

Characterizing Brain Tissue by Assessment of the Distribution of Anisotropic Microstructural Environments in Diffusion-Compartment Imaging (DIAMOND)

Benoit Scherrer,^{1*} Armin Schwartzman,² Maxime Taquet,¹ Mustafa Sahin,¹ Sanjay P. Prabhu,¹ and Simon K. Warfield¹

Purpose: To develop a statistical model for the tridimensional diffusion MRI signal at each voxel that describes the signal arising from each tissue compartment in each voxel.

Theory and Methods: In prior work, a statistical model of the apparent diffusion coefficient was shown to well-characterize the diffusivity and heterogeneity of the mono-directional diffusion MRI signal. However, this model was unable to characterize the three-dimensional anisotropic diffusion observed in the brain. We introduce a new model that extends the statistical distribution representation to be fully tridimensional, in which apparent diffusion coefficients are extended to be diffusion tensors. The set of compartments present at a voxel is modeled by a finite sum of unimodal continuous distributions of diffusion tensors. Each distribution provides measures of each compartment microstructural diffusivity and heterogeneity.

Results: The ability to estimate the tridimensional diffusivity and heterogeneity of multiple fascicles and of free diffusion is demonstrated.

Conclusion: Our novel tissue model allows for the characterization of the intra-voxel orientational heterogeneity, a prerequisite for accurate tractography while also characterizing the overall tridimensional diffusivity and heterogeneity of each tissue compartment. The model parameters can be estimated from short duration acquisitions. The diffusivity and heterogeneity microstructural parameters may provide novel indicator of the presence of disease or injury. *Magn Reson Med* 000:000–000, 2015. © 2015 Wiley Periodicals, Inc.

Key words: diffusion-weighted MRI; tissue microstructure; statistical distribution model; non-monoexponential decay; diffusion compartment imaging

INTRODUCTION

Diffusion-weighted imaging (DWI) enables non-invasive investigation of the brain microstructure by probing natural barriers to diffusion in tissues. As the DWI spatial resolution is typically on the order of 6–27 mm³, the measured DW signal in each voxel combines the signal arising from a variety of microstructural environments including multiple cell types, sizes, geometries and orientations, and extra-cellular space. This is well-known to give rise to an overall non-monoexponential decay in each voxel (1–4).

Multiple models have been proposed to account for the observed non-monoexponential decay, including fitting a “stretched-exponential model” (5) or a Kurtosis term (6). These models attempt to describe the signal arising from each entire voxel and focus on capturing the mathematical deviation from the monoexponential decay. However, they conflate the signal arising from multiple compartments and do not directly reflect the tissue microstructure.

A primary alternative is to focus on describing the MR signal formation with a model whose parameters reflect the tissue compartments present in each voxel. These so-called *diffusion compartment models* are based on underlying observations of the biological microstructure that alter the diffusion signal and are of great interest to characterize and compare tissue properties. For example, assessment of the isotropic diffusion arising from the extracellular space is useful to account for cerebrospinal fluid due to partial volume effect (7) and for the characterization of vasogenic edema, inflammation, and neurodegeneration (8,9). A model of the axonal dispersion is useful to capture information about the dendrite density (10,11). A model for each individual fascicle is useful to characterize the fascicle density (12), the axonal diameter distribution (12,13) or the myelin integrity. We propose to refer to the various diffusion compartments models as “diffusion compartment imaging” (DCI) techniques. Among DCI techniques, CHARMED (14) represents the intra-axonal diffusion by the analytic expression for diffusion within impermeable cylinders and the extracellular hindered diffusion with a full tensor. However, it was shown that *b*-values up to 10,000 s/mm² are necessary to reliably estimate the parameters of this parameterization (14), which is not achievable in routine clinical practice. NODDI (11) focuses on explicitly modeling the fascicle dispersion with a Watson distribution of sticks in each voxel. However, this is achieved at the cost of key assumptions that are inconsistent with the known tissue microstructure. The NODDI model relies on a prefixed representation of a

¹Department of Radiology, Boston Children’s Hospital and Harvard Medical School, 300 Longwood Avenue, Boston, Massachusetts, USA

²Department of Statistics, North Carolina State University, Raleigh, North Carolina, USA

Grant sponsor: NIH; Grant numbers: U01 NS082320, R01 NS079788, R01 EB013248, R01 EB019483, R21 EB012177, U54NS092090; Grant sponsor: DOD; Grant number: W81XWH-11-1-0365; Grant sponsors: BCH; Grant references: BCH TRP Pilot, BCH CTREC K-to-R Merit Award; Grant sponsors: Intel(c) Corporation; Grant reference: Intel(c) IPCC; Grant sponsor: Developmental Synaptopathies Consortium; Grant number: U54NS092090. The Developmental Synaptopathies Consortium is a part of the NCATS Rare Diseases Clinical Research Network (RDCRN). RDCRN is an initiative of the Office of Rare Diseases Research (ORDR), NCATS, funded through collaboration between NCATS, NIMH, NINDS and NICHD. Maxime Taquet was supported by F.R.S.-FNRS and WBI.WORLD.

*Correspondence to: Benoit Scherrer, Ph.D., Boston Children’s Hospital, 300 Longwood Avenue, Boston, MA 02115. E-mail: benoit.scherrer@childrens.harvard.edu, Twitter: @benoitscherrer

Received 8 September 2014; revised 17 July 2015; accepted 11 August 2015

DOI 10.1002/mrm.25912

Published online 00 Month 2015 in Wiley Online Library (wileyonlinelibrary.com).

© 2015 Wiley Periodicals, Inc.

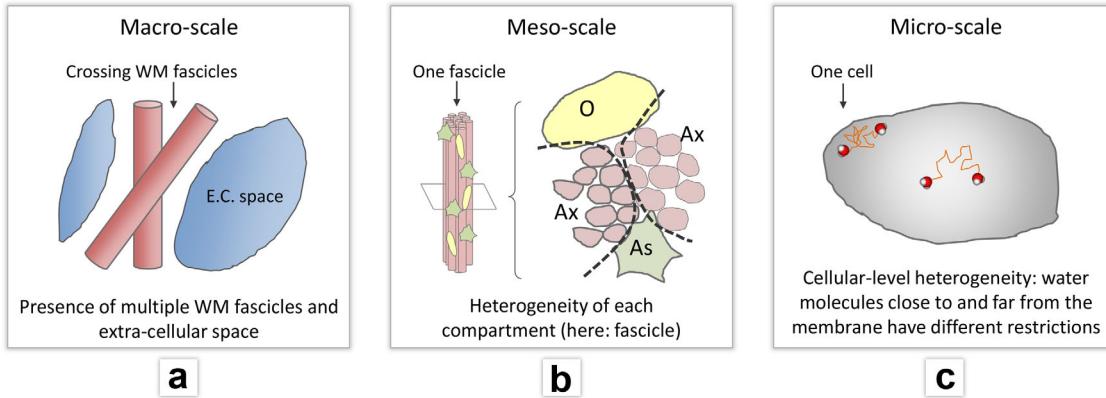


FIG. 1. Multiple scales of intra-voxel heterogeneity are responsible for the observed non-monoexponential decay. **a:** Large-scale heterogeneity includes the mixing of large-scale microstructural environments, also referred to as tissue compartments, such as the mixing of multiple WM (WM) fascicles with extra-cellular space. It also originates from partial volume averaging such as occurs when cerebrospinal fluid, gray matter and/or WM are present in a voxel. **b:** Each large-scale microstructural environment may contain a complex varying microstructure including axons with varying radii and degrees of myelination, or fascicles with a varying density of glial cells such as astrocytes and oligodendrocytes. It is likely that water molecules *interacting with an homogeneous portion of this structure* are well modeled by an exponential. However, the overall signal arising from the compartment may significantly deviate from a mono-exponential decay in the presence of heterogeneity of the spin packets composing the compartment. **c:** At an even smaller scale, other biophysical mechanisms such as intracellular heterogeneities and the proximity of cell membranes may locally restrict water molecule motion and contribute to the signal decay behavior. For example, Sehy et al. (4) observed a non-monoexponential decay within the intracellular space of a *single* cell, the frog oocyte. (Ax: axons with various degrees of myelination; As: Astrocyte; O: Oligodendrocyte).

white matter (WM) fascicle constant throughout the entire brain, while fascicles with various microstructures have been observed such as in the body of the corpus callosum (12,15). NODDI ignores the intra-axonal radial diffusivity and, similarly to diffusion tensor imaging (DTI) (16), consider only a single fascicle compartment per voxel while fascicles crossing with angle $>40^\circ$ occurs in 60 to 90% of the voxels in the human brain (17). NODDI may capture crossing fascicles as increased dispersion but cannot characterize each of them separately. More generally, various parametric models have been proposed in DCI. In their taxonomy of diffusion models, (18,19) ranked different combinations of parameterizations for each compartment when fitting a single fascicle region. They showed that the best performing models included three compartments to, respectively, represent freely diffusing water; water molecules restricted to the intra-axonal space; and water molecules hindered by axonal membranes. However, the best parameterization to describe the signal arising from each tissue compartment remains an open problem.

A solution may lie in considering a more detailed model of the tissue microstructure that gives rise to the signal attenuation. Particularly, the observed non-monoexponential decay in voxels likely arises from both *large-scale* and *small-scale* intra-voxel heterogeneities (Fig. 1). Yablonskiy et al. (20) proposed a statistical model of the apparent diffusion coefficient (ADC) that intrinsically reflects the presence of heterogeneous microstructural environments in each voxel. They considered measurements arising from the large number of spin packets in each voxel such that each homogeneous spin packet undergoes local isotropic Gaussian diffusion described by an ADC D and proposed to model the DW signal by integration of the contribution of all spin packets:

$$S_k = S_0 \int P(D) \exp(-b_k D) dD, \quad [1]$$

where S_k is the measured diffusion signal for the b -value b_k , S_0 is the signal with no diffusion applied and $P(D)$ is a probability distribution that describes the fraction of spin packets with an ADC D in the voxel (Fig. 2a–c). This statistical model of the ADC reflects, via the shape of the scalar-valued distribution $P(D)$, the overall diffusivity and heterogeneity in each voxel. Originally formulated as a mono-directional model, it was extended to the multi-directional case by estimating one ADC per direction. This model, however, cannot characterize the three-dimensional (3D) anisotropy of diffusion observed in the brain (21) and cannot describe the spin packets' distribution of oriented 3D large-scale microstructural environments such as WM fascicles. Jbabdi et al. (22) proposed to incorporate a 1D Gamma distribution of diffusivities in the ball-and-stick model. However, a Gamma distribution with same shape and scale parameters for all the compartments was used. A distribution model specific for each compartment was not investigated, and the technique was mostly used to "improve the fitting and reduce the overfitting" (22) in each voxel.

