Estimation of Tensors and Tensor-Derived Measures in Diffusional Kurtosis Imaging[1]

Ali Tabesh[a,2], **Jens H. Jensen**[a,b], **Babak A. Ardekani**[c,d], and **Joseph A. Helpern**[a,b,c,d]

[a]Center for Biomedical Imaging, Department of Radiology, New York University School of Medicine, 660 First Avenue New York, NY 10016 USA

[b]Department of Physiology and Neuroscience, New York University School of Medicine, 550 First Ave, New York, NY 10016 USA

[c]Center for Advanced Brain Imaging, The Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd, Orangeburg, NY 10962 USA

[d]Department of Psychiatry, New York University School of Medicine, 550 First Ave, New York, NY 10016 USA

Abstract

This paper presents two related advancements to the diffusional kurtosis imaging (DKI) estimation framework to increase its robustness to noise, motion, and imaging artifacts. The first advancement substantially improves the estimation of diffusion and kurtosis tensors parameterizing the DKI model. Rather than utilizing conventional unconstrained least squares (LS) methods, the tensor estimation problem is formulated as linearly constrained linear LS, where the constraints ensure physically and/or biologically plausible tensor estimates. The exact solution to the constrained problem is found via convex quadratic programming methods or, alternatively, an approximate solution is determined through a fast heuristic algorithm. The computationally more demanding quadratic programming-based method is more flexible, allowing for an arbitrary number of diffusion weightings and different gradient sets for each diffusion weighting. The heuristic algorithm is suitable for real-time settings such as on clinical scanners, where run time is crucial. The advantage offered by the proposed constrained algorithms is demonstrated using *in vivo* human brain images. The proposed constrained methods allow for shorter scan times and/or higher spatial resolution for a given fidelity of the DKI parametric maps. The second advancement increases the efficiency and accuracy of the estimation of mean and radial kurtoses by applying exact closed-form formulae.

Keywords

MRI; diffusion; kurtosis; non-Gaussian; brain

Introduction

Diffusion of water molecules in biological tissues is conventionally quantified via diffusion tensor imaging (DTI) (1). DTI is a valuable tool for noninvasive characterization of tissue microstructural properties (2–3). The DTI model uses a Gaussian approximation to the probability distribution governing the random displacement of water molecules. In many

[2] Corresponding author. Phone: +1-212-263-3344; fax: +1-212-263-4830; ali.tabesh@nyumc.org.

biological tissues, however, the displacement probability distribution can deviate considerably from a Gaussian form. Several techniques have been proposed to characterize non-Gaussian diffusion, with diffusion spectrum imaging (DSI) (4) being arguably the most comprehensive. Although it is being used in human research, the application of DSI as a routine clinical protocol has been limited by its extended acquisition time and hardware requirements.

Diffusional kurtosis imaging (DKI) is a clinically feasible extension of DTI which enables the characterization of non-Gaussian diffusion by estimating the kurtosis of the displacement distribution (5–8), in addition to the estimation of the standard DTI-derived parameters. DKI has shown promising results in studies of human brain aging (9), tumor characterization in gliomas (10) and head and neck cancers (11), and rodent brain maturation (12). Moreover, the additional information provided by DKI has been exploited to resolve intravoxel fiber crossings (13), which could help to improve upon DTI-based fiber tracking methods.

The DKI model is parameterized by the diffusion tensor (DT) and kurtosis tensor (KT) from which several rotationally-invariant scalar measures are extracted. The most common DT-derived measures are mean, axial, and radial diffusivity (2,14), as well as fractional anisotropy (FA) (2); and the KT-derived measures are axial, radial, and mean kurtoses (6–7,15). The interpretability of these metrics is influenced by the estimation accuracy of the tensors. Noise, motion, and imaging artifacts can introduce errors into the estimated tensors. Sufficiently large errors can cause the tensor estimates to be physically and/or biologically implausible. For instance, the directional diffusivities may become negative, that is, the DT may become non-positive definite (NPD). A well-known consequence of NPD DT estimates is that FA values, which in theory should range between 0 and 1, may exceed 1, particularly in high FA regions of the brain such as the corpus callosum (16). Inaccuracies in the estimated tensors may also drive the directional kurtoses outside of an acceptable range. Empirical evidence in the brain as well as idealized multi-compartment diffusion models suggest that directional kurtoses should typically be positive and should not exceed a certain level depending on tissue complexity (6). The maximum allowable kurtosis is also influenced by the maximum b-value used in image acquisition (8,13).

In our previous work, the DT and KT were estimated using unconstrained nonlinear least squares (UNLS) (7). In a related work, higher-order DTs were estimated using unconstrained linear least squares (ULLS) (17). These unconstrained schemes do not guarantee acceptable tensor estimates. To address this drawback in the context of DTI, the Cholesky decomposition has been utilized to impose the non-negative definiteness constraint on the DT (16,18) using either the ULLS or UNLS algorithms. Moreover, a parameterization has been proposed in a UNLS framework to guarantee a positive diffusivity function in a fourth-order tensor-only model of the diffusion signal (19). While these methods outperform other algorithms for imposing the positive diffusivity function in the context of a second order-only and a fourth order-only diffusion signal model, their generalization to impose the constraints on the DKI signal model is not straightforward.

We propose that the tensor estimation problem be cast as linear least squares (LS) subject to linear constraints. The constraints ensure that the directional diffusivities and kurtoses along the imaged gradient directions remain within a physically and biologically plausible range. The proposed constrained linear LS (CLLS) formulation yields a convex objective function that permits efficient solutions via convex quadratic programming or a fast heuristic algorithm. The two algorithms proposed for solving the CLLS problem strike different tradeoffs between the speed and exactness of the solution as well as algorithm flexibility. The quadratic programming-based (CLLS-QP) algorithm exactly satisfies the constraints and can also handle an arbitrary number of diffusion weightings and different gradient sets

for each diffusion weighting, but it does so at a moderately high computational cost. On the contrary, the heuristic (CLLS-H) algorithm produces an approximation to the optimal solution, but it accomplishes this at almost no computational overhead compared to the ULLS. Both constrained algorithms substantially improve the estimation fidelity of the tensors.

Once the tensors are estimated, they are utilized to determine the scalar diffusion and kurtosis measures. In our previous work, mean kurtosis (MK) was being approximated as the average of directional kurtoses (5–7). Here, we present an exact analytical formula for the MK, expressing it as an explicit function of the DT and KT. We also present an exact analytical expression for the radial kurtosis, defining it as the diffusional kurtosis averaged over all directions perpendicular to the DT eigenvector with the largest eigenvalue. Our definition is consistent with the definition of radial diffusivity in DTI and the definitions given in (8,20), but is different from that of (15). The analytic formulae for the MK and radial kurtosis were previously quoted in (8) but here for the first time we present their derivations and discuss important practical details of their implementation. The application of these formulae improves both the accuracy and efficiency of DKI parameter estimation.

Theory

Constrained LS Formulation

For direction vector **n** and diffusion weighting b, the DKI approximation to diffusion signal intensity $S(\mathbf{n},b)$ is given by (6)

$$\ln\left[\frac{S(\mathbf{n}, b)}{S_0}\right] = -b\sum_{i=1}^{3}\sum_{j=1}^{3}n_i n_j D_{ij} + \frac{1}{6}b^2\overline{D}^2\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}n_i n_j n_k n_l W_{ijkl},$$ [1]

where S_0 is the signal intensity for $b = 0$, D_{ij} and W_{ijkl} are the elements of the second-order DT **D** and fourth-order kurtosis tensor **W**, respectively, and $\overline{D} = (1/3)\mathrm{tr}(\mathbf{D})$ is mean diffusivity, where tr(.) denotes matrix trace. For the classic Stejskal-Tanner sequence (21), the b-value is given by $b = (\gamma\delta g)^2(\Delta - \delta/3)$, where Δ is the diffusion time, δ is the pulse duration, g is the diffusion-sensitizing gradient amplitude, and γ is the nuclear spin gyromagnetic ratio. Diffusivity $D(\mathbf{n})$ and kurtosis $K(\mathbf{n})$ along direction **n** are given by

$$D(\mathbf{n}) = \sum_{i=1}^{3}\sum_{j=1}^{3}n_i n_j D_{ij},$$ [2]

and

$$K(\mathbf{n}) = \frac{\overline{D}^2}{D(\mathbf{n})^2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3}n_i n_j n_k n_l W_{ijkl}.$$ [3]

Both **D** and **W** are symmetric, that is, they are invariant to permutations of their indices. Consequently, **D** has 6 and **W** has 15 independent parameters.

To ensure that the diffusivity and kurtosis parameters are physically and biologically plausible, $D(\mathbf{n})$ and $K(\mathbf{n})$ must satisfy

$$D(\mathbf{n}) \geq 0,$$ [4]

and

$$K_{\max}(\mathbf{n}) \geq K(\mathbf{n}) \geq K_{\min},$$ [5]

and

$$K_{\max}(\mathbf{n}) = \frac{C}{b_{\max} D(\mathbf{n})}, \quad 3 \geq C \geq 0,$$ [6]

where $\mathbf{n} \in \mathbf{N}$, with $\mathbf{N}$ denoting the set of gradient directions used in acquisition. The constraints in Eq. [4] ensure that $\mathbf{D}$ satisfies the necessary conditions for non-negative definiteness. The constraints in Eq. [5] arise from the fact that biologically relevant tissue geometries impose restrictions on physically acceptable directional kurtoses (6). While the theoretically feasible minimum kurtosis of a probability distribution is -2, multi-compartment diffusion models and empirical evidence in the brain suggest a super-Gaussian displacement distribution; that is, $K_{\min} = 0$ (6). Although scenarios are conceivable under which $S(\mathbf{n},b)$ can be an increasing function of b (22), this has never been observed in biological tissues. Thus, it is reasonable to require that the estimated $S(\mathbf{n},b)$ be strictly decreasing functions of b for all $\mathbf{n} \in \mathbf{N}$. The maximum $K(\mathbf{n})$ in Eq. [6] with $C = 3$ ensures that this condition is satisfied in the range of b-values used for data acquisition (13). Smaller C values may be used to further constrain the directional kurtoses.

The objective in the estimation problem is to find $\mathbf{D}$ and $\mathbf{W}$ such that the right-hand side of Eq. [1] closely matches its left-hand side, while the constraints in Eqs. [4] and [5] are satisfied. The estimation problem may therefore be formulated as

$$\begin{aligned} &\text{Minimize } \|\mathbf{AX} - \mathbf{B}\|^2 \\ &\text{such that } \mathbf{CX} \leq \mathbf{d}, \end{aligned}$$ [7]

where

$$\mathbf{X} = [D_{11}\ D_{22}\ D_{33}\ D_{12}\ D_{13}\ D_{23}\ V_{1111}\ \ldots\ V_{1112}\ \ldots\ V_{1122}\ \ldots\ V_{1233}]^T$$ [8]

is a 21×1 vector of unknowns, and

$$V_{ijkl} = \overline{D}^2 W_{ijkl}.$$ [9]

The change of variables in Eq. [9] is made to permit a *linear* LS formulation. Note that $\mathbf{V}$ is reminiscent of the fourth-order tensor in the formulation of (17).

Matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \begin{bmatrix} -b_1 \mathbf{A}_D^{(1)} & \frac{1}{6} b_1^2 \mathbf{A}_K^{(1)} \\ \vdots & \vdots \\ -b_M \mathbf{A}_D^{(M)} & \frac{1}{6} b_M^2 \mathbf{A}_K^{(M)} \end{bmatrix},$$ [10]

where M is the number of nonzero b-values, and row k of $\mathbf{A}_D^{(m)}$ and $\mathbf{A}_K^{(m)}$ are given by

$$\begin{aligned} [\mathbf{A}_D^{(m)}]_k &= [(n_{k1}^{(m)})^2\ (n_{k2}^{(m)})^2\ (n_{k3}^{(m)})^2\ n_{k1}^{(m)} n_{k2}^{(m)}\ n_{k1}^{(m)} n_{k3}^{(m)}\ n_{k2}^{(m)} n_{k3}^{(m)}], \\ [\mathbf{A}_K^{(m)}]_k &= [(n_{k1}^{(m)})^4\ \ldots\ 4(n_{k1}^{(m)})^3 n_{k2}\ \ldots\ 6(n_{k1}^{(m)})^2 (n_{k2}^{(m)})^2\ \ldots\ 12 n_{k1}^{(m)} n_{k2}^{(m)} (n_{k3}^{(m)})^2], \end{aligned}$$ [11]

where $\mathbf{n}_k^{(m)} = [n_{k1}^{(m)} \ n_{k2}^{(m)} \ n_{k3}^{(m)}]^T$. Note that each diffusion weighting b_m may correspond to a distinct gradient set $\mathbf{N}^{(m)} = \{\mathbf{n}_1^{(m)}, \ldots, \mathbf{n}_{N_m}^{(m)}\}$. The number of rows of $\mathbf{A}$ is thus given by

$$N = \sum_{m=1}^{M} N_m.$$

Vector $\mathbf{B}$ is given by

$$\mathbf{B} = [\ln(S_1/S_0) \ \ldots \ \ln(S_N/S_0)]^T, \qquad [12]$$

where $S_k = S(\mathbf{n}_i^{(m)}, b_m)$ is the diffusion signal intensity for gradient direction $\mathbf{n}_i^{(m)}$ and diffusion weighting b_m, and

$$k = \sum_{j=1}^{m-1} N_j + i.$$

Matrix $\mathbf{C}$ representing the linear constraints is given by

$$\mathbf{C} = \begin{bmatrix} -\mathbf{A}_D & \mathbf{0} \\ \mathbf{0} & -\mathbf{A}_K \\ -(C/b_{max})\mathbf{A}_D & \mathbf{A}_K \end{bmatrix}, \qquad [13]$$

where

$$\mathbf{A}_D = \begin{bmatrix} \mathbf{A}_D^{(1)} \\ \vdots \\ \mathbf{A}_D^{(M)} \end{bmatrix} \qquad [14]$$

and

$$\mathbf{A}_K = \begin{bmatrix} \mathbf{A}_K^{(1)} \\ \vdots \\ \mathbf{A}_K^{(M)} \end{bmatrix}. \qquad [15]$$

Vector $\mathbf{d}$ is given by

$$\mathbf{d} = \begin{bmatrix} \mathbf{0} & K_{min}D(\mathbf{n}_1^{(1)})^2 & \cdots & K_{min}D(\mathbf{n}_{N_1}^{(1)})^2 & \cdots & K_{min}D(\mathbf{n}_1^{(M)})^2 & \cdots & K_{min}D(\mathbf{n}_{N_M}^{(M)})^2 & \mathbf{0} \end{bmatrix}^T, \qquad [16]$$

where the zero vectors are $1 \times N$-dimensional.

Two comments regarding the above formulation are in order:

1. At least *two* nonzero b-values and 15 gradient directions per b-value are required for the solution of Eq. [7]. The need for two b-values can also be seen from Eq. [1], where to find the two unknowns $D(\mathbf{n})$ and $K(\mathbf{n})$, at least two signal measurements (obtained with two different b-values) are needed.

2. For the typical choice $K_{min} = 0$, Eq. [7] provides an exact formulation of the estimation problem. If $K_{min} \neq 0$, the $D(\mathbf{n}_n^{(m)})$ terms in **d** need to be approximated. This may be accomplished via estimating **D** once and plugging that estimate into Eq. [2]. The estimate of **D** may be obtained from the ULLS estimate of **X**, denoted as $\hat{\mathbf{X}}$, given by

$$\widehat{\mathbf{X}} = \mathbf{A}^+\mathbf{B},$$

[17]

where $\mathbf{A}^+$ denotes the pseudoinverse of **A**. The $D(\mathbf{n}_n^{(m)})$ terms can be further refined using the CLLS-based estimate of **D**.

Quadratic Programming (CLLS-QP) Algorithm

The problem defined by Eq. [7] is a convex quadratic programming problem whose solution can be determined via standard algorithms. Two common classes of algorithms for solving this problem are active set and interior-point methods (23). Here we used the active set method which works as follows. First, the algorithm is initialized with a feasible point, if one exists. A feasible point, defined as a point that satisfies the constraints in Eq. [7], may be found via linear programming. In our implementation, the ULLS solution of Eq. [17] is used first as a potential feasible point. If this point is feasible, it is also the solution to the CLLS problem. If the point is not feasible, then the algorithm uses the feasible point found via linear programming.

In the second phase, the feasible point is iteratively refined until it converges to the optimal solution. At each iteration, an estimate of the active constraints in Eq. [7] is maintained. Active constraints are a subset of the constraints that are actually enforced, that is, they are the ones that become equality constraints. As the algorithm proceeds, the set of active constraints is updated by sequentially adding the active and removing the inactive constraints from the set.

Heuristic (CLLS-H) Algorithm

Alternatively, the heuristic procedure described below may be used to estimate **D** and **W** such that the constraints of Eqs. [4] and [5] are approximately satisfied. This algorithm assumes that exactly two nonzero b-values are used, and that the gradient directions are the same for both b-values; that is, $\mathbf{N}^{(1)} = \mathbf{N}^{(2)} = \{\mathbf{n}_1,\ldots,\mathbf{n}_N\}$. The central idea in this algorithm is to first estimate $D(\mathbf{n}_i)$ and $K(\mathbf{n}_i)$ along individual gradient directions, and then to threshold the ones that violate the constraints of Eqs. [4] and [5] to within their prescribed limits (steps D1, K1, and K2 below). More specifically, $D(\mathbf{n}_i)$ and $K(\mathbf{n}_i)$ are first determined from $S(\mathbf{n}_i, b_1)$ and $S(\mathbf{n}_i, b_2)$ (Eqs. [18] and [19]). Then, $D(\mathbf{n}_i)$ are thresholded such that they satisfy the constraints on both $D(\mathbf{n}_i)$ and $K(\mathbf{n}_i)$ (steps D1-D4). Next, the constrained $D(\mathbf{n}_i)$ are used to estimate **D** (Eq. [21]), which in turn is used to re-estimate $D(\mathbf{n}_i)$ (Eq. [22]). The re-estimated diffusivities $D^{(R)}(\mathbf{n}_i)$ are then used to obtain the re-estimated $K^{(R)}(\mathbf{n}_i)$ (Eq. [23]). Finally, $K^{(R)}(\mathbf{n}_i)$ are thresholded (steps K1 and K2), and are then used along with $D^{(R)}(\mathbf{n}_i)$ to compute **W**.

Along direction $\mathbf{n}_i$, $D_i = D(\mathbf{n}_i)$ and $K_i = K(\mathbf{n}_i)$ can be estimated using $S(\mathbf{n}_i, b_1)$ and $S(\mathbf{n}_i, b_2)$ as

$$D_i = \frac{b_2 D_i^{(1)} - b_1 D_i^{(2)}}{b_2 - b_1},$$

[18]

and

$$K_i = 6\frac{D_i^{(1)} - D_i^{(2)}}{(b_2 - b_1)D(\mathbf{n}_i)^2},$$ [19]

where

$$D_i^{(1)} = \frac{-\ln(S(\mathbf{n}_i, b_1)/S_0)}{b_1}, \quad D_i^{(2)} = \frac{-\ln(S(\mathbf{n}_i, b_2)/S_0)}{b_2}.$$ [20]

Note that Eqs. [18] and [19] result when exactly two nonzero b-values are used.

Next, a set of rules is applied to constrain D_i:

D1. If $D_i \leq 0$, set $D_i = 0$.

D2. Otherwise, if $D_i^{(1)} < 0$, set $D_i = 0$.

D3. Otherwise, if $D_i > 0$ and $K_i < K_{min}$,

    a. If $K_{min} = 0$, set $D_i = D_i^{(1)}$.

    b. Otherwise, set

$$D_i = \frac{1}{x}[\sqrt{1 + 2xD_i^{(1)}} - 1] \text{ where } x = -\frac{1}{3}K_{min}b_1.$$

D4. Otherwise, if $D_i > 0$ and $K_i > K_{max_i}$, where $K_{max_i} = K_{max}(\mathbf{n}_i)$, set

$$D_i = \frac{D_i^{(1)}}{1 - Cb_1/6b_{max}}$$

where $b_{max} = \max(b_1, b_2)$. Note that the above equation results by plugging the expression for $K_{max_i} = K_{max}(\mathbf{n}_i)$ given by Eq. [6] into Eq. [1], setting $b = b_1$, and solving for D_i. The diffusion signal for $b = b_1$ is utilized here, as it is less sensitive to noise and is a stronger function of D_i.

From the constrained D_i set, **D** may be estimated as

$$\widehat{\mathbf{X}}_D = \mathbf{A}_D^+ \mathbf{b}_D.$$ [21]

where $\hat{\mathbf{X}}_D$ is the ULLS estimate of $\mathbf{X}_D = [D_{11}\ D_{22}\ D_{33}\ D_{12}\ D_{13}\ D_{23}]^T$, and $\mathbf{b}_D = [D_1 \ldots D_N]^T$.

The re-estimated $\mathbf{b}_D^{(R)}$ is then obtained as

$$\mathbf{b}_D^{(R)} = \mathbf{A}_D \widehat{\mathbf{X}}_D.$$ [22]

The re-estimation step has a noise removal effect on $\mathbf{b}_D$ by forcing it to be consistent with $\hat{\mathbf{X}}_D$. Next, $D^{(R)}(\mathbf{n}_i) = D_i^{(R)}$ are used to obtain $K^{(R)}(\mathbf{n}_i) = K_i^{(R)}$ as

$$K_i^{(R)}=\begin{cases} 6\frac{D_i^{(R)}-D_i^{(2)}}{b_2(D_i^{(R)})^2} & \text{if } D_i^{(R)}>0 \\ 0 & \text{otherwise} \end{cases},$$ [23]

Next, a set of rules is applied to constrain $K_i^{(R)}$:

K1. If $K_i^{(R)}>K_{\max_i}^{(R)}$, set $K_i^{(R)}=K_{\max_i}^{(R)}$.

K2. If $K_i^{(R)}<K_{\min}$, set $K_i^{(R)}=K_{\min}$.

Finally, the kurtosis tensor is estimated as

$$\widehat{\mathbf{X}}_K=\frac{1}{\overline{D}^2}\mathbf{A}_K^+\mathbf{b}_K$$ [24]

where $\hat{\mathbf{X}}_K$ is the ULLS estimate of $\mathbf{X}_K = [W_{1111} \ldots W_{1112} \ldots W_{1122} \ldots W_{1233}]^T$, and $\mathbf{b}_K=[(D_1^{(R)})^2 K_1^{(R)} \cdots (D_N^{(R)})^2 K_N^{(R)}]^T$.

