Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison

Eleftheria Panagiotaki [a,*], Torben Schneider [b], Bernard Siow [a,c], Matt G. Hall [a], Mark F. Lythgoe [c], Daniel C. Alexander [a]

[a] Centre for Medical Image Computing, Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK
[b] NMR Research Unit, Department of Neuroinflammation, UCL Institute of Neurology, University College London, WC1N 3BG, UK
[c] Centre for Advanced Biomedical Imaging, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

ABSTRACT

This paper aims to identify the minimum requirements for an accurate model of the diffusion MR signal in white matter of the brain. We construct a taxonomy of multi-compartment models of white matter from combinations of simple models for the intra- and the extra-axonal spaces. We devise a new diffusion MRI protocol that provides measurements with a wide range of imaging parameters for diffusion sensitization both parallel and perpendicular to white matter fibres. We use the protocol to acquire data from two fixed rat brains, which allows us to fit, study and compare the different models. The study examines a total of 47 analytic models, including several well-used models from the literature, which we place within the taxonomy. The results show that models that incorporate intra-axonal restriction, such as ball and stick or CHARMED, generally explain the data better than those that do not, such as the DT or the biexponential models. However, three-compartment models which account for restriction parallel to the axons and incorporate pore size explain the measurements most accurately. The best fit comes from combining a full diffusion tensor (DT) model of the extra-axonal space with a cylindrical intra-axonal component of single radius and a third spherical compartment of non-zero radius. We also measure the stability of the non-zero radius intra-axonal models and find that single radius intra-axonal models are more stable than gamma distributed radii models with similar fitting performance.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) uniquely captures the entangled structural organisation of brain white matter, identifying changes and alterations in microstructure. DW-MRI has become popular because it provides unique insight into brain tissue microstructure and connectivity non-invasively. The technique measures the displacement of particles, usually water molecules, which is affected by the presence of the surrounding structure in the environment, thus DW-MRI can be used to make inferences about the local structure. The simplest and most commonly used model for relating the DW-MRI signal to diffusion in tissue is the diffusion tensor (DT) (Basser et al., 1994). The model provides useful microstructural markers of tissue integrity such as fractional anisotropy (FA) and mean diffusivity (MD). However these indices lack specificity, because many features of the microstructure can affect them. To address this limitation, the recent trend in the diffusion MRI field (Alexander, 2008; Alexander et al., 2010; Assaf and Basser, 2005; Assaf et al., 2008; Barazany et al., 2009; Behrens et al., 2003; Jespersen et al., 2007, 2010; Zhang et al., 2011a,b) is towards more

direct microstructure imaging by fitting more descriptive models of tissue that relate specific parameters, such as cell size and density, directly to the signal.

Diffusion tensor imaging (DTI) (Basser et al., 1994) models the displacement of spins in 3D with a zero-mean Gaussian distribution by fitting the apparent diffusion tensor to six or more normalised DW images. A significant limitation of the DT model is that it does not account for restricted diffusion within cells so the signal departs from the model even in single fibre populations, especially as the diffusion weighting factor b becomes large. Multi-compartment models overcome some of the limitations of the DT model. Often they assume hindered diffusion in the extra-axonal space and restricted diffusion in the intra-axonal space.

Behrens et al. (2003) propose a simple alternative to the DT model that is able to account for one or more fibre directions within each voxel. They propose a partial volume model of local diffusion, the "ball and stick", which assumes that diffusion within axons is along a single direction and that outside axons is isotropic.

Alexander (2008) investigates the feasibility of using diffusion MRI to measure axon dimensions in white matter in vivo. To relate axon radius to the diffusion MRI signal he uses a two-compartment model that incorporates a pore radius. The model of the extra-axonal space uses a cylindrically symmetric tensor with the principal eigenvector indicating

the fibre direction. The model for the intra-axonal space is a cylinder with non-zero pore size, unlike Behrens' stick model, which is a cylinder with radius zero. The signal model uses the Gaussian phase distribution (GPD) approximation (Douglass and McCall, 1958) for the signal in cylinders of radius R (Gelderen et al., 1994; Stepisnik, 1993). The cylinder axis aligns with the principal eigenvector of the extra-axonal DT.

The model in Alexander (2008) is a simplification of Assaf's CHARMED model (Assaf and Basser, 2005), which also assumes cylindrical restriction in the intra-axonal space. They use Neuman's expression (Neuman, 1974) for diffusion in cylindrical confinement for a pulse-gradient spin-echo (PGSE) experiment, which satisfies the Short Gradient Pulse (SGP) approximation (Price, 1997). Unlike Alexander's model, which assumes a single cylinder radius, the model assumes cylinders with gamma-distributed radii, which introduces one extra parameter, although in Assaf and Basser (2005) both gamma parameters are fixed rather than estimated. The hindered compartment uses the full DT model constrained only to have principal direction aligned with the cylinder axis. In later work, Assaf et al. (2008) use this model in an NMR experiment to estimate distributions of axon diameters of bovine optic and sciatic nerve samples. They extend the CHARMED framework, by considering the diameter distribution of cylindrical axons as an unknown to be estimated from their data. They name this method AxCaliber. To gain sensitivity for estimating the axon diameter they use a fixed gradient direction perpendicular to the axons with a combination of different diffusion times and gradient strengths.

Other methods describe diffusion with three or more compartments and allow exchange between them. For example, Barazany et al. (2009) study the diameter distribution of axons in the rat corpus callosum in vivo using a three-compartment model. The model is an extension of the AxCaliber model (Assaf et al., 2008) with the addition of an isotropic-diffusion compartment to account for partial volume contributions from areas of cerebrospinal fluid (CSF).

Stanisz et al. (1997) construct a three-pool model with prolate ellipsoidal axons and spherical glial cells each with partially permeable membranes. Fitting such a complex model requires very high quality measurements, typically using NMR spectroscopy rather than MRI. Experiments on such a high quality data set however suggest that the full complexity of the model is required to explain the data.

Recently, Alexander et al. (2010) demonstrate orientationally invariant estimation of axon diameter and density in both fixed monkey brains and in vivo human brains with their ActiveAx technique. They use a four-compartment tissue model: a combination of the two-compartment model used in Alexander (2008) with a CSF compartment similar to Barazany et al. (2009) and a fourth compartment which accounts for water coming from stationary water molecules trapped in glial cells and other subcellular structures similar to Stanisz et al. (1997).

To trust the parameter estimates we obtain from fitting these models, we must ensure that the models include all the important physiological parameters of the tissue that affect the signal. The aim of this work is thus to compare a range of models to identify those that explain signals from brain white matter tissue the best. We focus on multi-compartment models with no exchange, which model the signal as the sum of signals from water inside and outside structures with impermeable walls. The class of models includes the ball and stick model (Behrens et al., 2003), a bitensor model (Alexander and Barker, 2005; Parker and Alexander, 2003; Peled et al., 2006; Tuch et al., 2002), CHARMED as used in AxCaliber (Assaf et al., 2008), the simplified CHARMED model in Alexander (2008) and the minimal model of white matter diffusion (MMWMD) in Alexander et al. (2010). Each model uses a different combination of axonal, extra-axonal and other compartments. We construct a taxonomy of multi-compartment models that includes these models as well as the extended set of combinations of compartments. Preliminary work studied two-compartment models (Panagiotaki et al., 2009).

Here we extend to include three compartments, which proves necessary to capture observed trends in the signal.

We compare models within the taxonomy experimentally to determine which explain the diffusion MR signal in white matter the best. We also include the DT model using linear and non-linear fitting in our comparison (Basser et al., 1994). In total we compare 47 models of the diffusion MRI signal in brain white matter. We define a diffusion MRI protocol to allow evaluation and comparison of the models for parallel and perpendicular signals in brain white matter over a wide range of scan parameters. Such a study is challenging in brain tissue because fibre orientation varies, so most previous studies use simpler tissue samples such as spinal cord (Assaf et al., 2008) and excised nerve (Stanisz et al., 1997). We acquire data from two ex vivo rat brains that have been perfusion fixed. The data sets contain a much more comprehensive set of measurements than we can acquire on live subjects but is rich enough to get good estimates from complex models, thus to identify important effects, and provide reliable selection of appropriate models eventually to use with sparser in vivo data.

The paper structure is as follows: the Methods section details the tissue models and constructs the model taxonomy. It also details the tissue sample preparation, MRI acquisition, model fitting and parameter estimation procedures. Experiments show microstructure parameter estimation from experiments on fixed rat brains and evaluation of the models using the Bayesian information criterion (BIC). The final section concludes and discusses limitations, applications, and areas for future work.

Methods

This section describes the tissue models and the tissue samples. It then details the MRI acquisition and the model-fitting procedure.

Tissue models

We model brain white matter with combinations of three types of compartments. Each of the compartments provides a separate normalised MR signal S_1, S_2, S_3. The signals come from a) water inside the axons, b) water from outside the axons and c) water from other cellular structures, such as glial cells, trapped water on membranes or non-parallel fibres. We refer to these three types of compartments as a) intra-axonal, which are restricted models of diffusion, b) extra-axonal, which include isotropic and anisotropic non-restricted models and c) compartments of isotropic restriction.

The total diffusion MR signal for a multi-compartment model is

$$S = \sum_{i=0}^{3} f_i S_i, \tag{1}$$

where f_i is the proportion of unweighted signal from water molecules in population $i, 0 \leq f_i \leq 1$ and

$$\sum_{i=0}^{3} f_i = 1. \tag{2}$$

Intra-axonal compartments

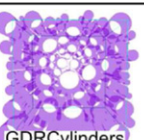We investigate three models for the intra-axonal compartment, as shown in Table 1:

1. The first is Behrens' "stick" model (Behrens et al., 2003) which has fibre direction $\mathbf{n}$ and diffusivity d as parameters. The "stick" model describes diffusion in an idealised cylinder with zero radius. The signal for the model is

$$S_1 = \exp\left(-bd(\mathbf{n} \cdot \mathbf{G})^2\right) \tag{3}$$

Table 1
Intra-axonal compartment models. The fibre orientation $\mathbf{n}$ is defined by the angles θ, φ.

