Fast and Robust Estimation of Diffusional Kurtosis Imaging (DKI) Parameters by General Closed-form Expressions and their Extensions

Yoshitaka Masutani* and Shigeki Aoki

1Department of Radiology, The University of Tokyo Hospital
7–3–1 Hongo, Bunkyo-ku, Tokyo 113–8655, Japan
2Department of Radiology, Juntendo Hospital

(Received August 19, 2013; Accepted December 11, 2013; published online April 28, 2014)

Diffusional kurtosis imaging (DKI) for clinical imaging involves time-consuming computation and demonstrates low robustness. Standard estimation of DKI parameters is based on an extension of Stejskal-Tanner’s signal model with squared b-value term and is a least-squares fitting problem. The use of numerical methods for computation requires time, and estimation of DKI parameters is noise sensitive and often produces noisy results, such as images with pepper noise.

In this study, we propose general closed-form solutions for DKI parameters to avoid numerical computation for least-squares fitting, solutions that can be applied to diffusion weighted imaging (DWI) datasets with any number of b-values more than three. Solutions are obtained through stationary-point conditions of an objective function that are minimized for fitting. We use 3 techniques to extend the solutions to increase robustness—b-value-dependent weighting in fitting, removal of outliers, and addition of neighbor sampling. Based on synthetic datasets and clinical datasets that both consist of 6 b-value and 3 b-value datasets, we detail and compare the 3 methods including a method by Jensen et al.

Keywords: closed-form solution, diffusional kurtosis imaging, least-squares fitting, MRI, synthetic data

Introduction

Diffusion-weighted imaging (DWI) including diffusion tensor imaging (DTI) assumes the Gaussian diffusion of water molecules, whereas diffusional kurtosis imaging (DKI) quantifies their non-Gaussianity, which is caused by various bi-structural constraints and applied clinically to characterize normal and abnormal tissues. Unlike DTI and its matured metrics, such as fractional anisotropy (FA), DKI and its derived parameters are in development. However, several reports of clinical applications reveal important features of DKI, such as its potential usefulness in diagnostic imaging to evaluate neurological issues in aging and disorders including stroke, Alzheimer’s disease, schizophrenia, glioma, Parkinson disease, and attention deficit hyperactivity disorder. Recent trials on DKI in such regions as the liver and spine have also been reported. DKI may also benefit conventional DTI by aiding more accurate esti-
mation of DTI parameters.\textsuperscript{14}

In DKI, non-Gaussianity is represented as a new parameter, kurtosis, in addition to the diffusion coefficient. The kurtosis parameter is used in basically 2 ways—in isotropic analysis of the mean kurtosis value among several directions of the motion-probing gradient (MPG) and in anisotropic analysis in which several kurtosis values obtained in different MPG directions in a manner similar to that of DTI make up the kurtosis tensor.\textsuperscript{15} Thus, DKI is expected to provide a new horizon of diagnostic imaging based on DWI by introducing new information on isotropic and anisotropic diffusion properties. Originally, the value of kurtosis is defined to characterize the form of the probability density function (PDF) of water molecules after diffusion and is therefore obtained most directly using the PDF estimated in q-space imaging (QSI),\textsuperscript{3} which assumes no models for the water diffusion process. However, in DKI, the kurtosis value can be estimated based on a simple signal model without PDF estimation as follows.

The DKI signal model can be regarded as an extension of the Stejskal-Tanner model\textsuperscript{16} for DWI, denoted as:

\[ S(b) \equiv S_0 \cdot \exp(-bD), \]  

in which \( S \) is a signal value of DWI data depending on \( b \), \( S_0 \) is a baseline signal value for \( b = 0 \), and \( b \) is the diffusion weighting factor, called the “b-value” and consisting of such physical parameters as strength of the MPG. In addition to the parameter of the diffusion coefficient, \( D \), we also consider the parameter of kurtosis, \( K \), for describing the non-Gaussianity. That is, the DKI signal model can be described by adding the squared b-value term with \( D \) and \( K \) to the Stejskal-Tanner model:\textsuperscript{3}

\[ S(b) \equiv S_0 \cdot \exp(-bD + \frac{1}{6} b^2 D^2 K). \]  

In both signal models, the parameters are estimated with a series of measured signal values, \( S \), and b-values, \( b \), used in the acquisition of DWI data. In addition to those 3 parameters of \( D \) and \( K \), \( S_0 \) is also obtained as a model parameter, although it is often approximated by \( S(0) \), which is the measured signal value with \( b = 0 \).

Several methods can be considered to compute DKI parameters based on Formula [2]. Because this signal model is not a linear combination of the DKI parameters, computation of DKI is a non-linear least-squares fitting (LSF) problem conventionally solved by numerical methods with iterative improvement of solutions.\textsuperscript{3} One current issue in clinical DKI is the cost of computation because it is a time-consuming process that requires numerical methods. Jensen and associates\textsuperscript{17,18} have proposed an alternative approach that employs simple closed-form expressions for a special case of only 3 b-values; for \( b_i \) (\( i = \{1, 2, 3\} \)):

\[ D \equiv \frac{(b_3 + b_1)D^{(12)} - (b_2 + b_1)D^{(13)}}{b_3 - b_2}, \]

and

\[ K \equiv 6 \frac{D^{(12)} - D^{(13)}}{(b_3 - b_2)D^2}. \]  

where \( D^{(ij)} \) is denoted as:

\[ D^{(ij)} \equiv \log S(b_i) - \log S(b_j). \]  

The \( S_0 \) value is also determined by substituting \( D, K, \) and \( S(b_i) \) into Formula [2]. This computation can be performed very fast without iterative computation. However, this solution is noise-sensitive because estimation of 3 parameters with 3 measurements mathematically becomes an even-determined problem with 3 measurements of signal values for 3 unknown parameters. Therefore, this method requires acquisition of images with less noise by multiple collections of signal values. Naturally, another apparent drawback of this approach is that it cannot be used directly for datasets with more than 3 b-values unless 3 pairs of b-values and signal values are selected. Considering that more samples generally yield a more reliable parameter estimation with least-squares fitting, we need other general purpose approaches without limitation in number of b-values.

Even with more than 3 b-values, DKI computing is often sensitive to noise in clinical data and suffers from low robustness of parameter estimation. DKI images obtained from clinical data often include salt-and-pepper or just pepper noise in kurtosis images in which DKI parameter computation has been unreliable (see examples in the Results section). Such poor parameter estimation in each MPG direction reduces the reliability of the analysis results of kurtosis anisotropy.

We undertook this study to address the 2 major problems in DKI computation—its high computational cost and low robustness. In this study, we extend our preliminary report,\textsuperscript{19} describing more mathematical details and further developments. We introduce our general closed-form expressions of DKI parameters for the non-linear LSF problem in the previous report\textsuperscript{19} and then propose new robust computation methods for DKI that extend the general closed-form solutions. Based on the notion that reliable parameter estimation in each MPG
direction improves analysis results of kurtosis anisotropy, we focus on robust estimation of DKI parameters in a single MPG direction. Finally, we compare the methods with synthetic data and clinical data based on several experiments.