Tensor-Derived Kurtosis Measures

The MK is defined as the diffusional kurtosis averaged over all gradient directions. In a reference frame that diagonalizes **D**, it may be calculated as (see Appendix for the derivation)

$$\begin{aligned}\overline{K} =F_1(\lambda_1, \\ \lambda_2, \lambda_3)\tilde{W}_{1111} \\ +F_1(\lambda_2, \lambda_1, \lambda_3)\tilde{W}_{2222} \\ +F_1(\lambda_3, \\ \lambda_2, \lambda_1)\tilde{W}_{3333} \\ + F_2(\lambda_1, \\ \lambda_2, \lambda_3)\tilde{W}_{2233} \\ +F_2(\lambda_2, \lambda_1, \lambda_3)\tilde{W}_{1133} \\ +F_2(\lambda_3, \\ \lambda_2, \lambda_1)\tilde{W}_{1122}\end{aligned}$$ [25]

where $\lambda_1 > \lambda_2 > \lambda_3$ denote the eigenvalues of **D**, and $\tilde{W}_{ijkl}$ are the elements of the kurtosis tensor in the rotated frame and are given by

$$\tilde{W}_{ijkl}=\sum_{i',j',k',l'=1}^{3} R_{ii'}R_{jj'}R_{kk'}R_{ll'}W_{i'j'k'l'},$$ [26]

with R_{ij} being the j-th component of the eigenvector of **D** corresponding to λ_i. Functions F_1 and F_2 are given by

$$F_1(\lambda_1,\lambda_2,\lambda_3)=\frac{(\lambda_1+\lambda_2+\lambda_3)^2}{18(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}\left| \frac{\sqrt{\lambda_2\lambda_3}}{\lambda_1}R_F\left(\frac{\lambda_1}{\lambda_2},\frac{\lambda_1}{\lambda_3},1\right)+\frac{3\lambda_1^2-\lambda_1\lambda_2-\lambda_1\lambda_3-\lambda_2\lambda_3}{3\lambda_1\sqrt{\lambda_2\lambda_3}}R_D\left(\frac{\lambda_1}{\lambda_2},\frac{\lambda_1}{\lambda_3},1\right)\right|$$ [27]

and

$$F_2(\lambda_1, \lambda_2, \lambda_3) = \frac{(\lambda_1+\lambda_2+\lambda_3)^2}{3(\lambda_2-\lambda_3)^2} \left[\frac{\lambda_2+\lambda_3}{\sqrt{\lambda_2\lambda_3}} R_F\left(\frac{\lambda_1}{\lambda_2}, \frac{\lambda_1}{\lambda_3}, 1\right) + \frac{2\lambda_1 - \lambda_2 - \lambda_3}{3\sqrt{\lambda_2\lambda_3}} R_D\left(\frac{\lambda_1}{\lambda_2}, \frac{\lambda_1}{\lambda_3}, 1\right) - 2 \right], \qquad [28]$$

where R_F and R_D are Carlson's elliptic integrals (24) defined as

$$R_F(x, y, z) = \frac{1}{2}\int_0^\infty (t+x)^{-1/2}(t+y)^{-1/2}(t+z)^{-1/2}d \qquad [29]$$

and

$$R_D(x, y, z) = \frac{3}{2}\int_0^\infty (t+x)^{-1/2}(t+y)^{-1/2}(t+z)^{-3/2}d. \qquad [30]$$

As there are highly efficient numerical algorithms for calculating Carlson's elliptic integrals (24–25), Eqs. [27] and [28] may be used for fast computation of the MK. Note that when two or more of the eigenvalues of **D** coincide, F_1 and F_2 have removable singularities whose treatment is discussed in the Appendix.

The axial kurtosis $K_\parallel$ is the diffusional kurtosis in the direction of the highest diffusion. In a reference frame that diagonalizes **D**, it is simply given by

$$K_\parallel = \frac{(\lambda_1+\lambda_2+\lambda_3)^2}{9\lambda_1^2}\tilde{W}_{1111}. \qquad [31]$$

The radial kurtosis $K_\perp$ is defined as the mean diffusional kurtosis perpendicular to the direction of highest diffusion, consistent with the definition of radial diffusivity. In a reference frame that diagonalizes **D**, it is given by (see the Appendix for the derivation)

$$K_\perp = G_1(\lambda_1, \lambda_2, \lambda_3)\tilde{W}_{2222} + G_1(\lambda_1, \lambda_3, \lambda_2)\tilde{W}_{3333} + G_2(\lambda_1, \lambda_2, \lambda_3)\tilde{W}_{2233}. \qquad [32]$$

where

$$G_1(\lambda_1, \lambda_2, \lambda_3) = \frac{(\lambda_1+\lambda_2+\lambda_3)^2}{18\lambda_2(\lambda_2-\lambda_3)^2}\left(2\lambda_2 + \frac{\lambda_3^2 - 3\lambda_2\lambda_3}{\sqrt{\lambda_2\lambda_3}}\right) \qquad [33]$$

and

$$G_2(\lambda_1, \lambda_2, \lambda_3) = \frac{(\lambda_1+\lambda_2+\lambda_3)^2}{3(\lambda_2-\lambda_3)^2}\left(\frac{\lambda_2+\lambda_3}{\sqrt{\lambda_2\lambda_3}} - 2\right). \qquad [34]$$

The treatment of the singularities of $G_1(\lambda_1,\lambda_2,\lambda_3)$ and $G_2(\lambda_1,\lambda_2,\lambda_3)$ is discussed in the Appendix.

Method

Magnetic Resonance Imaging Procedure

DKI scans were performed on three healthy volunteers using a 3 T Siemens Tim Trio system with a 12-channel head coil. This study was approved by the New York University School of Medicine Institutional Review Board and written informed consent was obtained from all subjects. Diffusion-weighted images (DWI's) were acquired along 30 gradient directions with a twice-refocused spin-echo sequence. Scan parameters were TR = 5900 ms, TE = 96

ms, matrix = 82 × 82, FOV = 222 × 222 mm^2, 45 slices, slice thickness = 2.7 mm with no gap, partial Fourier encoding = 3/4, NEX = 44 for $b = 0$ (NEX = 45 for Dataset 2), and NEX = 4 for $b = 500$, 1000, 1500, and 2000 s/mm^2 (NEX = 5 for Dataset 2). The total acquisition time was 53:20 minutes (65:32 minutes for Dataset 2). We emphasize that standard DKI scans (discussed below) require 7:29 minutes.

Processing Pipeline

In the first stage of processing, the DWI's are optionally smoothed using a Gaussian kernel to reduce the impact of noise and misregistration. Next, **D** and **W** are estimated using the CLLS-QP algorithm, and finally, the scalar measures are obtained from the tensors. The pipeline was implemented in-house in the MATLAB environment (http://www.mathworks.com). The program, designated as Diffusional Kurtosis Estimator (DKE), is available upon request from the corresponding author.

The parameters used with DKE were as follows: the full width at half maximum of the Gaussian kernel was 3.375 mm; $K_{min} = 0$; and $C = 3$. The tensors were also estimated using the UNLS, ULLS and CLLS-H algorithms following smoothing of DWI's. Three processing protocols were considered where different subsets of the acquired DWI's were used to estimate the parametric maps. In the *reference* protocol, the maps were obtained using all of the available DWI's. In the *standard* protocol, the maps were estimated using NEX = 11 for $b = 0$ images, and NEX = 1 for $b = 1000$, 2000 s/mm^2 DWI's. Finally, in the *fast* protocol, the maps were obtained with NEX = 11 for $b = 0$ images, and NEX = 1 with 15 of $b = 1000$ s/mm^2 and all of $b = 2000$ s/mm^2 DWI's. We used the maps obtained using the reference protocol and ULLS as baseline for evaluating the maps estimated with the other protocols. The standard protocol is currently used for clinical research acquisitions, and the fast protocol was conjured up as a potential candidate for future clinical research acquisitions. The estimated acquisition time for the fast protocol is 5:54 minutes.

Results

Table 1 shows the processing times for generating the scalar tensor-derived measures using different tensor estimation methods, and summarizes the features of the algorithms. Figure 1, Figure 2, and Figure 3 show the MK, MD, and FA maps obtained using the standard protocol. The regions with the most noticeable change in the estimated MK across all datasets are in the splenium and the genu of the corpus callosum, where the voxels with the incorrectly estimated MK values are in black. The CLLS-H algorithm corrects the MK values for some of these voxels, whereas the CLLS-QP method corrects all of the voxels. Although constrained methods also improve the estimates of the MD and FA, this improvement is visually less discernible. Figure 4 shows the voxels where tensor estimates obtained via ULLS and the standard and reference protocols violate the constraints in Eqs. [4] or [5]. With the standard protocol, very few voxels violate the Eq. [4] constraint in general. Such voxels are typically located in the corpus callosum. Minimum kurtosis constraint violations also typically occur in the corpus callosum. These voxels are less likely to violate the maximum kurtosis constraints, although this is not uncommon. The cerebrospinal fluid (CSF) voxels show a prominent pattern of violating the maximum kurtosis constraints.

Table 2 shows the root mean-square error (RMSE) between the maps obtained using the standard protocol and the maps estimated using the reference protocol. RMSE values are shown for the whole brain as well as for voxels with constraint violations. The whole-brain RMSE values averaged over all subjects indicate that the constrained methods outperform the unconstrained methods in estimating all three scalar quantities. The RMSE improvement offered by the constrained methods for the voxels with constraint violations is more

dramatic, and is more than 35% for the MK, 7% for the MD, and 8% for the FA. Note that a lower threshold of −2, the theoretical minimum kurtosis, was applied to MK values estimated with the unconstrained methods to prevent their RMSE values from becoming excessively large. The difference between the RMSE's of the two unconstrained methods is fairly small, and so is the difference between the constrained methods.

Figure 5, Figure 6, and Figure 7 show the MK, MD, and FA maps estimated using the fast protocol, while Figure 8 shows the voxels for which the constraints are violated. In this case, the voxels with constraint violations are distributed more widely in the brain parenchyma, although their most consistent location is the corpus callosum. Table 3 shows the RMSE values for the maps obtained using the fast protocol. As with the standard protocol, the RMSE values of the unconstrained methods are comparable. However, with the fast protocol, the whole-brain RMSE gain due to the CLLS-QP method is larger than with the standard protocol. For the voxels with constraint violations, the RMSE improvement due to the constrained methods exceeds 40% for the MK, 10% for the MD, and 19% for the FA.

Figure 9 shows the directional diffusivities and kurtoses for two Dataset 1 voxels with constraint violations. In the corpus callosum voxel, the UNLS solution violates 5 out of 90 constraints, and the ULLS solution violates 11 constraints. The CLLS-H solution reduces the total number of violations from 11 to 5, and the estimate based on CLLS-QP satisfies all constraints as expected. In the prefrontal white matter voxel, the unconstrained algorithms both violate one constraint, whereas both constrained algorithms satisfy all constraints.

Finally, Figure 10 shows the maps estimated using the new expressions for the MK and $K_\perp$, along with the map for $K_\parallel$.

Discussion

The voxels in Figure 1 and Figure 5 where CLLS-H and CLLS-QP produce improved MK estimates are those for which ULLS yields an NPD or a *nearly* NPD estimate of **D** and/or the constraints on the minimum and maximum kurtoses are violated. In these voxels, which most consistently occur in highly anisotropic regions such as the corpus callosum, λ_1 is large and λ_3 is small. This situation increases the sensitivity of unconstrained methods to noise. One effect of noise is to create the appearance of non-decreasing diffusion signal intensity with increasing b. This typically occurs along gradient directions with high diffusivity, but can also happen along directions with very low diffusivity. At high b-values, a small disturbance due to noise can artificially increase the signal intensity with increasing b-value, yielding falsely large estimates of directional kurtoses. A prominent instance of this effect can be observed in the CSF voxels in Figure 4 and Figure 8.

The other effect of noise is to create the appearance of the diffusion signal being a concave function of b along low-diffusivity directions, yielding negative estimates of directional kurtoses. In this case, the measured diffusion-weighted signal intensity may even exceed the signal intensity at $b = 0$, in which case a negative estimate of diffusivity, that is, an NPD DT, arises.

It is well-known that imaging noise causes λ_1 to be overestimated and λ_3 to be underestimated (26). The constraints in Eqs. [4] and [5] force a smaller λ_1 and larger λ_3, yielding more reasonable estimates of **D** and **W**. The top right plot of Figure 9(a) shows an instance of this, where the CLLS-QP estimates of directional diffusivities along directions with high diffusivity are slightly lower than those obtained with ULLS, whereas the opposite is the case along low-diffusivity directions.