Instead, Eq. [1] can be generalized by modeling the tri-dimensional diffusivity of each spin packet with a full diffusion tensor \mathbf{D} (Fig. 2d–f). This is analytically challenging because it implies the integration of a matrix-variate distribution $P(\mathbf{D})$ over the space of tensors, which are symmetric positive-definite (SPD) matrices. Basser et al. (23) proposed a normal distribution for symmetric matrices, but not restricted to SPD matrices. In contrast, a natural parametric distribution for SPD matrices is the matrix-variate Gamma ($mv\text{-}\Gamma$) distribution (24), which generalizes the Wishart distribution by allowing non-

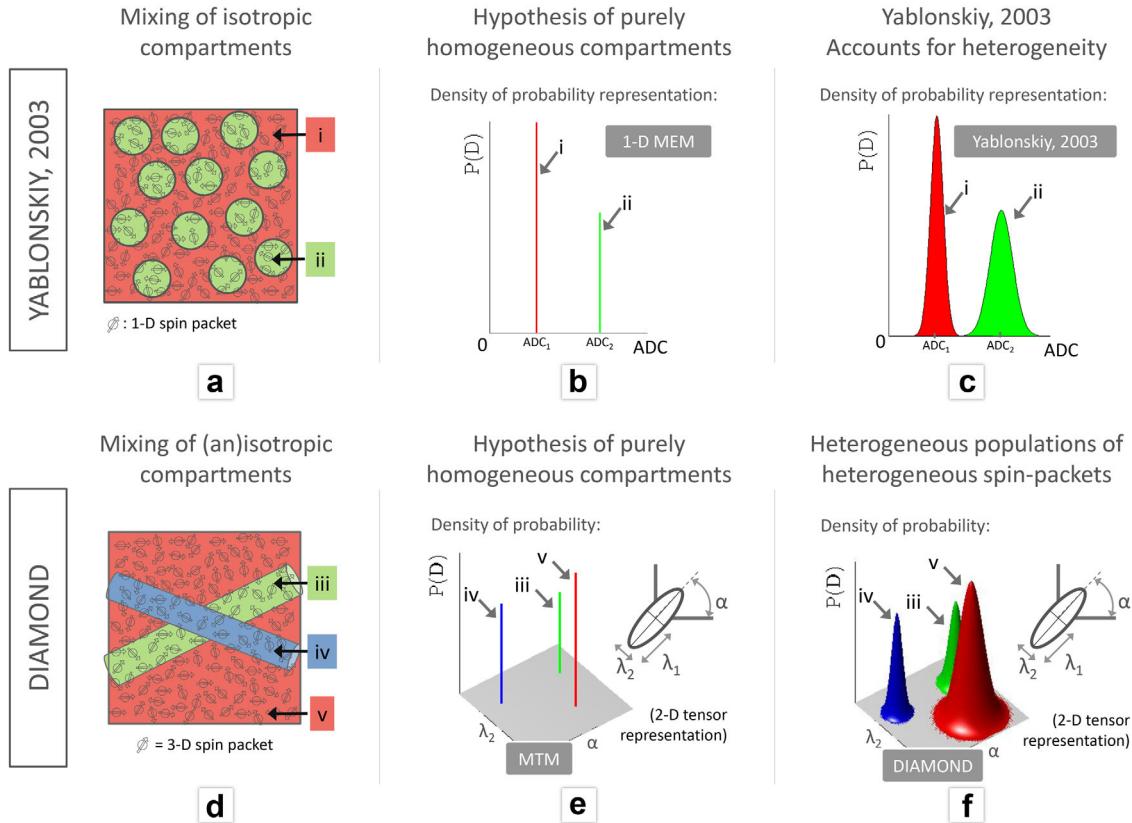


FIG. 2. DIAMOND generalizes the 1D ADC statistical distribution model of Yablonskiy et al. (20). **a:** Hypothetical isotropic two-population (or two-compartment) environment. **b:** Under the hypothesis of purely homogeneous compartments with no exchange, a bi-exponential decay is traditionally considered to describe the signal arising from (a). The probability density of diffusivities represents the fraction of all possible ADCs present in the voxel. For the bi-exponential model, it consists of two Diracs representing the ADC of each compartment. **c:** Because it is more realistic to consider that each compartment has some degree of heterogeneity, Yablonskiy et al. (20) proposed to consider a peak-shaped continuous distribution of ADCs for each compartment, the heterogeneity of a compartment being captured by the concentration of the corresponding peak. **d:** Hypothetical multi-compartment model, in which an isotropic (red) and two anisotropic (blue and green) compartments are mixed. For illustration purpose we represent here a 2D multicompartiment model. **e:** Under the hypothesis of purely homogeneous compartments with no exchange, a multitensor model is typically used to describe the signal arising from (d). The corresponding probability density of diffusivities is composed of a mixture of delta functions. **f:** DIAMOND captures the multidimensional diffusivity and heterogeneity of each compartment using peak-shaped distributions of multidimensional diffusivities. The expectation of each matrix-variate distribution captures the compartment overall diffusivity while the distribution concentration captures its microstructural heterogeneity. Specifically, a distribution with a broad peak indicates a highly heterogeneous compartment.

integer numbers of degrees of freedom. Jian et al. used a mixture of $N > 100$ Wishart distributions with fixed parameters and orientations in a spherical deconvolution approach (25). Leow et al. (26) extended this work by considering a nonparametric distribution of tensors on the sphere to avoid the assumption of a fixed anisotropy profile for all the WM fascicles. Similarly to Q-Ball imaging (27,28) and constrained spherical deconvolution (29), the models of (25) and (26) focus on describing the *orientation profile* of the diffusion at a single b -value without explicitly modeling the tissue microstructure. They correspond to voxel-wise *models of the signal*. They conflate the signal arising from each compartment and do not provide insight into each of them.

In this work, we generalize the approach of (20) and propose a novel DCI technique that characterizes the Distribution of 3D Anisotropic MicroStructural environments in Diffusion-compartment imaging (DIAMOND). Building on preliminary work (30), we consider measure-

ments arising from the large number of 3D spin packets such that each homogeneous spin packets undergoes local anisotropic 3D diffusion represented by a diffusion tensor \mathbf{D} . We consider that each voxel contains heterogeneous populations (e.g., arising from multiple fascicles and extra-cellular space) of heterogeneous spin packets and describe each population with a mv- Γ distribution of spin packets. The signal at each voxel is modeled by mathematical integration of the contributions of each 3D spin packet, which has an analytical solution. For each compartment, a priori information about the expected shape of the spin packets distribution can be introduced to represent free, isotropically restricted, intra-axonal, and hindered diffusion arising from each WM fascicle in each voxel. DIAMOND is an hybrid biophysical *model of the tissues* that combines multicompartiment and statistical modeling to provide insight into each compartment in each voxel. Unlike (22), it focuses on characterizing each compartment with a statistical distribution of

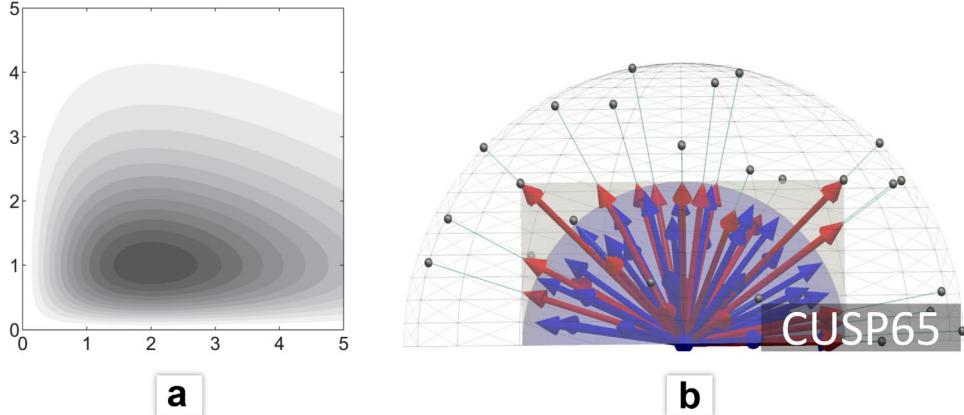


FIG. 3. a: Illustration of the density [3] on Sym_2^+ with concentration parameter $\kappa = 3$ and mean parameter $\mathbf{D}^0 = 1.5 \text{ Diag}(2, 1)$, so that the mode of the density is at $\mathbf{M}^0 = \text{Diag}(2, 1)$. Shown is a cross-section of the density for SPD matrices \mathbf{D} whose off-diagonal entries are equal to zero. The horizontal and vertical axes are the first and second diagonal entries of \mathbf{D} , respectively. b: The CUSP65 gradient set. CUSP is based on the modification of a 2-shell HARDI. In contrast to a multishell HARDI, the pulse duration and separation δ and Δ of the pulsed-gradient spin echo sequence are fixed to achieve the *b*-value of the *inner shell* (instead of the outer-shell for multi-shell HARDI), which requires a shorter TE and provides a significant SNR boost. The gradients of the outer shell have maximally separated gradients orientation with respect to the inner shell (32) but cannot be imaged with the fixed low δ and Δ . Instead, their strength is reduced (red gradients) so that they lie in the cube enclosing the inner shell. This cube is a cube of constant TE in q-space, in which any gradient can be imaged without modification of δ and Δ but by appropriate application of the gradient system (33). This enables imaging *b*-values up to three time the *b*-value of the inner-shell, corresponding to the diffusion gradients $\mathbf{g} = (\pm 1, \pm 1, \pm 1)$. CUSP enables imaging of multiple *b*-value with low TE, high SNR, and high angular coverage.

diffusivities. In contrast to models of the signal (25–29), it requires the estimation of the number of tissue compartments in each voxel, which in turn enables direct assessment of compartment-specific diffusion characteristics such as the compartment mean diffusivity, axial diffusivity and radial diffusivity (cMD, cAD, and cRD). It also provides an overall measure of heterogeneity for each compartment.

We evaluate our approach with numerous *in silico* and *in vivo* experiments using short-duration diffusion gradient sets. We demonstrate that the angular error of DIAMOND favorably compares to the ball-and-stick model (31) and Jbabdi's model (22). We provide insight into DIAMOND's model parameters via various numerical simulations of tissue microstructures. We show that DIAMOND better predicts the DW signal compared to NODDI (11) with both simulations and *in vivo* data, providing evidence that it better captures the underlying DW signal formation. Finally, we investigate the repeatability of DIAMOND's estimation *in vivo* with a test-retest experiment.

This article is organized as follows. We first describe how DIAMOND is derived by generalizing the ADC distribution model of Yablonskiy et al. (20). We then describe the evaluation method and report the results. Finally, we discuss how estimation of the distribution of 3D diffusivities within each compartment provides insight into the tissue microstructure and may provide novel markers of abnormal tissues in the brain from short duration acquisitions.

THEORY

DIAMOND: A Biophysical Model of the Contribution of 3D Spin Packets

Following the ADC approach of (20), we consider measurements of the signal arising from the large number of

spin packets within a voxel. In contrast to (20), we consider 3D spin-packets so that when interacting with an homogeneous portion of the microstructure they give rise to *anisotropic* 3D Gaussian diffusion represented by a *diffusion tensor* \mathbf{D} , whose contribution to the signal for a diffusion gradient \mathbf{g}_k is: $S_0 \exp(-b_k \mathbf{g}_k^T \mathbf{D} \mathbf{g}_k) d\mathbf{D}$. This enables us to capture the local 3D structure of the restriction and hindrance to diffusion of water molecules. Spin packets travel along different trajectories and are confronted with different barriers to displacement. The fraction of spin packets described by a 3D diffusivity \mathbf{D} in the voxel is given by a *matrix-variate* distribution $P(\mathbf{D})$ and the DW signal S_k is modeled by:

$$S_k = S_0 \int_{\mathbf{D} \in \text{Sym}^+(3)} P(\mathbf{D}) \exp(-b_k \mathbf{g}_k^T \mathbf{D} \mathbf{g}_k) d\mathbf{D}, \quad [2]$$

where $\text{Sym}^+(3)$ is the set of 3×3 SPD matrices. If a voxel consisted of exactly one *homogeneous* microstructural environment characterized by a tensor \mathbf{D}^0 , $P(\mathbf{D})$ would be a delta function $P(\mathbf{D}) = \delta(\mathbf{D} - \mathbf{D}^0)$ and our model would be equivalent to DTI. If it consisted of several identifiable homogeneous microstructural environments (i.e., compartments), a mixture of delta functions would be used (Fig. 2e). More realistically, each microstructural environment contains some degree of heterogeneity and is best described by a *population of spin packets*. This can be accounted for by modeling the composition of each microstructural environment with a peak-shaped matrix-variate distribution of spin packets centered around \mathbf{D}^0 and defined over the space of SPD matrices (Fig. 2f).

