of these parameters is the level of water mobility which

tunes the diffusivity and thereby the attenuation of the diffusion weighted signal. Water mobility is influenced by the level of fiber alignment. Furthermore, each AxTM itself is associated with different levels of water mobility due to its association to the axis of symmetry  $\vec{c}$ . Figure 5 shows simulated signals in three tissue types with varying degrees of fiber alignment as a function of angle  $\psi$  between  $\vec{c}$  and diffusion gradient  $\vec{g}$ . In HA for example, the diffusion signal is heavily diffusion weighted if measured along the main fiber orientation (small angle  $\psi$ ) and the Rician noise bias in these signals therefore strongest, see Figure 5A. Since the parallel AxTM ( $D_{\parallel}$  and  $W_{\parallel}$ ) predominantly depend on the signal along these directions, it can be expected that they, too, are more heavily biased in a high fiber alignment setting. Accordingly, we found that RBC turned out to be particularly important for the AxTM associated with parallel diffusion ( $D_{\parallel}$  and  $W_{\parallel}$ ) in highly aligned white matter (synthetic dataset and in-vivo white matter dataset). On the other hand, diffusion perpendicular to  $\vec{c}$  will be more restricted than the parallel diffusion and the SNR therefore higher along those directions.  $D_{\perp}$  and  $W_{\perp}$  should therefore be less affected by the Rician noise bias. Indeed, we found that RBC was not as effective for the perpendicular parameters  $D_{\perp}$  and  $W_{\perp}$  in in-vivo white matter and HA as for the parallel parameters. An apparent contradiction to this argument was found for  $W_{\perp}$  in the synthetic MA dataset. This contradictory finding, however, can be explained by the relatively high perpendicular diffusivity of the MA voxel (Table S1 of Appendix S1) causing a strong diffusion weighting and therefore smaller perpendicular signal in this case (Figure 5B).

RBC was furthermore ineffective in the synthetic LA voxel, see Figure 6. Here, the signal change as a function of angle  $\psi$  between  $\vec{c}$  and  $\vec{g}$  is smaller than the variation introduced by noise and the signals almost seem independent of the diffusion gradient direction, see Figure 5C. This is because the synthetic LA voxel does not possess a clearly distinguishable axis along which water mobility is significantly heightened compared to other directions. For the same SNR (in reference to the  $S_0$  signal), the Rician noise bias in such tissues is therefore less severe, compared to, for example, signals in HA acquired for diffusion gradients parallel to the axis of symmetry. Accordingly, we found that RBC had little to no effect on parameter estimation in the synthetic LA voxel. In addition to this, the original, seemingly constant signal pattern could be modified by the noise in such a way that the RBC might even worsen the bias if it acts on the noisy signals and further shifts them away from their original, seemingly constant pattern. For example, signals that were reduced due to noise might be further reduced by the RBC.

### 4.3 | Advantage of the axisymmetric DKI framework

In Section 1, we hypothesized that a reduction of the parameter space could make axisymmetric DKI more robust against the Rician noise bias. We observed that axisymmetric DKI predominantly improved accuracy of parameter estimation of the perpendicular AxTM  $D_{\perp}$  and  $W_{\perp}$  in both the synthetic dataset (HA and MA voxels) and the in-vivo white matter dataset, that is, A-MPE < 5% was achieved for lower SNRs when using axisymmetric DKI to estimate these metrics.

The increased accuracy for  $W_{\perp}$  when using the axisymmetric DKI framework can be understood with the information provided in Reference 3. Here an approach for estimation of  $W_{\perp}$  (equation 24 in Reference 3) with the axisymmetric DKI framework is outlined that only depends on measurements with diffusion gradients perpendicular to the principal diffusion axis (axis of symmetry  $\vec{c}$ ) and the square of the MD. The measurements perpendicular to the main diffusion axis are less strongly attenuated because diffusivity is lower in that direction, see Section 4.2. Therefore, these signals are less strongly affected by the Rician bias. Moreover the MD, which depends on diffusion gradients that are not perpendicular to the principal diffusion axis, can be estimated from the lower diffusion shells in a DKI acquisition. Thus, the influence of the Rician bias on the MD is lower. Taken together, the estimation of  $W_{\perp}$  with axisymmetric DKI can be done based on “perpendicular measurements” (for the higher b-shell) and the MD which are both less strongly affected by the Rician noise bias resulting in a reduced bias of  $W_{\perp}$ . The observed improvement for  $D_{\perp}$  can be understood following the same line of thought and equation 23 in Reference 3 which lays out a scheme for estimating  $D_{\perp}$  based on axisymmetric DKI and measurements with diffusion gradients perpendicular to the principal diffusion axis.

Another observation was that, accuracy of parameter estimation with axisymmetric DKI was substantially better in tissues with low fiber alignment (Figure 6) compared to standard DKI. We observed, for example, an SNR reduction of up to 70% when using axisymmetric DKI instead of standard DKI for estimation of  $W_{\perp}$  in the LA voxel, see Figure 6. This could be due to the seemingly constant and high signal, independent of the diffusion gradient direction, in the synthetic LA dataset. Since the variation in the almost constant diffusion signal is dominated by noise, the complex 22 parametric standard DKI framework is more likely to overfit the data than the eight parametric axisymmetric DKI framework, particularly in case of lower SNRs where noise has a greater impact. This could be the reason for the clear advantage of axisymmetric DKI over standard

DKI in this fiber configuration. However, axisymmetry in the underlying fibers could be a crucial prerequisite for observing this improvement in LA-type voxels using axisymmetric DKI, see Section 4.4.

## 4.4 | Extra analyses and considerations

### 4.4.1 | In-vivo gray matter simulation

In comparison to white matter, gray matter does not possess a clear axis of symmetry which makes it less suitable for axisymmetric DKI.<sup>3</sup> However, in practice axisymmetric DKI metrics can generally also be computed in gray matter which is why we also simulated and analyzed axisymmetric DKI in in-vivo gray matter, see Section S1.5 and Figure S1 of Appendix S1. Gray matter could potentially be modeled by the synthetic voxel of fibers with low alignment (LA) shown in Figure 4. Comparing the SNR needed by each fitting method to reach the A-MPE < 5% threshold in in-vivo gray matter with the synthetic LA configuration (Figure 6, main document or Figure S1 of Appendix S1) revealed that axisymmetric DKI is performing better than standard DKI in the synthetic LA voxel but worse in the in-vivo gray matter dataset.

An intuitive explanation for the observed deviation could be that violation of the assumption of axisymmetry in gray matter leads to an inherent bias of the axisymmetric DKI parameter estimates that cannot be neglected. This intuitive explanation is further supported by our recent study<sup>31</sup> where we experimentally observed a deviation between standard DKI and axisymmetric DKI parameter estimates in the brain. However, this intuitive explanation would be in contradiction to a recently introduced analytical proof<sup>32</sup> showing that axisymmetric DKI and standard DKI produce the same parameter estimates even if axisymmetry in the underlying tissue is not fulfilled. However, this proof only holds under the condition that the first eigenvector of the diffusion tensor aligns with the axis of symmetry. Therefore, a thorough analysis is required in future studies to improve our understanding of the observed deviations between the analytical proof and the experimental observations.

### 4.4.2 | Analysis of precision

We have additionally investigated the precision of the four proposed methods for in-vivo white matter by calculating the SD in reference to the ground truth (R-STD) of each method (see Section S1.6 of Appendix S1). Our main findings were: first, the precision was not improved by RBC. Second, the precision of axisymmetric DKI was

worse than that of standard DKI. The first finding was expected because RBC does not improve precision. RBC shifts the expectation value (modeled by the mean) of the distribution of fit results toward the ground truth; however, this does not influence the width of the distribution of fit results which defines the precision by which a quantity can be estimated. Improvement of precision can be achieved by either averaging multiple identically performed measurements or by applying de-noising methods like, for example, done in Reference 14. Our second finding is in contrast to what we have assumed in the introduction of our article, namely that axisymmetric DKI is less susceptible to the noise-induced variation of the acquired diffusion MRI signals compared to standard DKI due to its smaller parameter space. Our findings show that the SNR required to reach the  $R\text{-STD} \leq 5\%$  threshold was typically slightly and in some cases substantially (e.g., *in-vivo* white matter dataset for  $W_{\perp}$ ) higher when using axisymmetric DKI compared to standard DKI. The case of substantially higher SNRs required to reach the  $R\text{-STD} \leq 5\%$  threshold in the axisymmetric DKI fit results could be connected to outliers. The interquartile range was used to compute a robust, that is, less outlier sensitive, measure (“R-IQR”) for the SD, see Section S1.6 of Appendix S1. With respect to the R-IQR, standard DKI, and axisymmetric DKI performed almost identically in *in-vivo* white matter for  $W_{\perp}$ , making the outliers in axisymmetric DKI a highly likely explanation for the stark difference in precision between standard and axisymmetric DKI in this case. However, even when using the outlier-robust R-IQR measure, both DKI frameworks perform very similar, that is, axisymmetric DKI seems to not have an improved precision (compared to standard DKI) despite its smaller parameter space. One reason that might explain this finding could be that the axisymmetric DKI framework cannot be linearised (in contrast to standard DKI) and thus might be more susceptible to noise, for example, due to noise enhancement.

#### 4.4.3 | Error propagation in standard DKI

The five AxTM can be calculated from the standard DKI tensor metrics. The number of standard DKI metrics needed to compute each AxTM is listed in Table 1. This property illustrates the risk of error propagation of the base standard DKI tensor metrics into the AxTM. The uncertainty of each estimated standard DKI metric propagates into the resulting AxTM computed from them resulting in a dependency of the uncertainty on the number of the base standard DKI tensor metrics.  $D_{\parallel}$ , for example, relies on only one parameter while  $W_{\perp}$  relies on three, so that a

priori it can be assumed that  $W_{\perp}$  is more severely affected by uncertainty propagation in standard DKI.

#### 4.4.4 | Possible circularity of simulation study

Since the simulation studies were either based on the axisymmetric DKI (synthetic dataset) or the standard DKI framework (*in-vivo* white matter dataset), one might argue that the simulations will favor their respective signal frameworks. Here we rediscuss our signal-framework comparisons in the light of this potential circularity: We observed that using axisymmetric DKI was generally advantageous over standard DKI for the perpendicular AxTM. Axisymmetric DKI was also performing significantly better on all AxTM except  $\bar{W}$  in the LA dataset. Since the improvement of axisymmetric DKI over standard DKI for the perpendicular AxTM was observed across both simulations, the observation cannot be explained by a circularity argument and we believe that it is a genuine advantage of axisymmetric DKI. The have additionally investigated the synthetic LA dataset, however, is based on the axisymmetric DKI framework and the better results might well be confounded by the circularity argument. However, our noise-robustness argument is also a reasonable explanation for the superiority of axisymmetric DKI in this case. Thus, the truth might be in between, that is, the real improvement of estimation accuracy in the LA dataset when using axisymmetric DKI might be lower than in the simulation but we would expect to still observe an improvement in *in-vivo* data given a sufficient level of axisymmetry.

#### 4.4.5 | Limits of current measurements protocols

Looking at the estimation accuracy for each of the five AxTM individually revealed that each metric comes with different SNR requirements. Estimation of  $W_{\parallel}$  with an A-MPE  $< 5\%$ , for example, required an SNR of 81 in the HA voxel of the synthetic dataset and an SNR of 199 (standard DKI) or 200 (axisymmetric DKI) in the MA voxel of the synthetic dataset if RBC was not used. This reveals that using current measurement protocols could yield biased estimates under realistic conditions where the SNR is below 81 or 200 if RBC is not used. This underlines the importance of using RBC in cases where all five AxTM are of importance, for example, for estimation of the biophysical microstructure parameters.<sup>18,19</sup>

#### 4.4.6 | Limits of RBC for single voxel application

Similar to previous simulation studies on RBC,<sup>10,13,14</sup> we focused on the effects of RBC on the averaged estimated AxTM over the 2500 noise samples which is a stable estimator for the expectation value of the distribution of the fit results. The SNR at which that expectation value of the AxTM could be estimated with an A-MPE < 5% is reported in this article. If this expectation value is unbiased for a specific SNR, the underlying DKI framework is demonstrated to be unbiased, as well. Parameter estimation based on real, in-vivo measurements typically relies on only one noise realization. Therefore, the SNR results from our simulation study cannot be transferred directly to standard in-vivo DWI measurements. The SNRs found in our simulation study show when the estimator itself is unbiased given sufficient repetitions. We also investigated the SNR requirements for a high precision and found that much higher SNRs were required (see Section S1.6 of Appendix S1 or Section 4.4.2 above for a more general discussion of estimation precision).

## 5 | CONCLUSION

Our study revealed that axisymmetric DKI with RBC is the most SNR effective choice for estimating the AxTM in tissues that do not violate its assumptions because of two mutually supporting factors. First, RBC itself is most effective for the parallel diffusivity and kurtosis and the mean of the kurtosis tensor, however, it needs at least some level of fiber alignment to work. Second, compared to standard DKI, axisymmetric DKI is superior in axisymmetric fibers with low alignment and more effective for estimating the perpendicular diffusivity and kurtosis. This makes the combination of axisymmetric DKI with RBC a possibly valuable tool for neuroscience and clinical research studies focusing on white matter where a gain in SNR could either be used to reduce scan time or increase spatial resolution.

### ACKNOWLEDGMENTS

The implemented code for axisymmetric DKI fitting makes use of the following externally written tools:

- The Gauss Newton fit algorithm implementation used in this study was conceptualized and written by Jan Modersitzki<sup>27</sup> and expanded by Lars Ruthotto who, for example, implemented slice-wise parameter estimation and introduced an efficient, multi-voxel procedure to accelerate convergence; both improved the algorithm's run-time.

- For the initial guess of the axisymmetric DKI fit implementation, we used code from the repository of Sune Nørhøj Jespersen: <https://github.com/sunenj/Fast-diffusion-kurtosis-imaging-DKI>.<sup>3</sup>

Open Access funding enabled and organized by Projekt DEAL.

### FUNDING INFORMATION

This work was supported by the German Research Foundation (DFG Priority Program 2041 “Computational Connectomics”, [MO 2397/5-1; MO 2397/5-2], by the Emmy Noether Stipend: MO 2397/4-1) and by the BMBF (01EW1711A and B) in the framework of ERA-NET NEURON.

### AUTHORS' CONTRIBUTIONS

**Jan Malte Oeschger:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, **Karsten Tabelow:** Conceptualization, Methodology, Supervision, Writing – review & editing, **Siawoosh Mohammadi:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.


### DATA AVAILABILITY STATEMENT


The open-source ACID toolbox (<http://www.diffusiontools.com/>) for SPM contains the fitting algorithms for standard and axisymmetric DKI with and without RBC used in this study. Furthermore, a code repository for simulation and analysis of the data used in this study will be made available at [https://github.com/quantitative-mri-and-in-vivo-histology/axisymmetric\\_dki\\_with\\_rician\\_bias\\_correction\\_simulation\\_study](https://github.com/quantitative-mri-and-in-vivo-histology/axisymmetric_dki_with_rician_bias_correction_simulation_study).

### ETHICS STATEMENT

The in-vivo dMRI data used for this study were acquired with the help of a human research participant. The participant provided written informed consent and was compensated for its participation. The local ethics committees at University Medical Center Hamburg-Eppendorf approved the study (PV5141).

### ORCID

Jan Malte Oeschger  <https://orcid.org/0000-0003-0237-923X>

Karsten Tabelow  <https://orcid.org/0000-0003-1274-9951>

Siawoosh Mohammadi  <https://orcid.org/0000-0003-1311-9636>



## REFERENCES

1. Deppe M, Kellinghaus C, Duning T, et al. Nerve fiber impairment of anterior thalamocortical circuitry in juvenile myoclonic epilepsy. *Neurology*. 2008;71:1981-1985.
2. Rovira À, Wattjes MP, Tintoré M, et al. MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis—Clinical implementation in the diagnostic process. *Nat Rev Neurol*. 2015;11:471-482.
3. Hansen B, Shemesh N, Jespersen SN. Fast imaging of mean, axial and radial diffusion kurtosis. *NeuroImage*. 2016;142:381-393.
4. Hansen B, Jespersen SN. Recent developments in fast kurtosis imaging. *Front Phys*. 2017;5:40.
5. Jensen JH, Helpert JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med*. 2005;53:1432-1440.
6. Derek P, Jones K. *Diffusion MRI Theory, Methods, and Applications: Theory, Methods, and Applications*. Oxford University Press; 2012.
7. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. *Med Phys*. 1985;12:232-233.
8. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magn Reson Med*. 1995;34:910-914.
9. Sijbers J, den Dekker A, Scheunders P, Van Dyck D. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans Med Imaging*. 1998;17:357-361.
10. Veraart J, Rajan J, Peeters RR, Leemans A, Sunaert S, Sijbers J. Comprehensive framework for accurate diffusion MRI parameter estimation. *Magn Reson Med*. 2013;70:972-984.
11. Glenn GR, Tabesh A, Jensen JH. A simple noise correction scheme for diffusional kurtosis imaging. *Magn Reson Imaging*. 2015;33:124-133.
12. Koay CG, Özarslan E, Basser PJ. A signal-transformational framework for breaking the noise floor and its applications in MRI. *J Magn Reson*. 2009;197:108-119.
13. Veraart J, Hecke WV, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magn Reson Med*. 2011;66:678-686.
14. André ED, Grinberg F, Farrher E, et al. Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLoS One*. 2014;9:e94531.
15. Coutu J-P, Chen JJ, Rosas HD, Salat DH. Non-Gaussian water diffusion in aging white matter. *Neurobiol Aging*. 2014;35:1412-1421.
16. Genç E, Fraenz C, Schlüter C, et al. Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nat Commun*. 2018;9:1-11.
17. Donat CK, Yanez Lopez M, Sastre M, et al. From biomechanics to pathology: predicting axonal injury from patterns of strain after traumatic brain injury. *Brain*. 2021;144:70-91.
18. Novikov DS, Veraart J, Jelescu IO, Fieremans E. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *NeuroImage*. 2018;174:518-538.
19. Jespersen SN, Olesen JL, Hansen B, Shemesh N. Diffusion time dependence of microstructural parameters in fixed spinal cord. *NeuroImage*. 2018;182:329-342.
20. Jensen JH, Helpert JA. Quantifying non-Gaussian water diffusion by means of pulsed-field-gradient MRI. *Proceedings of ISMRM 2003*; 2003:2154.
21. Hansen B, Lund TE, Sangill R, Jespersen SN. Experimentally and computationally fast method for estimation of a mean kurtosis. *Magn Reson Med*. 2013;69:1754-1760.
22. Hansen B, Khan AR, Shemesh N, et al. White matter biomarkers from fast protocols using axially symmetric diffusion kurtosis imaging. *NMR Biomed*. 2017;30:e3741.
23. Tabesh A, Jensen JH, Ardekani BA, Helpert JA. Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magn Reson Med*. 2011;65:823-836.
24. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. *NeuroImage*. 2013;81:335-346.
25. Mohammadi S, Tabelow K, Ruthotto L, Feiweier T, Polzehl J, Weiskopf N. High-resolution diffusion kurtosis imaging at 3T enabled by advanced post-processing. *Front Neurosci*. 2015;8:1-14.
26. Polzehl J, Tabelow K. Low SNR in diffusion MRI models. *J Am Stat Assoc*. 2016;111:1480-1490.
27. Modersitzki J. *FAIR: Flexible Algorithms for Image Registration*. SIAM; 2009.
28. Coelho S, Pozo JM, Jespersen SN, Jones DK, Frangi AF. Resolving degeneracy in diffusion MRI biophysical model parameter estimation using double diffusion encoding. *Magn Reson Med*. 2019;82:395-410.
29. Benitez A, Fieremans E, Jensen JH, et al. White matter tract integrity metrics reflect the vulnerability of late-myelinating tracts in Alzheimer's disease. *NeuroImage Clin*. 2013;4:64-71.
30. Eickhoff SB, Stephan KE, Mohlberg H, et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*. 2005;25:1325-1335.
31. Oeschger JM, Tabelow K, Mohammadi S. Violation of axial-symmetric assumption in DKI and consequences for biophysical parameter estimates across white matter. *Proceedings of ISMRM 2022*; 2022:0982.
32. Nørhøj Jespersen S. White matter biomarkers from diffusion MRI. *J Magn Reson*. 2018;291:127-140.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

## Appendix S1. Supporting information

**How to cite this article:** Oeschger JM, Tabelow K, Mohammadi S. Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study. *Magn Reson Med*. 2023;89:787-799. doi: 10.1002/mrm.29474

Supporting Information: Axisymmetric diffusion kurtosis imaging  
with Rician bias correction: A simulation study

Jan Malte Oeschger<sup>1,\*</sup>

Karsten Tabelow<sup>2</sup>

Siawoosh Mohammadi<sup>1,3</sup>

September 1, 2022

**1** University Medical Center Hamburg Eppendorf, Institute of Systems Neuroscience, Hamburg,  
Germany

**2** Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

**3** Max Planck Institute for Human Cognitive and Brain Sciences, Department of Neurophysics,  
Leipzig, Germany

\* Corresponding author:

**Name** Jan Malte Oeschger

**Institute** University Medical Center Hamburg-Eppendorf, Institute of Systems Neuroscience

**Address** Martinstraße 52, 20246 Hamburg, Germany

**E-mail** j.oeschger@uke.de

**Phone** +49-40-7410-27301



# 1 S1 Supporting Information

## 2 S1.1 Parameter estimation and the Rician noise bias (detailed)

3 Standard DKI or axisymmetric DKI parameter estimation would typically be done using the ac-  
4 quired magnitude signals  $S_{b,\vec{g}}$  and Eq. (2.1a, main text) or (2.3, main text) in the least-squares  
5 approach<sup>1;2;3</sup> (found fit results are denoted with a hat, i.e.,  $\hat{S}_0, \hat{D}, \hat{W}$ ):

$$(\hat{S}_0, \hat{D}, \hat{W}) = \operatorname{argmin}_{\tilde{S}_0, D, W} \sum_i (S_{b,\vec{g}_i} - \tilde{S}_{b,\vec{g}_i}(\tilde{S}_0, D, W))^2 \quad [\text{S1.1}]$$

6 However, this least-squares approach is built on the assumption of Gaussian distributed noise in  $S_{b,\vec{g}_i}$   
7 which is not true in reality.  $S_{b,\vec{g}}$  is a magnitude signal computed from the noise contaminated k-space  
8 data described by a complex Gaussian (standard deviation  $\sigma$ ) as the sum of squares of the measured  
9 signal intensity<sup>4</sup> from the receiver coil after it was Fourier transformed into real space. Computing  
10 the sum of squares rectifies the composite magnitude signal and leads to Rician distributed noise  
11 for one receiver coil ( $L=1$ ). Therefore, assuming Gaussian noise in MRI magnitude signals leads  
12 to a bias that propagates into the estimated parameters which is referred to as the "Rician noise  
13 bias". Eq. (S1.1) is therefore biased.

14

15 More generally, if one assumes uncorrelated noise and statistically independent receiver coils with  
16 an equivalent noise variance<sup>5</sup>, the resulting probability density function of the noisy magnitude  
17 data is given by a non-central  $\chi$ -distribution<sup>4</sup>, where  $2L$  is the number of degrees of freedom of the  
18 distribution.  $L=1$  results in the Rician distribution<sup>6;7</sup>.

19

20 The severity of the Rician noise bias depends on the SNR<sup>8</sup> because the sum of squares rectifies

21 the composite magnitude signal: the lower the SNR, the larger the bias. For RBC, we rely on an  
 22 approach outlined in<sup>8</sup> that uses the expectation value  $\mathbf{E}(S_{b,\vec{g}})$  of the noisy composite magnitude  
 23 signal. The probability density function of  $S_{b,\vec{g}}$  is a non-central  $\chi$  distribution whose expectation  
 24 value  $\mathbf{E}(S_{b,\vec{g}})$  is given by<sup>8</sup>:

$$\mathbf{E}(S_{b,\vec{g}}) = \mathbf{E}(\tilde{S}_{b,\vec{g}}(\tilde{S}_0, D, W), \sigma) = \sigma \sqrt{\frac{\pi}{2}} \cdot \mathbf{L}_{1/2}^{(L-1)}\left(\frac{\tilde{S}_{b,\vec{g}}(\tilde{S}_0, D, W)^2}{2\sigma^2}\right) \quad [\text{S1.2}]$$

25 where  $\mathbf{L}_{1/2}^{(L-1)}(x) = \frac{\Gamma(L+1/2)}{\Gamma(3/2)\Gamma(L)} \mathbf{M}(-1/2, L, x)$  is the generalized Laguerre polynomial which can be  
 26 expressed using a confluent hypergeometric function  $\mathbf{M}$ , the Gamma function  $\Gamma$  and the number of  
 27 receiver coils  $L$ . Only for simplicity of notation, in the text we neglect any possible dependence of  $\sigma$   
 28 on  $b$ ,  $\vec{g}$  or location, the employed RBC algorithm used the same  $\sigma$  in every image voxel. The SNR  
 29 dependent expectation value Eq. (S1.2) differs from the noise-free signal,  $\mathbf{E}(\tilde{S}_{b,\vec{g}}(\tilde{S}_0, D, W), \sigma) >$   
 30  $\tilde{S}_{b,\vec{g}}(\tilde{S}_0, D, W)$  with the difference decreasing with increasing SNR. Following<sup>8</sup>, we implemented a  
 31 time-efficient fitting algorithm that, unlike Equation (S1.1), accounts for Rician noise in magnitude  
 32 dMRI data by solving the optimization problem:

$$(\hat{S}_0, \hat{D}, \hat{W}) = \operatorname{argmin}_{\tilde{S}_0, D, W} \sum_i (S_{b,\vec{g}_i} - \mathbf{E}(\tilde{S}_{b,\vec{g}_i}(\tilde{S}_0, D, W), \sigma))^2 \quad [\text{S1.3}]$$

33 Estimating parameters this way is referred to as "quasi-likelihood" estimation and is denoted as  
 34 "RBC ON" in this paper. It was shown, that parameter estimation using the non-central  $\chi$  noise  
 35 statistic in a quasi-likelihood framework yields asymptotically unbiased parameter estimates<sup>9:8</sup>.

36

37 Rician bias corrected, standard DKI or axisymmetric DKI parameter estimation can be done by

38 using Equation (2.1a, main text) or Equation (2.3, main text) to compute the noise-free signal  
39 predictions  $\tilde{S}_{b,\bar{g}}$ , then using Equation (S1.2) to compute  $\mathbf{E}(\tilde{S}_{b,\bar{g}}(\tilde{S}_0, D, W), \sigma)$  and finally minimize  
40 Equation (S1.3) to estimate the framework parameters  $(\hat{S}_0, \hat{D}, \hat{W})$  for standard DKI or  $\Omega$  for ax-  
41 isymmetric DKI.

42

43 In reality, noise correlations between receiver coils occur and are non-negligible, especially for a  
44 higher number of receiver coils (32 or 64). This affects the degrees of freedom of the underlying  
45 noise statistic. However, the non-central  $\chi$  distribution can still be used as a good approximation, if  
46 an effective number of coils  $L_{\text{eff}}$  and noise variance  $\sigma_{\text{eff}}^2$  are used<sup>5</sup> for which  $L \geq L_{\text{eff}}$  and  $\sigma^2 \leq \sigma_{\text{eff}}^2$   
47 can be shown. Similarly, the generalized autocalibrating partially parallel acquisition (GRAPPA)  
48 scheme can be accounted for by specifying an effective number of coils  $L_{\text{eff}}$ , while  $L = 1$  for sensitivity  
49 encoding (SENSE)<sup>5</sup>.

## 50 **S1.2 Parameter estimation with the Gauss-Newton algorithm**

51 To minimize Eq. (S1.1) or Eq. (S1.3) time-efficiently, we have implemented a Gauss-Newton  
52 minimization algorithm<sup>10</sup> in Matlab for slice-wise and parallelizable parameter estimation on MR-  
53 images instead of using standard Matlab optimization functions. The used tools are freely avail-  
54 able online within the ACID toolbox (<http://www.diffusio.tools.com/>) for SPM. Slice-wise fit-  
55 ting refers to fitting all voxels of an image-slice at the same time which improves run-time. The  
56 implemented algorithm is highly adaptable and can fit any signal model (especially non-linear  
57 models). Gauss Newton parameter estimation approximates the search direction in parameter  
58 space based on the Jacobian and is sensitive to the initial guess. For the initial guess of the ax-  
59 isymmetric DKI fit implementation, we used code from the repository of Sune Nørhøj Jespersen:  
60 <https://github.com/sunenj/Fast-diffusion-kurtosis-imaging-DKI><sup>11</sup>.

### 61 **S1.3 Simulation study: Datasets and overview (detailed)**

62 We assessed estimation accuracy of the five AxTM as a function of the SNR in a simulation study  
63 with two classes of datasets. The first class consisted of three synthetic voxels with varying fiber  
64 alignment (defined in<sup>12</sup>). This dataset is referred to as "synthetic dataset" because it was de-  
65 rived in the context of another study<sup>12</sup> by random sampling of the parameter space of biophysical  
66 parameters and consequent derivation of the corresponding AxTM. The other class of datasets was  
67 based on an in-vivo measurement and consisted of either twelve major white matter fiber tract vox-  
68 els ("in-vivo white matter dataset") or twelve voxels from typical gray matter areas ("in-vivo  
69 gray matter dataset"). Details on both classes of datasets are given below and in Figure 2 (main  
70 manuscript). For all datasets, magnitude diffusion MRI data were simulated for varying SNRs  
71 and fitted with standard DKI and axisymmetric DKI, with and without RBC (as described in  
72 Section 2.3, main manuscript or Section S1.1) to obtain estimates of the five AxTM. Accuracy of  
73 the obtained AxTM estimates were evaluated as the absolute value of the mean percentage error  
74 (A-MPE):

$$\text{A-MPE} = 100 \cdot \frac{|\text{GT} - \overline{\text{FitResults}}(\text{SNR})|}{\text{GT}} \quad [\text{S1.4}]$$

75 Here GT refers to the ground truth and  $\overline{\text{FitResults}}$  refers to the average of the fit results over the  
76 noise samples. We evaluated the accuracy of the AxTM estimates for each estimation method by  
77 looking for the SNR after which the A-MPE was smaller 5%. The 5% threshold was considered an  
78 acceptable error in a trade-off between estimation accuracy and SNR requirement. The different  
79 setup of both simulation studies enables an isolated investigation of the effectiveness and tissue  
80 dependence of the RBC and to test the fitting methods in in-vivo data. As a summary to compare  
81 each method, we looked at the maximum SNR needed across the five AxTM for which A-MPE  
82 consistently  $< 5\%$  for all AxTM ("Maximum" column in Figure 6, main manuscript).

83

84 **Datasets:** The **synthetic dataset** consisted of three synthetic sets of AxTM (from<sup>12</sup>) describing  
85 three voxels with varying fiber alignment, one with fibers with low alignment ("LA", FA=0.067), one  
86 with fibers with moderate alignment ("MA", FA=0.24) and one with highly aligned fibers ("HA",  
87 FA=0.86). The AxTM of the three synthetic voxels are summarized in Supporting Information  
88 Table S1. Figure 4 (main manuscript) shows two areas of typical brain regions in a map of the  
89 mean of the kurtosis tensor  $\overline{W}$  where LA and HA voxels can be found and the corresponding  
90 idealized fiber stick model.

91

92 The simulated **in-vivo white matter dataset** is based on an in-vivo DWI measurement with  
93 the following measurement parameters: The sequence was a mono-polar single-shot spin-echo EPI  
94 scheme, consisting of 16 non-diffusion-weighted images ( $b = 0$  image). The diffusion weighted  
95 images were acquired at three b values (500s/mm<sup>2</sup>, 1250s/mm<sup>2</sup>, 2500s/mm<sup>2</sup>), sampled for 60 unique  
96 diffusion-gradient directions for the 1250s/mm<sup>2</sup> and 2500s/mm<sup>2</sup> shells and 30 unique directions for  
97 the 500s/mm<sup>2</sup> shell. The entire protocol was repeated with reversed phase encoding directions  
98 ("blip-up", "blip-down" correction) to correct for susceptibility-related distortions so that in total  
99 166·2 images were acquired. Other acquisition parameters were: an isotropic voxel size of (1.6mm<sup>3</sup>),  
100 FoV of 240x230x154mm<sup>3</sup>, TE = 73ms, TE = 5300ms and 7/8 partial Fourier imaging. Signal  
101 simulation in our simulation study was done with only one  $b = 0$  signal, so that the simulated  
102 sequence consisted of 151 signals per noise realization.

103 The **in-vivo white matter dataset** consists of twelve voxels extracted from four major white  
104 matter tracts (three voxels from each of the four fiber tracts, see Figure 3, main manuscript) from  
105 an in-vivo brain measurement (SNR=23.4) of a healthy volunteer. The twelve voxels were extracted  
106 from the in-vivo measurement by fitting the standard DKI framework in 12 white matter voxels



107 of the acquired in-vivo DWI magnitude images to get the corresponding 22 standard DKI tensor  
108 metrics, the derived data are therefore referred to as "in-vivo white matter". Three voxels each with  
109 HA to MA (defined through the fractional anisotropy (FA) threshold  $FA \geq 0.4^{13}$ ) were extracted  
110 from these four major white matter fiber tracts based upon the Jülich fiber atlas: the callosum body  
111 (**cb**), the corticospinal tract (**ct**), the optic radiation (**or**) and the superior longitudinal fasciculus  
112 (**slf**), see Figure 3, main manuscript. The selected voxels differ from the synthetic voxels in that  
113 here only HA and MA voxels were selected. The sets of the 12x22 in-vivo white matter standard  
114 DKI tensor metrics are documented in Table S2, the derived AxTM are found in Table S3.  
115 The **in-vivo gray matter dataset** was produced according to the same procedure used for the  
116 **in-vivo white matter dataset**, only that the voxels were selected from typical gray matter areas.  
117 The sets of the 12x22 in-vivo gray matter standard DKI tensor metrics are documented in Supporting  
118 Information Table S4, the derived AxTM are found in Supporting Information Table S5. Since white  
119 matter is the focus of this manuscript, details and results on the **in-vivo gray matter dataset**  
120 can be found in Supporting Information Section S1.3.

121

122

123 **Signal framework used for simulation:** The three synthetic voxels of AxTM were simulated  
124 with the axisymmetric DKI framework to first obtain noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$ .  
125 The twelve in-vivo white matter and gray matter voxels were simulated with the standard DKI  
126 framework to first obtain noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$ .

127

128 **Contamination with noise:** For both the **synthetic** and the **in-vivo dataset** (white matter  
129 or gray matter), the noise-free diffusion MRI signals  $\tilde{S}_{\text{noise-free}}$  were contaminated with noise for  
130 SNRs [1, 2, 3...200] and magnitude signals  $S_{\text{cont}}$  were computed. The noisy magnitude signals were

131 computed according to  $S_{\text{cont}} = |\tilde{S}_{\text{noise-free}} + \alpha + \beta i|$ , where  $\alpha, \beta \in \mathcal{N}(0, \sigma)$  are drawn from a zero  
132 mean Gaussian with standard deviation  $\sigma$ , yielding different  $\text{SNR} = \sqrt{2}S_0/\sigma$  (for one receiver coil)  
133 for a given  $S_0 = 1$ .

134

135 **Estimating the five AxTM:** Both, the simulated signals  $S_{\text{cont}}$  from the `synthetic` and the  
136 `in-vivo dataset` were fitted with axisymmetric DKI and standard DKI, with and without RBC  
137 (Section 2.3, main manuscript or Section S1.1) to obtain estimates of the AxTM whose accuracy  
138 could then be investigated as a function of SNR.

#### 139 **S1.4 Simulation studies: Details**

140 We simulated 200 SNRs:  $\text{SNR} = [1, 2, 3, \dots, 200]$ . Noise was added according to  $S_{\text{cont}} = |\tilde{S}_{\text{noise-free}} + \alpha + \beta i|$ ,  
141 where  $\alpha, \beta \in \mathcal{N}(0, \sigma)$  are drawn from a zero mean Gaussian with standard deviation  $\sigma$ , yielding  
142 different  $\text{SNR} = \sqrt{2}\frac{S_0}{\sigma}$  (for one receiver coil) for a given  $S_0 = 1$ . For every SNR, 2500 noise sam-  
143 ples were realized, i.e., 2500 · 151 pairs  $(\alpha, \beta)$  were drawn and 2500 · 151  $S_{\text{cont}}$  were calculated per  
144 SNR for every simulated voxel. These diffusion MRI magnitude signals were then fitted with the  
145 four proposed methods. For each of the 2500 noise samples per SNR, 2500 parameter estimates  
146 of  $D_{\parallel}, D_{\perp}, W_{\parallel}, W_{\perp}, \overline{W}$  were obtained and averaged to find the SNR above which the average over  
147 these 2500 noise samples had a A-MPE < 5% (`synthetic dataset`). For the `in-vivo datasets`  
148 the A-MPE was averaged per SNR across the 12 simulated voxels and the SNR above which this  
149 averaged A-MPE < 5% is reported.

150

151 For simulation of the three synthetic voxels, the axis of symmetry  $\vec{c} = (1, 0, 0)^T$  was fixed throughout  
152 the study. For data fitting, the two angles  $\theta$  and  $\phi$  that define the axis of symmetry within the  
153 axisymmetric DKI framework were variable but constrained to  $\theta, \phi \in [-2\pi, 2\pi]$  which improved

154 convergence of the fitting algorithm. Data were simulated according to the simulation scheme  
 155 described in (Section 2.5, main text).

### 156 S1.5 Simulation of in-vivo gray matter

157 To test whether the results found for the "LA" voxel translates to in-vivo applications, we addi-  
 158 tionally performed a simulation and analysis of in-vivo gray matter voxels according to the same  
 159 procedure already used for the in-vivo white matter simulation. For this, 12 voxels were extracted  
 160 from four gray matter areas (three voxels from each gray matter area) analogously to extraction of  
 161 the white matter voxels described in Section 2.5 in the main text. The four gray matter areas were  
 the frontal cortex (fc), the motor cortex (mc), the thalamus (th) and the visual cortex (vc).