Table 2 and Table 3 provide further confirmation of the improvement offered by the constrained algorithms. Note that the whole-brain RMSE improvement is actually due to only 10–20% of voxels for which there are constraint violations, hence the RMSE improvement within the set of voxels with constraint violations is quite larger than the improvement over the whole brain. The constrained algorithms achieve comparable RMSE, and both outperform the unconstrained algorithms consistently across all subjects, measures and protocols. However, the RMSE improvement is more evident with the fast protocol. Note also that the average whole-brain RMSE in estimating the MK and MD with the fast protocol and the constrained methods is smaller than or is comparable to the whole-brain RMSE with the standard protocol and unconstrained methods. This means that the constrained methods can achieve a savings of more than 20% in acquisition time. (Note that a total of 56 $b = 0$ images and DWI's were utilized in the fast protocol, whereas in the standard protocol 71 images were used.) The RMSE improvement offered by the constrained methods is less pronounced in estimating the FA. A possible explanation for this may be that the FA is more heavily influenced by the Rician noise corrupting the DWI's than the MD, consistent with previous reports in the literature (27). Therefore, to obtain more reliable FA estimates, constrained methods may need to be combined with schemes to compensate for the Rician noise bias.

The CLLS-H algorithm strikes a balance between complexity and accuracy by correcting the MK in most of the voxels that violate the constraints, doing so at a negligible 2.5% overhead compared to the ULLS (cf. Table 1). Thus, the CLLS-H method can be beneficial in real-time situations where run time is crucial. On the other hand, the CLLS-QP approach removes all such voxels, but it has a much longer run time. It should be noted that while the improvement offered by CLLS-QP comes at the expense of increased algorithm complexity and consequently higher run time, it is straightforward to speed up the computations via parallelization.

A limitation of the formulation in Eq. [7] is that it ensures that constraints in Eqs. [4] and [5] are satisfied along all *imaged* gradient directions, rather than all *possible* directions. Imposing the latter constraints is more challenging and requires a deeper understanding of the kurtosis tensor. A practically viable alternative is to estimate the tensors using either of the constrained algorithms and then check if the solution satisfies the constraints along other directions. The directions of the eigenvectors of **D** are of the most interest as Eq. [25] and Eq. [32] are explicit functions of the kurtosis tensor elements along these directions. If the solution does not satisfy the constraints along these directions, then these constraints can be added to the set of constraints in Eqs. [4] and [5] and the problem re-solved. This procedure may be repeated until the solution satisfies the constraints.

Finally, as the expressions for $K_\perp$ and MK used for estimating the maps in Figure 10 represent the exact values of these parameters, they more accurately quantify the diffusional kurtosis characteristics of the tissue. The advantage of $K_\perp$ as defined here and in (8,20) over that of (15) is that ours gives all radial directions equal weight, rather than selecting out two particular radial directions.

Conclusions

We presented two related advancements to the DKI estimation framework. The first advancement significantly improves the fidelity of the tensors parameterizing the DKI model by casting the estimation problem as CLLS. The CLLS formulation ensures that the estimated diffusivities and kurtoses along the imaged gradient directions remain within a physically and biologically plausible range. Two algorithms were provided for minimizing the CLLS cost function: a convex quadratic programming (CLLS-QP) method and a

heuristic (CLLS-H) algorithm. The CLLS-H algorithm imposes a negligible computational overhead compared to the ULLS algorithm, making it suitable for real-time applications with two *b*-values and equal numbers of gradients per *b*-value, whereas the computationally more demanding CLLS-QP algorithm permits an arbitrary number of *b*-values and different gradient sets per each *b*-value. *In vivo* image data demonstrated the significant improvement offered by the proposed constrained tensor estimation methods over unconstrained algorithms. The advantage of the constrained methods is particularly evident when fewer DWI's are available. The RMSE improvement due to constrained methods may be translated into shorter scan times and/or increased spatial resolution for a given fidelity of the parametric maps; our results indicate a potential time savings of more than 20%. The second advancement allows for more accurate and efficient calculation of MK and $K_\perp$ by utilizing the exact closed-form formulae derived in this paper. The proposed tensor estimation methods along with the exact expressions for the kurtosis measures improve the robustness of the DKI metrics, thus facilitating the clinical application of DKI.

Acknowledgments

Appendix

The MK may be calculated from the formula

$$\overline{K} = \sum_{i,j,k,l=1}^{3} A_{ijkl} W_{ijkl},$$ [A1]

where

$$A_{ijkl} \equiv \frac{\overline{D}^2}{4\pi} \int_0^\pi d\theta \int_0^{2\pi} d\phi \frac{n_i n_j n_k n_l}{\left(\sum_{p,q=1}^{3} n_p n_q D_{pq}\right)^2} \sin\theta$$ [A2]

and n_i is the *i*-th element of unit direction vector **n**. MK may be obtained by numerically evaluating the integral in Eq. [A2]. However, this can be computationally inefficient if a high degree of accuracy is desired. Below we give an analytic formula for A_{ijkl} that enables fast computation of the MK.

In a reference frame that diagonalizes **D** (see Eq. [26]), we can write

$$\overline{K} = \sum_{i,j,k,l=1}^{3} A_{ijkl} \tilde{W}_{ijkl},$$ [A3]

where A_{ijkl} are given by

$$A_{ijkl}(\lambda_1, \lambda_2, \lambda_3) = \frac{\overline{D}^2}{4\pi} \int_0^\pi d\theta \int_0^{2\pi} d\phi \frac{n_i(\theta,\phi) n_j(\theta,\phi) n_k(\theta,\phi) n_l(\theta,\phi) \sin\theta}{\left(\lambda_1 \sin^2\theta \cos^2\phi + \lambda_2 \sin^2\theta \sin^2\phi + \lambda_3 \cos^2\theta\right)^2},$$ [A4]

and

$$n_1(\theta,\phi)=\sin\theta\cos\phi,$$
$$n_2(\theta,\phi)=\sin\theta\sin\phi,$$
$$n_3(\theta,\phi)=\cos\theta.$$

[A5]

It can be seen from Eq. [A4] that A_{ijkl} is invariant with respect to permutations of its indices. Thus, we need consider only the 15 components A_{1111}, A_{1112}, A_{1113}, A_{1122}, A_{1123}, A_{1133}, A_{1222}, A_{1223}, A_{1233}, A_{1333}, A_{2222}, A_{2223}, A_{2233}, A_{2333}, and A_{3333}. From symmetry of the integrand in Eq. [A4], one may show

$$A_{1112}=A_{1113}=A_{1123}=A_{1222}=A_{1223}=A_{1233}=A_{1333}=A_{2223}=A_{2333}=0.$$

[A6]

This leaves only 6 nonzero components to consider. We can further reduce this number by noting the symmetries

$$A_{1111}(\lambda_1,\lambda_2,\lambda_3)=A_{2222}(\lambda_2,\lambda_1,\lambda_3)=A_{3333}(\lambda_3,\lambda_2,\lambda_1)$$

[A7]

and

$$A_{2233}(\lambda_1,\lambda_2,\lambda_3)=A_{1133}(\lambda_2,\lambda_1,\lambda_3)=A_{1122}(\lambda_3,\lambda_2,\lambda_1).$$

[A8]

Hence we only need to find analytic expressions for A_{1111} and A_{2233}. From these, and by applying conventional integration techniques and also utilizing $\bar{D}=(\lambda_1+\lambda_2+\lambda_3)/3$, the MK can be calculated as prescribed by Eq. [25], where $F_1(\lambda_1,\lambda_2,\lambda_3)=A_{1111}(\lambda_1,\lambda_2,\lambda_3)$ and $F_2(\lambda_1,\lambda_2,\lambda_3)=6A_{2233}(\lambda_1,\lambda_2,\lambda_3)$.

When $\lambda_1=\lambda_2$ or $\lambda_1=\lambda_3$, Eq. [27] has a removable singularity, which can be resolved with

$$A_{1111}(\lambda_1,\lambda_1,\lambda_3)=3A_{2233}(\lambda_3,\lambda_1,\lambda_1) \quad\text{or}\quad A_{1111}(\lambda_1,\lambda_2,\lambda_1)=3A_{2233}(\lambda_2,\lambda_1,\lambda_1).$$

[A9]

When $\lambda_2=\lambda_3$, Eq. [28] has a removable singularity. This can be resolved with

$$A_{2233}(\lambda_1,\lambda_3,\lambda_3)=\frac{(\lambda_1+2\lambda_3)^2}{144\lambda_3^2(\lambda_1-\lambda_3)^2}\left[\lambda_3(\lambda_1+2\lambda_3)+\lambda_1(\lambda_1-4\lambda_3)\alpha\left(1-\frac{\lambda_1}{\lambda_3}\right)\right],$$

[A10]

where

$$\alpha(x)\equiv\begin{cases}\frac{1}{\sqrt{x}}\operatorname{arctanh}(\sqrt{x}), & \text{if } x>0,\\ \frac{1}{\sqrt{-x}}\arctan(\sqrt{-x}), & \text{if } x<0.\end{cases}$$

[A11]

Finally, for the special case when $\lambda_1=\lambda_2=\lambda_3$, we simply have

$$A_{1111}=\frac{1}{5} \quad\text{and}\quad A_{2233}=\frac{1}{15}.$$

[A12]

The radial kurtosis $K_\perp$ is the mean diffusional kurtosis perpendicular to the direction of the highest diffusion. In a reference frame that diagonalizes $\mathbf{D}$, it is given by

$$K_\perp=\sum_{i,j,k,l=1}^{3}C_{ijkl}\tilde{W}_{ijkl},$$

[A13]

where

$$C_{ijkl}(\lambda_1, \lambda_2, \lambda_3) = \frac{\overline{D}^2}{2\pi} \int_0^{2\pi} d\phi \frac{n_i(\phi) n_j(\phi) n_k(\phi) n_l(\phi)}{(\lambda_2 \cos^2\phi + \lambda_3 \sin^2\phi)^2},$$ [A14]

and

$$\begin{aligned} n_1(\phi) &= 0, \\ n_2(\phi) &= \cos\phi. \\ n_3(\phi) &= \sin\phi. \end{aligned}$$ [A15]

It can be seen from Eq. [A14] that C_{ijkl} is invariant with respect to permutations of its indices. Thus, we need consider only the 15 components C_{1111}, C_{1112}, C_{1113}, C_{1122}, C_{1123}, C_{1133}, C_{1222}, C_{1223}, C_{1233}, C_{1333}, C_{2222}, C_{2223}, C_{2233}, C_{2333}, and C_{3333}. From symmetry of the integrand in Eq. [A14] and the fact that $n_1(\phi) = 0$, one may show

$$C_{1111} = C_{1112} = C_{1113} = C_{1122} = C_{1123} = C_{1133} = C_{1222} = C_{1223} = C_{1233} = C_{1333} = C_{2223} = C_{2333} = 0.$$ [A16]

This leaves only 3 nonzero components to consider. We can further reduce this number by noting the symmetry

$$C_{2222}(\lambda_1, \lambda_2, \lambda_3) = C_{3333}(\lambda_1, \lambda_3, \lambda_2).$$ [A17]

Hence we only need to find analytic expressions for C_{2222} and C_{2233}. From these, $K_\perp$ can be obtained as prescribed by Eq. [32], where the integrals

$$C_{2222}(\lambda_1, \lambda_2, \lambda_3) = \frac{\overline{D}^2}{2\pi} \int_0^{2\pi} d\phi \frac{\cos^4\phi}{(\lambda_2 \cos^2\phi + \lambda_3 \sin^2\phi)^2}$$ [A18]

and

$$C_{2233}(\lambda_1, \lambda_2, \lambda_3) = \frac{\overline{D}^2}{2\pi} \int_0^{2\pi} d\phi \frac{\cos^2\phi \sin^2\phi}{(\lambda_2 \cos^2\phi + \lambda_3 \sin^2\phi)^2}$$ [A19]

are elementary forms which lead to Eqs. [33] and [34], and $G_1(\lambda_1,\lambda_2,\lambda_3) = C_{2222}(\lambda_1,\lambda_2,\lambda_3)$ and $G_2(\lambda_1,\lambda_2,\lambda_3) = 6C_{2233}(\lambda_1,\lambda_2,\lambda_3)$.