Intra-Axonal compartments		
Model	Form	Degrees of freedom
Stick	$S = \exp(-bd(\mathbf{n} \cdot \mathbf{G})^2)$	d, θ, ϕ
Cylinder	GPD approx.	d, θ, ϕ, R
GDRCylinders	$P(R; k, \vartheta) = \frac{R^{k-1} e^{\frac{-R}{\vartheta}}}{\Gamma(k)\vartheta^k}$	$d, \theta, \phi, k, \vartheta$

where $b = (\Delta - \delta/3)(\gamma \delta |\mathbf{G}|)^2$ for the PGSE sequence which we use throughout this work, Δ is the time between the onsets of the two pulses, δ is the pulse gradient duration, γ is the gyromagnetic ratio and $\mathbf{G}$ is the gradient vector.

2. The second is the "cylinder" model, as used in Alexander (2008), Alexander et al. (2010), which accounts for non-zero cylinder radius. This model has an extra parameter, R, representing a single axon radius. We use the GPD approximation for perpendicular signal for this model assuming the PGSE sequence (Stepisnik, 1993). Appendix A provides the signal for the "cylinder" model.

3. The third model is an extension of the "cylinder" but instead of a single radius it has gamma-distributed radii as used in Assaf and Basser (2005), Assaf et al. (2008) and Barazany et al. (2009). The radii are drawn from a gamma distribution

$$P(R; k, \vartheta) = \frac{R^{k-1} e^{\frac{-R}{\vartheta}}}{\Gamma(k)\vartheta^k}, \tag{4}$$

where Γ is the gamma function. We refer to this model as "GDRCylinders". Instead of the single parameter R, this model has two parameters to define the cylinders' radii, the shape parameter k and the scale parameter ϑ, where $k\vartheta$ is the mean and $k\vartheta^2$ is the variance.

Extra-axonal compartments

The candidate models for the extra-axonal compartment assume hindered diffusion and provide signal S_2. We investigate three models for the extra-axonal compartment. We illustrate the models in Table 2. Each one is a diffusion tensor (DT) model with different constraints. The normalised DT signal is

$$S_2 = \exp\left(-b\hat{\mathbf{G}}^T \mathbf{D} \hat{\mathbf{G}}\right), \tag{5}$$

where $\hat{\mathbf{G}}$ is the gradient direction. The candidate extra-axonal models are as follows:
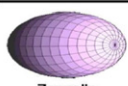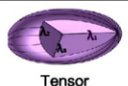
1. The first, referred to as a "ball", is isotropic, as in Behrens et al. (2003). The model has only one parameter, the diffusivity d, so that

$$\mathbf{D} = d\mathbf{I}, \tag{6}$$

where $\mathbf{I}$ is the identity tensor.

Table 2
Extra-axonal compartment models.

Intra-Axonal compartments		
Model	Form	Degrees of freedom
Ball	$\mathbf{D} = d\mathbf{I}$	d
Zeppelin	$\mathbf{D} = \alpha \mathbf{n}\mathbf{n}^T + \beta \mathbf{I}, \; d_\parallel = \alpha + \beta, d_\perp = \beta$	$d_\parallel, d_\perp, \theta, \phi$
Tensor	$\mathbf{D} = d_\parallel \mathbf{n}\mathbf{n}^T + d_{\perp_1} \mathbf{n}_{\perp_1} \mathbf{n}_{\perp_1}^T + d_{\perp_2} \mathbf{n}_{\perp_2} \mathbf{n}_{\perp_2}^T$	$d_\parallel, d_{\perp_1}, d_{\perp_2}, \theta, \phi, \alpha$

2. The second is an anisotropic, but cylindrically symmetric DT, as in Alexander (2008) and we refer to this as a "zeppelin". The "zeppelin" has

$$\mathbf{D} = \alpha \mathbf{n}\mathbf{n}^T + \beta \mathbf{I}, \tag{7}$$

where $d_\parallel = \alpha + \beta$. The model has the following parameters: the principal direction $\mathbf{n}$, $d_\parallel$ which is the diffusivity parallel to the principal direction, and $d_\perp = \beta$ perpendicular to the principal direction.

3. Finally we consider a full tensor, which we refer to as a "tensor". This model has three diffusivity parameters: parallel diffusivity $d_\parallel$ and $d_{\perp_1}$, $d_{\perp_2}$ perpendicular with $d_{\perp_1} \neq d_{\perp_2}$. It has an additional three degrees of freedom for the orthogonal eigenvectors $\mathbf{n}$, $\mathbf{n}_{\perp_1}$ and $\mathbf{n}_{\perp_2}$. The form of the DT is:

$$\mathbf{D} = d_\parallel \mathbf{n}\mathbf{n}^T + d_{\perp_1} \mathbf{n}_{\perp_1} \mathbf{n}_{\perp_1}^T + d_{\perp_2} \mathbf{n}_{\perp_2} \mathbf{n}_{\perp_2}^T. \tag{8}$$

Isotropic restriction compartments

We consider four extra models intended to capture other cellular structures that can be combined with two-compartment models from combinations of the intra- and extra-axonal models. All these models assume isotropic restriction. Two of the models, the "astrosticks" and the "astrocylinders" (see list below and Table 3), assume restriction from isotropically oriented cylinders. The other two, the "sphere" and "dot" (see list below and Table 3), assume restriction from spherical boundaries. The models of isotropic restriction are as follows:

1. The "astrosticks" model has cylinders with uniformly distributed orientation and zero diameter. The model represents signal coming from a type of glial cell called astrocytes, or populations of axons with arbitrary orientation.

2. The "astrocylinders" model has similar motivation to "astrosticks", assuming particles diffusing in a cylinder, but this time with non-zero radius R.

3. The "sphere" model has particles diffusing inside impermeable spherical boundaries with non-zero radius R_s. The "sphere" models signal coming from water molecules inside spherical glial cells as in Staniz et al. (1997).

4. The "dot" model is a special case of the "sphere" with zero radius in which particles do not move, as used in Alexander et al. (2010). Alexander et al. (2010) suggests that the "dot" models

Table 3
Isotropically restricted compartment models.

Isotropic restriction compartments		
Model	Form	Degrees of freedom
Astrosticks	$S_G = \int S_r p(\mathbf{n}) d\mathbf{n}, R = 0$	d
Astrocylinders	$S_G = \int S_r p(\mathbf{n}) d\mathbf{n}, R > 0$	d, R
Sphere	GPD approx. $R_s > 0$	d, R_s
Dot	$R = 0, S = 1$	–

Fig. 1. Plot of the synthesised diffusion signals from the "dot", "astrosticks", "astrocylinders" and "sphere" models using the imaging protocol described in the section MRI acquisition. We use $d_\parallel = 8 \times 10^{-10}$ m²/s and $R = Rs = 4 \times 10^{-6}$ m.

signal coming from molecules trapped in glial cells or stuck into cellular membranes in fixed tissue.

Appendix A derives expressions for the astrosticks and astrocylinders models. Next we present a summary for the signal of each isotropic restriction model:

1. The signal from "astrosticks" S_3 comes from substituting Eq. (A.3) and $L_\perp = 0$ in Eq. (A.7). The model has only one parameter: the diffusivity $d_\parallel$.
2. The signal from "astrocylinders" S_3 comes from substituting Eqs. (A.3) and (A.5) in Eq. (A.7). The model has two parameters: the diffusivity $d_\parallel$ and the radius R.
3. To model the signal from particles trapped inside the "sphere" we use Murday and Cotts (1984) expression, using the GPD approximation. The model has two parameters: the diffusivity $d_\parallel$ and the radius R_s.
4. The signal from the "dot" remains unattenuated so $S_3 = 1$ (no parameters).

Fig. 1 plots the synthesised diffusion signals from the "astrosticks", "astrocylinders", "dot" and the "sphere" models using the imaging protocol described in the section MRI acquisition.

Naming system

We use combined terms to refer to specific two- and three-compartment models. For example "ZeppelinStick" assumes zero radius cylinders for the intra-axonal space, cylindrical symmetry for the extra-axonal space and no isotropically restricted compartment. The taxonomy includes models from previous studies such as "BallStick" (Behrens et al., 2003), CHARMED (Assaf and Basser, 2005; Assaf et al., 2008) as used in AxCaliber, which is the "TensorGDRCylinders" model, the simplified version of CHARMED as used in Alexander (2008), which is the "ZeppelinCylinder" model, and the MMWMD in Alexander et al. (2010) without the CSF compartment, which is the "ZeppelinCylinderDot" model.
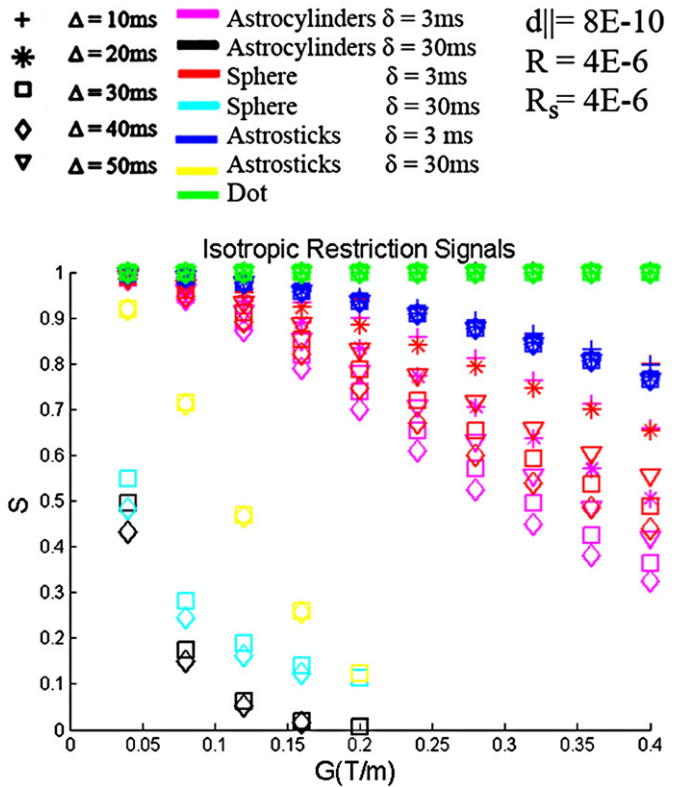
In this paper we study all the combinations of the two- and three-compartment models, as well as the DT model and a bitensor model. The single DT is the same as the extra-axonal "tensor" compartment, and the bitensor model is a mixture of two "zeppelin" compartments with the same principal direction, which we name "Bizeppelin". In total we consider 47 models.

Tissue samples

The brains of two male Sprague–Dawley rats were perfusion fixed through the left ventricle using isotonic saline, followed by 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). The rat brains were extracted and immersion-fixed in fresh 4% PFA for at least two weeks. The T_2 reduction due to fixation in PFA was reversed by immersion in 4% PBS with 1% Penicillin Streptomycin solution for two weeks (Shepherd et al., 2009). Finally, each brain was embedded in 1% agarose gel in PBS to prevent motion during the DW-MRI acquisition.