A natural peak-shaped distribution for SPD matrices is the mv- Γ distribution (Fig. 3). A $p \times p$ SPD random matrix $\mathbf{D} \in \text{Sym}^+(p)$ follows a mv- Γ distribution with shape parameter $\kappa > (p-1)/2$ and scale parameter $\Sigma \in \text{Sym}^+(p)$ if it has density (24,34,35):

$$P_{\kappa,\Sigma}(\mathbf{D}) = \frac{|\mathbf{D}|^{\kappa-(p+1)/2}}{|\Sigma|^{\kappa}\Gamma_p(\kappa)} \exp(-\text{trace}(\Sigma^{-1}\mathbf{D})), \quad [3]$$

where $|\cdot|$ is the matrix determinant and Γ_p is the multivariate gamma function:

$$\Gamma_p(\kappa) = \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma[\kappa - (j-1)/2]. \quad [4]$$

Its expectation is $\mathbf{D}^0 = \kappa\Sigma$ and its mode is $\mathbf{M}^0 = \mathbf{D}^0(\kappa-1)/\kappa$ (24). The shape parameter κ determines the concentration: for constant \mathbf{D}^0 , the density [3] becomes more concentrated around \mathbf{D}^0 as κ increases. We consider the presence of N_p populations of spin packets in slow exchange in a voxel and represent the composition of each population with a mv- Γ distribution of spin packets $P_{\kappa_j,\Sigma_j}(\mathbf{D})$ with parameters κ_j and Σ_j , $j \in [1, \dots, N_p]$. This amounts to the mixture:

$$P(\mathbf{D}) = \sum_{j=1}^{N_p} f_j P_{\kappa_j,\Sigma_j}(\mathbf{D}), \quad [5]$$

where $f_j \in [0, 1]$ are the fractions of occupancy and sum to one. Combining [2] and [5] yields:

$$S_k = S_0 \sum_{j=1}^{N_p} f_j \int_{\mathbf{D} \in \text{Sym}^+(3)} P_{\kappa_j,\Sigma_j}(\mathbf{D}) \exp(-b_k \mathbf{g}_k^T \mathbf{D} \mathbf{g}_k) d\mathbf{D} \quad [6]$$

The integrals in the right-hand side of [6] are Laplace transforms of $P_{\kappa_j,\Sigma_j}(\mathbf{D})$, which have a known analytical expression (24,25). This leads to the following model of the DW signal for a gradient direction \mathbf{g}_k :

$$\begin{aligned} S_k &= S_0 \sum_{j=1}^{N_p} f_j \left(1 + \frac{b_k \mathbf{g}_k^T \mathbf{D}^0 j \mathbf{g}_k}{\kappa_j} \right)^{-\kappa_j} \\ &= S_0 \sum_{j=1}^{N_p} f_j \exp \left(-\kappa_j \log \left(1 + \frac{b_k \mathbf{g}_k^T \mathbf{D}^0 j \mathbf{g}_k}{\kappa_j} \right) \right) \\ &= S_0 \sum_{j=1}^{N_p} f_j \mathcal{D}(\mathbf{D}^0 j, \kappa_j), \end{aligned} \quad [7]$$

where $\mathbf{D}^0 j = \kappa_j \Sigma_j$ is the expectation of the j^{th} mv- Γ distribution. $\mathcal{D}(\mathbf{D}^0, \kappa) = S_0 \left(1 + \frac{b_k \mathbf{g}_k^T \mathbf{D}^0 \mathbf{g}_k}{\kappa} \right)^{-\kappa}$ denotes the non-monoexponential decaying signal arising from a population of spin packets described by $P_{\kappa,\Sigma}(\mathbf{D}^0 = \kappa\Sigma)$. Finite values of κ_j capture the heterogeneity of each population of spin packets (Fig. 2). Specifically, a distribution with a sharp peak indicates a population with a highly homogeneous microstructure; a distribution with a broad peak indicates a highly heterogeneous population.

A DIAMOND Model That Captures Isotropic, Hindered, and Restricted Diffusion

Equation [7] provides the generic expression of the DW signal arising from heterogeneous populations (e.g., different fascicles) of heterogeneous spin packets in each voxel. One may consider particular compartments in voxels and evaluate their presence using model selection

techniques. For example, some voxels likely contain water molecules that undergo *free* diffusion at the measured diffusion scale due to partial volume effect with cerebrospinal fluid. This can be modeled with a mv- Γ distribution $P_{\kappa_{\text{free}}, \Sigma_{\text{free}}}$ with isotropic mode $\mathbf{M}^0_{\text{free}} = \text{diag}(3 \times 10^{-3}) \text{ mm}^2/\text{s}$ and two free parameters, the shape parameter κ_{free} and the volumic fraction f_{free} .

Cytotoxic brain edemas which occur in ischemic strokes (36,37) and in traumatic brain injuries (38) are characterized by intracellular water accumulation caused by an increased cell membrane permeability for ions and by ionic pump failure due to energy depletion (39). This leads to a greater proportion of water molecules inside glial cells, where diffusion is macroscopically isotropic but becomes restricted. This *isotropic restricted* diffusion may be captured by a second distribution with three free parameters: the isotropic mode $\mathbf{M}^0_{\text{iso,r}}$, the shape parameter $\kappa_{\text{iso,r}}$, and the volumic fraction $f_{\text{iso,r}}$.

The diffusion of water molecules restricted and hindered by a fascicle j may be *both* represented by a single mv- Γ distribution (30) with anisotropic cylindrical \mathbf{M}^0_j , shape parameter κ_j , and fractions f_j leading to the signal generation model:

$$\begin{aligned} (\mathcal{M}_1) \quad S_k &= f_{\text{free}} \mathcal{D}(\mathbf{D}^0_{\text{free}}, \kappa_{\text{free}}) + f_{\text{iso,r}} \mathcal{D}(\mathbf{D}^0_{\text{iso,r}}, \kappa_{\text{iso,r}}) \\ &\quad + \sum_{j=1}^{N_f} f_j \mathcal{D}(\mathbf{D}^0_j, \kappa_j), \end{aligned} \quad [8]$$

where $N_f = N_p - 2$ is the number of fascicles in the voxel. This model has $6N_f + 4$ free parameters[†].

Alternatively, the diffusion of water molecules restricted to the intra-axonal space of a fascicle j and the surrounding hindered, extra-axonal molecules may *each* be represented by a mv- Γ distribution with anisotropic cylindrical $\mathbf{M}^0_{j,r}$ (intra-axonal restricted) and $\mathbf{M}^0_{j,hin}$ (hindered) with identical eigenvectors and with shape parameters $(\kappa_{j,r}, \kappa_{j,hin})$ and volumic fractions $(f_{j,r}, f_{j,hin})$:

$$\begin{aligned} (\mathcal{M}_2) \quad S_k &= f_{\text{free}} \mathcal{D}(\mathbf{D}^0_{\text{free}}, \kappa_{\text{free}}) + f_{\text{iso,r}} \mathcal{D}(\mathbf{D}^0_{\text{iso,r}}, \kappa_{\text{iso,r}}) \\ &\quad + \sum_{j=1}^{N_f} [f_{j,r} \mathcal{D}(\mathbf{D}^0_{j,r}, \kappa_{j,r}) + f_{j,hin} \mathcal{D}(\mathbf{D}^0_{j,hin}, \kappa_{j,hin})]. \end{aligned} \quad [9]$$

A tortuosity model (40) may be used to constrain the diffusivities of water molecules inside and around axons (11,12), resulting in only two additional free parameters per fascicle compared to (\mathcal{M}_1): the intra-axonal volume fraction $\nu_j = \frac{f_{j,r}}{f_{j,r} + f_{j,hin}}$ and one additional concentration parameter per fascicle ($\kappa_{j,r}$ and $\kappa_{j,hin}$ instead of κ_j).

Parameter Estimation

The parameters of DIAMOND at each voxel are estimated using a maximum a posteriori approach. We denote by V the image domain, by \mathbf{y} the set of N_g DW images and by

[†]Four parameters for each \mathbf{D}^0_j ($j \in [1, N_f]$), one parameter for $\mathbf{M}^0_{\text{iso,r}}$, $N_f + 1$ fractions f_{\bullet} , $N_f + 2$ concentration parameters κ_{\bullet} .

y_k^i the i^{th} voxel of DW image k . We consider a variable number N_{pi} of populations of spin packets for each voxel i . We denote by $\mathbf{D} = (\mathbf{D}^{0,i}_1, \dots, \mathbf{D}^{0,i}_{N_{pi}}, i \in V)$ and $\kappa = (\kappa_1^i, \dots, \kappa_{N_{pi}}^i, i \in V)$ the parameters of the mv- Γ distributions for each population, and by $\mathbf{f} = (f_1^i, \dots, f_{N_{pi}}^i, i \in V)$ the corresponding fractions of occupancy. We parameterize the tensors \mathbf{D} in the log-domain by setting $\mathbf{L} = (\log(\mathbf{D}^{0,i}_1), \dots, \log(\mathbf{D}^{0,i}_{N_{pi}}), i \in V)$ to ensure the estimation of positive-definite matrices (41–(43)). The estimation of the model parameters is performed by maximizing:

$$\begin{aligned} \left\{ \hat{\mathbf{L}}_{\text{MAP}}, \hat{\kappa}_{\text{MAP}}, \hat{\mathbf{f}}_{\text{MAP}} \right\} &= \underset{\mathbf{L}, \kappa, \mathbf{f}}{\text{argmax}} p(\mathbf{L}, \kappa, \mathbf{f} | \mathbf{y}) \\ &= \underset{\mathbf{L}, \kappa, \mathbf{f}}{\text{argmax}} p(\mathbf{y} | \mathbf{L}, \kappa, \mathbf{f}) p(\mathbf{f} | \kappa, \mathbf{L}) p(\kappa | \mathbf{L}) p(\mathbf{L}). \end{aligned} \quad [10]$$

The parameters $\hat{\mathbf{D}}_{\text{MAP}}$ are computed as $\exp(\hat{\mathbf{L}}_{\text{MAP}})$. In this work, we did not consider any prior knowledge on the estimated fractions and shape parameters κ so that $p(\mathbf{f} | \kappa, \mathbf{L})$ and $p(\kappa | \mathbf{L})$ are uniform densities. We assume that, conditionally on $\mathbf{L}, \kappa, \mathbf{f}$, the intensities of the DW images \mathbf{y} are normally distributed around the unknown modeled signal S_k with variance σ^2 so that:

$$\begin{aligned} p(\mathbf{y} | \mathbf{L}, \kappa, \mathbf{f}) \\ = \prod_{i \in V} \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{\sum_{k=1}^{N_g} \|S_k(\exp(\mathbf{L}^i), \kappa^i, \mathbf{f}^i) - y_k^i\|^2}{2\sigma^2} \right), \end{aligned} \quad [11]$$

where S_k is given by [7]. We consider a regularization prior $p(\mathbf{D})$ that exploits spatial homogeneity by setting $p(\mathbf{L}) \propto \prod_{i \in V} \prod_{j=1}^{N_p} \exp(-\alpha_{\text{reg}} \phi(\|\nabla \mathbf{L}^{i,j}\|))$, where $\|\nabla \mathbf{L}^{i,j}\|$ is the norm of the spatial gradient of $\mathbf{L}^{i,j}$, α_{reg} is a parameter controlling the regularization strength, and $\phi(s) = \sqrt{1 + s^2/K_{\text{reg}}^2}$ to account for *anisotropic* regularization and to preserve sharp contours. Similarly to (33), $\|\nabla \mathbf{L}^{i,j}\|$ is approximated by finite difference after identifying the log-tensor most similar to $\mathbf{L}^{i,j}$ in each neighboring voxel

METHODS

Algorithmic settings. The DIAMOND estimation algorithm was implemented in C++ and parallelized over the image space. The numerical optimization (Eq. [10]) was achieved with BOBYQA (44), a derivative-free bound-constrained optimization technique. We found BOBYQA to converge faster than the Newton method and to be less sensitive to local minima compared to gradient descent algorithms such as conjugate gradient or Levenberg-Marquardt. We parameterized the orientation of the \mathbf{L}^i 's with the Euler angles, as they empirically led to a more efficient optimization. Tensors and fractions were initialized as in (33) and concentration parameters κ 's were initialized to 100 (high compartment homogeneity). At each voxel, DW intensities for which $y_k^i/S_0 < 0.1$ were considered below the noise floor and discarded in the objective function. The optimization was achieved by gradually increasing the model complexity, from a single stick model to the ball-and-stick to the estimation

of the full DIAMOND model. The spatial regularization was achieved by iteratively refining the model parameters over the image while progressively increasing the regularization parameter α_{reg} in $[0, 1]$ (45) with $K_{\text{reg}} = 0.01$ (33).