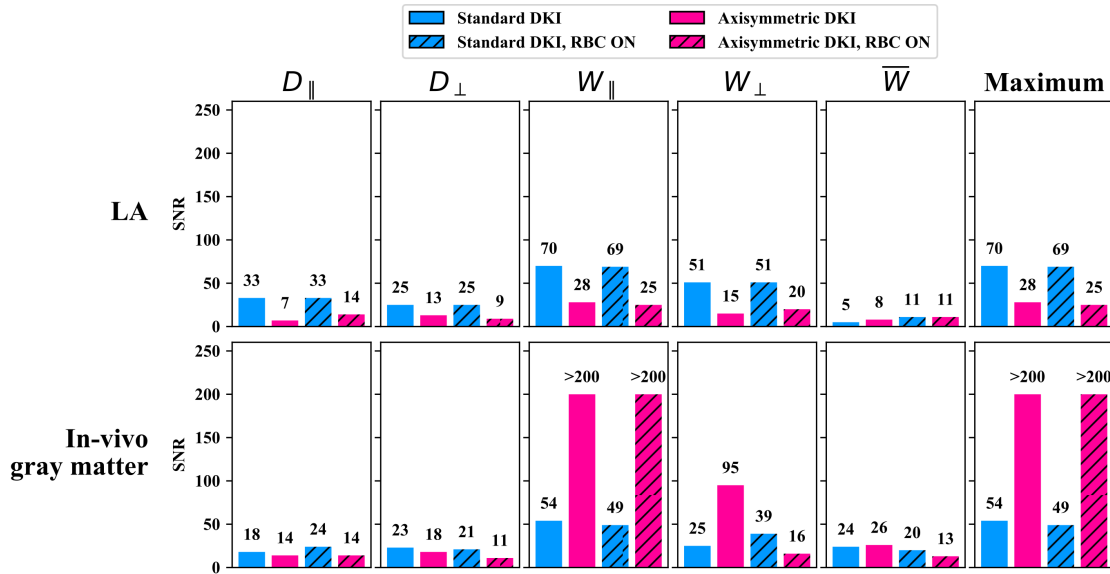


Figure S1: Signal-to-noise ratio (SNR) above which the absolute value of the mean percentage error (A-MPE, Eq. (2.8) in main text)  $< 5\%$  for the in-vivo gray matter dataset (bottom) and for the LA synthetic voxel (top). For the in-vivo gray matter dataset the A-MPE was computed in accordance with the procedure for the in-vivo white matter dataset, i.e., the A-MPE was averaged across the 12 simulated in-vivo gray matter voxels and the SNR above which this average A-MPE  $< 5\%$  is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve A-MPE  $< 5\%$  across all five AxTM.

162

163 **Results:**

164 **Axisymmetric DKI not clearly superior to standard DKI in in-vivo gray matter:** Esti-  
165 mation of  $D_{\parallel}$  and  $D_{\perp}$  was improved by using the axisymmetric DKI framework instead of standard  
166 DKI. E.g., it only required an SNR= 14 (axisymmetric DKI) instead of SNR= 18 (standard DKI)  
167 to achieve A-MPE <5% for  $D_{\parallel}$ . However, axisymmetric DKI performed much worse than standard  
168 DKI for  $W_{\parallel}$  where it needed SNRs above 200 to achieve A-MAPE <5% both with and without RBC  
169 which is in contrast to the results found for the synthetic "LA" dataset (see Figure S1). Another  
170 difference to the synthetic "LA" dataset is that RBC could substantially improve performance of  
171 the axisymmetric DKI framework for  $W_{\perp}$  where it reduced the SNR requirements from 95 without  
172 RBC to 16 with RBC.

173

## 174 **S1.6 Evaluation of precision**

175 Analogous the absolute value of the mean percentage error (A-MPE) for the bias, we have quantified  
176 the precision of the four investigated methods (standard DKI and axisymmetric DKI with and  
177 without RBC) by calculating the standard deviation in reference to the ground truth (R-STD):  
178  $R\text{-STD} = 100 \cdot \frac{\text{std}(\text{Distribution}_{\text{Estimator}})}{\text{GroundTruth}}$ . Here  $\text{std}(\text{Distribution}_{\text{Estimator}})$  is the standard deviation over  
179 the distribution of fit results for each AxTM per method and SNR. The distribution of fit results for  
180 a specific AxTM per method and SNR is made up of the 2500 fit results obtained from the simulated  
181 2500 noise samples per SNR. Analogous to the evaluation of the A-MPE, we were then interested  
182 to see at what SNR the  $R\text{-STD} < 5\%$ , i.e., at what SNR is the precision of a certain method within  
183 5% of the corresponding ground truth value. We did this analysis for the in-vivo white matter and  
184 synthetic voxels (Figure S2). We additionally calculated the outlier-robust version of the R-STD, the  
185 "R-IQR",  $R\text{-IQR} = 100 \cdot \frac{\text{IQR}(\text{Distribution}_{\text{Estimator}})}{\frac{1.3490}{\text{GroundTruth}}}$ . Here  $\text{IQR}(\text{Distribution}_{\text{Estimator}})$  is the interquartile  
186 range<sup>14</sup> of the distribution of fit results and the computed quantity  $\frac{\text{IQR}(\text{Distribution}_{\text{Estimator}})}{1.3490}$  is a robust

187 estimator for the standard deviation and hence the precision.

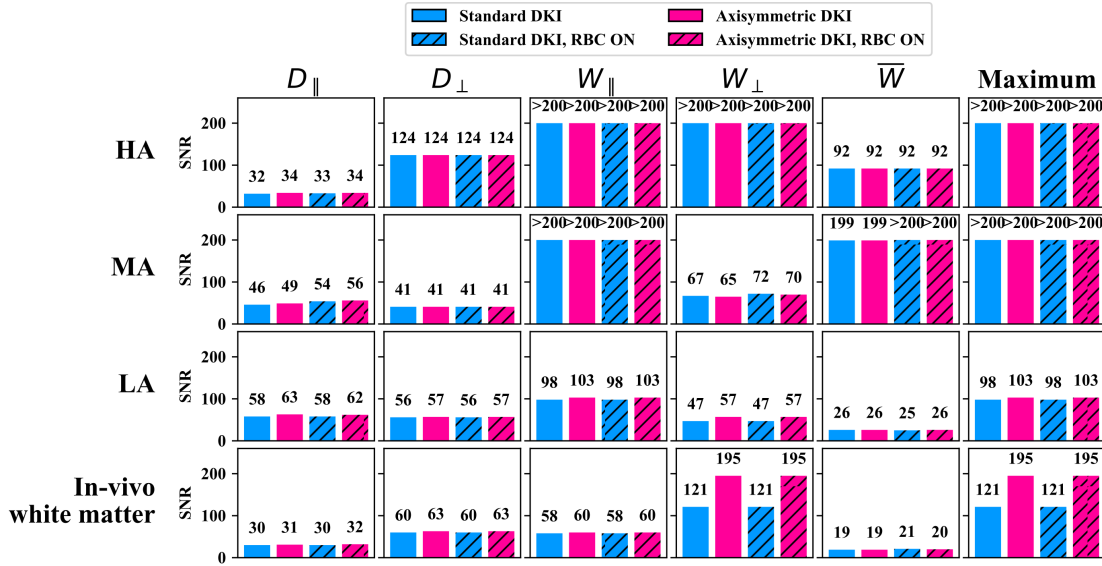


Figure S2: Signal-to-noise ratio (SNR) above which the standard deviation in reference to the ground truth (R-STD) < 5% for the synthetic dataset with high, medium and low fiber alignment ("HA", "MA", "LA") and the in-vivo white matter dataset. For the in-vivo white matter dataset, the R-STD was averaged across the 12 simulated voxels and the SNR above which this average R-STD < 5% is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve R-STD < 5% across all five AxTM.

## 188 Results:

189 **Precision is not improved by RBC or axisymmetric DKI:** Generally, higher SNRs were re-  
 190 quired to reach the R-STD < 5% threshold than reaching the A-MPE < 5% threshold. Within each  
 191 dataset (HA, MA, LA, in-vivo white matter) and for each AxTM, all four methods almost always  
 192 performed very similar to each other, regardless of RBC or DKI framework. A larger difference be-  
 193 tween methods was only observed for  $W_{\perp}$  of the in-vivo white matter dataset where standard  
 194 DKI both with and without RBC required an SNR of 121 to reach the R-STD < 5% threshold while  
 195 axisymmetric DKI both with and without RBC required an SNR of 195, see Figure S2. Further  
 196 investigation of this case revealed that the axisymmetric DKI fit results were affected by outliers for  
 197  $W_{\perp}$  in the in-vivo white matter dataset. Figure S3 shows the outlier-robust R-IQR computed

198 for the synthetic and the `in-vivo white matter` dataset. It can be seen that in this case the in-  
 199 vivo white matter results for  $W_{\perp}$  obtained with both DKI frameworks are similar (SNR of 120 for  
 200 standard DKI and 119 for axisymmetric DKI). Since estimation of the R-IQR is an outlier-robust  
 201 measure for the R-STD, this finding indicates that the observed difference in SNR requirements  
 202 between standard DKI and axisymmetric DKI for  $W_{\perp}$  in Figure S2 was caused by outliers in the  
 203 results of the axisymmetric DKI fit.

204

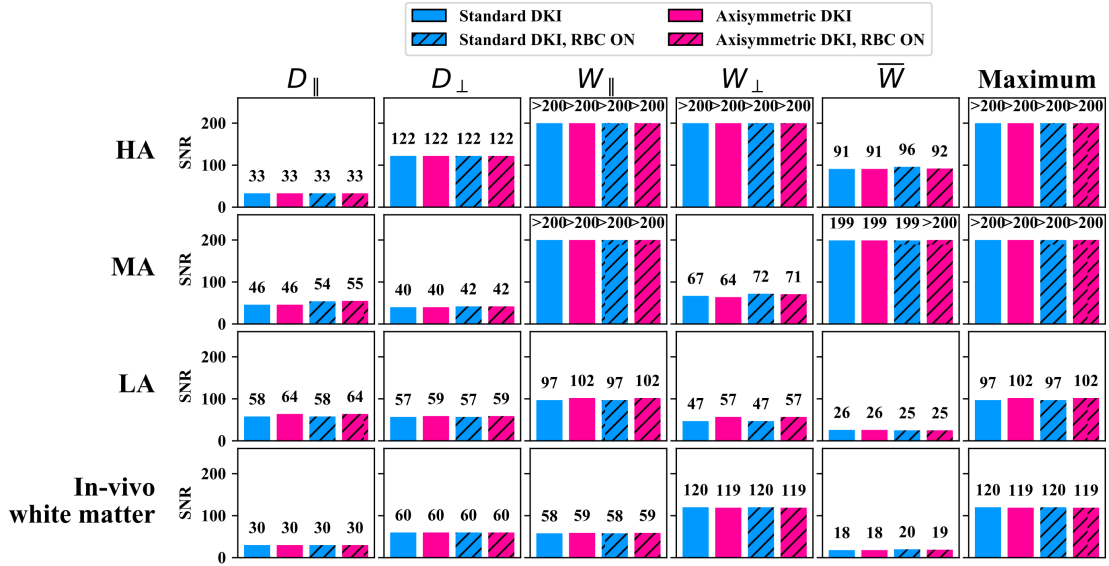


Figure S3: Signal-to-noise ratio (SNR) above which the R-IQR  $< 5\%$  for the `synthetic dataset` with high, medium and low fiber alignment ("HA", "MA", "LA") and the `in-vivo white matter dataset`. For the `in-vivo white matter dataset`, the R-IQR was averaged across the 12 simulated voxels and the SNR above which this average R-IQR  $< 5\%$  is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve R-IQR  $< 5\%$  across all five AxTM.

## 205 S1.7 Ground truth DKI datasets

Table S1: Set of synthetic AxTM,  $\tilde{S}_0$  and axis of symmetry  $\vec{c}$  used to simulate the **synthetic dataset** based on axisymmetric DKI. The **synthetic dataset** consisting of three voxels with sets of  $\{D_{\parallel}, D_{\perp}, W_{\parallel}, W_{\perp}, \bar{W}\}$  was taken from<sup>12</sup>, diffusivities are in [ $\frac{\mu\text{m}^2}{\text{ms}}$ ],  $S_0$  is in arbitrary units.

Dataset	$D_{\parallel}$	$D_{\perp}$	$W_{\parallel}$	$W_{\perp}$	$\bar{W}$	$\tilde{S}_0$	$\vec{c}$
Fibers with high alignment (HA)	1.503	0.195	1.456	0.291	0.926	1	$(1, 0, 0)^T$
Fibers with moderate alignment (MA)	1.557	1.048	0.396	0.708	0.330	1	$(1, 0, 0)^T$
Fibers with low alignment (LA)	0.457	0.408	2.901	2.702	2.770	1	$(1, 0, 0)^T$

Table S2: Ground truth in-vivo standard DKI voxels for the in-vivo white matter dataset (Figure 3, main text), shown are the diffusion and kurtosis tensor components and  $\tilde{S}_0$ , the diffusivities are in  $[\frac{\mu\text{m}^2}{\text{ms}}]$ .

Parameter	cb voxel 1	cb voxel 2	cb voxel 3	ct voxel 1	ct voxel 2	ct voxel 3
$D_{11}$	1.92726	1.62057	1.86790	0.53951	0.78384	0.73087
$D_{22}$	0.31583	0.38127	0.41435	0.36193	0.42447	0.49957
$D_{33}$	0.39808	0.39813	0.36425	1.54966	1.37428	1.29096
$D_{12}$	-0.00731	-0.26198	-0.11110	-0.15777	-0.08699	-0.05191
$D_{13}$	0.04079	-0.25240	0.09536	0.00094	-0.08181	-0.03291
$D_{23}$	0.03170	-0.09149	-0.03971	-0.04121	0.15430	0.03331
$\tilde{S}_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	4.26728	4.20207	3.73642	0.58735	1.16398	1.12725
$W_{2222}$	0.30183	0.43656	0.36961	0.22285	0.51330	0.74976
$W_{3333}$	0.43632	0.59204	0.26940	3.41702	2.14949	2.30314
$W_{1112}$	-0.15654	-0.41618	-0.18526	-0.07888	-0.06150	-0.08996
$W_{1113}$	0.10051	-0.64422	0.18910	0.00636	-0.02166	0.02767
$W_{2221}$	0.19206	-0.46340	-0.26578	-0.08910	-0.06061	-0.09517
$W_{3331}$	-0.08171	-0.15145	0.12274	-0.08214	-0.17008	-0.09078
$W_{2223}$	0.06537	-0.17548	-0.08403	0.09611	0.10937	-0.04710
$W_{3332}$	0.07430	-0.16995	-0.09129	-0.12567	0.28125	-0.09293
$W_{1122}$	0.45163	0.51896	0.50221	0.20913	0.21167	0.25976
$W_{1133}$	0.45000	0.53712	0.43661	0.52253	0.56751	0.30986
$W_{2233}$	0.15867	0.16724	0.07048	0.32234	0.42040	0.39227
$W_{1123}$	-0.00501	-0.08834	-0.03636	0.07669	0.03705	-0.00512
$W_{2213}$	-0.00262	0.00521	0.08195	-0.02098	-0.08206	-0.02479
$W_{3312}$	0.03732	-0.13410	-0.11029	-0.16710	-0.06077	-0.02733
Parameter	or voxel 1	or voxel 2	or voxel 3	slf voxel 1	slf voxel 2	slf voxel 3
$D_{11}$	1.06085	0.67273	0.69468	0.46699	0.68565	0.63911
$D_{22}$	0.59047	0.75000	1.79639	0.43803	0.48026	0.61127
$D_{33}$	1.36088	1.48673	0.49756	1.69614	1.23269	1.22297
$D_{12}$	0.12149	0.04143	-0.43159	0.05648	0.02784	0.08187
$D_{13}$	0.61736	0.22247	0.19896	0.07981	-0.07817	-0.04450
$D_{23}$	0.05184	0.26023	-0.21783	-0.20355	-0.15512	-0.09252
$\tilde{S}_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	1.18421	0.74417	0.60364	0.38485	1.27168	0.65976
$W_{2222}$	0.48582	0.71486	2.21601	0.51498	0.72106	0.83091
$W_{3333}$	1.72750	1.99702	0.37346	3.25042	2.56281	1.96113
$W_{1112}$	0.09557	0.04945	-0.18369	0.03519	0.11477	0.00824
$W_{1113}$	0.42320	0.11965	0.04012	0.01711	0.05484	0.06571
$W_{2221}$	0.22072	-0.01878	-0.49678	0.11049	-0.03851	-0.10122
$W_{3331}$	0.70421	0.13580	0.03943	0.15730	-0.32397	-0.13608
$W_{2223}$	0.13593	-0.00500	-0.40378	-0.02510	0.06858	0.06213
$W_{3332}$	0.15332	0.36637	0.09475	-0.41687	-0.24668	-0.29428
$W_{1122}$	0.19154	0.25614	0.52836	0.12737	0.22065	0.18565
$W_{1133}$	0.61697	0.26980	0.15419	0.49145	0.36651	0.53366
$W_{2233}$	0.26455	0.40726	0.34763	0.42856	0.14162	0.27141
$W_{1123}$	0.04012	0.00642	-0.11345	-0.04016	-0.01867	-0.08256
$W_{2213}$	0.06872	0.07859	0.24902	-0.01085	-0.04099	0.02417
$W_{3312}$	0.02283	0.04014	-0.03715	0.06156	0.07493	0.17797



Table S3: Ground truth AxTM of the **in-vivo dataset**, corresponding to the tensor components listed in Table S2, the diffusivities are in [ $\frac{\mu\text{m}^2}{\text{ms}}$ ]. Additionally, the deviation from axial symmetry is listed as  $\frac{|\lambda_2 - \lambda_3|}{MD}$ , where  $\lambda$  are the diffusion tensor eigenvalues and MD is the mean diffusivity.

Voxel	$D_{\parallel}$	$D_{\perp}$	$W_{\parallel}$	$W_{\perp}$	$\bar{W}$	$\frac{ \lambda_2 - \lambda_3 }{MD}$
cb voxel 1	1.928	0.356	4.276	0.401	1.425	0.117
cb voxel 2	1.714	0.343	4.549	0.387	1.535	0.346
cb voxel 3	1.883	0.382	3.798	0.240	1.279	0.091
ct voxel 1	1.551	0.450	3.427	0.471	1.267	0.444
ct voxel 2	1.413	0.585	2.373	0.762	1.245	0.461
ct voxel 3	1.295	0.613	2.294	0.903	1.221	0.299
or voxel 1	1.857	0.578	2.891	0.463	1.109	0.126
or voxel 2	1.623	0.643	2.244	0.706	1.064	0.074
or voxel 3	1.995	0.497	2.959	0.498	1.051	0.251
slf voxel 1	1.732	0.435	3.421	0.439	1.249	0.170
slf voxel 2	1.275	0.562	2.715	0.919	1.203	0.283
slf voxel 3	1.242	0.616	2.153	0.725	1.087	0.185

Table S4: Ground truth in-vivo standard DKI voxels for the in-vivo gray matter dataset, shown are the diffusion and kurtosis tensor components and  $\tilde{S}_0$ , the diffusivities are in  $[\frac{\mu\text{m}^2}{\text{ms}}]$ .

Parameter	fc voxel 1	fc voxel 2	fc voxel 3	mc voxel 1	mc voxel 2	mc voxel 3
$D_{11}$	2.50277	1.31597	2.79424	1.61384	1.11080	1.47339
$D_{22}$	2.63215	1.42647	2.91701	1.54804	1.08889	1.53996
$D_{33}$	2.64935	1.40526	3.05176	1.51739	0.97417	1.40661
$D_{12}$	-0.03462	-0.00900	0.01455	0.01249	0.02530	0.02278
$D_{13}$	0.07854	0.04254	-0.02458	-0.06448	0.03418	-0.02653
$D_{23}$	-0.06748	0.01904	0.06747	-0.06417	0.02150	0.00347
$\tilde{S}_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	0.56370	0.71667	0.42047	0.79404	0.89151	0.74619
$W_{2222}$	0.58092	0.93726	0.54043	0.65753	0.91672	0.86487
$W_{3333}$	0.62256	0.86199	0.53516	0.72181	0.84637	0.62263
$W_{1112}$	-0.00158	0.06715	-0.01741	-0.05272	0.05849	-0.02434
$W_{1113}$	0.02441	0.03391	0.00434	-0.02477	-0.03365	-0.02214
$W_{2221}$	-0.00470	-0.03735	0.02543	-0.02886	0.01141	0.03757
$W_{3331}$	0.01530	-0.01568	-0.01119	-0.01303	0.04189	0.04724
$W_{2223}$	-0.00684	0.02774	0.00758	-0.02444	0.03554	-0.02079
$W_{3332}$	-0.00801	-0.00802	0.01668	-0.00738	0.10307	0.05838
$W_{1122}$	0.19703	0.32119	0.15041	0.19543	0.33747	0.27489
$W_{1133}$	0.18887	0.32222	0.16831	0.21814	0.35201	0.23772
$W_{2233}$	0.19620	0.36126	0.15564	0.22378	0.34395	0.30244
$W_{1123}$	-0.00581	-0.02205	-0.00321	-0.02123	0.00832	-0.04963
$W_{2213}$	0.00504	0.02109	0.00011	-0.01980	0.04589	-0.01298
$W_{3312}$	0.00350	-0.02690	-0.00496	0.01040	0.02448	-0.00446
Parameter	th voxel 1	th voxel 2	th voxel 3	vc voxel 1	vc voxel 2	vc voxel 3
$D_{11}$	1.55114	0.85394	0.66375	1.11690	1.33449	1.58681
$D_{22}$	1.85125	0.83582	0.83513	1.12416	1.20188	1.61950
$D_{33}$	1.73899	0.99126	0.72202	1.19876	1.35827	1.53273
$D_{12}$	0.09370	-0.02131	0.04164	0.04721	-0.09738	-0.04635
$D_{13}$	0.07596	-0.07940	-0.03502	-0.03314	-0.05108	0.03868
$D_{23}$	-0.09738	-0.06603	0.03045	-0.05430	0.00351	0.02261
$\tilde{S}_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	0.74150	1.02877	1.07671	0.90363	0.74407	0.63246
$W_{2222}$	0.88619	1.40701	1.34001	0.69061	0.57022	0.82642
$W_{3333}$	0.84407	1.28545	1.50206	0.80784	0.83729	0.60003
$W_{1112}$	0.02028	0.14173	0.27765	0.05292	-0.09488	-0.09342
$W_{1113}$	0.03574	-0.04782	-0.01630	-0.03856	-0.07960	-0.02566
$W_{2221}$	0.05168	-0.09366	0.09460	-0.08445	-0.07263	0.02135
$W_{3331}$	0.03040	-0.12668	-0.18875	0.00338	-0.01054	0.03591
$W_{2223}$	0.01033	-0.04656	0.19234	-0.02794	-0.00807	-0.00253
$W_{3332}$	-0.06420	-0.16355	0.09208	-0.00390	-0.02928	0.06912
$W_{1122}$	0.27604	0.37859	0.75919	0.27710	0.32099	0.20954
$W_{1133}$	0.26147	0.48093	0.47265	0.27969	0.29104	0.27671
$W_{2233}$	0.29981	0.31864	0.61072	0.25239	0.26290	0.24716
$W_{1123}$	-0.01479	0.02698	-0.04394	0.03306	-0.05787	-0.04289
$W_{2213}$	-0.00818	-0.05220	-0.18916	-0.00695	0.01136	-0.01380
$W_{3312}$	0.01688	-0.03173	0.06075	0.00558	-0.03849	0.00490

Table S5: Ground truth AxTM of the **gray matter in-vivo dataset**, corresponding to the tensor components listed in Table S2, the diffusivities are in [ $\frac{\mu\text{m}^2}{\text{ms}}$ ]. Additionally, the deviation from axial symmetry is listed as  $\frac{|\lambda_2 - \lambda_3|}{MD}$ , where  $\lambda$  are the diffusion tensor eigenvalues and MD is the mean diffusivity.

Voxel	$D_{\parallel}$	$D_{\perp}$	$W_{\parallel}$	$W_{\perp}$	$\bar{W}$	$\frac{ \lambda_2 - \lambda_3 }{MD}$
fc Dataset 1	2.738	2.523	0.623	0.560	0.586	0.042
fc Dataset 2	1.440	1.354	1.004	0.880	0.905	0.082
fc Dataset 3	3.081	2.841	0.543	0.466	0.489	0.036
mc Dataset 1	1.662	1.509	0.743	0.669	0.690	0.070
mc Dataset 2	1.137	1.018	1.066	0.878	0.944	0.102
mc Dataset 3	1.547	1.436	0.900	0.714	0.773	0.054
th Dataset 1	1.914	1.614	0.922	0.764	0.829	0.153
th Dataset 2	1.041	0.820	1.519	1.130	1.216	0.102
th Dataset 3	0.849	0.686	1.556	1.257	1.521	0.139
vc Dataset 1	1.248	1.096	0.796	0.799	0.804	0.043
vc Dataset 2	1.422	1.237	0.920	0.701	0.780	0.136
vc Dataset 3	1.652	1.543	0.751	0.638	0.705	0.058

## 206 References

- 207 [1] A. Tabesh, J. H. Jensen, B. A. Ardekani, and J. A. Helpert, “Estimation of tensors and tensor-  
208 derived measures in diffusional kurtosis imaging,” *Magnetic Resonance in Medicine*, vol. 65,  
209 no. 3, pp. 823–836, 2011.
- 210 [2] J. Veraart, J. Sijbers, S. Sunaert, A. Leemans, and B. Jeurissen, “Weighted linear least squares  
211 estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls,” *NeuroImage*,  
212 vol. 81, pp. 335–346, Nov. 2013.
- 213 [3] S. Mohammadi, K. Tabelow, L. Ruthotto, T. Feiweier, J. Polzehl, and N. Weiskopf, “High-  
214 resolution diffusion kurtosis imaging at 3T enabled by advanced post-processing,” *Frontiers in*  
215 *Neuroscience*, vol. 8, 2015.
- 216 [4] C. D. Constantinides, E. Atalar, and E. R. McVeigh, “Signal-to-noise measurements in magni-  
217 tude images from NMR phased arrays,” *Magnetic Resonance in Medicine*, vol. 38, pp. 852–857,  
218 12 1997.
- 219 [5] S. Aja-Fernández and A. Tristán-Vega, “Influence of noise correlation in multiple-coil statistical  
220 models with sum of squares reconstruction,” *Magnetic Resonance in Medicine*, vol. 67, no. 2,  
221 pp. 580–585, 2012.
- 222 [6] S. O. Rice, “Mathematical analysis of random noise,” *Bell System Technical Journal*, vol. 23,  
223 no. 3, pp. 282–332, 1944.
- 224 [7] H. Gudbjartsson and S. Patz, “The Rician distribution of noisy MRI data,” *Magnetic Resonance*  
225 *in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.

- 226 [8] J. Polzehl and K. Tabelow, “Low SNR in Diffusion MRI Models,” *Journal of the American*  
227 *Statistical Association*, vol. 111, no. 516, pp. 1480–1490, 2016.
- 228 [9] H. Bunke and W. Schmidt, “Asymptotic results on nonlinear approximation of regression func-  
229 tions and weighted least squares,” *Series Statistics*, vol. 11, no. 1, pp. 3–22, 1980.
- 230 [10] J. Modersitzki, *FAIR: Flexible Algorithms for Image Registration*. Philadelphia: SIAM, 2009.
- 231 [11] B. Hansen, N. Shemesh, and S. N. Jespersen, “Fast imaging of mean, axial and radial diffusion  
232 kurtosis,” *NeuroImage*, vol. 142, pp. 381–393, Nov. 2016.
- 233 [12] S. Coelho, J. M. Pozo, S. N. Jespersen, D. K. Jones, and A. F. Frangi, “Resolving degeneracy  
234 in diffusion MRI biophysical model parameter estimation using double diffusion encoding,”  
235 *Magnetic Resonance in Medicine*, vol. 82, no. 1, pp. 395–410, 2019.
- 236 [13] A. Benitez, E. Fieremans, J. H. Jensen, M. F. Falangola, A. Tabesh, S. H. Ferris, and J. A.  
237 Helpert, “White matter tract integrity metrics reflect the vulnerability of late-myelinating  
238 tracts in Alzheimer’s disease,” *NeuroImage : Clinical*, vol. 4, pp. 64–71, Nov. 2013.
- 239 [14] P. J. Rousseeuw and C. Croux, “Alternatives to the Median Absolute Deviation,” *Journal of the*  
240 *American Statistical Association*, vol. 88, no. 424, pp. 1273–1283, 1993. Publisher: [American  
241 Statistical Association, Taylor & Francis, Ltd.].

## **6.2 Paper 2: Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates**

The scientific article "Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates" was published in "Magnetic Resonance in Medicine" (Wiley) in 2024.

## RESEARCH ARTICLE

# Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates

Jan Malte Oeschger<sup>1</sup>  | Karsten Tabelow<sup>2</sup>  | Siawoosh Mohammadi<sup>1,3,4</sup>  

<sup>1</sup>Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>2</sup>Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

<sup>3</sup>Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Sachsen Germany

<sup>4</sup>Max Planck Research Group MR Physics, Max Planck Institute for Human Development, Berlin, Germany

## Correspondence

Siawoosh Mohammadi, Max Planck Research Group MR Physics, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.  
Email: [mohammadi@mpib-berlin.mpg.de](mailto:mohammadi@mpib-berlin.mpg.de)

## Funding information

Bundesministerium für Bildung und Forschung, Grant/Award Numbers: 01EW1711A, 01EW1711B; Deutsche Forschungsgemeinschaft, Grant/Award Numbers: MO 2397/5-1, MO 2397/5-2; Deutsche Forschungsgemeinschaft, Grant/Award Numbers: MO 2397/4-1, MO 2397/4-2

## Abstract

**Purpose:** The purpose of the study is to identify differences between axisymmetric diffusion kurtosis imaging (DKI) and standard DKI, their consequences for biophysical parameter estimates, and the protocol choice influence on parameter estimation.

**Methods:** Noise-free and noisy, synthetic diffusion MRI human brain data is simulated using standard DKI for a standard and the fast “199” acquisition protocol. First the noise-free “baseline” difference between both DKI models is estimated and the influence of fiber complexity is investigated. Noisy data is used to establish the signal-to-noise ratio at which the baseline difference exceeds noise variability. The influence of protocol choices and denoising is investigated. The five axisymmetric DKI tensor metrics ( $AxTM$ ), the parallel and perpendicular diffusivity and kurtosis and mean of the kurtosis tensor are used to compare both DKI models. Additionally, the baseline difference is also estimated for the five parameters of the WMTI-Watson model.

**Results:** The parallel and perpendicular kurtosis and all of the WMTI-Watson parameters had large baseline differences. Using a Westin or FA mask reduced the number of voxels with large baseline difference, that is, by selecting voxels with less complex fibers. For the noisy data, precision was worsened by the fast “199” protocol but adaptive denoising can help counteract these effects.

**Conclusion:** For the diffusivities and mean of the kurtosis tensor, axisymmetric DKI with a standard protocol delivers similar results as standard DKI. Fiber complexity is one main driver of the baseline differences. Using the “199” protocol worsens precision in noisy data but adaptive denoising mitigates these effects.

## KEYWORDS

apparent differences, axisymmetric DKI, bias, biophysical parameters, standard DKI, white matter

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Magnetic Resonance in Medicine* published by Wiley Periodicals LLC on behalf of International Society for Magnetic Resonance in Medicine.

## 1 | INTRODUCTION

Diffusion kurtosis imaging (DKI) has increasingly been used to study the neuronal tissue microstructure and derive biophysical parameters relevant for understanding brain function and impact of disease<sup>1-3</sup> during the past ten years. DKI is a more complex extension of the diffusion tensor imaging (DTI) framework and provides diffusion kurtosis metrics that can provide complementary information<sup>4-6</sup> to DTI. However, the increased complexity comes with an increase in acquisition time. Since time is a limited resource in scientific and especially clinical settings, needing more time poses a major hurdle for a more extensive implementation and application of DKI.

Axisymmetric DKI was recently introduced as a more acquisition time efficient DKI model.<sup>7-9</sup> Axisymmetric DKI reduces the parameter space by imposing additional symmetry assumptions, that is, axisymmetrically organized fibers in the imaged tissue structure and axisymmetric DKI (eight parameters) can be fitted with less data than is required by standard DKI (22 parameters). One time-efficient acquisition scheme that leverages axisymmetric DKI's data demand advantages is the fast "199" scheme<sup>7,10</sup> that relies on a total of 19 images (18 diffusion weighted images and one nondiffusion weighted image), only. However, parameter estimation from fewer diffusion weighted images is more susceptible to noise and it can be expected that the signal-to-noise ratio (SNR) requirements for accurate parameter estimates are higher when fitting data acquired with the fast "199" protocol compared to a standard diffusion MRI (dMRI) protocol. To mitigate the influence of noise, denoising algorithms can be used.<sup>11,12</sup>

An important parameter subset existing in both axisymmetric DKI and standard DKI are the five axisymmetric DKI tensor metrics (AxTM), the parallel and perpendicular diffusivity ( $D_{\parallel}$  and  $D_{\perp}$ ) and kurtosis ( $W_{\parallel}$  and  $W_{\perp}$ ) and mean of the kurtosis tensor ( $\overline{W}$ ). Since the AxTM are attainable with both DKI models, they are perfectly suited to be used for a model comparison. The five AxTM are also directly related to the five parameters of the biophysical standard model,<sup>13</sup> the axon water fraction AWF, axon dispersion  $\kappa$ , parallel and perpendicular extra-axonal diffusivities  $D_{e,\parallel}$  and  $D_{e,\perp}$  and intra axonal diffusivity  $D_a$ , here estimated with the WMTI-Watson model.<sup>14,15</sup>

It has been suggested that the symmetry assumptions made in axisymmetric DKI are likely a reasonable approximation to diffusion in major white matter fiber bundles, for example, occurring in white matter.<sup>7</sup> Violation of these additional symmetry assumptions, for example, in white matter voxels containing crossing fibers, might lead to a

deviation of axisymmetric DKI fit results from their standard DKI reference counterpart.

In this work, we simulate synthetic white matter data using standard DKI as forward model to explicitly include complex, non-axisymmetric fiber configurations. We hypothesize that any observed deviation between both DKI model variants is due to an error in axisymmetric DKI rooted in an underlying complex and thus non-axisymmetric fiber configuration. To test this hypothesis, voxel selection masks are used to filter out voxels with high fiber complexity that likely break axisymmetric DKI's symmetry assumptions. Furthermore, we establish the "baseline difference" between axisymmetric DKI and standard DKI AxTM estimates from noise-free data. The baseline differences are inherent to axisymmetric DKI and will always be there with respect to standard DKI. Another topic is the influence of noise on the standard deviation (SD) of axisymmetric DKI fit results. Here the question at which SNR the baseline difference between both DKI models becomes bigger than axisymmetric DKI's SD is investigated. This will establish an SNR regime where the standard deviation caused by noise dominates parameter estimation and at what SNR the baseline difference between both DKI models becomes visible which is here referred to as the "tipping point". To investigate this, first a noise simulation study is performed where the SNR space is densely sampled to precisely establish the tipping point without using denoising methods. Then, we use Multi-shell Position-Orientation Adaptive Smoothing (msPOAS), an adaptive denoising algorithm that is preserving tissue boundaries without introducing blurring<sup>11,16,17</sup> to establish if the tipping point can be reached for realistic SNRs in voxels where the axisymmetric conditions were best fulfilled. In both noise studies, the influence of the acquisition protocol (fast "199" protocol vs. a standard dMRI protocol) is investigated as an additional variable. Finally, the median bias based on the two acquisition protocols is quantified in five well known fiber tracts for a typical SNR found in dMRI datasets.<sup>18</sup> Throughout the work, the noise-free estimates of the five AxTM or the biophysical parameters based on standard DKI are used as a ground truth reference and axisymmetric DKI fit results are compared to them on a voxel-wise basis, hence the baseline difference between axisymmetric DKI and standard DKI is also referred to as "bias".

## 2 | METHODS

A detailed description of the standard DKI model and the axisymmetric DKI model is provided in the Sections S1.1 and S1.2 but can also be found in References 7 and 19.



## 2.1 | Dataset

### 2.1.1 | Acquisition

Multishell, in vivo dMRI data with 153 diffusion gradient directions and  $b$ -values of  $0 \frac{\text{s}}{\text{mm}^2}$  (18 images),  $550 \frac{\text{s}}{\text{mm}^2}$  (30 directions),  $1100 \frac{\text{s}}{\text{mm}^2}$  (45 directions), and  $2500 \frac{\text{s}}{\text{mm}^2}$  (60 directions) was acquired from a healthy volunteer at 3T with: FOV of  $200 \times 203 \times 170 \text{ mm}^3$  at 1.7 mm isotropic resolution, TE = 75 ms, TR = 5800 ms, gradient separation  $\Delta = 38.8 \text{ ms}$  and gradient pulse duration  $\delta = 13.3 \text{ ms}$ . Image reconstruction at the scanner from the multi-channel information was done using a sum of squares algorithm where all channels had the same weights. The following pre-processing was applied to the dMRI data using the ACID toolbox<sup>20</sup> (in that order): eddy current and motion artifact correction using the ACID “ECMO-CO” module, susceptibility artifact correction using the ACID “HySCO” module and Rician bias correction using the ACID “Rician bias correction” module that uses a second moment approach and acts on the signals of the individual diffusion weighted images as described in Reference 21. The noise level  $\sigma$  was calculated using the ACID “Noise estimation” module with the “repeated measures” method where  $\sigma$  is estimated as the averaged SD over the diffusion measurements of the highest diffusion shell in the ventricles, see Reference 20. Additionally, we acquired multiparameter mapping data<sup>22</sup> on the same subject and calculated the  $R_1$  map using the hMRI toolbox,<sup>23</sup> for anatomical visualization.

The in vivo dMRI data used for this study was acquired with the help of a human research participant. The participant provided written informed consent. The local ethics committees at University Medical Center Hamburg-Eppendorf approved the study (PV5141).

