

A General Framework for Experiment Design in Diffusion MRI and Its Application in Measuring Direct Tissue-Microstructure Features

Daniel C. Alexander*

This article introduces a new and general framework for optimizing the experiment design for diffusion MRI of samples with unknown orientation. An illustration then uses the framework to study the feasibility of measuring direct features of brain-tissue microstructure *in vivo*. The study investigates the accuracy and precision with which we can estimate potentially important new biomarkers such as axon density and radius in white matter. Simulation experiments use a simple model of white matter based on CHARMED (composite hindered and restricted model of diffusion). The optimization finds acquisition protocols achievable on modern human and animal systems that consist of 120 measurements with fixed maximum gradient strengths. Axon radii in brain tissue are typically in the range 0.25–10 μm . Simulations suggest that estimates of radii in the range 5–10 μm have highest precision and that a maximum gradient strength of 0.07 T m⁻¹ is sufficient to distinguish radii of 5, 10, and 20 μm . Smaller radii are more difficult to distinguish from one another but are identifiable as small. A maximum gradient strength of 0.2 T m⁻¹ distinguishes radii of 1 and 2 μm . The simulations also suggest that axon densities and diffusivity parameters in the normal range for white matter are recoverable. The experiment-design optimization has applications well beyond the current work to optimize the protocol for fitting any model of the diffusion process. *Magn Reson Med* 60:439–448, 2008. © 2008 Wiley-Liss, Inc.

Key words: quantitative MRI; diffusion MRI; microstructure; axon density; axon diameter; experiment design; Cramer-Rao lower bound; Rician noise; active imaging

INTRODUCTION

Simple biomarkers derived from diffusion-tensor (DT) MRI, such as mean diffusivity (MD) and fractional anisotropy (FA) (1), are useful as indicators of major microstructural changes, such as brain damage through stroke (2). They also correlate with cognitive ability, e.g. (3), highlight diseased tissue, e.g. (4), and allow monitoring of development (5) and aging (6). A limitation of biomarkers derived from the DT is that they are non-specific. Changes in FA accompany changes in axon density, radius distribution, orientation distribution, or permeability and cannot distinguish between them. Very different combinations of these microstructure features may give the

same FA. Direct measurements of these features would provide much greater insight into disease mechanisms, recovery, and development. Such measurements would also provide basic neuroscientific insight and may predict cognitive performance better than nonspecific indices like FA.

Previous work, e.g. (7–10), shows that diffusion MRI can provide estimates of features such as pore sizes in minerals and axon sizes, density and permeability in excised biological tissue. Most methods use measurements from a pulsed-gradient spin-echo (PGSE) (11) or stimulated-echo (STEAM) (12) sequence. Both pulse sequences have three main tunable parameters: the gradient vector \mathbf{G} , the length of the gradient pulses δ , and the time between the onsets of the two pulses Δ .

One approach to measure pore sizes or axon radii uses the diffraction pattern of the diffusion MRI signal, as a function of $|\mathbf{G}|$, from restricted water molecules in the long-diffusion-time limit ($\Delta \gg R^2/2d$, where R is the pore radius and d is the diffusivity) and assuming short gradient pulses ($\Delta \gg \delta$). Callaghan et al. (13) first noticed the diffraction pattern and related its frequency to the sample pore size. Avram et al. (14) recovered the radii of cylindrical polymer tubes with radii 20 μm and 9 μm from the diffraction pattern using STEAM. They also studied a related approach that estimates R using a two-compartment model. In their model, R is a linear function of the root mean square (RMS) displacement of the slower diffusing component perpendicular to the fiber direction. Topgaard and Soderman (15) recover the spacing of parallel glass plates from the diffraction pattern. However, Weng et al. (10) find that the diffraction pattern in measurements from an excised rat brain does not reflect the mean axon radius measured by microscope. Peled et al. (16) noted that measurements from samples with a range of pore sizes, as rat-brain tissue is likely to have, do not exhibit diffraction. However, Ong et al. (17) showed that the RMS displacement predicts the mean axon radius well in several distinct pathways in the mouse spinal cord, where the distribution of radii may be less broad. Lätt et al. (18) used simulations to study the diffraction pattern as a predictor of pore size on clinical MR systems. They concluded that $R = 10 \mu\text{m}$ is the lowest identifiable radius using current hardware.

An alternative approach is to construct a geometric model of microstructure that predicts the MR signal from water diffusing within. Model-based techniques potentially account for factors such as distributions of cell sizes more robustly and naturally provide estimates of several interesting features simultaneously. Stanisz et al. (9) constructed a geometric model of bovine optic-nerve tissue

Department of Computer Science, University College London, London, United Kingdom

*Correspondence to: Daniel Alexander, Department of Computer Science, UCL, Gower Street, London WC1E 6BT, United Kingdom. E-mail: d.alexander@cs.ucl.ac.uk

Received 3 August 2007; revised 5 March 2008; accepted 12 March 2008.

DOI 10.1002/mrm.21646

Published online in Wiley InterScience (www.interscience.wiley.com).

© 2008 Wiley-Liss, Inc.

consisting of non abutting spherical glial cells and ellipsoidal axons with semi permeable walls embedded in a homogeneous substrate. The model assumes short gradient pulses. They acquire 800 PGSE measurements on an excised sample: four repeats of each combination of $\Delta \in \{0.008, 0.01, 0.02, 0.03\}$ s and 50 settings of $|\mathbf{G}|$ in the range 0–1.4 T m⁻¹; \mathbf{G} is normal to the fiber direction and $\delta = 0.003$ s throughout. The fitted sizes, volume fractions and permeabilities of each cell type agree qualitatively with histology consistently in experiments with three separate samples.

Assaf et al.'s CHARMED model (19) of white matter contains impermeable parallel cylindrical non abutting axons with a gamma distribution of radii embedded in a homogeneous substrate. The model for the signal from the extra cellular compartment uses the DT model with principal direction aligned with the cylinder axes. The intra cellular signal uses Neuman's model of the MR signal from spins diffusing within a cylinder in the presence of a constant diffusion gradient (20). Later work (7) extends the CHARMED model to use Van Gelderen's extension (21) of Murday and Cott's model (22), which approximates the MR signal of water diffusing in cylinders with finite-length gradient pulses. In (7), Assaf et al. acquired 1,024 measurements using a similar protocol to (9): eight repeats of each combination of eight settings of Δ in the range [0.02, 0.15] s and 16 settings of $|\mathbf{G}|$ in the range [0, 1.2] T m⁻¹ with $\delta = 0.0025$ sec throughout. The acquisition takes about 12 h. The test sample is excised nerve-tissue with known fiber orientation and \mathbf{G} is perpendicular to the fiber direction for all the measurements. They fit the parameters of the gamma distribution of axon radii and show that the maximum likelihood distribution agrees closely with a histogram of radii measured by hand on a histological image of the same sample. Recent work (23) relaxes the constraint on the shape of the radius distribution to some extent.