We estimated the number of fascicles N_{fi} at each voxel i by first considering $N_{fi} = 0$ and by increasing N_{fi} (up to three) as long as it significantly decreases the generalization error (GE), as proposed in (46). We currently considered the model DIAMOND^{M1} (Eq. [8]) with $f_{\text{iso,r}} = 0$, leading to $6N_{fi} + 2$ free parameters per voxel. The total running time for a typical DWI acquisition with matrix size 128×128 was approximately 7 min per slice with a Intel® Xeon® E5-2687W processor with 8 cores.

DW gradient set. The estimation of DIAMOND requires an acquisition that achieves multiple non-zero b -values. While the full study of acquisition scheme is out of scope of this work, we chose here to use the Cube and Sphere (CUSP) (33) gradient set. CUSP combines spherical and cubic sampling in q-space (Fig. 3b), achieving a large number of non-zero b -values with high angular coverage, short TE and high SNR. We considered a CUSP65 set with 5 $b = 0$ s/mm² images, 30 gradients on the inner shell at $b = 1000$ s/mm² and 30 gradients on the cube of constant TE corresponding to b -values between 1000 s/mm² and 3000 s/mm².

Angular Detection Accuracy

We simulated fascicles crossing at varying angles (0° – 90°) and compared the angular reconstruction error of DIAMOND to that of the ball-and-sticks model (31) and Jbabdi's model (22) as implemented in FSL. For each crossing angle, we simulated the DW signal arising from the multi-tensor model:

$$S_k = S_0 [f_{\text{free}} \exp(-b_k D_{\text{free}}) + f_1 \exp(-b_k \mathbf{g}_k^T \mathbf{D}_1 \mathbf{g}_k) + f_2 \exp(-b_k \mathbf{g}_k^T \mathbf{D}_2 \mathbf{g}_k)] \quad [12]$$

with $D_{\text{free}} = 3 \times 10^{-3}$ mm²/s (47); $f_{\text{free}} = 0.15$; $f_1 = 0.60$; $f_2 = 0.25$; $\text{Trace}(\mathbf{D}_1) = \text{Trace}(\mathbf{D}_2) = 2.1 \times 10^{-3}$ mm²/s (48); $\text{FA}(\mathbf{D}_1) = \text{FA}(\mathbf{D}_2) = 0.9$; CUSP65 gradient set; Rician-noise corruption (SNR = 30dB on $b = 0$ s/mm²). We independently repeated the simulations 1000 times and estimated, for each of them, the DIAMOND, ball-and-sticks and Jbabdi's models. We compared the average and 95% confidence interval of the angular error between the estimated fascicle orientations.

Assessment of the Modeled Diffusion Profile and Interpretability of Model Parameters

We compared the ability of DIAMOND and NODDI (11) to capture the diffusion profile. We generated realistic synthetic diffusion data with Monte-Carlo simulations (49) using the Camino toolkit (200,000 walkers, 5000 time points). We considered two axonal geometries: (1) aligned cylinders of varying radii and (2) cylinders crossing at 45° . The signal was simulated with CUSP65 and parameters typically achievable with a clinical scanner: pulse duration $\delta = 30$ ms, pulse separation $\Delta = 40$ ms,

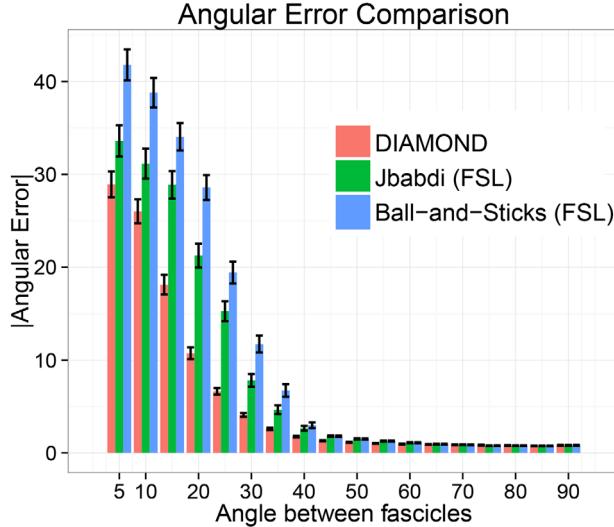


FIG. 4. Angular detection accuracy achieved by DIAMOND, the ball-and-sticks model and Jbabdi's model. The plot shows, for each crossing angle, the average and 95% confidence interval of the angular error between the estimated fascicle orientations for 1000 independent model estimations in a single voxel (SNR on $b=0 \text{ s/mm}^2$: 30 dB).

$\text{TE}=90 \text{ ms}$, maximum gradient strength $G=0.040 \text{ T/m}$ and Rician noise corruption ($\text{SNR}=30 \text{ dB}$ on the $b=0 \text{ s/mm}^2$ image). NODDI parameters were estimated using the publicly available NODDI toolbox.[‡] The mean and standard deviation of the diffusion profiles of DIAMOND and NODDI over 100 independent repetitions of each experiment in a single voxel were plotted and compared. Variations of DIAMOND's and NODDI's model parameters for increasing axonal radii were investigated.

We also simulated voxels with various axonal orientation dispersion by simulating in each voxel 10,000 cylinders with orientations drawn from a Watson distribution with increasing dispersion indices. The signal was simulated with the CUSP65 scheme ($\delta=30 \text{ ms}$; $\Delta=40 \text{ ms}$; $\text{TE}=90 \text{ ms}$; $G_{\max}=0.040 \text{ T/m}$) using the analytical expression of the diffusion in sticks (11,19).

Model Comparison with *In Vivo* Data

We compared DIAMOND to various diffusion models with *in vivo* acquisitions by assessing their GE. The GE quantifies the capability of a model to accurately predict the DW signal for unobserved gradient directions and strengths. Unfortunately, the GE cannot be computed exactly when a finite number of observations (i.e., DW images) is available. We used the 0.632 bootstrap method (using $B=500$ bootstrap iterations) which provides an estimate of GE with low bias and low variance (50). Acquisitions were carried out using a Siemens 3T Trio scanner with a 32 channel head coil and a pulsed-gradient spin echo DWI sequence with echo-planar imaging readout. Eddy current distortion was minimized by utilizing a twice-refocused spin echo sequence (51). We used a CUSP90 scheme that augments CUSP65 with

three inner-shells of six directions at $b=400, 600, 800 \text{ s/mm}^2$, and seven $b=0 \text{ s/mm}^2$ images to assess the signal prediction with a wide range of b -values. A healthy volunteer was scanned using two repetitions of CUSP90 with the following imaging parameters: $\text{FOV}=220 \text{ mm}$, matrix = 128×128 , 71 slices, resolution = $1.7 \times 1.7 \times 2 \text{ mm}^3$, $\text{TR}=10704 \text{ ms}$, $\text{TE}=78 \text{ ms}$, Partial Fourier = 6/8, iPAT = 2, Pixel band width = 1396 Hz/pixel. The scan duration was 16 min for each CUSP90. The SNR measured over a whole brain WM mask in a $b=0 \text{ s/mm}^2$ image was 30.4 (29.6dB). A T1-weighted MPRAGE image was also acquired with 160 slices, $\text{FOV}=256 \text{ mm}$, matrix = 256×256 , resolution = $1 \times 1 \times 1 \text{ mm}^3$, $\text{TE}=2.27 \text{ ms}$, $\text{TR}=1410 \text{ ms}$, 3 min 50 s. The DW images were corrected for potential head motion during the scan by rigid registration of the DW-images to the $b=0 \text{ s/mm}^2$ image (52). The gradient orientations were compensated for the rotation component of the transformation for each image (53).

We considered five tissue models: (1) the single tensor model (16), (2) NODDI (11), (3) a multi-tensor model with one isotropic ($\mathbf{M}^0_{\text{free}}=3 \times 10^{-3} \text{ mm}^2/\text{s}$) and one anisotropic tensor (1T+iso), (4) a DIAMOND model with one isotropic and a *single* anisotropic compartment (1DIAMOND+iso), and (5) the complete DIAMOND model with up to three cylindrical anisotropic compartments (DIAMOND+iso).

Assessment of the Repeatability of Estimated Model Parameters with Test-Retest

We used the two repetitions of CUSP90 acquisitions (see Model Comparison with *In Vivo* Data section) to evaluate the repeatability of DIAMOND's estimation with *in vivo* data. We estimated DIAMOND and NODDI for each CUSP90 and compared the corresponding model parameters in two ROIs.

Extraction of Fascicle-Based Diffusion Characteristics

A multi-fascicle DCI model such as DIAMOND enables characterization of diffusion parameters for each fascicle in each voxel. This, however, requires the delineation of the fascicle(s) of interest to identify, at each voxel, the anisotropic compartment corresponding to the local fascicle orientation. We spatially aligned one CUSP90 acquisition to the T1-weighted scan automatically parcelated using (54). We then achieved dense whole brain tractography (parameters: $n=15$ streamlines per WM voxel, stopping criteria: touching a gray matter voxel, $cFA < 0.15$; angle $> 40^\circ$) and extracted the corticospinal tracts (CST) by selecting only streamlines touching the internal capsule and either the precentral gyrus (primary motor cortex) or the postcentral gyrus (primary somatosensory cortex). Additional exclusion regions were manually drawn to filter out spurious tracts. The CST was also extracted from a DTI whole brain tractography using same tractography parameters and regions of interest. We extracted the cFA of the CST by computing, at each streamline point, the cFA of the anisotropic compartment most aligned with the local streamline orientation. We compared results to (1) DTI tractography and projection of the DTI FA on the corresponding streamlines;

[‡]<http://cmic.cs.ucl.ac.uk/mig/>.