Methods and Materials

The generalized problem formulation for DKI parameter estimation in a single MPG direction using Formula [2] follows. When we use a series of b-values, \( b_\alpha (\alpha = 1 \ldots N) \), in multiple DWI acquisitions, we obtain corresponding signal values, \( S_\alpha = S(b_\alpha) \), at each voxel. Then, we estimate the DKI parameters \( D, K, \) and \( S_0 \) at the voxel location so that the signal model of Formula [2] describes their relationships well. This is a typical curve-fitting problem with those pairs of b-values and signal values (Fig. 1). Including the even-determined problem solved by Formulas [3], we have to obtain \( N \) pairs of b-values and signal values \((b_\alpha, S_\alpha)\) with at least 3 different b-values \((N \geq 3)\).

1. Non-linear LSF formulation and general closed-form (GCF) expressions

First, we describe the most basic solution presented in our previous report.\(^{19}\) When an over-determined problem is solved by LSF, an objective function that represents the fitting residual is generally employed to make an optimization problem. Generally, numerical methods that include the LM algorithm iteratively improve solutions to minimize objective functions by using the gradient and Hessian of \(J\). In our earlier study,\(^{19}\) we showed that the least-squares solutions can also be obtained as closed-form expressions for any number of b-values more than three. The idea is based on use of the 3 conditions on partial derivatives of \(J\) at stationary points below:

\[
\frac{\partial J}{\partial D} = \frac{\partial J}{\partial K} = \frac{\partial J}{\partial S_0} = 0,
\]

where

\[
\frac{\partial J}{\partial D} = \frac{D^3K^2\Lambda_{40} - 9D^2K\Lambda_{30} + 6DK\log S_0\Lambda_{20} - 6DK\Lambda_{21} + 18D\Lambda_{20} - 18\log S_0\Lambda_{10} + 18\Lambda_{10}}{9},
\]

\[
\frac{\partial J}{\partial K} = \frac{\Lambda_{40} - 6D^3\Lambda_{30} - 6D^2\Lambda_{21} + 6D^2\log S_0\Lambda_{20}}{18},
\]

and

\[
\frac{\partial J}{\partial S_0} = \frac{6\log S_0 + D^2K\Lambda_{20} - 6D\Lambda_{10} - 6\Lambda_{01}}{3S_0}.
\]

Here, \( \Lambda_{ij} (i = \{0, 1, 2, 3, 4\}, j = \{0, 1, 2\}) \) consists of b-values and signal values as below:

\[
\Lambda_{ij} \equiv \frac{1}{N} \left\{ \sum_{\alpha=1}^{N} b_\alpha^i \cdot (\log S_\alpha)^j \right\}.
\]
Then, we obtain only a unique stationary point, and the estimated parameters $\hat{D}$, $\hat{K}$, and $\hat{S}_0$ of the LSF solutions in closed-form expressions are:

$$\hat{D} = \frac{Q_1}{P_1}, \quad \hat{K} = 6\frac{P_1Q_2}{Q_1^2}, \quad \text{and} \quad \hat{S}_0 = \exp\left(\frac{Q_1}{P_1}\right).$$

The values $P_1$, $Q_1$, $Q_2$, and $Q_3$ consisting of $A_{ij}$ are defined as:

$$P_1 \equiv -\Lambda_{10}^2\Lambda_{40} + 2\Lambda_{10}\Lambda_{20}\Lambda_{30} - \Lambda_{20}^3 + \Lambda_{20}\Lambda_{40} - \Lambda_{30}^2,$$

$$Q_1 \equiv (\Lambda_{10}\Lambda_{40} - \Lambda_{20}\Lambda_{30})\Lambda_{01} + (\Lambda_{20}^2 - \Lambda_{40})\Lambda_{11} - (\Lambda_{10}\Lambda_{20} - \Lambda_{30})\Lambda_{21},$$

$$Q_2 \equiv (\Lambda_{20}^2 - \Lambda_{10}\Lambda_{30})\Lambda_{01} - (\Lambda_{10}\Lambda_{20} - \Lambda_{30})\Lambda_{11} + (\Lambda_{10}^2 - \Lambda_{20})\Lambda_{21},$$

and

$$Q_3 \equiv (\Lambda_{20}\Lambda_{40} - \Lambda_{30}^2)\Lambda_{01} - (\Lambda_{10}\Lambda_{40} - \Lambda_{20}\Lambda_{30})\Lambda_{11} + (\Lambda_{10}\Lambda_{30} - \Lambda_{20}^2)\Lambda_{21}.$$ 

Note that here we define 2 different types of values $P_i$ and $Q_i$, so that the former shows signal-independent values consisting of $A_{ij}$ with $j = 0$, and the latter is all signal dependent by including $A_{ij}$ with $j \neq 0$.

As well as DKI parameters, we can obtain the value of the objective function, $J$, at the unique stationary point, $J_{SP}$, by substituting $\hat{D}$, $\hat{K}$, and $\hat{S}_0$ into $J$ as:

$$J_{SP} = \frac{Q_1^2}{P_1P_2} - \frac{Q_4}{P_2}.$$ 

where additional definitions for $P_2$ and $Q_4$ are:

$$P_2 \equiv \Lambda_{20}^2 - \Lambda_{40}$$

and

$$Q_4 \equiv -\Lambda_{40}\Lambda_{01}^2 + 2\Lambda_{20}\Lambda_{01}\Lambda_{21} - \Lambda_{21}^2 - (\Lambda_{20}^2 - \Lambda_{40})\Lambda_{02}.$$ 

The value $J_{SP}$ is important for evaluating the quality of fitting because it represents the MSE as mentioned earlier and is utilized in robust methods described later.

Using those closed-form expressions, we can avoid the process of numerical optimization in estimating DKI parameters. More strictly, in a mathematical sense, however, the closed-form expressions of DKI parameters here are never guaranteed for minimization of $J$. Generally, stationary points can be local minima/maxima, global minimum/maximum, saddle point. The $J_{SP}$ value here can be regarded as the global minimum of $J$ only if the Hessian of $J$ shown below is positive-definite at the stationary point, which is:

$$H = \begin{pmatrix}
\frac{\partial^2 J}{\partial D^2} & \frac{\partial^2 J}{\partial D \partial K} & \frac{\partial^2 J}{\partial D \partial S_0} \\
\frac{\partial^2 J}{\partial K \partial D} & \frac{\partial^2 J}{\partial K^2} & \frac{\partial^2 J}{\partial K \partial S_0} \\
\frac{\partial^2 J}{\partial S_0 \partial D} & \frac{\partial^2 J}{\partial S_0 \partial K} & \frac{\partial^2 J}{\partial S_0^2}
\end{pmatrix}. \quad [19]$$