The direct microstructure-imaging techniques reviewed above all rely on both prior knowledge of the pore or axon orientation and lengthy data acquisitions (many hours or even days) with high magnetic-field gradient strengths. The techniques are therefore practical only for non-live samples with a single pore orientation. To realize the full potential of direct microstructure features as imaging biomarkers requires a technique to measure them over the whole brain in live human subjects. In vivo neuroimaging imposes much lower limits on gradient strengths (modern human systems have maximum gradient strengths in the range 0.04–0.08 T m⁻¹) and the tolerable acquisition time is an order of magnitude lower (less than 1 h) than previous in vitro experiments. Moreover, the axon orientation is unknown in general.

This study develops a new algorithm for experiment-design optimization using stochastic optimization of the Cramer-Rao lower bound (CRLB) and a procedure to sample the posterior distribution of model parameters using Markov Chain Monte Carlo (MCMC). Both methods use the appropriate Rician noise model. They adapt naturally to a wide variety of alternative models, pulse sequences, and applications. This study optimizes an in vivo imaging protocol for a simple model of the diffusion MRI signal from white-matter tissue that includes axon density and

radius as parameters and recovers posterior distributions on the parameters. The Methods section outlines the simple model and introduces the new method for experiment-design optimization. The Experiments and Results section evaluates the precision and accuracy with which we can estimate the axon radius, density and other direct microstructure features, using the optimized acquisition protocols. The discussion section concludes.

METHODS

This section outlines a simple model for the MR signal from water diffusing in white matter and develops the experiment-design optimization.

Simple Model

The model is based on Assaf et al.'s CHARMED model (19), but distilled to the smallest number of parameters possible for testing feasibility of recovering features of this kind of system. The key differences are:

1. The simplified model assumes a single axon radius rather than a distribution.
2. The simplified model assumes cylindrical symmetry of the apparent DT in the extra cellular space. The DT is $(d_{\parallel} - d_{\perp})\mathbf{nn}^T + d_{\perp}\mathbf{I}$, which has major eigenvector \mathbf{n} in the fiber direction with corresponding eigenvalue d_{\parallel} (the diffusivity parallel to \mathbf{n}) and minor eigenvalues d_{\perp} (the apparent diffusivity perpendicular to \mathbf{n}); \mathbf{I} is the identity tensor.
3. The intrinsic diffusivities of the intra and extra cellular compartments are the same and equal to d_{\parallel} .

We use Van Gelderen's model (21) for the signal attenuation inside the cylinders from the component of \mathbf{G} perpendicular to \mathbf{n} , as in (7).

The dependent parameters of the model are as follows:

- The volume fraction f of the intra cellular compartment, which relates simply to the axon density.
- The fiber direction, \mathbf{n} .
- The axon radius, R .
- The intrinsic diffusivity, d_{\parallel} .
- The apparent diffusion coefficient, d_{\perp} , perpendicular to \mathbf{n} in the extra cellular space.

This study concentrates on measurements from the PGSE sequence, so the tunable, independent variables are \mathbf{G} , Δ , and δ .

Reference (24) gives a more complete description of the model. It is an experimental model only and does not account for many of the microstructural variables in real brain tissue. Some other variables are straightforward to incorporate. Integration over a model distribution for R , as in (7), can include a distribution of axon radii. Axon wall permeability allows exchange of particles between the intra and extra cellular compartments, which a pair of coupled differential equations (9) can model. White matter also contains glial cells, which a third compartment with spherical restriction might model, as in (9). Integration over a

model distribution for \mathbf{n} incorporates a distribution of fiber directions. Sen and Basser (25) show how to incorporate thickness of the axon walls and different concentrations among compartments. Relaxation times may vary between compartments, as in (26). Models of tortuosity (27) that estimate d_{\perp} from d_{\parallel} , f and R potentially reduce the number of parameters in the model by one, but such models are approximate only and break down at high f . Other effects, such as abutting cells and loss of percolation of the extra cellular space are difficult to incorporate in analytic models.

Experiment-Design Optimization

Each measurement in the acquisition may use a different combination of \mathbf{G} , Δ , and δ . We seek the set of combinations that give the best parameter estimates. This section first outlines the requirements and constraints of an in vivo imaging protocol that reduce the set of combinations the optimization can consider. We then construct an objective function that reflects the precision of model-parameter estimates for a particular protocol. The experiment-design optimization aims to find the protocol that minimizes the objective function, and the final subsection outlines the optimization procedure that performs the minimization.

Requirements and Constraints of In Vivo Imaging

In vivo human imaging places three key constraints on the protocol:

1. The tolerable acquisition time limits the number of measurements.
2. Both power and safety constraints limit the maximum gradient strength.
3. Model-parameter estimation must be orientationally invariant.

To avoid orientational dependence, DT-MRI experiments commonly use “high-angular resolution” acquisition protocols in which $|\mathbf{G}|$, Δ , and δ are the same for each measurement, but each has a unique gradient direction with the whole set distributed evenly on a hemisphere (28). A similar strategy works here. The optimization considers only the class of protocol that acquires the same number M of measurements in each of N gradient directions. The M combinations of $|\mathbf{G}|$, δ , and Δ are the same in each direction. We choose the N directions by electrostatic minimization (28), fix them, and optimize the M combinations of $|\mathbf{G}|$, δ , and Δ . This constraint reduces the dimensionality of the optimization significantly from $5NM$ (each PGSE measurement has 5 degrees of freedom) to $3M$ (fixing the gradient direction reduces the degrees of freedom of each measurement to 3).

Objective Function

The Fisher information matrix and the CRLB are standard tools for experiment design. The CRLB provides a lower bound on the variance of a fitted model parameter that often correlates closely with the true variance. To optimize the

protocol, we aim to minimize the sum of the coefficients of variation of the model parameters

$$\tilde{F} = \sum_{i=1}^W \sigma_i^2 / w_i^2, \quad [1]$$

where w_i , $i = 1, \dots, W$ are the model parameters and σ_i is the standard deviation of w_i . The σ_i^2 are unknown, so we use the CRLBs in their place:

$$F = \sum_{i=1}^W (J^{-1})_{ii} / w_i^2, \quad [2]$$

where $(J^{-1})_{ii}$ is the CRLB for w_i . The Appendix defines the CRLB and provides formulae for both the familiar form, which assumes a Gaussian noise model, and the CRLB under the assumption of Rician noise, which proves beneficial here. We limit the sum in Eq. [2] to terms for the scalar model parameters f , d_{\parallel} , d_{\perp} , and R only.

The function F provides the basis of an objective function to minimize with respect to the M combinations of $|\mathbf{G}|$, δ , and Δ . However, F depends on specific settings of the parameters, w_1, \dots, w_W , which take a range of values. The full objective function therefore integrates F over prior distributions on each w_i . This study assumes δ -function priors on the model parameters f , d_{\parallel} , and d_{\perp} : $f = 0.7$, $d_{\parallel} = 1.7 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, and $d_{\perp} = 0.2 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, unless stated otherwise. The next section uses a variety of single and multiple settings for R .

A δ -function prior for \mathbf{n} is not sufficient to ensure orientational invariance. The full objective function averages F over a set S of sample directions. The algorithm to construct S proceeds as follows:

1. Initialize S to contain one sample from a uniform distribution on the sphere.
2. Minimize F averaged over S .
3. Find the $\tilde{\mathbf{n}} \in S_{500}$, where S_{500} is a set of 500 sample directions evenly distributed over the sphere, which has the largest F with the optimized protocol.
4. IF $\tilde{\mathbf{n}} \notin S$, add $\tilde{\mathbf{n}}$ to S and GOTO 2, ELSE END.