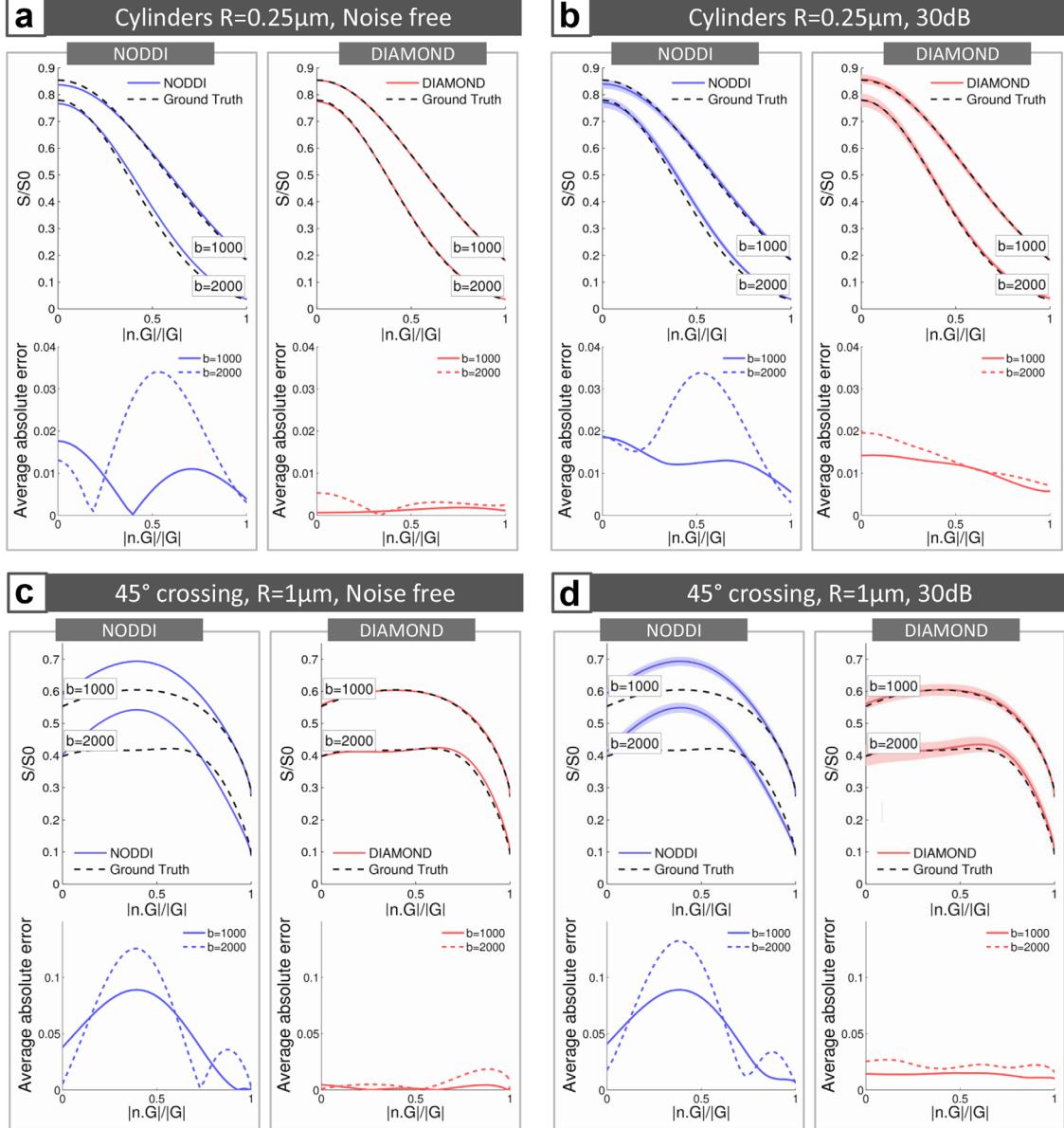


FIG. 5. Monte-Carlo numerical simulations with Camino for infinite SNR and SNR = 30 dB (CUSP65 gradient set). **a** and **b**: simulations with cylinders of radius $R = 0.25 \mu\text{m}$ oriented along the z axis. **c** and **d**: simulations with crossing cylinders (45°). The plots report the average signal attenuation (S/S₀) modeled by DIAMOND (red line) and NODDI (blue line) for gradients distributed in the x-z plane obtained by a rotation around the y axis, for $b = 1000 \text{ s/mm}^2$ and $b = 2000 \text{ s/mm}^2$. The average was computed over 100 independent repetitions of signal simulation and model estimation in a single voxel. The standard deviation over the 100 repetitions is reported in light red and blue. The black dashed line shows the ground truth. We also report a plot of the average absolute error with the ground truth. It shows that the signal attenuation is better captured by DIAMOND.

and (2) DIAMOND tractography and projection of the DTI FA. We also compared volumetric maps of extracted FA and cFA. The volumetric cFA map was computed by selecting, at each voxel, the cFA of the compartment whose orientation was most aligned with the largest number of streamlines.

High-Resolution and High-SNR DWI

We achieved DIAMOND reconstruction from high-resolution and high SNR super-resolved DWI (55). We acquired three orthogonal anisotropic CUSP65 scans (axial, coronal, and sagittal) with FOV = 240 mm,

matrix = 192×192 , resolution = $1.25 \times 1.25 \times 2 \text{ mm}^3$, a number of slices between 70 and 90 depending on the acquisition orientation and TE = 91 ms. For each orientation, we acquired two $b = 0 \text{ s/mm}^2$ with same parameters as above but opposite phase encoding directions to correct the scans for geometric and intensity distortion (56,57). The total scan duration was lower than 45 min. We reconstructed the underlying high-resolution isotropic DW images at $1 \times 1 \times 1 \text{ mm}^3$ using the quantitative super-resolution technique described in (55). DIAMOND's parameters were then estimated at each voxel.

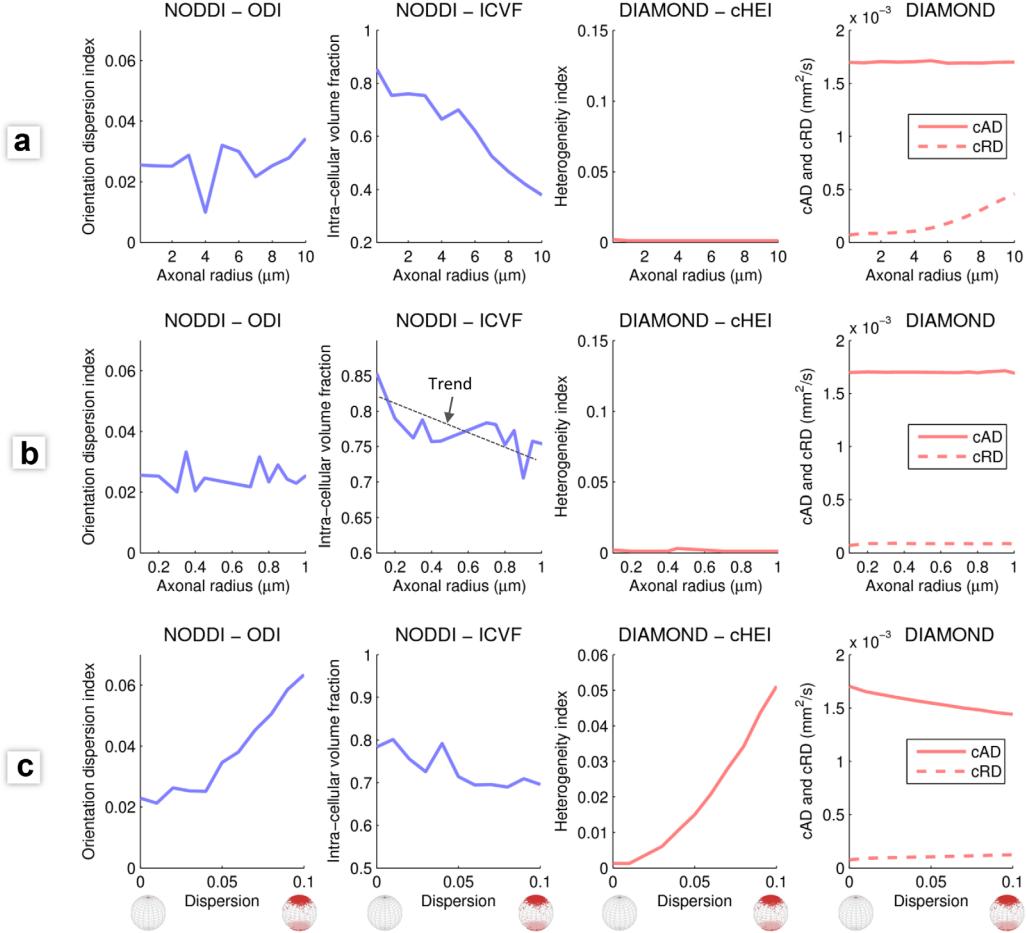


FIG. 6. Investigation of the variations of DIAMOND and NODDI parameters for various microstructures. **a** and **b**: Simulations for various axonal radii and fixed intra-cellular volume fraction ($ICVF = 0.7$). **c**: Simulations for various fascicle dispersion indexes and fixed $ICVF = 0.7$. **NODDI**: we report the dispersion index (ODI) and the $ICVF$ parameters. The variability of the curves shows the high sensitivity of the numerical optimizer to local minima. It shows that increasing axonal radius is characterized by a decreasing $ICVF$ (a and b, blue), while a fixed $ICVF$ was simulated. Increasing dispersion is correctly characterized by an increasing ODI index (c, blue). **DIAMOND**: we report the axial diffusivity and radial diffusivity of the fascicle compartment (cAD / cRD) and report an heterogeneity index ($cHEI$) based on the same transform than ODI : $cHEI(\kappa) = 2/\pi \arctan(1/\kappa)$ (small values = low heterogeneity; large values = high heterogeneity). It shows that increasing axonal radius is characterized by increasing cRD with constant cAD and $cHEI$ for large radii (a and b, red), while increasing axonal dispersion is characterized by increasing $cHEI$, decreasing cAD and constant cRD (c, red).

RESULTS

Figure 4 shows that the angular accuracy of DIAMOND favorably compares to that of the ball-and-sticks model and Jbabdi's model (22), particularly for angles $<30^\circ$.

Figure 5 reports the diffusion profiles of DIAMOND and NODDI computed from our Monte-Carlo simulations. Figure 5a–b shows that NODDI's diffusion profile deviates from the ground truth, even for parallel cylinders and without noise (Fig. 5a). In contrast, the diffusion profile of DIAMOND is closer to the ground truth. Figure 5c–d shows that the dispersion model used in NODDI is unable to capture the signal attenuation in voxels with crossing fascicles. This is an important limitation as fascicles crossing with angle $>40^\circ$ occurs in 60 to 90% of the voxels in the human brain (17). In contrast, the diffusion profile is better captured with DIAMOND.

