Journal of Magnetic Resonance 213 (2011) 136–144

Contents lists available at SciVerse ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Robust tensor estimation in diffusion tensor imaging

Ivan I. Maximov [a,*], Farida Grinberg [a], N. Jon Shah [a,b]

[a] Institute of Neuroscience and Medicine – 4, Forschungszentrum Juelich GmbH, 52425 Juelich, Germany
[b] Department of Neurology, Faculty of Medicine, RWTH Aachen University, JARA, 52074 Aachen, Germany

ARTICLE INFO

Article history:
Received 26 April 2011
Revised 7 September 2011
Available online 22 September 2011

Keywords:
Diffusion tensor imaging
Noise correction
Least median squares
Robust tensor estimation
Non-linear constrained robust regression

ABSTRACT

The signal response measured in diffusion tensor imaging is subject to detrimental influences caused by noise. Noise fields arise due to various contributions such as thermal and physiological noise and sources related to the hardware imperfection. As a result, diffusion tensors estimated by different linear and non-linear least squares methods in absence of a proper noise correction tend to be substantially corrupted. In this work, we propose an advanced tensor estimation approach based on the least median squares method of the robust statistics. Both constrained and non-constrained versions of the method are considered. The performance of the developed algorithm is compared to that of the conventional least squares method and of the alternative robust methods proposed in the literature. Two examples of simulated diffusion attenuations and experimental *in vivo* diffusion data sets were used as a basis for comparison. The robust algorithms were shown to be advantageous compared to the least squares method in the cases where elimination of the outliers is desirable. Additionally, the constraints were applied in order to prevent generation of the non-positive definite tensors and reduce related artefacts in the maps of fractional anisotropy. The developed method can potentially be exploited also by other MR techniques where a robust regression or outlier localisation is required.

1. Introduction

Diffusion tensor imaging (DTI) is an outstanding non-invasive technique providing valuable information on tissue microstructure and is well established in clinical studies. DTI has been successfully applied in diagnosis and evaluation of acute stroke, tumours and various neurological disorders [1–3]. Most frequently, conventional diagnostic DTI tools utilise the maps of diffusion scalar metrics such as the apparent diffusion coefficient, the fractional anisotropy (FA), a colour FA and other rotational invariants. The quality of these maps and, thus, the consistency of the diagnostic outcome is dependent on the evaluation accuracy of the diffusion tensors from the experimental raw data [4–7]. In turn, evaluation precision is influenced by the noise level and the prescribed diffusion gradient directions [8].

The signal in DTI experiments is distorted by the noise which becomes especially crucial in the range of high diffusion weightings (the so-called "*b*-factors"). Severe artefacts may also originate from the sources of physiological noise such as a cardiac pulsation, bulk head motion, and respiratory motion. Hardware instabilities related to the technical imperfections produce additional contributions. As a consequence, diffusion tensor estimation becomes unstable and may even give rise to non-positive definite tensors [9].

Basically, modern methods of diffusion tensor estimation can be classified into those based on the linear least squares (LLS), the non-linear least squares (NLLS) and their counterparts using constraints (cLLS and cNLLS, respectively) [10,11]. These methods usually produce reliable results as long as raw data sets are not contaminated by the outliers originating from noise. Measured signal attenuations corrupted by the outliers cannot be efficiently corrected using either the Gaussian or Rician noise distributions. Such error sets that cannot be treated by any known distributions and are, generally, subject to investigation by the robust statistics [12]. Numerous robust algorithms exist based on various robust estimators such as M-, S-, Q-estimators, least median/trimmed squares, repeated medians, etc. [12].

The first applications of robust estimators in DTI were demonstrated by Mangin et al. [13] and Chang et al. [14]. The robust algorithm proposed in Ref. [14], the so-called RESTORE, is based on the well-known Geman–McClure M-estimator [15]. The M-estimator allows one to localise outliers and to exclude them from the subsequent data fitting procedure. More recently, the problem of the outlier localisation has been addressed in a dedicated study [16]. Using statistical analysis over a large data set, Walker et al. [16] have shown that certain outliers in diffusion weighted images of the human brain arise as a result of cardiac gating applied in the acquisition scheme and tend, therefore, to have a common localisation within the images. It was concluded that a general assumption postulating the homogeneity of the statistical properties of DTI

* Corresponding author. Fax: +49 2461 61 2820.
E-mail address: i.maximov@fz-juelich.de (I.I. Maximov).

1090-7807/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2011.09.035

across the brain is not valid. A disadvantage of the M-estimator used in Refs. [13,14,16] is, however, that it is characteristic of a rather low breakdown point. This means that in the cases in which the number of outliers is close to the validation limit of the robust estimator, it can lead to the wrong outlier detection. Unstable robust estimations become especially pronounced in the case of non-linear regression or in the minimisation problem with constraints [17].

Improved approaches for diffusion tensor estimation were proposed by Landman et al. [18,19]. The authors hypothesised that noise fields are spatially variable and suggested an application of the Q-estimator. Utilisation of the robust Q-estimator allowed one to substantially reduce the inaccuracy of the diffusion tensor evaluation by accounting for spatial variation of the noise fields and by the implied procedure of outlier detection.

In this work, we propose a new, more robust and more stable approach to diffusion tensor estimation based on the least median squares (LMS) [12,20]. The LMS method and its modification the least trimmed squares (LTS) algorithm, allows one to exclude the artefacts related to noise straightforwardly during the minimisation procedure. This is in contrast to the RESTORE algorithm in which the outlier detection and the fitting procedure are performed as two subsequent steps. We developed the LMS algorithm for the non-linear minimisation function using a constrained condition of positive definite diffusion tensors that allows one to prevent physically unjustified solutions.