The process usually converges with S containing only three or four elements.

Optimization

The objective function has many local minima. We use a stochastic optimization algorithm, SOMA (self-organizing migratory algorithm) (29), with population size of 50,500 migrations and otherwise default settings, to perform the minimization. Preliminary experiments (not shown) found that SOMA consistently outperforms a variety of other standard optimization algorithms for this task. The full optimization runs SOMA five times and picks the result with the smallest final value of the objective function.

During minimization, the optimization enforces the constraint $0 \leq |\mathbf{G}_k| \leq G_{\max}$, as well as $\delta_k \geq 0$ and $\Delta_k \geq \delta_k + \text{P180}$, where P180 is the time required for the 180° pulse. The optimization must also account for changes to the echo time TE required for the sequence, since T_2 effects reduce

Table 1
Optimized Combinations of $|\mathbf{G}|$, δ , and Δ from the Gaussian CRLB Objective Function with $G_{\max} = 0.2 \text{ T m}^{-1}$, $R = 2 \mu\text{m}$, and $\sigma_0 = 0.02$

| $ \mathbf{G} \text{ (T m}^{-1}\text{)}$ | $\Delta \text{ (s)}$ | $\delta \text{ (s)}$ | $ \mathbf{q} \text{ (mm}^{-1}\text{)}$ | $t \text{ (s)}$ | $b \text{ (s mm}^{-2}\text{)}$ |
|------------------------------------------|----------------------|----------------------|-----------------------------------------|-----------------|--------------------------------|
| 0.200 | 0.025 | 0.020 | 1052.8 | 0.018 | 20,087 |
| 0.200 | 0.026 | 0.018 | 956.2 | 0.021 | 18,771 |
| 0.121 | 0.029 | 0.016 | 507.8 | 0.023 | 6,035 |
| 0.200 | 0.013 | 0.008 | 411.4 | 0.010 | 1,744 |

The table includes the more familiar quantities $|\mathbf{q}|$, $t = \Delta - \delta/3$ and $b = t|\mathbf{q}|^2$.

the signal more for longer TE, which increases the significance of the noise. To avoid the need for precise knowledge of T_2 , TE is the same for all measurements and thus depends on the largest $\delta + \Delta$, i.e. $\text{TE} = \max_{k=1,\dots,M}(\delta_k + \Delta_k) + C$, where C is a constant depending on the time required for the 90° pulse and for read out. Equations [A2] and [A6] show that the CRLB depends on the noise level σ in the measurements. To account for the effects of varying TE, we set

$$\sigma = \sigma_0 \exp((\text{TE} - \text{TE}_0)/T_2) \quad [3]$$

in the Fisher information matrix in Eqs. [A1] and [A6], where σ_0 is the inverse of the signal to noise ratio (SNR) at a fixed echo time TE_0 . Unless otherwise stated, we set $T_2 = 0.07 \text{ s}$, which is typical for white matter, and assume a base setting of $\text{TE}_0 = 0.08 \text{ s}$. Thus, if $\sigma_0 = 0.02$, the SNR of the unweighted (i.e. $|\mathbf{G}| = 0$) signal is 50 when $\text{TE} = 0.08 \text{ s}$.

EXPERIMENTS AND RESULTS

This section contains simulation experiments that test the feasibility of measuring direct-microstructure features using the experiment-design optimization. All experiments use the simple model from the Methods section. The experiments test two settings of G_{\max} . One setting is $G_{\max} = 0.07 \text{ T m}^{-1}$, which is at the limit of what modern human scanners can achieve on live subjects. The other is $G_{\max} = 0.2 \text{ T m}^{-1}$, which is easily achievable in modern small-bore animal scanners. We limit the number of measurements to 120, which modern scanners can acquire in less than 45 min using a fast acquisition such as echo-planar imaging. Unless otherwise stated, we set $P180 = 0.005\text{s}$.

The first subsection shows example output from the experiment-design optimization and the second investigates the dependence of F on the ratio of N to M . The key results of the study are in the “Parameter Estimation” section, which shows how well we can recover the parameters of the model in the “Simple Model” section using the optimized protocols. Finally, the “Dependencies” section studies the dependence of the precision and accuracy of parameter estimates on various settings of the system.

Optimized Acquisition Protocols

This section compares qualitatively some protocols that the optimization procedure produces. For illustration, we set $R = 2 \mu\text{m}$, $G_{\max} = 0.2 \text{ T m}^{-1}$, $\sigma_0 = 0.02$, $N = 30$, and $M = 4$. Table 1 shows the optimized protocol for the objective function based on the Gaussian CRLB and Table 2 shows the optimized protocol for the Rician CRLB.

The objective function that uses the Gaussian CRLB tends to favor higher levels of diffusion weighting that results in measurements with very low signal to noise. The Rician version penalizes high diffusion weighting more. The difference is marked in this example where R is close to the limit of the measurable range (see later sections). At higher R , optimized protocols become similar for the two noise models. Simulation experiments (not shown) demonstrate clear superiority of protocols optimized using the Rician CRLB, and we shall not consider the Gaussian CRLB, further.

Choice of M

This section compares different combinations of N and M . The experiment optimizes the protocol separately for each $R \in \{1, 2, 5, 10, 20\} \mu\text{m}$ with each $M \in \{2, 3, 4, 5, 6, 8, 10, 12\}$ and $N = 120/M$. We then compute F , in Eq. [2], for the optimized protocol with each $\mathbf{n} \in S_{500}$. Figure 1 plots the mean F over S_{500} for each combination of R and M at each G_{\max} . The mean F is lowest for $R = 10 \mu\text{m}$, but comparable for 5 and $20 \mu\text{m}$ and, for $G_{\max} = 0.2 \text{ T m}^{-1}$, $2 \mu\text{m}$. The mean F is much higher for $R = 1 \mu\text{m}$ (note the scaling factors at lower R in the graphs), suggesting that lower R is harder to estimate. Values of $M \in \{3, 4, 5, 6\}$ give the lowest F and $M = 4$ is best most often. All the remaining tests use $N = 30$ and $M = 4$.

Parameter Estimation

This section shows simulation experiments that measure the ability to recover known parameter settings using the optimized protocols. Unless explicitly stated otherwise, all the model parameters take the values listed in the “Objective Function” section. Each experiment synthesizes data from the model in section “Simple Model” using the protocol under investigation and adds synthetic Rician noise assuming $\sigma_0 = 0.02$.