### 2.1.2 | Generation of synthetic data

The acquired multishell, in vivo dMRI data was fitted with standard DKI to obtain the 22 standard DKI tensor metrics per voxel which were then used for generation of noise-free, synthetic dMRI data using standard DKI as a forward model with the same diffusion shells and gradient directions that were used to acquire the data.

## 2.2 | Biophysical parameters

The framework presented in Reference 13,31, here referred to as “WMTI-Watson,” was used to establish an analytical connection between the five AxTM  $\Omega = \{\bar{W}, W_{\parallel}, W_{\perp}, D_{\parallel}, D_{\perp}\}$  and the biophysical parameters  $\beta = \{\text{AWF}, \kappa, D_{e,\perp}, D_{e,\parallel}, D_a\}$ . A detailed description of the framework can be found in Section S1.3.

We neglected a potential model error of the WMTI-Watson model with respect to the biological tissue ground truth and assumed the biophysical parameters based on AxTM estimates from standard DKI to be the ground truth. In this study, we focused on quantifying the propagation of the error in the AxTM estimates into the respective WMTI-Watson parameter estimates introduced by axisymmetric DKI.

## 2.3 | Computation of difference between both DKI models and substantially differing voxels

The estimated parameters, either the set of AxTM  $\Omega$  or the set of biophysical parameters  $\beta$  were estimated based on standard DKI and axisymmetric DKI and compared using the voxel-wise absolute percentage error (A-PE):

$$\text{A-PE} = 100 \cdot \frac{|\theta_{\text{standardDKI}} - \theta_{\text{axisymmetricDKI}}|}{\theta_{\text{standardDKI}}} \quad (1)$$

Here  $\theta$  is an element of either  $\Omega$  or  $\beta$  and the subscript indicates whether the parameter was estimated based on standard DKI or axisymmetric DKI. If  $\text{A-PE} > 5\%$ , the corresponding voxel was classified as a “substantially differing voxel” (SDV). The study focused on white matter only, to obtain the white matter mask, we segmented the  $R_1$  map into tissue probability maps (TPM) and thresholded the white matter TPM ( $\text{TPM} > 0.9$ ), see green contour in Figure 2. To summarize the results we estimated a) the number of SDV in white matter in percent and b) the median A-PE in the population of SDV. We implemented the condition  $\theta \in \Omega \geq 0$  and  $\theta \in \beta \geq 0$  because negative diffusivities are non-physical and AWF and  $\kappa$  are  $\geq 0$  by definition. Furthermore, kurtosis estimates in the healthy brain have been found<sup>24</sup> well above 0.

### 2.3.1 | Influence of voxel selection masks on number of substantially differing voxels

To identify voxels in white matter that likely break the axisymmetric tissue symmetry assumption,<sup>25,26</sup> two different masks were used: a fractional anisotropy (FA) mask and a so-called “Westin mask.” To generate the FA mask, white matter voxels with  $\text{FA} \geq 0.55$  were selected based on the FA of an unidirectional phantom, see Reference 27. This voxel selection is referred to as “FA mask.” To generate the Westin mask, white matter voxels that fulfill the conditions imposed by a threshold for the Westin indices computed via the diffusion tensor eigenvalues  $\lambda$ ,  $C_L = \frac{\lambda_1 - \lambda_2}{\lambda_1} \geq 0.4$ ,  $C_P = \frac{\lambda_2 - \lambda_3}{\lambda_1} \leq 0.2$  and  $C_S = \frac{\lambda_3}{\lambda_1} \leq 0.35$ <sup>26</sup> were

investigated, this voxel selection is referred to as “Westin mask.” For comparison, a third mask containing the entire white matter was used.

## 2.4 | Noise simulations

### 2.4.1 | Simulation of noisy dMRI data and fast “199” protocol

The noise-free, in vivo human brain data  $\tilde{S}_{\text{noise-free}}$  were noise contaminated according to  $S_{\text{cont}} = |\tilde{S}_{\text{noise-free}} + \alpha + \beta i|$ , where  $\alpha, \beta \in \mathcal{N}(0, \sigma)$  were drawn from a zero mean Gaussian with SD  $\sigma$ . A special property of axisymmetric DKI is that due to its reduced parameter space, it needs fewer measurements and—at a minimum—all five AxTM can be estimated from a fast two shell setup with nine distinct gradient directions  $n^{(i+/-)}$ <sup>7,10</sup> and one additional  $b = 0$  measurement. Here, the fast protocol was simulated for b-values of  $b = 1100 \frac{\text{s}}{\text{mm}^2}$  and  $b = 2500 \frac{\text{s}}{\text{mm}^2}$ .

### 2.4.2 | Single voxel analysis

Here we asked: at what SNR is the baseline difference between standard DKI and axisymmetric DKI bigger than the SD of axisymmetric DKI? The SNR at which this happens is referred to as “tipping point.” To investigate this, we simulated 12 single voxels (see Section S1.7) selected from the white matter of the in vivo human brain data set for densely sampled SNRs = [1, 2, 3, ...140] according to the procedure reported above for 2500 noise samples for the standard MRI protocol described in Section 2.1 and additionally for the fast “199” protocol.<sup>7,10</sup> Then, axisymmetric DKI was fitted to the simulated data. To evaluate the simulation and make it comparable to the noise-free study regarding the SDV, the differences between standard DKI and axisymmetric DKI were normalized to the corresponding average AxTM values in white matter and expressed as a percentage thereof.

### 2.4.3 | White matter Westin mask analysis with adaptive denoising

To investigate whether the tipping point can be reached for typical SNRs of published DKI protocols, we performed another noise simulation study that used the adaptive denoising “msPOAS” module of the ACID toolbox with two set-ups for the adaptation parameter  $\lambda$ ,  $\lambda = 10$  and  $\lambda = 100$  on noisy, whole brain simulation data. The adaptation bandwidth  $\lambda$  controls the adaptivity of msPOAS,<sup>11</sup> ranging from  $\lambda = 0$  (complete adaptation)

where the original image is unchanged, to  $\lambda = \infty$  (non-adaptive denoising) similar to, for example, Gaussian smoothing. Smoothing was performed in the white matter mask because msPOAS needs coherent regions to work properly, while parameter estimation was done in the Westin mask. For the simulated SNRs, values of published protocols were chosen: SNR = [5, 15, 30, 52] (SNR: 52, see Reference 28, SNRs: 30, 15, 5, see Reference 29) and SNR 100. That data was then fitted with axisymmetric DKI in the Westin mask (see Section 2.3), since the Westin mask turned out to be most effective in reducing the number of SDV (see Section 3.2). To evaluate this analysis,  $n = 100$  noise samples were simulated, the SD per voxel was calculated and the median was computed. Second, the percentage of voxels reaching the tipping point was calculated.

### 2.4.4 | Quantification of bias in fiber tracts

To quantify the difference between both DKI models based on the two acquisition protocols under realistic conditions, the whole brain human dMRI data was simulated and fitted with the fast “199” and the standard acquisition protocol for SNR = 39, which was the SNR reported in one of the original fast “199” protocol studies.<sup>18</sup> Data were smoothed using msPOAS and  $\lambda = 100$ . “Bias” in this context refers to the difference of the noisy, voxel-wise fit results to the ground truth (noise-free, standard DKI fit results), again computed as the A-PE, see Equation (1). Results were evaluated with a specific focus on five well-known fiber tracts: corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf), and posterior corona radiata (pcr). The fiber tracts were identified with the JHU-ICBM-DTI-81 white matter atlas<sup>30</sup> for which it was nonlinearly registered to the subject space of the in vivo dMRI data using the spatial normalization tool in SPM12, see Figure 1.

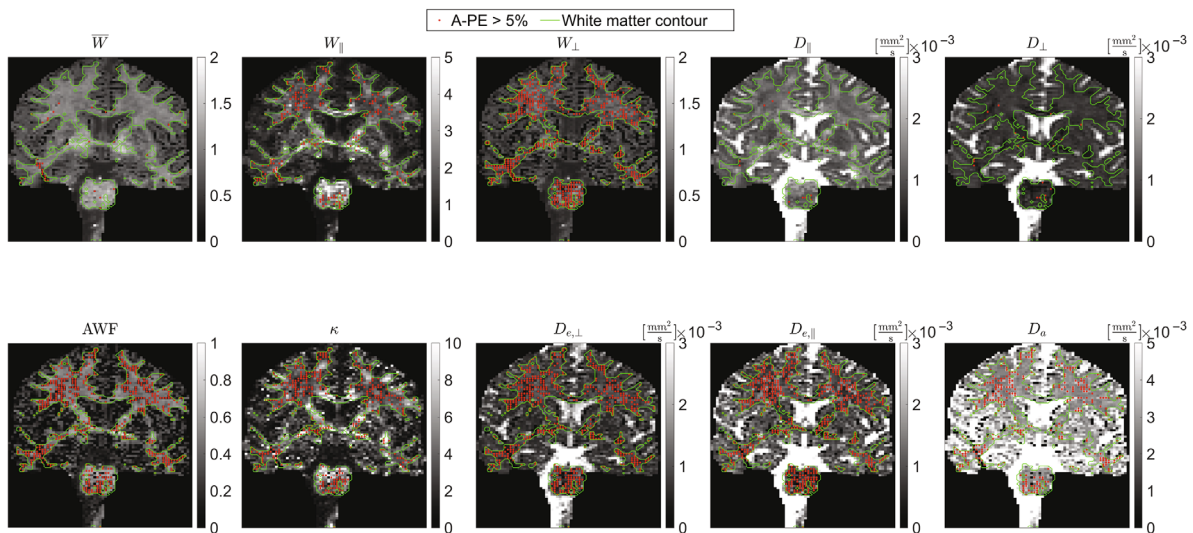
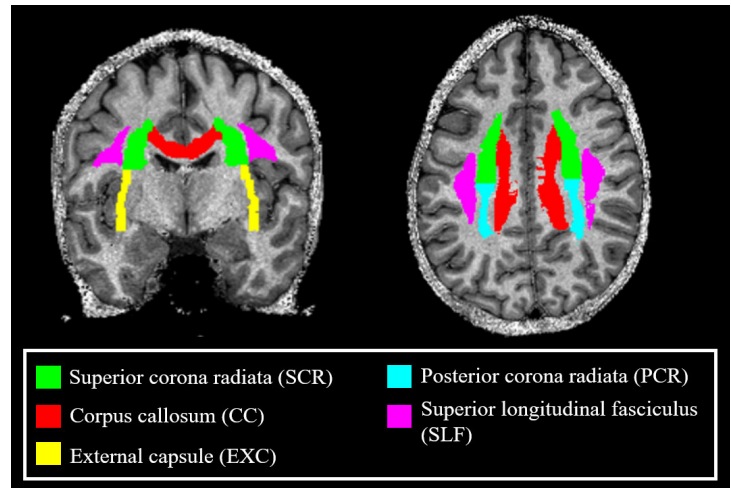
## 3 | RESULTS

### 3.1 | Summary measures: Number of SDV highly parameter dependent and biophysical parameters affected the most. Median A-PE similar across all parameters

#### 3.1.1 | Differences between AxTM across the white matter using the two DKI models

Figure 2 shows the spatial distribution of SDV (see Section 2.3) as red dotted voxels in a slice of the AxTM and biophysical parameters, Figure 3 summarizes the

**FIGURE 1** The five regions of interest corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf) and posterior corona radiata (pcr) identified with the JHU-ICBM-DTI-81 white matter atlas indicated in the coregistered  $R_1$  image of the human brain in vivo dataset used in this study.



**FIGURE 2** Examples of the axisymmetric DKI tensor metrics (top) and biophysical parameters (bottom) in a slice of the human brain data used in this study. The green contour outlines the white matter, the red dots indicate voxels where the A-PE  $\geq 5\%$  (“substantially differing voxels”). The red barplots of the top row in Figure 3 quantify the percentage of substantially differing voxels in the whole white matter.

number of SDV and the median A-PE in that population using barplots.

The number of SDV in the white matter mask is highly parameter dependent (Figure 3A.1), for example, only 1% of  $D_{\parallel}$ ,  $D_{\perp}$  and 2% of the mean of the kurtosis tensor  $\bar{W}$ . The  $W_{\parallel}$  (22%) and  $W_{\perp}$  (51%) were affected much more, see also spatial distribution of SDV. However, the median difference in the SDV across the AxTM was more similar and ranged between 7% ( $\bar{W}$ ) and 11% ( $W_{\perp}$ ).

### 3.1.2 | Differences between biophysical parameters across the white matter based on the two DKI models

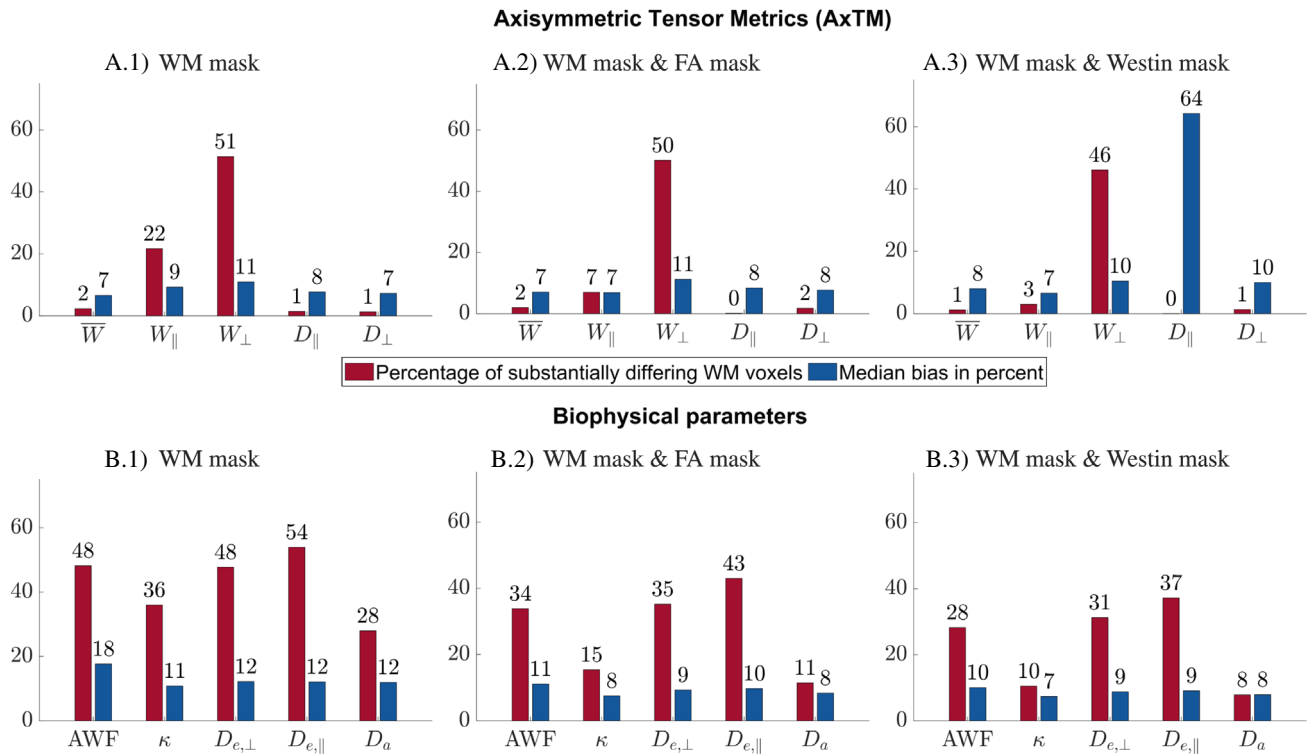
Both the number of SDV and the median A-PE in the SDV population (11%–18%) was higher for the biophysical

parameters than the AxTM. Again, the percentage of SDV was parameter dependent and spanned from 28% ( $D_a$ ) to 54% ( $D_{e,\parallel}$ ), see Figure 3B.1.  $D_{e,\parallel}$ ,  $D_{e,\perp}$  and AWF had most SDV while  $D_a$  and  $\kappa$  had the least. Figure S1 in Section S1.4 documents the underlying A-PE histogram distributions.

For neither the AxTM nor the biophysical parameters Figure 2 revealed a spatial distribution pattern of the SDV in the depicted white matter slice.

### 3.2 | Influence of voxel selection mask on number of SDV and median bias

Confining the analysis to white matter voxels in the Westin mask reduces the number of SDV more than confining the analysis to white matter voxels in the FA mask. For the AxTM, the median bias remained approximately the same



**FIGURE 3** Barplots summarizing the number of substantially differing voxels (SDV) of the axisymmetric DKI tensor metrics (AxTM) (top) and biophysical parameters (bottom). Shown is the rounded number of SDV (red barplots) and the median absolute percentage error (A-PE) or “bias” in those voxels (blue barplots). The subplot’s titles indicate the subset of voxels that were analyzed, that is, (A.1, B.1) “WM mask” = white matter voxels were analyzed, (A.2, B.2) “WM mask & FA mask” = white matter voxels in the FA mask were analyzed and (A.3, B.3) “WM mask & Westin mask” = white matter voxels in the Westin mask were analyzed. Number of voxels in the white matter mask: 101 521, number of voxels in the WM mask & FA mask analysis: 28 741, number of voxels in the WM mask & Westin mask analysis: 20 527.

for all three voxel masks, except for  $D_{\parallel}$  in the Westin mask analysis that was heavily affected by very few outliers with high median bias (64%), while in the case of the biophysical parameters the Westin mask and FA mask could reduce the median bias.

### 3.2.1 | AxTM

Both, the FA mask and the Westin mask, reduced the number of SDV for all AxTM, compare Figure 3A.2 and A.3. They were most effective for  $W_{\parallel}$  and least effective for  $W_{\perp}$ . For example, for  $W_{\perp}$  the number of SDV could be reduced from 51% in the whole white matter to 46% in the Westin mask while for  $W_{\parallel}$  the number of SDV could be reduced from 22% to 3%. In general, the Westin mask was more effective in reducing the SDV than the FA mask. Interestingly, the spread of the median A-PE remained similar between  $\approx 7\%$  to  $\approx 11\%$  when applying both masks, see Figure 3A.2, 3A.3, except for  $D_{\parallel}$  in the Westin mask that was affected by outliers.

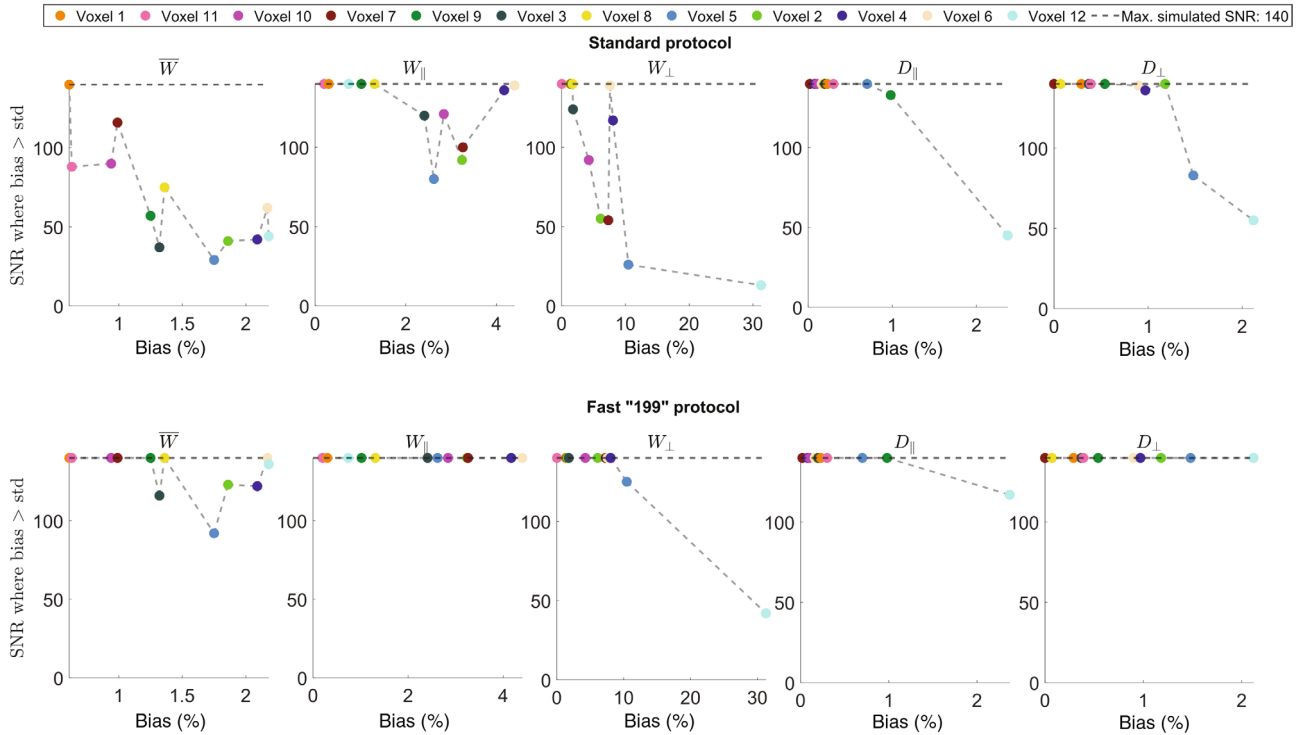
### 3.2.2 | Biophysical parameters

The same trend as for the AxTM was also observed for the biophysical parameters with regards to the number of SDV: both masks reduced the number of SDV in all parameters significantly and the Westin mask was more effective than the FA mask. But, as opposed to the AxTM, the spread of the median A-PE was also reduced from (11%–18%) to (8%–10%) for the Westin mask and (8%–11%) for the FA mask, see Figure 3B.2, B.3.

## 3.3 | Noise simulations

### 3.3.1 | Single voxel analysis

Figure 4 shows the SNR of the tipping point for each of the 12 simulated voxels. The term “bias” here refers to the normalized difference between standard DKI and axisymmetric DKI in %.



**FIGURE 4** Signal-to-noise ratio (SNR) at which the tipping point (bias > std) is reached, computed for  $n = 2500$  noise realizations for the 12 simulated voxels. Top shows the results for the standard diffusion kurtosis imaging (DKI) protocol, the bottom shows the results for the fast “199” protocol. The connecting gray line’s purpose is to indicate the trend that for higher biases the SNR required to reach the tipping point tends to be lower.

### 3.3.2 | Standard protocol

The likelihood of finding an SNR threshold  $\leq 140$  increases for increasing axisymmetric DKI biases, that is, from left to right on the x-axis.  $W_{\perp}$  and  $\overline{W}$  had the highest number of voxels for which a tipping point below the simulation limit (SNR 140) could be found, where the lowest tipping point was found at SNR = 13 for the  $W_{\perp}$  voxel with the highest bias (voxel 12). For the diffusion parameters  $D_{\perp}$  and  $D_{\parallel}$ , most voxels did not reach the tipping point up to SNRs = 140.

### 3.3.3 | Fast “199” protocol

For the fast “199” protocol, too, the likelihood of finding an SNR tipping point  $\leq 140$  increases for increasing biases. However, much fewer voxels reached the tipping point below the simulation limit compared to the standard protocol. Again the lowest tipping point was found for the  $W_{\perp}$  voxel with the highest bias (voxel 12).

### 3.3.4 | White matter Westin mask analysis with adaptive denoising

Figure 5 shows the median SD in the Westin mask (blue and cyan markers), the median difference between

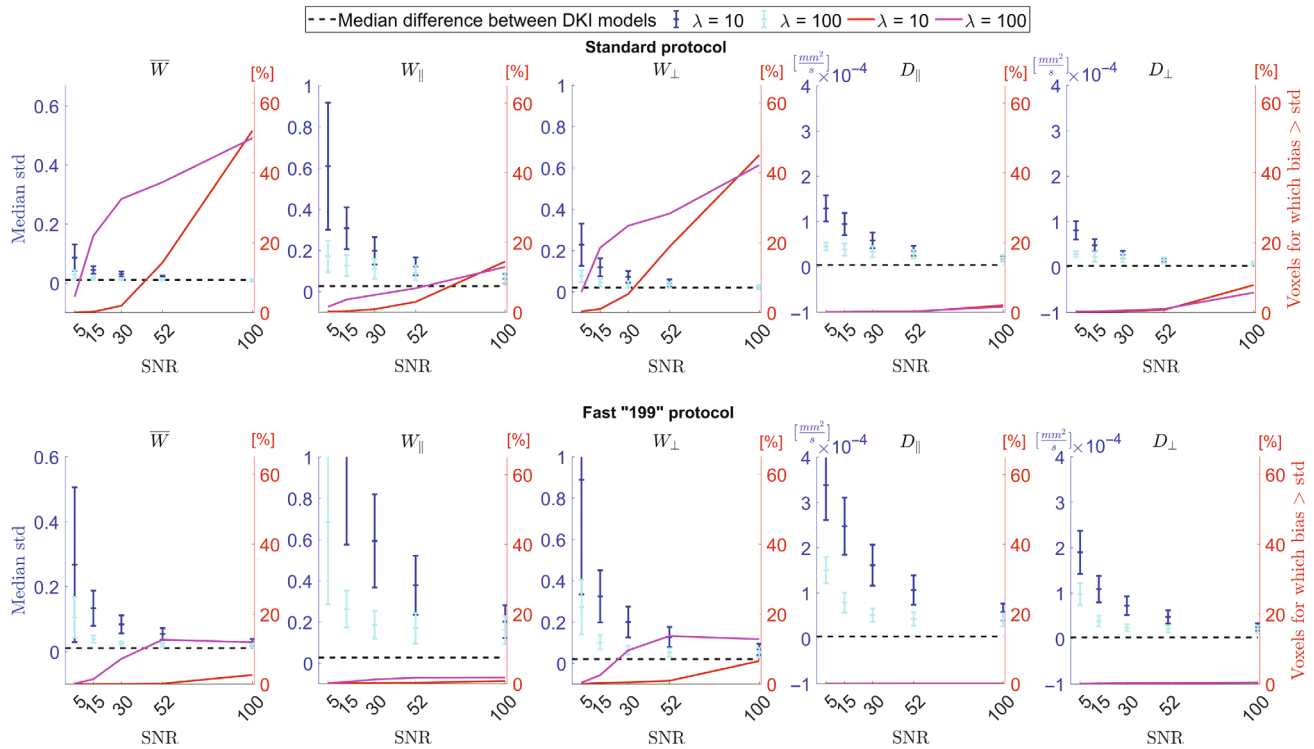
standard DKI and axisymmetric DKI in the Westin mask (dashed black line) and the percentage of voxels crossing the voxel-wise tipping point (red and magenta markers, right y-axis).

### 3.3.5 | Standard protocol

The voxel-wise tipping point was reached for realistic, higher SNRs up to 52 predominantly for the kurtosis parameters but the percentage depended on the adaptation parameter  $\lambda$ . For example, for  $\overline{W}$  at SNR = 52,  $\approx 37\%$  of voxels reached the tipping point for  $\lambda = 100$  while it was only  $\approx 14\%$  for  $\lambda = 10$ . A similar pattern was observed for  $W_{\perp}$  where for SNR = 52,  $\approx 28\%$  of voxels reached the tipping point for  $\lambda = 100$  and only  $\approx 19\%$  in case of  $\lambda = 10$ . For  $W_{\parallel}$  at SNR = 52 the number of voxels reaching the tipping point was  $\approx 7\%$  for  $\lambda = 100$  and below 5% for  $\lambda = 10$ . For the diffusion parameters the number of voxels was close to zero for all realistic SNRs up to 52.

### 3.3.6 | Fast “199” protocol

The number of voxels reaching the tipping point was lower when the fast protocol was used compared to the standard protocol. For SNR = 52 and  $\lambda = 100$  the tipping point was



**FIGURE 5** Median SD in subset of white matter voxels in Westin mask, computed for  $n = 100$  noise samples (blue and cyan datapoints, left y-axis), median difference between standard diffusion kurtosis imaging (DKI) and axisymmetric DKI in that subsample of voxels (dashed black line) and percentage of voxels crossing the voxel-wise tipping point (red and magenta data, right y-axis). Shown are the results when fitting the standard MRI protocol (top) and the fast “199” protocol (bottom) data, see Sections 2.1 and 2.4.

reached by  $\approx 13\%$  of voxels in case of  $\bar{W}$  and  $\approx 14\%$  in case of  $W_{\perp}$ . For SNRs up to 52 and in case of  $\lambda = 10$  the number of voxels reaching the tipping point was always close to zero for all parameters.

### 3.3.7 | Median SD

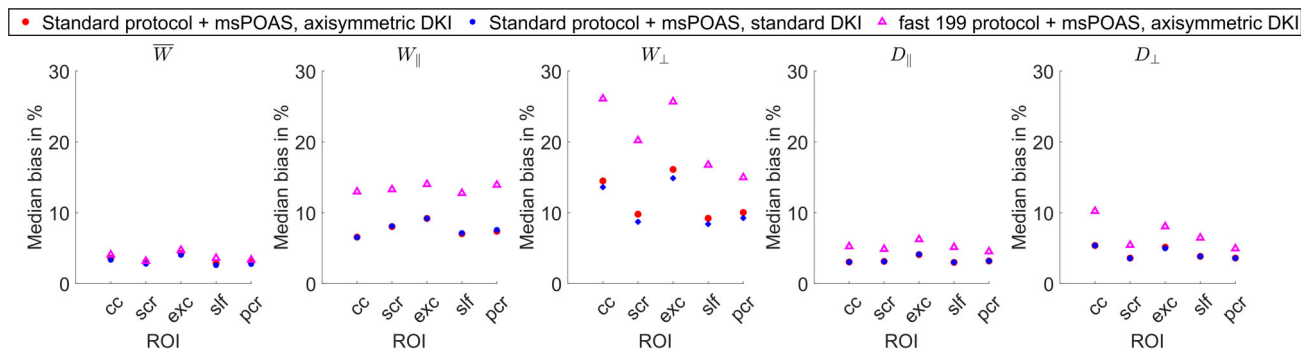
Independent of the protocol, the median SD always became smaller for increasing SNRs and was also smaller for a higher adaptation parameter  $\lambda$ .

### 3.3.8 | Quantification of bias in fiber tracts

Figure 6 shows the median bias computed in the five fiber tracts corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf), and posterior corona radiata (pcr). The median bias of standard DKI and axisymmetric DKI is always very similar when a standard MRI protocol is used. For  $\bar{W}$  the combination of the fast protocol and axisymmetric DKI performs very similarly to when a standard protocol is used while for the other parameters it performs worse.

## 4 | DISCUSSION

In this work, we found that depending on the parameter, there is a significant number of voxels where axisymmetric DKI and standard DKI inherently differ more than 5%. For  $D_{\parallel}$ ,  $D_{\perp}$  and  $\bar{W}$  we found little differences between the two DKI models, while  $W_{\perp}$  and  $W_{\parallel}$  showed larger differences. All five axisymmetric DKI based biophysical parameters were strongly different from their standard DKI based counterparts. Introduction of two voxel selection masks that reduced the number of voxels containing more complex fiber configurations mitigated the observed differences between both DKI models (especially for  $W_{\parallel}$  and the biophysical parameters) suggesting that fiber complexity is one cause of the observed baseline differences between both DKI models. The noise simulations revealed the challenges of the fast “199” protocol. For a higher but still realistic SNR of 52, the single voxel noise simulation without denoising showed that the baseline differences between axisymmetric DKI and standard DKI are often invisible when a standard MRI acquisition protocol is used (the most tipping points were reached for  $\bar{W}$ : 5/12 voxels, that is,  $\approx 42\%$  and  $W_{\perp}$ : 2/12 voxels, that is,  $\approx 17\%$ ) and that they are almost always invisible when the fast “199” protocol is used (only 1/12 voxels for  $W_{\perp}$  reached the



**FIGURE 6** Median bias for noisy human brain data at a signal-to-noise ratio (SNR) = 39, smoothed with multi-shell position-orientation adaptive Smoothing (msPOAS) using  $\lambda = 100$  in the five white matter fiber tracts corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf) and posterior corona radiata (pcr). The median bias was computed for a simulation of the standard protocol that was fitted with standard diffusion kurtosis imaging (DKI) (blue data points) and axisymmetric DKI (red data points) and for a simulation of the fast “199” protocol that was fitted with axisymmetric DKI (pink data points).

tipping point, that is,  $\approx 8\%$ ). We also found that the SNR of the tipping point depended on the absolute value of the bias. However, in another noise simulation, adaptive denoising improved axisymmetric DKI parameter estimation from the fast protocol to a point where  $\approx 13\%$  of  $\bar{W}$  voxels and  $\approx 14\%$   $W_{\perp}$  voxels reached the tipping point for realistic SNRs while it did not substantially improve estimation from the standard protocol. Furthermore, our fiber tract analysis revealed that under realistic experimental conditions (realistic SNR and one repeated measurement only), the combination of the fast protocol with adaptive denoising performed similar to standard protocol estimates for  $\bar{W}$ .

#### 4.1 | Noise-free simulation: baseline differences between AxTM across the white matter using the two DKI models

The AxTM capture different properties of diffusion in tissue and it is not surprising that the observed differences between both DKI models is AxTM dependent. We used the number of SDV to quantify the baseline differences between axisymmetric DKI and standard DKI. We found that the diffusion parameters  $D_{\parallel}$  and  $D_{\perp}$  and the mean of the kurtosis tensor  $\bar{W}$  have very few SDV compared to  $W_{\parallel}$  and  $W_{\perp}$ .

With a median A-PE of 7%–11% in white matter, the error made in axisymmetric DKI might be acceptable depending on the application. Purely judging from the number of SDV, the diffusion parameters were “safest” where only 1% of white matter voxels were SDV, followed by the mean of the kurtosis tensor  $\bar{W}$  with 2%. It can generally be expected that the kurtosis parameters are more sensitive to a model error since they are quadratic in the  $b$ -value  $b$  compared to the linear diffusion parameter counterparts. The propagation of error when fitting the axisymmetric DKI model to dMRI data therefore will

be more severe for the kurtosis parameters. Interestingly, the number of SDV increases tenfold from  $\bar{W}$  to  $W_{\parallel}$  and roughly doubles from  $W_{\parallel}$  to  $W_{\perp}$ . The reason for this trend still needs to be explored.

#### 4.2 | Noise-free simulation: differences between biophysical parameters across the white matter based on the two DKI models

The number of SDV were significantly enhanced in the biophysical parameters (in case of the whole white matter mask up to 54%, see Figure 3) compared to the AxTM from which they were computed. A reason for the enhancement might be that all five AxTM are required to estimate the biophysical parameters, see Section S1.3, and that the connection is complex and nonlinear. The observed AxTM differences could therefore be amplified due to nonlinear effects but also synergistically enhance the number of SDV in the biophysical parameters. However, the found median A-PE of 11%–18% might be acceptable depending on the study.

Also, similar to Reference 31, here the “–” branch tended to yield physically unfeasible, constant high (50)  $\kappa$  values where the objective function did not have a well defined minimum for  $\kappa$ . Furthermore, for this branch  $D_{e,\parallel} > D_a$  which in healthy white matter was found to be the biologically invalid solution by most studies.<sup>32</sup> We therefore did not report the results of the – branch.

#### 4.3 | Noise-free simulation: influence of voxel selection mask on number of SDV

The difference between standard DKI and axisymmetric DKI is likely linked to fiber complexity that brakes the symmetry assumptions of axisymmetric DKI. Reducing the

number of voxels with complex fiber structures using voxel selection masks reduced the number of SDV.

The Westin mask reduced the number of SDV in both the AxTM and the biophysical parameters, supporting the hypothesis that differences between both DKI models are linked to the underlying fiber complexity. However, the Westin mask effectiveness was parameter dependent and, for example, worked particularly well for  $W_{\parallel}$  where it more than quartered the number of SDV while it had much smaller effects on  $W_{\perp}$ . The FA mask performed very similarly but reduced the number of SDV slightly less than the Westin mask. In particular, both masks struggled with reducing the number of SDV of  $W_{\perp}$ .

These findings indicate that next to fiber complexity other factors may play a role in determining the difference between axisymmetric DKI and standard DKI estimates. Other factors that could play a role are differences in glia-cell density,<sup>33</sup> the axonal diameters distribution or the fiber dispersion. With regards to the FA, for example, in vivo tissue FA could also be influenced by factors like the degree of myelination or axon density. This means that the conclusion “if  $FA \geq 0.55$  then the voxel has a unidirectional fiber configuration” is not necessarily strictly true and voxels in the FA mask might still have a complex fiber structure. The smaller effectiveness of both the FA mask and the Westin mask on  $W_{\perp}$  could originate from axisymmetric DKI generally oversimplifying estimation of  $W_{\perp}$  since it is directly estimated as one model parameter instead of calculated from three separate tensor metrics as in standard DKI, see, for example, Reference 34.

#### 4.4 | Inter-dependence of A-PE and difference in main fiber orientation

It was shown that axisymmetric DKI produces the same results as standard DKI if two requirements<sup>35</sup> are met, see Section S1.6. Fulfillment of condition b) was not explicitly checked in Reference 35 where differences between axisymmetric DKI and standard DKI were reported. The degree to which the main fiber orientations, estimated with both DKI models, differ can be quantified with the angle  $\phi$  between them. The majority of angles  $\phi$  were between  $\approx 1$  and 5 degrees in white matter (Figure S3) demonstrating that condition b) is not fulfilled in most cases. Investigating the dependency of the A-PE on angle  $\phi$  using density scatter plots showed an inter-dependency predominantly for  $W_{\perp}$  and  $W_{\parallel}$ , Figure S3. For these parameters these findings indicate that at least to some extent, there is a causal relationship between  $\phi$  and the A-PE.

It could be ruled out that the observed differences between both DKI models in this study is only due to violating condition a) of Reference 35 by implementing a

log-of-signals fit demonstrating that this fit implementation still produced different fit results for both DKI models, see Section S1.5.

## 4.5 | Noise simulations

### 4.5.1 | Single voxel analysis

Simulating 12 single voxels for  $n = 2500$  noise realizations revealed that the tipping point at which the axisymmetric DKI inherent bias is larger than the SD of the axisymmetric DKI parameter estimates was only reached for high, unrealistic SNRs  $\geq 52$  in most cases. Therefore, the differences between both DKI models will likely not be visible under realistic experimental conditions if techniques like denoising are not used. We found that the tipping point depended on the axisymmetric DKI bias with the strong tendency that the higher the bias, the earlier the tipping point is reached. Furthermore and in line with the main findings that were dealing with the number of SDV per AxTM, it was harder to find the tipping point for the diffusion parameters than the kurtosis parameters, most likely because the diffusion parameters were less biased than the kurtosis parameters and therefore the requirements for the SD to reach the tipping point were too high.

