NODDI-DTI: extracting neurite orientation and dispersion parameters from a diffusion tensor

Luke J. Edwards\textsuperscript{a,b,∗}, Kerrin J. Pine\textsuperscript{a,b}, Nikolaus Weiskopf\textsuperscript{a,b}, Siawoosh Mohammadi\textsuperscript{a,c}

\textsuperscript{a}Wellcome Trust Centre for Neuroimaging, UCL Institute of Neurology, UCL, 12 Queen Square, London, WC1N 3BG, UK.
\textsuperscript{b}Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany.
\textsuperscript{c}Institut für Systemische Neurowissenschaften, Universitätsklinikum Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany.

Abstract

Purpose: To allow efficient extraction of NODDI parameters from data intended for DTI analysis, permitting biophysical analysis of DTI datasets.

Theory: Simple relations between NODDI parameters (representing axon density, $\nu$, and dispersion, $\tau$) and DTI invariants (MD and FA) were derived through moment expansion of the NODDI signal model with no CSF compartment. NODDI-DTI uses these relations to extract NODDI parameters from DTI data. Diffusional kurtosis strongly biased MD estimates, thus a novel heuristic correction requiring only DTI data was derived and used.

Methods: NODDI-DTI parameter estimates using the first shell of data were compared to parameters extracted by fitting the NODDI model to (i) both shells (recommended) and (ii) the first shell (as for NODDI-DTI) of data in white matter of three different in vivo datasets, with CSF volume fraction fixed at zero.

Results: NODDI-DTI and one-shell NODDI parameter estimates gave similar errors compared to two-shell NODDI estimates. NODDI-DTI gave unphysical parameter estimates in a small percentage of voxels, reflecting voxelwise DTI estimation error or NODDI model invalidity.

Conclusion: NODDI-DTI is a promising technique to interpret restricted datasets acquired for DTI analysis biophysically, though its limitations must be borne in mind.

Keywords: diffusion MRI, NODDI, DTI, neurite density, neurite orientation dispersion

Word count: 4997
Number of figures: 7

*Corresponding author
Email address: ledwards@cbs.mpg.de (Luke J. Edwards)
Introduction

The white matter (WM) of the human brain consists of dense bundles of neuronal axons connecting its functional areas. Neural circuits thus formed allow these areas to work together as a coherent entity. Changes in WM impact these neural circuits, and are thus the subject of studies investigating pathology [1, 2, 3], and cognition and learning [4, 5].

Diffusion tensor imaging (DTI) [6] is, at present, the most commonly used method to observe WM changes in-vivo [4, 5, 7]. This is because DTI is simply implemented and time efficient while allowing robust extraction of complementary parameters (e.g. ‘fractional anisotropy’ (FA) and ‘mean diffusivity’ (MD) [8]) sensitive to microstructural WM changes [9], even in clinical contexts (see e.g. References [1, 3]). Despite its microstructural sensitivity, the underlying model of DTI, gaussian anisotropic diffusion [6], is unspecific to biological changes. Numerous studies show MD and FA change in white matter (e.g. due to learning a new skill [4] or the pathology of Alzheimer’s disease [2]), but cannot, in the absence of further information, distinguish e.g. changes in axon density from changes in axon arrangement.

In order to extract parameters of direct neurobiological relevance from diffusion MRI, we need biophysical models [10, 11, 9]. The majority of biophysical models (including the NODDI model investigated below) are ‘multicompartment models’. Such models assume voxelwise diffusion contrast arises from linear combination of diffusion signals from distinguishable water compartments. Numerous multicompartment models have been proposed (see e.g. References [12, 13, 14, 15, 16, 17, 18, 19, 20, 21]), but the complexity and lack of robustness of most of these models hinder their routine use in neuroscientific and clinical studies.

The NODDI (neurite orientation dispersion and density imaging) model [17] is a multicompartment model allowing robust and time-efficient extraction of maps of parameters representing neurite (in WM: axon) density and dispersion, and represents a trade-off between complexity, robustness, and acquisition-time duration. Robustness is achieved by fixing the values of several model parameters from earlier models [22, 18],...
reducing the number of fitted parameters. As a result, the amount of data required to invert the model is reduced, giving acquisition-time durations approaching those available in clinical settings [17]. NODDI is thus gaining popularity in diffusion application studies [23][24][15][25][26][20], though the potential of fixed parameters leading to bias in the fitted parameters has been a source of criticism [15][27][21].

Previous studies have shown correlations between NODDI parameters and MD and FA [17][28][25][29][30]. Below, we derive relations demonstrating their origin: in the absence of ‘free water’ (see below), MD and FA are formally sufficient to uniquely define parameters fitted in the NODDI model. Application of these relations to extract NODDI parameters from DTI parameters constitutes ‘NODDI-DTI’. Below, we demonstrate that NODDI-DTI gives reasonable estimates of NODDI parameters in WM, and characterise its limitations.