MRI acquisition

We are interested in diffusion in directions parallel and perpendicular to the fibre orientation, since these directions reveal most about the underlying brain microstructure. However, it is challenging to acquire such measurements from brain tissue, because fibre orientation varies throughout the white matter. Our approach is to select one central parallel direction and several perpendicular directions and identify voxels in which the fibres align with the central parallel direction after imaging. We focus analysis only on these voxels and discard all others.

We acquire diffusion-weighted MR images of two perfusion-fixed rat brains, using a small bore 9.4T Varian Inc. scanner equipped with gradients capable of 1 T/m and a 26 mm Rapid Biomedical GmBH
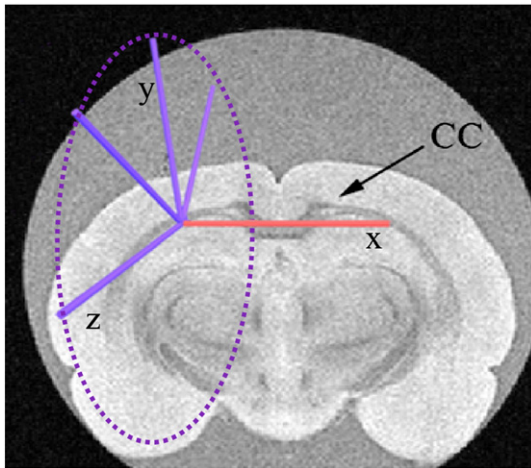
Fig. 2. Diffusion-weighted MR image of a fixed rat brain. The red bar indicates the central gradient direction used for the encoding scheme and the blue bars indicate the four directions perpendicular to the central one.

birdcage r.f. coil. We use a five-direction encoding-scheme and place the sample in the scanner oriented to ensure that some fibres in the corpus callosum (CC) are parallel to our central direction. We choose four evenly spaced directions perpendicular to the central direction in our chosen voxels (see Fig. 2). We use the PGSE sequence for 70 diffusion weightings: five diffusion times $\Delta = 10, 20, 30, 40, 50$ ms, gradient durations $\delta = 3$ ms for all Δ and $\delta = 30$ ms for $\Delta = 40, 50$ ms, gradient strength $|\mathbf{G}|$ varied from 40 to 400 mT/m in ten steps of 40 mT/m. Measurements with b value greater than 7.7×10^{10} sm^{-2} were not performed due to poor SNR (<2.6), i.e., all combinations with $|\mathbf{G}| = 200$ to 400 mT/m and $\delta = 30$ ms for both rat brains. In total we acquired images with 59 diffusion weightings in each direction. We refer to the first rat brain dataset as dataset A and the second as dataset B. In Fig. 3 we plot the parallel and the mean of the four perpendicular directions of the log normalised signal from voxels in the region of interest in the CC and show MR images for various b values from both datasets.

We use minimum echo times (TE) to maximise SNR and choose repetition times (TR) long enough to ensure no gradient heating effects. For each combination of diffusion-weighting parameters we acquire separate $b = 0$ images to correct for T_1 and T_2 dependence. We also perform a separate DTI acquisition using a 42-direction scheme with b value 4.5×10^9 sm^{-2}, as suggested by Dyrby et al. (2011) for

fixed brain tissue, and six $b = 0$ measurements. The imaging parameters for the DTI and DW-MRI acquisitions are as follows: in-plane field of view $= 2 \times 2$ cm, matrix size $= 256 \times 256$, 10 slices, slice thickness 0.5 mm. In total we acquire 343 images in approximately 65 h for each sample.

Since water diffusion in tissue is highly temperature sensitive, it is critical to avoid temperature changes. We use an airflow mechanism to maintain the samples' temperature during the acquisitions. To ensure that the sample was scanned at a consistent temperature while using a wide range of b values, we test the temperature fluctuation at 1.5 times the maximum b value of our imaging protocol. Results show that the temperature at which it reached a plateau was less than 1° above the 'baseline' start temperature. The temperature measurements in the protocol with the maximum gradient duty cycle provide further evidence that the temperature did not fluctuate more than ± 1 °C.

Model fitting

We fit each model to the data using an iterative optimization procedure and synthesise diffusion-weighted data from the fitted models. The model parameters do not relate linearly to the measurements, therefore we have to fit the models to the data by non-linear optimization. Here we use a Levenberg–Marquardt algorithm (Press et al., 1988) for minimising a chi-squared objective function with the offset Gaussian noise model, which adds an expected offset to each measurement from Rician noise bias (Alexander, 2009; Jones and Basser, 2004). The objective function is:

$$L_{offG} = \sum_{n=1}^{M} \left(\tilde{S}_n(\delta, \Delta, \mathbf{G}) - \sqrt{S_n(\delta, \Delta, \mathbf{G})^2 + \sigma^2} \right)^2, \tag{9}$$

where M is the number of measurements, $S_n(\delta, \Delta, \mathbf{G})$ is the model predicted signal and $\tilde{S}_n(\delta, \Delta, \mathbf{G})$ is the signal for the nth measurement.

Fits of the simplest models are relatively independent of starting position. More complex models are more sensitive to the starting position therefore we use parameter estimates of simpler models to provide initial estimates. For example, to get the starting point for the three-compartment "TensorGDRCylindersSphere" we use the estimate from the two-compartment "TensorCylinder". In turn, the starting point for "TensorCylinder" uses fits from two simpler models: the "BallCylinder" and the "TensorStick". The "TensorStick" starting point depends on the "BallStick" and the linear DT estimation. The "BallCylinder" depends on the "BallStick". Finally the "BallStick"
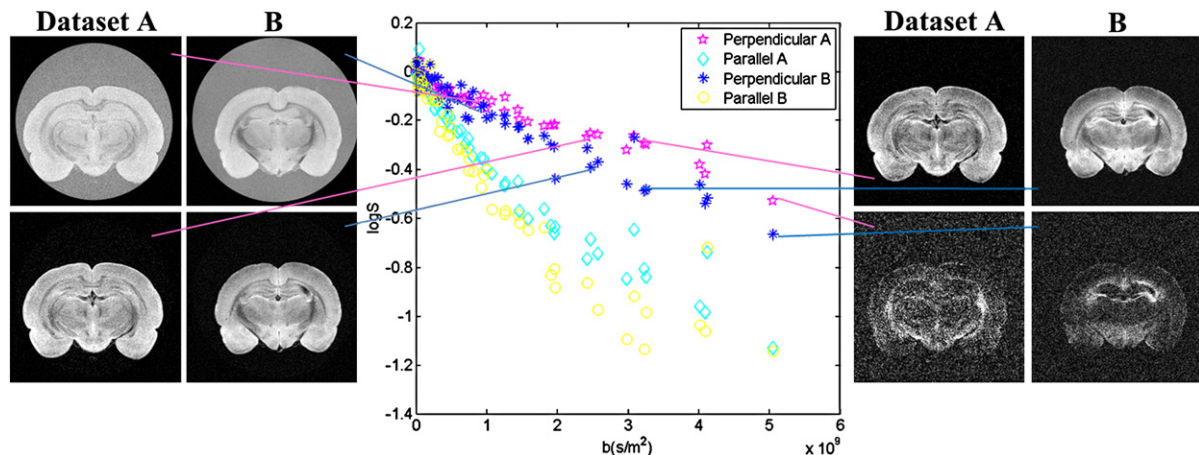
Fig. 3. Plot of the parallel and the mean of the four perpendicular directions of the log signal from voxels in a region of interest in the CC and demonstration of MR images (perpendicular direction) for various b values from both datasets. Top left example has $\Delta = 10$ ms, $\delta = 3$ ms,$|G| = 360$ mT/m, bottom left example has $\Delta = 30$ ms,$\delta = 3$ ms,$|G| = 360$ mT/m, top right example has $\Delta = 40$ ms,$\delta = 30$ ms,$|G| = 40$ mT/m, bottom right example has $\Delta = 50$ ms,$\delta = 30$ ms,$|G| = 40$ mT/m.

depends on the estimate of the linear DT estimation. Despite the simplicity of the "BallStick" model, we still initialise it using the DT model, since the fibre orientation parameter occasionally ends in a local minimum with the non-linear optimization from random starting points. We do not list full initialization details for all models here, but more information is in Panagiotaki (2011) or in the implementation in the open source Camino toolkit (Cook et al., 2006).[1]

For all models, we choose the best fit parameters from the models after 1000 perturbations of the starting parameters to ensure a good minimum. We also calculate the histogram of the objective function over all 1000 runs to monitor the frequency of the best parameter estimation and the stability of the fitting procedure.

Parameter constraints

We constrain a set of model parameters in the optimization using transformations that limit the range of each parameter to biophysical meaningful values. All the parameter transformations are presented in Appendix B. For the "cylinder" model we constrain the single axon radius R to be within biologically plausible limits so that $0.1\ \mu m \leq R \leq 20.1\ \mu m$. For the "GDRCylinders" model we constrain the scale parameter ϑ so that $0.1\ \mu m \leq \vartheta \leq 20\ \mu m$ and allow a broad range for the shape parameter k so that $0 \leq k \leq 20$. For the isotropic restriction model "astrocylinders" we constrain the radius to be the same as the intra-axonal "cylinder" radius when the model includes that parameter. For the "sphere" model we allow the sphere radius R_s twice the limit of the axon radius R so that $0.1\ \mu m \leq R_s \leq 40.1\ \mu m$. For all the compartments we ensure that the parallel diffusivity $d_\parallel$, and perpendicular diffusivities $d\perp_1$, $d\perp_2$ are always positive. For the extra-axonal "zeppelin" we constrain the perpendicular diffusivity $d\perp \leq d_\parallel$, and for the "tensor" the perpendicular diffusivity $d\perp_2 \leq d\perp_1 \leq d_\parallel$. We constrain all volume fractions to be in $[0,1]$ and sum to 1. Finally, for the "tensor" and "zeppelin" models, we constrain the principal direction to the orientation of the intra-axonal compartment.

Experiments and results

To study parallel and perpendicular signal attenuation we choose a set of voxels with fibre direction in alignment with the central direction. To achieve this, we manually segment the CC on a FA map from the DTI acquisition and threshold for voxels with FA >0.5 in which the principal direction of the DT is parallel to the central gradient direction within a tolerance of 2°. We average the data contained within all the resulting 21 and 36 voxels from datasets A and B respectively. The intention is to obtain a high signal to noise set of measurements with varying diffusion parameters representative of coherent white matter.