Figure 6 shows that, with NODDI, variations of the axonal radii are captured by variations of the intra-cellular volume fraction ($ICVF$), although a constant $ICVF = 0.7$ was simulated in CAMINO. This is true for very large axons (Fig. 6a) but also for $R < 1 \mu\text{m}$ (Fig. 6b). In contrast, DIAMOND captures varying large axonal radii via cRD (Fig. 6a), while other parameters ($cHEI$ and cAD) remains approximately constant. These changes in cRD are consistent with the known physical behavior of water molecules diffusing more freely in larger axons. Small axonal radii could not be identified using cRD , however (Fig. 6b). Figure 6c shows that NODDI correctly captures the increased axonal dispersion via the ODI parameter, while a slight decrease in $ICVF$ is observed. DIAMOND captures the increased axonal dispersion as an increased heterogeneity of 3D diffusivities ($cHEI$), which is an expected behavior. Interestingly, cRD remains constant with increasing dispersion while cAD

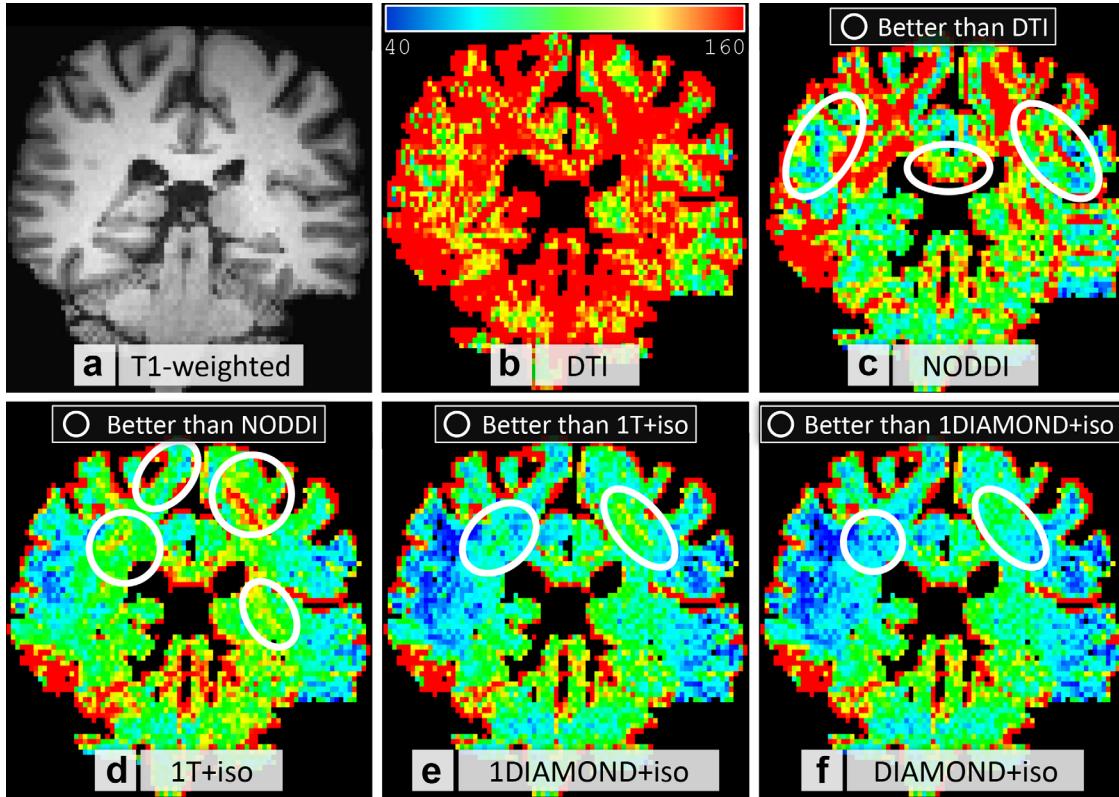


FIG. 7. Comparison of five diffusion models with *in vivo* data by assessment of their GE. It shows that DIAMOND+iso better predicts the *in vivo* DW signal, supporting the fact that it better captures the underlying diffusion decay.

decreases. This is consistent with the known physical behavior of water molecules whose diffusion endures more restrictions along the average fascicle orientation when the dispersion increases.

Figure 7 compares the GE maps of DIAMOND and various other diffusion models with *in vivo* data. DTI is the worst predictor of the diffusion signal (Fig. 7a), likely because it does not account for the non-monoeponential decay and assumes a single compartment per voxel. NODDI (Fig. 7b) provides a lower GE in regions of crossing and close to the cortex because it models the fascicle dispersion and accounts for freely diffusing water. However, Fig. 7c shows that estimation of all the parameters of 1T+iso ultimately provides better prediction of the signal compared to NODDI. This is likely because 1T+iso does not rely on prefixed parameters and can model anisotropic dispersion via an oblate tensor compartment. Figure 7d shows that accounting for the heterogeneity of 3D diffusivities (DIAMOND) slightly improves the GE in regions of crossings. Finally, Fig. 7e shows that DIAMOND+iso provides the best GE. This is because it accounts for *each* fascicle in *each* voxel and models each compartment heterogeneity.

Figure 8 reports the results of the test-retest experiment. It shows that, for fascicle compartments, the correlation of DIAMOND parameters (cAD, cRD, cHEI, f_j) between test and retest is comparable to NODDI, even in a crossing fascicle region. It indicates that DIAMOND's compartment-specific diffusivity characteristics (cAD,

cRD, cHEI) can be used to characterize fascicles with an acquisition with 90 DW images.

Figure 9a-g shows whole brain DIAMOND maps estimated from a single CUSP90 acquisition. Importantly, it must be understood that the k^{th} anisotropic compartment among neighboring voxels may not represent the same underlying biological compartment (i.e., the same fascicle), making the interpretation of whole brain maps challenging. In contrast, Figures 9h-j report the FA and cFA extracted for the CST. It shows that cFA is less impacted by crossings than is FA (Fig. 9i,j). It also demonstrates that DIAMOND tractography better resolves the extent of the CST than does DTI tractography (Fig. 9h,i).

Figure 10 reports the DIAMOND reconstruction from high-resolution super-resolved DWI. Especially, it shows the orientation of the mode of each mv-G distribution. The fascicles' orientation estimated in the corona radiata matches the known anatomy (Fig. 10b) and the projection of the fascicles into the gray matter of the cortex are well-captured (Fig. 10c).

DISCUSSION

We proposed a novel diffusion model that is motivated by biophysical considerations of the microstructure giving rise to the DW signal and that can be applied in clinical practice. Inspired by the statistical framework of (20), we considered measurements of the signal arising from 3D spin packets within each voxel and proposed to characterize the population of 3D spin packets of each

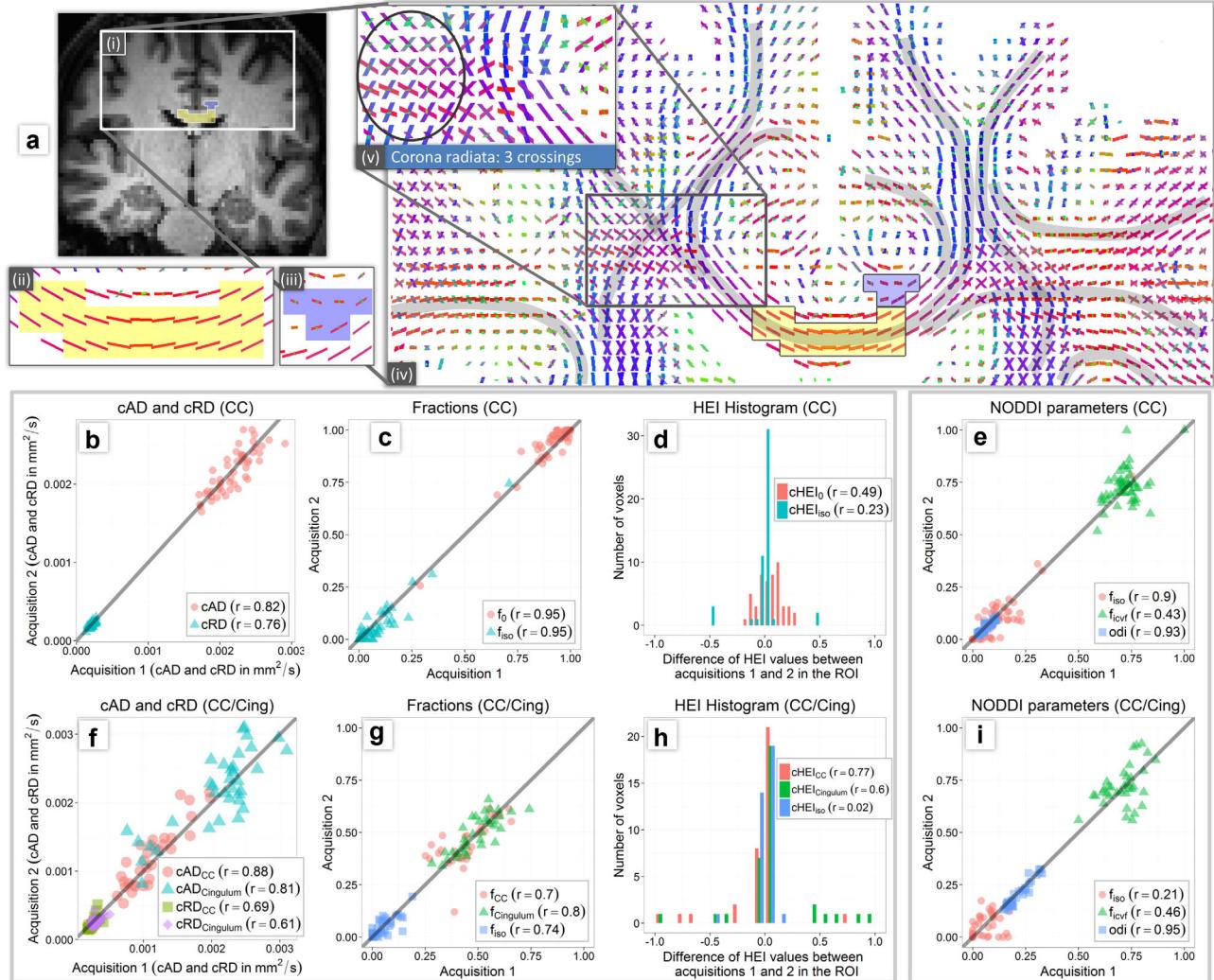


FIG. 8. Test-retest comparison of DIAMOND and NODDI. **a-i:** Illustration of the two ROIs in which the repeatability of model parameters was investigated. **a-ii:** Zoom on the single fascicle ROI, drawn in the body of the corpus callosum (50 voxels). **a-iii:** Zoom on the two-fascicles ROI in which the corpus callosum (red) and the cingulum (green) are crossing (35 voxels). **a-iv:** Estimated DIAMOND's fascicle orientations from a single CUSP90 scans, successfully matching the known anatomy. Particularly, three fascicles are detected in the corona radiata in which the corpus callosum, the corticospinal tract and the superior longitudinal fasciculus intersect. **b-d:** DIAMOND and **e:** NODDI parameters' repeatability in the single-fascicle ROI. **f-h:** DIAMOND and **i:** NODDI parameters' repeatability in the two-fascicles ROI. We report the Pearson correlation coefficient r between the two acquisitions for each parameter. This figure shows that the correlation between test and retest acquisitions is comparable between DIAMOND and NODDI for fascicle compartments and that DIAMOND's compartment-specific diffusivity characteristics (cAD, cRD, cHEI) can be used to characterize tissues with an acquisition with 90 DW images.

compartment by estimating a peak-shaped statistical distribution of diffusion tensors (Fig. 2f). In this work, we considered the peak-shaped mv- Γ distribution $P_{\kappa,\Sigma}$ which enables computation of an analytical solution to (Eq. [2]).