Theory

The NODDI signal model supposes three compartments: intraneurite water, extraneurite water, and free water [17]. The biophysical parameters fitted in the model are neurite density (volume fraction of the intraneurite compartment), \(\nu\); a measure of neurite dispersion, \(\kappa\); a vector giving the main neurite orientation; and a volume fraction accounting for partial-volume effects with free water (nominally CSF) [31][32][17]. An important fixed parameter is the intrinsic diffusivity of the intraneurite compartment, \(d = 1.7 \times 10^{-3} \text{ mm}^2 \text{ s}^{-1}[17]\). The primary neurite orientation [17] is formally equivalent to the principal eigenvector of the diffusion tensor (DT) (see Appendix A), as previously observed empirically [33].

We examine a reduced form of the NODDI model with no CSF volume fraction. In order to avoid overestimating CSF volume fraction during NODDI model fitting [15][34], we fix this parameter at zero in our NODDI fits. For ease of computation we use \(\tau\) as our measure of dispersion, where [15][17][18]

\[
\tau = \frac{1}{\sqrt{\pi \kappa}} \exp(-\kappa) \text{erf}(\sqrt{\kappa}) - \frac{1}{2\kappa}, \quad \tau \in [1/3, 1],
\]

(1)
and erfi is the imaginary error function. \( \tau \) ranges from 1/3 (isotropically distributed neurites) to 1 (perfectly aligned neurites)—increasing \( \tau \) corresponds to increasing neurite alignment—and is the average of \( \cos^2(\psi) \) over the neurite distribution, where \( \psi \) is the angle between a given neurite and the main neurite orientation [15].

By expanding the NODDI signal model in moments, one can derive a DT corresponding to the NODDI model [18]. As shown in Appendix A, appropriate combination of the eigenvalues of this DT allows expression of \( \nu \) and \( \tau \) in terms of MD and FA of this DT:

\[
\nu = 1 - \sqrt{\frac{1}{2} \left( \frac{3\text{MD} \cdot \text{FA}}{d} - 1 \right)},
\]

\[
\tau = \frac{1}{3} \left( 1 + \frac{4}{|d - \text{MD}|} \frac{\text{MD} \cdot \text{FA}}{\sqrt{3 - 2\text{FA}^2}} \right).
\]

These relations are exemplified in Figure 1.

Eqs. (2) and (3) demonstrate a one-to-one mapping from (MD,FA) to \((\nu, \tau)\). We can predict domains within which MD and FA should lie if the NODDI model provides a valid representation: substituting \( \nu \in [0, 1] \) and \( \tau \in [1/3, 1] \) into Eqs. (2) and (3) gives the domains

\[
\text{MD} \in [d/3, d], \ \text{FA} \in \left[ 0, \sqrt{\frac{3}{2}} \frac{|d - \text{MD}|}{\sqrt{2\text{MD}^2 + (d - \text{MD})^2}} \right].
\]

Exceptions will arise in practice either from DT measurement error or the invalidity of the NODDI model as a representation in a given voxel.

Estimation of a quantitative DT (i.e. the first moment of the diffusion signal) is non-trivial, as immediately apparent from examination of Figure 1: the main bulk of the MD values lies below the prediction of Eq. (2). A large part of this bias is due to neglecting higher order moments in estimating MD [35]. In order to reduce this bias, we define the heuristically corrected MD,

\[
\text{MD}_h = \text{MD} + \frac{1}{6} \left( \sum_{i,j=1}^{3} \frac{1 + 2\delta_{ij}}{15} \lambda_i \lambda_j \right),
\]

where \( \lambda_i \) is the \( i \)th eigenvalue of the measured DT and \( \delta_{ij} \) is the Kronecker delta. Eq. (5),
derived in Appendix B, pragmatically assumes that: only the first higher moment, diffusional kurtosis [36], contributes; the square of the apparent diffusion coefficient is uncorrelated with the apparent diffusional kurtosis; the mean diffusional kurtosis can be taken to be unity (approximately true over much healthy human brain WM [36, 37, 38]); and the effect of diffusional kurtosis on each individual eigenvalue is negligible. Figure 1 shows much closer agreement between the prediction of Eq. (2) and the experimental data when using MD$_h$; further justification of this correction is given below.

Substituting Eq. (5) into Eq. (2) gives the relation used in the following to compute $\nu$ from DT invariants:

$$\nu = 1 - \sqrt{\frac{3MD_h}{2d} - \frac{1}{2}} = 1 - \sqrt{\frac{3}{2d} \left( MD + \frac{b}{6} \left( \sum_{i,j=1}^{3} \frac{1 + 2\delta_{ij}}{15} \lambda_i \lambda_j \right) \right) - \frac{1}{2}}.$$  \hspace{1cm} (6)

The effect of failing to correct for diffusional kurtosis is much less pronounced for FA [35], and preliminary experiments (data not shown) showed that applying diffusional kurtosis correction to only MD in Eq. (3) resulted in a modest increase in the number of unphysical $\tau$ parameter estimates. This latter observation can be explained using Eq. (4): whenever heuristic diffusional kurtosis correction leads to overestimation of MD, the upper bound for allowed FA values is artificially decreased, potentially leading to unphysical $\tau$ estimates. We therefore apply no correction to Eq. (3).