Model ranking

The Bayesian information criterion (BIC) (Schwarz, 1978) evaluates the models. BIC chooses the most economical analytic model by rewarding those that minimise the objective function, while simultaneously penalising increasing numbers of model parameters:

$$BIC = -2ln(L) + kln(n), \qquad (10)$$

where L is the likelihood of the estimated model, n is the number of measurements and k is the number of free model parameters to be estimated. The model which provides the lower value of the BIC is the one to be preferred. Table 4 presents the ranking of the models, according to BIC for datasets A and B.

Table 4
Table of the models ranked by the Bayesian information criterion (BIC) for dataset A, the ranking of the models according to BIC for dataset B and the number of parameters (counting S_0) for each model.

Models	BIC				No. parameters
	Dataset A $\times 10^3$	Rank A	Dataset B $\times 10^3$	Rank B	
TensorCylinderSphere	−2.587	1	−2.636	1	11
TensorStickSphere	−2.558	2	−2.586	2	10
ZeppelinCylinderSphere	−2.495	3	−2.542	5	9
ZeppelinGDRCylindersSphere	−2.489	4	−2.507	7	10
ZeppelinStickSphere	−2.482	5	−2.488	9	8
TensorCylinderDot	−2.481	6	−2.561	3	10
TensorGDRCylindersDot	−2.446	7	−2.557	4	11
TensorGDRCylindersSphere	−2.431	8	−2.518	6	12
ZeppelinCylinderDot	−2.429	9	−2.499	8	8
ZeppelinGDRCylindersDot	−2.412	10	−2.484	10	9
BallGDRCylindersDot	−2.382	11	−2.466	11	8
BallCylinderSphere	−2.377	12	−2.448	13	8
BallCylinderDot	−2.375	13	−2.454	12	7
BallGDRCylindersSphere	−2.366	14	−2.420	14	9
ZeppelinStickDot	−2.329	15	−2.396	15	7
TensorCylinderAstrocylinders	−2.290	16	−2.331	17	10
TensorCylinderAstrosticks	−2.289	17	−2.331	18	10
TensorGDRCylindersAstrosticks	−2.286	18	−2.329	19	11
TensorStickAstrocylinders	−2.282	19	−2.324	20	10
TensorStickAstrosticks	−2.279	20	−2.324	21	9
ZeppelinGDRCylindersAstrocylinders	−2.278	21	−2.230	36	9
BallStickSphere	−2.272	22	−2.332	16	7
BallCylinderAstrocylinders	−2.243	23	−2.294	23	7
BallCylinderAstrosticks	−2.239	24	−2.294	24	7
ZeppelinCylinderAstrocylinders	−2.238	25	−2.288	25	8
BallStickAstrocylinders	−2.237	26	−2.285	29	7
ZeppelinCylinderAstrosticks	−2.234	27	−2.288	26	8
TensorStickDot	−2.233	28	−2.276	33	9
BallStickAstrosticks	−2.232	29	−2.28	30	6
BallGDRCylindersAstrosticks	−2.231	30	−2.282	31	8
ZeppelinStickAstrocylinders	−2.231	31	−2.280	31	8
ZeppelinStickAstrosticks	−2.226	32	−2.279	32	7
BallGDRCylindersAstrocylinders	−2.219	33	−2.271	35	8
BallStickDot	−2.218	34	−2.274	34	6
TensorGDRCylindersAstrocylinders	−2.209	35	−2.314	22	11
ZeppelinGDRCylindersAstrosticks	−2.174	36	−2.286	28	9
TensorCylinder	−2.066	37	−2.130	37	9
TensorGDRCylinders	−2.051	38	−2.107	38	10
BallCylinder	−2.050	39	−2.103	39	6
ZeppelinCylinder	−2.044	40	−2.101	40	7
BallGDRCylinders	−2.044	41	−2.100	41	7
TensorStick	−2.043	42	−2.100	42	8
ZeppelinGDRCylinders	−2.040	43	−2.097	43	8
Bizeppelin	−2.028	44	−2.085	44	7
BallStick	−2.020	45	−2.067	45	5
ZeppelinStick	−2.019	46	−2.065	46	6
DT	−1.597	47	−1.677	47	7

The ranking of the models according to BIC is similar for datasets A and B: both the 2 highest-performing and the 11 lowest-performing models are identical between the two datasets. The ranking of the remaining models exhibits only small variations between datasets. The "sphere" and "dot" models, consistently perform the best. The first 12 highest-ranked models in both datasets are combined with non-zero radius intra-axonal compartments, with the exception of the "TensorStickSphere" and "ZeppelinStickSphere". We also observe that the extra-axonal compartments of the "tensor" and the "zeppelin" are favoured over the isotropic "ball" compartment ranking these models also in the top 10 best-performing. Following the "sphere" and "dot" models, are the "astrocylinders". The models that are ranked the lowest, are the DT and two-compartment models with the intra-axonal "stick" model. Another important observation is that the "Bizeppelin" model is ranked very low as well, highlighting the importance of restriction in describing these data. Finally, the "TensorCylinder" minimises the BIC for the two-compartment models and the "TensorCylinderSphere" minimises the BIC overall.

---

[1] www.camino.org.uk.

Parameter estimation

The best fit microstructure parameters for a selection of models are summarized in Table 5. The parameter estimates for the full set of models are provided as on line material. Here we concentrate on a set of key models including representatives from the literature and models that perform particularly well or poorly in our comparison. We will discuss the variation of fitted estimates of corresponding parameters across models in turn, first for two-compartment models, then three-compartment.

In the class of two-compartment models, we see that "cylinder" models give higher values for the intra-axonal volume fraction f_1 than the "stick" models. The "GDRCylinders" models give higher values for f_1 than the "cylinder" models. This is as expected, because "GDRCylinders" accommodates more intra-axonal mobility than "cylinder", which accommodates more than "stick". Thus for fixed perpendicular apparent diffusion coefficient (ADC), f_1 increases from "stick" to "cylinder" to "GDRCylinders". The "cylinder" models consistently estimate R around 2 μm. To compare the radius estimate between the "GDRCylinders" and the "cylinder" model we calculate the mean, $k\vartheta$, and compare it to the R. The "GDRCylinders" models predict $k\vartheta$ around 2.5 μm, a little higher than the "cylinder" models'R. Both non-zero radius intra-axonal models estimate the diffusivity $d_\parallel$ at similar levels, while they estimate it higher than the "stick", which again is a likely consequence of the "cylinder" models accommodating intra-axonal perpendicular mobility.

Considering three-compartment models, we start with general trends independent of the choice of third compartment, then make comparisons between choices of third compartment. The additions of any of the third compartments retain the same trends in f_1 and $d_\parallel$ as two-compartment models: f_1 and $d_\parallel$ increase for models that accommodate intra-axonal mobility. The estimates for the volume fraction of the third compartment, f_3, decrease as the mobility in the intra-axonal compartments increases. Also, the two-compartment models estimate $d_\parallel$ lower than all the three-compartment models. This makes sense, because the two-compartment models can only accommodate restriction parallel to the fibres by reducing the $d_\parallel$.

Regarding the "dot" models, "cylinder" models estimate the radius R around 3.8 μm, while "GDRCylinders" models give a higher estimate at approximately 4.4 μm. The increased $d_\parallel$ estimate compared to two-compartment models corresponds with the increase in R.

However, the "astrosticks" models, "cylinder" and "GDRCylinders" models estimate the axon radius R around 2 μm, in agreement with

the two-compartment models despite still having higher $d_\parallel$. The "astrocylinders" models present the same trends for all the parameter estimates as "astrosticks". However, neither "astro" model seems to explain the data as well as the spherical third compartments.

As for the three-compartment "sphere" models, the radius of the "cylinder" and "GDRCylinders" models is estimated at approximately 5.8 μm. The radius of the sphere R_s, when the model is combined with a "cylinder", is approximately 2.5 μm, and around 3 μm when combined with a "stick". We note here that the previously observed trend for the f_1 estimate is now disrupted and the "stick" models provide higher f_1 than both non-zero radius intra-axonal compartments. This may occur because the "sphere" can accommodate the mobility that the "cylinder" or "GDRCylinders" capture just as well, so some redundancy between the intra-axonal and third compartment arises.

Since there is no available ground truth for the true fibre orientation within the CC, to compare how well the models estimate the orientation we assume that the parallel direction of our imaging scheme is representative of the left-right fibre direction in the rat CC. All two-compartment models approximate the orientation well. The DT using the non-linear fitting estimates the orientation the least closely compared to the rest of the models. The DT with the linear fitting however, provides the closest estimate of the parallel direction. The somewhat anomalously good performance of the linear DT for this parameter most likely arises because the voxels we analyse are chosen by thresholding the orientation from the same model with linear fitting.

The three-compartment models also estimate the left-right fibre orientation closely; however the estimated orientation from the "Dot" models in combination with the extra-axonal "ball" model show greater discrepancy from the specified direction compared with the rest of the three-compartment models. The trends we describe here are consistent between datasets.