Generally, the constrained LMS/LTS method is associated with certain mathematical and numerical difficulties [17,21]. In particular, solutions obtained by the LMS/LTS algorithm are not unique and might be potentially prone to local instabilities. The basic advantages of the LMS approach consist, however, in a sufficiently high breakdown point level and a relatively simple realisation of the numerical algorithm in comparison to other robust approaches exploited previously in DTI data evaluation [13,14,18,19].

The robust algorithm based on the LMS approach developed herein is strictly non-linear and applied with constraints relevant to the problem of diffusion tensor reconstruction. An application of the new method is demonstrated with respect to simulated data and *in vivo* experiments. The results are compared with the standard NLLS, RESTORE and their constrained counterparts.

2. Theory

In DTI experiments, attenuation of the measured signal, S_i, in the ith diffusion gradient direction is described as follows [4,9–11]:

$$S_i = S_0 \exp(-b\mathbf{g}_i^T \mathbf{D} \mathbf{g}_i) \tag{1}$$

where S_0 is the signal amplitude in absence of diffusion weighting, b is the diffusion-weighting factor [4,5], $\mathbf{g}_i$ is the diffusion-encoding unit vector, and $\mathbf{D}$ is a second order symmetric diffusion tensor. Measurements in at least six non-collinear directions are required in order to enable evaluation of tensor components. In this case, a solution of Eq. (1) can be obtained analytically by solving the corresponding system of the linear equations. In practice, more than six gradient directions are applied and the components of the diffusion tensor are evaluated using the NLLS method with the minimising objective, f_{NLLS}, expressed as follows [10,11]:

$$f_{\mathrm{NLLS}}(\mathbf{d}) = \frac{1}{2}\sum_{i=1}^{N}\left[S_i - S_0 \exp\left(\sum_{j=1}^{6} X_{ij} d_j\right)\right]^2 \tag{2}$$

where N is the number of diffusion encoding gradients, $\mathbf{d} = [D_{xx}, D_{xy}, D_{xz}, D_{yy}, D_{yz}, D_{zz}]$ is a vector representation of the diffusion tensor and $\mathbf{X}$ is an encoding gradient design matrix:

$$X = -\begin{pmatrix} b_1\mathbf{g}_{1x}^2 & 2b_1\mathbf{g}_{1x}\mathbf{g}_{1y} & 2b_1\mathbf{g}_{1x}\mathbf{g}_{1z} & b_1\mathbf{g}_{1y}^2 & 2b_1\mathbf{g}_{1y}\mathbf{g}_{1z} & b_1\mathbf{g}_{1z}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_N\mathbf{g}_{Nx}^2 & 2b_N\mathbf{g}_{N_x}\mathbf{g}_{N_y} & 2b_N\mathbf{g}_{N_x}\mathbf{g}_{N_z} & b_N\mathbf{g}_{Ny}^2 & 2b_N\mathbf{g}_{N_y}\mathbf{g}_{N_z} & b_N\mathbf{g}_{Nz}^2 \end{pmatrix} \tag{3}$$

The application of the Geman–McClure M-estimator algorithm (RESTORE) [14] to the minimisation procedure requires an introduction of a weighting function ω_i:

$$\omega_i = \frac{1}{r_i^2 + C^2} \tag{4}$$

where r_i is a residual, that is, a difference between the experimentally measured signal, S_i, and its estimation; C is a scale factor estimated through the median absolute deviation (MAD): $C = 1.4826 \times \mathrm{MAD}$. The weighting function ω_i allows one to detect and reject the outliers from the DTI raw data under an assumption that errors follow a Gaussian distribution. Alternative weighting functions were suggested by other authors [12,17,20].

3. Constrained non-linear LMS/LTS approaches (cLMS/cLTS)

The classical least squares method exploits the sum of squared residuals. In the least median squares method, the sum of squared residuals is replaced by the median of squared residuals. As a result, one obtains a robust estimator with a high resistance to contamination of the raw data sets by outliers. The amount of outliers can reach up to 50% of the data set. For linear functions, the efficiency of the LMS algorithm can be improved with the help of the "least trimmed squares" [12,20]. In this case, the following minimisation function is evaluated:

$$\min \sum_{i=1}^{h} (r_i^2)_{1:N} \tag{5}$$

where residuals are arranged in such a way that $r_1^2 < r_2^2 < \cdots < r_N^2$ and h is a truncation factor [12,20,21]. With respect to DTI, the LTS operates on logarithms of signal amplitudes rather than signal amplitudes themselves. During linear regression, the LTS method tends to converge much faster than the LMS method. In turn, an application of linear regression algorithms in DTI evaluation leads to substantial deviations in the estimation of the diffusion tensor due to the noise effects [22]. The application of the LTS algorithm can be easily generalised for the constrained non-linear optimisation [17,21] using Eq. (5). However, in the case of the non-linear problem, the LTS tends to lose its former facility of faster convergence, in contrast to the LMS.