Because the elements of the Hessian derived from $J$ consist of measured signal values contaminated with noise, it is practically impossible to guarantee the positive definiteness of the Hessian. Conversely, positive definiteness is a criterion that reflects whether the solution is reliable based on the minimization of $J$. Because the Hessian of $J$ is real and symmetric, the positive definiteness of $H$ can be judged efficiently using Sylvester’s criterion.$^{22}$ The conditions are that the leading principal minors (LPM) of $H$, $H_k (k = 1, 2, 3)$, are all positive; these are the determinants of the partial matrix at the upper left $k$-by-$k$ corner of $H$ as follows:

$$H_1 \equiv \frac{\partial^2 J}{\partial D^2}, \quad H_2 \equiv \begin{vmatrix}
\frac{\partial^2 J}{\partial D^2} & \frac{\partial^2 J}{\partial D \partial K} \\
\frac{\partial^2 J}{\partial K \partial D} & \frac{\partial^2 J}{\partial K^2}
\end{vmatrix},$$

and

$$H_3 \equiv \begin{vmatrix}
\frac{\partial^2 J}{\partial D^2} & \frac{\partial^2 J}{\partial D \partial K} & \frac{\partial^2 J}{\partial D \partial S_0} \\
\frac{\partial^2 J}{\partial K \partial D} & \frac{\partial^2 J}{\partial K^2} & \frac{\partial^2 J}{\partial K \partial S_0} \\
\frac{\partial^2 J}{\partial S_0 \partial D} & \frac{\partial^2 J}{\partial S_0 \partial K} & \frac{\partial^2 J}{\partial S_0^2}
\end{vmatrix}. \quad [20]$$

Those values at the stationary point are also expressed in closed forms using Formulas [5] and [11]. That is, the positive-definiteness test (PDT) is performed by checking $H_1 > 0$, $H_2 > 0$, and $H_3 > 0$ to confirm that the fitting problem is for the residual minimization.

2. Modified GCF for robust DKI parameter estimation

In this section, we show 3 different ideas for extending the GCF to increase robustness in estimating DKI parameters—weighted GCF, outlier removal in LSF, and addition of neighbor sampling.

**Weighted GCF:** The significance of the fitting residual may depend on the $b$-value of the sample because higher $b$-values yield low DWI signals and, consequently, a low signal-to-noise ratio (SNR). Therefore, it is natural to weight the elements of residuals in the LSF objective functions, such as
General Closed-form Expressions for DKI Parameters

\[ J_w \equiv \sum_{a=1}^{N} W_a \left\{ \frac{1}{6} b_a^2 D^2 K - b_a D - \log \left( \frac{S_a}{S_0} \right) \right\}^2, \]  \hfill [21]

where \( W_a \) is the weighting factor for each sample pair of \((b_a, S_a)\), so that lower values of \( W_a \) are for higher \( b \)-values. Then, we obtain closed-form expressions for \( \hat{D}, \hat{K}, \) and \( \hat{S}_0 \) using a formula identical to Formulas [11] but replacing \( \Lambda_{ij} \) with their weighted versions:

\[ \Lambda^w_{ij} \equiv \sum_{a=1}^{N} W_a \cdot b_a^i \cdot (\log S_a)^j. \]  \hfill [22]

This idea of weighting is already employed in a method for estimating DTI parameters that uses a signal-dependent weighting factor. We used \( b \)-value-dependent weights to avoid overweighting low signals with higher level noise. That is, higher signal values are weighted with higher ratios. Through preliminary experiments, however, we confirmed better fitting results using \( b \)-value-dependent than signal-dependent weighting. It is obvious that signal-dependent weighting yields higher weights for signals with higher noise levels. Consequently, the use of signal-dependent weighting leads to undesirable overweighting of signals contaminated with high level noise. Therefore, we weighted signals of lower \( b \)-values with a higher ratio in estimating DKI parameters. Details are described later.

Outlier removal in LSF: Generally, for LSF problems in which a certain amount of measured samples are available, removal of outliers among samples often yields better fitting. Based on this notion, several algorithms, such as the random sample consensus (RANSAC) algorithm, have been employed mainly in computer vision problems. In this algorithm, removed samples are selected randomly and iteratively in searching for better fitting of the smaller fitting residual. Moreover, when we obtain a DKI dataset with several \( b \)-values, it is possible to remove some outlier sample pairs \((b_a, S_a)\) to find better fitting. However, when only 3 samples remain, it becomes an even-determined problem so that fitting evaluation by residual is no longer possible.

Our algorithm for computing DKI parameters with outlier removal is based on trials of fitting with all the possible combinations of samples to find the minimum MSE of residuals at the stationary point. When \( n \) samples are removable out of \( N \) (but \( n < N - 3 \)), the MSEs of residuals are evaluated \( \sum_{k=0}^{n} (N) \) times, including the case of no sample removal. For example, when \( N = 6 \) (and therefore \( n = 2 \)), \( J_{SP} \) in Equation [16] must be calculated 22 times. In addition, the PDT is used to check estimates and minimize the fitting residual.

One potential problem in applying this method based on minimizing the residual is over-fitting. An extreme example of over-fitting is the even-determined case with only 3 \( b \)-values remaining. That is, removing too many samples may yield an undesired solution. However, over-fitting cannot be judged by the residual. We attempted 2 practical solutions for this problem. One was to limit the number of removed samples, which we determined before parameter estimation. In the Experiment section, we examine robustness depending on the number of removed samples. The other solution is to validate the value ranges of estimated parameters; that is, we used estimated parameters only if the values were within practical ranges. We describe how to determine the range in the next section.

Addition of neighbor sampling: The algorithms described so far are applied at each voxel location. If we also use signal values from neighboring voxels where the spatial changes of signals are small, we expect better estimation based on the increase in samples. That is, we expect a sort of regularization effect using this technique. Another advantage of this algorithm is that we can apply the GCF for the datasets with only 3 \( b \)-values because samples are virtually increased by neighboring voxels. For example, if we use 3 \( b \)-values and consider 4 in-plane neighbors, the sample number reaches 15. This approach can be simply realized just by adding neighboring pairs of \( b \)-values and signals to the sample series. We used 2 patterns of in-plane additional samples, as shown in Fig. 2. The details are described later.

3. Noise model and practical value ranges of DKI parameters

In the experiments shown in the next section, we used both synthetic datasets with several amounts of noise and clinical datasets. Because DWI signals are formed as magnitudes of complex valued signals, noise distribution is not simply additive but appears as Rician. A simple DKI signal model with such a noise term for measured signal value
$S_m(b)$ is expressed as\(^3\):

$$S_m(b) \equiv \sqrt{\left\{ S_0 \cdot \exp\left(-bD + \frac{1}{6}b^2D^2K\right)\right\}^2 + \eta^2}. \quad [23]$$

Therefore, one can replace the estimated signal value without noise as:

$$S(b) \equiv \sqrt{S_m(b)^2 - \eta^2}. \quad [24]$$

In the synthetic data studies, we used several noise amounts to evaluate the robustness of each method.