**Methods**

*Data collection and preprocessing*

All data were collected by scanning healthy volunteers in a MAGNETOM Tim Trio 3 T MRI system (Siemens AG, Healthcare Sector, Erlangen, Germany) as part of a study approved by a local ethics committee. In each case informed written consent was obtained prior to scanning.

The first two datasets (subject 1 and subject 2) used a 2D multiband spin-echo echo-planar imaging (EPI) sequence supplied by the Center for Magnetic Resonance Research, University of Minnesota [39, 40, 41]. Sequence parameters: field of view (FoV):
220 × 220 mm², 81 slices, 1.7 mm isotropic resolution, echo time: TE = 112 ms, volume repetition time: TR = 4835 ms, partial Fourier factor: 6/8, multiband factor: 3, total 4 × 66 EPI images with 60 diffusion weighted images per shell using $b$-values of $b = \{1000, 2500\}$ s mm$^{-2}$, and 4 × 6 interleaved non-diffusion weighted ($b = 0$) images, 2 × phase encoding polarities (Anterior → Posterior / Posterior → Anterior).

The third dataset (subject 3) used a 2D spin-echo EPI sequence. Sequence parameters: FoV: 192 × 189 mm², 63 slices, 2.0 mm isotropic resolution, TE = 100 ms, TR = 11700 ms, parallel imaging factor: 2, phase encoding polarity (Anterior → Posterior). Additional parameters for high $b$-value shell: partial Fourier factor: 6/8, total 111 EPI images with 100 diffusion weighted images with $b = 2800$ s mm$^{-2}$ and 11 interleaved $b = 0$ images. Additional parameters for low $b$-value shell: no partial Fourier, total 110 EPI images with 100 diffusion weighted images with $b = 800$ s mm$^{-2}$ and 10 interleaved $b = 0$ images.

In all cases (subjects 1–3) subject motion, eddy currents, and susceptibility distortions were corrected for using the ACID toolbox [42]; for details see References [43, 37, 44, 45]. Additionally for the multiband data (subjects 1 and 2), after the above corrections were applied, the corrected data from the two phase encoding directions were summed for use in subsequent analysis. This final step was unnecessary for subject 3.

**Parameter estimation and comparison**

Parameters were estimated only in WM voxels determined to be unaffected by CSF or grey matter partial volume effects. This determination was made by thresholding at 50% probability a WM probability map obtained by segmenting the first $b = 0$ image of each respective dataset in SPM 12 [46].

The ACID toolbox was used to compute FA, MD, and the eigenvalues of the DT from the low $b$-value shell of each dataset, and home-written SPM scripts were then used to generate $\nu$ and $\tau$ using Eqs. (6) and (3), respectively, from these DT parameters. For subjects 1 and 2, the ACID toolbox was also used to simultaneously estimate the diffusion and kurtosis tensors [37], giving silver standard mean diffusivity estimates (MD$_{\text{DKI}}$) less biased by the effects of diffusional kurtosis [35] allowing evaluation of the validity of Eq. 5. MD$_{\text{DKI}}$ was not computed for subject 3 as the second $b$-value was deemed too large to
allow accurate estimation of the kurtosis tensor [36].

NODDI-silver standard results were obtained by fitting both shells of data using the NODDI toolbox [17, 47] with CSF volume fraction fixed at zero, followed by conversion of \( \kappa \) into \( \tau \) using a home-written SPM script implementing Eq. (1). We refer to these silver-standard results as the ‘two-shell NODDI’ results. In order to investigate the magnitude of the differences between NODDI and NODDI-DTI, fits were also made of the subset of the diffusion data used for the DTI fitting using the NODDI toolbox. The term ‘one-shell NODDI’ distinguishes these results from NODDI-DTI and silver-standard NODDI results. As the CSF volume fraction is fixed at zero, one-shell NODDI is well-posed [48].

Parameter estimate comparisons were quantified using means and standard deviations of the differences, visualised using Bland–Altman (BA) plots [49]. Voxels where NODDI-DTI gave unphysical parameter estimates (\( \nu \in [0, 1], \tau \in [1/3, 1] \)) were excluded from analysis.

**Results**

Before examining the results of NODDI-DTI, we examine the heuristic correction of MD for diffusional kurtosis (Eq. (5)) used to calculate \( \nu \). Figure 1 and the BA plots in Figure 2 show that the heuristically corrected MD, MD\(_h\), is less biased compared to the uncorrected MD, justifying use of the correction. Numerical values (mean ± one standard deviation) of the differences in Figure 2 are for subject 1: 0.117 ± 0.040 (MD), 0.025 ± 0.046 (MD\(_h\)); and for subject 2: 0.119 ± 0.041 (MD), 0.029 ± 0.041 (MD\(_h\)).

The similarity of the NODDI and NODDI-DTI results can be seen in Figure 3, which shows parameter maps computed with each method, along with maps showing the differences between the parameter estimates. The differences between the NODDI and NODDI-DTI results are further presented in several complementary ways: BA plots in Figure 4 show general behaviour, plots of the means and standard deviations of the differences in Figure 5 compare this general behaviour across subjects, and the series of slices in Figures 6 and 7 show the behaviour of NODDI-DTI estimates throughout the brain. Subjects 1 and 2, measured using the same protocol, show the best agreement, but all
three subjects behave similarly, demonstrating robustness of NODDI-DTI.