A simple MCMC procedure provides samples of the posterior distributions of the model parameters given the data. We use broad uniform priors for all the scalar model parameters. The prior on \mathbf{n} is uniform on the sphere. The prior on the noise-level parameter σ is a broad uninformative gamma distribution. Proposal distributions are Gaussians with standard deviations chosen manually to give suitable acceptance rates. The likelihood ratio at each iteration uses the Rician noise model. Samples of the model parameters come from a Metropolis-Hastings sampler; samples of σ from Gibbs sampling. Rough initial estimates for f , \mathbf{n} , d_{\parallel} , and d_{\perp} come from the DT fitted to the log measurements. We initialize R to the true value to speed up convergence. The burn-in period is 10,000 iterations, which ensures convergence, and the sampling interval is

Table 2
As Table 1, but from the Rician CRLB Objective Function

| $ \mathbf{G} \text{ (T m}^{-1}\text{)}$ | $\Delta \text{ (s)}$ | $\delta \text{ (s)}$ | $ \mathbf{q} \text{ (mm}^{-1}\text{)}$ | $t \text{ (s)}$ | $b \text{ (s mm}^{-2}\text{)}$ |
|------------------------------------------|----------------------|----------------------|-----------------------------------------|-----------------|--------------------------------|
| 0.200 | 0.024 | 0.019 | 998.1 | 0.017 | 17,370 |
| 0.097 | 0.027 | 0.016 | 408.4 | 0.021 | 3,580 |
| 0.200 | 0.012 | 0.007 | 358.4 | 0.009 | 1,216 |
| 0.200 | 0.012 | 0.007 | 357.1 | 0.009 | 1,205 |

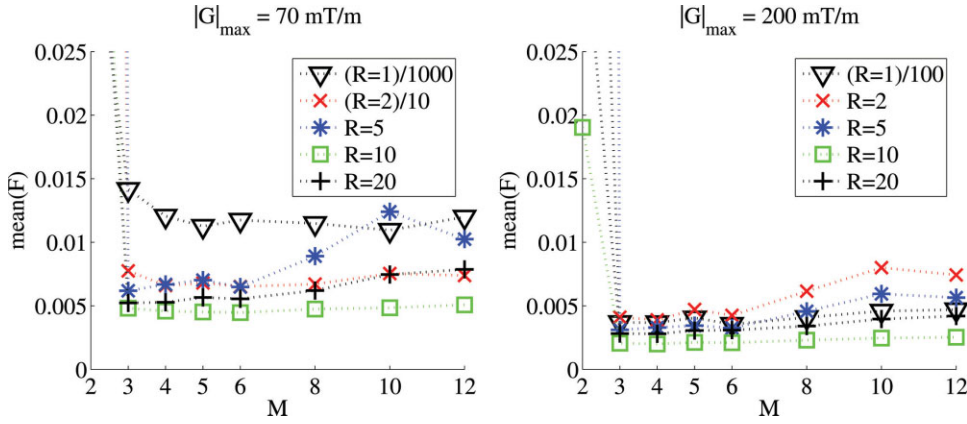


FIG. 1. Plots of the mean F over 500 orientations for optimized protocols against M at various R with $G_{\max} = 0.07$ (left) and 0.2 T m^{-1} (right). Note the scaling factors (see legends) required to reduce the values of $\text{mean}(F)$ to the range of the graphs at lower R ; for example, $\text{mean}(F) = 14$ for $G_{\max} = 0.07 \text{ T m}^{-1}$, $R = 1 \mu\text{m}$, and $M = 3$.

1,000, which gives approximately independent samples. Each run gathers 100 samples.

Figure 2 is the key result. The figure plots histograms of samples of the posterior distributions on R for each of the various true R at each G_{\max} . Each histogram combines output from 10 separate MCMC runs each with

a different setting of \mathbf{n} and independent noise, so the total number of samples is 1,000. In the top row of Fig. 2, each histogram comes from a separate protocol optimized for the specific value of R , which should give the best possible chance of recovering R precisely and accurately.

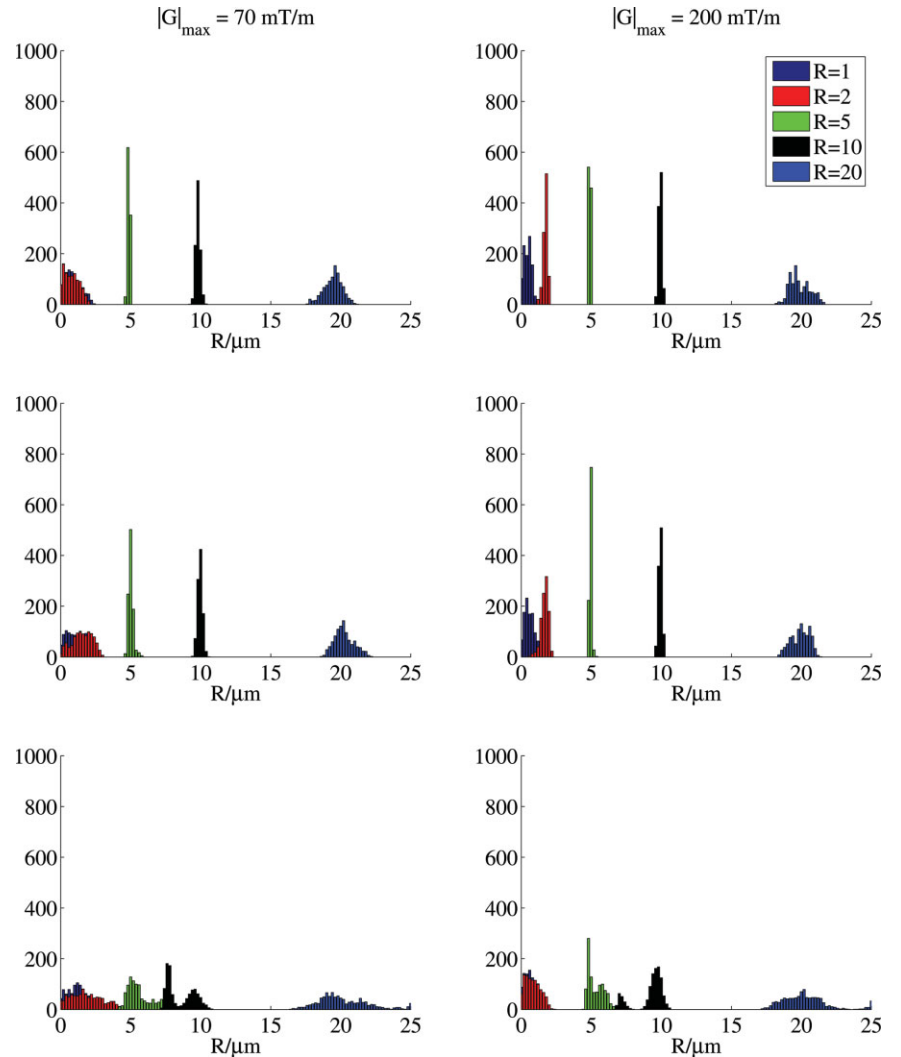


FIG. 2. Histograms of samples drawn from posterior distributions on R using MCMC. Top row: $\sigma_0 = 0.02$ and experiments with each setting of R use protocols optimized for that specific R . Middle row: as top row, but using a single protocol for all R . The protocol comes from minimizing the Rician CRLB objective function, F , averaged over a range of R . Bottom row: as middle row, but with $\sigma_0 = 0.05$.