For the special case of $\lambda_2 = \lambda_3$, Eq. [33] and Eq. [34] reduce to

$$C_{2222}(\lambda_1, \lambda_2, \lambda_2) = \frac{(\lambda_1 + 2\lambda_2)^2}{24\lambda_2^2}$$ [A20]

and

$$C_{2233}(\lambda_1, \lambda_2, \lambda_2) = \frac{(\lambda_1 + 2\lambda_2)^2}{72\lambda_2^2}.$$ [A21]

References

1. Basser PJ, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. Biophysical Journal. 1994; 66:259–267. [PubMed: 8130344]

2. Basser PJ. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. NMR in Biomedicine. 1995; 8:333–344. [PubMed: 8739270]

3. Basser PJ, Pierpaoli C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. Journal of Magnetic Resonance B. 1996; 111:209–219.

4. Wedeen VJ, Hagmann P, Tseng W-YI, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magnetic Resonance in Medicine. 2005; 54(6):1377–1386. [PubMed: 16247738]

5. Jensen, JH.; Helpern, JA. Quantifying non-Gaussian water diffusion by means of pulsed-field-gradient MRI; Toronto, Canada. Proceedings of the International Society for Magnetic Resonance in Medicine Annual Meeting; 2003. p. 2154

6. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magnetic Resonance in Medicine. 2005; 53:1432–1440. [PubMed: 15906300]

7. Lu H, Jensen JH, Ramani A, Helpern JA. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. NMR in Biomedicine. 2006; 19:236–247. [PubMed: 16521095]

8. Jensen JH, Helpern JA. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR in Biomedicine. 2010 Published online: May 19, 2010. DOI:10.1002/nbm.1518.

9. Falangola MF, Jensen JH, Babb JS, Hu C, Castellanos FX, Di Martino A, Ferris SH, Helpern JA. Age-related non-Gaussian diffusion patterns in the prefrontal brain. Journal of Magnetic Resonance Imaging. 2008; 28:1345–1350. [PubMed: 19025941]

10. Raab P, Hattingen E, Franz K, Zanella FE, Lanfermann H. Cerebral gliomas: Diffusional kurtosis imaging analysis of microstructural differences. Radiology. 2010; 254:876–881. [PubMed: 20089718]

11. Jansen JFA, Stambuk HE, Koutcher JA, Shukla-Dave A. Non-Gaussian analysis of diffusion-weighted MR imaging in head and neck squamous cell carcinoma: A feasibility study. AJNR Am J Neuroradiol. 2009 ajnr.A1919.

12. Cheung MM, Hui ES, Chan KC, Helpern JA, Qi L, Wu EX. Does diffusion kurtosis imaging lead to better neural tissue characterization? A rodent brain maturation study. NeuroImage. 2009; 45:386–392. [PubMed: 19150655]

13. Lazar M, Jensen JH, Xuan L, Helpern JA. Estimation of the orientation distribution function from diffusional kurtosis imaging. Magnetic Resonance in Medicine. 2008; 60:774–781. [PubMed: 18816827]

14. Song S-K, Sun S-W, Ramsbottom MJ, Chang C, Russell J, Cross AH. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. NeuroImage. 2002; 17(3):1429–1436. [PubMed: 12414282]

15. Hui ES, Cheung MM, Qi L, Wu EX. Towards better MR characterization of neural tissues using directional diffusion kurtosis analysis. NeuroImage. 2008; 42:122–134. [PubMed: 18524628]

16. Koay CG, Carew JD, Alexander AL, Basser PJ, Meyerand ME. Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging. Magnetic Resonance in Medicine. 2006; 55:930–936. [PubMed: 16526013]

17. Liu C, Bammer R, Acar B, Moseley ME. Characterizing non-Gaussian diffusion by using generalized diffusion tensors. Magnetic Resonance in Medicine. 2004; 51:924–937. [PubMed: 15122674]

18. Wang Z, Vemuri BC, Chen Y, Mareci TH. A Constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. IEEE Transactions on Medical Imaging. 2004; 23:930–939. [PubMed: 15338727]

19. Barmpoutis A, Hwang MS, Howland D, Forder JR, Vemuri BC. Regularized positive-definite fourth order tensor field estimation from DW-MRI. NeuroImage. 2009; 45:S153–S162. [PubMed: 19063978]

20. Poot DHJ, den Dekker AJ, Achten E, Verhoye M, Sijbers J. Optimal Experimental Design for Diffusion Kurtosis Imaging. IEEE Transactions on Medical Imaging. 2010; 29:819–829. [PubMed: 20199917]

21. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. Journal of Chemical Physics. 1965; 42:288–292.

22. 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.

23. Nocedal, J.; Wright, SJ. Numerical Optimization. New York: Springer; 1999.

24. Carlson BC. Computing elliptic integrals by duplication. Numerische Mathematik. 1979; 33:1–16.

25. Carlson BC, Notis EM. Algorithm 577: Algorithms for incomplete elliptic integrals [S21]. ACM Transactions on Mathematical Software. 1981; 7(3):398–403.

26. Anderson AW. Theoretical analysis of the effects of noise on diffusion tensor imaging. Magnetic Resonance in Medicine. 2001; 46(6):1174–1188. [PubMed: 11746585]