In the real datasets, the noise amount, $\eta$, can be estimated from the mean value of signals at air region in the image.\(^3\) Unlike synthetic data in which we can determine the “true” value of signals, it is difficult to evaluate the robustness of a DKI parameter computation in clinical data. However, we can consider DKI parameter estimation to have failed when one of the estimated values is outside the reasonable range, such as in the case of negative $D$ values. We determined ranges for DKI parameters for clinical datasets and used them as a criterion for the parameter estimation error, a process termed parameter range check (PRC). We then counted the parameter estimation error, a process termed parameter range check (PRC). We then counted for clinical datasets and used them as a criterion for the parameter estimation error, a process termed parameter range check (PRC). We then counted for clinical datasets and used them as a criterion for the parameter estimation error, a process termed parameter range check (PRC). We then counted.

$\hat{\theta} \equiv \min_{\theta} \sum_{i=1}^{N_{\text{samples}}} \left( \frac{S_i - S_i(\theta)}{\eta_i} \right)^2$. \quad [11]$

In the real datasets, the noise amount, $\eta$, can be estimated from the mean value of signals at air region in the image.\(^3\)

General closed form (GCF): The GCF solution of Formulas [11] is simply used as a baseline method.

Weighted GCF (WGCF): For weighting factors, we used the following formulation consisting of 2 control parameters, $\lambda$ and $\gamma$:

$$W_\alpha = \frac{1}{1 + (\lambda \cdot b_\alpha)^\gamma}. \quad [25]$$

We used $\lambda = 1000^{-1}$ mm$^2$/sec and $\gamma = 0.5$, which we obtained through our preliminary experiment to determine the optimal combination of control parameters.

Single outlier removal (SOR): Among $N$ samples, only a single sample can be removed so that the parameter estimation with the minimum residual is selected.

Multiple outlier removal (MOR): Among $N$ samples, $n (N = n)$ samples are removed while keeping the problem over-determined. As with the SOR method, the estimation with the minimum residual is selected.

Neighbor sampling (NS): This method is identical to the GCF method except that samples are increased 5 times by using signal values at the neighboring pixels, which are shown in Fig. 2(a).

Blocked neighbor sampling (BNS): BNS is similar to the NS method but considers DKI parameter estimation at the boundaries of different diffusion characteristics. Figure 2(b) shows such a configuration of 4 blocks containing a center voxel; among the 4 blocks of sample combinations, parameters of the block with the minimum residual are selected.

Jensen's closed forms (JCF): In processing 3 b-value datasets, we also used the closed-form solutions of Formulas [3] by Jensen and colleagues\(^17,18\) for comparison.

These methods can be categorized into 2 classes according to the number of estimation candidates of the method. In one group, GCF, WGCF, NS, and JCF, the estimation is uniquely determined; in the other, which includes such methods as SOR, MOR, and BNS, there may be multiple estimation candidates. For the latter group, the PDT and PRC are applied to eliminate candidates to select solutions. That is, we choose the estimation of the minimum residual from within the estimations that satisfy the conditions of both PDT and PRC. However, we must consider the case in which the tests have failed. There can be 4 types of estimations by success or failure of the PDT and PRC. Therefore, we determined the priority for searching the minimum residual as follows. First, we search within the estimations that clear both tests and then among the estimations that clear only the PRC.
Next, we search within the estimations that clear
only the PDT and then
finally among the estima-
tions that have failed both tests.

5. Materials

Synthetic datasets: By using Equation [23], 2
types of synthetic datasets are prepared to simulate
acquisition of DKI of white matter brain tissue.
One simulates imaging with 6 b-values—0, 500, 1000, 1500, 2000, and 2500 s/mm²—and the other,
with three—0, 1000, and 2000 s/mm². As Fig. 3(a)
shows, each axial slice of the datasets consists of
regions with 4 variations of \( S_0 \) values (0, 240, 320, and 400), 3 variations of \( D \) values (0.56 \( \times \) \( 10^{-3} \) mm²/s, 0.64 \( \times \) \( 10^{-3} \) mm²/s, and 0.72 \( \times \) \( 10^{-3} \) mm²/s), and 5 variations of \( K \) values (0, 0.4, 0.8, 1.2, and 1.6). The datasets have a matrix size of
128 \( \times \) 128 \( \times \) 128 and identical patterns of axial
slices except for noise. The datasets assume isotrop-
ic DKI parameters in 6 different MPG directions.
In addition, each type has 5 variations of noise
amount, \( \eta \), in Equation [23], which is approxi-
mated by 0-centered Gaussian distribution with
standard deviation, \( \sigma = 10, 20, 30, 40, \) and 50.
All signal values are quantized to integers to sim-
ulate real data from clinical scanners. The variation
of the \( S_0 \) value for fixed noise amount simulates the
relative change of not only SNR but also the quan-
tization effect of signal values.

Figures 3(b) and 4 show examples of data with
\( \sigma = 40 \) and results of DKI parameter estimation in a
single MPG direction using 2 baseline methods
(GCF and JCF) for 6 b-value and 3 b-value data.
Generally, higher noise levels yield more salt-and-
pepper noise on \( D \) and \( K \) images. It is interesting
that estimation of \( D \) is slightly better in 3 b-values
by the JCF method, and estimation of \( K \) is better in
6 b-values by the GCF method.

In the experiments based on synthetic datasets,
we applied the 6 methods of GCF, WGCF, SOR,
MOR, NS, and BNS for the 6 b-value data, but
we used only NS, BNS, and JCF for the 3 b-value
data. We evaluated the methods from 4 viewpoints—
visual assessment of images, results of PRC and
PDT, root-mean-square (RMS) error, and estima-
tion bias. We obtained the RMS error at \( S_0 > 0 \) area
by comparing the estimated parameter and true val-
The estimation bias is defined as the difference between the true and median values in each region of D and K combinations.

Clinical datasets: The institutional review board of Juntendo Hospital, where we acquired volunteer datasets, approved this study. We used a clinical 3-tesla MR scanner (Philips Achieva, Philips Medical Systems, Best, The Netherlands) and varied the number of b-values to prepare 2 groups of brain datasets. We used 6 b-values, 0, 500, 1000, 1500, 2000, and 2500 s/mm², in the first group, and 6 datasets are included. Pulse duration, δ, was 28.0 ms, and pulse interval, Δ, was 39.0 ms. In the second group, we used 3 b-values, 0, 1000, and 2000 s/mm², for 10 datasets. Both groups of datasets were obtained from healthy volunteers using repetition time [TR]/echo time [TE], 3000/80 ms, and 6 to 32 MPG directions depending on the dataset. In the first group, the numbers of MPG directions were 6, 15, 20, 24, 28, 30, and 32. The second group also consisted of datasets of 6, 15, 20, 24, 28, 30, and 32 directions and included 5 datasets of 20 directions. The time parameters for the δ and Δ pulses were 13.3 and 45.3 ms. The matrix size was 128 × 128 for the first group of 6 b-values and 80 × 80 for the second group. Figure 5 shows examples. For all the datasets, the estimated level of SNR was equivalent to that of σ = 20–40 of the synthetic datasets. We estimated SNR in a clinical dataset at several regions of interest (ROIs) in the S₀ images and compared the values with those of synthetic datasets to find the closest SNR level.