Quality of fit

Fig. 4 visualises the fit of the data synthesised from the DT fit using linear estimation, the DT fit with non-linear estimation, the "BallStick" and the "Bizeppelin" model to the scan data. Fig. 5 presents data synthesised from the simplified version of CHARMED (ZeppelinCylinder), the CHARMED (TensorGDRCylinders), the MMWMD model (ZeppelinCylinderDot) and the best two- and three-compartment (TensorCylinder and TensorCylinderSphere) models. To account for possible offsets that can arise from Rician noise, the synthesised data plots actually show the mean signals

Table 5
Fitted parameters for the DT, "Bizeppelin", "BallStick", the simplified version of CHARMED (ZeppelinCylinder), the MMWMD model (ZeppelinCylinderDot) the best two-compartment models "TensorCylinder", CHARMED (TensorGDRCylinders) and the best three-compartment model "TensorCylinderSphere". The units for the diffusivities $d_\parallel, d\perp_1, d\perp_2$ are in μm²/s, for the radii R, R_s, ϑ in μm and for the angles θ, φ, α are in radians.

| Models		Best parameter estimates													
		S_0	f_1	f_2	f_3	$d_\parallel$	$d\perp_1$	$d\perp_2$	θ	ϕ	R	R_s	k	ϑ	α
DT linear	A	0.824	1	–	–	24.9	17.8	16.6	1.570	1.571	–	–	–	–	0.45
	B	0.861	1	–	–	24.2	17.3	13.9	1.571	1.571	–	–	–	–	0.39
DT	A	0.954	1	–	–	383	75.5	63.5	1.94	1.893	–	–	–	–	1.268
	B	0.935	1	–	–	271	41.5	17.2	1.418	1.753	–	–	–	–	0.718
Bizeppelin	A	0.985	0.525	0.475	–	344	a = 344 b = 4.9	–	1.628	1.514	–	–	–	–	–
	B	0.987	0.566	0.434	–	293	a = 5.26 b = 289	–	1.69	1.521	–	–	–	–	–
BallStick	A	0.985	0.443	0.557	–	342	–	–	1.46	1.473	–	–	–	–	–
	B	0.988	0.523	0.477	–	293	–	–	1.39	1.521	–	–	–	–	–
ZeppelinCylinder	A	0.986	0.504	0.496	–	344	341	–	1.521	1.621	2.06	–	–	–	–
	B	0.988	0.565	0.434	–	295	256	–	1.45	1.521	1.89	–	–	–	–
TensorCylinder	A	0.982	0.45	0.55	–	352	351	219	1.59	1.55	1.95	–	–	–	0.78
	B	0.988	0.53	0.47	–	298	297	161	1.66	1.56	1.83	–	–	–	0.62
TensorGDRCylinders	A	1.11	0.540	0.460	–	346	346	254	1.49	1.54	–	–	1.77	1.28	1.183
	B	1.205	0.604	0.396	–	302	1.8	0.23	1.71	1.52	–	–	2.6	0.876	2.03
ZeppelinCylinderDot	A	0.996	0.346	0.331	0.323	829	359	–	1.29	1.831	4.3	–	–	–	–
	B	0.996	0.393	0.282	0.325	661	299	–	1.29	1.41	3.5	–	–	–	–
TensorCylinderSphere	A	0.994	0.101	0.516	0.382	802	321	169	1.57	1.57	8.0	2.94	–	–	2.318
	B	0.995	0.155	0.490	0.355	615	224	108	1.45	1.57	5.5	2.5	–	–	0.733
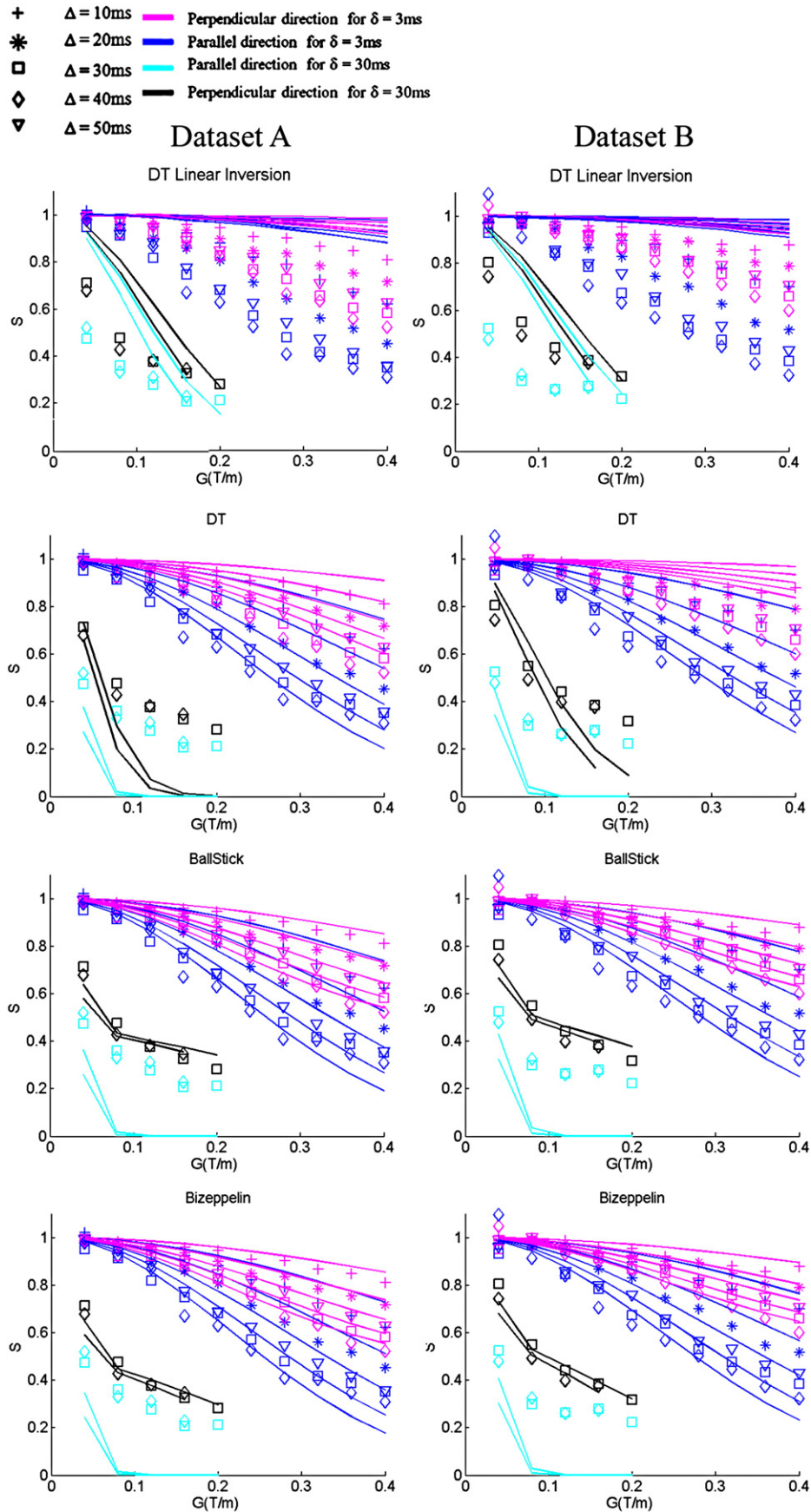
Fig. 4. Data synthesised from the DT, using both linear and non-linear fitting, the "BallStick", the "Bizeppelin" model and the scan data from the PGSE experiment for both data sets. The symbols represent the scan data and the lines show the corresponding measurements predicted by each model with the fitted parameters. The normalised signal S is plotted for all the values of Δ, δ as a function of the gradient strength |G| for the parallel and the mean over the four perpendicular directions.

over 500 trials adding independent Rician noise at approximately the level in the scan data. However, this procedure only significantly affects measurements with very low signal. We compare data synthesised from the analytical models by plotting the signal S for all values of Δ and δ as a function of the gradient strength $|G|$ for the parallel and the mean of the four perpendicular directions.

Fig. 4 shows that the DT model using the linear estimation significantly departs from the scan data in both directions. Non-linear estimation improves the fit in both directions, however discrepancies from the scan data are still obvious. In contrast, all two-compartment models capture the broad trends of the data and the anisotropy that separates the parallel and perpendicular signals. The simplest five-parameter "BallStick" model captures the broad trends in the data better than the six-parameter DT model, performing similar to the "Bizeppelin" model. The subtle variations that improve the fit for the "ZeppelinCylinder", "TensorGDRCylinders" and "TensorCylinder" models (see Fig. 5) over the "BallStick" and "Bizeppelin" models are difficult to observe qualitatively. In all two-compartment models we observe the biggest departures for large δ in the parallel direction.

All three-compartment models capture the observed departure in the parallel direction better than the two-compartments. That is the key difference that ranks the three-compartment models over the two-compartment. The best performing "sphere" and "dot" models have greater restriction than the "astro" third compartments and capture the trends in the data better. We also observe that for $\delta = 30$ ms the "sphere" models predict the parallel signal closer to the scan data than the "dot" models and underestimate the perpendicular signal, while the "dot" signal estimation is closer to the perpendicular direction and overestimates the parallel. On balance, the BIC prefers the "sphere" models, but both have limitations.

Stability

To evaluate the stability of the fit of the models to the data we compute the histogram of the final objective function obtained by the Levenberg–Marquardt algorithm from each of the 1000 perturbations of the starting parameters. In general we find that the "cylinder" models are much more stable than "GDRCylinders" models i.e., the fitting finds the lowest objective function value more frequently. Fig. 6 shows examples for the key models with both non-zero radius intra-axonal compartments using dataset A.

To quantify this difference more precisely, we estimate the number of runs required for each of these models to ensure obtaining the lowest objective function in at least one run with probability $P > 0.99$. From the histograms we estimate the probability of finding the best solution in a single run as the fraction of the 1000 runs in the lowest histogram bin. The probability of not obtaining the best solution in N runs is then $(1-p_{model})^N$. To get the number of runs for $P > 0.99$ we take the smallest integer N for which $(1-p_{model})^N < 0.01$.

Table 6 lists N for the key models with both non-zero radius intra-axonal compartments. The table suggests that stable models are the "BallStick" and the three-compartment "ZeppelinCylinderDot". Generally the number of required runs increases with complexity and the "GDRCylinders" models require many more runs than "cylinder". The DT model is surprisingly unstable since it has only 6 parameters, but most likely because even the global minimum does not fit the data well.

Discussion and conclusions

We have constructed, studied, evaluated and compared 47 analytic models for the diffusion-weighted MR signal in brain white matter. The study includes widely used models such as:

• the DT (Basser and Jones, 2002; Basser et al., 1994),
• the "BallStick" model (Behrens et al., 2003),

• the "Bizeppelin" model (Alexander and Barker, 2005; Parker and Alexander, 2003; Peled et al., 2006; Tuch et al., 2002), which is a three-dimensional analogue of the biexponential model (Clark and Le Bihan, 2000; Niendorf et al., 1996)
• CHARMED as used in AxCaliber (Assaf et al., 2008),
• the simplified CHARMED model in Alexander (2008),
• the MMWMD in ActiveAx (Alexander et al., 2010) without the CSF compartment.

It also includes new compartment models with novel combinations of two and three compartments. The intention here is to identify the model components required to capture the signal in brain white matter. Thus we use a model selection procedure to choose the model which best describes the data, while minimising the number of parameters. We also devise an imaging protocol designed to cover as broad a sampling of the space of measurements of diffusion in white matter as possible. This provides high signal to noise measurements representative of a coherent white matter voxel in the brain in a similar way to previous studies on nerve tissue (Stanisz et al., 1997) or spinal cord (Assaf et al., 2004) where orientation is simpler to handle. The dataset supports the model comparison and selection to identify good models to fit in future work to sparser data voxel by voxel to produce parameter maps.