In order to guarantee a positive definite solution an additional constraint in the minimisation algorithm is required. One of the most efficient ways of applying the constraint is based on Cholesky decomposition [11,23,24]. The latter exploits the fact that the matrix $\mathbf{D}$ can be represented as a multiplication of two triangular matrices $\mathbf{L}$:

$$\mathbf{D} = \mathbf{L}^T\mathbf{L}, \quad \mathbf{L} = \begin{pmatrix} \rho_1 & \rho_2 & \rho_3 \\ 0 & \rho_4 & \rho_5 \\ 0 & 0 & \rho_6 \end{pmatrix} \tag{6}$$

where the elements of the matrix $\mathbf{L}$ must obey the following conditions:

$$\rho_1 > 0, \quad \rho_4 > 0, \quad \rho_6 > 0 \tag{7}$$

The vector representation of the diffusion tensor $\mathbf{D}$ in Eq. (2) can now be easily rewritten via the elements of the triangular matrix $\mathbf{L}$:

$$\mathbf{d} = [\rho_1^2, \rho_1\rho_2, \rho_1\rho_3, \rho_1^2 + \rho_4^2, \rho_2\rho_3 + \rho_4\rho_5, \rho_3^2 + \rho_5^2 + \rho_6^2] \tag{8}$$

The advantage of Cholesky decomposition is that an applied constraint has a very simple form, see Eqs. (6) and (7). However, the original function in Eq. (2) becomes more complicated for numerical computations since the vector **d** in Eq. (2) is now substituted by a more complex expression, Eq. (8). Alternatively, a non-linear constraint can be expressed using the Sylvester criterion [23]:

$$D_{xx} > 0, \quad D_{xx}D_{yy} - D_{xy}^2, \quad \det(\mathbf{D} > 0) \tag{9}$$

It is noteworthy that for the constrained minimisation problem, the conditions described by Eq. (9) are fully equivalent to those expressed by Eq. (7). The detailed description of different parametrisation schemes of the diffusion tensor representations can be found in Ref. [11].

Due to limitations on the acquisition time, DTI is usually performed either with multiple *b*-values but only a few gradient directions or with a large number of gradient directions but only a few *b*-values. The multiple *b*-value experiments allow one to obtain more detailed information about the tissue microstructure using different non-Gaussian models of diffusion such as the biexponential fitting model [25,27], diffusion kurtosis imaging [26,27], and others. In turn, the multiple gradient directions allow one to substitute the recurring acquisitions by redundant multiple coding gradients, and the latter to apply more sophisticated algorithms reconstructing the orientation distribution function of the diffusion tensors and fibre tracking such as the Q-ball [28], high angular resolution diffusion imaging [29], constrained spherical deconvolution [30], etc. We provide two possible objective functions in order to obtain the maximal efficiency in both cases: multiple *b*-values and multiple gradient directions.

The experiments with multiple *b*-values usually are very time consuming, especially in the case of high *b*-values (in range of [0; 7000] s mm^{-2}). In order to reduce the experimental time frequently only six gradient directions are applied. With the constraints determined by Eq. (7) and the diffusion tensor elements expressed by Eq. (8), the minimising objective, f_{cNLMS}, in the constrained non-linear least median squares algorithm can be formulated as:

$$f_{\text{cNLMS}} = \sum_{i=1}^{N} \text{median}_j [S_{ij} - S_0 \exp(-b_j \mathbf{g}_i^T \mathbf{D} \mathbf{g}_i)]^2 \tag{10}$$

the indexes *i* and *j* refer to the direction of magnetic field gradients and *b*-values, respectively. In the case of more complicated diffusion model, for example, biexponential fitting, Eq. (10) can be easily modified into following:

$$f_{biexp} = \sum_{i=1}^{N} \text{median}_j [S_{ij} - a \exp(-b_j \mathbf{g}_i^T \mathbf{D}_{\text{fast}} \mathbf{g}_i) - (1-a) \exp(-b_j \mathbf{g}_i^T \mathbf{D}_{\text{slow}} \mathbf{g}_i)]^2 \tag{11}$$

where $\mathbf{D}_{\text{fast}}$ and $\mathbf{D}_{\text{slow}}$ are the fast and the slow components of the diffusion tensor, respectively; *a* is the relative fraction of the fast component.

In multiple gradient direction experiments the most relevant is the LTS algorithm of Eq. (5) and the objective function of Eq. (2) with or without constrained conditions of Eqs. (7) and (8) due to its faster convergence feature. The truncation factor *h* has to belong to $\{[N/2]+4; N\}$ interval. The detailed algorithm diagram is described in Appendix A.

4. Experimental

In vivo diffusion studies were carried out with a whole-body 3T Siemens MAGNETOM Tim-Trio scanner (Siemens Medical Systems, Erlangen, Germany). The gradient system provided a maximal gradient strength of 40 mT m^{-1}. Diffusion-weighted images were acquired using a bipolar gradient double spin-echo echo-planar imaging pulse sequence provided by the manufacturer. In the experiments with multiple *b*-values [0, 200, 400, 600, 800, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 7000] s mm^{-2} we used six non-collinear directions of the diffusion encoding gradients. For six gradient directions we used dual gradient scheme. In the case of 30 gradient directions the standard manufacturer scheme and one diffusion weighting $b = 1000$ s mm^{-2} were applied. We studied two healthy volunteers 27 and 37 years old. In the case of the 27 years old volunteer the voxel size in six directional experiments was $2 \times 2 \times 2$ mm^3 and for 37 years old volunteer the voxel size in 30 directional experiments was $2.4 \times 2.4 \times 2.4$ mm^3. The echo times were 113 and 112 ms and the repetition times were 1000 and 8900 ms, respectively. The study was performed in accordance with the ethical approval from the local ethics committee.