As well as the synthetic data experiments, the 6 methods of GCF, WGCF, SOR, MOR, NS, and BNS are applied for the clinical data with 6 b-values, whereas NS, BNS, and JCF are used for the data with 3 b-values. For the clinical datasets, we evaluated the 6 methods visually and by PRC and PDT results.

Experimental Results

1. Synthetic data experiments

In the synthetic datasets, signal values are corrected based on Equation [24] for bias removal of Rician noise before subsequent experiments to compute DKI parameters.

Visual assessments: Fig. 6(a) shows examples of images of estimated D and K parameters in a single MPG direction for σ = 30 data with 6 b-values. It is natural that the lower S₀ yields the worst estimation of D and K as a result of the relatively low SNR and the integer quantization effect. In addition, lower D is found to make slightly more errors in the K estimation. This can also be attributed to a quantiza-

Fig. 5. Clinical data sample. (a) With 6 b-values and (b) with 3 b-values. Axial sections of a single motion-probing gradient (MPG) direction are displayed.

Fig. 6. Parameter estimation results in σ = 30 data (left: D, right: K). General closed form (GCF), weighted GCF (WGCF), single outlier removal (SOR), multiple outlier removal (MOR), neighbor sampling (NS), and blocked neighbor sampling (BNS) results. (a) Six b-value data and (b) 3 b-value data.
tion effect. At the border of areas of \( S_0 > 0 \) and \( S_0 = 0 \), parameter estimation errors by the NS and BNS methods are noticeable, although the quality of parameter estimation in the regions inside the border seems better than that by the other 4 methods. Compared to BNS, NS showed a smoothing effect at the borders between the \( K = 0 \) and \( K > 0 \) regions.

We obtained similar results for the 3 b-value datasets shown in Fig. 6(b), which displays the estimated parameter images for the data with \( \sigma = 30 \) noise. The baseline JCF method seems to yield a noisy image, whereas the other two showed smoother results.

**On PRC/PDT results:** First, we checked the PDT results by each method and found very few PDT errors, except when the NS or BNS method was used. The BNS technique yielded the most PDT errors; Fig. 7 shows a typical example for 6 b-value data, in which neither of the estimated parameters passed the PRC \( (D = 1.34 \times 10^{-3} \text{mm}^2/\text{s} \text{ and } K = 82.9) \). The fitting result seems to achieve the minimum residual for the sample pairs, but the Hessian is not positive definite for these samples; consequently, the stationary point conditions yield no more minimization of the objective function \[5\]. We confirmed that these errors are found at the border of \( S_0 > 0 \) and \( S_0 = 0 \), where the BNS method picks up signals across the border. The frequency of PDT error was 0.0001% at maximum by BNS among the 5 synthetic datasets at the different noise levels. The PDT error results mostly from the negative values of the first LPM \( H_1 \) in Equations [20]. The frequency of PDT errors gradually increases when we add more noise. We also examined the PDT errors by BNS with an extra noise level of \( \sigma = 100 \) and obtained an error frequency lower than 0.02%. These facts imply that the positive definiteness of our objective function \[5\] is virtually guaranteed within the practical range of image quality.

We examined the success rates of PRC for each method for the 6 b-value datasets, and Fig. 8(a) shows the results. Basically, success was less with more noise. Among the 6 methods, success was highest with the MOR method followed by the BNS method. These results are understandable because the possibility for avoiding PRC errors depends on the numbers of samples and combinations employed in each method. As shown in Table 1, the PRC yields higher success rates when the numbers of samples and combinations are larger. For the 3 b-value datasets, BNS showed the best performance among the 3 methods, as shown in Fig. 8(b).

**On RMS error:** When we observe the RMS errors of the estimated parameters within the dataset, higher PRC success rates do not necessarily guarantee low RMS errors. Figure 9 shows the RMS
errors of $D$ and $K$ estimation at the different levels of noise. The RMS was computed within the area of the signal $S(0)$ that is higher than 160. In the 6 b-value data results, the worst RMS for $K$ estimation was achieved by the MOR method, which showed the best performance for PRC. As described earlier, we infer that it is caused by overfitting. Figure 10(a–c) shows the results of outlier removal by the SOR and MOR methods. As Fig. 10(a) shows, samples in DWI of higher b-values are mainly removed in SOR results. For the MOR method, Fig. 10(b, c) indicates that 2 samples are mostly removed. Among the removal of 2 samples, pairs of such high b-values as 2000 and 2500 are often removed. These tendencies can be caused by lower SNR in higher b-values and seem very reasonable in considering the principle for sample removal. However, as seen in the RMS analysis, the 2 conditions, limiting the number of removed samples and PRC, for sample re-

### Table 1. Number of combinations and signal samples

|                | General closed form (GCF) | Weighted GCF (WGCF) | Single outlier removal (SOR) | Multiple outlier removal (MOR) | Neighbor sampling (NS) | Blocked neighbor sampling (BNS) |
|----------------|---------------------------|---------------------|-------------------------------|-------------------------------|-----------------------|-------------------------------|
| Combinations   | 1                         | 1                   | 7                             | 22                            | 1                     | 4                             |
| Samples        | 6                         | 6                   | 6                             | 6                             | 30                    | 24                            |

|                | NS             | BNS            | Jensen’s closed form (JCF)   |
|----------------|----------------|----------------|-----------------------------|
| Combinations   | 1              | 4              | 1                           |
| Samples        | 15             | 12             | 3                           |

![Fig. 9. Root mean square (RMS) errors in $D$ and $K$ estimation by 6 methods. (a) For $D$ in 6 b-value data, (b) for $K$ in 6 b-value data, (c) for $D$ in 3 b-value data, and (d) for $K$ in 3 b-value data.](image)
moval with simple minimization of the fitting residual are not necessarily sufficient to avoid the over-fitting problem.

In the data for 3 b-values, the JCF method suffers from errors even with less noise than with the other 2 methods. It is noticeable that the NS method can outperform the other two only for estimating kurtosis when the noise level is \( \sigma = 50 \).

On the whole, for the methods with neighbor sampling, NS and BNS showed better RMS results than others at higher noise levels. Especially, the BNS method shows the lowest errors for both \( D \) and \( K \), whereas the performance of others differs depending on the parameter to be estimated. Results for the most basic method of GCF and the slightly better WGCF were reasonable when the noise level was lower than \( \sigma = 30 \) in the data for 6 b-values. In that for 3 b-values, JCF also displayed similar quality when \( \sigma < 30 \). However, the results drastically worsened at the higher noise level. Therefore, if the quality of images is somewhat guaranteed, we suppose that these methods can also be chosen because of their computational simplicity.