NODDI-DTI gave unphysical parameter estimates for some voxels: for $\nu$ such voxels constituted 2.07 %, 2.19 %, and 6.83 % of the total WM voxels for subjects 1–3 respectively; for $\tau$ such voxels constituted 0.27 %, 0.74 %, and 5.39 %. The proportion of unphysical parameter estimates was higher for $\nu$ than $\tau$, and greater for subject 3 than for subjects 1 and 2.

The magnitudes of the means and standard deviations of the differences between one- and two-shell NODDI are shown in Figure 5. One-shell NODDI gave stable fits in this case because the CSF compartment fraction was fixed at zero [48]. NODDI-DTI showed smaller mean differences and one-shell NODDI smaller standard deviations of the differences. Overall, however, both methods were comparable, further demonstrating the validity of NODDI-DTI.

**Discussion**

This work has demonstrated that, with caveats to be discussed below, parameters of potential neurobiological relevance from the NODDI signal model can be extracted from DTI parameters. The improved interpretability thus gained is not only applicable to future DTI studies, but also existing DTI studies.

As an example of using NODDI-DTI to reinterpret existing DTI studies, we apply the method superficially to the study of Scholz, et al. [4], which demonstrated a statistically significant FA increase in WM “underlying the intraparietal sulcus” after participants learned to juggle [4]. Assuming no concomitant change in MD (as suggested by the authors not reporting any significant change), this FA increase could be interpreted, using Eq. (3), as an increase in $\tau$, i.e. an increase in alignment of neuronal axons in this area with training. This result is much more specific than a change in FA, though we note that quantitative analysis would require reanalysis of the original data. A follow up study allowing for a full NODDI analysis, combined with proper mechanistic analysis of the WM plasticity mechanisms, would allow investigation of this effect in more detail.

Unfortunately, NODDI-DTI did not always give physically plausible parameter esti-
mates for the datasets studied herein. We posit four overlapping explanations for these unphysical estimates, related to NODDI-DTI assumptions.

Assumption 1: DTI parameters can be accurately estimated from the diffusion signal. Errors in DTI parameter estimation will lead to errors in parameters estimated using Eqs. (3) and (6), potentially giving unphysical parameter estimates. The $b$-values used likely resulted in overestimates of FA in regions of high anisotropy due to poor estimates of the low DT eigenvalues [8, 50], explaining why many of the unphysical $\tau$ estimates were found in the highly anisotropic corpus callosum (Figure 7). This hypothesis is bolstered by the results of subject 3: a lower $b$-value increased the proportion of unphysical estimates, especially in highly anisotropic regions.

Assumption 2: CSF can be ignored in voxels with a high probability of being WM. NODDI-DTI could give unphysical parameter estimates whenever a voxel contains a significant amount of CSF: the high diffusivity of CSF [17] can take MD$_b$ outside the limits of NODDI-DTI (Eqs. (4)). Figures 6 and 7 show that many of the voxels where NODDI-DTI gave unphysical parameter estimates are close to the edge of the WM mask, in line with partial volume effects being important. The larger voxel size used for scanning subject 3 means that partial volume effects are more prominent, partially explaining the poorer $\nu$ performance in this case. Because CSF volume fraction was fixed at zero in our NODDI fits, residual partial volume effects may also have affected our NODDI parameter estimates.

Assumption 3: MD can be heuristically corrected for diffusional kurtosis bias. Figure 1 shows that diffusional kurtosis affects our estimates of MD; such effects could take MD estimates out of the range of applicability for NODDI-DTI. While our heuristic correction (Eq. (5)) substantially mitigates this issue, it does not eliminate it. This is evident in the $\nu$ Bland–Altman plots (Figure 4), where the mean and standard deviation of the differences visibly vary with the mean $\nu$ estimate, implying (via Eq. (6)) residual correlation between the errors and the corrected MD.

Residual bias in the region around the corpus callosum of the $\nu$ difference maps in Figure 6 can be explained by incomplete diffusional kurtosis correction. Here the mean
diffusional kurtosis is greater than unity \[36\], and the low DT eigenvalues are poorly estimated (see above), meaning the heuristic correction is not entirely valid.

Assumption 4: the NODDI model is a valid representation of the diffusion signal. Assumptions underlying the NODDI signal model can be considered overly restrictive: the model cannot formally represent WM voxels containing perpendicularly crossing fibre bundles \[51\], it is controversial whether intra- and extraneurite intrinsic diffusivities can be taken to be equal \[27\], and the extraneurite signal averaging is unrealistic \[21\]. NODDI will also be unrepresentative whenever a voxel contains a significant amount of iron (e.g. through partial voluming with iron rich grey matter nuclei); susceptibility effects artificially lower diffusivities \[52\], and so lower \(d\), assumed to be a fixed value in NODDI. In voxels where the NODDI signal model is unrepresentative, NODDI-DTI may give unphysical parameter estimates. Such failures will not be immediately apparent in the NODDI fitted parameters because constraints in the fitting procedure mean parameters outside the physical range can never be returned, regardless of the model’s biological plausibility in a given voxel.