The new imaging protocol provides parallel and perpendicular signals from brain tissue. Although the models we present here are applicable and feasible for whole brain imaging, the acquisition protocol purposefully is not. The aim of this work is to compare models using a sample with known orientation, which allows many more measurements to support the model comparison than if we try to accommodate arbitrary orientation. Once we have established appropriate models, we can subsequently find more economical and orientationally invariant protocols for whole brain imaging, for example, using the experiment design optimization in Alexander (2008). We note that the more complex models require rich imaging protocols with different b values, to provide feasible and stable microstructural estimates and are not identifiable with more standard single-shell acquisitions, where only simpler models such as the DT and "BallStick" are viable.

The results reveal that the DT model shows a significant departure from the scan data, especially using the linear estimation, and confirms expectations that the model is poor at high b value, because it does not account for restriction. Even the simplest two-compartment models, such as the "BallStick" capture the signal much better than the DT model. The "Bizeppelin" model has similar BIC to the simplest two-compartment models, but more complex two-compartment models outperform it. This suggests that explicit modelling of restriction is important in white matter.

The departure of the signals from the two-compartment models in the parallel direction most likely comes from a small amount of restriction parallel to the fibres from glial cells and/or non-parallel fibres. This supports observations in both Stanisz et al. (1997) and Alexander et al. (2010). Three compartment models capture this effect and thus consistently outperform the two-compartment models. The BIC in Table 4 consistently prefers the spherical compartments "sphere" and "dot" over uniformly oriented cylinders. Also the BIC steadily supports the non-zero axon radius parameters: both the "cylinder" and the "GDRCylinders" models rank higher than the "stick" models. This adds credence to techniques that estimate axon diameter parameters (Alexander, 2008; Alexander et al., 2010; Assaf and Basser, 2005; Barazany et al., 2009; Stanisz et al., 1997).

Concerning the extra-axonal compartments, the BIC shows that the "tensor" and "zeppelin" models consistently perform better than the isotropic "ball", despite the increased number of parameters. This demonstrates that the data support the anisotropy in the extra-axonal space.
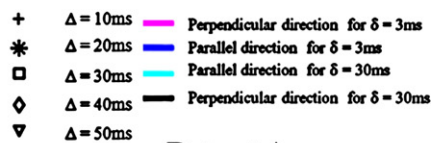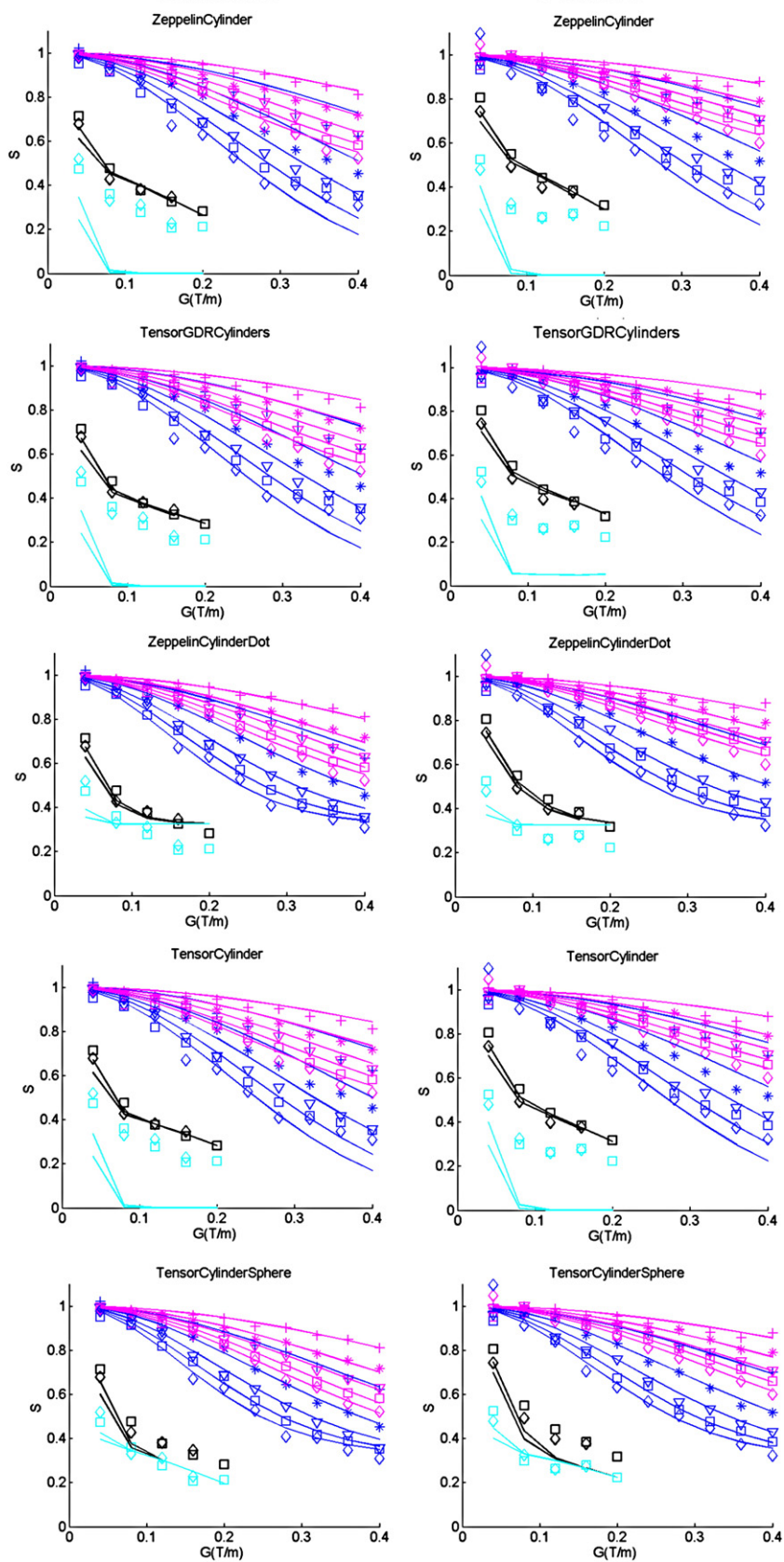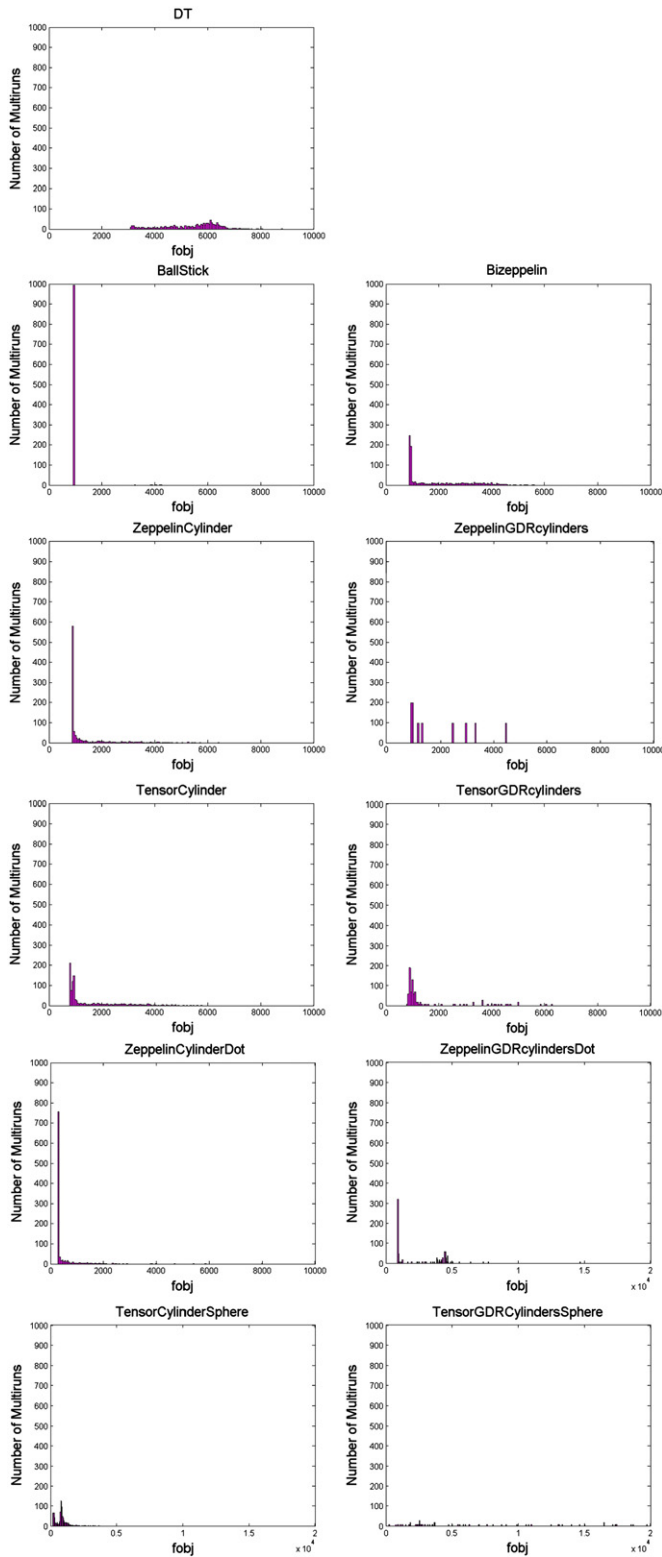
Fig. 6. Histograms of the final objective function of 1000 runs for the DT, "BallStick", "ZeppelinCylinder", "ZeppelinGDRCylinders", "TensorCylinder" and "TensorGDRCylinders" using dataset A.

According to BIC both non-zero intra-axonal models perform well. However, the best fit parameters are much harder to locate for the "GDRCylinders" compared to the "cylinder". According to Table 6,

Table 6
Table indicating the number of runs required to obtain the best solution in each model with probability $P > 0.99$ for datasets A and B.