On estimation bias: Fig. 11 shows the box plot of estimated DKI parameters at the region of \( S_0 = 320, D = 0.64 \times 10^{-3} \text{ mm}^2/\text{s}, K = 1.2 \) within the dataset of \( \sigma = 30 \). In addition to the 5 levels of John Tukey, values of true minimum (\( * \)), true maximum (\( + \)), and mean (\( \times \)) are displayed, whereas other outliers are omitted from display. The digits in parentheses indicate the minimum and maximum values extraordinarily out of range. In comparison with GCF, other methods generally improve estimation quality. However, in the 6 b-value data results, SOR and MOR again seem to suffer from over-fitting to yield extraordinary values, though the smaller box ranges (from the 25th to the 75th percentiles) are obtained.

As well as the RMS results, results of the NS and BNS methods showed better quality here. From the viewpoint of bias of median values of estimated parameters, however, BNS can outperform NS, especially in kurtosis estimation, as seen in Fig. 11(b, d). In each method, more bias is observed with more noise (Fig. 12). It is interesting that each method has a different tendency to increase positive or negative bias. For further investigation, Fig. 13 summarizes the dependency of estimation bias on \( D \) and \( K \) values at the noise level \( \sigma = 30 \). Whereas low values of the 2 parameters basically yield more bias, biases for both \( D \) and \( K \) are suppressed most in the BNS method. Another advantage of BNS is that its bias is consistently negative, whereas the bias of other techniques, such as NS and MOR, is inconsistent.

2. Clinical data

Figure 14(a) shows examples of \( D \) and \( K \) images estimated for a 6 b-value dataset by each method in an MPG direction, and Fig. 14(b) shows their mean images among MPG directions. We observed pep-
Fig. 11. Box-plot for parameter estimation for $D = 0.64 \times 10^{-3}$ mm$^2$/s and $K = 1.2$ ($\sigma = 30$). (a) For $D$ in 6 b-value data, (b) for $K$ in 6 b-value data, (c) for $D$ in 3 b-value data, and (d) for $K$ in 3 b-value data.

Fig. 12. Estimation bias as average of the difference between true value and medians among all estimated parameters. (a) For $D$ in 6 b-value data, (b) for $K$ in 6 b-value data, (c) for $D$ in 3 b-value data, and (d) for $K$ in 3 b-value data.
per noise by each method, but images by NS and BNS offered less noise. Especially in the mean images, NS and BNS methods yielded smooth results, although NS showed smoother or blurred results. Similar results were observed for the datasets of 3 b-values, as shown by an example in Fig. 14(c, d). Here, the result by BNS was blocky in a single MPG direction compared with the smoother result by NS. Another difference is noticeable around the ventricle in the mean kurtosis image with zooming; that is, the size of the ventricle seems different among the methods. This observation can result from estimation bias. For example, the NS method with a larger negative bias than the other techniques, as shown in Fig. 13, may enlarge such low kurtosis structures as the ventricle. For this reason, we should be cautious when identifying anatomical structures based on kurtosis values.

Figure 15 and Table 2 show the results of PRC and PDT for the 6 b-value and 3 b-value datasets. As well as the results for the synthetic datasets, PDT errors in clinical datasets were rarely observed (mostly less than 1.0%) for any of the methods. Except for over-fitting errors by the SOR and MOR methods, BNS yielded the least errors in both PDT and PRC, followed by NS. However, NS yielded the most PDT errors, at about 3.6%, compared with less than 1.0% by the other methods. This can be attributable to the error at the border between 2 regions with certain contrast, such as the brain surface.

As a whole, the NS and BNS methods performed better, results consistent with the synthetic data results.

Discussion

One of our purposes for introducing GCF-based methods was to expedite computation by replacing numerical methods. Using our in-house software implemented in C/C++ language on a standard PC workstation with Core i7 CPU 920 2.67 GHz (Intel, Santa Clara, CA, USA), the computation time by the baseline GCF for the synthetic dataset was about 30 seconds. Although computation time depends on implementation of the computation platform, parallelization, and other factors, computation was 10 times faster using the baseline GCF method compared to our implementation of the numerical LM method despite identical estimated parameters in most cases. In a few cases, obtained parameters differed between the GCF and numerical methods. We confirmed that this was caused by the non-positive definiteness of the Hessian; that is, the numerical method could not make the estimated parameters converge. This can also be a motivation for replacing the numerical method with the GCF and derived methods. To our surprise, computation time differed only slightly among the derived methods for 6 b-values—GCF, WGCF, SOR, MOR, NS, and BNS. This can be due to pre-computation of several constant values, \( \lambda_{0} \), of [10], which are independent of signal values. Such methods as SOR, MOR, and BNS with several choices for minimizing residuals took slightly more time than the baseline methods of GCF and JCF, but the increase was less than 10%. Thus, the GCF-based methods can improve robustness by keeping computation time within a practical range.
Generally, including numerical methods, it is important to reduce computation time by avoiding redundant computations because it may provide a key for new computation algorithms. Actually, while trying to reduce signal-dependent computation in the numerical computation, we obtained one trigger for our GCF solution by separating pre-computable constants. In addition, recent software tools for symbolic mathematics can help us obtain these kinds of analytic solution. In our studies, we first obtained GCFs manually and then checked them with a symbolic math tool in Matlab (MathWorks Inc., Natick, MA, USA). Appendix A shows our Matlab code for obtaining the GCF.

Our GCF and extended methods are based on the signal model expressed in Formula [2], which is identical to a cumulant expansion for the diffusion MR signal. Otherwise, there are several ways to obtain DKI parameters. For example, another solution can be obtained via reparameterization (shown in Appendix B) with the same signal model of Formula [2]. Although we confirmed similar performances for this solution and GCF, results were slightly worse in our preliminary experiment.

Fig. 14. Parameter estimation results for clinical dataset. (a) $D$ and $K$ images in a single motion-probing gradient (MPG) direction of 6 b-value data, (b) mean $D$ and mean $K$ images of 6 b-value data, (c) $D$ and $K$ images in a single MPG direction of 3 b-value data, and (d) mean $D$ and mean $K$ images of 3 b-value data.
We are currently investigating the relationship between our GCF and the linear solution by reparameterization, including computational efficiency. On the other hand, the kurtosis parameter is originally defined for the non-Gaussian form of the probability density function (PDF) of displaced water molecules. Therefore, it seems more natural to obtain the PDF first and then measure directly the kurtosis of its form. Based on the notion of Q-space imaging (QSI), the PDF is obtained via Fourier transform of the normalized signal decay curve, and several reports utilize direct estimation from the PDF.29,30 The potentials of those methods and our GCF may be comparable, but generally, reliable PDF estimation with fewer b-values, such as for clinical DKI data, is quite difficult. Synthetic data studies such as ours but with more variations in the number of b-values may help performance comparison.