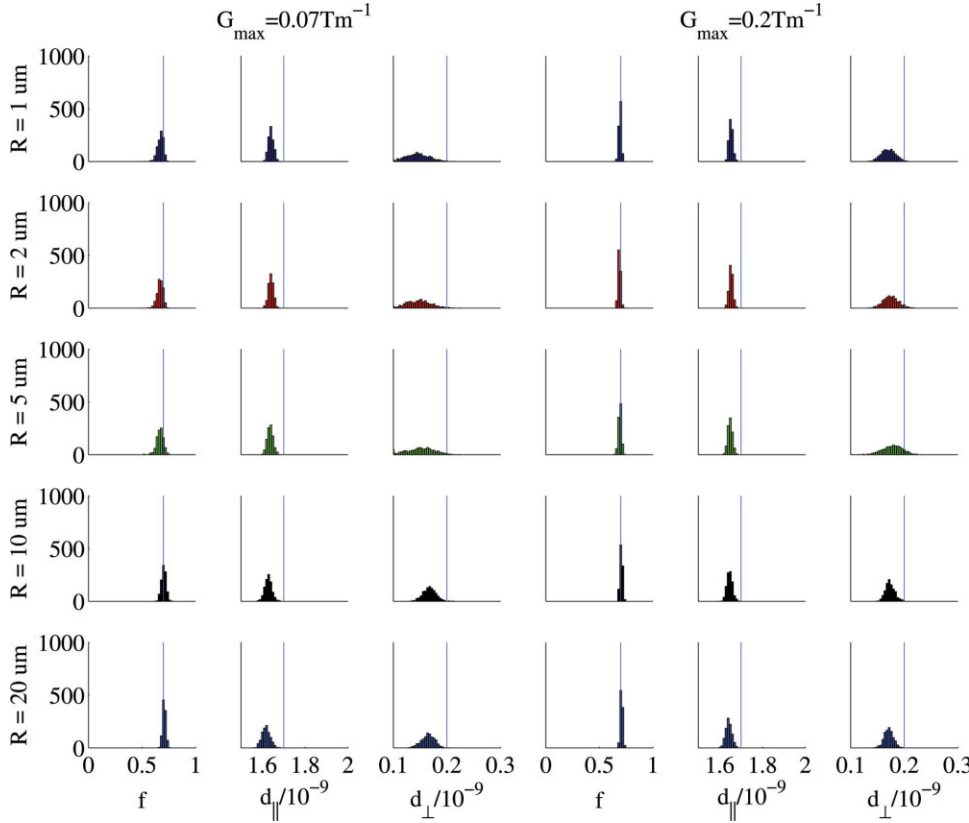


FIG. 3. Histograms of samples drawn from posterior distributions on the other scalar model parameters at various R using the single protocol used for the middle row of Fig. 2 with $\sigma_0 = 0.02$ throughout. The left three columns show f (left), d_{\parallel} (middle), and d_{\perp} (right) for $G_{\max} = 0.07 \text{ T m}^{-1}$ and the right three columns show the same for $G_{\max} = 0.2 \text{ T m}^{-1}$. The blue vertical line in each figure shows the true setting.

The axon radius is easiest to estimate when it is in the range 5–10 μm . Variance of the posterior increases as R increases or decreases from this range. For $G_{\max} = 0.07 \text{ T m}^{-1}$, $R = 1 \mu\text{m}$ and $R = 2 \mu\text{m}$ are identifiable as $R \in [0, 2] \mu\text{m}$, but are indistinguishable from each other and smaller radii. However, at $G_{\max} = 0.2 \text{ T m}^{-1}$, the posterior distributions for $R = 1$ and $2 \mu\text{m}$ separate and center on the correct values. The variance is higher at $20 \mu\text{m}$ than at 5 or $10 \mu\text{m}$ despite the value of F being similar (see Figure 1). However, F does not include terms for the orientation parameter \mathbf{n} , which becomes harder to estimate as R increases.

The results in the top row of Fig. 2 are an upper bound on what is achievable, since they come from protocols optimized individually for each R . The middle row of Fig. 2 shows posterior distributions using only one protocol at each G_{\max} . The single protocols come from minimizing F averaged over both S and each $R \in \{5, 10, 20\} \mu\text{m}$. For higher R , histograms are slightly broader, but the trends remain similar. The single protocols give slightly better results at low R than the protocols optimized for those specific R and the posterior distributions for $R = 1$ and $2 \mu\text{m}$ separate slightly even with $G_{\max} = 0.07 \text{ T m}^{-1}$. The CRLBs for R are much larger than those for the other model parameters when R is low and that term dominates the sum in Eq. [2]. At higher R , all the CRLBs have similar magnitude, so the optimization also minimizes the variance of the other model parameters, which appears to provide some advantage even for lower R . For comparison, the bottom row of Fig. 2 shows posterior distributions with $\sigma_0 = 0.05$, i.e. unweighted SNR of 20 when $\text{TE} = 0.08 \text{ sec}$, using a single protocol at each G_{\max} obtained by optimizing F averaged over $R \in \{5, 10, 20\} \mu\text{m}$.

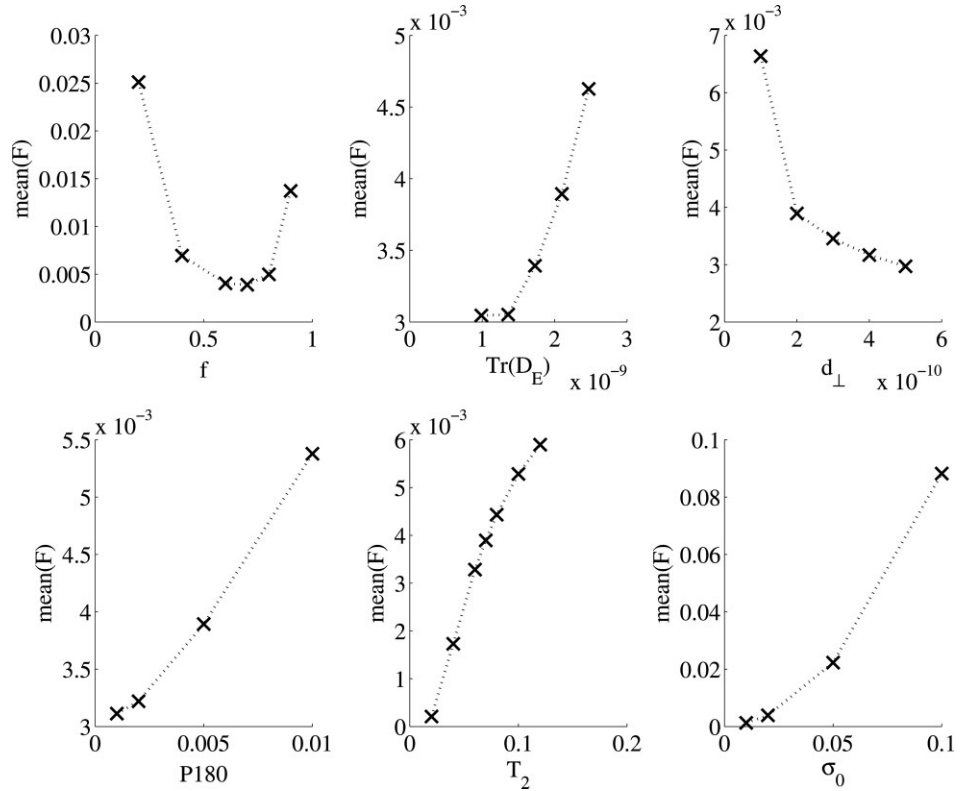
Figure 3 plots posterior distributions of the other scalar model parameters f , d_{\parallel} , and d_{\perp} for $G_{\max} = 0.07$ and 0.2 T m^{-1} . Posterior distributions on f are tight about the correct value of 0.7 at all radii. Both diffusivities have downward bias from their true settings of 1.7 and $0.2 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$; d_{\parallel} has lower standard error than d_{\perp} . The bias in recovered diffusivities is a consistent feature that becomes more marked as SNR decreases; it is a consequence of the Rician noise.