The greater number of unphysical NODDI-DTI parameter estimates for \(\nu\) as compared to \(\tau\) can be explained by \(\nu\) estimation being more sensitive to partial volume and diffusional kurtosis effects. This is borne out by the locations of the failures (Figure 6): mainly either close to the edge of the WM mask (implying partial volume effects), or in regions of high anisotropy (implying residual diffusional kurtosis effects).

Pathology could further undermine the assumptions underlying NODDI-DTI: pathological processes can lead to free water located far from CSF compartments \[53\], can affect mean kurtosis values \[54\], and could affect the ‘true’ value of \(d\) \[15\].

NODDI-DTI could be improved and made more appropriate for clinical studies through investigation of the following points. Unphysical parameter estimates could be pragmatically eliminated by constraining DT fitting using Eqs. (4) (appropriately corrected using Eq. (5)). Estimates of CSF volume fraction could be incorporated into NODDI-DTI using the free water elimination method \[53, 32, 55\]. Known values of mean diffusional kurtosis in the brain \[36, 37, 38\] could be used to construct mean diffusional kurtosis Bayesian
priors \[56, 57\], or diffusional kurtosis corrected MD and FA could be measured directly using time-efficient methods \[58, 59, 60\].

We finish by providing practical recommendations for a minimal NODDI-DTI acquisition scheme. Results at \(b = 1000 \text{ mm}^2 \text{s}^{-1}\) were reasonable, and we would recommend this as a lower \(b\)-value bound; \(b = 800 \text{ mm}^2 \text{s}^{-1}\) unfortunately gave many unphysical parameter estimates in the corpus callosum. An upper bound on \(b\)-value comes from ensuring diffusional kurtosis does not constitute the majority of the diffusion contrast. Eq. (B.3) shows that (assuming the apparent diffusional kurtosis is approximately unity \[36, 37, 38\]) choosing \(b \ll 6/d \approx 3500 \text{ mm}^2 \text{s}^{-1}\) means that the DT dominates diffusion contrast, giving an upper \(b\)-value bound. High resolution acquisitions which maintain good signal to noise ratio \[50\] are recommended to reduce partial volume effects. Accurate DT estimation requires measurement of at least 30 distinct diffusion directions \[6\]; we recommend at least this number for application of NODDI-DTI, although the lowest number of orientations tested was 60.

Conclusions

We have estimated NODDI biophysical parameters representing neurite density and dispersion with reasonable accuracy from diffusion tensor parameters extracted from single-shell diffusion data. Heuristic kurtosis correction of MD was necessary to remove diffusional kurtosis bias; use of corrections such as that derived here could improve other analyses of single-shell diffusion data requiring quantitative MD estimates.

NODDI-DTI opens up two new opportunities: (a) more direct neurobiological interpretation of observed microstructural changes in DTI data (including interpretation of existing datasets), and (b) simple and time efficient estimation of biophysical parameters from smaller diffusion datasets, despite limitations due to the underlying NODDI model and difficulties estimating accurate diffusion tensors.
Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013) / ERC grant agreement n° 616905. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 658589. The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust 0915/Z/10/Z.

Appendix A. Derivation of NODDI-DTI relations (Eqs. (2) and (3))

Appendix A.1. Diffusion tensor of the NODDI signal model

We begin derivation of the NODDI-DTI relations by deriving the DT arising from the NODDI signal model. This derivation is similar to that of Reference [18], where the DT of a precursor to the NODDI model [22] was derived.

The normalised signal arising from the NODDI signal model can be written [17]

\[ S = \nu \int p(\kappa, \mu, \vec{n}) \exp\{-bd\vec{q}^t\vec{n}\vec{n}^t\vec{q}\} \, d\vec{n} + (1 - \nu) \exp\{-b\vec{q}^tD_{ec}\vec{q}\} \]  
(A.1)

where the first term represents the intraneurite water compartment with diffusivity \( d \) parallel to the neurite and zero perpendicular to it; the second term represents the extraneurite water compartment; arrows denote normalised vectors; \( \cdot^t \) denotes transposition; \( \vec{q} \) is the diffusion gradient vector; \( \nu \) represents neurite density; and

\[ D_{ec} = d \int p(\kappa, \mu, \vec{n}) \left( \vec{n}\vec{n}^t + (1 - \nu)(1 - \vec{n}\vec{n}^t) \right) \, d\vec{n}, \]  
(A.2)

the DT of the extraneurite compartment. The form of the extraneurite DT arises from assuming that: the diffusivity of the extraneurite space in the absence of neurites is equal to the intraneurite diffusivity along the direction of the neurite [17], the neurites reduce the diffusivity in a long-time-limit tortuous manner [17], and extracellular water is in fast
exchange among all neurite orientations [21]. The probability density
\[ p(\kappa, \vec{\mu}, \vec{n}) = \frac{\exp\{-\kappa(\vec{\mu}^t\vec{n})^2\}}{\int \exp\{-\kappa(\vec{\mu}^t\vec{n})^2\} \, d\vec{n}} \]  
(A.3)
is a Watson distribution giving the distribution of neurites about main orientation \( \vec{\mu} \) with dispersion parameter \( \kappa \) [17]. Isotropically distributed neurites correspond to \( \kappa = 0 \), neurites perfectly aligned along \( \vec{\mu} \) correspond to \( \kappa \to \infty \).