### 4.5.2 | White matter Westin mask analysis with adaptive denoising

Intuitively one would expect axisymmetric DKI to have better precision than standard DKI because of its reduced parameter space compared to standard DKI. However, in an earlier study performed on simulations of noisy dMRI data<sup>19</sup> we have found that axisymmetric DKI does not improve precision compared to standard DKI. Furthermore, in the “Single voxel analysis” we have found that without denoising, the tipping point at which the difference between axisymmetric DKI and standard DKI becomes visible typically requires very high SNRs, often above 140. We therefore were interested to see if the tipping point could be reached for realistic SNRs (up to 52) in combination with additional adaptive denoising (msPOAS) with two settings of adaptation parameter  $\lambda$ .<sup>11</sup> The influence of  $\lambda$  was significant for low SNRs while its effect was decreasing the higher the SNR was. We found that for a realistic but high SNR 52, the tipping point was only reached for  $\bar{W}$  and  $W_{\perp}$  in roughly 1/3 of voxels in the Westin mask if a standard MRI protocol is used and for a high adaptation parameter  $\lambda = 100$ . In all other cases the number of voxels reaching the tipping point was significantly lower and especially for the diffusion parameters it was close to 0 for realistic SNRs. This indicates that



the bias introduced by the symmetry assumption made in axisymmetric DKI is not going to be visible under realistic, experimental conditions in most of the voxels.

#### 4.5.3 | Quantification of bias in fiber tracts

The noise simulations are more realistic than the noise-free simulation but introduce additional factors that could cause a difference between standard DKI and axisymmetric DKI, for example, model dependent noise susceptibility and the Rician bias. Noisy data was simulated at SNR = 39 and in an earlier study<sup>19</sup> this SNR was found to be sufficient to reduce the difference between both DKI models below 5% in in vivo white matter, see Fig. 6 in Reference 19.

If the noise induced Rician noise bias can be considered negligible if it is below 5% at SNR = 39, the observed difference between axisymmetric DKI and standard DKI estimated for the standard protocol is most likely the axisymmetric DKI inherent bias. Consistent with the findings concerned with the median bias found in Figure 3, the fiber tract analysis, Figure 6, showed that the median bias of the kurtosis AxTM per fiber tract was highest for  $W_{\perp}$ , followed by  $W_{\parallel}$  and lowest for  $\bar{W}$ . Also, both DKI models based on the standard protocol performed very similarly (except for  $W_{\perp}$ ) which again demonstrates that the tipping point is not reached in most cases under realistic conditions, that is, noise is dominating the bias and not the model differences.

However, some of the findings in the fiber tract analysis were counter-intuitive. In an other study, the corpus callosum was found to predominantly have voxels with a single fiber orientation<sup>36</sup> and could therefore be considered to have less of a complex structure than other tracts making it more likely to fulfill the assumptions made in axisymmetric DKI. However, the corpus callosum had the highest median bias for 2 of 5 AxTM. Furthermore, in Reference 36 the slf tract was reported to host many voxels with two fiber orientations, that is, fiber crossings which would make it prone to failing the assumption of axisymmetric DKI but the slf was one of the tracts with the lowest differences between both DKI models. This finding, again, points to other factors than fiber complexity, here in the form of fiber crossings, playing a role in determining the difference between standard DKI and axisymmetric DKI, as already discussed in Section 4.3.

#### 4.5.4 | Fast “199” protocol versus standard protocol

Axisymmetric DKI’s major advantage is its reduced data demand which can save acquisition time if it is used with

a fast acquisition protocol. Combination with additional, well known image acceleration techniques like multislice and GRAPPA could speed up acquisition even more. Moreover, the shorter sampling scheme was reported to have better contrast-to-noise ratio compared to standard DKI in rat brains with induced stroke<sup>37</sup> and it is less susceptible to motion artifacts which is even more important in clinical settings than for basic research where the subjects are more used to scanning and thus less probable to move during the acquisition. We set out to investigate potential problems going along with axisymmetric DKI and its assumptions and the fast “199” image acquisition scheme.

For the single voxel analysis, investigating the influence of the MRI protocol on the tipping point revealed that the fast “199” protocol had even higher SNR requirements to reach the tipping point than the standard protocol. This is plausible since reducing the number of diffusion gradients in the MRI protocol makes parameter estimation more prone to the influence of noise because fitting is done with fewer data points. Furthermore, the effective SNR,  $SNR_{\text{eff}} = SNR \cdot \sqrt{\text{number of } b = 0 \text{ images}}$  is higher for the standard acquisition protocol, since here more and optimally distributed  $b = 0$  images are acquired compared to the fast “199” protocol where only one  $b = 0$  image is acquired. In the “white matter Westin mask analysis with adaptive denoising” analysis we found that when switching to the fast “199” measurement protocol, the number of individual voxels that reached the tipping point in total was smaller compared to the standard protocol which is plausible as described above. It is also important to consider that fewer acquired diffusion gradients can go along with disadvantages, for example, FSL’s eddy current correction (“eddy”) as a preprocessing step will be challenging since it ideally requires to densely sample the whole sphere with the diffusion gradients and a minimum of approximately 10–15 directions for a  $b = 1500$  shell with increasing demands for higher shells, see <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/eddy>. Note that these requirements do not apply to the eddy current and motion correction algorithm implemented in the ACID toolbox that was used to pre-process the measured dMRI data used for synthetic data generation since here the eddy-current field is estimated independent of the diffusion gradient directions.<sup>38</sup> Furthermore, Rician bias correction depends on an accurate estimate of the noise level  $\sigma$  and the chosen “repeated measures” method (see Section 2.1) works better with more acquired diffusion gradients since it  $\sigma$  is estimated as the SD over the highest diffusion shell. All in all, parameter estimation with the fast “199” protocol is challenged by a higher bias and the baseline difference between both DKI models is likely not visible for realistic SNRs.

## 5 | CONCLUSION

Axisymmetric DKI offers advantages like a reduced data demand that is relevant in scientific and clinical practice. We asked the question whether these advantages are counteracted by an error related to the intrinsic simplification of the model. We found that in the noise-free case and when using a standard MRI acquisition protocol  $D_{\perp}$ ,  $D_{\parallel}$ , and  $\overline{W}$  could be estimated with few SDV with respect to their standard DKI counterpart. For all other parameters, that is,  $W_{\perp}$ ,  $W_{\parallel}$  and the biophysical parameters, the number of SDV was high. However, the number of SDV can be reduced if a Westin mask is used, suggesting that fiber complexity is one main driver for the differences between both DKI models. Under realistic conditions with noise in the acquired data, the model inherent baseline difference between both DKI variants requires very high SNRs to become visible since it is hidden in the SD otherwise. Furthermore, our results showed that the fast “199” protocol is particularly vulnerable to the effects of noise in the dMRI data but also that adaptive denosing can help counteract these effects. Here, we only contrasted two extreme acquisition protocols, one suited for standard DKI and one optimized for fast axisymmetric DKI. To find the “sweet spot” between noise susceptibility and time reduction in the acquisition protocol a follow-up study is required.

### ACKNOWLEDGEMENT

Open Access funding enabled and organized by Projekt DEAL.

### FUNDING INFORMATION


This work was supported by the German Research Foundation (DFG Priority Program 2041 “Computational Connectomics”, [MO 2397/5-1; MO 2397/5-2], by the Emmy Noether Stipend: (MO 2397/4-1 and MO 2397/4-2) and by the BMBF (01EW1711A and B) in the framework of ERA-NET NEURON.

### DATA AVAILABILITY STATEMENT

The open-source ACID toolbox<sup>20</sup> (<http://www.diffusiontools.org/>) for SPM contains the estimation methods for standard and axisymmetric DKI used in this study (commit hash: 73e1d23dc). Furthermore, a repository hosting the code used for simulation and analysis of the data in this study is available on Github at [https://github.com/quantitative-mri-and-in-vivo-histology/apparent\\_axDKI\\_differences](https://github.com/quantitative-mri-and-in-vivo-histology/apparent_axDKI_differences).


### ORCID

Jan Malte Oeschger  <https://orcid.org/0000-0003-0237-923X>

Karsten Tabelow  <https://orcid.org/0000-0003-1274-9951>

Siawoosh Mohammadi  <https://orcid.org/0000-0003-1311-9636>

### TWITTER

Siawoosh Mohammadi  [siawooshmn](https://twitter.com/siawooshmn)

### REFERENCES

- Coutu J-P, Chen JJ, Rosas HD, Salat DH. Non-Gaussian water diffusion in aging white matter. *Neurobiol Aging*. 2014;35:1412-1421.
- Genç E, Fraenz C, Schlüter C, et al. Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nat Commun*. 2018;9:1905.
- Donat CK, Yanez Lopez M, Sastre M, et al. From biomechanics to pathology: predicting axonal injury from patterns of strain after traumatic brain injury. *Brain*. 2021;144:70-91.
- Zhuo J, Xu S, Proctor JL, et al. Diffusion kurtosis as an in vivo imaging marker for reactive astrogliosis in traumatic brain injury. *Neuroimage*. 2012;59:467-477. Neuroergonomics: The human brain in action and at work.
- Steven AJ, Zhuo J, Melhem ER. Diffusion kurtosis imaging: an emerging technique for evaluating the microstructural environment of the brain. *Am J Roentgenol*. 2014;202:W26-W33.
- Taha HT, Chad JA, Chen JJ. DKI enhances the sensitivity and interpretability of age-related DTI patterns in the white matter of UK biobank participants. *Neurobiol Aging*. 2022;115:39-49.
- Hansen B, Shemesh N, Jespersen SN. Fast imaging of mean, axial and radial diffusion kurtosis. *Neuroimage*. 2016;142:381-393.
- Hansen B, Khan AR, Shemesh N, et al. White matter biomarkers from fast protocols using axially symmetric diffusion kurtosis imaging. *NMR Biomed*. 2017;30:e3741.
- Hansen B, Jespersen SN. Recent developments in fast kurtosis imaging. *Front Phys*. 2017;5:40.
- Hansen B, Lund TE, Sangill R, Jespersen SN. Experimentally and computationally fast method for estimation of a mean kurtosis. *Magn Reson Med*. 2013;69:1754-1760.
- Tabelow K, Mohammadi S, Weiskopf N, Polzehl J. POAS4SPM: a toolbox for SPM to denoise diffusion MRI data. *Neuroinformatics*. 2015;13:19-29.
- Tournier J-D, Smith R, Raffelt D, et al. Mrtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage*. 2019;202:116137.
- Novikov DS, Veraart J, Jelescu IO, Fieremans E. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *Neuroimage*. 2018;174:518-538.
- Alexander DC, Dyrby TB, Nilsson M, Zhang H. Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR Biomed*. 2019;32:e3841.
- Novikov DS, Fieremans E, Jespersen SN, Kiselev VG. Quantifying brain microstructure with diffusion MRI: theory and parameter estimation. *NMR Biomed*. 2019;32:e3998.
- Becker S, Tabelow K, Voss H, Anwander A, Heidemann R, Polzehl J. Position-orientation adaptive smoothing of diffusion

- weighted magnetic resonance data (poas). *Med Image Anal.* 2012;16:1142-1155.
17. Becker S, Tabelow K, Mohammadi S, Weiskopf N, Polzehl J. Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by mspos. *Neuroimage.* 2014;95:90-105.
  18. Hansen B, Jespersen SN. Data for evaluation of fast kurtosis strategies, b-value optimization and exploration of diffusion MRI contrast. *Sci Data.* 2016;3:160072.
  19. Oeschger JM, Tabelow K, Mohammadi S. Axisymmetric diffusion kurtosis imaging with Rician bias correction: a simulation study. *Magn Reson Med.* 2022;89:787-799. doi:10.1002/mrm.29474
  20. David G, Fricke B, Oeschger JM, et al. Acid: a comprehensive toolbox for image processing and modeling of brain, spinal cord, and post-mortem diffusion mri data. *bioRxiv.* 2023.
  21. André ED, Grinberg F, Farrher E, et al. Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLoS One.* 2014;9:1-15.
  22. Callaghan MF, Josephs O, Herbst M, Zaitsev M, Todd N, Weiskopf N. An evaluation of prospective motion correction (pmc) for high resolution quantitative mri. *Front Neurosci.* 2015;9:1-9.
  23. Tabelow K, Balteau E, Ashburner J, et al. Hmri – a toolbox for quantitative mri in neuroscience and clinical research. *Neuroimage.* 2019;194:191-210.
  24. Lätt J, Nilsson M, Wirestam R, et al. Regional values of diffusional kurtosis estimates in the healthy brain. *J Magn Reson Imaging.* 2013;37:610-618.
  25. Westin CF, Maier SE, Mamata H, Nabavi A, Jolesz FA, Kikinis R. Processing and visualization for diffusion tensor MRI. *Med Image Anal.* 2002;6:93-108.
  26. Fieremans E, Jensen JH, Helpert JA. White matter characterization with diffusional kurtosis imaging. *Neuroimage.* 2011;58:177-188.
  27. Tournier JD, Yeh C-H, Calamante F, Cho K-H, Connelly A, Lin C-P. Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. *Neuroimage.* 2008;42:617-625.
  28. Veraart J, Nunes D, Rudrapatna U, et al. Noninvasive quantification of axon radii using diffusion MRI. *Elife.* 2020;9:e49855.
  29. Manzano-Patron J-P, Moeller S, Andersson JL, Yacoub E, Sotiropoulos SN. Denoising diffusion MRI: considerations and implications for analysis. *Imag Neurosci.* 2023;2:1-29.
  30. Hua K, Zhang J, Wakana S, et al. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage.* 2008;39:336-347.
  31. Jespersen SN, Olesen JL, Hansen B, Shemesh N. Diffusion time dependence of microstructural parameters in fixed spinal cord. *Neuroimage.* 2018;182:329-342.
  32. Jelescu IO, Palombo M, Bagnato F, Schilling KG. Challenges for biophysical modeling of microstructure. *J Neurosci Methods.* 2020;344:108861.
  33. Schurr R, Mezer AA. The glial framework reveals white matter fiber architecture in human and primate brains. *Science.* 2021;374:762-767.
  34. Tabesh A, Jensen JH, Ardekani BA, Helpert JA. Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magn Reson Med.* 2011;65:823-836.
  35. Nørhøj Jespersen S. White matter biomarkers from diffusion MRI. *J Magn Reson.* 2018;291:127-140.
  36. Jeurissen B, Leemans A, Tournier J-D, Jones DK, Sijbers J. Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Hum Brain Mapp.* 2013;34:2747-2766. doi:10.1002/hbm.22099
  37. Wu Y, Kim J, Chan S-T, et al. Comparison of image sensitivity between conventional tensor-based and fast diffusion kurtosis imaging protocols in a rodent model of acute ischemic stroke. *NMR Biomed.* 2016;29:625-630. doi:10.1002/nbm.3506
  38. Mohammadi S, Möller HE, Kugel H, Müller DK, Deppe M. Correcting eddy current and motion effects by affine whole-brain registrations: evaluation of three-dimensional distortions and comparison with slice-wise correction. *Magn Reson Med.* 2010;64:1047-1056.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**Appendix S1.** Supporting information: Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates.

**How to cite this article:** Oeschger JM, Tabelow K, Mohammadi S. Investigating apparent differences between standard DKI and axisymmetric DKI and its consequences for biophysical parameter estimates. *Magn Reson Med.* 2024;1-13. doi: 10.1002/mrm.30034

Supporting Information: Investigating apparent differences  
between standard DKI and axisymmetric DKI and its  
consequences for biophysical parameter estimates

Jan Malte Oeschger<sup>1</sup>      Karsten Tabelow<sup>2</sup>      Siawoosh Mohammadi<sup>1,3,4,\*</sup>

January 12, 2024

**1** University Medical Center Hamburg-Eppendorf, Institute of Systems Neuroscience, Hamburg,  
Germany

**2** Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

**3** Max Planck Research Group MR Physics, Max Planck Institute for Human Development,  
Berlin, Germany

\* Corresponding author:

**Name** Siawoosh Mohammadi

**Institute** Max Planck Institute for Human Development, Max Planck Research Group MR Physics

**Address** Lentzeallee 94, 14195, Berlin

**E-mail** mohammadi@mpib-berlin.mpg.de

**Phone** +49 30 82406-0

## Supporting Information

### S1.1 Standard DKI signal representation

For a given diffusion weighting  $b$  and diffusion gradient  $\vec{g} = (g_1, g_2, g_3)^\top$ , the noise-free DKI signal can be represented as in<sup>1,2</sup>:

$$S_{b,\vec{g}}(S_0, D, W) = S_0 \exp \left[ -bD + \frac{b^2}{6} \left( \frac{\text{Tr}(D)}{3} \right)^2 W \right] \quad [\text{S1.1a}]$$

$$D = \sum_{i,j=1}^3 g_i g_j D_{ij} \quad [\text{S1.1b}]$$

$$W = \sum_{i,j,k,l=1}^3 g_i g_j g_k g_l W_{ijkl} \quad [\text{S1.1c}]$$

where  $D_{ij}$  are the diffusion tensor entries,  $W_{ijkl}$  are the kurtosis tensor entries and  $S_0$  is the non-diffusion-weighted signal ( $b = 0 \frac{\text{s}}{\text{mm}^2}$ ).

From the tensors metrics in D and W, the AxTM can be directly computed:  $D_{\parallel} = \lambda_1$  where  $\lambda_1$  is the first eigenvalue of the diffusion tensor D,  $D_{\perp} = \frac{\lambda_2 + \lambda_3}{2}$ .  $W_{\parallel}$  and  $W_{\perp}$  can be computed from the fitted W tensor according to formulas 11 and 12 from<sup>4</sup>:  $W_{\parallel} = W(\nu_1) = W_{1111}$ , where  $\nu_1$  is the first eigenvector of the corresponding diffusion tensor and  $W_{\perp} = 3/8(W_{2222} + W_{3333} + 2W_{2233})$ .  $\bar{W}$  can be computed according to Eq. 10 from<sup>5</sup>:  $\bar{W} = 1/5(W_{1111} + W_{2222} + W_{3333} + 2W_{1122} + 2W_{1133} + 2W_{2233})$  (in<sup>5</sup>: 1=x, 2=y, 3=z).

## S1.2 Axisymmetric DKI

Axisymmetric DKI<sup>4</sup> assumes symmetric diffusion around an axis of symmetry  $\vec{c}$  inside an imaging voxel. Mathematically, this assumption leads to axisymmetric diffusion and kurtosis tensors with a drastically reduced number of independent tensor parameters compared to standard DKI (from 15 to 3 parameters for the kurtosis tensor and from 6 to 2 parameters for the diffusion tensor). Apart from the tensors, axisymmetric DKI additionally contains two parameters for the axis of symmetry.

With the axis of symmetry  $\vec{c}$  parameterized by the inclination  $\theta$  and azimuth  $\phi$ :  $\vec{c} = \begin{pmatrix} \sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta \end{pmatrix}$ ,

the diffusion and kurtosis tensors can be determined according to<sup>4</sup>:

$$D = D_{\perp} \mathbf{I} + (D_{\parallel} - D_{\perp}) \vec{c} \vec{c}^T \quad [\text{S1.2}]$$

and

$$W = \frac{1}{2} (10W_{\perp} + 5W_{\parallel} - 15\overline{W}) \mathbf{P} + W_{\perp} \mathbf{\Lambda} + \frac{3}{2} (5\overline{W} - W_{\parallel} - 4W_{\perp}) \mathbf{Q}$$

where  $\Psi = \{D_{\parallel}, D_{\perp}, W_{\parallel}, W_{\perp}, \overline{W}, S_0, \theta, \phi\}$  are the 8 framework's parameters ( $S_0$  is the non diffusion-

weighted signal) and  $\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  is the identity matrix. The tensors  $\mathbf{P}$ ,  $\mathbf{\Lambda}$  and  $\mathbf{Q}$  can be

computed with the Kronecker delta  $\delta_{xy}$  and the components of the axis of symmetry  $c_x$  ( $x, y \in$

1, 2, 3) as:  $\mathbf{P}_{ijkl} = c_i c_j c_k c_l$ ,  $\mathbf{Q}_{ijkl} = \frac{1}{6}(c_i c_j \delta_{kl} + c_i c_k \delta_{jl} + c_i c_l \delta_{jk} + c_j c_k \delta_{il} + c_j c_l \delta_{ik} + c_k c_l \delta_{ij})$  and  $\Lambda_{ijkl} = \frac{1}{3}(\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})^4$ . The according noise-free signal  $S_{b,\vec{g}}(\Psi)$  can then be computed and fitted to dMRI data based upon the axisymmetric tensors<sup>6</sup>:

$$S_{b,\vec{g}}(\Psi) = S_0 \exp(-B_{ij} D_{ij} + \frac{1}{6} \overline{D}^2 B_{ij} B_{kl} W_{ijkl}) \quad [\text{S1.3}]$$

where

$$B_{ij} D_{ij} = \text{Tr}(B) D_{\perp} + (D_{\parallel} - D_{\perp}) \vec{c}^T B \vec{c} \quad [\text{S1.4}]$$

and

$$B_{ij} B_{kl} W_{ijkl} = \frac{1}{2} (10W_{\perp} + 5W_{\parallel} - 15\overline{W}) (\vec{c}^T B \vec{c})^2 \quad [\text{S1.5}]$$

$$+ \frac{1}{2} (5\overline{W} - W_{\parallel} - 4W_{\perp}) (\vec{c}^T B \vec{c} \text{Tr}(B)) \quad [\text{S1.6}]$$

$$+ 2\vec{c}^T B B \vec{c} + \frac{W_{\perp}}{3} (\text{Tr}(B)^2 + 2 \text{Tr}(B \otimes B)) \quad [\text{S1.7}]$$

with

$$B = b \begin{pmatrix} g_x^2 & g_x g_y & g_x g_z \\ g_x g_y & g_y^2 & g_y g_z \\ g_x g_z & g_y g_z & g_z^2 \end{pmatrix}$$

Note that the AxTM can also be computed from the standard DKI tensor metrics assuming axial symmetry, see Section S1.1.

### S1.3 Derivation of the relationship between the axisymmetric DKI tensor metrics and the biophysical parameters

The derivation of the relationship between the axisymmetric DKI tensor metrics and the biophysical parameters is based upon the work by<sup>7;8</sup>. Starting point for the derivation are the formulas found in<sup>7</sup> that establish a connection between the axisymmetric DKI tensor metrics and the biophysical parameters, based upon the assumption of an axially symmetric fiber orientation distribution function (ODF):

$$M_1 = 3D_0 = AWF \cdot D_a + (1 - AWF)(2D_{e,\perp} + D_{e,\parallel})$$

$$M_2 = \frac{3}{2}D_2 \frac{1}{p_2} = AWF \cdot D_a + (1 - AWF)(D_{e,\parallel} - D_{e,\perp})$$

$$M_3 = D_2^2 + 5D_0^2 \left(1 + \frac{W_0}{3}\right) = AWF \cdot D_a^2 + (1 - AWF)[5D_{e,\perp}^2 + (D_{e,\parallel} - D_{e,\perp})^2 + \frac{10}{3}D_{e,\perp}(D_{e,\parallel} - D_{e,\perp})]$$

$$M_4 = \frac{1}{2}D_2(D_2 + 7D_0) \frac{1}{p_2} + \frac{7}{12} \frac{1}{p_2} W_2 D_0^2 = AWF \cdot D_a^2 + (1 - AWF)((D_{e,\parallel} - D_{e,\perp})^2 + \frac{7}{3}D_{e,\perp}(D_{e,\parallel} - D_{e,\perp}))$$

$$M_5 = \frac{9}{4}D_2^2 + \frac{35}{24}W_4 D_0^2 = p_4(AWF \cdot D_a^2 + (1 - AWF)(D_{e,\parallel} - D_{e,\perp})^2)$$

M1, M2, M3, M4 and M5 only depend on the biophysical parameter  $\kappa$  via the functions  $p_2$  and  $p_4$  and the axisymmetric DKI tensor metrics  $\overline{W}$ ,  $W_{\parallel}$ ,  $W_{\perp}$ ,  $D_{\parallel}$ ,  $D_{\perp}$ :

$$D_0 = \frac{1}{3}(2D_{\perp} + D_{\parallel}) \qquad D_2 = \frac{2}{3}(D_{\parallel} - D_{\perp})$$

$$W_0 = \overline{W} \qquad W_2 = \frac{1}{7}(3W_{\parallel} + 5\overline{W} - 8W_{\perp})$$

$$W_4 = \frac{4}{7}(W_{\parallel} - 3\overline{W} + 2W_{\perp})$$



$$p_2 = \frac{1}{4} \left( \frac{3}{\sqrt{\kappa} F(\sqrt{\kappa})} - 2 - \frac{3}{\kappa} \right)$$

$$p_4 = \frac{1}{32\kappa^2} (105 + 12\kappa(5 + \kappa) + \frac{5\sqrt{\kappa}(2\kappa - 21)}{F(\sqrt{\kappa})})$$

here,  $F$  is Dawsons function. A quadratic equation for  $AWF$  can be found (for detailed derivation see<sup>8</sup>):

$$0 = a \cdot AWF^2 - \left( a + c - \frac{40}{3} \right) AWF + c \quad [\text{S1.8}]$$

where  $a$  and  $c$  are:

$$a = (\Delta m)^2 - \left( \frac{7}{3} + 2d_2 \right) \Delta m + m_2 \quad [\text{S1.9}]$$

and:

$$c = (\Delta m - 5 - d_2)^2 \quad [\text{S1.10}]$$

that depend on  $d_2$ ,  $m_2$ ,  $\bar{D}$  and  $\Delta m$  which can be computed with the axisymmetric DKI tensor metrics and  $\kappa$ :

$$\bar{D} = \frac{1}{3} (M_1 - M_2) \quad \Delta m = \frac{M_3}{\bar{D}^2} - \frac{M_4}{\bar{D}^2}$$

$$d_2 := \frac{M_2}{\bar{D}} \quad m_2 := \frac{M_4}{\bar{D}^2}$$

Eq. (S1.8) has two solutions referred to as "branches", which, in turn, can be computed with  $a$  and

$c$ :

$$AWF_- = \frac{-\frac{40}{3} + a + c - \sqrt{-4ac + (\frac{40}{3} - a - c)^2}}{2a} \quad [\text{S1.11}]$$

$$AWF_+ = \frac{-\frac{40}{3} + a + c + \sqrt{-4ac + (\frac{40}{3} - a - c)^2}}{2a} \quad [\text{S1.12}]$$

The solution for  $AWF$  (either branch "+" or "-") can then be used to compute the diffusivities  $D_{e,\perp}$ ,  $D_{e,\parallel}$  and  $D_a$  analytically.

$$D_{e,\perp} = \frac{\bar{D}}{(1 - AWF)} \quad D_a = \left( \frac{\Delta m(1 - AWF) - 5 - d_2}{-AWF} \right) \bar{D}$$

$$D_{e,\parallel} = \left( \frac{d_2 - AWF \cdot \frac{D_a}{\bar{D}}}{(1 - AWF)} \right) \bar{D} + D_{e,\perp}$$

However, at this point  $\kappa$  is still unknown and needed to estimate  $p_2$  and  $p_4$ . All the biophysical diffusion parameters and  $AWF$  can now be expressed in terms of  $\kappa$  and the axisymmetric DKI tensor metrics. This is used to define an objective function where  $\kappa$  is the only unknown parameter, since the axisymmetric DKI tensor metrics have previously been estimated:

$$0 = [p_4(AWF \cdot D_a^2 + (1 - AWF)(D_{e,\parallel} - D_{e,\perp})^2)] - M_5 \quad [\text{S1.13}]$$

This objective function is then minimized to find  $\kappa$  with which first  $AWF$  and then the biophysical diffusivities can be found as described. It was feasible to optimize this problem over a discrete, linearly sampled range of  $[0 \leq \kappa \leq 50]$  because it depends on one non-negative parameter  $\kappa$ . This procedure was faster and more precise compared to using the available MATLAB solvers. There are at least two solutions<sup>9;8</sup> ("branches") to the optimization problem, but in the main paper only the results of the branch labeled "+" is reported. This branch choice corresponds to assuming

$4 - \sqrt{\frac{40}{3}} < \frac{D_a - D_{e,\parallel}}{D_{e,\perp}} < 4 + \sqrt{\frac{40}{3}}$  associated with  $D_a > D_{e,\parallel}$  and also labeled  $\eta = 1$  in<sup>7</sup> or  $\zeta = +$  in<sup>8</sup> and is in line with existing literature<sup>7;10;11;12</sup> on the branch selection using the intra-axonal diffusivity as a deciding factor, see Section 4.2, main document, for a further discussion.

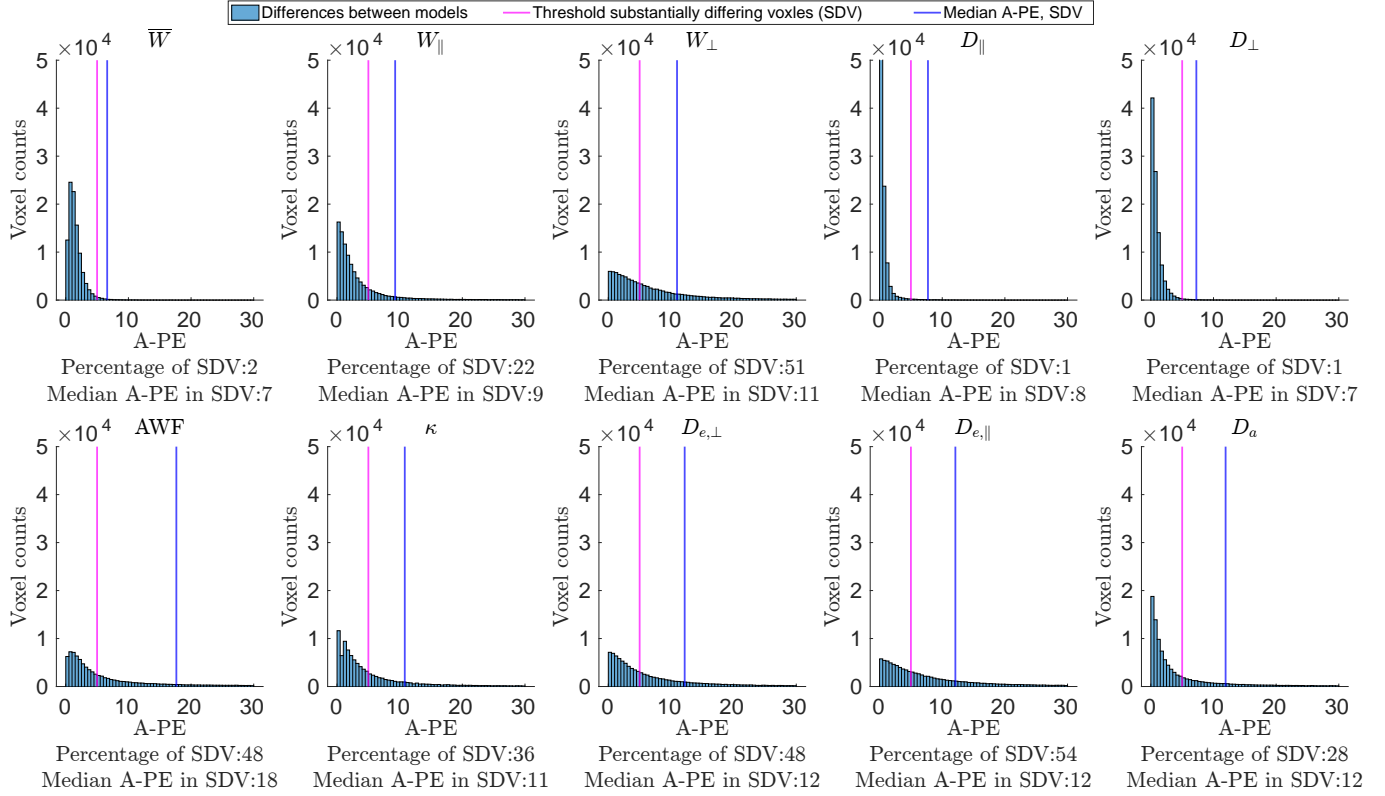


Figure S1: Histograms of underlying A-PE distribution for the AxTM and biophysical parameters. The pink vertical lines indicate the threshold after which axisymmetric DKI parameter estimation results were considered "substantially differing" ( $A-PE \geq 5\%$ ), the blue vertical lines indicate the median difference in the population of substantially differing voxels. The x-axis was confined to  $[0, 30]$ .

#### S1.4 Histograms of A-PE for the AxTM and biophysical parameters

Because the main document cites the summary measures number of substantially differing voxels (SDV) and median bias in the population of SDV for brevity, Figure S1 shows the actual distributions of A-PE of the AxTM and biophysical parameters.

#### S1.5 Comparison of fit of log of signals with NLLS fit

An earlier work<sup>13</sup> has analytically shown that axisymmetric DKI and standard DKI should produce the same results if two pre-conditions are fulfilled: a) the log of the signals is being fitted and b) the axis of symmetry ( $\vec{c}$ ) and the first eigenvector ( $\vec{v}_1$ ), two measures for the main fiber orientation in

both DKI models, are identical. To rule out the possibility that the observed differences as measured by the A-PE between standard DKI and axisymmetric DKI in this study are caused by fitting the non-linear signals (pre-condition a)), a log-of-signals fit was implemented for both standard DKI and axisymmetric DKI and used to fit the same dMRI data described in Section 2.1 (main document). Figure S2 shows the difference between standard DKI and axisymmetric DKI when fitting the log-of-signals (orange histograms) versus the differences when using the NLLS fit implementation (blue histograms) at the top and the scatter density plots between the results obtained with both methods at the bottom.

Fitting the log of the signals still went along with substantial differences between both DKI models that in some cases, e.g.,  $W_{\perp}$ , showed a close relation with the NLLS fit results. It can therefore be ruled out that the observed differences in the main study are purely caused by not fitting the log of the signals.

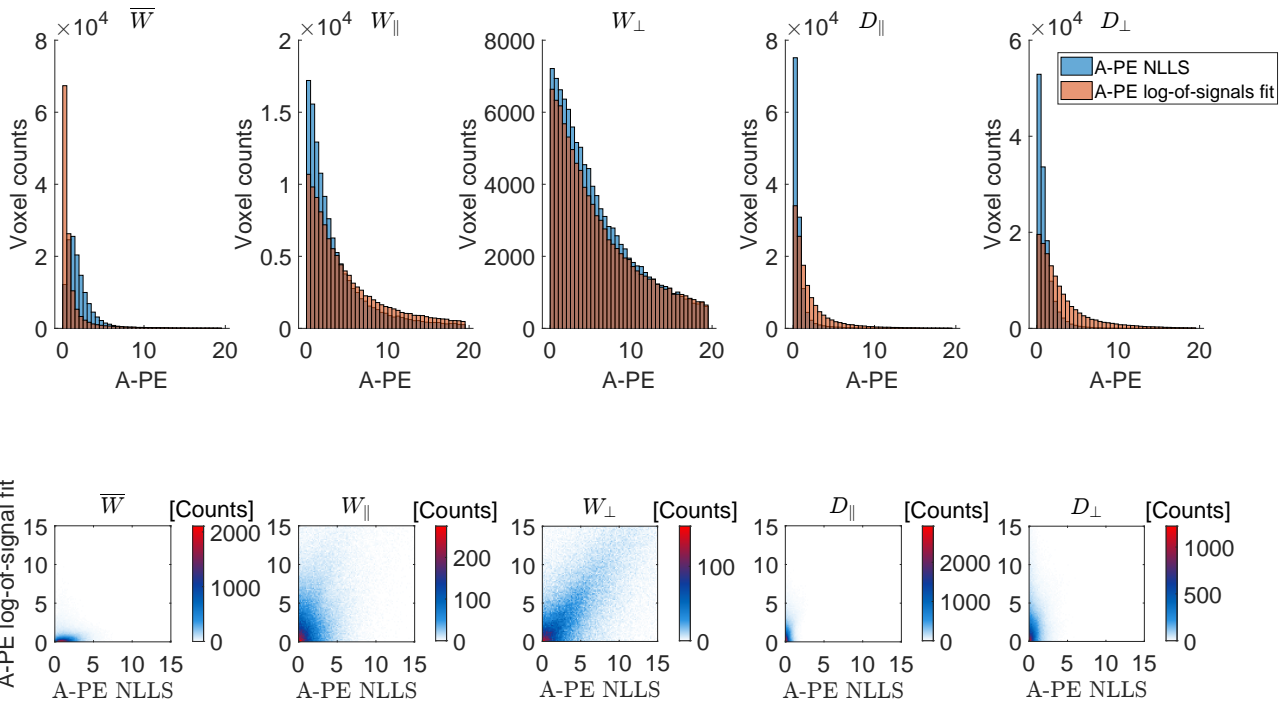


Figure S2: A-PE estimated based on a log-of-signals fit and the non-linear least squares (NLLS) fit used for this study. Top: histograms of A-PE distributions when fitting the log-of-signals (orange histograms) versus the NLLS fit implementation (blue histograms). Bottom: Scatter density plots of the A-PE estimated with a log-of-signals fit (y-axis) versus the NLLS fit (x-axis).

## S1.6 Inter-dependence of A-PE and difference in main fiber orientation

It was analytically shown in<sup>13</sup> that standard DKI and axisymmetric DKI produce the same results under two conditions (see Section S1.5). Here, the pre-condition b) from Section S1.5 is investigated. To investigate a possible inter-dependency between the A-PE and the difference between  $\vec{v}_1$  and  $\vec{c}$ , the angle  $\phi$  between  $\vec{v}_1$  and  $\vec{c}$  was calculated according to:  $\phi = \cos^{-1}(\text{abs}(\frac{\vec{v}_1 \cdot \vec{c}}{|\vec{v}_1| |\vec{c}|}))$  and plotted against the A-PE as a scatter density plot for each parameter, see Figure S3.