To increase the robustness of computation, the PDT is useful for checking if the parameter estimation is a residual minimizing problem. In our experiments, the clinical datasets yielded more PDT errors than the synthetic data. We obtained PDT

![Figure 15](image.png)

**Figure 15.** Results for positive-definiteness test (PDT) and parameter range check (PRC) in clinical data (a) with 6 b-value data and (b) with 3 b-value data.
errors at the brain surface mainly by the NS method. Further observation revealed distortion from residual eddy current around the brain surface and other regions. Because of the very low signal contrast in higher b-value DWI, the use of image registration techniques based on signal intensity is not enough to correct distortion. Therefore, it is worth considering a new distortion correction technique that is more robust and usable for high-b-value DWI.

Our experiments over a wide spectrum revealed interesting characteristics of the methods. As far as we observe the RMS in kurtosis estimation, there seems to be a critical noise level at which the quantity of RMS errors changes drastically depending on the method employed. For example, in Fig. 9(b), RMS errors increase suddenly between $\sigma = 30$ and $\sigma = 40$ in the baseline GCF and WGCF methods and between $\sigma = 20$ and $\sigma = 30$ in the SOR and MOR methods. Accordingly, when the noise level of datasets is known to be low, we can implement DKI analysis software using simple GCF or WGCF rather than BNS, which is more complex. In the axial layout of our synthetic datasets, we assigned 3 variations of $S_0$ values (Fig. 3). When we observe such parameter estimation results as those in the images in Figs. 4 and 6, it is obvious that regions of low $S_0$ showed noisier results. We attribute this to the natural degradation of relative SNR with lower $S_0$ and a quantization effect for integer signal values. Because most clinical scanners yield DWI data in a DICOM format, signal values are converted to integers. This fact implies that the range of $S_0$ should be set as wide as possible in observing the maximum signal value that is not saturated.

An apparent limitation of this study is that we focused on robust DKI parameter estimation in a single MPG direction. In our next step, we will extend parameter estimation in multiple MPG directions; that is, anisotropic analysis of non-Gaussianity. As well as DTI treating orientational anisotropy of diffusion coefficient by second-order diffusion tensor (DT), the diffusion kurtosis tensor was introduced to describe anisotropy of kurtosis by a fourth-order tensor. The diffusion kurtosis tensor (DKT) can be obtained from a series of DWI data in 2 ways—by computing $D$ and $K$ in each MPG direction and then approximating DT and DKT and by the concurrent estimation of each element of DT and DKT. For the former approach, our GCF methods can be employed directly, but the latter approach requires our modifying the objective function and tackling different conditions on positive definiteness of its Hessian. We are currently working on both approaches in addition to a non-tensor approach based on direct analysis of the orientational kurtosis profile using spherical harmonics.

From the experimental results for evaluation of PRC and PDT errors, RMS error, and parameter estimation bias, we may tentatively conclude that BNS performs best among the 6 methods. As mentioned earlier, however, this study is intended to clarify the characteristics of each method; thus, combinations of the methods should be considered for building a more robust method. For example, the SOR and MOR outlier removal methods are expected to be improved by combination with other methods. It is clear that outlier samples negatively affect curve-fitting, and outlier removal must contribute to good parameter estimation. Compared with the other problems such as problems in computer vision, DKI parameter fitting provides a very limited number of samples, so a combination with the NS or BNS method is expected to avoid over-fitting by increasing the number of samples. In such a combination, other criteria in addition to our PDT and PRC can be introduced for judging the reliability of estimated parameters and curve. For example, the location of the vertex of a parabola of the estimated curve in log-signal domain may also help judge reliability, and the estimated $S_0$ parameter can be used for judging the reliability of the estimated curve because the measured $S_0$ value is most reliable among the measured signal sample with the lowest b-value.

Another extension of our GCF study can cover extension to DKI parameters in higher order signal models. In our preliminary trials with symbolic math software, the stationary point conditions on objective function could yield multiple solutions. Though the usefulness of the kurtosis parameter by higher order signal models remains unclear, robust estimation is necessary to prove it. Therefore, we also intend to develop their robust estimation method including its validation in clinical data.

In summary, we showed that the general closed expressions for DKI parameters, including diffusion coefficient and diffusional kurtosis, provide faster computation than numerical methods in DKI with any number of b-values more than three and can be extended using several techniques, such as neighbor sampling, for more robust estimation of parameters.

Acknowledgements
This research is partially supported by the Japan Society for the Promotion of Science (JSPS) Grant-
in-aid for Scientific Research (B) #21300185 during 2009 to 2011. The authors are grateful to Dr. Masaaki Hori and Mr. Issei Fukunaga of Juntendo Hospital for MR data acquisition from volunteers and for discussions in clinical viewpoints. We also thank Professor Jens H. Jensen of the Medical University of South Carolina, Charleston, SC, USA, for encouraging us to publish this study and providing useful information.

Appendix

A. Matlab code for the GCF solutions
This code was verified with version 2012a including symbolic math tool. Note that the closed form expression of the objective function $J$ and its Hessian at the stationary point can be also easily obtained by substituting the solutions of $D, K,$ and $S_0$.

```matlab
% symbolic variables
syms F J D K So n A b(n) S(n) L10 L20 L30 L40 L01 L02 L11 L21 sol;

% difference between model and signal value in log domain
F = log(So) - D * b(k) + D^2 * K * b(k)^2 / 6 - log(S(k));

% objective function to be minimized (average of squared F over samples = RMS)
J = subs(symsum(expand(F^2 / N), k, 1, N), ...
   {symsum(b(k), k, 1, N), ...
    symsum(b(k)^2, k, 1, N), ...
    symsum(b(k)^3, k, 1, N), ...
    symsum(b(k)^4, k, 1, N), ...
    symsum(log(S(k)), k, 1, N), ...
    symsum(log(S(k))^2, k, 1, N), ...
    symsum(b(k)*log(S(k)), k, 1, N), ...
    symsum(b(k)^2*log(S(k)), k, 1, N), ...
    {N*L10, N*L20, N*L30, N*L40, N*L01, N*L02, N*L11, N*L21});
J = simplify(J);

% D, K, So, solutions by solving equations of stationary point conditions
sol = solve(diff(J, D), diff(J, K), diff(J, So), D, K, So);
sol.D = simple(expand(sol.D, 'IgnoreAnalyticConstraints', true));
sol.K = simple(expand(sol.K, 'IgnoreAnalyticConstraints', true));
sol.So = simple(expand(sol.So, 'IgnoreAnalyticConstraints', true));
```

B. Linear solution with reparameterization
When we reparameterize the three variables; $D, K, S_0$, as

$$
P = \begin{pmatrix}
p \\ q \\ r
\end{pmatrix} = \begin{pmatrix}
\log S_0 \\
-D \\
-D^2 K / 6
\end{pmatrix} \quad [B.1]
$$
we obtain the LSF problem in linear equations as

$$
BP \cong \begin{pmatrix}
\log S_1 \\
\vdots \\
\log S_N
\end{pmatrix} \quad [B.2]
$$
where $B$ is a matrix consisting of polynomials of b-values as follows:

$$
B = \begin{pmatrix}
1 & b_1 & b_1^2 \\
& \vdots & \vdots \\
1 & b_N & b_N^2
\end{pmatrix} \quad [B.3]
$$