All results were post-processed using the in-house Matlab scripts (Matlab, The MathWorks, Natick, MA, USA) and the robust statistics package LIBRA: a MATLAB Library for Robust Analysis (http://wis.kuleuven.be/stat/robust/LIBRA.html) [31].

5. Results

Below, the efficiency of the developed robust LTS/LMS algorithms is compared with the least squares methods, NLLS and cNLLS, and the alternative robust algorithm, RESTORE. A comparison is performed using two simulated data sets and experimental *in vivo* diffusion data.

5.1. Example I: diffusion attenuation in the case of isotropic diffusion

We start here with a simple example in which the performance of the developed LMS method is compared to the NLLS algorithm described in Refs. [10,11] and the RESTORE algorithm [14] in the case of multiple *b*-value simulations. The normalised reference signal attenuation was simulated as an exponential function for isotropic diffusion with the diffusion coefficient $D = 3 \times 10^{-3}$ mm^2 s^{-1}. Rician noise and outliers were added. The simulations were performed for four values of signal-to-noise ratio: SNR = 3, SNR = 5, SNR = 10 and SNR = 20. Diffusion weightings were equal to [0, 250, 500, 750, 1000] s mm^{-2}. Two outliers with random amplitudes were added at $b = 250$ s mm^{-2} and $b = 750$ s mm^{-2}.

The results of the simulations are presented in Fig. 1. The reliability of the robust algorithms, RESTORE and LMS, increased with increasing SNR. The values of diffusivity obtained by the RESTORE and LMS algorithms were: $(3.8 \pm 1.1) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.2 \pm 0.9) \times 10^{-3}$ mm^2 s^{-1} (LMS) at SNR = 3 and $(3.0 \pm 0.2) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.0 \pm 0.2) \times 10^{-3}$ mm^2 s^{-1} (LMS) at SNR = 20. Since the NLLS method has 0% breakdown criterion to the outliers all estimations obtained by the least squares have a substantial deviation from the original diffusion coefficient: the estimated value is 4.6 times higher at SNR = 3 and by more than 60% higher at SNR = 20.

Fig. 2 demonstrates the results of statistical simulations for various outlier amplitudes and SNR levels. A total of 2800 attenuation samples for isotropic diffusion were simulated for SNRs from 1 to 28 units with the added Rician noise distribution. Diffusion weightings were equal to [0; 250; 500; 750; 1000] s mm^{-2}. One outlier was added at $b = 750$ s mm^{-2}. In each simulation, the amplitude of the outlier was randomly chosen from the interval [0; 1]. The simulated data sets were fitted using three methods, the LMS, NLLS, and RESTORE. The curves in Fig. 2 represent the averages, $\langle D \rangle$, of the fitted values of *D* as a function of the SNR for each of these methods. The bars show the standard deviations.

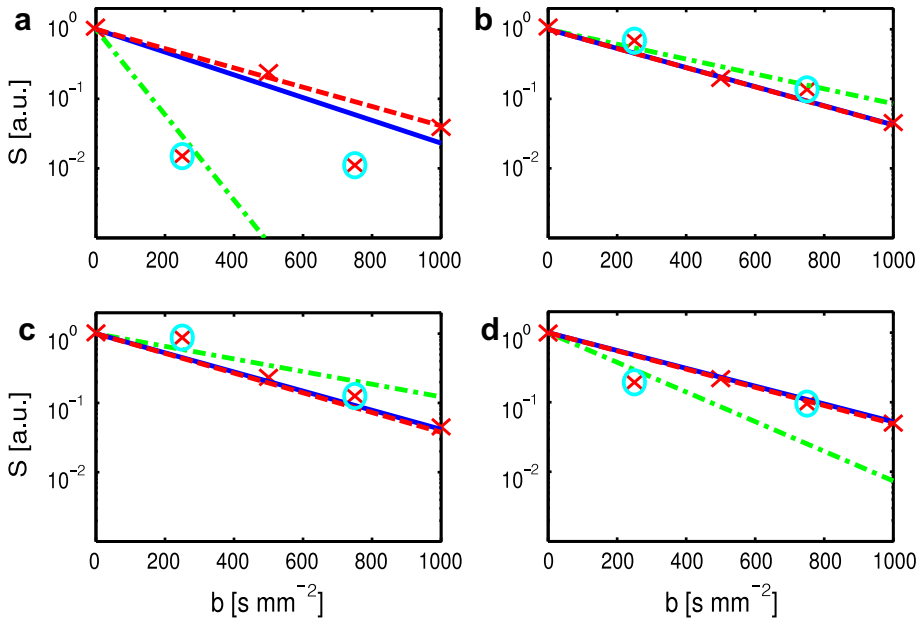

Fig. 1. Signal attenuation and fits for isotropic diffusion with diffusion coefficient $D = 3 \times 10^{-3}$ mm^2 s^{-1}. Fits were performed by the NLLS (green dotted line), RESTORE (blue solid line), and LMS (red dashed line) algorithms. The data points corrupted by outliers are marked by cyan circles. (a) SNR = 3. The corresponding reconstructed diffusion coefficients were equal to $(14.1 \pm 6.5) \times 10^{-3}$ mm^2 s^{-1} (NLLS), $(3.8 \pm 1.1) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.2 \pm 0.9) \times 10^{-3}$ mm^2 s^{-1} (LMS) (b) SNR = 5. The corresponding reconstructed diffusion coefficients were equal to $(2.5 \pm 3.5) \times 10^{-3}$ mm^2 s^{-1} (NLLS), $(3.2 \pm 1.1) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.2 \pm 1.1) \times 10^{-3}$ mm^2 s^{-1} (LMS) (c) SNR = 10. The corresponding reconstructed diffusion coefficients were equal to $(2.1 \pm 1.3) \times 10^{-3}$ mm^2 s^{-1} (NLLS), $(3.2 \pm 0.7) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.2 \pm 0.7) \times 10^{-3}$ mm^2 s^{-1} (LMS) (d) SNR = 20. The corresponding reconstructed diffusion coefficients were equal to $(4.9 \pm 1.5) \times 10^{-3}$ mm^2 s^{-1} (NLLS), $(3.0 \pm 0.2) \times 10^{-3}$ mm^2 s^{-1} (RESTORE), $(3.0 \pm 0.2) \times 10^{-3}$ mm^2 s^{-1} (LMS).