In almost all AxTM voxels the angle  $\phi$  was greater  $0^\circ$  degree, see x-axes of Figure S3, indicating that the main fiber orientations estimated by both DKI models were almost never identical. This violates one of the necessary presumptions named in<sup>13</sup>. The scatter density plots in Figure S3 indicate that the kurtosis metrics  $W_\perp$  and  $W_\parallel$  had the highest inter-dependency between A-PE and  $\phi$  allowing to establish at least a partial causality in this case.

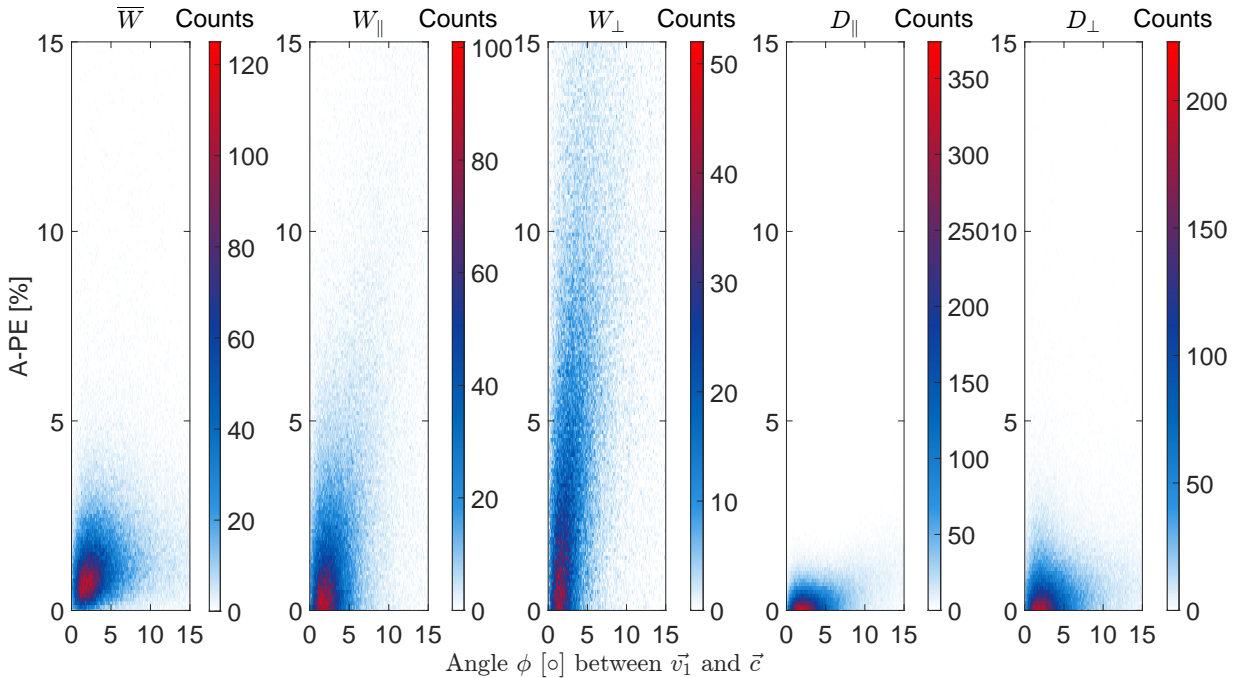


Figure S3: Scatter density plots between A-PE and angle  $\phi$  computed for voxels in the white matter mask.

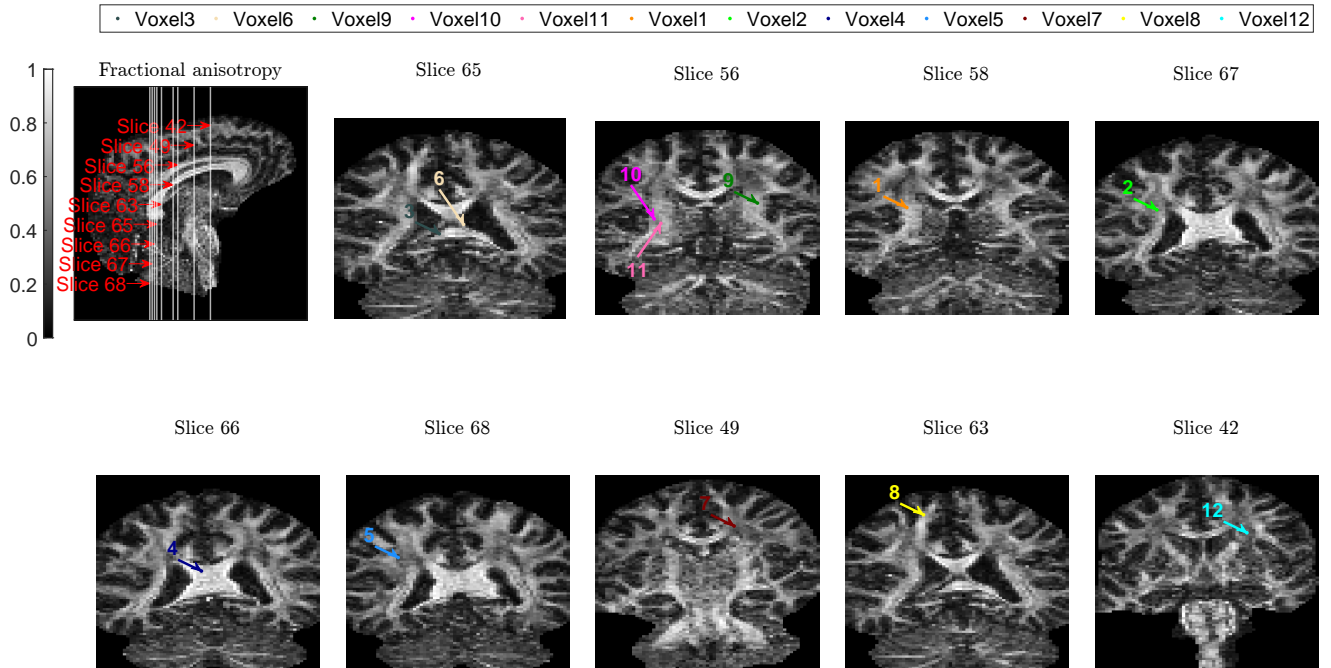


Figure S4: Voxel positions of the simulated voxels in the "single voxel analysis" corresponding to the parameters listed in Table S1 and Table S2. Top left image shows a sagittal overview of a fractional anisotropy image identifying the slices from which the voxels were extracted. The other images show the concrete voxel position in the according coronal slices.

### S1.7 Ground truth datasets, single voxel analysis

Here, the ground truth datasets used to simulate the data of the single voxel analysis (see Section 2.4, main manuscript) are documented. Figure S4 shows the 12 selected voxels in the FA image of the human brain data used in this study. Table S1 and Table S2 document the corresponding ground truth diffusion and kurtosis tensor components and ground truth AxTM values.



Table S1: Ground truth datasets for the voxels simulated in the single voxel analysis (see Section 2.4, main document), shown are the diffusion and kurtosis tensor components and  $S_0$ , the diffusivities are in [ $\frac{\mu\text{m}^2}{\text{ms}}$ ].

Parameter	Voxel 1	Voxel 2	Voxel 3	Voxel 4	Voxel 5	Voxel 6
$D_{11}$	0.52518	0.73186	1.47547	1.59281	0.94294	1.35323
$D_{22}$	0.78697	0.74288	1.04926	0.28458	0.90303	0.00005
$D_{33}$	1.23249	1.18773	0.81426	0.42473	1.31293	0.66199
$D_{12}$	-0.11517	-0.03012	0.71930	0.19176	0.01800	0.11456
$D_{13}$	0.07573	0.17281	0.46296	0.54313	0.06349	-0.69321
$D_{23}$	-0.63971	-0.24029	0.17883	0.05975	0.17136	-0.10751
$S_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	0.67775	1.07929	1.22721	3.92558	1.16881	4.03335
$W_{2222}$	0.65740	0.93806	0.61860	0.51298	0.69336	-1.23230
$W_{3333}$	1.71966	1.25306	0.60559	-0.00206	1.34117	1.08392
$W_{1112}$	0.00302	0.06835	0.56475	0.30247	-0.02502	0.11231
$W_{1113}$	-0.02877	0.16450	0.25254	0.95462	0.21085	-2.10340
$W_{2221}$	-0.16901	-0.09296	0.48807	0.07772	0.04203	0.13599
$W_{3331}$	0.05064	0.04273	0.30247	0.38988	0.10736	-0.96426
$W_{2223}$	-0.21847	0.01396	0.02596	-0.00270	0.11575	0.30162
$W_{3332}$	-0.79489	-0.40981	0.13169	0.18746	0.19683	-0.26582
$W_{1122}$	0.27296	0.37678	0.69557	0.59812	0.18333	0.09223
$W_{1133}$	0.37894	0.17585	0.36908	0.50056	0.40884	1.18587
$W_{2233}$	0.63105	0.52844	0.42591	0.02421	0.48544	-0.12350
$W_{1123}$	-0.15648	-0.12087	0.11193	-0.00798	0.00510	-0.40817
$W_{2213}$	0.06741	0.20863	0.21401	0.14025	0.00056	0.06386
$W_{3312}$	-0.07199	-0.14695	0.18450	0.06740	0.14811	0.21569
Parameter	Voxel 7	Voxel 8	Voxel 9	Voxel 10	Voxel 11	Voxel 12
$D_{11}$	0.70021	0.37055	0.45038	0.71309	0.68214	0.62777
$D_{22}$	1.23004	1.09191	0.91982	0.67809	0.73508	1.18568
$D_{33}$	0.74326	1.01422	1.48194	1.29819	1.65466	0.64001
$D_{12}$	-0.03462	0.07976	0.02886	-0.08694	-0.14949	-0.00692
$D_{13}$	-0.05384	-0.09113	0.14072	-0.33201	-0.08331	0.16751
$D_{23}$	-0.14196	-0.35298	-0.62415	-0.24086	-0.40065	0.06892
$S_0$	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
$W_{1111}$	0.25736	0.29625	0.59465	0.61617	0.90990	0.72413
$W_{2222}$	1.29478	1.19521	0.79010	0.94335	0.57185	1.94369
$W_{3333}$	0.84182	1.13021	2.11612	1.71149	2.21282	0.89254
$W_{1112}$	-0.20589	-0.15870	-0.07156	-0.20480	-0.10260	-0.45005
$W_{1113}$	-0.09519	0.04410	0.11754	0.06422	-0.11515	0.23148
$W_{2221}$	0.08161	0.12204	0.14148	-0.00514	-0.15031	0.09221
$W_{3331}$	0.08907	-0.08658	0.16565	-0.48564	-0.10745	0.04541
$W_{2223}$	0.04681	-0.45644	-0.15422	-0.18872	-0.09809	0.17618
$W_{3332}$	-0.00168	-0.30874	-0.67320	-0.12712	-0.45940	0.09544
$W_{1122}$	0.48547	0.30077	0.17656	0.23455	0.26919	0.26294
$W_{1133}$	0.16228	0.17903	0.30837	0.44797	0.22822	0.04977
$W_{2233}$	0.19544	0.81838	0.57202	0.34830	0.49117	0.54980
$W_{1123}$	-0.07992	0.01297	-0.12377	-0.09888	-0.05498	-0.10805
$W_{2213}$	-0.11182	-0.16823	-0.01215	-0.06127	0.07760	0.10235
$W_{3312}$	-0.07009	0.00660	0.01548	-0.17444	-0.12436	0.00486

Table S2: Ground truth AxTM of the single voxel analysis, corresponding to the tensor components listed in Table S1, the diffusivities are in [ $\frac{\mu\text{m}^2}{\text{ms}}$ ].

Voxel	$D_{\perp}$	$D_{\parallel}$	$W_{\perp}$	$W_{\parallel}$	$\overline{W}$
Voxel 1	0.42174	1.70116	0.58537	2.86371	1.12414
Voxel 2	0.66092	1.34064	0.90280	1.96013	1.08651
Voxel 3	0.57898	2.18102	0.45733	2.90384	1.08651
Voxel 4	0.23493	1.83226	0.09464	4.54115	1.33646
Voxel 5	0.88698	1.38493	0.82815	1.70754	1.07171
Voxel 6	0.10993	1.79541	-0.43832	6.96149	1.23884
Voxel 7	0.70221	1.26909	0.49554	1.14350	0.81607
Voxel 8	0.52738	1.42193	0.71850	2.65236	1.04360
Voxel 9	0.47971	1.89272	0.56223	2.78642	1.12295
Voxel 10	0.59895	1.49148	0.72198	2.03711	1.06653
Voxel 11	0.63328	1.80533	0.73628	2.56648	1.13435
Voxel 12	0.62942	1.19462	0.67564	2.03629	1.05708

## References

- [1] J. H. Jensen and J. A. Helpert, “Quantifying non-Gaussian water diffusion by means of pulsed-field-gradient MRI,” *Proceedings of ISMRM 2003, program number 2154*, 2003.
- [2] J. H. Jensen, J. A. Helpert, A. Ramani, H. Lu, and K. Kaczynski, “Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of Magnetic Resonance Imaging,” *Magnetic Resonance in Medicine*, vol. 53, no. 6, pp. 1432–1440, 2005.
- [3] A. Tabesh, J. H. Jensen, B. A. Ardekani, and J. A. Helpert, “Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging,” *Magnetic Resonance in Medicine*, vol. 65, no. 3, pp. 823–836, 2011.
- [4] B. Hansen, N. Shemesh, and S. N. Jespersen, “Fast imaging of mean, axial and radial diffusion kurtosis,” *NeuroImage*, vol. 142, pp. 381–393, Nov. 2016.
- [5] B. Hansen, T. E. Lund, R. Sangill, and S. N. Jespersen, “Experimentally and computationally fast method for estimation of a mean kurtosis,” *Magnetic Resonance in Medicine*, vol. 69, no. 6, pp. 1754–1760, 2013.
- [6] B. Hansen, A. R. Khan, N. Shemesh, T. E. Lund, R. Sangill, S. F. Eskildsen, L. Østergaard, and S. N. Jespersen, “White matter biomarkers from fast protocols using axially symmetric diffusion kurtosis imaging,” *NMR in Biomedicine*, vol. 30, no. 9, p. e3741, 2017.
- [7] S. N. Jespersen, J. L. Olesen, B. Hansen, and N. Shemesh, “Diffusion time dependence of microstructural parameters in fixed spinal cord,” *NeuroImage*, vol. 182, pp. 329–342, 2018.
- [8] D. S. Novikov, J. Veraart, I. O. Jelescu, and E. Fieremans, “Rotationally-invariant mapping of

- scalar and orientational metrics of neuronal microstructure with diffusion MRI,” *NeuroImage*, vol. 174, pp. 518 – 538, 2018.
- [9] I. O. Jelescu, J. Veraart, E. Fieremans, and D. S. Novikov, “Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue,” *NMR in biomedicine*, vol. 29, pp. 33–47, Jan. 2016.
- [10] N. Kunz, A. R. da Silva, and I. O. Jelescu, “Intra- and extra-axonal axial diffusivities in the white matter: Which one is faster?,” *NeuroImage*, vol. 181, pp. 314–322, Nov. 2018.
- [11] A. F. Howard, M. Cottaar, M. Drakesmith, Q. Fan, S. Y. Huang, D. K. Jones, F. J. Lange, J. Mollink, S. U. Rudrapatna, Q. Tian, K. L. Miller, and S. Jbabdi, “Estimating axial diffusivity in the NODDI model,” *NeuroImage*, vol. 262, p. 119535, Nov. 2022.
- [12] B. Dhital, M. Reisert, E. Kellner, and V. G. Kiselev, “Intra-axonal diffusivity in brain white matter,” *NeuroImage*, vol. 189, pp. 543–550, Apr. 2019.
- [13] S. Nørhøj Jespersen, “White matter biomarkers from diffusion MRI,” *Journal of Magnetic Resonance*, vol. 291, pp. 127–140, June 2018.

### **6.3 Paper 3: Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy**

The work "Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy" was published in "NMR in Biomedicine" (Wiley) in 2023.

## RESEARCH ARTICLE

# Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy

Sebastian Papazoglou<sup>1,2</sup> | Mohammad Ashtarayeh<sup>1</sup> | Jan Malte Oeschger<sup>1</sup>  |  
Martina F. Callaghan<sup>3</sup> | Mark D. Does<sup>4,5,6,7</sup>  | Siwoosh Mohammadi<sup>1,2,8</sup>

<sup>1</sup>Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>2</sup>Max Planck Research Group MR Physics, Max Planck Institute for Human Development, Berlin, Germany

<sup>3</sup>Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, London, UK

<sup>4</sup>Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee, USA

<sup>5</sup>Institute of Imaging Science, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>6</sup>Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center, Nashville, Tennessee, USA

<sup>7</sup>Department of Electrical Engineering, Vanderbilt University, Nashville, Tennessee, USA

<sup>8</sup>Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

## Correspondence

Sebastian Papazoglou, Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.  
Email: [papazoglou@mpib-berlin.mpg.de](mailto:papazoglou@mpib-berlin.mpg.de)

## Abstract

Biophysical diffusion-weighted imaging (DWI) models are increasingly used in neuroscience to estimate the axonal water fraction ( $f_{AW}$ ), which in turn is key for noninvasive estimation of the axonal volume fraction ( $f_A$ ). These models require thorough validation by comparison with a reference method, for example, electron microscopy (EM). While EM studies often neglect the unmyelinated axons and solely report the fraction of myelinated axons, in DWI both myelinated and unmyelinated axons contribute to the DWI signal. However, DWI models often include simplifications, for example, the neglect of differences in the compartmental relaxation times or fixed diffusivities, which in turn might affect the estimation of  $f_{AW}$ . We investigate whether linear calibration parameters (scaling and offset) can improve the comparability between EM- and DWI-based metrics of  $f_A$ . To this end, we (a) used six DWI models based on the so-called standard model of white matter (WM), including two models with fixed compartmental diffusivities (e.g., neurite orientation dispersion and density imaging, NODDI) and four models that fitted the compartmental diffusivities (e.g., white matter tract integrity, WMTI), and (b) used a multimodal data set including ex vivo diffusion DWI and EM data in mice with a broad dynamic range of fibre volume metrics. We demonstrated that the offset is associated with the volume fraction of unmyelinated axons and the scaling factor is associated with different compartmental  $T_2$  and can substantially enhance the comparability between EM- and DWI-based metrics of  $f_A$ . We found that DWI models that fitted compartmental diffusivities provided the most accurate estimates of the EM-based  $f_A$ . Finally, we introduced a more efficient hybrid calibration approach, where only the offset is estimated but

**Abbreviations:** ACID, artefact correction in diffusion MRI; ANOVA, analysis of variance; BAYDIFF, Bayesian estimation of microstructural diffusion parameters based on single and multiple MRI diffusion encodings; BIC, Bayesian information criterion; CKO, conditional knockout; DWI, diffusion-weighted imaging; EM, electron microscopy; Gd-DTPA, gadolinium diethylenetriamine penta-acetic acid; NODDI, neurite orientation dispersion and density imaging; PBS, phosphate-buffered saline; Pten, phosphatase and tensin homolog; Rictor, rapamycin-insensitive companion of mTOR; ROI, region of interest; RSS, residual sum of squares; SD, standard deviation; SMI, standard model imaging; Tsc2, tuberous sclerosis complex subunit 2; WM, white matter; WMTI, white matter tract integrity; WMTI-W, WMTI-Watson.

First authorship is shared by Sebastian Papazoglou and Mohammad Ashtarayeh.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *NMR in Biomedicine* published by John Wiley & Sons Ltd.

**Funding information**

German Research Foundation (DFG Priority Program 2041 “Computational Connectomics”, Grant/Award Numbers: MO2397/5-1, MO2397/52; DFG Emmy Noether Stipend, Grant/Award Numbers: MO2397/4-1, MO2397/4-2; BMBF, Grant/Award Numbers: 01EW1711A, 01EW1711B; Forschungszentrum Medizintechnik Hamburg, Grant/Award Number: 01fmthh2017; National Institute of Health, Grant/Award Number: EB019980; MRC and Spinal Research Charity, Grant/Award Number: MR/R000050/1; Wellcome, Grant/Award Number: 203147/Z/16/Z

the scaling is fixed to a theoretically predicted value. Using this approach, a similar one-to-one correspondence to EM was achieved for WMTI. The method presented can pave the way for use of validated DWI-based models in clinical research and neuroscience.

**KEYWORDS**

axonal volume fraction, axonal water fraction, biophysical model, calibration, diffusion-weighted imaging, g ratio, histology reference, unmyelinated axons

**1 | INTRODUCTION**

Diffusion-weighted imaging (DWI) is frequently used by neuroscientists as a noninvasive tool to infer microstructural tissue features. A range of biophysical multicompartiment DWI models have been proposed to connect the diffusion-weighted signal to the axonal water fraction,<sup>1–5</sup> inspired by early biophysical models such as those of Assef et al. and Jespersen et al.<sup>6,7</sup> Most of these models are variants of the so-called standard model of white matter.<sup>8</sup> Such models are increasingly used in clinical research and neuroscience,<sup>9–14</sup> where the most widely employed DWI models are neurite orientation dispersion and density imaging (NODDI)<sup>15</sup> and white matter tract integrity (WMTI).<sup>16</sup> They have been employed, for example, for the estimation of the MR g ratio—a measure that is indicative of neuronal conduction velocity and thus of the functional integrity of white matter (WM) fibres.<sup>4,17,18</sup>

However, these biophysical DWI models include certain simplifying assumptions about the underlying tissue microstructure in order to meet the demand for reasonable measurement times and numerical stability of parameter estimation. One important limitation is that these models neglect different compartmental  $T_2$  values in the intra- and extracellular signal.<sup>19–22</sup> The signals of the aforementioned multicompartiment DWI models are modelled as the sum of signal contributions from the individual compartments (e.g. axonal, extracellular, isotropic compartments). The fraction of the signal of the axonal compartment is then usually directly related to the metric for the axonal water fraction by multiplication by a factor that accounts for the low sensitivity to the myelin water signal.<sup>17</sup> Despite this correction factor, signal fraction and axonal water fraction are not truly interchangeable, because of the different transverse relaxation times ( $T_2$ ) in the compartments. Instead, the signal fractions are weighted fractions, the weights of which depend on the compartmental  $R_2$  differences ( $R_2 = 1/T_2$ ) and the echo time  $T_E$  employed.<sup>19,21</sup>

Although challenges and limitations of these biophysical DWI models are well-known,<sup>3</sup> their accuracy has been investigated only to a limited extent. An analysis of the accuracy of DWI-based axonal metrics requires an accurate reference. A frequently used method for measuring the axonal volume fraction is electron microscopy (EM), because its resolution allows one to distinguish between the myelin sheath and axonal body of single axons. While the fraction of unmyelinated axons can, in principle, also be assessed with EM,<sup>23–25</sup> analyses typically focus solely on the fraction of myelinated axons, especially when performing EM on human brain tissue.<sup>26–28</sup> This is because, compared with myelinated axons, unmyelinated axons are more difficult to distinguish from other entities such as, for example, glial cells, and hence more prone to misclassification than myelinated axons.<sup>29–31</sup> On the other hand, DWI-based estimates of the axonal volume fraction are not only sensitive to the volume fraction of myelinated axons but, presumably to a lesser degree, also affected by the unmyelinated axons.<sup>32</sup> Therefore, testing the accuracy of DWI-based estimates of axon volume fractions of myelinated axons by comparison with the EM reference would require a calibration step that corrects for potential differences in the sensitivity to unmyelinated axons and for potential limitations of DWI models.

In this study we demonstrate that linear calibration including an offset and a scaling can improve the comparability of DWI- and EM-based axonal volume metrics and allows us to assess the accuracy of DWI-based models of the volume fraction of myelinated axons. We hypothesise that the offset accounts for the differential sensitivity of our EM and DWI to the fraction of unmyelinated axons. Moreover, the linear calibration includes a scaling factor to account for compartmental  $T_2$  differences. We compare different models with varying degrees of complexity. We investigate six DWI models based on the so-called standard model of WM,<sup>8</sup> including two models with fixed compartmental diffusivities, NODDI<sup>15</sup> and NODDI-DTI,<sup>33</sup> and four models that fitted the compartmental diffusivities. The latter included two relatively novel implementations, standard model imaging (SMI)<sup>34</sup> and a Bayesian variant fitting the standard-model parameters (BAYDIFF),<sup>35</sup> as well as an older model (WMTI)<sup>16</sup> and a variant of it, WMTI-Watson (WMTI-W).<sup>36</sup> For the comparison of DWI- and EM-based models of the axonal volume fraction, we use a multimodal, ex vivo dataset of DWI and EM data of mouse WM from the corpus callosum and fornix.<sup>37,38</sup> Before comparing the DWI models with EM, we first perform a group selection based on mouse models using the EM-based axon volume fraction as the selection criterion. Then, we determine the best combination of calibration parameters for each DWI model required to establish comparability with the EM reference. This combination of calibration parameters is also compared with a hybrid calibration approach, where the scaling calibration factor is

determined analytically using a newly derived analytical approximation that relates the scaling parameter to the compartmental  $T_2$  differences given  $T_E$  while the offset parameter is estimated from the data. Finally, we assess the accuracy of the DWI-based models of the volume fraction of myelinated axons achieved through the proposed, purely data-driven and hybrid calibration approaches by comparison with their EM-based counterpart.

## 2 | BACKGROUND

### 2.1 | DWI- and EM-based metrics of axonal volume

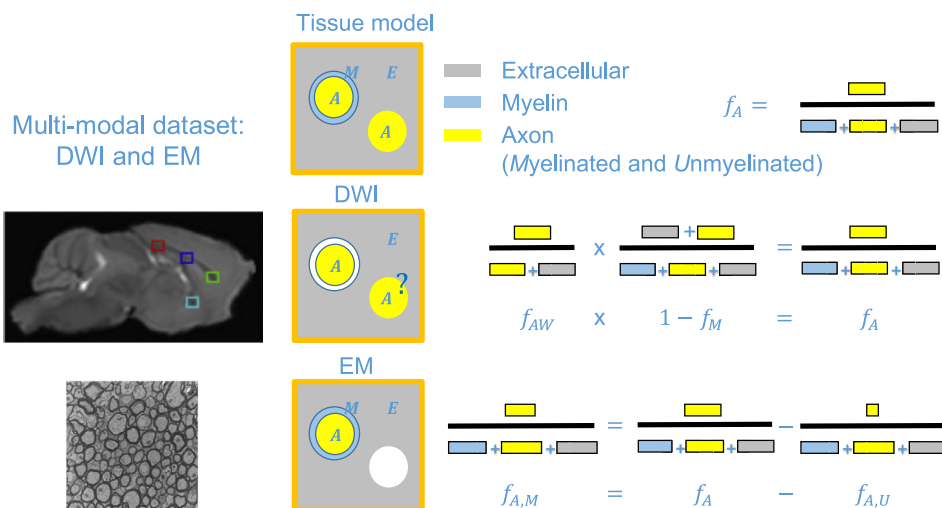
White matter tissue is typically modelled as being composed of three distinct, nonoverlapping compartments quantified by the axonal ( $f_A$ ), myelin ( $f_M$ ), and extracellular volume fraction ( $f_E$ ), with

$$f_A + f_M + f_E = 1 \tag{1}$$

in every WM voxel. A schematic description of the modelled volume fractions is shown in Figure 1. In DWI, in practice the myelin compartment is not affecting the signal, due to the short relaxation time of the myelin water. Therefore, the DWI signal is determined by the axonal water fraction  $f_{AW}$ . In the two-compartment standard model of WM<sup>8</sup> shown in the second row of Figure 1,  $f_{AW}$  is given by

$$f_{AW} = \frac{f_A}{f_A + f_E} \tag{2}$$

Some DWI models also include an optional isotropic or CSF (cerebrospinal fluid) compartment ( $f_{iso}$ ) and for the ex vivo case an additional dot compartment ( $f_{dot}$ ) accounting for fixation effects, which would then have to be included in the sum on the left of Equation (1). In order to distinguish DWI-based metrics from the “true” axonal water fraction  $f_{AW}$  given by Equation (2), we denote them in the following by  $f_{AW}^{(DWI)}$  ( $\approx f_{AW}$ ). To convert  $f_{AW}$  (Equation 2) into the axonal volume fraction  $f_A$ , it has to be rescaled by  $1 - f_M$  (Figure 1, second row). We used electron microscopy metrics as the reference for both the axonal volume fraction of myelinated axons, denoted by  $f_A^{(EM)}$ , and the myelin volume fraction, denoted by  $f_M^{(EM)}$ . By rescaling  $f_{AW}^{(DWI)}$  with the EM-based prefactor  $1 - f_M^{(EM)}$ , we obtain a DWI-based estimate of the axonal volume fraction  $f_A^{(DWI)}$ :



**FIGURE 1** Schematic relating the tissue volume fractions of the three-compartment tissue model to their counterpart from the multimodal dataset including DWI (top left) and EM (bottom left) data. Note that areas presented in white are not observable with the corresponding technique, that is, myelin in the case of DWI and unmyelinated axons in the case of EM. The question mark indicates that it is not known to what extent the fraction of unmyelinated axons can be estimated by the DWI models. DWI and EM images were taken from Kelm et al.<sup>37</sup> and modified. Coloured boxes in the DWI image indicate regions of interest (ROIs) in which DWI metrics were available in this study (green: genu, blue: midbody, red: splenium, cyan: fornix).



$$f_A^{(DWI)} = (1 - f_M^{(EM)}) f_{AW}^{(DWI)}. \quad (3)$$

In practice, one would choose an MRI-based metric for the myelin volume fraction to rescale the axonal water fraction. This typically requires additional calibration parameters to generate a volume fraction from the MRI-based myelin marker. However, in this study we are interested in the effect of calibration parameters on the axonal volume fraction. Thus, to avoid ambiguity, we used the EM reference instead of an MRI-based myelin marker as the metric for the myelin volume fraction. This ensures that any calibration is correcting for differences between EM and DWI due to the DWI-based axonal water fraction only. Furthermore, we distinguish between myelinated ( $f_{A,M}$ ) and unmyelinated ( $f_{A,U}$ ) axon volume fractions (with  $f_A = f_{A,M} + f_{A,U}$ ) to account for the fact that in this study only the first was assessed by EM, that is,

$$f_A^{(EM)} = f_{A,M}. \quad (4)$$

A glossary of the metrics is given in Table 1.

## 2.2 | Calibration parameters

In order to establish comparability between the EM reference ( $f_A^{(EM)}$ ) and the DWI-based axonal volume fraction ( $f_A^{(DWI)}$ ), it is necessary to account for the differential sensitivity of the DWI models to the fraction of unmyelinated axons, which we model here by an additional offset calibration parameter  $U$  to  $f_A^{(DWI)}$  (Equation 3):

$$f_A^{(DWI)} \mapsto f_A^{(DWI)} - U \quad (\text{offset calibration}). \quad (5)$$

Furthermore, in Equation (3) it is assumed that, apart from measurement error,  $f_{AW}^{(DWI)}$  is equal to the true  $f_{AW}$ , which is a common assumption implying that compartmental  $T_2$  differences in transverse relaxation are negligible. If compartmental differences cannot be neglected, the fraction of the axonal signal would become a function of the employed echo time  $T_E$  with  $f_{AW}^{(DWI)}(T_E = 0) = f_{AW}$  (see, e.g., Gong et al.<sup>21</sup>). In the subsequent section it will be shown analytically that this  $T_E$  dependence can be separated into a scaling calibration factor:

$$f_{AW}^{(DWI)} \mapsto s(T_E) f_{AW}^{(DWI)} \quad (\text{scaling calibration}). \quad (6)$$

## 2.3 | Analytical derivation of the scaling calibration

All DWI models tested in this study (SMI, BAYDIFF, WMTI, WMTI-W, NODDI, and NODDI-DTI) are variants of the standard model.<sup>8</sup> They can all be derived on the basis of the four-compartment, ex vivo NODDI signal model, composed of axonal (with index  $a$ ), extracellular, isotropic (iso), and so-called dot (dot) compartments, the last of which accounts for water trapped inside small cavities in fixed tissue with effectively no diffusivity.<sup>39,40</sup> It is given by

$$\frac{S_{DWI}}{S_0} = (1 - \nu_{dot}) \{ (1 - \nu_{iso}) [\nu S_A + (1 - \nu) S_E] + \nu_{iso} S_{iso} \} + \nu_{dot} S_{dot}, \quad (7)$$

where  $\nu_{dot}$ ,  $\nu_{iso}$ , and  $\nu$  are the signal fractions of the dot, isotropic, and axonal compartments, respectively. The usual in vivo NODDI model is obtained by setting  $\nu_{dot} = 0$  and the signal models for SMI, BAYDIFF, WMTI-W, WMTI, and NODDI-DTI are retrieved setting  $\nu_{dot} = \nu_{iso} = 0$ . The

**TABLE 1** Summary of the employed DWI and EM metrics and their relation with the tissue compartment model volume fractions.

Description	Tissue model metric	DWI metric	EM metric
Axonal volume fraction	$f_A$	$f_A^{(DWI)}$	-
Axonal volume fraction (unmyelinated)	$f_{A,U}$	-	-
Axonal volume fraction (myelinated)	$f_{A,M}$	-	$f_A^{(EM)}$
Axonal water fraction	$f_{AW}$	$f_{AW}^{(DWI)}$	-
Myelin volume fraction	$f_M$	-	$f_M^{(EM)}$
Extracellular volume fraction	$f_E$	-	-

compartmental signals ( $S_{\text{dot}}$ ,  $S_{\text{iso}}$ , and  $S_a$ ) are functions of the diffusion vector  $\mathbf{b}$  and a set of biophysical parameters  $\{p_i\}$ , which depend on the DWI model (see Table 2), with the assumption that, at  $\mathbf{b} = 0$ ,  $S_A = S_E = S_{\text{iso}} = S_{\text{dot}} = 1$ . In analogy to the volume fractions defined in the tissue model shown in Figure 1, the signal fractions in Equation (7) are  $\nu = f_A / (f_A + f_E)$ ,  $\nu_{\text{iso}} = f_{\text{iso}} / (f_A + f_E + f_{\text{iso}})$ , and  $\nu_{\text{dot}} = f_{\text{dot}} / (f_A + f_E + f_{\text{iso}} + f_{\text{dot}})$ . In this signal model, the axonal water fraction is given directly by the corresponding coefficient of  $S_A$ : (ex vivo NODDI)  $f_{\text{AW}}^{(\text{DWI})} = (1 - \nu_{\text{dot}})(1 - \nu_{\text{iso}})\nu$ , (in vivo NODDI and BAYDIFF, i.e.,  $\nu_{\text{dot}} = 0$ )  $f_{\text{AW}}^{(\text{DWI})} = (1 - \nu_{\text{iso}})\nu$ , and (SMI, WMTI-W, WMTI and NODDI-DTI, i.e.,  $\nu_{\text{dot}} = \nu_{\text{iso}} = 0$ )  $f_{\text{AW}}^{(\text{DWI})} = \nu$ . This changes if the compartmental signals are functions of the corresponding compartmental transverse relaxation times. In that case  $S_0$  becomes  $S_0 = \rho B_1^+ B_1^- \widehat{S}_0(t)^{41}$  and  $S_{\text{DWI}}$  changes correspondingly, where  $B_1^+$  is the transmit profile,  $B_1^-$  the receive profile, and  $\rho$  includes the proton density, which is assumed to be the same for all compartments.  $\widehat{S}_0(t)$  denotes the time-dependent part of  $S_0$  that remains after cancelling parts common with  $S_{\text{DWI}}$  related to, for example, the head coil profile or proton density:

$$\widehat{S}_0(t) = (1 - \nu_{\text{dot}}) \left\{ (1 - \nu_{\text{iso}}) \left[ \nu e^{-T_E/T_{2,a}} + (1 - \nu) e^{-T_E/T_{2,e}} \right] + \nu_{\text{iso}} e^{-T_E/T_{2,\text{iso}}} \right\} + \nu_{\text{dot}} e^{-T_E/T_{2,\text{dot}}} \quad (8)$$

A simple signal model accounting for compartmental relaxation times is then given by

$$\frac{S_{\text{DWI}}}{S_0} = \frac{1}{\widehat{S}_0(t)} \left( (1 - \nu_{\text{dot}}) \left\{ (1 - \nu_{\text{iso}}) \left[ \nu S_a e^{-T_E/T_{2,a}} + (1 - \nu) S_e e^{-T_E/T_{2,e}} \right] + \nu_{\text{iso}} S_{\text{iso}} e^{-T_E/T_{2,\text{iso}}} \right\} + \nu_{\text{dot}} e^{-T_E/T_{2,\text{dot}}} S_{\text{dot}} \right), \quad (9)$$

where  $T_E$  is the echo time and  $T_{2,a}$ ,  $T_{2,e}$ ,  $T_{2,\text{iso}}$ , and  $T_{2,\text{dot}}$  are the transverse relaxation times in the axonal, extracellular, isotropic, and dot compartments. For this signal model  $S_0 \neq 1$  and hence the coefficient of  $S_a$  is now  $f_{\text{AW}}^{(\text{DWI})}(T_E) = (1 - \nu_{\text{dot}})(1 - \nu_{\text{iso}})\nu e^{-T_E/T_{2,a}} / \widehat{S}_0(t)$  and hence  $T_E$ -dependent. The application of a model that does not account for compartmental relaxation to diffusion MRI data will therefore require a calibration scaling factor  $s(T_E)$  in order to retrieve the desired axonal water fraction from the coefficient of the axonal signal. As can be seen directly from the ex vivo NODDI signal model (Equation 9),

$$s_{\text{pred}} \equiv \frac{\widehat{S}_0(t)}{e^{-T_E/T_{2,a}}} = (1 - \nu_{\text{dot}}) \left\{ (1 - \nu_{\text{iso}}) \left[ \nu + (1 - \nu) e^{-T_E \Delta_e} \right] + \nu_{\text{iso}} e^{-T_E \Delta_{\text{iso}}} \right\} + \nu_{\text{dot}} e^{-T_E \Delta_{\text{dot}}}, \quad (10)$$

where  $\Delta_e = 1/(1/T_{2,e} - T_{2,a})$ ,  $\Delta_{\text{iso}} = 1/(1/T_{2,\text{iso}} - T_{2,a})$ , and  $\Delta_{\text{dot}} = 1/(1/T_{2,\text{dot}} - T_{2,a})$ . Then we have  $s_{\text{pred}} \cdot f_{\text{AW}}^{(\text{DWI})} = (1 - \nu_{\text{dot}})(1 - \nu_{\text{iso}})\nu \equiv f_A^{(\text{DWI})}$  again. For the usual NODDI signal model  $\nu_{\text{dot}} = 0$  and for the remaining two-compartment models of this study  $\nu_{\text{dot}} = \nu_{\text{iso}} = 0$ . Hence the scaling

**TABLE 2** Summary by DWI model of the input data, the free biophysical parameters  $p_j$ , and the assumptions on them for the four validated DWI models. Symbols are as follows: ( $D_{\parallel}$ ) parallel diffusivity, ( $D_{\perp}$ ) perpendicular diffusivity, ( $W_{\parallel}$ ) parallel kurtosis, ( $W_{\perp}$ ) perpendicular kurtosis, ( $\langle W \rangle$ ) mean kurtosis, ( $FA$ ) fractional anisotropy, ( $MD$ ) mean diffusivity, ( $D_e^{\parallel}, D_e^{\perp}$ ) diffusivities in the extracellular compartment, ( $D_a^{\parallel}$ ) diffusivity in the axonal compartment, ( $\kappa$ ) fibre dispersion, ( $\nu$ ) axonal signal fraction, ( $\nu_{\text{iso}}$ ) and ( $\nu_{\text{dot}}$ ) signal fractions of the isotropic and dot compartments, respectively. Finally,  $p_2$  is a rotational invariant of the fibre orientation distribution function and represents an anisotropy metric.<sup>45</sup>

DWI model	Input	Biophysical parameters $p_j$	Assumptions
SMI	DWI data, noise map	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, \nu, p_2$	$\nu_{\text{dot}} = \nu_{\text{iso}} = 0$
BAYDIFF	DWI data, noise map	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, \nu, \nu_{\text{iso}}$	$\nu_{\text{dot}} = 0$
WMTI-W <sup>+</sup>	$D_{\parallel}, D_{\perp}, \langle W \rangle, W_{\parallel}, W_{\perp}$	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, \kappa, \nu$	$\nu_{\text{dot}} = \nu_{\text{iso}} = 0$ $D_a^{\parallel} > D_e^{\parallel}$
WMTI	All 21 DKI parameters	$D_e^{\parallel}, D_e^{\perp}, D_a^{\parallel}, \nu$	$\nu_{\text{dot}} = \nu_{\text{iso}} = 0, \kappa \rightarrow \infty$
NODDI	DWI data	in vivo: $\kappa, \nu, \nu_{\text{iso}}$ ex vivo: $\kappa, \nu, \nu_{\text{iso}}, \nu_{\text{dot}}$	$D_e^{\parallel} = D_a^{\parallel}$ $D_e^{\perp} = (1 - \nu) D_e^{\parallel}$ $D_{\text{iso}} = 2.0 \mu\text{m}^2/\text{ms}$ $D_e^{\parallel} = 0.35 \mu\text{m}^2/\text{ms}$
NODDI-DTI	$FA, MD$	$\kappa, \nu$	$\nu_{\text{dot}} = \nu_{\text{iso}} = 0$ $D_e^{\parallel} = D_a^{\parallel}$ $D_e^{\perp} = (1 - \nu) D_e^{\parallel}$ $D_e^{\parallel} = 0.35 \mu\text{m}^2/\text{ms}$

can be predicted for the DWI models using Equation (10) once the compartmental signal fraction, compartmental  $T_2$ , and echo times are known from the literature or estimated from multi-echo measurements such as described, for example, in Appendix B.