Dependencies

The final experiment investigates the dependence of the ability to measure direct microstructure features on the previously fixed model-parameter settings and other system variables. Figure 1 shows the dependence of F on both R and M ; F becomes larger as R decreases, which suggests that lower R is harder to recover. Figure 4 shows how F , averaged over S_{500} , depends on other variables of the model and system. The default settings in all the simulations are $G_{\max} = 0.2 \text{ T m}^{-1}$, $R = 2 \mu\text{m}$, $N = 30$, $M = 4$, $f = 0.7$, $d_{\parallel} = 1.7 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, $d_{\perp} = 0.2 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$, $\sigma_0 = 0.02$, $P180 = 0.005 \text{ s}$, and $T_2 = 0.07 \text{ s}$. Each plot in Fig. 4 shows the average F over S_{500} as one setting varies with the others fixed. Values of F are from protocols optimized separately at each setting of the parameter under investigation.

The value of F is minimum when f is close to 0.7 and increases sharply as f decreases (Fig. 4 top left). Examination of the CRLBs for each individual parameter (not shown) reveals that, although d_{\perp} becomes easier to estimate as f decreases, the dominant effect is that both f and R become harder to estimate. As the intrinsic diffusivity

FIG. 4. Plots of the mean F over 500 orientations as different parameters of the model and system vary with all the other settings fixed (see text). Values of $\text{mean}(F)$ are for protocols optimized separately for each individual data point. Top row: the volume fraction, f (left). The intrinsic diffusivity (middle); D_E is the extra cellular diffusion tensor, which has eigenvalues d_{\parallel} , d_{\perp} , and d_{\perp} , and d_{\parallel}/d_{\perp} remains fixed. The tortuosity in the extra cellular compartment (right). Bottom row: the length of the 180° pulse (left), T_2 (middle), and σ_0 (right).



of the system decreases, i.e., both d_{\parallel} and d_{\perp} decrease with d_{\parallel}/d_{\perp} constant (top middle in Fig. 4), F decreases, because R becomes easier to measure. A reduction in diffusivity has a very similar effect in increasing R , since both affect the timescale over which spins interact with restricting boundaries. As diffusivity decreases, the window of R we can recover also decreases; we can measure smaller axon radii more accurately in tissue with lower diffusivity, such as excised tissue samples. In contrast, as d_{\perp} increases with d_{\parallel} fixed (top right in Fig. 4), F comes down, which is mainly because d_{\perp} itself becomes more identifiable. Figure 4 (bottom left) shows the effect of reducing the length of the 180° pulse. Shorter P180 allows shorter diffusion times, which may give access to measurements that are more informative about smaller R . Dependence of F on P180 is not dramatic; f and R become slightly more identifiable as P180 decreases. Precision appears to increase as T_2 decreases (Fig. 4 bottom middle), because the SNR at $TE = 0.08$ s is fixed at 50 so the optimization finds greater benefit in estimating all parameters by reducing TE. Finally, Fig. 4 (bottom right) shows that F is almost exactly proportional to σ_0^2 , which suggests that the effect of the Rician noise does not vary greatly over the range of σ_0 .

DISCUSSION

The experiment-design optimization is essential to obtain the kinds of acquisition protocols this article studies. It is difficult to imagine how we might choose an orientationally invariant protocol for the model studied here or similar models, without such an optimization. However, suppose, for example, we take $M = 4$ and construct a protocol in a similar way to (7). If we take each combination of

$|\mathbf{G}| \in \{G_{\max}/2, G_{\max}\}$ and $\Delta \in \{\Delta_{\max}/2, \Delta_{\max}\}$ with δ fixed short at, say, 0.005 s, the mean F over 500 orientations, at the best setting of Δ_{\max} , is around 30 times larger than for the optimized protocol with the same G_{\max} . The optimization consistently makes major improvements to parameter estimates over ad-hoc protocols.

The experiment-design optimization adapts naturally to other models of diffusion in brain tissue. The same procedure can optimize sampling for more complex models incorporating other effects, as discussed at the end of the ‘‘Simple Model’’ section. Moreover, optimal sampling for simpler and more familiar models, such as single or multiple DT models, is still the subject of debate and the same technique can optimize sampling strategies for them.

For the simple model we investigate here, protocols with $M = 4$ consistently give lower F than other M . It seems likely that the optimal M depends on the number of model parameters and this model has four scalar parameters and \mathbf{n} . One possible refinement of the protocols we consider here, which may improve orientational invariance at no cost, is to separate the gradient directions slightly for each of the N sets of M measurements. One possible implementation of this idea uses recent work by Cook et al. (30), which shows how to divide MN evenly distributed directions into M evenly distributed subsets of N directions. We can assign a separate combination of $|\mathbf{G}|$, Δ , and δ to each subset. Experiments (not shown) with such protocols show little improvement over the class of protocols used here. This suggests that $N = 30$ is sufficient for orientational invariance. However, for more complex models that require larger M , or smaller MN , these subset-based protocols may prove advantageous. Further improvement may come from allowing different numbers of measurements

in each subset. For example, high angular resolution may be more important for measurements with higher diffusion weighting.

The simulation results suggest feasibility of measuring direct microstructure features, such as axon density and radius distribution, *in vivo*. With typical diffusivities and packing densities in brain tissue, radii of 5–10 μm are most identifiable. At higher R , \mathbf{n} becomes harder to estimate, which makes estimates of R less precise. At lower R , we can identify R as low, but cannot distinguish low values of R from one another. The simulations use only single values of R , but real brain tissue has a distribution in the range 0.25–10 μm . The actual distribution of radii varies among white matter structures. Assaf et al. (7,23) fit simple continuous models for the radius distribution. The simulations using single radii reveal which parts of fitted distributions are likely to be reliable and which rely on the shape of the model. To optimize the acquisition for recovering a continuous distribution, one option is to replace the term in R in Eq. [2] with terms for the parameters of the radius distribution and optimize in the same way. The optimum M may increase to accommodate the extra parameters. Dependence of F on the mean of a continuous distribution model is likely to be similar to the dependence on the single R in the current model. Dependence of F on the distribution variance will reflect the proportion of the distribution that lies in the range of recoverable values of R . However, simpler optimizations like those we perform here may be sufficient to optimize the acquisition for fitting continuous distributions. Protocols optimized for any R in a range, say 0.25–10 μm , (by averaging F over the range) should provide the sensitivity required to fit any continuous distribution within that same range. However, the optimization can become unstable for low settings of R , so the best strategy may be to optimize for the range of most recoverable R only, as for the results in Fig. 2 (middle and bottom).

Higher gradient strength increases the ability to distinguish smaller axon radii. However, inspection of the optimal protocols at various G_{max} suggests that the main advantage of higher gradient strength is to allow acquisition of similar measurements with shorter TE and hence higher SNR. The settings of other parameters of the system also affect the precision and accuracy of estimates, as Fig. 4 shows. The most profound effect is the dependence of the range of measurable R on the intrinsic diffusivity. As diffusivity increases, the values of R we can estimate also increase. The simulations suggest that the lower end of the axon-radius range is difficult to distinguish *in vivo*. However, excised tissue samples have lower diffusivity, mainly because they are usually at room, rather than body, temperature, so it may be easier to measure the full range of radii.