Eq. (A.1) can be equated with an expansion of the normalised diffusion signal in \( b \) [61],
\[ \log(S) = -b\vec{q}^t D_{\text{NODDI}} \vec{q} + O(b^2), \]  
(A.4)
such that the DT can be extracted by inspection from
\[ \frac{d \log(S)}{db} \bigg|_{b=0} = -\nu d \int p(\kappa, \vec{\mu}, \vec{n}) \vec{q}^t \vec{n} \vec{n}^t \vec{q} \, d\vec{n} - (1 - \nu) \vec{q}^t D_{\text{ec}} \vec{q} \]  
(A.5)
\[ = -\vec{q}^t \left[ d \int p(\kappa, \vec{\mu}, \vec{n}) (\vec{n} \vec{n}^t + (1 - \nu)^2 (1 - \vec{n} \vec{n}^t)) \, d\vec{n} \right] \vec{q}, \]  
(A.6)
as
\[ D_{\text{NODDI}} = d \int p(\kappa, \vec{\mu}, \vec{n}) (\vec{n} \vec{n}^t + (1 - \nu)^2 (1 - \vec{n} \vec{n}^t)) \, d\vec{n} \]  
(A.7)
\[ = (1 - \nu)^2 d + (2 - \nu) \nu d \int p(\kappa, \vec{\mu}, \vec{n}) \vec{n} \vec{n}^t \, d\vec{n}. \]  
(A.8)

The integral appearing on the right-hand side of Eq. (A.8) is given by [18]
\[ \int p(\kappa, \vec{\mu}, \vec{n}) \vec{n} \vec{n}^t \, d\vec{n} = \tau \vec{\mu} \vec{\mu}^t + \frac{(1 - \tau)}{2} (1 - \vec{\mu} \vec{\mu}^t), \]  
(A.9)
where \( \tau \) is defined in Eq. (1). Inserting Eq. (A.9) into Eq. (A.8) gives
\[ D_{\text{NODDI}} = ((1 - \nu)^2 d + (2 - \nu) \nu \tau d) \vec{\mu} \vec{\mu}^t + ((1 - \nu)^2 d + (2 - \nu) \nu \frac{(1 - \tau)}{2} d (1 - \vec{\mu} \vec{\mu}^t), \]  
(A.10)
from which, by inspection, the largest eigenvalue (corresponding to an eigenvector co-
linear with the main neurite orientation) is

\[ \lambda_1 = \mu^T D_{\text{NODDI}} \mu = (2 - \nu) \nu \tau d + (1 - \nu)^2 d, \]  

(A.11)

and the other two eigenvalues are degenerate (with respective eigenvectors arbitrarily defined in the plane perpendicular to \( \mu \)):

\[ \lambda_2 = \lambda_3 = (2 - \nu) \nu \frac{(1 - \tau)}{2} d + (1 - \nu)^2 d. \]  

(A.12)

Because \( \lambda_1 \geq \lambda_2 \), when the primary eigenvector of \( D_{\text{NODDI}} \) is well-defined (i.e. when \( D_{\text{NODDI}} \) is not isotropic), this eigenvector is formally equivalent to the main neurite orientation, as observed empirically [33].

Appendix A.2. Relation of \( \nu \) to MD (Eq. (2))

MD is defined in terms of the eigenvalues of a DT as [6]

\[ \text{MD} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3}. \]  

(A.13)

Inserting the eigenvalues from Eqs. (A.11) and (A.12) results in

\[ \frac{3\text{MD}}{d} = (2 - \nu) \nu + 3(1 - \nu)^2. \]  

(A.14)

Solving this quadratic equation for \( \nu \) one obtains

\[ \nu = 1 \pm \sqrt{\frac{1}{2} \left( \frac{3\text{MD}}{d} - 1 \right)}, \]  

(A.15)

where the sign ambiguity is resolved by recalling \( \nu \leq 1 \) to give Eq. (2).

Appendix A.3. Relation of \( \tau \) to MD and FA (Eq. (3))

A convenient definition of FA in terms of the eigenvalues of a DT is [6]:

\[ \text{FA} = \sqrt{\frac{3}{2} \sqrt{\left( \frac{(\lambda_1 - \text{MD})^2 + (\lambda_2 - \text{MD})^2 + (\lambda_3 - \text{MD})^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \right)^2}}. \]  

(A.16)
Because the eigenvalues are linear functions of $\tau$ (Eqs. (A.11) and (A.12)) and there is symmetry between them, it is convenient to simplify this equation by solving for $\lambda_2$ before proceeding further. Utilising the identities $\lambda_1 = 3MD - \lambda_2 - \lambda_3$ (Eq. (A.13)) and $\lambda_2 = \lambda_3$ (Eq. (A.12)), Eq. (A.16) becomes:

$$FA^2 = \frac{9(\lambda_2 - MD)^2}{(3MD - 2\lambda_2)^2 + 2\lambda_2^2},$$

which can be rearranged into the quadratic equation

$$\lambda_2^2 - 2MD\lambda_2 + 3MD^2 \frac{1 - FA^2}{3 - 2FA^2} = 0 \quad (A.18)$$

for which the solutions are:

$$\lambda_2 = MD \left( 1 \pm \frac{FA}{\sqrt{3} - 2FA^2} \right). \quad (A.19)$$

Eq. (A.19), reveals that all we must do to express $\tau$ in terms of MD and FA is (i) express $\tau$ in terms of MD and $\lambda_2$ and then (ii) substitute Eq. (A.19) into the resulting expression.