### 3 | METHODS AND MATERIALS

We divided our analysis into three steps. First, we statistically assessed differences between the mouse models with respect to the EM-based axonal volume fraction  $f_A^{(EM)}$ . In a second step, we determined, via data fitting, which combination of calibration parameters improves the one-to-one correspondence between the DWI-based axonal volume fraction  $f_A^{(DWI)}$  and its EM-based counterpart  $f_A^{(EM)}$  the most. Finally, we assessed the error and bias of the DWI models relative to the dynamic range in the EM data for the combinations found in the second analysis.

#### 3.1 | Dataset

The dataset used in this study is described in detail in Kelm et al. and West et al.<sup>37,38</sup> The data included DWI and EM histology data in an ex vivo cohort of  $N = 15$  mice. Six were healthy controls (i.e.,  $N_{\text{Controls}} = 6$ ) and nine were genetically modified mouse models: three Pten CKO (hypermyelinated), three Rictor CKO (hypomyelinated), and three Tsc2 CKO (severely hypomyelinated), that is, ( $N_{\text{Pten}} = N_{\text{Rictor}} = N_{\text{Tsc2}} = 3$ ). In total, for each mouse, diffusion kurtosis imaging (DKI) data (see, e.g., Kelm et al.<sup>37</sup> for further details) and EM metrics  $f_A^{(EM)}$  and  $f_M^{(EM)}$  were available in the four aforementioned regions of interest (ROIs). This resulted in  $N = 60$  numerical values ( $15 \text{ mice} \times \text{four ROIs}$ ) for each EM and DWI metric.

##### 3.1.1 | Tissue preparation for DWI and EM

Tissue treatment for DWI and EM was as follows: in situ mouse brains were perfusion-fixed using 2.5% glutaraldehyde and 2% paraformaldehyde + 1 mM Gd-DTPA (Magnevist, Bayer HealthCare, Wayne, NJ, USA) in phosphate-buffered saline (PBS). After excision, mouse brains were postfixed in the aforementioned fixative solution at 4 °C for one week. After that the brains were washed thoroughly with PBS + 1 mM Gd-DTPA at 4 °C for at least one week, with the solution being changed three times in order to wash out residual fixative that would reduce the tissue  $T_2$ .<sup>42</sup>

##### 3.1.2 | DWI

For DWI imaging, the mouse brains were placed in MR-compatible, perfluoropolyether liquid-filled tubes (Fomblin, Solvay Solexis, Thorofare, NJ, USA). Further DWI and DKI parameters were as follows: all DWI was performed on a 15.2T 11-cm horizontal bore Bruker Biospec scanner (Bruker BioSpin, Billerica, MA, USA) at bore temperature ( $17 \pm 0.5$  °C), FOV =  $19.2 \times 14.4 \times 10.8 \text{ mm}^3$ , matrix size =  $128 \times 96 \times 72$ , at an isotropic resolution of  $150 \mu\text{m}$ , that is,  $22\,500 \mu\text{m}^2$  cross-sectional voxel area. DKI was performed using a 3D diffusion-weighted fast spin-echo sequence. Further parameters were as follows: repetition time  $T_R = 200 \text{ ms}$ , echo time  $T_E = 19.0 \text{ ms}$ , gradient pulse duration  $\delta = 5 \text{ ms}$ , diffusion time  $\Delta = 12 \text{ ms}$ ,  $b$  values = 3000 and  $6000 \text{ s/mm}^2$ , 30 directions, and two signal averages with reversed gradient polarity.

##### 3.1.3 | EM

After DWI, the brains were prepared for EM. To this end, thick midsagittal tissue sections were cut from the brains in four ROIs, three in the corpus callosum (genu, midbody, splenium) and one ROI in the fornix, as indicated by the coloured boxes in Figure 1. The sections were then placed in 1% osmium tetroxide in cacodylate buffer for one hour and dehydrated in graded ethanol. Then the tissue sections were embedded in epoxy resin and  $1\text{-}\mu\text{m}$  thick sections were cut and stained with 1% toluidine blue.<sup>37</sup> Finally, from the thick sections ROIs were selected using a standard mouse brain atlas and then ultrathin sections were cut for EM. EM-based tissue metrics were assessed on images of size  $2304 \times 1888$  pixels (controls) and  $2048 \times 1632$  pixels (Pten, Rictor, and Tsc2) at a resolution of  $0.004 \times 0.009 \mu\text{m}^2$ , that is, with a total area of  $\approx 156 \mu\text{m}^2$  or  $\approx 126 \mu\text{m}^2$ , respectively (see Vanderbilt University data at <https://osf.io/yp4qg/>).<sup>43</sup> EM section size ranged between  $\approx 10 \times 10$  and  $40 \times 40 \mu\text{m}^2$ , with  $0.022 \mu\text{m}$  thickness.

### 3.2 | DWI model fitting

The DWI models were variants of the standard model of two nonexchanging compartments in which fibres are assumed to be impermeable sticks with no diffusion perpendicular to their orientation.<sup>1,8</sup> For a summary of the input parameters, biophysical (output) parameters, and assumptions of the DWI models, see Table 2. All DWI models except NODDI took as input all or a combination of the 21 standard DKI parameters, that is, the six independent elements of the diffusion tensor and 15 independent elements of the kurtosis tensor. The DKI parameters were estimated using the nonlinear least-squares DKI framework implemented in the ACID toolbox (<https://diffusioontools.com/>), for further details see Appendix F.<sup>44</sup> In each of the four ROIs,  $f_{AW}^{(DWI)}$  was determined voxelwise. Overall, the number of voxels in the manually delineated ROIs ranged between six and 12. Then the mean value of  $f_{AW}^{(DWI)}$  in the ROI was calculated, whereby voxels in which  $f_{AW}^{(DWI)} < 0$ ,  $f_{AW}^{(DWI)} > 1$ , or  $f_{AW}^{(DWI)} = \text{NaN}$  were discarded. This resulted in a reduced number of valid voxels only for a few DWI models, mouse individuals, and ROIs. All ROIs had at least four valid voxels, except for one which had only three valid voxels. Averaged over all ROIs and mouse individuals there were, per DWI model: BAYDIFF: 5% outliers ( $>1$ ), WMTI-W<sup>+</sup> 1% (NaN), and NODDI-DTI 5% ( $<0$ ). SMI, WMTI, and NODDI had no outliers.

SMI:  $f_{AW}^{(DWI)}$  was estimated using the standard model of WM as implemented and described at <https://github.com/NYU-DiffusionMRI/SMI>.<sup>45</sup> In principle, this implementation of SMI allows modelling of two or three compartments including extra- and intracellular compartments and an isotropic compartment. Here, we chose the option to model only two compartments, which corresponds to discarding the isotropic compartment. To generate the noise map, we divided the signal at  $b=0$  by the reported  $\text{SNR} = 150$ .<sup>37</sup> Furthermore, the machine-learning bounds of the diffusion parameters were adjusted to fit the ex vivo situation of the mouse models (using the notation introduced in the present study): signal fraction of the axonal compartment  $\nu \in (0.05, 0.95)$ , axonal diffusivity  $D_a^{\parallel} \in (0.05, 0.7) \mu\text{m}^2/\text{ms}$ , extracellular parallel diffusivity  $D_e^{\parallel} \in (0.05, 0.7) \mu\text{m}^2/\text{ms}$ , extracellular perpendicular diffusivity  $D_e^{\perp} \in (0.05, 0.7) \mu\text{m}^2/\text{ms}$ , and free water (isotropic) compartment signal fraction  $\nu_{\text{iso}} \in (0, 0.5)$ . No further parameters were fitted because of the single echo experiments used in this study.

BAYDIFF:  $f_{AW}^{(DWI)}$  was estimated using the code from <https://bitbucket.org/reisert/baydiff/wiki/Home>.<sup>35</sup> The prior distributions of the diffusivities used in the simulations for the initial training were adjusted to fit the ex vivo situation: all intra- and extra-axonal diffusivities were assumed to be uniformly distributed in the interval  $(0.05, 0.7) \mu\text{m}^2/\text{ms}$ . The same noise map as for SMI was used.

WMTI:  $f_{AW}^{(DWI)}$  was estimated using the WMTI model<sup>16</sup> implemented at <https://github.com/NYU-DiffusionMRI/DESIGNER>. WMTI has four free parameters, two extracellular diffusivities, one parallel and one perpendicular to the fibres, one intracellular diffusivity parallel to the WM fibres, and the axonal water fraction. Fibres are assumed to be parallel.

WMTI-W:  $f_{AW}^{(DWI)}$  was estimated using an in-house fitting algorithm implementation of the biophysical model introduced by Jespersen et al.<sup>36</sup> The model has five free parameters, two extracellular diffusivities, one parallel and one perpendicular to the fibres, one intracellular diffusivity parallel to the WM fibres, the dispersion of fibres, and the axonal water fraction. Furthermore, due to the degeneracy of its solution, WMTI-W possesses two branches WMTI-W<sup>+</sup> and WMTI-W<sup>-</sup>. The two branches include different assumptions on the compartmental diffusivities parallel to the direction of fibres:  $D_a^{\parallel} > D_e^{\parallel}$  (WMTI-W<sup>+</sup>) and  $D_a^{\parallel} < D_e^{\parallel}$  (WMTI-W<sup>-</sup>). Since the negative branch is known to yield unphysical results related to the absence of a proper minimum in its objective function<sup>36</sup> and because, for in vivo application, the plus branch has been shown to be preferable,<sup>46</sup> we discarded the negative branch from our analyses.

NODDI:  $f_{AW}^{(DWI)}$  was estimated from the ex vivo NODDI model<sup>15</sup> implemented in the NODDI MATLAB toolbox ([https://www.nitrc.org/projects/noddi\\_toolbox](https://www.nitrc.org/projects/noddi_toolbox)). NODDI is the only three-compartment (four compartments in ex vivo) model. In addition to the extracellular and axonal compartments that are shared by all DWI models of this study, it includes isotropic and cerebrospinal fluid compartments (and a dot compartment for the ex vivo case). The free parameters of in vivo NODDI are fibre dispersion, axonal water fraction, and isotropic water fraction and the ex vivo NODDI model features an additional signal fraction of the dot compartment (restricted water pool<sup>39,40</sup>). The diffusivities for the isotropic compartment ( $D_{\text{iso}}$ ) and extracellular compartment, parallel to the fibre direction ( $D_e^{\parallel}$ ), were set to 2 and  $0.35 \mu\text{m}^2/\text{ms}$ , respectively, as proposed in West et al.<sup>38</sup>

NODDI-DTI:  $f_{AW}^{(DWI)}$  was determined from the aforementioned DKI fit using the fractional anisotropy (FA) and mean diffusivity (MD) from the standard DKI model as input. The FA and MD maps used as input for NODDI-DTI were calculated from the DKI fit as recommended in Edwards et al.<sup>33</sup> to avoid a kurtosis bias in MD. NODDI-DTI features only two free parameters, fibre dispersion and axonal water fraction. The compartmental diffusivities are fixed as with NODDI.

### 3.3 | Statistical group selection

To prevent calibration parameter fitting from modelling noisy data, we assessed differences in  $f_A^{(EM)}$  across the mouse models in terms of analysis of variance (ANOVA) with the null hypothesis that the mean value of  $f_A^{(EM)}$  was the same across all models.

### 3.4 | Best calibration parameter combinations

#### 3.4.1 | Calibration parameter combinations

The combinations of linear calibration parameters that could potentially improve the one-to-one correspondence between DWI and EM were determined as follows: the case without any additional parameters corresponding to the assumption that  $f_A^{(DWI)} \equiv f_A$  as given in Equation (3) defined the baseline. We then pooled the 15 individual mice and four ROIs into two groups according to the results from an ANOVA. Group 1 included healthy controls and only moderately hyper- or hypomyelinated mice (Pten or Rictor mouse models) and group 2 only included heavily hypomyelinated mice (Tsc2 mouse model), respectively. This choice was based on our finding that only between these two groups could a significant difference in the EM-based axonal volume fraction  $f_A^{(EM)}$  be observed, and not between any of the mouse models in the first group. Then we allowed for the estimation of individual offsets in each of the two groups (1: Controls, Pten, Rictor, and 2: Tsc2). For the purpose of optimisation they were written as column vectors  $\mathbf{U}_j = U_j \mathbf{e}_j$  (with  $U_j$  being the offset of group  $j$  and  $\mathbf{e}_j$  being a  $N_j \times 1$  vector of ones) with  $j \in \{1, 2\}$ ,  $N_1 = 48$ , and  $N_2 = 12$  (for all DWI models). In this notation, Equation (3) including an offset calibration (Equation 5) and scalar scaling calibration (Equation 6) reads

$$\mathbf{f}_A^{(DWI)} = \left(1 - \mathbf{f}_M^{(EM)}\right) \cdot s \cdot \mathbf{f}_{AW}^{(DWI)} - \mathbf{U}, \quad (11)$$

where  $\mathbf{U} = [\mathbf{U}_1^T, \mathbf{U}_2^T]^T \in (N_1 + N_2, 1)$ ,  $\mathbf{f}_M^{(EM)}$  and  $\mathbf{f}_{AW}^{(DWI)}$  indicate column vectors with components sorted in agreement with  $\mathbf{U}$ , and  $s$  is the scaling calibration parameter. We considered the following combinations of calibration parameters:  $U_1 = 0$  and  $U_2 \neq 0$  or  $U_1 \neq 0$  and  $U_2 = 0$ , and also both offsets were allowed to vary individually at the same time, that is,  $U_1 \neq U_2 \neq 0$ . In total, for each DWI model, we analysed seven combinations of the calibration parameters, denoted in the following as  $\{U_1\}$ ,  $\{U_2\}$ ,  $\{s\}$ ,  $\{U_1, U_2\}$ ,  $\{U_1, s\}$ ,  $\{U_2, s\}$ ,  $\{U_1, U_2, s\}$ , and the baseline  $\{\}$ . For a summary of the calibration parameters tested, see also Table 3.

#### 3.4.2 | Data-driven calibration parameter estimation

The offsets  $U_j$  and the scaling parameter  $s$  were estimated by minimising the residual sum of squares (RSS) between the DWI-based estimate for the axonal volume (Equation 11) and the EM-based gold standard (defined in Equation 4):

$$RSS = \left(\mathbf{f}_A^{(EM)} - \mathbf{f}_A^{(DWI)}\right)^T \cdot \left(\mathbf{f}_A^{(EM)} - \mathbf{f}_A^{(DWI)}\right), \quad (12)$$

where again the bold-faced quantities represent vectors including all available numerical values assembled into column vectors. In order to ensure physically reasonable estimates of the axon volume fraction, the optimization function (Equation 12) had to be complemented by a boundary condition. The constraint concerns the upper limit of the sum of volume fractions, that is,  $\min\left(\mathbf{f}_A^{(DWI)} + \mathbf{f}_M^{(EM)}\right) - 1 < 0$ . Further constraints were lower and upper bounds for the calibration parameters:  $U_j \in [0, 1]$  and  $s \in [0, 2]$ . All parameter estimations were performed using the nonlinear equation solver `fmincon` as implemented in Matlab 2020a (Mathworks, CA, USA).

To quantify the intramodel performance improvement of each DWI model due to the calibration parameters  $U_j$  and  $s$ , we used the Bayesian information criterion (BIC)<sup>47</sup>:

$$BIC = k \ln n + n \ln \frac{RSS}{n}, \quad (13)$$

where  $k$  is the number of model parameters, which varied between zero (baseline  $\{\}$ ) and three depending on the combination of  $U_j$ ,  $n$  is the number of evaluated data points, and  $RSS$  is defined in Equation (12). The BIC measures a model's capability of explaining given data while penalising

**TABLE 3** Summary of the tested calibration parameters. Groups of mice are defined in Figure 2.

Description	Calibration parameter
Estimate of $f_{A,U}$ of the first group of mice	$U_1$
Estimate of $f_{A,U}$ of the second group of mice	$U_2$
Scaling accounting for differences in compartmental $T_2$	$s$

overfitting. A lower  $BIC$  indicates less information loss, meaning that the model with the lowest  $BIC$  explains the data best. Since we employed the uncalibrated case  $\{\}$  as baseline, we only report differences  $\Delta BIC$  with respect to this case, that is,

$$\Delta BIC = BIC - BIC_{\{\}}. \quad (14)$$

$\Delta BIC$  was always calculated using all available data. For assessing the variation in  $U_j$  and  $s$ , we performed a leave-one-out analysis by successively discarding the data of one mouse until each mouse was excluded once.

### 3.4.3 | Hybrid calibration parameter estimation

To simplify the demand on the distribution of the calibration data, we introduced a hybrid calibration approach. To this end, we used the best calibration parameter combination determined by the data-driven approach described in Section 3.4.2 and estimated the offset calibration parameter  $U_j$ , while fixing the scaling parameter. In this approach, the scaling parameter was fixed to the theoretically predicted value (Equation 10) using compartmental  $T_2$  estimates derived in Appendix B and only the offset  $U_j$  was estimated using Equation (12).

## 3.5 | Assessment of bias and error

For comparison of the accuracy achieved by the data-driven (Section 3.4.2) and hybrid calibration (Section 3.4.3) approaches, we performed a Bland–Altman (BA) analysis<sup>48</sup> of the differences

$$\delta = f_A^{(DWI)} - f_A^{(EM)} \quad (15)$$

versus the mean

$$m = \frac{1}{2} (f_A^{(DWI)} + f_A^{(EM)}). \quad (16)$$

The error was given by

$$\epsilon = 1.96 \sqrt{\langle \delta^2 \rangle - \langle \delta \rangle^2}. \quad (17)$$

We also report the mean difference  $\langle \delta \rangle$ , that is, the bias, and bias  $\bar{\delta}$  and error  $\bar{\epsilon}$  relative to the dynamic range in the EM-based axonal metric, that is,

$$\bar{\delta} = \frac{\langle \delta \rangle}{\Delta f_A^{(EM)}} \quad (18)$$

and

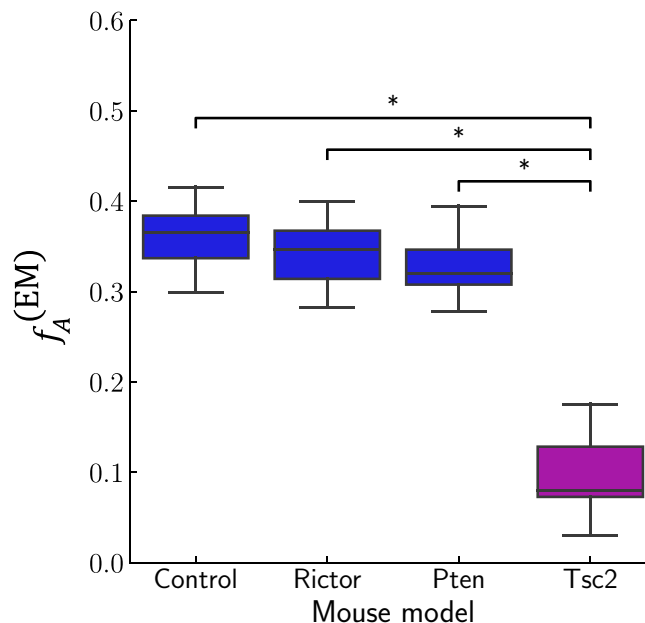
$$\bar{\epsilon} = \frac{\epsilon}{\Delta f_A^{(EM)}}, \quad (19)$$

where  $\Delta f_A^{(EM)} = f_{A,\max}^{(EM)} - f_{A,\min}^{(EM)}$  and angled brackets indicate an average.

## 4 | RESULTS

### 4.1 | Statistical group selection

The result of the ANOVA analysis (Section 3.3) shown in Figure 2 revealed a significant ( $p < 0.05$ ) difference between the EM-based axonal volume fraction ( $f_A^{(EM)}$ ) of Tsc2 mice and any of the other models, while no significant differences were observed among Pten, Rictor, and Controls. As a consequence, the four mouse models were pooled into two groups: (1) Controls, Pten, Rictor, and (2) Tsc2 for further analysis.



**FIGURE 2** EM-based axonal volume fraction  $f_A^{(EM)}$  for the four mouse models: Controls, Pten, Rictor, and Tsc2. An ANOVA revealed significant differences ( $p < 0.05$ ) only between Tsc2 and Rictor, Controls, or Pten, respectively. No further significant differences were observed. This motivated the pooling of the data into two groups: (1) Controls, Rictor, and Pten, and (2) Tsc2.

## 4.2 | Best calibration parameter combinations

### 4.2.1 | Data-driven calibration parameter estimation

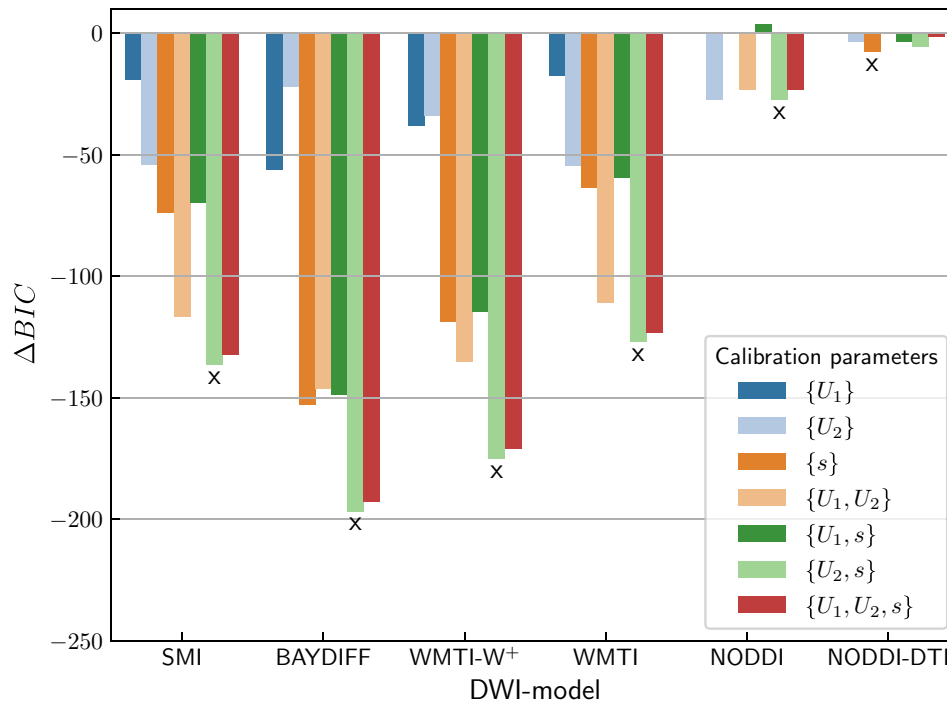
In order to determine the best combination of calibration parameters, we performed a  $BIC$  (Equation 13) analysis.  $\Delta BIC$  (Equation 14) for the tested parameter combinations are shown in Figure 3 for each DWI model separately. For all DWI models except NODDI-DTI, greatest evidence for improvement was achieved for the  $\{U_2, s\}$  set of calibration parameters, that is, when the offset for the severely hypomyelinated group 2 ( $U_2$ ) was combined with the scaling ( $s$ ). For NODDI-DTI,  $\{s\}$  had the lowest  $\Delta BIC$ . Table 4 summarises the offsets  $U_2$  and scaling  $s$  for the best combination of calibration parameters as indicated in Figure 3. The estimated offset  $U_2$  for the Tsc2 mouse model varied between 0.18 and 0.24 for SMI, BAYDIFF, WMTI, and WMTI- $W^+$ . The scaling varied between 0.52 and 1.11 (note that scaling  $s = 1$  is equivalent to no additional scaling calibration). For a summary of all tested calibration parameter combinations see Table A1 in the Appendix.

### 4.2.2 | Estimation of the theoretically predicted scaling calibration $s_{pred}$

Using the derived expression for the scaling calibration in Equation (10) together with the rescaled compartmental  $T_2$  times (Equations B2),  $T_E = 19$  ms, and the mean  $\nu = 0.475$  from Veraart et al.<sup>19</sup> and Gong et al.<sup>21</sup> and  $\nu_{iso} = 0.05$ ,<sup>21</sup> we found  $s_{pred} \approx 0.93$  for BAYDIFF,  $s_{pred} \approx 0.9$  for NODDI, and  $s_{pred} \approx 0.89$  for the other DWI models. Table 4 shows that the smallest relative difference between fitted and predicted scaling  $\Delta s$  was found for NODDI-DTI (-2%) and the largest relative difference was found for BAYDIFF (-44%).

## 4.3 | Bias and error of the best parameter combinations

In Figure 4 we compare scatter plots of the histological reference  $f_A^{(EM)}$  versus its DWI-based counterpart  $f_A^{(DWI)}$ . The first row of the figure shows the baseline (i). SMI, BAYDIFF, WMTI- $W^+$ , and WMTI clearly overestimated the axonal volume fraction of myelinated axons in both groups of mice, indicated by the global offset from the line of unity. For NODDI, only the Tsc2 mice featured an obvious positive offset, while for NODDI-DTI all mouse models showed considerably better one-to-one correspondence, although with large variance along  $f_A^{(DWI)}$ . The second row shows the best calibration parameter combinations (ii). A substantially improved one-to-one correspondence was observed only for SMI, BAYDIFF, WMTI- $W^+$ , and WMTI, while NODDI and NODDI-DTI only improved a little or not visibly at all. The third row shows the scatter plots for the case in which the scaling was fixed to its predicted value and the offset was determined by fitting to the data (iii). There, the one-to-one



**FIGURE 3** Contribution of calibration parameters to DWI model improvement. Shown are the differences  $\Delta BIC$  (Equation 14) with respect to the parameter combination with the smallest  $BIC$  in each DWI model. A lower value indicates better model performance. Values for the uncalibrated case  $\{\}$  served as baseline, that is, for this case  $\Delta BIC = 0$ . An x indicates the calibration parameter combination with the largest evidence of improvement with respect to the baseline without further calibration for each DWI model, respectively.

**TABLE 4** Summary of volume fraction of unmyelinated axons  $U_2$ , scaling  $s$ , and relative difference between fit  $s$  and theory  $s_{pred}$ , that is,  $\Delta s = (s - s_{pred})/s_{pred}$ , for the best parameter combinations from the first analysis (see also Figure 3).  $U_2$  and  $s$  were estimated in a leave-one-out fashion, in which each mouse individual was excluded from the computation once in order to get an estimate of the standard deviation (see also the final paragraph in Section 3.4.2). Note that a value of 0 corresponds to exactly zero, while 0.0 corresponds to  $<0.005$ .

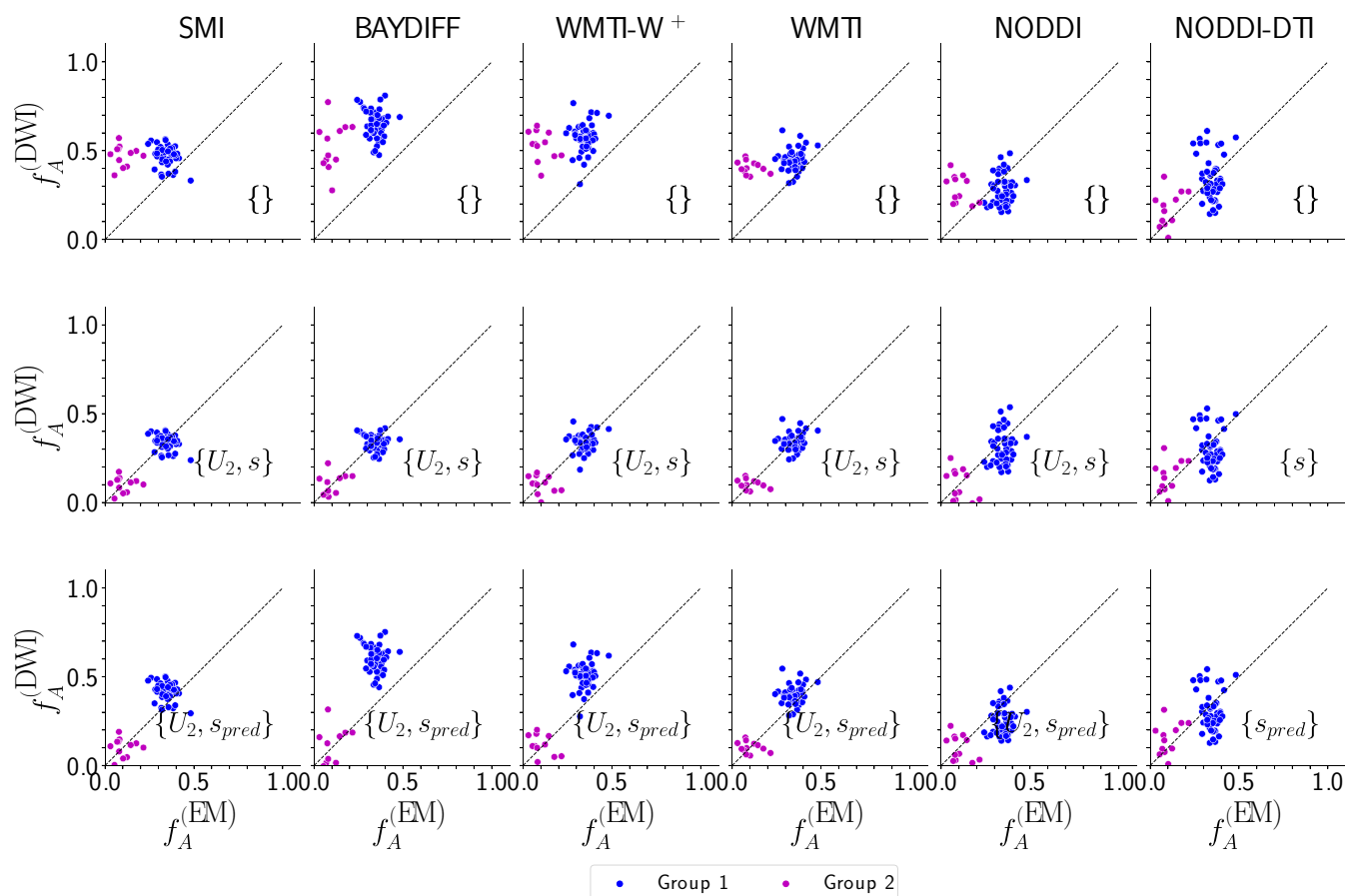
DWI model	Calibr. parameters	$U_2$ (SD)	$s$ (SD)	$s_{pred}$	Rel. $\Delta s$ [%]
SMI	$\{U_2, s\}$	0.24 (0.01)	0.72 (0.01)	0.89	-19.0
BAYDIFF	$\{U_2, s\}$	0.18 (0.0)	0.52 (0.0)	0.93	-44.0
WMTI-W <sup>+</sup>	$\{U_2, s\}$	0.21 (0.0)	0.59 (0.01)	0.89	-33.0
WMTI	$\{U_2, s\}$	0.21 (0.0)	0.76 (0.01)	0.89	-14.0
NODDI	$\{U_2, s\}$	0.21 (0.01)	1.11 (0.01)	0.9	22.0
NODDI-DTI	$\{s\}$	0	0.87 (0.02)	0.89	-2.0

correspondence is similar to the second row (ii) only for SMI and WMTI, while all other DWI models except NODDI-DTI show a less good one-to-one correspondence. Interestingly, it was mainly the control group for which the correspondence achieved between DWI and EM was less good than in case (ii). Again, NODDI-DTI displayed no visible changes compared with either (i) or (ii).

The capability to predict the EM-based reference is quantified in terms of BA plots, shown in Figure 5, and bias and error relative to the dynamic range of the EM reference, summarised in Table 5. The results for the same calibration parameter combinations (i)–(iii) as in Figure 4 are shown. The BA plots show a substantial reduction in bias and error for the best calibration parameter combinations only for SMI, BAYDIFF, WMTI-W<sup>+</sup>, and WMTI. For NODDI, only the error was reduced, and for NODDI-DTI no improvement was observed at all.

Error and bias relative to the dynamic range in the reference  $f_A^{(EM)}$  were substantially reduced for the best combination of calibration parameters (ii) only for SMI, BAYDIFF, WMTI-W<sup>+</sup>, and WMTI, whereby the relative bias was close to zero after calibration (see Table 5). BAYDIFF benefited the most in terms of relative bias, showing a reduction of 75% (from -73% to 2%). The largest reduction of relative error was observed for SMI with 26% (from 55% to 29%). NODDI and NODDI-DTI benefited much less from calibration. Their relative errors could be improved by 15% (NODDI) and 5% (NODDI-DTI). Their relative biases, however, increased slightly by 3% or 10%, respectively. Overall, the lowest relative error after calibration of all DWI models was observed for WMTI (26%). When the scaling was fixed to the theoretically predicted values and only





**FIGURE 4** Scatter plots showing gold standard  $f_A^{(EM)}$  (EM) versus DWI-based estimates of the axonal volume fraction ( $f_A^{(DWI)}$ ). The first row shows the baseline, that is, without additional calibration parameters, and subsequent rows show the best (in terms of  $\Delta BIC$ ) calibration parameter combinations (see also Figure 3 and Table 4) with both parameters determined based on data (second row) and with the scaling fixed to the predicted values (last row). Data shown pool four ROIs per mouse individual. The data were divided into two groups determined by statistical distinguishability observed in the EM gold standard  $f_A^{(EM)}$  (see Figure 2). The two groups are: (1) controls, Rictor, and Pten (blue), and (2) Tsc2 (magenta). Note that a corresponding comparison of the DWI-based axonal water fraction  $f_{AW}^{(DWI)}$  with the EM-based axonal volume fraction  $f_A^{(EM)}$  is shown in Figure C1 in the Appendix.