The experiment-design optimization presents a variety of new opportunities for further work. As discussed earlier, the optimization procedures adapt simply to a variety of alternative models. Moreover, it is simple to incorporate prior information on the distribution of parameter settings to further fine tune the optimization. In particular, in applications where information about the distribution of fiber orientations is available, such as *ex vivo* nerve-tissue imaging or *in vivo* optic-nerve or spinal-cord imaging, it is simple to adapt the optimization to exploit that knowledge.

The method should improve future imaging protocols and studies that use them in a wide range of applications.

The feasibility study into measuring direct-microstructure features provides compelling results, but clearly requires further study. The next step is to acquire data from *ex vivo* and *in vivo* samples and compare with histology in a similar way to (7,10). The simple model is an extreme simplification of the geometry of real brain tissue. However, its estimated parameters may still provide crude measurements that correlate closely enough with cell dimensions to make useful biomarkers. Preliminary experiments (not shown) suggest that the model behaves sensibly under departures. In particular, if the system contains two distinct axon radii or a gamma distribution of radii, the posterior on the single R tends to have a single peak near the mean of the true R distribution. The posterior distributions on a single R (as shown in Fig. 2) tend to be tighter about the mean than non trivial true distributions of R . If the true distribution of fiber orientations departs from the δ -function the model assumes, the fitted R becomes biased upward. Compartmental T_2 introduces bias in the fitted f , but has less effect on the estimates of R and the diffusivity parameters.

Microstructure features such as axon-radius distribution, orientation distribution, permeability, glial-cell compartments, compartmentalized T_2 , and compartmentalized intrinsic diffusivity are simple to incorporate into the model. Future experiments with scanner data will determine which parameters are required to model brain tissue sufficiently well.

Direct-microstructure biomarkers have clear potential benefits for monitoring tissue changes in development, disease or recovery over non specific biomarkers derived from the DT. Long term, these kinds of features may also benefit tractography and connectivity mapping. For example, knowledge of the axon-radius distribution in each fiber tract could help resolve ambiguities at fiber crossings and may provide better definitions of connectivity based on information-transfer potential. Conversely, tractography may assist estimation of direct microstructure features. Microstructure is likely to remain consistent along fibers, so we can exploit spatial coherence along reconstructed pathways to provide extra signal for fitting.

This study considers only the PGSE pulse sequence, but other pulse sequences can provide diffusion MRI measurements that may be better for estimating certain features of the microstructure. Both STEAM (12) and steady-state free precession (31) allow measurements with much longer diffusion times by avoiding T_2 decay outside the gradient pulses. Twice-refocused spin-echo sequences (32) can be more economic than PGSE and produce less eddy-current distortions. Other sequences that vary the gradient orientation during the gradient pulses (33) or use multiple pulse pairs (34) may also prove advantageous for measuring direct microstructure features. With the exception of STEAM, however, all the pulse sequences above have different tunable settings to the PGSE sequence and models that relate those settings to the parameters of complex particle-displacement patterns still require development. The experiment-design optimization developed here adapts simply to these alternative pulse sequences once models are available.

APPENDIX : CRAMER-RAO LOWER BOUND

The general form of the Fisher information matrix J has (i, j) -th element

$$J_{ij} = E \left(\frac{\partial^2 L}{\partial w_i \partial w_j} \right), \quad [A1]$$

where L is the log likelihood of the measurement, given the model parameters, for an appropriate noise model and E denotes expectation given that noise model. For Gaussian noise,

$$J_{ij} = \sigma^{-2} \sum_{k=1}^K \frac{\partial A_k}{\partial w_i} \frac{\partial A_k}{\partial w_j}, \quad [A2]$$

where $A_k, k = 1, \dots, K$, ($K = NM$ in this application) is the model's prediction of the signal for the k -th combination of independent variables (\mathbf{G}_k, δ_k and Δ_k in this application) for parameter settings w_1, \dots, w_W . Equation [A2] is simple to derive from Eq. [A1] (35). The CRLB for w_i is the i -th diagonal element of J^{-1} , $(J^{-1})_{ii}$, which we use in Eq. [2] to replace σ_i^2 in Eq. [1].

The noise on MRI measurements is not Gaussian, but Rician (36), so that

$$P(\tilde{A}) = \frac{\tilde{A}}{\sigma^2} I_0 \left(\frac{A\tilde{A}}{\sigma^2} \right) \exp \left(-\frac{A^2 + \tilde{A}^2}{2\sigma^2} \right), \quad [A3]$$

where $I_i, i = 0, 1, 2, \dots$, are the modified Bessel functions of the first kind; \tilde{A} and A are the measurement and signal (predicted by the model), respectively. The log likelihood of a set of K measurements comes from taking logs of Eq. [A3] and summing over all measurements to obtain

$$L_R = \sum_{k=1}^K \left[\log \tilde{A}_k - 2 \log \sigma + \log I_0 \left(\frac{A_k \tilde{A}_k}{\sigma^2} \right) - \frac{A_k^2 + \tilde{A}_k^2}{2\sigma^2} \right] \quad [A4]$$

A Rician CRLB comes from the general expression for the Fisher information matrix in Eq. [A1] using the Rician log likelihood, L_R in Eq. [A4]. The substitution requires the expectations of the second derivatives of L_R with respect to each w_i and σ . For example,

$$E \left(\frac{\partial^2 L_R}{\partial w_j \partial w_k} \right) = \int_0^\infty \frac{\partial^2 L_R}{\partial w_j \partial w_k} P(\tilde{A}) d\tilde{A}. \quad [A5]$$

The following expectations are straightforward to derive, although the algebra is tedious:

$$\begin{aligned} E \left(\frac{\partial^2 L_R}{\partial w_i \partial w_j} \right) &= \sum_{k=1}^{NM} \frac{1}{\sigma^4} \frac{\partial A_k}{\partial w_i} \frac{\partial A_k}{\partial w_j} (Z_k - A_k^2) \\ E \left(\frac{\partial^2 L_R}{\partial w_i \partial \sigma} \right) &= \sum_{k=1}^{NM} \frac{2A_k}{\sigma^5} \frac{\partial A_k}{\partial w_i} (Z_k - A_k^2 - \sigma^2) \\ E \left(\frac{\partial^2 L_R}{\partial \sigma^2} \right) &= \sum_{k=1}^{NM} \frac{4A_k^2}{\sigma^6} \left(\frac{\sigma^4}{A_k^2} - \sigma^2 - A_k^2 + Z_k \right), \end{aligned} \quad [A6]$$

where

$$Z_k = \int_0^\infty A_k^2 I_1^2 \left(\frac{A_k \tilde{A}}{\sigma^2} \right) I_0^{-2} \left(\frac{A_k \tilde{A}}{\sigma^2} \right) P(\tilde{A}) d\tilde{A}. \quad [A7]$$

The function Z_k does not have closed form and requires numerical computation. The optimization procedure in the "Experiment Design Optimization" section precomputes a look-up table of sampled values and uses linear interpolation to estimate Z_k during optimization. Andersson (37) derives the Rician CRLB in a similar way for the specific case of the DT model. The new expressions earlier are for the general case and do not depend on a specific model.