Models	No of runs for $P > 0.99$	
	Dataset A	Dataset B
DT	510	500
BallStick	1	1
Bizeppelin	22	9
ZeppelinCylinder	5	3
ZeppelinGDRCylinders	24	19
TensorCylinder	18	10
TensorGDRCylinders	458	52
ZeppelinCylinderDot	3	3
ZeppelinGDRCylindersDot	17	40
TensorCylinderSphere	68	74
TensorGDRCylindersSphere	510	458

the fitting finds the minimum objective function for the "cylinder" model much more frequently than the "GDRCylinders" model and "cylinder" models are less sensitive to the starting parameters, revealing greater efficiency. For example, the duration of a single run when fitting the "ZeppelinCylinder" model is 6 s, while for the "ZeppelinGDRCylinders" is 36 s. Furthermore, Table 6 suggests that the "ZeppelinGDRCylinders" would require about 5 times as many runs per voxel so the total computation time is approximately 30 times longer. Thus using standard model fitting techniques the simpler "cylinder" models are a more practical choice, although better fitting procedures may improve the practicality of the "GDRCylinders".

The most realistic values for R come from the models "ZeppelinCylinder", "TensorCylinder", "ZeppelinGDRCylinders", "TensorGDRCylinders", "BallCylinderAstrocylinders" and "ZeppelinCylinderAstrocylinders". Although we have not performed histology to measure the axon diameter distribution in our samples, we get a rough idea from the histology data from the rat CC in Barazany et al. (2009). The axon diameter histograms in Fig. 1 of Barazany et al. (2009) give mean axon radius, weighted by volume (Alexander et al., 2010) and scaled by 1.5 to correct for shrinkage (Aboitiz et al., 1992) of 1.07 μm in the genu, 1.4 μm in the midbody, 1 μm in the splenium and 1.25 μm over the whole CC. The closest approximations are 1.73 μm from model "ZeppelinCylinderAstrocylinders", 1.74 μm from "BallCylinderAstrocylinders", 1.83 μm from models "TensorCylinder" and "TensorGDRCylinders" and 1.89 μm from "ZeppelinCylinder" and "ZeppelinGDRCylinders". In the parameter estimation we find that the axon radius estimation from the "GDRCylinders" model is generally in agreement with the "cylinder" model for both datasets. However, we must emphasize that the axon radius is not a quantitative measurement, because it is affected by the acquisition protocol and axon diameter distribution, although it does provide contrast between regions of tissue with different axon diameter distributions. This is why Alexander et al. (2010) refer to the estimate as the axon diameter index. We do not attempt to base the model comparison and ranking on the axon diameter estimate.

The diffusivity is low compared to reported estimates in in-vivo studies (Barazany et al., 2009), however this is to be expected due to the lower temperature, changes in the chemical environment during fixation and loss of active function after death. The values we estimate are similar to those used and estimated in other fixed tissue diffusion studies (Alexander et al., 2010; McNab et al., 2009). A translation of the parameter estimates and the model ranking to what we might expect in vivo is not simple because discrepancies arise from the fixation, the death process and temperature changes (Widjaja et al., 2009). Alexander et al. (2010) confirm this expectation finding that the addition of the third compartment is less necessary in vivo.

Fig. 5. As Fig. 4 but data synthesised from the "ZeppelinCylinder", "TensorGDRCylinders", "ZeppelinCylinderDot", "TensorCylinder" and "TensorCylinderSphere" models and the scan data from the PGSE experiment for both data sets.

This study uses the offset Gaussian noise model for parameter estimation in all experiments. We have repeated all the fitting experiments with both Gaussian and Rician noise models. We omit the other two types since the model comparison results are similar although minor variations arise in the parameter estimates. We choose to present the offset Gaussian fitting as it improves the quality of fit compared to the Gaussian model and can be more stable than the Rician model (Alexander, 2009). The parameter estimates using a Gaussian noise model during fitting can be found in Panagiotaki (2011).

As a final observation, the best ranked models are consistently those with the most parameters (degrees of freedom). This suggests that other effects not captured by the multi-compartment models presented here may be significant so useful avenues for further work arise to extend the set of models to include such effects. One kind of extension is simply to relax the constraints we impose on our models. For example, we might have different diffusivities or orientation parameters in each compartment. It is also possible to allow T_2 to vary among compartments, as currently we assume it is the same. Additionally we could add permeability parameters to capture exchange between compartments, although the results in Stanisz et al. (1997) suggest that longer diffusion times are necessary to see these effects. We note that both Stanisz et al. (1997) and Alexander et al. (2010) use a tortuosity model to estimate $d_\perp$ from $d_\parallel$ and we do not consider that here. We also assume here that the axons are straight and parallel. Although this assumption should hold fairly well in the midsagittal CC, where we take the data from, even small amounts of orientation dispersion may affect the measurements, particularly as we average over voxels from different locations. Models that accommodate fibre dispersion are available (Jespersen et al., 2007; Kaden et al., 2007; Zhang et al., 2011b) and would be interesting to compare in the future, although here the aim is to identify a model for a single orientation from which to derive dispersed orientation models. To further test the orientation estimates we repeat the experiment for the two-compartment models relaxing the constraint that equates the orientation of the intra- with the extra-axonal compartments. The results show that the extra orientation does reduce BIC slightly, however we do not include the models in the main results because the imaging protocol does not support non-aligned fibre orientation. Further work could assess these models with more gradient orientations. Future work could also incorporate a CSF compartment in the model to account for CSF contamination that can occur due to the small volume of the CC structure, as in Alexander et al.(2010) and Barazany et al.(2009).

All experiments in this paper are restricted to a controlled and "ideal" configuration, that of a single fibre orientation in the corpus callosum. The study has relevance to multi-fibre techniques (Seunarine and Alexander, 2009) because multi-fibre models require good single-fibre models either as a kernel for deconvolution or as a component in mixture models. Some potential complications arise in extending the set of models to multiple fibre populations. Although multiple intra-axonal components are straightforward to model, the effect of multiple fibre populations on the extra-axonal compartment is less clear: a single oblate tensor could be used for interdigitating orthogonal fibres, whereas a mixture of two prolate tensors might be more suitable an interface. Zhang et al. (2011a) discuss such issues further. Future work to study the extension to multiple fibres or fibre dispersion will require higher angular resolution acquisitions, which increase data requirements.

We limit the investigation here to the standard PGSE sequence. Another exciting area for further work is to substitute the standard PGSE sequence with other diffusion-sensitive sequences, such as oscillating gradients (Callaghan and Stepisnik, 1995; Does et al., 2003; Drobnjak et al., 2010), twice-refocused spin echo sequences (Clayden et al., 2009; Reese et al., 2003), or double wave vector sequences (Koch and Finsterbusch, 2008; Mitra, 1995; Shemesh et al., 2009).

These sequences probe different time scales and features of water dispersion to offer increased and new sensitivity in comparison to PGSE. They may require entirely different model components.

White matter microstructure imaging is a useful alternative to the classical histological techniques. However, identifying the appropriate model is critical to providing useful and reliable information. This study provides a way to achieve that and makes progress towards finding a suitable model for white matter diffusion microscopy. The microstructural parameter estimates could potentially be used as biomarkers for studying normal development as well as brain diseases. Accurate estimation of the variation of indices such as the axon diameter and density as well as diffusivity could help diagnosis and prognosis. Finally these analytic models can be used in combination with tractography algorithms, as in Sherbondy et al. (2010) for whole brain microstructure estimates. This could help resolve ambiguities that occur from problematic regions such as crossing, kissing and fanning fibres. All the models are implemented in the open-source diffusion MRI toolkit Camino (Cook et al., 2006) and are available using Gaussian, offset Gaussian and Rician noise models.

Acknowledgments

Appendix A. Signal models

This appendix states the signal model for the "cylinder" model and derives the models of isotropic restriction "astrosticks" and "astrocylinders". First we state the signal model for the cylinder and then derive the signal model for isotropically oriented cylinders.

To model the restricted signal S_1 from particles diffusing within a cylinder we assume that it can be regarded as the product of the signals parallel and perpendicular to the cylinder axis, $\mathbf{n}$ (Assaf et al., 2004)

$$S_1 = S_{1\parallel} S_{1\perp}. \tag{A.1}$$

In the direction parallel to the fibres we assume free diffusion, so that

$$S_{1\parallel} = \exp\left(-(\Delta - \delta/3)(\gamma\delta\mathbf{G}\cdot\mathbf{n})^2 d\right) = \exp\left(L_\parallel(\mathbf{G}\cdot\mathbf{n})^2\right), \tag{A.2}$$

where

$$L_\parallel = -(\Delta - \delta/3)(\gamma\delta)^2 d_\parallel, \tag{A.3}$$

$\mathbf{G}\cdot\mathbf{n}$ is the component of the gradient parallel to the fibres and $d_\parallel$ is the free diffusivity within the cylinder.

To model $S_{1\perp}$ we can consider the following generalised form:

$$S_{1\perp} = \exp\left(L_\perp\left(\mathbf{G}\cdot\mathbf{G} - (\mathbf{G}\cdot\mathbf{n})^2\right)\right), \tag{A.4}$$

where $\mathbf{G}\cdot\mathbf{G} - (\mathbf{G}\cdot\mathbf{n})^2$ is the square of the magnitude of the component of the gradient perpendicular to $\mathbf{n}$. The GPD approximation (Gelderen et al., 1994; Stepisnik, 1993) provides a model for $S_{1\perp}$ by substituting

$$L_\perp = -2\gamma^2 \sum_{m=1}^{\infty} \frac{2d_\parallel\beta_m^2\delta - 2 + 2Y(\delta) + 2Y(\Delta) - Y(\Delta - \delta) - Y(\Delta + \delta)}{d_\parallel^2\beta_m^6((\alpha/2)^2\beta_m^2 - 1)} \tag{A.5}$$

where $Y(x) = \exp\left(d_\parallel\beta_m^2 x\right)$, β_m is the mth root of equation $J_1'(\beta\alpha/2) = 0$ and J_1' is the derivative of the Bessel function of the first kind, order one.

Table B.7

Table presenting the model parameters, their range, the optimised parameters (each has range $(-\infty,\infty)$) and the transformation used in the optimization.