27. Basu, S.; Fletcher, T.; Whitaker, R. Lecture Notes in Computer Science. Vol. Volume 4190. Berlin: Springer; 2006. Rician noise removal in diffusion tensor MRI; p. 117-125.
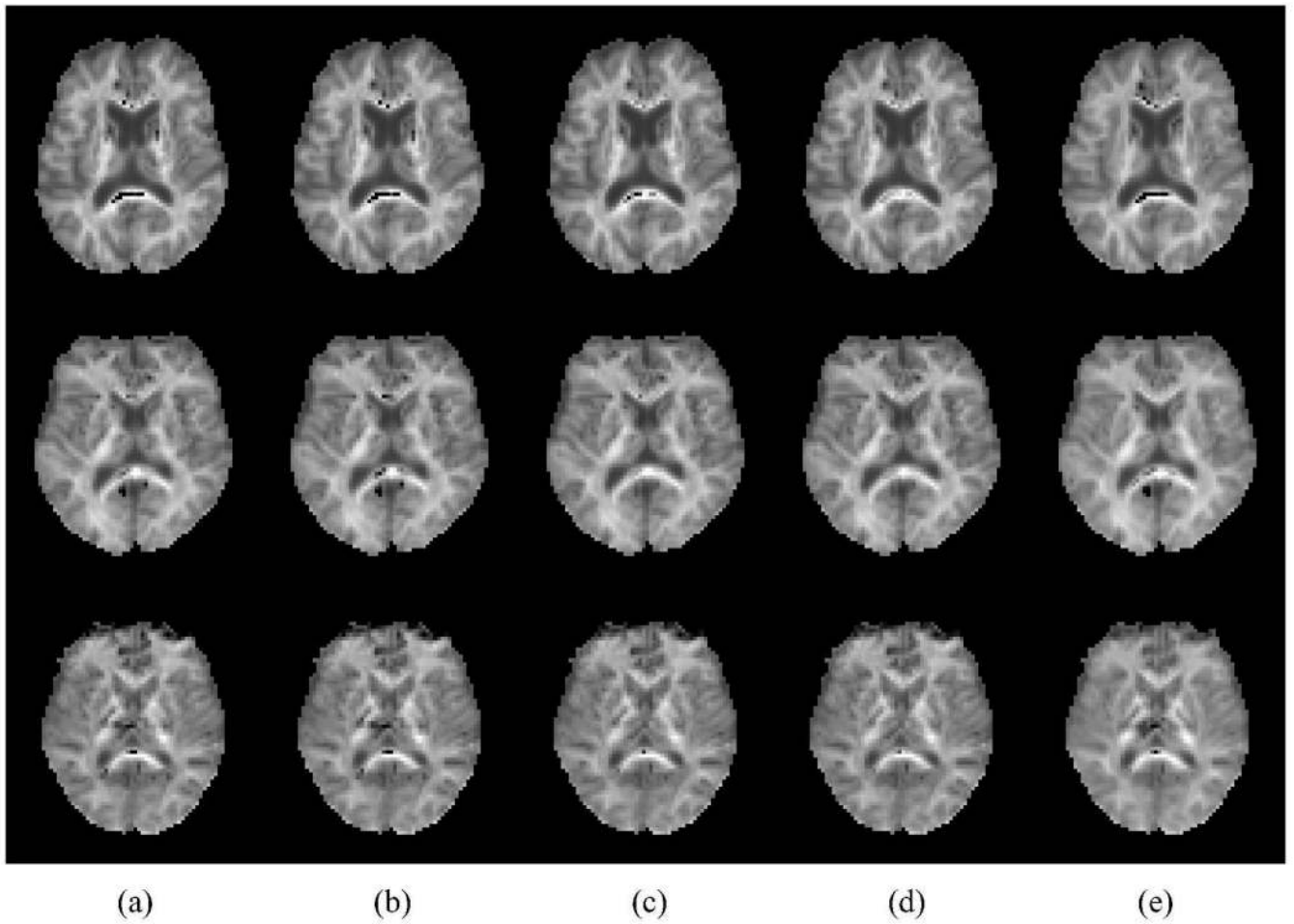
Figure 1.
MK maps obtained using the standard protocol and (a) UNLS; (b) ULLS; (c) CLLS-H; (d) CLLS-QP; and (e) the reference protocol and ULLS.
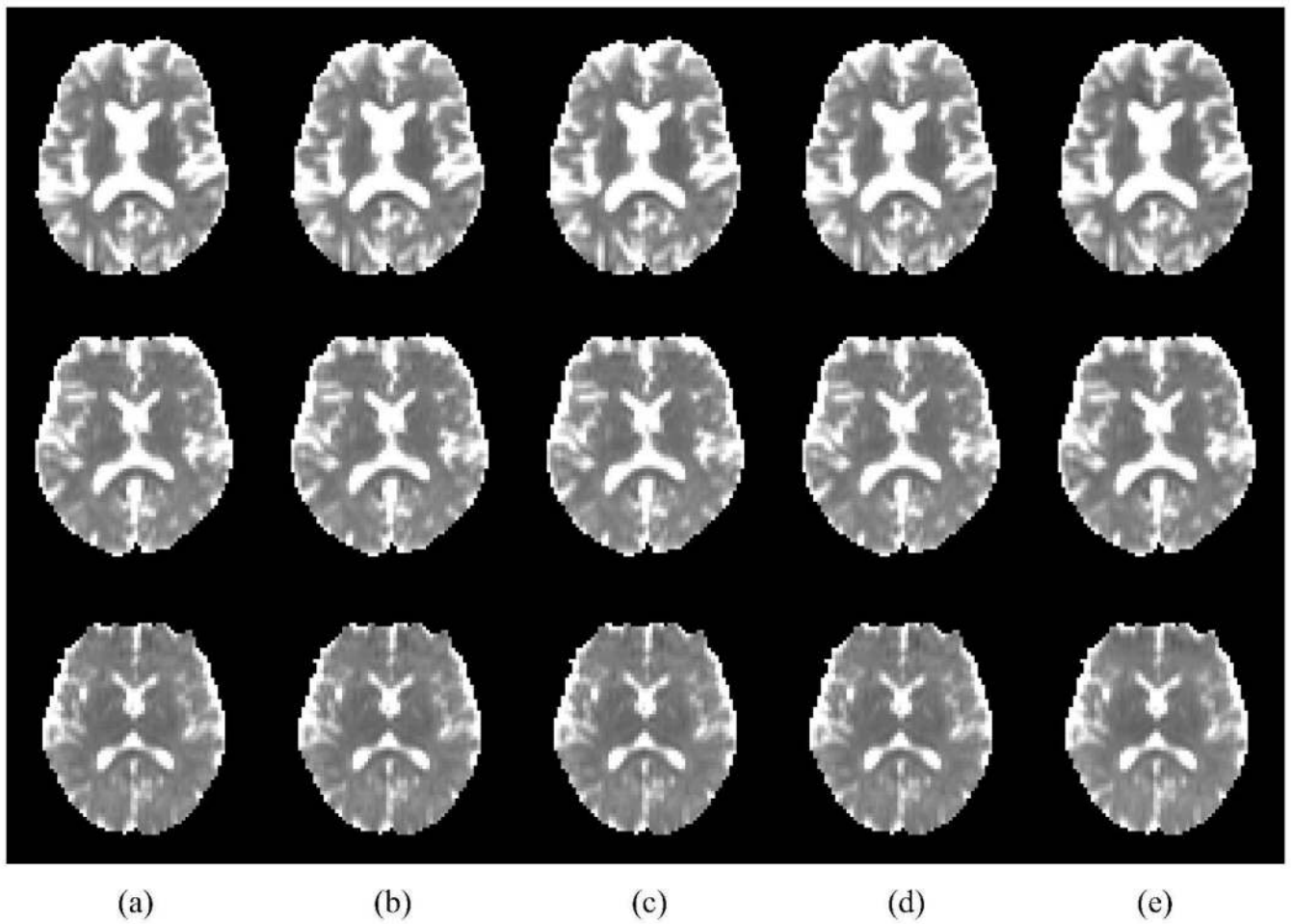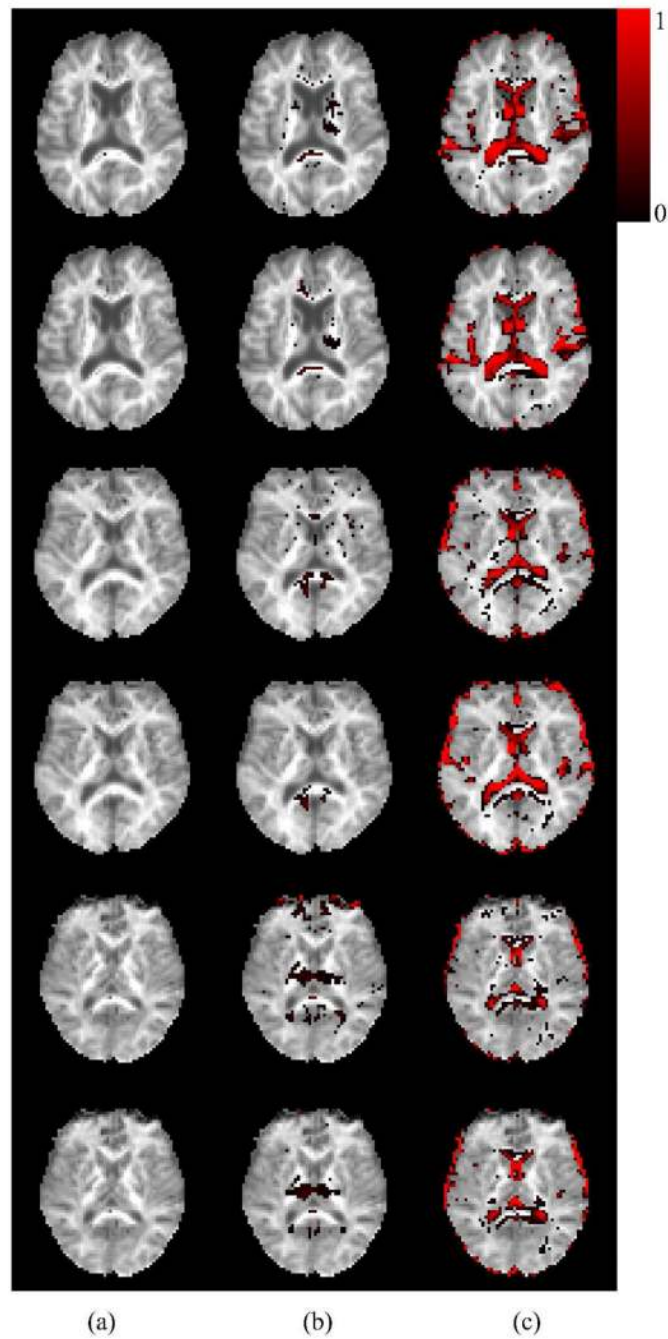
Figure 2.
MD maps obtained using the standard protocol and (a) UNLS; (b) ULLS; (c) CLLS-H; (d) CLLS-QP; and (e) the reference protocol and ULLS.
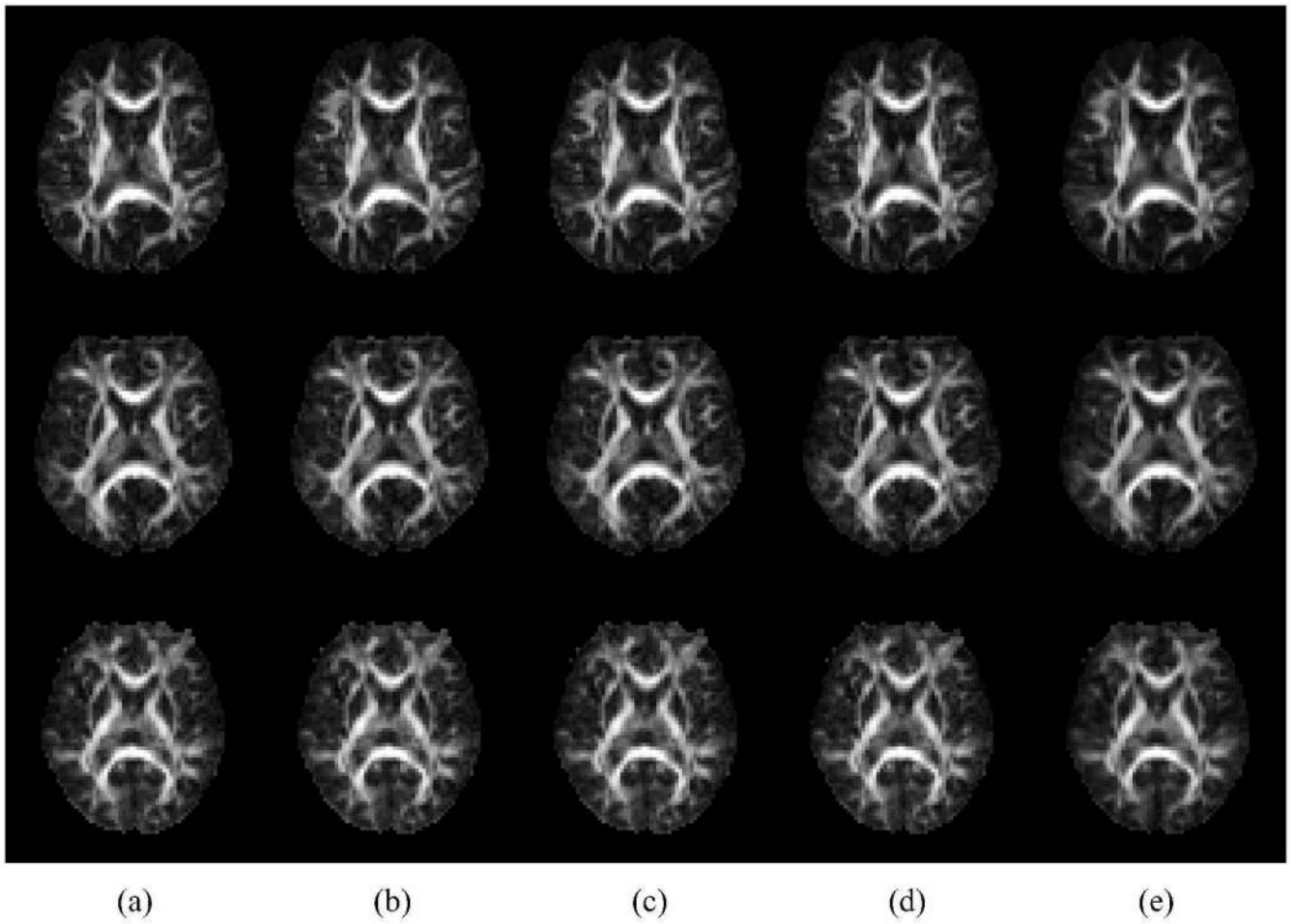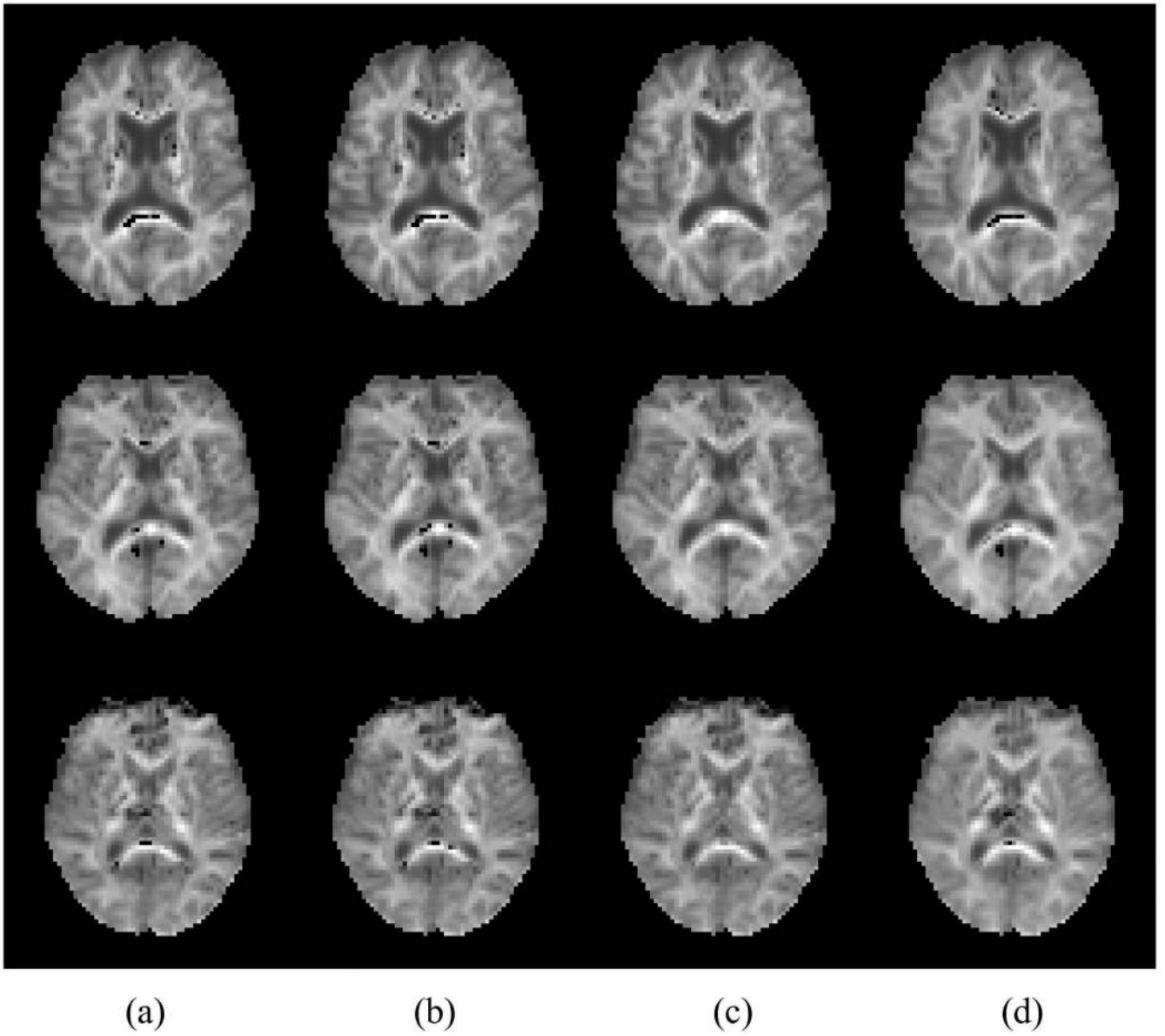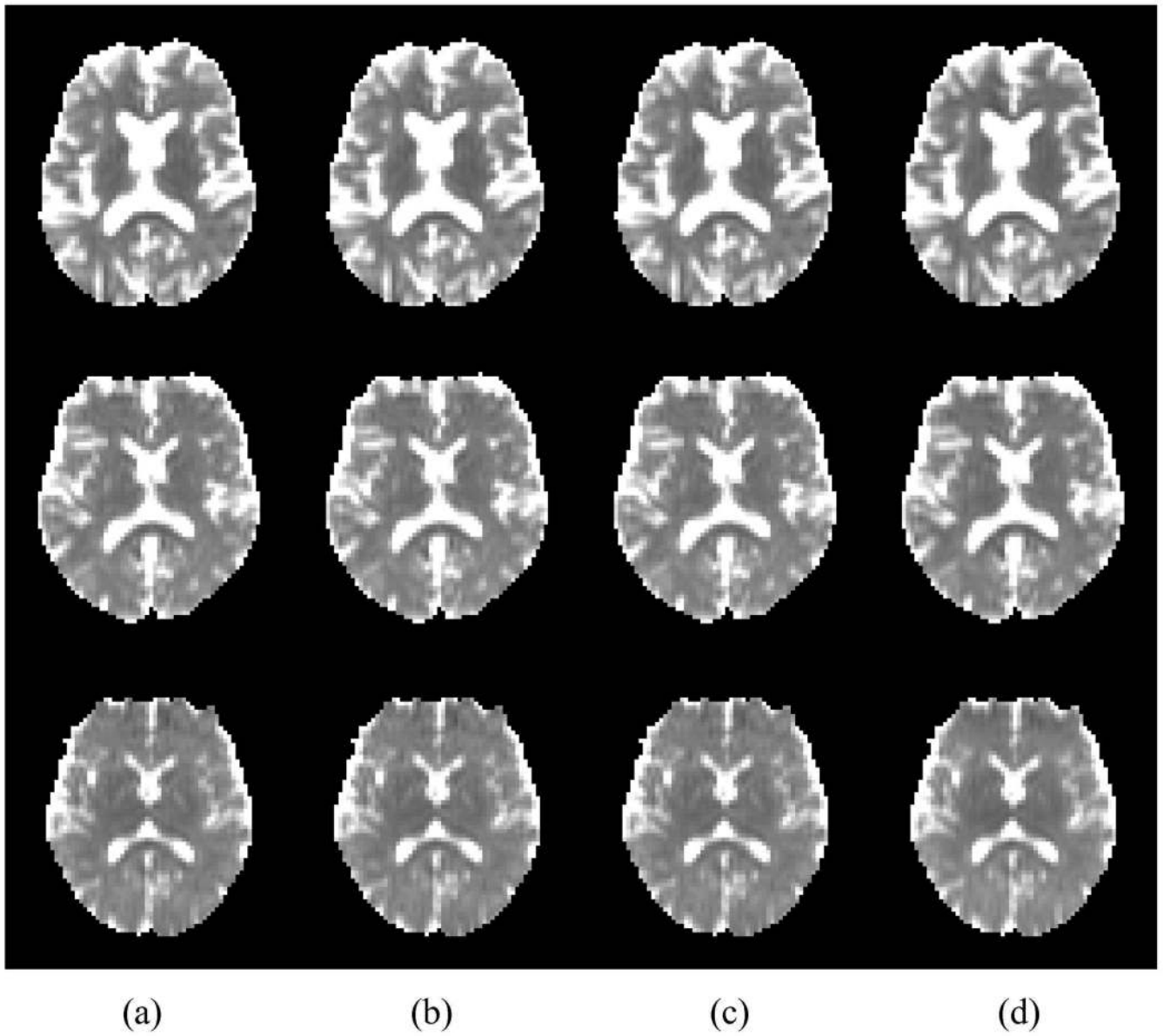
Figure 3.
FA maps obtained using the standard protocol and (a) UNLS; (b) ULLS; (c) CLLS-H; (d) CLLS-QP; and (e) the reference protocol and ULLS.
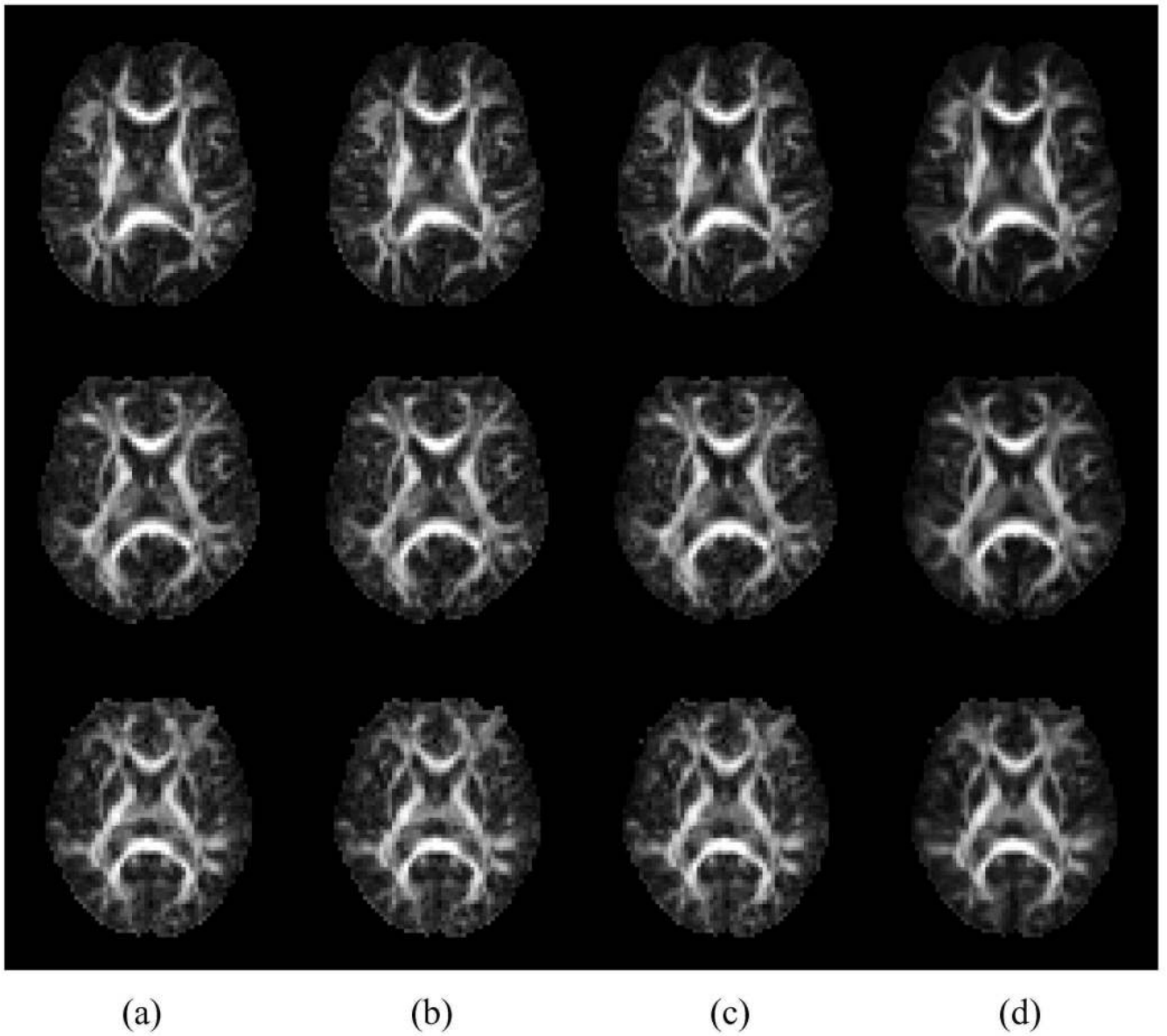
Figure 4.
Voxels (indicated in red) for which the diffusion and kurtosis tensors estimated using standard (rows 1, 3, and 5) and reference (rows 2, 4, and 6) protocols and ULLS violate the constraints on (a) directional diffusivities; (b) minimum directional kurtoses; and (c) maximum directional kurtoses. The shade of red for each voxel indicates the fraction of constraints that are violated.