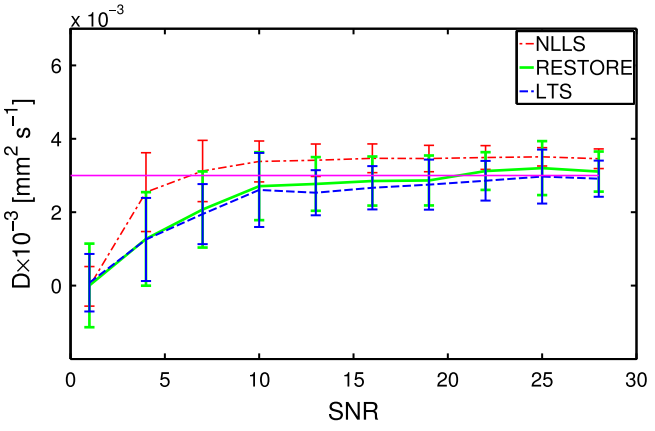

Fig. 2. Average fitted values of the diffusivity as a function of SNR for the NLLS, RESTORE and LMS methods. The fits were applied to simulated signal attenuations distorted by the Rician noise and one outlier at $b = 750$ s mm^{-2} with random amplitude from the range [0; 1]. For each SNR value, 100 samples were simulated for the purpose of statistical evaluation.

Fig. 2 shows that the NLLS (red dotted curve) tends to considerably overestimate the original diffusivity of 3×10^{-3} mm^2 s^{-1} (solid magenta line) in the range of the relatively low noise levels in comparison to LMS (dashed blue curve) and RESTORE (solid green curve). The standard deviation becomes extremely large at high noise levels. In contrast, the LMS and RESTORE produce more reasonable fits for high and moderate SNR levels. With increasing SNR, both LMS and RESTORE exhibit the good fits, however, RESTORE has a tendency to overestimate the original diffusion coefficient (see Fig. 2).

5.2. Example II: diffusion tensors

In this example, diffusion attenuations were generated for 6 and 30 orientations of the diffusion encoding gradients. Gradient orientations were the same as those used for *in vivo* measurements, see the Experimental section. The values of b were equal to [0; 200; 400; 600; 800; 1000] s mm^{-2} for six gradient directions and [0; 1000] s mm^{-2} for 30 gradient directions. The original vector **d** was formed for anisotropic tensor with vector elements equal to $[1.5 \times 10^{-3};\ 0;\ 0;\ 1.5 \times 10^{-3};\ 0;\ 3.0 \times 10^{-3}]$ mm^2 s^{-1} and FA = 0.41. The mean diffusivity was equal to 2.0×10^{-3} mm^2 s^{-1}. The simulated attenuations were subjected to the influence of Rician noise for different values of SNR: 3, 5, 10, 20, and 40 units. In the case of six-gradient-simulations one outlier was applied at $b = 800$ s mm^{-2} to all six directions; random outlier amplitudes were chosen from the range [0; $3S_0$]. In the case of 30-gradient-simulations, we distorted by an outlier each fifth direction at $b = 1000$ s mm^{-2}; random outlier amplitudes were chosen from the range [0; $3S_0$]. In order to obtain the statistical estimation of the algorithm efficiency we provided 200 samples for each SNR unit and each algorithm. The results of the simulations for 6 and 30 gradient directions are presented in Fig. 3a and b, respectively.

Fig. 3a and b demonstrate the principal difference between robust algorithms and the least squares method. In the case of the six gradient directions (see Fig. 3a) the SNR plays a crucial role for diffusion tensor estimation. At low SNR (equal to 3 or 5 units) a scattering of the eigenvalues and eigenvectors is enormous for all methods. In turn, at high SNR the robust algorithms, RESTORE and LMS, demonstrate the substantial improving in contrast to the classical NLLS one. The 30 gradient-direction simulations have a better representation of the robust estimated tensors even at low SNR (>5 units) (see Fig. 3b).