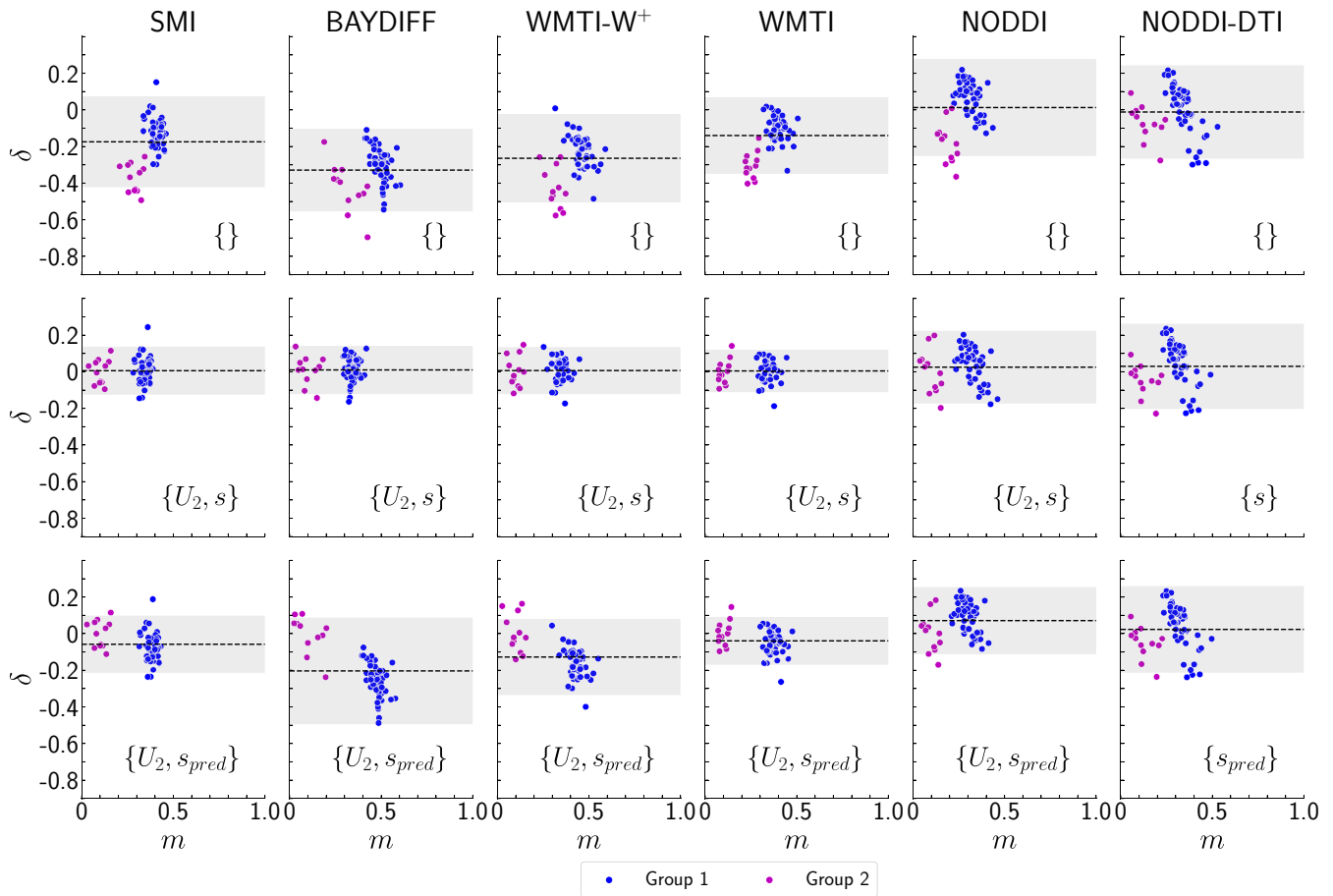
the offset  $U_2$  was determined on the basis of data (iii), similar improvement of the relative error to before, that is, for (ii), could only be achieved for SMI and WMTI. Improvement of the relative bias was less than in the case of the purely data-driven approach (ii).

## 5 | DISCUSSION

In this study, we demonstrated that the one-to-one correspondence between EM- and DWI-based metrics of the axonal volume fraction could be improved by biophysically motivated, linear calibration with an offset associated with the volume fraction of unmyelinated axons and a scaling factor correcting for global compartmental  $T_2$  relaxation differences. Using these calibration parameters, we found the best one-to-one correspondence between the EM-based axon-volume fraction and its WMTI-based counterpart, closely followed by the SMI, BAYDIFF, and WMTI- $W^+$  counterparts. Moreover, we proposed a method to predict the scaling parameter from known compartmental  $T_2$  values. We showed that using a hybrid method that uses the predicted scaling parameter and only estimates the offset parameters achieved similar one-to-one correspondence between the EM-based AVF and the DWI-based counterparts when using the WMTI and SMI models.

### 5.1 | Calibration parameters

The biophysical interpretation of the fitted calibration parameters potentially provides new insights into the DWI models investigated. Our hypothesis that the offset could be interpreted as a measure associated with the fraction of unmyelinated axons is supported by the observation



**FIGURE 5** Bland–Altman plots of differences  $\delta$  (Equation 15) versus mean  $m$  (Equation 16) between EM and DWI for the baseline (first row) and the best-performing calibration parameter combination, where all parameters were estimated on the basis of data (second row) and the scaling was fixed to the theoretically predicted values (last row). The dashed line corresponds to the bias  $\langle \delta \rangle$ , while the shaded region corresponds to  $\langle \delta \rangle \pm \epsilon$  (see Section 3.5). Individual data shown are in correspondence with Figure 4. Note that a corresponding comparison of the DWI-based axonal water fraction  $f_{AW}^{(DWI)}$  with the EM-based axonal volume fraction  $f_A^{(EM)}$  is shown in Figure C2 in the Appendix.

that for all models with a high one-to-one correspondence the offset is particularly large for the Tsc2 mouse model and negligible for the other mouse models. This trend in the offset parameter follows the change in the fraction of unmyelinated axons between the Tsc2 mice and the other mouse models.

To estimate the fraction of unmyelinated axons per mouse model, we use the approximation that the total axonal volume fraction (i.e., the sum of myelinated and unmyelinated axons) is the same for all mouse models, although their relative proportion might change across mouse models. Given this approximation, the volume fraction of unmyelinated axons can be estimated as follows: observing from Figure 2 that the volume fraction of myelinated axons is about 0.35 and taking the percentage (of the fraction of all axons) of unmyelinated axons reported in the literature (e.g., 33% in Abdollahzadeh et al.<sup>25</sup> or 30% in Jelescu et al.<sup>49</sup>), the total axonal volume would be given approximately by  $\approx 0.35 / (1.0 - 0.33) \approx 0.52$  (or 0.5 for Jelescu et al.). Assuming this total axonal volume fraction in Tsc2 mice to be the same as in control mice (i.e., 0.5), we can estimate from Figure 2 the volume fraction of unmyelinated axons to be 0.4 for Tsc2 mice.

Including an offset associated with the fraction of unmyelinated axons led to a substantial improvement in terms of  $\Delta BIC$  for all models, except for the NODDI–DTI model. On the basis of the aforementioned simple approximation, we expect that the fraction of unmyelinated axons in Tsc2 mice is about 0.4, indicating that the offset  $U_2$  in hypomyelinated mice of 0.18–0.24 (see Table 4) underestimates the volume fraction of unmyelinated axons.

Moreover, the fitted scaling parameter was smaller than its predicted counterpart for all two-compartment models. For the four-compartment NODDI model, however, the fitted scaling parameter was larger than its predicted counterpart. One reason for the mismatch between fitted and predicted scaling parameters might be that the proposed model for the predicted scaling parameter is not covering all the mechanisms that are driving the actual scaling process. Possible other factors could be, for example, nonmyelin macromolecules, which could also lead to an additional scaling effect ( $f_A = (1 - f_{NM} - f_M)f_{AW}$ ).<sup>4</sup>

**TABLE 5** Summary of the metrics assessed to validate the capability of the recalibrated DWI models to predict the EM-based gold standard. Shown are bias (relative bias, Equation 17) and error (relative error, Equation 18) for the baseline and the best-performing single- and multiparameter combinations. Corresponding values for the comparison of EM-based axonal volume fraction  $f_A^{(EM)}$  and DWI-based axonal water fraction  $f_{AW}^{(DWI)}$  are given in Appendix Table C1. Note that in Table D1 we provide an alternative version of this table including confidence intervals of the relative biases and error.

DWI model	Calibration parameters	Bias $\langle \delta \rangle$ (rel. bias $\bar{\delta}$ [%])	Error $\epsilon$ (rel. error $\bar{\epsilon}$ [%])
SMI	{}	-0.17 (-39)	0.25 (55)
	{ $U_{2,S}$ }	0.01 (1)	0.13 (29)
	{ $U_{2,S_{pred}}$ }	-0.06 (-13)	0.16 (35)
BAYDIFF	{}	-0.33 (-73)	0.23 (50)
	{ $U_{2,S}$ }	0.01 (2)	0.13 (29)
	{ $U_{2,S_{pred}}$ }	-0.2 (-45)	0.29 (64)
WMTI-W <sup>+</sup>	{}	-0.26 (-58)	0.24 (53)
	{ $U_{2,S}$ }	0.01 (1)	0.13 (28)
	{ $U_{2,S_{pred}}$ }	-0.13 (-28)	0.21 (46)
WMTI	{}	-0.14 (-31)	0.21 (46)
	{ $U_{2,S}$ }	0.0 (1)	0.12 (26)
	{ $U_{2,S_{pred}}$ }	-0.04 (-9)	0.13 (29)
NODDI	{}	0.01 (3)	0.27 (59)
	{ $U_{2,S}$ }	0.03 (6)	0.2 (44)
	{ $U_{2,S_{pred}}$ }	0.07 (16)	0.18 (41)
NODDI-DTI	{}	-0.01 (-3)	0.26 (56)
	{ $S$ }	0.03 (7)	0.23 (51)
	{ $S_{pred}$ }	0.02 (5)	0.24 (52)

Another study that compared WMTI with EM also estimated an offset and a scaling parameter.<sup>49</sup> However, these results were not directly comparable with ours due to several experimental differences: (1) they compared in vivo DWI with ex vivo EM, whereas our DWI measurements were performed on ex vivo tissue samples, (2) in EM they assessed both unmyelinated and myelinated axons, while in our data only the myelinated axons were assessed, (3) they estimated the slope and offset for the axonal water fractions and not, as we did here, for the axonal volume fractions. Ignoring issue (1), one could convert their *slope* and *offset* into our parameters using the following equations:  $s_{pred}^{(Jelescu)} = (1 - f_M^{(EM)})/slope$  and  $U^{(Jelescu)} = offset \cdot (1 - f_M^{(EM)})/slope - f_{A,U}$ . Here, we could not compare their slope and offset with ours because we did not have access to the individual myelin volume fractions  $f_M^{(EM)}$  and the volume fractions of unmyelinated axons  $f_{A,U}$  from Jelescu et al.<sup>49</sup>

## 5.2 | Impact of calibration parameters on bias and error between EM and DWI

All DWI models deviate clearly from the EM reference without additional calibration ( $\bar{\epsilon} \approx 46\%$  (WMTI) – 59% (NODDI), see Table 5). One reason for the observed difference is probably that the unmyelinated axons were not assessed with EM in this study, whereas the axonal water fraction measured with DWI is expected to be affected by both myelinated and unmyelinated axons. A comparison of the DWI-based axon volume fraction with our EM data therefore requires accounting for this difference, which we sought to achieve through additional calibration parameters. While SMI benefits the most from calibration (by 26%), WMTI-W<sup>+</sup>, BAYDIFF, and WMTI followed closely (by 25%, 21%, and 20%, respectively), while calibration had much less of an impact on NODDI and NODDI-DTI (15% and 5%). In particular, DWI models with fewer free model parameters improved less through calibration. It is striking that relative bias and relative error were reduced most substantially for the DWI models that fitted the compartmental diffusivities (see Table 5). Although the relative errors of these DWI models are all within their mutual confidence intervals, as given in Table D1 in the Appendix, there is an apparent trend that WMTI is somewhat more accurate. In particular, that WMTI was slightly more accurate than WMTI-W<sup>+</sup> was surprising, because WMTI-W<sup>+</sup> also accounts for fibre dispersion whereas WMTI does not. This may be partly attributed to chance, but our observation that WMTI-W<sup>+</sup> is more prone to noise (see simulations in Figure E1) is also in agreement with this trend. The larger number of free parameters in WMTI-W<sup>+</sup> as compared with WMTI can lead to a less well-conditioned optimisation problem, which, in turn, could explain the higher noise susceptibility of WMTI-W<sup>+</sup>. Moreover, the corpus callosum and fornix, which were the focus of this study, have highly aligned fibres, making the neglect of fibre dispersion less relevant than in other areas of the brain with higher fibre dispersion.

### 5.3 | Practical impact

The estimation of two calibration parameters, scaling and offset, is challenging, necessitating a well-distributed set of data points with a wide dynamic range. Acquiring such a dataset presents a practical challenge, as achieving a wide dynamic range within the same tissue type (e.g., white matter) requires significant modulation of tissue composition. In our study, we addressed this challenge by utilizing a multimodal dataset,<sup>38</sup> which included data from three mouse models with myelination ranging from hypo- to hypermyelinated, along with control mice.

However, applying this approach to humans is even more challenging, due to the limited availability of multimodal human data acquired with EM and DWI. To overcome this, we proposed a hybrid calibration approach that involves predicting a scaling factor and only fitting the remaining offset calibration parameter. We demonstrated that, for the standard and NODDI signal models with compartmental  $T_2$  dependence (Equation 10), the scaling parameter for axonal volume calibration can be predicted independently of the MRI protocol using known compartmental  $T_2$  values. Furthermore, for WMTI (and to a lesser degree for SMI) the hybrid calibration approach yielded results comparable with the control calibration method, where both calibration parameters were estimated. The practical relevance of the hybrid calibration method lies in the fact that, at least for the WMTI model, it can be nearly as accurate as the control calibration method, enabling more efficient estimation of the remaining offset calibration parameter. This will enable future studies to estimate the remaining offset calibration parameter from a limited amount of multimodal data, and is thus particularly relevant when used for human DWI applications where histological reference data are limited.

Overall, we found that the DWI model based axonal water fractions ( $f_{AW}^{(DWI)}$ ) (shown in Figure C1), and consequently also the axonal volume fractions  $f_A^{(DWI)}$ , are clearly less sensitive to demyelinating disease processes than a measure that is specifically assessing the fraction of myelinated axons. We believe that this is due to the sensitivity of the DWI-based signal to both myelinated and unmyelinated axons. Thus, it might be relevant for the planning of future clinical studies investigating demyelination processes using MRI to complement the DWI with other imaging contrasts that are more specific to the myelin pool, for example, the magnetisation transfer saturation rate as obtained from multiparameter mapping<sup>50</sup>

### 5.4 | Limitations

A number of limiting factors need to be considered when interpreting the results of our study. We made the strong but plausible assumption that the axon volume fraction across different mouse types is approximately constant. Since our reference EM measurements contained only the fraction of myelinated axons, we had to make this assumption to be able to test whether the offset is related to the fraction of unmyelinated axons. In principle, it is also possible to estimate the fraction of unmyelinated axons with EM.<sup>25,49</sup> However, those estimates have to be treated with caution because unmyelinated axons are more difficult to detect than myelinated axons even in high-quality EM data. This is because unmyelinated axons have lower contrast in EM and are often smaller in size than myelinated axons, making it more challenging to estimate the fraction of unmyelinated axons accurately.<sup>51</sup> This makes it likely that the estimated volume fraction of unmyelinated axons is less accurate than that of myelinated axons.

The assumption of no exchange between axonal and extracellular compartments implies that unmyelinated axons are impermeable, that is, the axonal cell membrane fully separates the intra- from the extra-axonal water. Note that, even if the unmyelinated axons were partly permeable, the diffusion time used in the present study is short enough (12 ms) to justify the assumption of reduced exchange between intra- and extracellular water. In both cases, the fraction of unmyelinated axons would contribute to the DWI-based axonal water fraction  $f_{AW}^{(DWI)}$ .

Furthermore, the assumption of highly aligned axons in WMTI might be violated because even the most aligned axons have an appreciable angular dispersion.<sup>52</sup> However, it was proposed that WMTI can be used for voxels with an FA threshold larger than 0.4.<sup>16</sup> This condition was violated in eight mouse-ROI combinations out of the total number of 60 combinations of our study. Despite these violations, the WMTI-based  $f_A^{(DWI)}$  showed the highest correspondence to its EM-based counterpart  $f_A^{(EM)}$ .

We used in vivo  $T_2$  values estimated across the entire human brain. However, compartmental relaxation times are likely to vary across fibre tracts, age, pathology, and between species. The proposed approach to estimate  $s_{pred}$  is a first-order approximation to correct for global compartmental  $T_2$  differences. Despite these simplifications, the hybrid calibration approach using  $s_{pred}$  worked almost as well for WMTI as the data-driven calibration approach.

Of note was the fact that the remaining error was also rather large (26%) for the DWI models that fitted the compartmental diffusivities. This may be attributed partly to a potentially large variance in the reference, originating in the relatively small EM section size, which is probably not sufficient to capture the distribution of axons in the MRI voxels representatively. The cross-sectional area of MRI voxels was  $\approx 144$  times larger for controls, and  $\approx 187$  for Pten, Rictor, and Tsc.

We estimated the theoretical scaling factors on the basis of in vivo compartmental  $T_2$  values rescaled from 3T to 15.2T by a factor estimated from ex vivo values in human brain and finally compared them with ex vivo mouse models, which were thoroughly washed in PBS and Gd-DTPA. At least for WMTI, the aforementioned limitations appear to be less relevant, since, for this DWI model,  $\{U_{2,s_{pred}}\}$  achieved similar accuracy to the combination  $\{U_{2,s}\}$ .

Furthermore, our results are based on a multimodal dataset of fixed tissue acquired in three different mouse models. This might be a problem, since the fraction of unmyelinated axons might be different between humans and mice. However, it was shown that the fraction of unmyelinated axons is constant across species.<sup>23</sup>

A potential, unexplained factor affecting the calibration parameters in this study is represented by uncontrolled tissue deformations due to chemical and physical treatment of the tissue samples. Tissue shrinkage due to chemical fixation can be ruled out as a limiting factor for the comparability of DWI and EM data, since this study was performed on ex vivo DWI and EM data that underwent common chemical fixation procedures (see Section 3.1.1). However, preparation of the tissue samples for EM required additional steps such as dehydration in graded ethanol and the cutting of sections (Section 3.1.3). Dehydration in ethanol has been identified as a source for shrinkage, varying across whole-brain samples between 2% and 3%.<sup>53</sup> However, shrinkage at tissue surfaces can be larger. Shrinkage and expansion of tissue structures due to the cutting and unfolding of sections for EM also cannot be ruled out and their magnitude is difficult to assess. Visual inspection of the EM sections, however, suggested that such deformations were rather small.

Finally, some models account for the effect of fixation by incorporating an additional dot compartment. In our study, only the NODDI model accounted for the dot compartment explicitly. In vivo, it has been shown that the dot compartment can be neglected,<sup>54,55</sup> but it is debatable whether this applies to the ex vivo case as well.<sup>19</sup> Although the other models neglected the dot compartment, the SMI, BAYDIFF, WMTI, and WMTI- $W^+$  models described the EM data better when using calibration parameters.

The translation of our findings to the in vivo human situation might be confounded by the effect of the fixative on the DWI signal. Future studies assessing the effect of fixation on DWI data might help to translate the estimated calibration parameters into the in vivo situation.

## 5.5 | Conclusion

In summary, we demonstrated that linear calibration with two biophysically motivated calibration parameters, an offset accounting for the volume fraction of unmyelinated axons and a scaling factor accounting for global compartmental  $T_2$  differences, enhances agreement between EM- and DWI-based estimates of the volume fraction of myelinated axons. Our findings suggest that, after calibration, the DWI models that fitted the compartmental diffusivities, that is, WMTI, BAYDIFF, WMTI- $W^+$ , and SMI, were the most accurate. Finally, we introduced a more efficient hybrid calibration approach, where only the offset is estimated but the scaling is predicted theoretically, and found that it was particularly accurate for WMTI, for which a similar one-to-one correspondence to EM was achieved using this approach. This makes the hybrid approach particularly appealing for usage in human brain data, where multimodal data are less common than for animals.

## ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG Priority Program 2041 “Computational Connectomics” [MO 2397/5-1; MO 2397/52]), the DFG Emmy Noether Stipend [MO 2397/4-1 and MO 2397/4-2], the BMBF (01EW1711A and 01EW1711B) in the framework of ERA-NET NEURON and the Forschungszentrum Medizintechnik Hamburg (fmthh; grant 01fmthh2017). MDD is supported by the National Institute of Health EB019980. MFC is supported by the MRC and Spinal Research Charity through the ERA-NET Neuron joint call (MR/R000050/1), the Wellcome Centre for Human Neuroimaging is supported by core funding from the Wellcome (203147/Z/16/Z).

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

Jan Malte Oeschger  <https://orcid.org/0000-0003-0237-923X>

Mark D. Does  <https://orcid.org/0000-0003-2701-6208>

## REFERENCES

- Jelescu IO, Budde MD. Design and validation of diffusion MRI Models of white matter. *Front Phys*. 2017;5:61. <https://www.frontiersin.org/articles/10.3389/fphy.2017.00061>
- Alexander DC, Dyrby TB, Nilsson M, Zhang H. Imaging brain microstructure with diffusion MRI: practicality and applications. *NMR Biomed*. 2019; 32(4):e3841. <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3841>
- Jelescu IO, Palombo M, Bagnato F, Schilling KG. Challenges for biophysical modeling of microstructure. *J Neurosci Methods*. 2020;344:108861. <https://www.sciencedirect.com/science/article/pii/S0165027020302843>
- Mohammadi S, Callaghan MF. Towards in vivo g-ratio mapping using MRI: Unifying myelin and diffusion imaging. *J Neurosci Methods*. 2021;348: 108990. <https://linkinghub.elsevier.com/retrieve/pii/S0165027020304131>

5. Kraguljac NV, Guerreri M, Strickland MJ, Zhang H. Neurite orientation dispersion and density imaging in psychiatric disorders: A systematic literature review and a technical note. *Biolog Psychiat Global Open Sci*. 2023;3(1):10-21. <https://www.sciencedirect.com/science/article/pii/S2667174322000064>
6. Assaf Y, Freidlin RZ, Rohde GK, Basser PJ. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnet Reson Med*. 2004;52(5):965-978. <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.20274>
7. Jespersen SN, Kroenke CD, Østergaard L, Ackerman JJH, Yablonskiy DA. Modeling dendrite density from magnetic resonance diffusion measurements. *NeuroImage*. 2007;34(4):1473-1486. <https://www.sciencedirect.com/science/article/pii/S1053811906010950>
8. Novikov DS, Fieremans E, Jespersen SN, Kiselev VG. Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR Biomed*. 2019;32(4):e3998. <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3998>
9. Cox SR, Ritchie SJ, Tucker-Drob EM, et al. Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nat Commun*. 2016;7(1):13629. <http://www.nature.com/articles/ncomms13629>
10. Elliott LT, Sharp K, Alfaro-Almagro F, et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature*. 2018;562(7726):210-216. <http://www.nature.com/articles/s41586-018-0571-7>
11. Genç E, Fraenz C, Schlüter C, et al. Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nat Commun*. 2018;9(1):1905. <https://www.nature.com/articles/s41467-018-04268-8>
12. Donat CK, Yanez Lopez M, Sastre M, et al. From biomechanics to pathology: Predicting axonal injury from patterns of strain after traumatic brain injury. *Brain*. 2021;144(1):70-91. doi:10.1093/brain/awaa336
13. Coutu J-P, Chen JJ, Rosas HD, Salat DH. Non-Gaussian water diffusion in aging white matter. *Neurobiol Aging*. 2014;35(6):1412-1421 en. <https://www.sciencedirect.com/science/article/pii/S0197458013006192>
14. Karolis VR, Corbetta M, Thiebaut de Schotten M. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nat Commun*. 2019;10(1):1417. <https://www.nature.com/articles/s41467-019-09344-1>
15. Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*. 2012;61(4):1000-1016. <http://www.sciencedirect.com/science/article/pii/S1053811912003539>
16. Fieremans E, Jensen JH, Helpert JA. White matter characterization with diffusional kurtosis imaging. *NeuroImage*. 2011;58(1):177-188. <http://www.sciencedirect.com/science/article/pii/S1053811911006148>
17. Stikov N, Campbell JSW, Stroh T, et al. In vivo histology of the myelin g-ratio with magnetic resonance imaging. *NeuroImage*. 2015;118:397-405.
18. Campbell JSW, Leppert IR, Narayanan S, et al. Promise and pitfalls of g-ratio estimation with MRI. *NeuroImage*. 2018;182:80-96. <https://www.sciencedirect.com/science/article/pii/S1053811917306857>
19. Veraart J, Novikov DS, Fieremans E. TE dependent Diffusion Imaging (TEdDI) distinguishes between compartmental T2 relaxation times. *NeuroImage*. 2018;182:360-369. <https://www.sciencedirect.com/science/article/pii/S1053811917307784>
20. Lampinen B, Szczepankiewicz F, Novén M, et al. Searching for the neurite density with diffusion MRI: Challenges for biophysical modeling. *Human Brain Mapping*. 2019;40(8):2529-2545. <https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.24542>
21. Gong T, Tong Q, He H, Sun Y, Zhong J, Zhang H. MTE-NODDI: Multi-TE NODDI for disentangling non-T2-weighted signal fractions from compartment-specific T2 relaxation times. *NeuroImage*. 2020;217:116906. <https://www.sciencedirect.com/science/article/pii/S105381192030392X>
22. Frigo M, Fick RHJ, Zucchelli M, Deslauriers-Gauthier S, Deriche R. Multi-tissue multi-compartment models of diffusion MRI. *bioRxiv*. 2021. <https://www.biorxiv.org/content/10.1101/2021.01.29.428843v1>
23. Olivares R, Montiel J, Aboitiz F. Species differences and similarities in the fine structure of the mammalian corpus callosum. *Brain, Behav Evol*. 2001;57(2):98-105. <https://www.karger.com/Article/FullText/47229>
24. Jelescu IO, Veraart J, Fieremans E, Novikov DS. Degeneracy in model parameter estimation for multicompartmental diffusion in neuronal tissue. *NMR Biomed*. 2016;29(1):33-47. <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3450>
25. Abdollahzadeh A, Belevich I, Jokitalo E, Tohka J, Sierra A. Automated 3D axonal morphometry of white matter. *Scientif Rep*. 2019;9(1):6084.
26. Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. *Brain Res*. 1992;598(1):143-153. <https://www.sciencedirect.com/science/article/pii/000689939290178C>
27. In: Johansen-Berg H, Behrens TEJ, eds. *Diffusion MRI*. San Diego: Academic Press; 2009:481-490. <https://www.sciencedirect.com/science/article/pii/B9780123747099000249>
28. Gibson EM, Purger D, Mount CW, et al. Neuronal activity promotes oligodendrogenesis and adaptive myelination in the mammalian brain. *Science*. 2014;344(6183):1252304. <https://www.science.org/doi/10.1126/science.1252304>
29. Lee H-H, Yaros K, Veraart J, et al. Along-axon diameter variation and axonal orientation dispersion revealed with 3D electron microscopy: Implications for quantifying brain white matter microstructure with histology and diffusion MRI. *Brain Struct Funct*. 2019;224(4):1469-1488. doi:10.1007/s00429-019-01844-6
30. West KL, Kelm ND, Carson RP, Gochberg DF, Ess KC, Does MD. Myelin volume fraction imaging with MRI. *NeuroImage*. 2018;182:511-521. <https://www.sciencedirect.com/science/article/pii/S1053811916307935>
31. Zaimi A, Wabartha M, Herman V, Antonsanti P-L, Perone CS, Cohen-Adad J. AxonDeepSeg: Automatic axon and myelin segmentation from microscopy data using convolutional neural networks. *Scientif Rep*. 2018;8(1):3816. <http://www.nature.com/articles/s41598-018-22181-4>
32. Beaulieu C, Allen PS. Determinants of anisotropic water diffusion in nerves. *Magnet Reson Med*. 1994;31(4):394-400. <https://onlinelibrary.wiley.com/doi/10.1002/mrm.1910310408>
33. Edwards LJ, Pine KJ, Ellerbrock I, Weiskopf N, Mohammadi S. NODDI-DTI: Estimating neurite orientation and dispersion parameters from a diffusion tensor in healthy white matter. *Front Neurosci*. 2017;11:720. <https://www.frontiersin.org/article/10.3389/fnins.2017.00720>
34. Novikov DS, Veraart J, Jelescu IO, Fieremans E. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI. *NeuroImage*. 2018;174:518-538. <https://www.sciencedirect.com/science/article/pii/S1053811918301915>
35. Reisert M, Kellner E, Dhital B, Hennig J, Kiselev VG. Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. *NeuroImage*. 2017;147:964-975. <https://www.sciencedirect.com/science/article/pii/S1053811916305353>
36. Jespersen SN, Olesen JL, Hansen B, Shemesh N. Diffusion time dependence of microstructural parameters in fixed spinal cord. *NeuroImage*. 2018;182:329-342. <https://www.sciencedirect.com/science/article/pii/S1053811917306869>

37. Kelm ND, West KL, Carson RP, Gochberg DF, Ess KC, Does MD. Evaluation of diffusion kurtosis imaging in ex vivo hypomyelinated mouse brains. *NeuroImage*. 2016;124:612-626. <https://www.sciencedirect.com/science/article/pii/S1053811915008411>
38. West KL, Kelm ND, Carson RP, Alexander DC, Gochberg DF, Does MD. Experimental studies of g-ratio MRI in ex vivo mouse brain. *NeuroImage*. 2018;167:366-371. <https://www.sciencedirect.com/science/article/pii/S1053811917310108>
39. Alexander DC, Hubbard PL, Hall MG, et al. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage*. 2010;52(4):1374-1389. <https://www.sciencedirect.com/science/article/pii/S1053811910007755>
40. Panagiotaki E, Schneider T, Siow B, Hall MG, Lythgoe MF, Alexander DC. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. *NeuroImage*. 2012;59(3):2241-2254. <https://www.sciencedirect.com/science/article/pii/S1053811911011566>
41. Tofts P. *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*. 1. Auflage: John Wiley & Sons; 2004.
42. Shepherd TM, Thelwall PE, Stanisz GJ, Blackband SJ. Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. *Magnet Reson Med*. 2009;62(1):26-34. <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.21977>
43. Cohen-Adad J, Does M, Duval T, et al. White Matter Microscopy Database. <https://osf.io/yp4qg/>; 2016.
44. David G, Fricke B, Oeschger JM, Ruthotto L, et al. ACID: A comprehensive toolbox for image processing and modeling of brain, spinal cord, and post-mortem diffusion MRI data. *bioRxiv*. 2023. <https://www.biorxiv.org/content/10.1101/2023.10.13.562027v1>
45. Coelho S, Baete SH, Lemberskiy G, et al. Reproducibility of the standard model of diffusion in white matter on clinical MRI systems. *NeuroImage*. 2022;257:119290. <https://www.sciencedirect.com/science/article/pii/S1053811922004104>
46. Dhital B, Reiser M, Kellner E, Kiselev VG. Intra-axonal diffusivity in brain white matter. *NeuroImage*. 2019;189:543-550. <https://www.sciencedirect.com/science/article/pii/S1053811919300151>
47. Burnham KP, Anderson DR, Burnham KP. *Model selection and multimodel inference: A practical information-theoretic approach*. 2nd ed.: Springer; 2002.
48. Altman DG, Bland JM. Measurement in medicine: The analysis of method comparison studies. *Statistician*. 1983;32(3):307. <https://www.jstor.org/stable/2987937?origin%3Dcrossref>
49. Jelescu IO, Zurek M, Winters KV, et al. In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. *NeuroImage*. 2016;132:104-114. <https://www.sciencedirect.com/science/article/pii/S1053811916001026>
50. Tabelow K, Balteau E, Ashburner J, et al. hMRI - A toolbox for quantitative MRI in neuroscience and clinical research. *NeuroImage*. 2019;194:191-210.
51. Seppeband F, Alexander DC, Kurniawan ND, Reutens DC, Yang Z. Towards higher sensitivity and stability of axon diameter estimation with diffusion-weighted MRI. *NMR Biomed*. 2016;29(3):293-308. <https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.3462>
52. Ronen I, Budde M, Ercan E, Annese J, Techawiboonwong A, Webb A. Microstructural organization of axons in the human corpus callosum quantified by diffusion-weighted magnetic resonance spectroscopy of N-acetylaspartate and post-mortem histology. *Brain Struct Funct*. 2014;219(5):1773-1785. doi:10.1007/s00429-013-0600-0
53. Rodgers G, Tanner C, Schulz G, et al. Virtual histology of an entire mouse brain from formalin fixation to paraffin embedding. Part 2: Volumetric strain fields and local contrast changes. *J Neurosci Methods*. 2022;365:109385. <https://linkinghub.elsevier.com/retrieve/pii/S0165027021003204>
54. Dhital B, Kellner E, Kiselev VG, Reiser M. The absence of restricted water pool in brain white matter. *NeuroImage*. 2018;182:398-406. <https://www.sciencedirect.com/science/article/pii/S1053811917308856>
55. Tax CMW, Kleban E, Chamberland M, Baraković M, Rudrapatna U, Jones DK. Measuring compartmental T2-orientational dependence in human brain white matter using a tiltable RF coil and diffusion-T2 correlation MRI. *NeuroImage*. 2021;236:117967. <https://www.sciencedirect.com/science/article/pii/S1053811921002445>
56. Birkel C, Langkammer C, Haybaeck J, et al. Temperature-induced changes of magnetic resonance relaxation times in the human brain: A postmortem study. *Magnet Reson Med*. 2014;71(4):1575-1580. <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.24799>
57. Coelho S, Pozo JM, Jespersen SN, Jones DK, Frangi AF. Resolving degeneracy in diffusion MRI biophysical model parameter estimation using double diffusion encoding. *Magnet Reson Med*. 2019;82(1):395-410. <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27714>
58. Oeschger JM, Tabelow K, Mohammadi S. Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study. *Magnet Reson Med*. 2023;89(2):787-799. <https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.29474>

**How to cite this article:** Papazoglou S, Ashtarayeh M, Oeschger JM, Callaghan MF, Does MD, Mohammadi S. Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted MRI and electron microscopy. *NMR in Biomedicine*. 2023;e5070. doi:10.1002/nbm.5070

## APPENDIX A: ADDITIONAL DATA

**TABLE A1** Summary of all fitted models and calibration parameter combinations as shown in Figure 3. The *BIC* was determined on the basis of all available data. In contrast,  $U_1$ ,  $U_2$ , and  $s$  were estimated in a leave-one-out fashion, in which each mouse individual was excluded from the computation once in order to get an estimate of the standard deviation (see also Table 4). The predicted scaling,  $s_{\text{pred}}$ , was calculated as described in Section 2.3.