The mv- Γ distribution generalizes the Wishart (58) by allowing a non-integer number of degrees of freedom (24,32,33). Jian et al. (25) were the first to examine a model of mixture of Wishart distributions in DWI. They investigated a spherical deconvolution approach in which the basis functions were Wishart distributions with fixed diffusivity, fixed number of degrees of freedom ($m = 4$) and fixed orientations distributed over the unit sphere. As underlined in (25), the number of Wishart components ($N > 100$) was linked to the discretization

resolution of the fascicle orientation distribution function, irrespective of the number of underlying tissue compartments present in each voxel. This corresponds to a *model of the signal*. It focuses on estimating in each voxel a distribution of orientations at a single non-zero b -value without assumption about the underlying tissue microstructure, and does not provide insight into the tissue compartments giving rise to the signal. Recent works have investigated subsequent extraction of tissue information from the shape of the fascicle orientation distribution function such as the apparent fiber density (59,60). However, because the fascicle orientation distribution function is defined at a single b -value, the information about various diffusion scales is missed and the provided tissue information is limited. In addition,

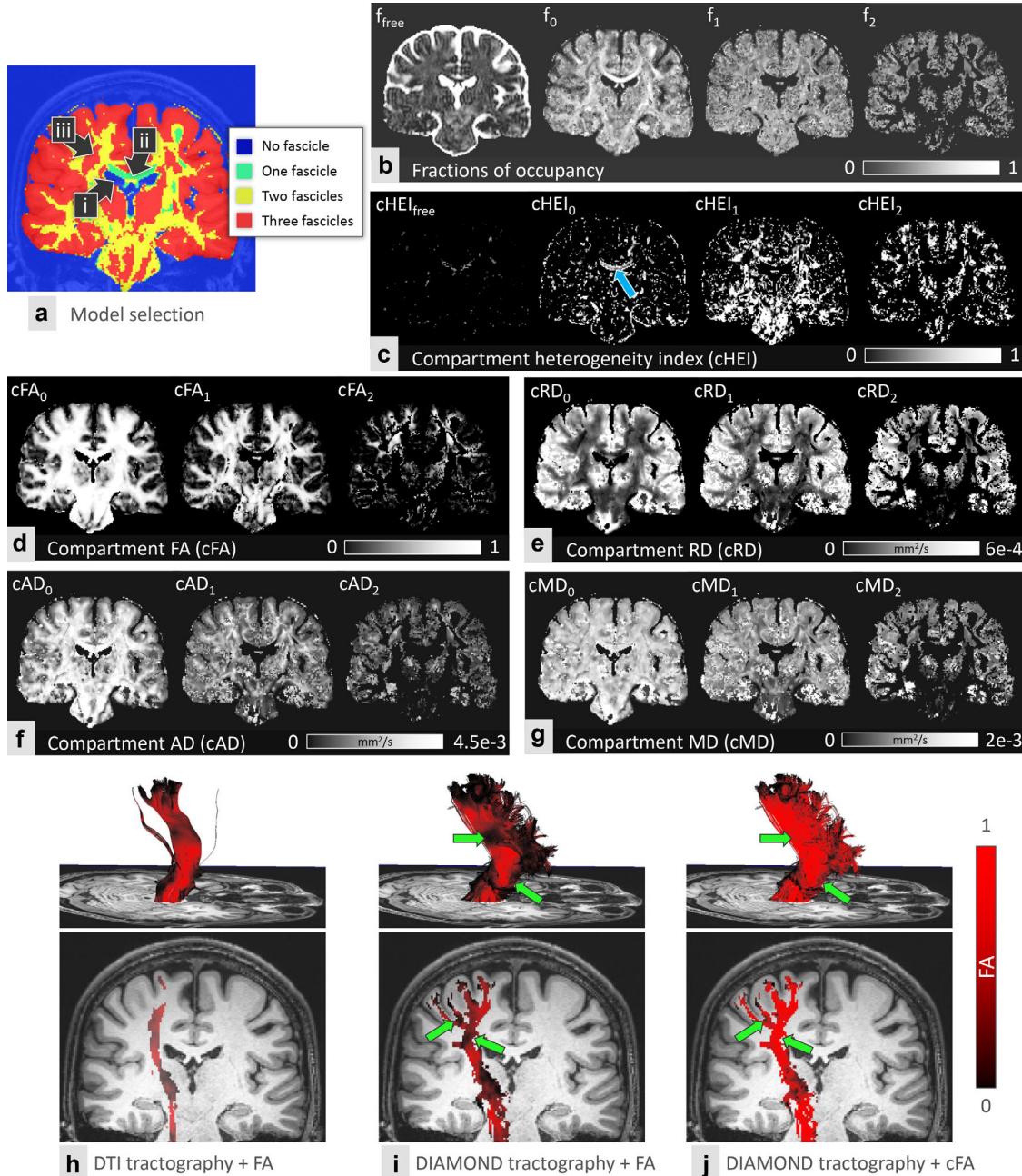


FIG. 9. Diffusion parameter maps estimated from one CUSP90 acquisition. The anisotropic compartments were ordered by decreasing cFA. **a:** Estimated number of fascicles at each voxel, correctly detecting no fascicle in the ventricles (a.i), a single fascicle in the body of the corpus callosum (a.ii) and up to three fascicles in the corona radiata (a.iii). **b:** Fractions of occupancy of each compartment. Particularity, f_{free} matches the known anatomical distribution of cerebrospinal fluid. **c:** cHEI for each compartment. Especially, high heterogeneity is found in the body of the corpus callosum (blue arrow). **d-g:** Whole brain maps of per-compartment diffusion characteristics (cFA, cRD, cAD, cMD). **h-j:** Extraction of the FA and cFA for the CST showing that DIAMOND tractography better resolves the extent of the CST than does DTI tractography (h-i) and that cFA is less impacted by crossings than is FA (green arrows, i-j). Images were created using the Misterl software (<http://www.benoitscherrer.com/Misterl/>).

similarly to NODDI (11) and the ball-and-sticks model (22,31), spherical deconvolution approaches (25,29) rely on the fundamental hypothesis that a single pre-fixed fascicle response function can well model the entire brain WM. This is inconsistent with the known presence of many axonal diameters and many degrees of myelination throughout the brain (12,15), and the interpretation of extracted tissue parameters

from the fascicle orientation distribution function remains unclear. While extracting a fascicle response function in each voxel was investigated recently (61), only a single fascicle response function can be used to deconvolve the signal in a voxel and the signal arising from two crossing fascicles with different characteristics (e.g., healthy and not healthy) cannot be modeled.

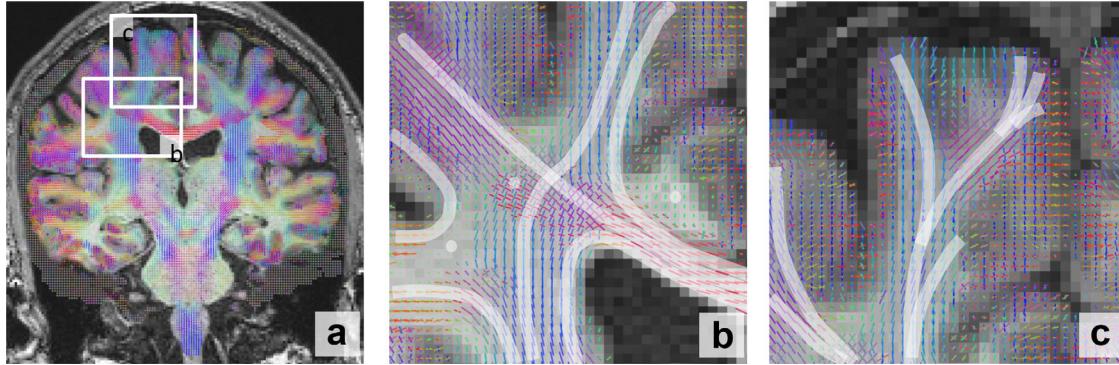


FIG. 10. DIAMOND reconstruction from high-resolution super-resolved DWI. **a:** Coronal slice of the DIAMOND model super-imposed on the T1-weighted image. **b:** DIAMOND correctly captures three crossing fascicles in the corona radiata. **c:** DIAMOND well captures the projection of the fascicles into the gray matter of the cortex.

In contrast, DIAMOND focuses on capturing the distributions of 3D diffusivities arising from each tissue compartment (see Fig. 2) and is an hybrid *model of the tissues* that combines both biophysical and statistical modeling. In contrast to *models of the signal* (25,26,28,29), it requires the estimation of the number of tissue compartments in each voxel, that is, the number of mv- Γ components. This was achieved by assessing the GE of models of increasing complexity, as proposed in (45). Unlike (25,26,28,29), DIAMOND also requires the acquisition of multiple non-zero b -values to disentangle the decay curves arising from each tissue compartment. Importantly, estimating a distribution of diffusivities within voxels requires acquiring all b -values with the same mixing times δ and Δ , ruling out the use of multiple separate single shells with optimized TE for each. We used the CUSP gradient scheme (Fig. 3b) which images a large number of different b -values with constant mixing times, uniform angular coverage and low TE, providing a substantial SNR boost compared to a multi-shell HARDI (33).

The parameters of DIAMOND provide a macroscopic statistical description of the tissue microstructure in each compartment. The expectation $E[P_{\kappa,\Sigma}] = \mathbf{D}^0 = \kappa\Sigma$ of each distribution $P_{\kappa,\Sigma}$ is a tensor that describes the average 3D diffusivity from which fractional anisotropy and axial and radial diffusivity of each compartment (cFA, cAD, and cRD) can be derived. The concentration parameter κ of $P_{\kappa,\Sigma}$ captures the overall compartment's heterogeneity that is consistent with the (oriented) 3D compartment. Such heterogeneity may result from heterogeneity in fascicle orientation, in axonal diameter, in axonal density or from undulation of axons. Modeling together multiple sources of heterogeneity is not a limitation. There are many possible sources of heterogeneity at different diffusion scales and it is not clear whether they can be captured separately when using a clinical scanner with long δ and $\delta \approx \Delta$ and clinically compatible scan times. The NODDI approach suggested that the fascicle dispersion and the ICVF could be specifically assessed by relying on a fixed representation of a WM fascicle throughout the brain. Our simulations illustrate that, when the simulated microstructure differs from the pre-fixed fascicle microstructure, NODDI's estimated diffusion profile substantially deviates from the true profile,

even without noise (Fig. 5). In addition, we demonstrate that when the fascicle's diffusivity is fixed, variations of the signal due to a varying microstructure are captured by the only remaining free parameters and changes in those parameters may be not consistent with the underlying microstructure. Particularly, we showed that variation of the axonal radius was captured by variations of the ICVF with NODDI, while a constant ICVF was simulated (Fig. 6a,b).

Instead, what we propose in DIAMOND is to estimate for each tissue compartment that can be detected its fraction of occupancy and a unimodal continuous distribution of 3D diffusivities, from which measures of the overall 3D diffusivity and overall microstructural heterogeneity can be derived. We show through simulations that axonal dispersion is captured by an increased heterogeneity of 3D diffusivities and a decreased cAD (Fig. 6b), as expected. We also show that the cRD increases with increasing axonal radii (Fig. 6a). We observed a 70% difference in cRD between $R = 0.01 \mu\text{m}$ ($\text{cRD} = 7 \times 10^{-5} \text{ mm}^2/\text{s}$) and $R = 4.5 \mu\text{m}$ ($\text{cRD} = 12 \times 10^{-5} \text{ mm}^2/\text{s}$), the largest axonal radius observed in the human brain (62). We demonstrate using *in vivo* data that accounting for each compartment heterogeneity substantially reduces the error we make at predicting the signal (Fig. 7), indicating that DIAMOND better captures the biophysical mechanisms underlying the DW signal formation. We show that the correlation of DIAMOND's tissue parameters (cAD, cRD, cHEI) between two separate acquisitions of a same subject is comparable to that of NODDI for fascicle compartments, even in the case of crossings (Fig. 8). It indicates that these compartment-specific diffusivity characteristics may be used to reliably characterize fascicles.

In contrast to DTI (16) and NODDI (11), we have considered the presence of multiple fascicles per voxel, that is, the presence of multiple anisotropic compartments. Our evaluation shows that the estimated number of fascicles and the estimated fascicle orientations matches the known anatomy (Figs. 9 and 10), leading to improved tractography compared to single-fascicle DTI tractography (Fig. 9h-j). In contrast to CHARMED (14), our results indicate that the fascicle orientations can be estimated with b -values $b \leq 3000 \text{ s/mm}^2$.

Modeling each fascicle in each voxel and assessing their respective diffusion parameters (cFA, cRD, cAD, cM, cHEI) may create a paradigm shift in the way normal and abnormal tissue are characterized. Some approaches such as TRACULA (63) have proposed to delineate WM fascicles using a multi-sticks tractography algorithm (64) and to subsequently characterize WM properties with DTI parameters. Unfortunately, this method remains highly sensitive to partial voluming and to crossings (Fig. 9i) because DTI conflates all the microstructural environments in each voxel. More generally, a large number of DCI models focus on characterizing properties of voxels (e.g., the voxel-wise neurite ODI in Ref. 11). Instead, a *multi-fascicle* DCI model such as DIAMOND aims at shifting the current focus on the voxel to a new focus on each compartment in each voxel, including each fascicle in each voxel, with the ultimate goal of better characterizing *each* neural circuit separately. While this makes whole brain parameter maps more challenging to visually interpret (Fig. 9c–g), group differences can be identified using the Bayesian framework described in (65). We also proposed to assess diffusion characteristics *along* fascicles, by identifying at each voxel the anisotropic compartment most aligned with the local streamlines' orientation. Microstructural properties may further be compared using the fascicle-based spatial statistics framework (66). Finally, we showed that corresponding *fascicle-specific* diffusion properties volumetric maps can be derived as well (Fig. 9j).