In Fig. 4a and b, we presented the averaged deviation of the angle of the evaluated main eigenvector from the original direction (see Fig. 3) obtained by NLLS, RESTORE and LMS/LTS algorithms. The data is presented as a function of SNR. Fig. 4a refers to the case of six-gradient-direction simulations. The original signals were distorted by the Rician noise and outliers at $b = 800$ s mm^{-2} with random amplitudes from the range [0, $3S_0$]. In the case of the 30

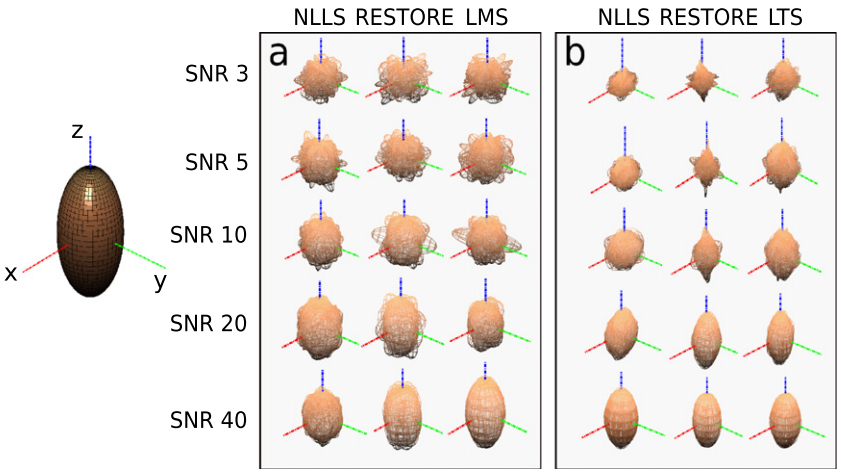

Fig. 3. The statistical estimations of the evaluated diffusion tensor presented in the form of the ellipsoidal skeleton for 6 (a) and 30 (b) gradient directions. The left solid ellipsoid exhibits the original diffusion tensor and shows the coordinate axes. In each statistical sample, the original signal is distorted by the Rician noise and outliers.

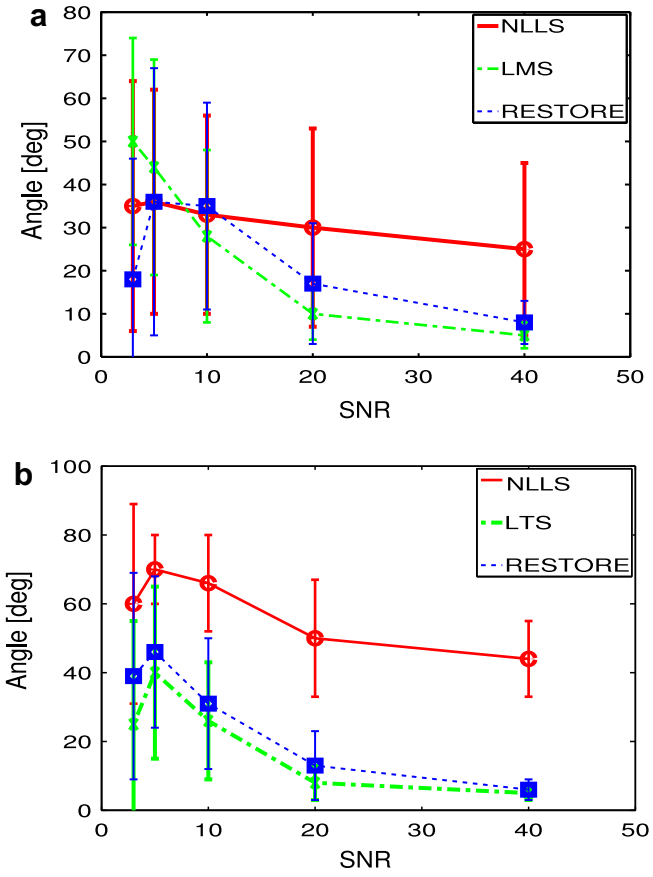

Fig. 4. The averaged angular deviation of the evaluated main eigenvector from the original diffusion tensor obtained by the NLLS, RESTORE and LMS/LTS algorithms. The curves represent a dependence of the angular deviation on the SNR: (a) averaged angular deviation in the case of six gradient directions and one outlier at $b = 800$ s mm^{-2}. (b) Averaged angular deviation in the case of 30 gradient directions and outliers posed in each fifth directions at $b = 1000$ s mm^{-2}.

gradient directions the averaged angle deviations are presented in Fig. 4b. The original signals were corrupted by the Rician noise and outliers posed at each fifth direction with random amplitudes from the range $[0, 3S_0]$. Note that, according to Fig. 4a and b, the NLLS algorithm cannot perform the reasonable assessments even at SNR = 40.

5.3. Example III: in vivo study

In this example we investigate how various numerical algorithms influence evaluation of the tensor eigenvalues and of the related FA maps in the human brain. The results are compared for classical and constrained NLLS, original and constrained RESTORE (about the constrained RESTORE see Appendix B), and LMS/LTS algorithms. The fits of experimental diffusion attenuations were performed on a voxel-by-voxel basis in the range of quasi-exponential attenuation behaviour limited to $b \leqslant 1000$ s mm^{-2} in the case of 30 gradient directions. The b-values in the case of six gradient-direction experiments are defined in Experimental section.

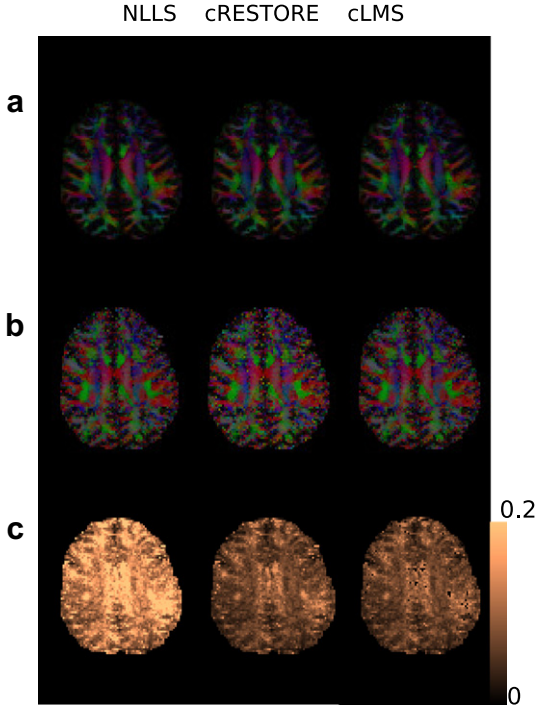

Fig. 5. Colour coded fractional anisotropy maps estimated in the biexponential fitting model (see Eq. (11)) by the NLLS, constrained RESTORE and constrained LMS algorithms. (a) Colour FA maps of the fast component. (b) Colour FA maps of the slow component. (c) Error maps of the biexponential objective function, Eq. (11).