Figure 5.
MK maps obtained using the fast protocol and (a) UNLS; (b) ULLS; (c) CLLS-QP; and (d) the reference protocol and ULLS. Note that due to the different numbers of gradient directions, the CLLS-H algorithm is not applicable.

(a) (b) (c) (d)

Figure 6.
MD maps obtained using the fast protocol and (a) UNLS; (b) ULLS; (c) CLLS-QP; and (d) the reference protocol and ULLS.

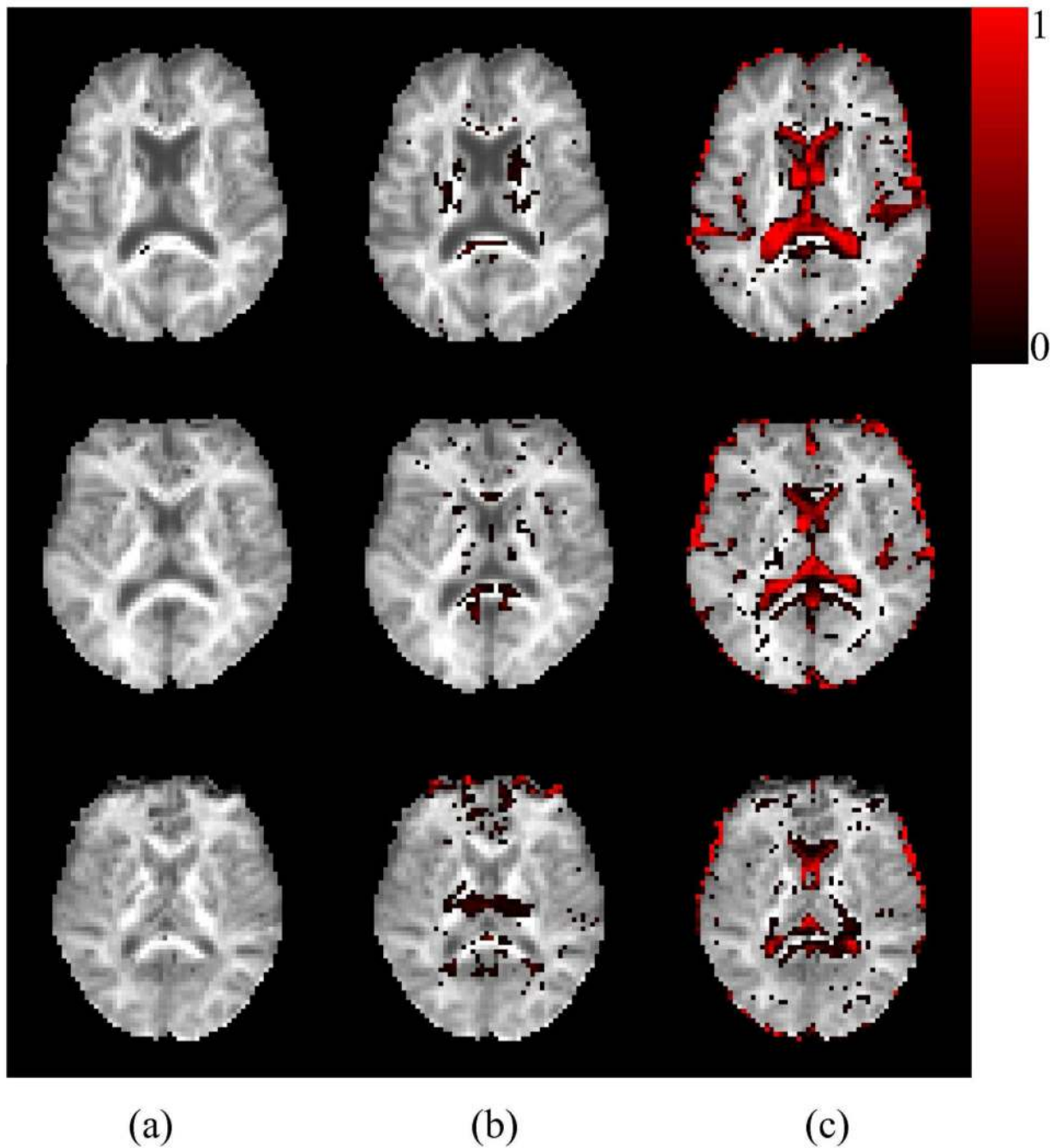(a)   (b)   (c)   (d)