Part (i) is achieved by substituting Eq. (2) into Eq. (A.12), then simplifying to give

$$\lambda_2 = \frac{1}{4}(d + 3MD - 3(d - MD)\tau), \quad (A.20)$$

which, after rearranging for $\tau$, reveals

$$\tau = \frac{1}{3} \left( \frac{1}{d - MD}(d + 3MD - 4\lambda_2) \right). \quad (A.21)$$

We can now perform part (ii): inserting Eq. (A.19) into Eq. (A.21) and simplifying gives the result

$$\tau = \frac{1}{3} \left( 1 \mp \frac{4}{d - MD} \frac{MD \cdot FA}{\sqrt{3} - 2FA^2} \right). \quad (A.22)$$

The sign ambiguity is resolved by recalling that $\tau \geq 1/3$ and that both MD and FA are nonnegative, resulting in Eq. (3).
Eq. (3) is ill-defined at MD = d (the denominator of the second term goes to zero); we classify values at this point as ‘unphysical’ unless FA is also zero. This latter situation corresponds to the complete absence of fibres (as confirmed by inserting MD = d into Eq. (2)), and so τ is taken to equal its isotropic value, 1/3.

Appendix B. Heuristic correction of MD for diffusional kurtosis (Eq. (5))

When diffusional kurtosis and higher order moments are zero, the normalised diffusion signal $S$ is related to the apparent diffusivity, $D_{app}$, by [62]

$$\log(S) = -bD_{app}, \quad (B.1)$$

giving

$$\text{MD} = \langle D_{app} \rangle = \frac{\langle \log(S) \rangle}{-b}, \quad (B.2)$$

where $\langle \cdot \rangle$ denotes averaging over all diffusion directions.

The complicated microstructure of white matter requires higher order moments to represent the diffusion signal [61, 36, 35]. To the order of the diffusional kurtosis the normalised diffusion signal is

$$\log(S) = -bD_{app} + \frac{b^2}{6} D_{app}^2 K_{app}, \quad (B.3)$$

where $K_{app}$ is the apparent diffusional kurtosis [61]. The effective mean diffusivity $\text{MD}_{eff}$ derived from this signal (as per Eq. (B.2)) would be

$$\text{MD}_{eff} = \frac{\langle \log(S) \rangle}{-b} = \text{MD} - \frac{b}{6} \langle D_{app}^2 K_{app} \rangle, \quad (B.4)$$

which differs from the true MD by a term which we call ‘diffusional kurtosis bias’. While good estimates of unbiased MD can be obtained from multi-$b$-value data [35], such extra data is not available for most DTI acquisitions, and so we derive and use an heuristic correction to mitigate diffusional kurtosis bias.
Defining the covariance of $D_{\text{app}}^2$ and $K_{\text{app}}$:

$$\text{cov} \left( D_{\text{app}}^2, K_{\text{app}} \right) = \langle (D_{\text{app}}^2 - \langle D_{\text{app}}^2 \rangle)(K_{\text{app}} - MK) \rangle$$  \hspace{1cm} (B.5)

$$= \langle D_{\text{app}}^2 K_{\text{app}} \rangle - \langle D_{\text{app}}^2 \rangle MK,$$  \hspace{1cm} (B.6)

where $MK = \langle K_{\text{app}} \rangle$, the mean kurtosis, we can write Eq. (B.4) in the form

$$MD_{\text{eff}} = MD - b \frac{6}{6} (\langle D_{\text{app}}^2 \rangle MK + \text{cov} \left( D_{\text{app}}^2, K_{\text{app}} \right)).$$  \hspace{1cm} (B.7)

We pragmatically assert that $\text{cov} (D_{\text{app}}^2, K_{\text{app}}) = 0$, i.e. assume that the apparent diffusivity and apparent diffusional kurtosis are uncorrelated. This assertion results in

$$MD_{\text{eff}} \approx MD - b \frac{6}{6} \langle D_{\text{app}}^2 \rangle MK,$$  \hspace{1cm} (B.8)

which is further simplified by assuming $MK = 1$ (true in much healthy WM [36, 37, 38]):

$$MD_{\text{eff}} \approx MD - b \frac{6}{6} \langle D_{\text{app}}^2 \rangle.$$  \hspace{1cm} (B.9)

To compute the average in Eq. (B.9), we express $D_{\text{app}}$ in components of the DT, $D$, and orientation vector, $\vec{q}$, i.e. [36]

$$\langle D_{\text{app}}^2 \rangle = \sum_{i,j,k,l=1}^{3} \langle q_i q_j q_k q_l D_{ij} D_{kl} \rangle.$$  \hspace{1cm} (B.10)