In future work, we will investigate and test a DIAMOND fitting procedure that accounts for non-Gaussian noise. We will also investigate the use of simultaneous multi-slice echo-planar imaging (67) to acquire CUSP90 scans in less than 8 min as well as to achieve high-resolution and high-SNR super-resolved DIAMOND imaging in clinically compatible scan time. Finally, we will compare variations of fascicle-specific diffusivity parameters (cAD, cRD, cHEI) in abnormal tissues with histological validation. DIAMOND may enable novel investigations in both normal development and in clinical practice.

REFERENCES

- Norris DG. The effects of microscopic tissue parameters on the diffusion weighted magnetic resonance imaging experiment. NMR Biomed 2001;14:77–93.
- Minati L, Wylam WP. Physical foundations, models, and methods of diffusion magnetic resonance imaging of the brain: A review. Concepts Magn Reson 2007;30A:278–307.
- Mulkern R, Haker S, Maier S. On high b diffusion imaging in the human brain: Ruminations and experimental insights. Magn Reson Imaging 2009;27:1151–1162.
- Sehy JV, Ackerman JJ, Neil JJ. Evidence that both fast and slow water ADC components arise from intracellular space. Magn Reson Med 2004;48:765–770.
- Bennett KM, Schmainda KM, Tong RB, Rowe DB, Lu H, Hyde JS. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. Magn Reson Med 2003;50:727–734.
- Jensen JH, Helpern 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.
- Metzler-Baddeley C, O'Sullivan MJ, Bells S, Pasternak O, Jones DK. How and how not to correct for CSF-contamination in diffusion MRI. Neuroimage 2012;59:1394–1403.
- Wang Y, Wang Q, Haldar JP, et al. Quantification of increased cellularity during inflammatory demyelination. Brain 2011;134:3590–3601.
- Pasternak O, Westin C, Bouix S, et al. Excessive extracellular volume reveals a neurodegenerative pattern in schizophrenia onset. J Neurosci 2012;32:17365–17372.
- Jespersen SN, Kroenke CD, Østergaard L, Ackerman JJ, Yablonskiy DA. Modeling dendrite density from magnetic resonance diffusion measurements. Neuroimage 2007;34:1473–1486.
- 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;61:1000–1016.
- Alexander DC, Hubbard PL, Hall MG, Moore EA, Ptito M, Parker GJ, Dyrbø TB. Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage 2010;52:1374–1389.
- Assaf Y, Blumenfeld-Katzir T, Yovel Y, Basser PJ. AxCaliber: A method for measuring axon diameter distribution from diffusion MRI. Magn Reson Med 2008;59:1347–1354.
- Assaf Y, Basser PJ. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage 2005;27:48–58.
- Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. Brain Res 1992;598:143–153.
- Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J Magn Reson B 1994;103:247–254.
- Jeurissen B, Leemans A, Tournier JD, 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.
- 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;59:2241–2254.
- Ferizi U, Schneider T, Tariq M, Wheeler-Kingshott CAM, Zhang H, Alexander DC. The importance of being dispersed: A ranking of diffusion MRI models for fibre dispersion using in vivo human brain data. In Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), LNCS, Nagoya, Japan, 8149, 2013. pp 74–81.
- Yablonskiy DA, Bretthorst GL, Ackerman JJ. Statistical model for diffusion attenuated MR signal. Magn Reson Med 2003;50:664–669.
- Moseley ME, Cohen Y, Kucharczyk J, Mintorovitch J, Asgari HS, Wendland MF, Tsuruda J, Norman D. Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system. Radiology 1990;176:439–445.
- Jbabdi S, Sotropoulos SN, Savio AM, Graña M, Behrens TE. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. Magn Reson Med 2012;68:1846–1855.
- Basser PJ, Pajevic S. A normal distribution for tensor-valued random variables: Applications to diffusion tensor MRI. IEEE Trans Med Imaging 2003;22:785–794.
- Gupta AK, Nagar DK. Matrix variate distributions. Boca Raton, FL: Chapman & Hall/CRC; 2000.
- Jian B, Vemuri BC, Ozarslan E, Carney PR, Mareci TH. A novel tensor distribution model for the diffusion-weighted MR signal. Neuroimage 2007;37:164–176.
- Leow AD, Zhu S, Zhan L, McMahon K, de Zubicaray GI, Meredith M, Wright MJ, Toga AW, Thompson PM. The tensor distribution function. Magn Reson Med 2009;61:205–214.
- Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magn Reson Med 2006;56:104–117.
- Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Regularized, fast, and robust analytical q-ball imaging. Magn Reson Med 2007;58:497–510.
- Tournier JD, Calamante F, Gadian DG, Connelly A. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage 2004;23:1176–1185.
- Scherrer B, Schwartzman A, Taquet M, Prabhu SP, Sahin M, Akhondi-Asl A, Warfield SK. Characterizing the Distribution of Anisotropic Microstructural Environments with Diffusion-weighted imaging (DIAMOND). In Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Vol. 8151, Nagoya, Japan: Springer; 2013. pp 518–526.

31. Behrens TE, Woolrich MW, Jenkinson M, Johansen-Berg H, Nunes RG, Clare S, Matthews PM, Brady JM, Smith SM. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn Reson Med* 2003;50:1077–1088.
32. Cook PA, Symms M, Boulby PA, Alexander DC. Optimal acquisition orders of diffusion-weighted MRI measurements. *J Magn Reson Imaging* 2007;25: 1051–1058.
33. Scherrer B, Warfield SK. Parametric representation of multiple white matter fascicles from cube and sphere diffusion MRI. *PLoS One* 2012; 7:e48232.
34. Anderson TH. An introduction to multivariate statistical analysis, 3rd ed. Hoboken, NJ: Wiley; 2003.
35. Muirhead RJ. Aspects of multivariate statistical theory. Hoboken, NJ: Wiley; 1982.
36. Mintorovitch, J, Yang, G, Shimizu, Y, Kucharczyk, HJ, Chan, PH, Weinstein, PR. Diffusion-weighted magnetic resonance imaging of acute focal cerebral ischemia: Comparison of signal intensity with changes in brain water and $\text{Na}^+, \text{k}(+)$ -atpase activity. *J Cereb Blood Flow Metab* 1994;14:332–336.
37. Schaefer PW, Grant PE, Gonzalez RG. Diffusion-weighted MR imaging of the brain. *Radiology* 2000;217:331–345.
38. Newcombe VF, Williams GB, Outtrim JG, et al. Microstructural basis of contusion expansion in traumatic brain injury: Insights from diffusion tensor imaging. *J Cereb Blood Flow Metab* 2013;33:855–862.
39. Unterberg AW, Stover J, Kress B, Kiening KL. Edema and brain trauma. *Neuroscience* 2004;129:1021–1029.
40. Szafer A, Zhong J, Gore JC. Theoretical model for water diffusion in tissues. *Magn Reson Med* 1995;33:697–712.
41. Schwartzman A. Random ellipsoids and false discovery rates: Statistics for diffusion tensor imaging data, Ph.D. thesis, USA: Stanford University; 2005.
42. Arsigny V, Fillard P, Pennec X, Ayache N. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magn Reson Med* 2006; 56:411–421.
43. Fillard P, Pennec X, Arsigny V, Ayache N. Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics. *IEEE Trans Med Imaging* 2007;26:1472–1482.
44. Powell MJD. The BOBYQA algorithm for bound constrained optimization without derivatives. In Technical report NA2009/06, Department of Applied Mathematics and Theoretical Physics, Cambridge, England, 2009.
45. Besag J. On the statistical analysis of dirty pictures. *J Roy Statist Soc Ser B* 1986;48:259–302.
46. Scherrer B, Taquet M, Warfield SK. Reliable selection of the number of fascicles in diffusion images by estimation of the generalization error. In: Wells WM, Joshi S, Pohl KM, editors. Information processing in medical imaging (IPMI), LNCS 7917. Berlin: Springer; 2013. pp 742–753.
47. Mills R. Self-diffusion in normal and heavy water in the range 1–45.deg. *J Phys Chem* 1973;77:685–688.
48. Jones DK. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study. *Magn Reson Med* 2004;51:807–815.
49. Hall MG, Alexander DC. Convergence and parameter choice for monte-carlo simulations of diffusion mri. *IEEE Trans Med Imaging* 2009;28:1354–1364.
50. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc* 1983;78:316–331.
51. Reese TG, Heid O, Weisskoff RM, Wedeen VJ. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn Reson Med* 2003;49:177–182.
52. Ruiz-Alzola J, Westin CF, Warfield SK, Alberola C, Maier S, Kikinis R. Nonrigid registration of 3D tensor medical data. *Med Image Anal* 2002;6:143–161.
53. Leemans A, Jones DK. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med* 2009;61:1336–1349.
54. Akhondi Asl A, Warfield SK. Simultaneous truth and performance level estimation through fusion of probabilistic segmentations. *IEEE Trans Med Imaging* 2013;32:1840–1852.
55. Scherrer B, Gholipour A, Warfield SK. Super-resolution reconstruction to increase the spatial resolution of diffusion weighted images from orthogonal anisotropic acquisitions. *Med Image Anal* 2012;16: 1465–1476.
56. Andersson JL, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *Neuroimage* 2003;20:870–888.
57. Holland D, Kuperman JM, Dale AM. Efficient correction of inhomogeneous static magnetic field-induced distortion in echo planar imaging. *NeuroImage* 2010;50:175–183.
58. Wishart J. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* 1928;20A:32–52.
59. Raffelt D, Tournier JD, Rose S, Ridgway GR, Henderson R, Crozier S, Salvado O, Connell A. Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *Neuroimage* 2012;59:3976–3994.
60. Dell’acqua F, Simmons A, Williams SC, Catani M. Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. *Hum Brain Mapp* 2013;34: 2464–2483.
61. Schultz T, Gröschel S. Auto-calibrating spherical deconvolution based on odf sparsity. In Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Vol. 8149, Nagoya, Japan: Springer; 2013. pp 663–670.
62. Liewald D, Miller R, Logothetis N, Wagner HJ, Schüz A. Distribution of axon diameters in cortical white matter: An electron-microscopic study on three human brains and a macaque. *Biol Cybern* 2014;108: 541–557.
63. Yendiki A, Panneck P, Srinivasan P, et al. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front Neuroinform* 2011;5:23.
64. Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* 2007;34:144–155.
65. Taquet M, Scherrer B, Peters JM, Prabhu SP, Warfield SK. A fully bayesian inference framework for population studies of the brain microstructure. In Proceedings of the 17th Int Conf Med Image Comput Comput Assist Interv (MICCAI), Boston, USA, no. 8673, 2014. pp 25–32.
66. Taquet M, Scherrer B, Commissick O, Peters JM, Sahin M, Macq B, Warfield SK. A mathematical framework for the registration and analysis of multi-fascicle models for population studies of the brain microstructure. *IEEE Trans Med Imaging* 2014;2:504–517. doi: 10.1109/TMI.2013.2289381.
67. Setsompop K, Cohen-Adad J, Gagoski BA, Raji T, Yendiki A, Keil B, Wedeen VJ, Wald LL. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *Neuroimage* 2012;63:569–580.