Figure 7.
FA maps obtained using the fast protocol and (a) UNLS; (b) ULLS; (c) CLLS-QP; and (d) the reference protocol and ULLS.
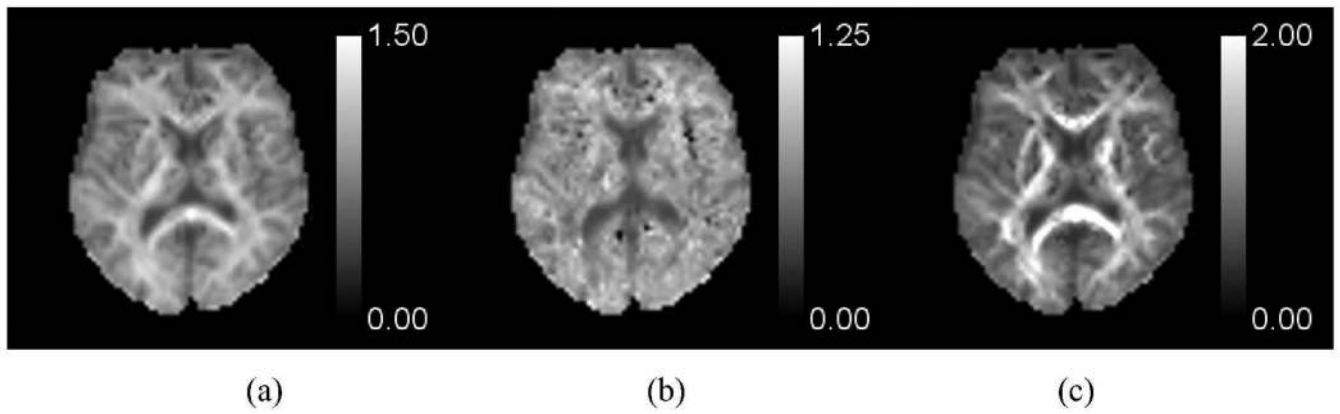
Figure 8.
Voxels (indicated in red) for which the diffusion and kurtosis tensors estimated using the fast protocol and ULLS violate the constraints on (a) directional diffusivities; (b) minimum directional kurtoses; and (c) maximum directional kurtoses. The shade of red for each voxel indicates the fraction of constraints that are violated.
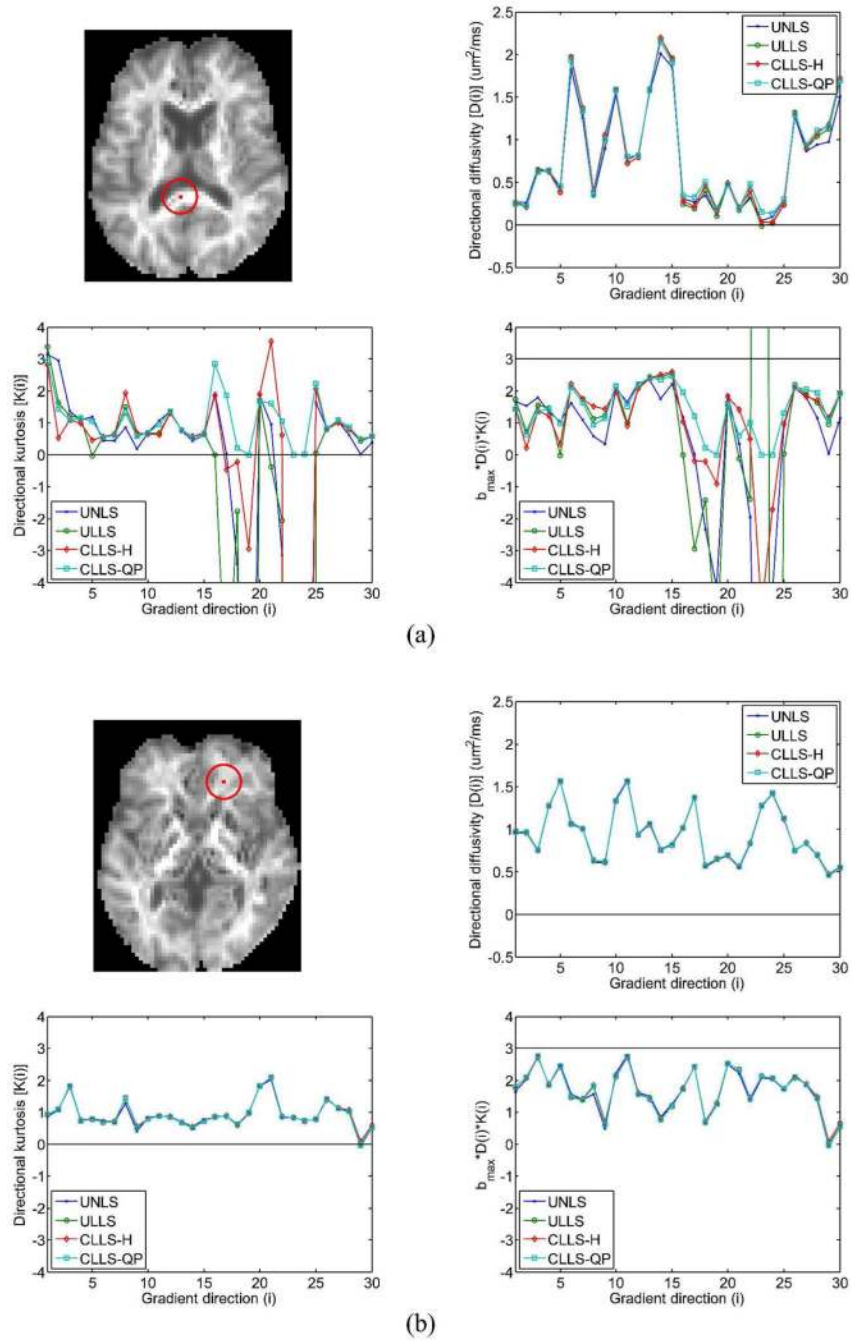
Figure 9.
Voxels in the (a) corpus callosum and (b) prefrontal white matter from Dataset 1 and their corresponding directional diffusivities (top right), kurtoses (bottom left), and the product $b_{max}D(\mathbf{n})K(\mathbf{n})$ (bottom right), estimated using different unconstrained and constrained methods. Parameters controlling the minimum and maximum kurtoses were set to $K_{min} = 0$ and $C = 3$. Note that the directional kurtoses must satisfy $b_{max}D(\mathbf{n})K(\mathbf{n}) < C$. Note that the lines connecting the data points are intended to aid with visualizing the points that fall outside of the plotted range, and do not imply any ordering among the gradient directions.

Figure 10.
Kurtosis maps for Dataset 2 estimated using the standard protocol, utilizing the CLLS-QP algorithm and the proposed exact expressions; (a) MK; (b) $K_\parallel$; and (c) $K_\perp$. Note that the maps are displayed with different scales.

Table 1

Feature comparison between the UNLS, ULLS, CLLS-H, and CLLS-QP algorithms. Run times are for all slices in Dataset 2 obtained with the standard protocol and include the input/output time.

	UNLS	ULLS	CLLS-H	CLLS-QP
Run time (seconds)[a]	594	158	162	531
Number of *b*-values	Any	Any	2 nonzero	Any
Number of gradient directions per *b*-value	Any	Any	Equal	Any

[a]Obtained on a Dell Precision workstation equipped with a 2.8 GHz Pentium Core 2 Duo processor.

Table 2

RMSE between the parametric maps estimated using the standard protocol and different estimation methods, and the maps obtained using the reference protocol and ULLS; (a) MK; (b) MD; and (c) FA. In each cell, two RMSE values are reported: the first RMSE value was obtained for the whole brain, whereas the second value was calculated for voxels for which the unconstrained methods violated at least one of the constraints on directional diffusivities or kurtoses. The whole-brain MK and MD RMSE values were computed for voxels in the brain parenchyma (MD < 2) for which the reference protocol and ULLS satisfied all the constraints on directional diffusivities and kurtoses. The FA RMSE values were obtained for the subset of the above-mentioned voxels that corresponded to the white matter (FA > 0.25). The RMSE values for MK and FA are unitless, and for MD they are in units of $\mu m^2/ms$.

(a)

Algorithm	UNLS	ULLS	CLLS-H	CLLS-QP
Dataset 1	0.101 / 0.239	0.102 / 0.239	0.092 / 0.137	0.093 / 0.137
Dataset 2	0.055 / 0.108	0.056 / 0.115	0.053 / 0.082	0.054 / 0.082
Dataset 3	0.086 / 0.198	0.086 / 0.196	0.078 / 0.128	0.079 / 0.132
Average	0.081 / 0.182	0.082 / 0.183	0.074 / 0.116	0.075 / 0.117

(b)

Algorithm	UNLS	ULLS	CLLS-H	CLLS-QP
Dataset 1	0.082 / 0.101	0.082 / 0.101	0.082 / 0.095	0.081 / 0.092
Dataset 2	0.055 / 0.082	0.055 / 0.083	0.055 / 0.078	0.055 / 0.077
Dataset 3	0.080 / 0.112	0.080 / 0.110	0.078 / 0.100	0.078 / 0.098
Average	0.073 / 0.098	0.072 / 0.098	0.072 / 0.091	0.072 / 0.089

(c)

Algorithm	UNLS	ULLS	CLLS-H	CLLS-QP
Dataset 1	0.045 / 0.056	0.045 / 0.058	0.043 / 0.050	0.043 / 0.046
Dataset 2	0.037 / 0.048	0.037 / 0.052	0.036 / 0.046	0.036 / 0.043
Dataset 3	0.045 / 0.064	0.045 / 0.067	0.044 / 0.058	0.044 / 0.053
Average	0.042 / 0.056	0.042 / 0.059	0.041 / 0.051	0.041 / 0.047

Table 3

RMSE between the parametric maps estimated using the fast protocol and different estimation methods, and the maps obtained using the reference protocol and ULLS; (a) MK; (b) MD; and (c) FA. In each cell, two RMSE values are reported: the first RMSE value was obtained for the whole brain, whereas the second value was calculated for voxels for which the unconstrained methods violated at least one of the constraints on directional diffusivities or kurtoses. The whole-brain MK and MD RMSE values were computed for voxels in the brain parenchyma (MD < 2) for which the reference protocol and ULLS satisfied all the constraints on directional diffusivities and kurtoses. The FA RMSE values were obtained for the subset of the above-mentioned voxels that corresponded to the white matter (FA > 0.25). The RMSE values for MK and FA are unitless, and for MD they are in units of $\mu m^2/ms$.

(a)			
Algorithm	UNLS	ULLS	CLLS-QP
Dataset 1	0.109 / 0.256	0.111 / 0.254	0.096 / 0.135
Dataset 2	0.060 / 0.114	0.062 / 0.122	0.058 / 0.084
Dataset 3	0.098 / 0.211	0.099 / 0.221	0.085 / 0.127
Average	0.089 / 0.194	0.091 / 0.199	0.080 / 0.115

(b)			
Algorithm	UNLS	ULLS	CLLS-QP
Dataset 1	0.084 / 0.096	0.084 / 0.096	0.083 / 0.086
Dataset 2	0.058 / 0.080	0.058 / 0.081	0.057 / 0.073
Dataset 3	0.082 / 0.105	0.082 / 0.104	0.080 / 0.092
Average	0.075 / 0.093	0.075 / 0.093	0.073 / 0.084

(c)			
Algorithm	UNLS	ULLS	CLLS-QP
Dataset 1	0.056 / 0.075	0.056 / 0.079	0.053 / 0.060
Dataset 2	0.048 / 0.068	0.048 / 0.073	0.047 / 0.060
Dataset 3	0.057 / 0.091	0.058 / 0.094	0.054 / 0.069
Average	0.053 / 0.078	0.054 / 0.082	0.051 / 0.063