In Fig. 5a–c we show the colour coded FA maps estimated in the biexponential fitting model by NLLS, constrained RESTORE and constrained LMS algorithms. Fig. 5a presents the colour FA maps of the fast components of the biexponential fitting estimated by NLLS, cRESTORE and cLMS algorithms. Fig. 5b presents the colour FA maps of the slow component of the biexponential fitting model estimated by the same algorithms. In Fig. 5c we plotted the error maps of the objective function for the biexponential fitting model, Eq. (11).

Fig. 6a and b presents the quantitative comparison between three methods: NLLS, constrained RESTORE and constrained LMS,

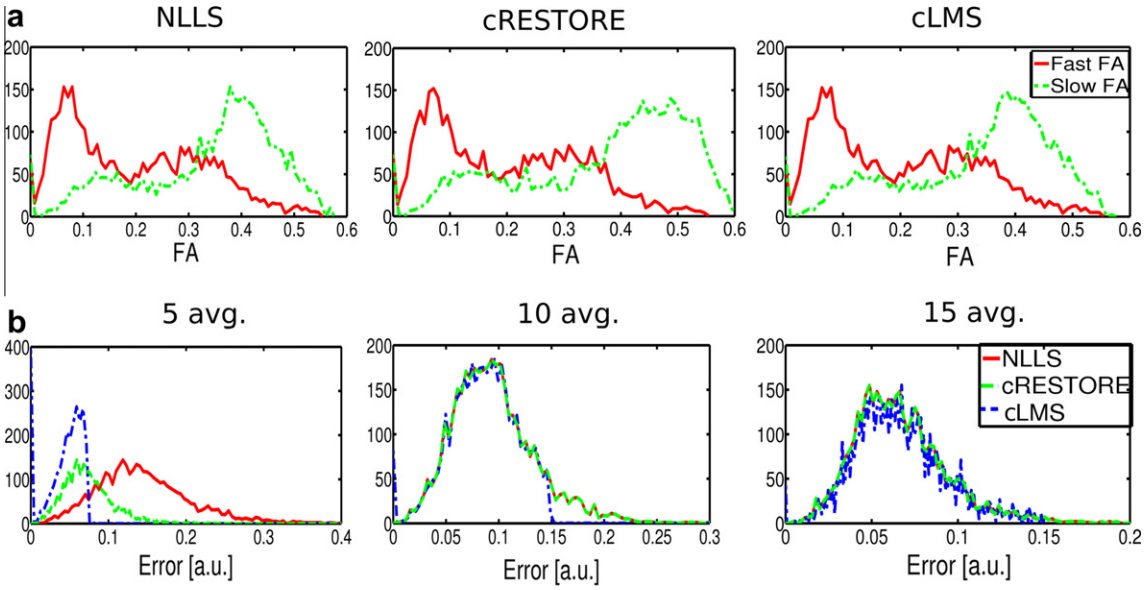

Fig. 6. The quantitative comparison of the NLLS, constrained RESTORE and constrained LMS algorithms in the biexponential fitting model (see Eq. (11)). (a) The histograms of FA maps of the fast (red) and slow (green) components of the diffusion tensor. (b) The dependence of the error-map histograms on the number of acquisitions/averages for the NLLS (red), cRESTORE (green) and cLMS (blue).