Model parameter	Range	Optimised parameter	Transformation
S_0	$[0,\infty)$	θ_1	$S_0 = \theta_1^2$
f_1	$[0,1]$	θ_2	$f_1 = cos^2\theta_2$
f_2	$[0,1-f_1]$	θ_3	$f_2 = (1-f_1)cos^2\theta_3$
$d_\parallel$	$[0,\infty)$	θ_4	$d_\parallel = \theta_4^2$
$d\perp$	$[0,d_\parallel]$	θ_5	$d\perp = d_\parallel sin^2\theta_5$
$d\perp_2$	$[0,d\perp]$	θ_6	$d\perp_2 = d\perp sin^2\theta_6$
R	$[1e^{-7}, 20.1e^{-6}]$	θ_7	$R = 1e^{-7} + 20e^{-6}cos^2\theta_7$
R_s	$[1e^{-7}, 40.1e^{-6}]$	θ_8	$R_s = 1e^{-7} + 40e^{-6}cos^2\theta_8$

To obtain the collective signal S_3 from water inside cylinders with a distribution of orientations p (Zhang et al., 2011b), we integrate the signal S_1 for a cylinder with axis $\mathbf{n}$ over all $\mathbf{n}$:

$$S_3 = \int S_{1\perp} p(\mathbf{n}) d\mathbf{n}. \tag{A.6}$$

Here we assume uniform p so $p(\mathbf{n}) = (4\pi)^{-1}$ and (He and Yablonskiy, 2009; Jespersen et al., 2007)

$$S_3 = \pi^{1/2} (2|\mathbf{G}|(L_\perp - L_\parallel)^{1/2})^{-1} exp(|\mathbf{G}|^2 L_\perp) \phi(|\mathbf{G}|(L_\perp - L_\parallel)^{1/2}), \tag{A.7}$$

where ϕ is the error function

$$\phi(z) = 2\pi^{-1/2} \int_0^z exp(-t^2) dt. \tag{A.8}$$

Appendix B. Parameter transformation

R1R3This appendix details the mathematical transformations between the model parameters and optimization parameters. Table B.7 presents the model parameters, their constrained ranges, the corresponding optimization parameters and the encoding function between the two.

Appendix C. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2011.09.081.

References

Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. Biophys. J. 66, 259–267.

Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen, H., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.

Callaghan, P.T., Stepisnik, J., 1995. Frequency-domain analysis of spin motion using modulated-gradient NMR. J. Magn. Reson., Ser. A 117, 118–122.

Clark, C.A., Le Bihan, D., 2000. Water diffusion compartmentation and anisotropy at high b values in the human brain. Magn. Reson. Med. 44 (6), 852–859.

Clayden, J., Nagy, Z., Hall, M.G., Clark, C., Alexander, D.C., 2009. Active Imaging with Dual Spin-Echo Diffusion MRI, pp. 264–275.

Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., Alexander, D.C., 2006. Camino: Open source diffusion-MRI reconstruction and processing. Proc 14th Scientific Meeting of ISMRM, Seattle, WA, USA, p. 2759.

Does, M.D., Parsons, E.C., Gore, J.C., 2003. Oscillating gradient measurements of water diffusion in normal and globally ischemic rat brain. Magn. Reson. Med. 49 (2), 206–215.

Douglass, D.C., McCall, D.W., 1958. Diffusion in paraffin hydrocarbons. J. Phys. Chem. 62, 1102.

Drobnjak, I., Siow, B., Alexander, D.C., 2010. Optimizing gradient waveforms for microstructure sensitivity in diffusion-weighted MR. J. Magn. Reson. 206, 41–51.

Dyrby, T.B., Baaré, W.F.C., Alexander, D.C., Jelsing, J., Garde, E., Søgaard, L.V., 2011. An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. Hum. Brain Mapp. 32, 544–563.

Gelderen, P.V., Despres, D., Zijl, P.C.M.V., Moonen, C.T.W., 1994. Evaluation of restricted diffusion in cylinders phosphocreatine in rabbit leg muscle. J. Magn. Reson. B 103, 255–260.

He, X., Yablonskiy, D.A., 2009. Biophysical mechanisms of phase contrast in gradient echo mri. Proc. Natl. Acad. Sci. 106 (32), 13558.

Jespersen, S., Kroenke, C., Østergaard, L., Ackerman, J., Yablonskiy, D., 2007. Modeling dendrite density from magnetic resonance diffusion measurements. Neuroimage 34 (4), 1473–1486.

Jespersen, S.N., Bjarkam, C.R., Nyengaard, J.R., Chakravarty, M.M., Hansen, B., Vosegaard, T., Østergaard, L., Yablonskiy, D., Nielsen, N.C., Vestergaard-Poulsen, P., 2010. Neurite density from magnetic resonance diffusion measurements at ultrahigh field: comparison with light microscopy and electron microscopy. Neuroimage 49 (1), 205–216.

Jones, D.K., Basser, P.J., 2004. "Squashing peanuts and smashing pumpkins": how noise distorts diffusion-weighted MR data. Magn. Reson. Med. 52 (5), 979–993.

Kaden, E., Knösche, T.R., Anwander, A., 2007. Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging. Neuroimage 37 (2), 474–488.

Koch, M., Finsterbusch, J., 2008. Compartment size estimation with double wave vector diffusion-weighted imaging. Magn. Reson. Med. 60 (1), 90–101.

McNab, J., Jbabdi, S., Deoni, S., Douaud, G., Behrens, T., Miller, K., 2009. High resolution diffusion-weighted imaging in fixed human brain using diffusion-weighted steady state free precession. Neuroimage 46 (3), 775–785.

Mitra, P.P., 1995. Multiple wave-vector extensions of the nmr pulsed-field-gradient spin-echo diffusion measurement (Jun) Phys. Rev. B. Condens Matter 51, 15074–15078 (URL) http://link.aps.org/doi/10.1103/PhysRevB.51.15074.

Murday, J.S., Cotts, R.M., 1984. Self-diffusion coefficient of liquid lithium. J. Chem. Phys. 48, 4938–4945.

Neuman, C.H., 1974. Spin echo of spins diffusing in a bounded medium. J. Chem. Phys. 60, 4508–4511.

Niendorf, T., Dijkhuizen, R., Norris, D.G., van Lookeren Campagne, M., Nicolay, K., 1996. Biexponential diffusion attenuation in various states of brain tissue: implications for diffusion-weighted imaging. Magn. Reson. Med. 36 (6), 847–857.

Panagiotaki, E., 2011. Geometric models of brain white matter for microstructure imaging with diffusion MRI. Ph.D. thesis, Department of Computer Science, University College London.

Panagiotaki, E., Fonteijn, H., Siow, B., Hall, M.G., Price, A., Lythgoe, M.F., Alexander, D.C., 2009. Two-compartment models of the diffusion MR signal in brain white matter. Med Image Comput Comput Assist Interv (MICCAI): Lecture Notes in Computer Science, pp. 329–336.

Parker, G.J.M., Alexander, D.C., 2003. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. Proc of 18th IPMI conference (Ambleside). Springer: LNCS 2732, pp. 684–695.

Peled, S., Friman, O., Jolesz, F., Westin, C.-F., 2006. Geometrically constrained two-tensor model for crossing tracts in dwi. Magn. Reson. Imaging 24 (9), 1263–1270.

Press, W.S., Flannery, B., Teukolsky, S., 1988. WT Vetterling WT Numerical Recipes in C.

Price, W.S., 1997. Pulse-Field Gradient Nuclear Magnetic Resonance as a tool for studying translational diffusion: part 1. basic theory. Magn. Reson. Educ. J. 9, 299–336.

Reese, T., Heid, O., Weisskoff, R., Wedeen, V., 2003. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn. Reson. Med. 49 (1), 177–182.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6, 461–464.

Seunarine, K.K., Alexander, D.C., 2009. Multiple Fibers: Beyond the Diffusion Tensor. Academic Press.

Shemesh, N., Ozarslan, E., Basser, P.J., Cohen, Y., 2009. Measuring small compartmental dimensions with low-q angular double-pgse nmr: the effect of experimental parameters on signal decay. J. Magn. Reson. 198 (1), 15–23.

Shepherd, T., Thelwall, P., Stanisz, G., Blackband, S., 2009. Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. Magn. Reson. Med. 62 (1), 26–34.

Sherbondy, A.J., Rowe, M., Alexander, D.C., 2010. MicroTrack: an algorithm for concurrent projectome and microstructure estimation. Med Image Comput Comput Assist Interv (MICCAI): Lecture Notes in Computer Science, pp. 183–190.

Aboitiz, F., Scheibel, A.B., Fisher, R.S., Zaidel, E., 1992. Fiber composition of the human corpus callosum. Brain Res. 598, 143–153.

Alexander, D.C., 2008. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. Magn. Reson. Med. 60, 439–448.

Alexander, D.C., 2009. Modelling, fitting and sampling in diffusion MRI. Visualization and Processing of Tensor Fields, pp. 3–20.

Alexander, D.C., Barker, G.J., 2005. Optical imaging parameters for fibre-orientation estimation in diffusion MRI. Neuroimage 27, 357–367.

Alexander, D.C., Hubbard, P.L., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J.M., Dyrby, T.B., 2010. Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage 52, 1374–1389.

Assaf, Y., Basser, P.J., 2005. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage 27 (1), 48–58.

Assaf, Y., Freidlin, R.Z., Rohde, G.K., Basser, P.J., 2004. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. Magn. Reson. Med. 52 (5), 965–978.

Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y., Basser, P.J., 2008. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. Magn. Reson. Med. 59, 1347–1354.

Barazany, D., Basser, P., Assaf, Y., 2009. In-vivo measurement of the axon diameter distribution in the corpus callosum of a rat brain. Brain 132, 1210–1220.

Basser, P.J., Jones, D.K., 2002. Diffusion-tensor MRI: theory, experimental design and data analysis — a technical review. NMR Biomed. 15, 456–467.

Stanisz, G.J., Szafer, A., Wright, G.A., Henkelman, R.M., 1997. An analytical model of restricted diffusion in bovine optic nerve. Magn. Reson. Med. 37, 103–111.

Stepisnik, J., 1993. Time-dependent self-diffusion by NMR spin-echo. Physica B 183 (4), 343–350.

Tuch, D.S., Reese, T.G., Wiegell, M.R., Belliveau, N.M.J.W., Wedeen, V.J., 2002. High angular resolution diffusion imaging reveals intravoxel white matter fibre heterogeneity. Magn. Reson. Med. 48, 577–582.

Widjaja, E., Wei, X., Vidarsson, L., Moineddin, R., Macgowan, C., Nilsson, D., 2009. Alteration of diffusion tensor parameters in postmortem brain. Magn. Reson. Imaging 27, 865–870.

Zhang, H., Dyrby, T.B., Alexander, D.C., 2011a. Axon diameter mapping in crossing fibers with diffusion MRI. Med Image Comput Comput Assist Interv (MICCAI): Lecture Notes in Computer Science, pp. 82–89.

Zhang, H., Hubbard, P.L., Parker, G.J.M., Alexander, D.C., 2011b. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. Neuroimage 56, 1301–1315.