Then, we obtain the estimated parameters $P$, including cases of $N > 3$ as:

$$
\hat{P} = \begin{pmatrix}
\hat{p} \\ \hat{q} \\ \hat{r}
\end{pmatrix} = (B^T B)^{-1} B^T \begin{pmatrix}
\log S_1 \\
\vdots \\
\log S_N
\end{pmatrix} \quad [B.4]
$$
and then we finally obtain the estimated DKI parameters by the inverse reparameterization as follows.
\[
\begin{pmatrix}
\hat{S}_0 \\
\hat{D} \\
\hat{K}
\end{pmatrix} = \begin{pmatrix}
\exp(\hat{p}) \\
\hat{q} \\
\hat{r}
\end{pmatrix}
\quad \quad \text{[B.5]}
\]

This solution is also faster than the numerical methods and is even faster by precomputing the combined matrix; \((B^TB)^{-1}B^T\) in advance, which is signal independent. Generally, the reparameterization technique is often used as an option for solving a non-linear LSF problem. One of the potential issues in this approach when applied in DKI might be that the optimization process is indirect regarding the parameter estimation due to the reparameterization. It is necessary to investigate the relationship between our GCF and this linear solution.

References

1. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 1986; 161:401–407.
2. Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from NMR spin echo. J Magn Reson B 1994; 103:247–254.
3. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczkynski K. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 2005; 53:1432–1440.
4. Falangola MF, Jensen JH, Babb JS, et al. Age-related non-Gaussian diffusion patterns in the prefrontal brain. J Magn Reson Imaging 2008; 28:1345–1350.
5. Jensen JH, Helpern JA. Quantifying non-Gaussian water diffusion by means of pulsed-field-gradient MRI. Proc 11th Ann Mtg ISMRM Toronto, Canada 2003; 2154.
6. Lu H, Jensen JH, Hu C, Falangola MF, Ramani A, Ferris S, Helpern JA. Alterations in cerebral microstructural integrity in normal aging and in Alzheimer’s disease: a multi-contrast diffusion MRI study. Proc 14th Ann Mtg ISMRM Seattle, WA, USA 2006; 723.
7. Ramani A, Jensen JH, Szulc KU, et al. Assessment of abnormalities in the cerebral microstructure of schizophrenia patients: a diffusional kurtosis imaging study. Proc 15th Ann Mtg ISMRM Berlin, Germany 2007; 648.
8. Raab P, Hattingen E, Franz K, Zanella FE, Lannermann H. Cerebral gliomas: diffusional kurtosis imaging analysis of microstructural differences. Radiology 2010; 254:876–881.
9. Van Cauter S, Vercaut J, Sijbers J, et al. Gliomas: diffusion kurtosis MR imaging in grading. Radiology 2012; 263:492–501.
10. Wang JJ, Lin WY, Lu CS, et al. Parkinson disease: diagnostic utility of diffusion kurtosis imaging. Radiology 2011; 261:210–217.
11. Helpern JA, Falangola MF, Di Martino A, et al. Alterations in brain microstructure in ADHD by diffusional kurtosis imaging. Proc 15th Ann Mtg ISMRM Berlin, Germany 2007; 1580.
12. Rosenkrantz AB, Sigmund EE, Winnick A, et al. Assessment of hepatocellular carcinoma using apparent diffusion coefficient and diffusion kurtosis indices: preliminary experience in fresh liver explants. Magn Reson Imaging 2012; 30:1534–1540.
13. Hori M, Fukunaga I, Masutani Y, et al. Visualizing non-Gaussian diffusion: clinical application of q-space imaging and diffusional kurtosis imaging of the brain and spine. Magn Reson Med Sci 2012; 11:221–233.
14. 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.
15. Veraart J, Poot DH, Van Hecke W, et al. More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging. Magn Reson Med 2011; 65:138–145.
16. Tanner JE, Stejskal EO. Restricted self-diffusion of protons in colloidal systems by the pulsed-gradient, spin-echo method. J Chem Phys 1968; 49:1768–1777.
17. Jensen JH, Hu C, Helpern JA. Rapid data acquisition and postprocessing for diffusional kurtosis imaging. Proc 17th Ann Mtg ISMRM Honolulu, HI, USA 2009; 1403.
18. Jensen JH, Helpern JA. MRI quantification of non-Gaussian water diffusion by kurtosis analysis. NMR Biomed 2010; 23:698–710.
19. Masutani Y, Aoki S. General closed-form expressions for DKI parameters and their application to fast and robust DKI computation based on outlier removal. Proc 20th Ann Mtg ISMRM Melbourne, Australia 2012; 20:3622.
20. Levenberg K. A method for the solution of certain non-linear problems in least squares. Quarterly of Applied Mathematics 1944; 2:164–168.
21. Marquardt D. An algorithm for least-squares estimation of nonlinear parameters. SIAM J Appl Math 1963; 11:431–441.
22. Gilbert GT. Positive definite matrices and Sylvester’s criterion. Am Math Month 1991; 98:44–46.
23. Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Comm ACM 1981; 24:381–395.
24. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med 1995; 34:910–914.
25. www.mathworks.com/products/matlab/
26. Liu C, Bammer R, Acar B, Moseley ME. Characterizing non–Gaussian diffusion by using generalized diffusion tensors. Magn Reson Med 2004; 51:924–937.
27. Mitra PP, Sen PN. Effects of microgeometry and surface relaxation on NMR pulsed-field-gradient experiments: simple pore geometries. Phys Rev B Condens Matter 1992; 45:143–156.
28. Frøhlich AF, Ostergaard L, Kiselev VG. Effect of impermeable boundaries on diffusion-attenuated MR signal. J Magn Reson 2006; 179:223–233.
29. Umezawa E, Yoshikawa M, Yamaguchi K, Ueoku S, Tanaka E. q-Space imaging using small magnetic field gradient. Magn Reson Med Sci 2006; 5:179–189.
30. Lätt J, Nilsson M, Wirestam R, et al. In vivo visualization of displacement-distribution-derived parameters in q-space imaging. Magn Reson Imaging 2008; 26:77–87.
31. Haselgrove JC, Moore JR. Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient. Magn Reson Med 1996; 36:960–964.
32. Barmpoutis A, Zhuo J. Diffusion kurtosis imaging: robust estimation from DW-MRI using homogeneous polynomials. Proc 8th IEEE Int Symp Biomed Imag 2011:262–265.
33. Veraart J, Van Hecke W, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. Magn Reson Med 2011; 66:678–686.
34. Lu H, Jensen JH, Ramani A, Helpern JA. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. NMR Biomed 2006; 19:236–247.
35. Suzuki Y, Van Cauteren M. Investigation of liver diffusion kurtosis imaging and perfusion effect. Proc 20th Ann Mtg Int Soc Magn Reson Med, Melbourne, Australia 2012; 91:4119.