As we integrate over all $\vec{q}$ on the sphere, we can freely choose the basis of $\vec{q}$. We thus choose the diagonal basis of $D$, simplifying Eq. (B.10) to:

$$\langle D_{\text{app}}^2 \rangle = \sum_{i,k=1}^{3} \langle q_i^2 q_k^2 \lambda_i \lambda_k \rangle = \sum_{i,k=1}^{3} \langle q_i^2 q_k^2 \rangle \lambda_i \lambda_k,$$  \hspace{1cm} (B.11)

where $\lambda_i$ is the $i$th eigenvalue of $D$ and is independent of orientation. Averages over the products of the components $q_i$, evaluate to $\langle q_i^2 q_k^2 \rangle = (1 + 2\delta_{ik})/15$, where $\delta_{ik}$ is the
Kronecker delta, thus
\[
\langle D_{\text{app}}^2 \rangle = \sum_{i,k=1}^{3} \frac{1 + 2\delta_{ik}}{15} \lambda_i \lambda_k.
\]  
(B.12)

Inserting Eq. (B.12) into Eq. (B.9) and rearranging gives the heuristically corrected MD:
\[
\text{MD} \approx \text{MD}_h = \text{MD}_{\text{eff}} + \frac{b}{6} \left( \sum_{i,k=1}^{3} \frac{1 + 2\delta_{ik}}{15} \lambda_i \lambda_k \right),
\]  
(B.13)

which becomes independent of diffusional kurtosis upon assuming the measured (diffusional kurtosis biased) eigenvalues can be substituted for the ‘true’ eigenvalues, giving Eq. (5).

**References**


[42] ACID toolbox for SPM
URL http://www.diffusiontools.com/


[46] SPM 12
URL http://www.fil.ion.ucl.ac.uk/spm/

[47] NODDI toolbox v0.9
URL http://www.nitrc.org/projects/noddi_toolbox/


Figure 1: Log density scatter plots comparing DTI invariants (computed using only the low $b$-value shell of data) and NODDI parameters (fitted using both shells of data). Each row shows a different subject as labelled. Overlaid red lines in the first and second columns show values of $\nu$ for given values of MD and MD$_h$, respectively, computed using Eq. (2). The overlaid red lines in the third column show $\tau$ computed using Eq. (3) for given FA, with MD set to the mean value in the WM of each subject.
Figure 2: Log density Bland–Altman plots comparing MD_{DKI} computed via simultaneous fit of the kurtosis tensor and DT using both shells of data, and mean diffusivity from a DT fit of the low-\(b\)-value shell without (MD, left) and with (MD_{h}, right) heuristic diffusional kurtosis correction. Simultaneous fit of the kurtosis tensor and DT was performed as per Reference 37. Differences are defined as MD_{DKI} − (MD or MD_{h}). Each row shows a different subject as labelled; results were not computed for subject 3 as the second \(b\)-value was deemed too large to allow accurate estimation of the kurtosis tensor 36. Red lines show mean difference, blue lines show ± two standard deviations of the difference.
Figure 3: Comparison of maps of parameters computed using NODDI-DTI and two-shell NODDI. Voxels where NODDI-DTI gave an unphysical parameter estimate are shown in blue. Windows are as per the limits of the colour scales beside each map, and slice number is given at the top left of the row for each subject, to allow for cross-referencing with Figures 6 and 7.
Figure 4: Bland–Altman plots comparing NODDI-DTI and two-shell NODDI results. Plotted is the log-density, and differences are defined as (two-shell NODDI parameter) – (NODDI-DTI parameter). Red lines show mean difference, blue lines show ± two standard deviations of the difference; the numerical values of the means and standard deviations of the differences are given in Figure 5. Axis ranges show bounds of means and differences in each case.
Figure 5: Plots of the mean differences between NODDI-DTI and two-shell NODDI (×), and between one-shell NODDI (using the NODDI toolbox to fit the shell of data used by NODDI-DTI) and two-shell NODDI (○) parameter estimates. Error bars show ± one standard deviation of the differences. Differences are defined as (two-shell NODDI parameter) – (estimated parameter). Numerical mean values are given beside each plotted point, and numerical values for the standard deviations are given beside the upper error bar.
Figure 6: Absolute value maps of the difference between $\nu$ computed using NODDI-DTI and two-shell NODDI. Data from all three subjects are shown, and slice numbers are given for each row (slice) and column (subject). The extent of the colour scale at the top right shows the windowing for all slices. Blue denotes voxels where NODDI-DTI gave an unphysical parameter estimate.
Figure 7: Absolute value maps of the difference between $\tau$ computed using NODDI-DTI and two-shell NODDI. Data from all three subjects are shown, and slice numbers are given for each row (slice) and column (subject). The extent of the colour scale at the top right shows the windowing for all slices. Blue denotes voxels where NODDI-DTI gave an unphysical parameter estimate. The large number of unphysical $\tau$ values in the corpus callosum of subject 3 is discussed in the text.