ACKNOWLEDGMENTS

The author thanks Derek Jones and Yaniv Assaf for useful discussions on this material. The EPSRC support Dr. Alexander with grants GRIT22858101 and EPIE06459011.

REFERENCES

1. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative diffusion tensor MRI. *J Magn Reson Ser B* 1996;111:209–219.
2. Sotak CH. The role of diffusion tensor imaging in the evaluation of ischemic brain injury—A review. *NMR Biomed* 2002;15:561–569.
3. Klingberg T, Hedehus M, Temple E, Salz T, Gabrieli JDE, Moseley ME, Poldrack RA. Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. *Neuron* 2000;25:493–500.
4. Toosy T, Werring DJ, Orrell RW, Howard RS, King MD, Barker GJ, Miller DH, Thompson AJ. Diffusion tensor imaging detects corticospinal tract involvement at multiple levels in amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* 2003;74:1250–1257.
5. McKinstry RC, Mathur A, Miller JH, Ozcan A, Snyder AZ, Scheff GL, Alml CR, Shiran SI, Conturo TE, Neil JJ. Radial organization of developing preterm human cerebral cortex revealed by non-invasive water diffusion anisotropy MRI. *Cereb Cortex* 2002;12:1237–1243.
6. Pfefferbaum A, Sullivan EV, Hedehus M, Lim KO, Adalsteinsson E, Moseley M. Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. *Magn Reson Med* 2000;44:259–268.
7. Assaf Y, Freidlin RZ, Basser PJ. The measurement of the axon diameter distribution in white matter using diffusion MR methods. In *Proceedings of ISMRM, Miami*. ISMRM; 2005, p 842.
8. Callaghan PT, Eccles CD, Xia Y. NMR microscopy of dynamic displacements: k -space and q -space imaging. *J Phys E Sci Instrum* 1988;21:820.
9. Stanisz GJ, Szafer A, Wright GA, Henkelman M. An analytical model of restricted diffusion in bovine optic nerve. *Magn Reson Med* 1997;37:103–111.
10. Weng, J-C, Chen, J-H, Kuo, L-W, Wedeen VJ, Tseng, W-Y. Maturation-dependent microstructure length scale in corpus callosum of fixed rat brain by magnetic resonance diffusion-diffraction. *Magn Reson Imag* 2007;25:78–86.
11. Stejskal EO, Tanner TE. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J Chem Phys* 1965;42:288–292.
12. Merboldt KD, Hanicke W, Frahm J. Diffusion imaging using stimulated echoes. *Magn Reson Med* 1992;19:233–239.
13. Callaghan PT, Coy A, MacGowan D, Packer KJ, Zelaya FO. Diffraction-like effects in NMR diffusion studies of fluids in porous solids. *Nature* 1991;351:467–469.
14. Avram L, Assaf Y, Cohen Y. The effect of rotational angle and experimental parameters on the diffraction pattern and micro-structural information obtained from q -space diffusion NMR: Implication for diffusion in white matter fibers. *J Magn Reson* 2004;169:30–38.
15. Topgaard D, Soderman O. Experimental determination of pore shape and size using q -space NMR microscopy in the long diffusion-time limit. *Magn Reson Imag* 2003;21:69–76.
16. Peled S, Cory DG, Raymond SA, Kirschner DA, Jolesz FA. Water diffusion, T_2 , and compartmentation in frog sciatic nerve. *Magn Reson Med* 1999;42:911–918.
17. Ong HH, Wright AC, Wehrli SL, Souza A, Schwartz ED, Saha PK, Wehrli FW. Q -space propagator maps of mouse spinal cord provide insight into regional axonal architecture. In *Proceedings of ISMRM, Seattle*. ISMRM; 2006, p 640.

18. Lätt J, Nilsson M, Malmberg C, Rosquist H, Wirestam R, Ståhlberg F, Topgaard D, Brockstedt S. Accuracy of q -space related parameters in MRI: Simulations and phantom measurements. *IEEE Trans Med Imag* 2007;26:1437–1447.
19. Assaf Y, Freidlin RZ, Rohde GK, Basser PJ. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magn Reson Med* 2004;52:965–978.
20. Neuman CH. Spin echo of spins diffusing in a bounded medium. *J Chem Phys* 1974;60:4508–4511.
21. Van Gelderen P, DesPres D, van Zijl PCM, Moonen CTW. Evaluation of restricted diffusion in cylinders. Phosphocreatine in rabbit leg muscle. *J Magn Reson Ser B* 1994;103:255–260.
22. Murday JS, Cotts RM. Self-diffusion coefficient of liquid lithium. *J Chem Phys* 1968;48:4938–4945.
23. Assaf Y, Basser PJ. Non parametric approach for axon diameter distribution estimation from diffusion measurements. In *Proceedings of ISMRM, Miami*. ISMRM; 2007, p 1536.
24. Alexander DC. Axon radius measurements in vivo from diffusion MRI: a feasibility study. In *Proceedings of ICCV Workshop on MMBIA, Rio de Janeiro*, 2007, IEEE.
25. Sen PN, Basser PJ. A model for diffusion in white matter. *Bio J* 2005;89:2927–2938.
26. Hansen B, Vestergaard-Poulsen P. Mapping the parameter space of a T_2 -dependent model of water diffusion MR in brain tissue. *Magn Reson Imag* 2006;24:1031–1038.
27. Nicholson C. Diffusion and related transport mechanisms in brain tissue. *Rep Prog Phys* 2001;64:815–884.
28. Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn Reson Med* 1999;42:515–525.
29. Zelinka I. SOMA—Self organizing migrating algorithm. In Babu BV and Onwubolu G, editors, *New optimization techniques in engineering*. Springer, 2004.
30. Cook PA, Boulby PA, Symms MR, Alexander DC. Optimal acquisition orders of diffusion-weighted MRI measurements. *J Magn Reson Imag* 2007;25:1051–1058.
31. Miller KL, Hargreaves BA, Gold GE, Pauly JM. Steady-state diffusion-weighted imaging of in vivo knee cartilage. *Magn Reson Med* 2004;51:394–398.
32. Reese TG, Heid O, Weisskoff RM, Wedeen VJ. Reduction of eddy-current-induced distortion using a twice-refocussed spin echo. *Magn Reson Med* 2003;49:177–182.
33. Wedeen VJ, Dai G, Tseng I, Wang R, Benner T. Diffusion encoding with 2D gradient trajectories yields natural contrast for 3D fiber orientation. In *Proceedings of ISMRM, Seattle*. Berkeley, USA: ISMRM; 2006, p 851.
34. Ozarslan E, Basser PJ. MR diffusion/diffraction phenomenon in multi-pulse-field-gradient experiments. *J Magn Reson* 2007;188:285–294.
35. Lindgren BW. *Statistical theory*. New York: Chapman and Hall; 1993.
36. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. *Med Phys* 1985;12:232–233.
37. Andersson J. Maximum likelihood estimation of diffusion parameters with a Rician noise model. In *Proceedings of the ISMRM, Berlin*: ISMRM; 2007, p 1881.