DWI model	Calibr. parameters	BIC	$U_1$ (SD)	$U_2$ (SD)	$s$ (SD)	$s_{\text{pred}}$	Rel. $\Delta s$ [%]
SMI	{}	-176	0 (-)	0 (-)	1 (-)	-	-
	{ $U_1$ }	-195	0.13 (0.0)	- (-)	1 (-)	-	-
	{ $U_2$ }	-230	- (-)	0.37 (0.01)	1 (-)	-	-
	{ $U_1, U_2$ }	-292	0.13 (0.0)	0.37 (0.01)	1 (-)	-	-
	{ $s$ }	-250	- (-)	- (-)	0.62 (0.02)	0.89	-30.0
	{ $U_1, s$ }	-246	0.0 (0.0)	- (-)	0.62 (0.02)	0.89	-30.0
	{ $U_2, s$ }	-312	- (-)	0.24 (0.01)	0.72 (0.01)	0.89	-19.0
	{ $U_1, U_2, s$ }	-308	0.0 (0.0)	0.24 (0.01)	0.72 (0.01)	0.89	-19.0
BAYDIFF	{}	-116	0 (-)	0 (-)	1 (-)	-	-
	{ $U_1$ }	-172	0.31 (0.0)	- (-)	1 (-)	-	-
	{ $U_2$ }	-138	- (-)	0.44 (0.01)	1 (-)	-	-
	{ $U_1, U_2$ }	-262	0.31 (0.0)	0.44 (0.01)	1 (-)	-	-
	{ $s$ }	-268	- (-)	- (-)	0.47 (0.01)	0.93	-49.0
	{ $U_1, s$ }	-264	0.0 (0.0)	- (-)	0.47 (0.01)	0.93	-49.0
	{ $U_2, s$ }	-313	- (-)	0.18 (0.0)	0.52 (0.0)	0.93	-44.0
	{ $U_1, U_2, s$ }	-309	0.0 (0.0)	0.18 (0.0)	0.52 (0.0)	0.93	-44.0
WMTI-W <sup>+</sup>	{}	-139	0 (-)	0 (-)	1 (-)	-	-
	{ $U_1$ }	-178	0.22 (0.0)	- (-)	1 (-)	-	-
	{ $U_2$ }	-173	- (-)	0.43 (0.0)	1 (-)	-	-
	{ $U_1, U_2$ }	-275	0.22 (0.0)	0.43 (0.0)	1 (-)	-	-
	{ $s$ }	-258	- (-)	- (-)	0.52 (0.01)	0.89	-41.0
	{ $U_1, s$ }	-254	0.0 (0.0)	- (-)	0.52 (0.01)	0.89	-41.0
	{ $U_2, s$ }	-315	- (-)	0.21 (0.0)	0.59 (0.01)	0.89	-33.0
	{ $U_1, U_2, s$ }	-311	0.0 (0.0)	0.21 (0.0)	0.59 (0.01)	0.89	-33.0
WMTI	{}	-200	0 (-)	0 (-)	1 (-)	-	-
	{ $U_1$ }	-218	0.1 (0.0)	- (-)	1 (-)	-	-
	{ $U_2$ }	-255	- (-)	0.3 (0.0)	1 (-)	-	-
	{ $U_1, U_2$ }	-311	0.1 (0.0)	0.3 (0.0)	1 (-)	-	-
	{ $s$ }	-264	- (-)	- (-)	0.68 (0.01)	0.89	-24.0
	{ $U_1, s$ }	-260	0.0 (0.0)	- (-)	0.68 (0.01)	0.89	-24.0
	{ $U_2, s$ }	-327	- (-)	0.21 (0.0)	0.76 (0.01)	0.89	-14.0
	{ $U_1, U_2, s$ }	-323	0.0 (0.0)	0.21 (0.0)	0.76 (0.01)	0.89	-14.0
NODDI	{}	-232	0 (-)	0 (-)	1 (-)	-	-
	{ $U_1$ }	-232	0.0 (0.0)	- (-)	1 (-)	-	-
	{ $U_2$ }	-259	- (-)	0.18 (0.01)	1 (-)	-	-
	{ $U_1, U_2$ }	-255	0.0 (0.0)	0.18 (0.01)	1 (-)	-	-
	{ $s$ }	-232	- (-)	- (-)	0.97 (0.03)	0.9	7.0
	{ $U_1, s$ }	-228	0.0 (0.0)	- (-)	0.97 (0.03)	0.9	7.0
	{ $U_2, s$ }	-259	- (-)	0.21 (0.01)	1.11 (0.01)	0.9	22.0
	{ $U_1, U_2, s$ }	-255	0.0 (0.0)	0.21 (0.01)	1.11 (0.01)	0.9	22.0

(Continues)



TABLE A1 (Continued)

DWI model	Calibr. parameters	BIC	$U_1$ (SD)	$U_2$ (SD)	$s$ (SD)	$s_{\text{pred}}$	Rel. $\Delta s$ [%]
NODDI-DTI	{}	-237	0 (-)	0 (-)	1 (-)	-	-
	$\{U_1\}$	-237	0.0 (0.0)	- (-)	1 (-)	-	-
	$\{U_2\}$	-240	- (-)	0.07 (0.0)	1 (-)	-	-
	$\{U_1, U_2\}$	-236	0.0 (0.0)	0.07 (0.0)	1 (-)	-	-
	$\{s\}$	-244	0.0 (0.0)	0.0 (0.0)	0.87 (0.02)	0.89	-2.0
	$\{U_1, s\}$	-240	0.0 (0.0)	- (-)	0.87 (0.02)	0.89	-2.0
	$\{U_2, s\}$	-242	- (-)	0.05 (0.0)	0.88 (0.02)	0.89	-1.0
	$\{U_1, U_2, s\}$	-238	0.0 (0.0)	0.05 (0.0)	0.88 (0.02)	0.89	-1.0

APPENDIX B: ESTIMATION OF COMPARTMENTAL  $T_2$  FROM THE LITERATURE

For estimating  $s_{\text{pred}}$ , we converted the compartmental  $T_2$  values  $T_{2,a}(3T) \approx 83\text{ms}$  and  $T_{2,e}(3T) \approx 59\text{ms}$  from Tax et al.<sup>55</sup> from 3T to 15.2T. The conversion factor  $r$  was estimated from average values for the transverse relaxation time in ex vivo human white matter on the basis of mono-exponential models:  $T_2(3T) = 83.8\text{ms}$ <sup>56</sup> and  $T_2(15.2T) \approx 33\text{ms}$  (from the supplementary material of West et al.<sup>30</sup>). The decrease of the relaxation time from 3T to 15.2T can then be estimated by the ratio

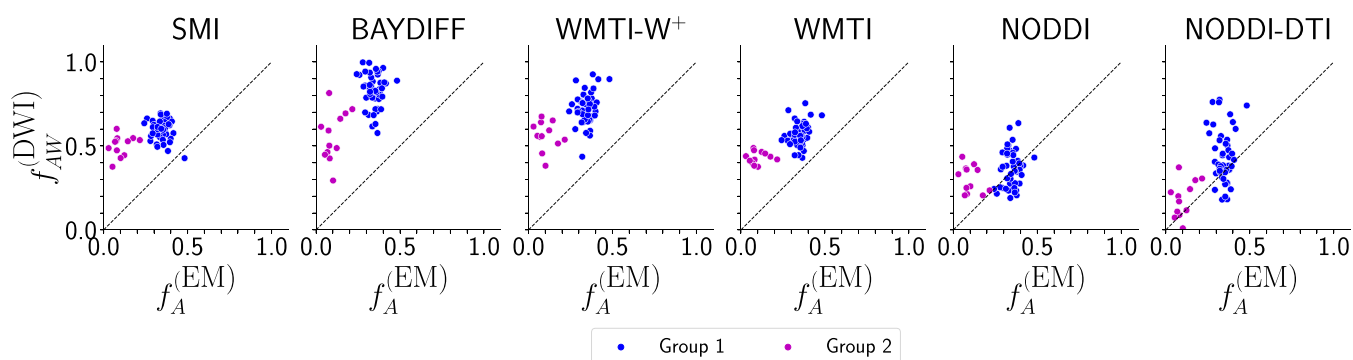
$$r = T_2(15.2T)/T_2(3T) \approx 0.4. \quad (\text{B1})$$

With Equation (B1), the relaxation rates of the individual compartments can be estimated as

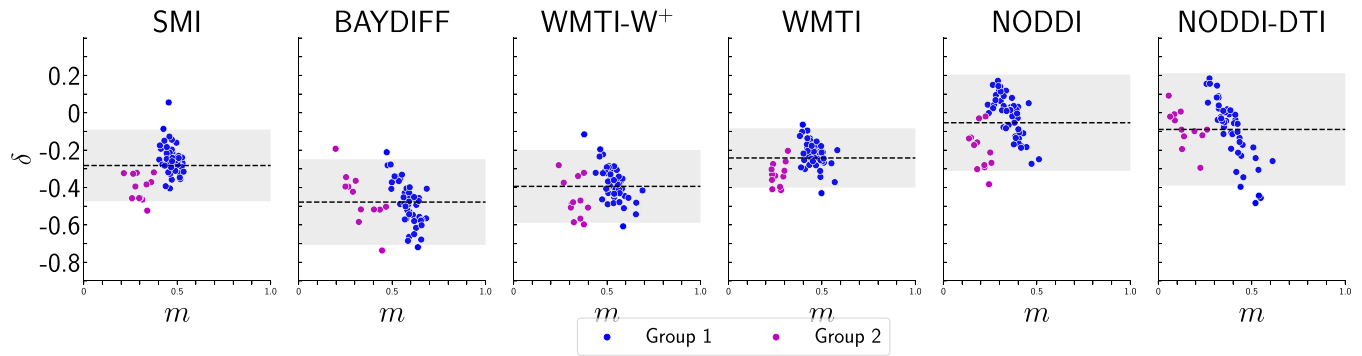
$$\begin{aligned} T_{2,a}(15.2T) &= rT_{2,a}(3T) \approx 33\text{ms}, \\ T_{2,e}(15.2T) &= rT_{2,e}(3T) \approx 23\text{ms}. \end{aligned} \quad (\text{B2})$$

Note that we assumed that the compartmental  $T_2$  times in the fixed tissue were the same as for in vivo tissue. This assumption is based on the observation that washing the samples in PBS retrieves  $T_2$  values similar to the in vivo case.<sup>42</sup> From Equations (B2), the compartmental differences are then  $\Delta_e \approx 12.6\text{s}^{-1}$  and  $\Delta_{\text{iso}} \approx -32.0\text{s}^{-1}$ , where  $T_{2,\text{iso}} = 1\text{ms}$  was assumed.<sup>21</sup> For  $T_{2,\text{dot}}$  no values could be found to the best of our knowledge. We assumed  $T_{2,\text{dot}} = T_{2,e}$ , that is,  $\Delta_{\text{dot}} = \Delta_e$  in Equation (10). Estimates of the compartmental values were employed in the hybrid approach described in Section 3.4.3.

## APPENDIX C: COMPARISON OF DWI-BASED AXON WATER FRACTION AND EM-BASED AXON VOLUME FRACTION



**FIGURE C1** The scatter plots depict the comparison of EM-based axonal volume ( $f_A^{(\text{EM})}$ ) and the axonal water fraction of the corresponding DWI models without rescaling by  $1 - f_M^{(\text{EM})}$  as in Equation (3) (compare with Figure 4).



**FIGURE C2** Bland–Altman plots of differences  $\delta$  (Equation 15) versus mean  $m$  (Equation 16) between EM-based axonal volume fraction ( $f_A^{(EM)}$ ) and DWI-based axonal water fraction ( $f_{AW}^{(DWI)}$ ) corresponding to Figure 5 (individual data shown are in correspondence with Figure C1). The corresponding biases and errors are summarised in Table C1

**TABLE C1** Bias (relative bias) and error (relative error) for the data points shown in Figure C2 (compare with Table 5). The definition of relative bias and error is given in Equations (18) and (19), respectively. Note that, for this table,  $f_A^{(DWI)}$  was replaced by  $f_{AW}^{(DWI)}$  in the equations.

DWI model	Bias ( $\delta$ ) (rel. bias $\bar{\delta}$ [%])	Error $\epsilon$ (rel. error $\bar{\epsilon}$ [%])
SMI	−0.28 (−62)	0.19 (42)
BAYDIFF	−0.39 (−87)	0.23 (51)
WMTI- $W^+$	−0.48 (−106)	0.19 (43)
WMTI	−0.24 (−53)	0.16 (35)
NODDI	−0.05 (−12)	0.26 (57)
NODDI-DTI	−0.09 (−20)	0.3 (67)

In Figure C1 we compare the histological reference  $f_A^{(EM)}$  with its DWI-based counterpart  $f_{AW}^{(DWI)}$ . As previously reported,  $f_{AW}^{(DWI)}$  correlates with  $f_A^{(EM)}$  across all DWI models,<sup>37</sup> but  $f_A^{(EM)}$  is more sensitive to the demyelination process than  $f_{AW}^{(DWI)}$  (see also the corresponding discussion in Section 5.3). Another interesting observation is that  $f_{AW}^{(DWI)}$  from NODDI–DTI shows almost a one-to-one correspondence with  $f_A^{(EM)}$ . The corresponding Bland–Altman plots are shown in Figure C2 and the relative biases and errors are summarised in Table C1.

#### APPENDIX D: CONFIDENCE INTERVALS FOR RELATIVE BIASES AND ERRORS

**TABLE D1** Table analogous to Table 5 but including confidence intervals ( $\pm 1.96 \cdot SD$ ) for the relative biases  $\bar{\delta}$  and relative errors  $\bar{\epsilon}$  estimated as described in Appendix D.

DWI model	Calibration parameters	Bias ( $\delta$ ) (rel. bias $\bar{\delta} \pm 1.96 \cdot SD$ [%])	Error $\epsilon$ (rel. error $\bar{\epsilon} \pm 1.96 \cdot SD$ [%])
SMI	{}	−0.17 (−39±5)	0.25 (56±5)
	{ $U_{2,s}$ }	0.01 (1±1)	0.13 (29±1)
	{ $U_{2,s_{pred}}$ }	−0.06 (−13±2)	0.16 (35±2)
BAYDIFF	{}	−0.33 (−74±7)	0.23 (51±4)
	{ $U_{2,s}$ }	0.01 (2±1)	0.13 (30±2)
	{ $U_{2,s_{pred}}$ }	−0.2 (−46±6)	0.29 (65±8)
WMTI- $W^+$	{}	−0.26 (−59±6)	0.24 (54±6)
	{ $U_{2,s}$ }	0.01 (1±1)	0.13 (29±3)
	{ $U_{2,s_{pred}}$ }	−0.13 (−29±4)	0.21 (47±5)
WMTI	{}	−0.14 (−32±4)	0.21 (47±5)
	{ $U_{2,s}$ }	0.0 (1±1)	0.12 (26±2)
	{ $U_{2,s_{pred}}$ }	−0.04 (−9±1)	0.13 (29±3)

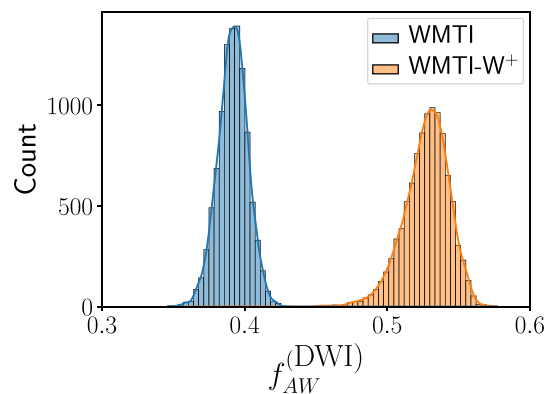
(Continues)

TABLE D1 (Continued)

DWI model	Calibration parameters	Bias $\langle \delta \rangle$ (rel. bias $\bar{\delta} \pm 1.96 \cdot SD$ [%])	Error $\epsilon$ (rel. error $\bar{\epsilon} \pm 1.96 \cdot SD$ [%])
NODDI	{}	0.01 (3 $\pm$ 3)	0.26 (60 $\pm$ 6)
	{ $U_2, s$ }	0.03 (6 $\pm$ 1)	0.2 (45 $\pm$ 4)
	{ $U_2, s_{\text{pred}}$ }	0.07 (16 $\pm$ 2)	0.18 (41 $\pm$ 3)
NODDI-DTI	{}	-0.01 (-3 $\pm$ 2)	0.26 (57 $\pm$ 5)
	{ $s$ }	0.03 (7 $\pm$ 2)	0.23 (52 $\pm$ 5)
	{ $s_{\text{pred}}$ }	0.02 (5 $\pm$ 2)	0.24 (53 $\pm$ 5)

In order to estimate confidence intervals for the relative biases and errors, we had to perform a slightly modified leave-one-out analysis compared with the one before (see also Table 5). There, the offsets  $U_j$  and the scaling  $s$  were determined as averages over 15 samples each with one distinct mouse individual excluded. Then, biases and errors as given in Table 5 were determined. Here, we again computed offsets  $U_j$  and scaling  $s$  15 times each with one distinct mouse individual excluded, but now each of the 15  $U_j$  and  $s$  were used to compute 15 biases and errors in order to estimate their variances as given in Table D1. For this reason, the biases and errors in Tables 5 and D1 differ slightly.

#### APPENDIX E: NOISE ANALYSIS OF WMTI AND WMTI-W<sup>+</sup>



**FIGURE E1** Histograms of results of axon water fraction estimated with WMTI-W<sup>+</sup> (orange) and WMTI (blue) based on a noise simulation of 10 000 noise samples simulated for SNR = 150.

To test the hypothesis of whether WMTI-W<sup>+</sup> is more prone to noise than WMTI, a noise simulation based on a variant of the standard model that uses the Watson distribution to model neurite dispersion (equation 1 in Jespersen et al.<sup>36</sup>) was performed. The signal model was used as a forward model and the integral was solved numerically (Matlab, Lebedev quadrature). First, noise-free signals were simulated for the 61 gradient directions and  $b$  values that were used to acquire the mouse dataset. To simulate the noise-free data, a set of biophysical parameters ( $\nu = 0.73, K = 8, D_a = 2.0, D_e^{\parallel} = 1.0, D_e^{\perp} = 0.3$ ) describing highly aligned fibres was taken from Coelho et al.<sup>57</sup> The noise-free signals were then used to compute noisy magnitude signals according to  $S_{\text{cont}} = |S_{\text{noise-free}} + a + b_i|$  (similar to Oeschger et al.<sup>58</sup>), where  $a, b_i \in N(0, \sigma)$  were each randomly drawn from a zero-mean Gaussian with standard deviation  $\sigma = \sqrt{2}S_0/\text{SNR}$ , where SNR is the signal-to-noise ratio and  $S_0 = 1$ . Following this procedure, 10 000 noise samples for SNR = 150 were simulated and fitted with standard DKI (the NLLS algorithm from the ACID toolbox was used), the results of which could then be used to compute the axon water fraction  $f_{AW}^{(DWI)}$  for WMTI-W<sup>+</sup> and WMTI. Histograms of the results are shown in Figure E1. The histograms in Figure E1 were compared using the standard deviation, which served as a measure to quantify the spread of results of both methods induced by noise. The standard deviation of WMTI-W<sup>+</sup> was 0.016, while the standard deviation of WMTI was 0.01. The fact that the standard deviation of WMTI-W<sup>+</sup> was 60% larger than for WMTI supports the hypothesis that WMTI-W<sup>+</sup> is more prone to noise than WMTI (see Section 5.2).

**APPENDIX F: INSTRUCTIONS FOR DKI FITTING USING THE ACID TOOLBOX**

- Installation instructions for the ACID toolbox are described in detail here: <https://bitbucket.org/siawoosh/acid-artefact-correction-in-diffusion-mri/wiki/Home>
- In this article, we used this tagged version of the ACID toolbox: <https://bitbucket.org/siawoosh/acid-artefact-correction-in-diffusion-mri/commits/d5ce665d709647aa1122c7b8b0b71420bc15e6e9>
- To estimate DKI parameters we used the following batch: [https://drive.google.com/file/d/1lw-4EWk57IoMH5DWovzxz8Ru8f05U3-h/view?usp=drive\\_link](https://drive.google.com/file/d/1lw-4EWk57IoMH5DWovzxz8Ru8f05U3-h/view?usp=drive_link). Please note that the corresponding paths might have to be adjusted to your local paths and that the data need to be requested from Mark D. Does.

## References

- [1] D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet, “MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders,” *Radiology*, vol. 161, pp. 401–407, Nov. 1986.
- [2] J.-P. Coutu, J. J. Chen, H. D. Rosas, and D. H. Salat, “Non-Gaussian water diffusion in aging white matter,” *Neurobiology of Aging*, vol. 35, pp. 1412–1421, June 2014.
- [3] E. Genç, C. Fraenz, C. Schlüter, P. Friedrich, R. Hossiep, M. C. Voelke, J. M. Ling, O. Güntürkün, and R. E. Jung, “Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence,” *Nature Communications*, vol. 9, May 2018.
- [4] C. K. Donat, M. Yanez Lopez, M. Sastre, N. Baxan, M. Goldfinger, R. Seeamber, F. Müller, P. Davies, P. Hellyer, P. Siegkas, S. Gentleman, D. J. Sharp, and M. Ghajari, “From biomechanics to pathology: predicting axonal injury from patterns of strain after traumatic brain injury,” *Brain*, vol. 144, pp. 70–91, Jan. 2021.
- [5] D. Liewald, R. Miller, N. Logothetis, H.-J. Wagner, and A. Schüz, “Distribution of axon diameters in cortical white matter: an electron-microscopic study on three human brains and a macaque,” *Biological Cybernetics*, vol. 108, no. 5, p. 541, 2014.
- [6] E. O. Stejskal and J. E. Tanner, “Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient,” *The journal of chemical physics*, vol. 42, no. 1, pp. 288–292, 1965.
- [7] P. J. Basser, J. Mattiello, and D. LeBihan, “Estimation of the effective self-diffusion tensor from the NMR spin echo,” *Journal of Magnetic Resonance. Series B*, vol. 103, pp. 247–254, Mar. 1994.
- [8] J. H. Jensen, J. A. Helpert, A. Ramani, H. Lu, and K. Kaczynski, “Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of Magnetic Resonance Imaging,” *Magnetic Resonance in Medicine*, vol. 53, no. 6, pp. 1432–1440, 2005.
- [9] J. H. Jensen and J. A. Helpert, “MRI quantification of non-Gaussian water diffusion by kurtosis analysis,” *NMR in Biomedicine*, vol. 23, no. 7, pp. 698–710, 2010.
- [10] D. K. Jones, PhD, *Diffusion MRI: Theory, Methods, and Applications*. Oxford University Press, Nov. 2010.
- [11] B. Hansen, N. Shemesh, and S. N. Jespersen, “Fast imaging of mean, axial and radial diffusion kurtosis,” *NeuroImage*, vol. 142, pp. 381–393, Nov. 2016.

- [12] B. Hansen, A. R. Khan, N. Shemesh, T. E. Lund, R. Sangill, S. F. Eskildsen, L. Østergaard, and S. N. Jespersen, “White matter biomarkers from fast protocols using axially symmetric diffusion kurtosis imaging,” *NMR in Biomedicine*, vol. 30, no. 9, p. e3741, 2017.
- [13] B. Hansen and S. N. Jespersen, “Recent developments in fast kurtosis imaging,” *Frontiers in Physics*, vol. 5, p. 40, 2017.
- [14] S. N. Jespersen, C. D. Kroenke, L. Østergaard, J. J. H. Ackerman, and D. A. Yablonskiy, “Modeling dendrite density from magnetic resonance diffusion measurements,” *NeuroImage*, vol. 34, pp. 1473–1486, Feb. 2007.
- [15] E. Fieremans, J. H. Jensen, and J. A. Helpert, “White matter characterization with diffusional kurtosis imaging,” *NeuroImage*, vol. 58, pp. 177–188, Sept. 2011.
- [16] I. O. Jelescu, J. Veraart, E. Fieremans, and D. S. Novikov, “Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue,” *NMR in biomedicine*, vol. 29, pp. 33–47, Jan. 2016.
- [17] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander, “NODDI: Practical *in vivo* neurite orientation dispersion and density imaging of the human brain,” *NeuroImage*, vol. 61, pp. 1000–1016, July 2012.
- [18] L. J. Edwards, K. J. Pine, I. Ellerbrock, N. Weiskopf, and S. Mohammadi, “NODDI-DTI: Estimating Neurite Orientation and Dispersion Parameters from a Diffusion Tensor in Healthy White Matter,” *Frontiers in Neuroscience*, vol. 11, p. 720, Dec. 2017.
- [19] D. S. Novikov, J. Veraart, I. O. Jelescu, and E. Fieremans, “Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion MRI,” *NeuroImage*, vol. 174, pp. 518 – 538, 2018.
- [20] S. N. Jespersen, J. L. Olesen, B. Hansen, and N. Shemesh, “Diffusion time dependence of microstructural parameters in fixed spinal cord,” *NeuroImage*, vol. 182, pp. 329–342, 2018.
- [21] M. Reisert, E. Kellner, B. Dhital, J. Hennig, and V. G. Kiselev, “Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach,” *NeuroImage*, vol. 147, pp. 964–975, Feb. 2017.
- [22] S. Coelho, S. H. Baete, G. Lemberskiy, B. Ades-Aron, G. Barrol, J. Veraart, D. S. Novikov, and E. Fieremans, “Reproducibility of the Standard Model of diffusion in white matter on clinical MRI systems,” *NeuroImage*, vol. 257, p. 119290, Aug. 2022.
- [23] D. S. Novikov, E. Fieremans, S. N. Jespersen, and V. G. Kiselev, “Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation,” *NMR in Biomedicine*, vol. 32, no. 4, p. e3998, 2019. e3998 nbm.3998.

- [24] I. O. Jelescu, M. Zurek, K. V. Winters, J. Veraart, A. Rajaratnam, N. S. Kim, J. S. Babb, T. M. Shepherd, D. S. Novikov, S. G. Kim, and E. Fieremans, “In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy,” *NeuroImage*, vol. 132, pp. 104–114, May 2016.
- [25] A. R. Khan, A. Chuhutin, O. Wiborg, C. D. Kroenke, J. R. Nyengaard, B. Hansen, and S. N. Jespersen, “Biophysical modeling of high field diffusion MRI demonstrates microstructural aberration in chronic mild stress rat brain,” *NeuroImage*, vol. 142, pp. 421–430, Nov. 2016.
- [26] A. R. Khan, A. Chuhutin, O. Wiborg, C. D. Kroenke, J. R. Nyengaard, B. Hansen, and S. N. Jespersen, “Summary of high field diffusion MRI and microscopy data demonstrate microstructural aberration in chronic mild stress rat brain,” *Data in Brief*, vol. 8, pp. 934–937, Sept. 2016.
- [27] P. Vestergaard-Poulsen, G. Wegener, B. Hansen, C. R. Bjarkam, S. J. Blackband, N. C. Nielsen, and S. N. Jespersen, “Diffusion-Weighted MRI and Quantitative Biophysical Modeling of Hippocampal Neurite Loss in Chronic Stress,” *PLOS ONE*, vol. 6, p. e20653, July 2011.
- [28] E. Fieremans, J. H. Jensen, J. A. Helpert, S. Kim, R. I. Grossman, M. Inglese, and D. S. Novikov, “Diffusion Distinguishes Between Axonal Loss and Demyelination in Brain White Matter,” *Proceedings of the International Society for Magnetic Resonance in Medicine*, vol. 20, p. 465, 2012.
- [29] D. S. Novikov and E. Fieremans, “Relating extracellular diffusivity to cell size distribution and packing density as applied to white matter,” *Proceedings of the International Society for Magnetic Resonance in Medicine*, vol. 20, p. 1829, 2012.
- [30] H. Gudbjartsson and S. Patz, “The Rician Distribution of Noisy MRI Data,” *Magnetic resonance in medicine*, vol. 34, pp. 910–914, Dec. 1995.
- [31] P. T. Callaghan, C. D. Eccles, and Y. Xia, “NMR microscopy of dynamic displacements: k-space and q-space imaging,” *Journal of Physics E: Scientific Instruments*, vol. 21, p. 820, Aug. 1988.
- [32] R. M. Henkelman, “Measurement of signal intensities in the presence of noise in MR images,” *Medical Physics*, vol. 12, no. 2, pp. 232–233, 1985.
- [33] J. Sijbers, A. den Dekker, P. Scheunders, and D. Van Dyck, “Maximum-likelihood estimation of Rician distribution parameters,” *IEEE Transactions on Medical Imaging*, vol. 17, pp. 357–361, June 1998.

- [34] O. Dietrich, J. G. Raya, S. B. Reeder, M. F. Reiser, and S. O. Schoenberg, “Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters,” *Journal of Magnetic Resonance Imaging*, vol. 26, no. 2, pp. 375–385, 2007.
- [35] W. A. Edelstein, G. H. Glover, C. J. Hardy, and R. W. Redington, “The intrinsic signal-to-noise ratio in nmr imaging,” *Magnetic Resonance in Medicine*, vol. 3, no. 4, pp. 604–618, 1986.
- [36] H. C. Torrey, “Bloch equations with diffusion terms,” *Physical Review*, vol. 104, no. 3, p. 563, 1956.
- [37] P. Mansfield, “Multi-planar image formation using nmr spin echoes,” *Journal of Physics C: Solid State Physics*, vol. 10, no. 3, p. L55, 1977.
- [38] R. Ordidge, “The development of echo-planar imaging (epi): 1977–1982,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 9, no. 3, pp. 117–121, 1999.
- [39] R. T. MK. Stehling and P. Mansfield, “Echo-planar imaging: Magnetic resonance imaging in a fraction of a second,” *Science*, vol. 254, pp. 43–50, 1991.
- [40] D. S. Novikov, “The Mesoscopic Physics of Diffusion MRI,” *Expanded lecture notes for ESMRM/ISMRM course: Diffusion: from acquisition to tissue microstructure*, 2018.
- [41] V. G. Kiselev, “Fundamentals of diffusion MRI physics,” *NMR in Biomedicine*, vol. 30, no. 3, p. e3602, 2017.
- [42] M. N. Berberan-Santos, “Expressing a probability density function in terms of another pdf: A generalized gram-charlier expansion,” *Journal of Mathematical Chemistry*, vol. 42, no. 3, pp. 585–594, 2007.
- [43] C. G. Koay, E. Özarslan, and P. J. Basser, “A signal-transformational framework for breaking the noise floor and its applications in MRI,” *Journal of Magnetic Resonance*, vol. 197, pp. 108–119, Apr. 2009.
- [44] J. Veraart, W. V. Hecke, and J. Sijbers, “Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model,” *Magnetic Resonance in Medicine*, vol. 66, no. 3, pp. 678–686, 2011.
- [45] J. Veraart, J. Rajan, R. R. Peeters, A. Leemans, S. Sunaert, and J. Sijbers, “Comprehensive framework for accurate diffusion MRI parameter estimation,” *Magnetic Resonance in Medicine*, vol. 70, pp. 972–984, Oct. 2013.



- [46] E. D. André, F. Grinberg, E. Farrher, I. I. Maximov, N. J. Shah, C. Meyer, M. Jaspar, V. Muto, C. Phillips, and E. Balteau, “Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging,” *PLOS ONE*, vol. 9, pp. 1–15, 04 2014.
- [47] K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl, “POAS4SPM: A Toolbox for SPM to Denoise Diffusion MRI Data,” *Neuroinformatics*, vol. 13, pp. 19–29, Jan. 2015.
- [48] D. S. Novikov, V. G. Kiselev, and S. N. Jespersen, “On modeling,” *Magnetic Resonance in Medicine*, vol. 79, no. 6, pp. 3172–3193, 2018.
- [49] D. C. Alexander, T. B. Dyrby, M. Nilsson, and H. Zhang, “Imaging brain microstructure with diffusion MRI: practicality and applications,” *NMR in Biomedicine*, vol. 32, no. 4, p. e3841, 2019.
- [50] C. M. W. Tax, F. Szczepankiewicz, M. Nilsson, and D. K. Jones, “The dot-compartment revealed? Diffusion MRI with ultra-strong gradients and spherical tensor encoding in the living human brain,” *NeuroImage*, vol. 210, p. 116534, Apr. 2020.
- [51] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [52] S. Sra and D. Karp, “The multivariate Watson distribution: Maximum-likelihood estimation and other aspects,” *Journal of Multivariate Analysis*, vol. 114, pp. 256–269, 2013.
- [53] S. Papazoglou, M. Ashtarayeh, J. M. Oeschger, M. F. Callaghan, M. D. Does, and S. Mohammadi, “Insights and improvements in correspondence between axonal volume fraction measured with diffusion-weighted mri and electron microscopy,” *NMR in Biomedicine*, vol. 37, no. 3, p. e5070, 2024.
- [54] G. David, B. Fricke, J. M. Oeschger, L. Ruthotto, P. Freund, K. Tabelow, and S. Mohammadi, “ACID: A Comprehensive Toolbox for Image Processing and Modeling of Brain, Spinal Cord, and Post-mortem Diffusion MRI Data,” *bioRxiv*, 2023.
- [55] J. M. Oeschger, K. Tabelow, and S. Mohammadi, “Axisymmetric diffusion kurtosis imaging with Rician bias correction: A simulation study,” *Magnetic Resonance in Medicine*, vol. 89, no. 2, pp. 787–799, 2022.
- [56] J. M. Oeschger, K. Tabelow, and S. Mohammadi, “Investigating apparent differences between standard dki and axisymmetric dki and its consequences for biophysical parameter estimates,” *Magnetic Resonance in Medicine*, vol. 92, no. 1, pp. 69–81, 2024.
- [57] J. Polzehl and K. Tabelow, “Low SNR in Diffusion MRI Models,” *Journal of the American Statistical Association*, vol. 111, no. 516, pp. 1480–1490, 2016.

- [58] J. D. Tournier, C.-H. Yeh, F. Calamante, K.-H. Cho, A. Connelly, and C.-P. Lin, “Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data,” *NeuroImage*, vol. 42, pp. 617–625, Aug. 2008.
- [59] E. Fieremans, J. H. Jensen, and J. A. Helpert, “White matter characterization with diffusional kurtosis imaging,” *NeuroImage*, vol. 58, no. 1, pp. 177–188, 2011.
- [60] C. F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis, “Processing and visualization for diffusion tensor MRI,” *Medical Image Analysis*, vol. 6, pp. 93–108, June 2002.
- [61] B. Hansen, T. E. Lund, R. Sangill, E. Stubbe, J. Finsterbusch, and S. N. Jespersen, “Experimental considerations for fast kurtosis imaging,” *Magnetic Resonance in Medicine*, vol. 76, pp. 1455–1468, Nov. 2016.
- [62] S. Becker, K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl, “Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by msPOAS,” *NeuroImage*, vol. 95, pp. 90 – 105, 2014.
- [63] S. Becker, K. Tabelow, S. Mohammadi, N. Weiskopf, and J. Polzehl, “Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by msPOAS,” *NeuroImage*, vol. 95, pp. 90–105, 2014.
- [64] B. Hansen and S. N. Jespersen, “Data for evaluation of fast kurtosis strategies, b-value optimization and exploration of diffusion MRI contrast,” *Scientific Data*, vol. 3, p. 160072, Aug. 2016.
- [65] K. L. West, N. D. Kelm, R. P. Carson, D. C. Alexander, D. F. Gochberg, and M. D. Does, “Experimental studies of g-ratio MRI in ex vivo mouse brain,” *NeuroImage*, vol. 167, pp. 366–371, Feb. 2018.
- [66] N. D. Kelm, K. L. West, R. P. Carson, D. F. Gochberg, K. C. Ess, and M. D. Does, “Evaluation of diffusion kurtosis imaging in ex vivo hypomyelinated mouse brains,” *NeuroImage*, vol. 124, pp. 612–626, Jan. 2016.
- [67] I. O. Jelescu, J. Veraart, V. Adisetiyo, S. S. Milla, D. S. Novikov, and E. Fieremans, “One diffusion acquisition and different white matter models: How does microstructure change in human early development based on WMTI and NODDI?,” *NeuroImage*, vol. 107, pp. 242–256, Feb. 2015.
- [68] N. Stikov, J. S. W. Campbell, T. Stroh, M. Lavelée, S. Frey, J. Novek, S. Nuara, M.-K. Ho, B. J. Bedell, R. F. Dougherty, I. R. Leppert, M. Boudreau, S. Narayanan, T. Duval, J. Cohen-Adad, P.-A. Picard, A. Gasecka, D. Côté, and G. B. Pike, “In vivo histology of the myelin g-ratio with magnetic resonance imaging,” *NeuroImage*, vol. 118, pp. 397–405, Sept. 2015.

[69] S. Mohammadi and M. F. Callaghan, "Towards in vivo g-ratio mapping using MRI: Unifying myelin and diffusion imaging," *Journal of Neuroscience Methods*, vol. 348, p. 108990, 2021.

## List of Figures

- 1 Graphic depicting the difference between accuracy and precision in diffusion weighted MRI. The bullseye and the red dots represent repeated measurements and estimation of a desired parameter. Shown are examples of: A) An axial view of a noise-free diffusion weighted MRI image as a reference, B) the same image shown in A) at SNR = 5 with low accuracy, i.e., a bias and low precision, C) the same image after Rician bias correction which improves accuracy by removing the bias but does not improve precision and D) the same image after using the adaptive denoising method "Multi-shell Position-Orientation Adaptive Smoothing" (msPOAS) which improves precision but does not remove the bias. . . . . 3
  
- 2 Shown is Figure 6 from<sup>55</sup>, showing the signal-to-noise ratio (SNR) above which the absolute value of the mean percentage error ("A-MPE") is < 5% for the synthetic dataset with high, medium and low fiber alignment ("HA", "MA", "LA") and the "in-vivo white matter" dataset. For the "in-vivo white matter" dataset, the A-MPE was averaged across the 12 simulated *in vivo* voxels and the SNR above which this average A-MPE < 5% is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve A-MPE < 5% across all five AxTM. . . . . 17
  
- 3 Shown is Figure S2 from<sup>55</sup>, showing the signal-to-noise ratio (SNR) above which the the standard deviation in reference to the ground truth ("R-STD") is < 5% for the synthetic dataset with high, medium and low fiber alignment ("HA", "MA", "LA") and the "in-vivo white matter" dataset. For the "in-vivo white matter" dataset, the R-STD was averaged across the 12 simulated *in vivo* voxels and the SNR above which this average R-STD < 5% is shown. The number above the barplots indicates the barplot's height. Blue encodes standard DKI, red encodes axisymmetric DKI, the hatched barplots show the results if RBC is used. "Maximum" shows the maximum SNR needed to achieve A-MPE < 5% across all five AxTM. . . . . 18

- 4 Shown is Figure 6 from<sup>56</sup>, showing the median total bias (axisymmetric DKI inherent bias and Rician bias) observed for noisy human brain data at an SNR = 39, smoothed with msPOAS in the five white matter fiber tracts corpus callosum (cc), superior corona radiata (scr), external capsule (exc), superior longitudinal fasciculus (slf) and posterior corona radiata (pcr). The median bias was computed for a simulation of the standard protocol that was fitted with standard DKI (red data points) and axisymmetric DKI (blue data points) and for a simulation of the fast "199" protocol that was fitted with axisymmetric DKI (pink data points). . . . . 21
- 5 Shown is Figure 3 from<sup>56</sup>, showing barplots summarizing the number of substantially differing voxels (SDV) of the axisymmetric DKI tensor Metrics (AxTM) (top) and biophysical parameters (bottom), estimated in a noise-free simulation of healthy white matter. Shown are the number of SDV (red barplots) and the median difference ("bias") in those voxels (blue barplots). The subplot's titles indicate the subset of voxels that were analyzed, i.e., "WM mask" = white matter voxels were analyzed, "WM mask & FA mask" = white matter voxels in the FA mask were analyzed and "WM mask & Westin mask" = white matter voxels in Westin mask were analyzed. Number of voxels in the white matter mask: 101521, number of voxels in the WM mask & FA mask analysis: 28741, number of voxels in the WM mask & Westin mask analysis: 20527. . . . . 23
- 6 Figure 1 from<sup>65</sup> showing electron microscopy from the mid-body of the corpus callosum (MCC) of the four different mouse models "Control", "Rictor CKO", "Tsc3 CKO" and "Pten CKO" simulated in this study (top), the corresponding g-ratios (see, e.g., Section 3.3.1) for each axon after segmentation (middle) and histograms of the g-ratio of all MCC axons corresponding to the top image. . . . 25
- 7 Results of biophysical model parameter estimation for simulation data based on the two groups of mice "1" and "2". The data points of each model were slightly shifted on the x-axis to make comparison easier. Shown is the average normalized root mean square error (NRMSE) and standard deviation across regions of interest (ROI) and the mice of each group for every simulated SNR. The groups of mice contained 12 (group 1) or 3 (group 2) mice . . . . . 26

- 8 Shown is Figure 3 from<sup>53</sup>, showing the contribution of calibration parameters to DWI model improvement where 1 and 2 are the two offsets adapted to two groups of mice with different levels of myelination and  $s$  is the scaling. Shown are the differences  $\Delta\text{BIC}$  with respect to the parameter combination with the smallest BIC in each DWI model. A lower  $\Delta\text{BIC}$  value indicates better model performance. Values for the uncalibrated case (i.e.,  $\text{AWF} \equiv \text{AVF}$ ) served as baseline, that is, for this case  $\Delta\text{BIC} = 0$ . An  $x$  indicates the calibration parameter combination with the largest evidence of improvement with respect to the baseline without further calibration for each DWI model, respectively. . . . . 30
- 9 Normalized root mean square error (NRMSE) of MRI based AVF estimates computed for the two mouse groups 1 and 2. Shown is the average normalized root mean square error (NRMSE) and standard deviation across regions of interest (ROI) and the mice of each group for every simulated SNR. The groups of mice contained 12 (group 1) or 3 (group 2) mice. Data points of each model were slightly shifted on the x-axis to make comparison easier. Note the different scaling of the y-axis of group 1 (left) and 2 (right). . . . . 31

### **Eidesstattliche Versicherung / Declaration on oath**

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Ort: Hamburg

Datum: 03.05.2024

Unterschrift: \_\_\_\_\_