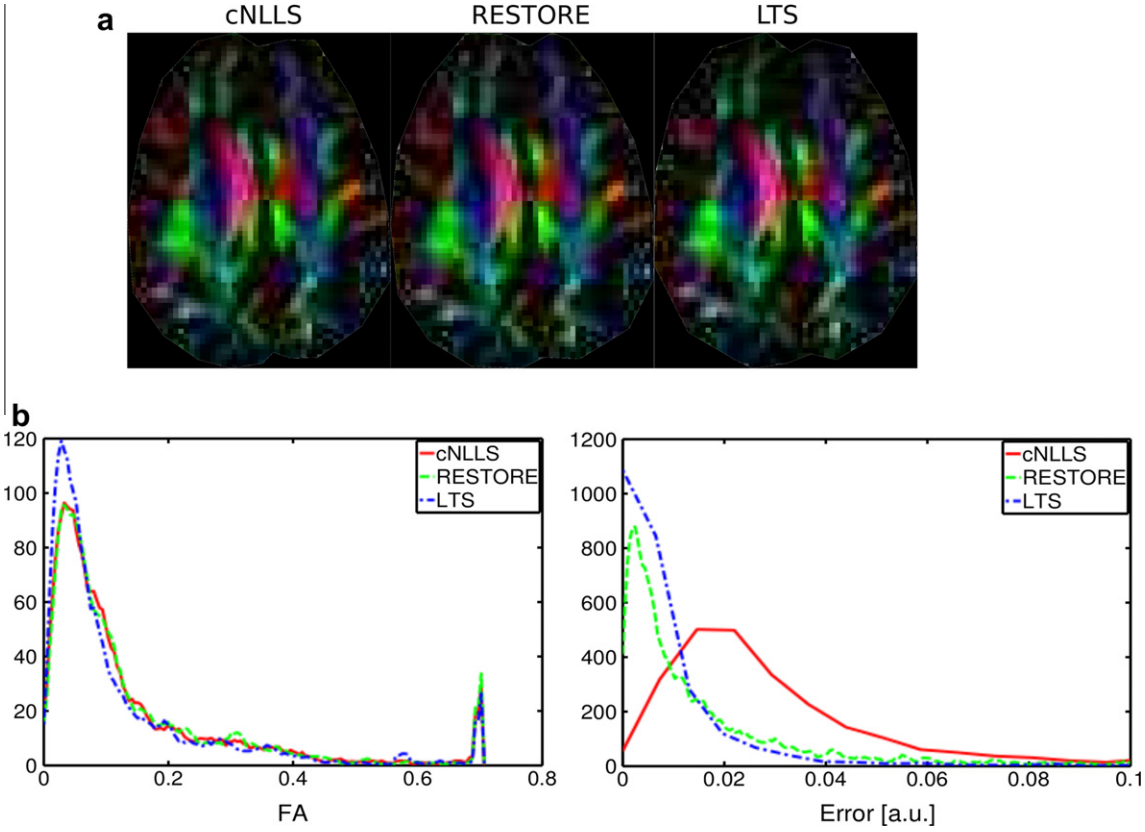

Fig. 7. The quantitative comparison of the constrained NLLS, RESTORE and LTS algorithms in the experiments using 30 gradient directions. (a) Colour FA maps obtained by the cNLLS, RESTORE and LTS algorithms. (b) FA and error-map histograms of the cNLLS (red), RESTORE (green) and LTS (blue) algorithms.

in the case of the biexponential fitting. In Fig. 6a we demonstrate the histograms obtained for the fast (red curve) and slow (green curve) FA maps. In Fig. 6b, the histograms of error maps are presented as a function of a number of acquisitions/averages.

Fig. 7a and b exhibits the difference between well known least squares and RESTORE algorithms and the developed LTS for 30 gradient-direction experiments. In Fig. 7a we presented the colour coded FA maps obtained by constrained NLLS, RESTORE and LTS algorithms. Fig. 7b shows the quantitative difference between three algorithms via the FA (left) and error (right) map histograms. Note that, although the visual difference between the histograms of FA obtained by all methods is negligible, the error-map histograms show significant differences, that is, the histogram obtained by the cNLLS algorithm is much broader than the histograms obtained by the robust approaches.

6. Discussion and conclusion

The three above examples were used to examine the differences in performance of various approaches to diffusion tensor evaluation. In particular, example I depicted in Figs. 1 and 2 demonstrates that the methods based on the robust statistics, that is, the LMS proposed in this work and the RESTORE proposed earlier in Ref. [14] are less sensitive to the presence of outliers than the standard least squares method, NLLS. This is especially obvious in the range of relatively high SNR (>10 units) in which both the LMS and RESTORE provide a satisfactory mean estimation of the original ADC value. This is in contrast to the NLLS approach which overestimates the true value by about 20%. Thus, the robust statistics methods are expected to be more advantageous whenever a suppression of the outliers is required. In particular, a stability against the outliers might be important for eliminating the outliers caused by the physiological noise [16,18,19] in *in vivo* studies. Any particular advantages of either of the two methods, the LMS and RESTORE, were not detected in this example. The main disadvantage of the above methods is, however, that they are not protective with respect to generation of the negative eigenvalues when reconstructing diffusion tensors and are very sensitive to the noise level.

This problem is usually overcome in the extended versions of both the robust and standard approaches by introducing constraints into the minimisation procedures [11]. However, a numerical solution of non-linear minimisation problems, where linear or non-linear constraints are included, becomes a complicated mathematical problem. In the case of the robust methods, it requires that the algorithm be numerically stable; the properties of the classical approaches such as precision, computational speed, simplicity, etc. are to be taken into account.

The efficiency of the least median versus the least squares method, the LMS/LTS and NLLS, respectively, and RESTORE with respect to estimation of tensor eigenvalues/eigenvectors was demonstrated for simulated data corrupted by Rician noise and outliers, Example II. Satisfactory fits were produced by the LMS/LTS and RESTORE methods in contrast to the NLLS method. In the case of multiple *b*-values, the advantages of the LMS algorithm are seen in Fig. 3a. The statistical deviations of the diffusion tensor skeletons are much smaller at SNR > 10 than they are for NLLS or RESTORE. The angular deviations of the orientation of the main eigenvector from the original diffusion tensor have a minimal values for LMS ($5^\circ \pm 3^\circ$) and RESTORE ($8^\circ \pm 5^\circ$) whereas the NLLS has a mean angle of about 30° even for SNR = 40. These simulations allow us to claim that the experiments with multiple *b*-values demands the application of the robust approaches in order to provide more precise evaluation of the main eigenvector.

In current studies, the multiple gradient directions are widely used. Excess gradient directional experiments allow one to increase the SNR and, at the same time, to minimize a variance in tensor estimations. The averaged statistical deviations of the diffusion tensor skeletons in Example II show only very small discrepancies for the RESTORE and LTS algorithms (see Figs. 4b and 5b). It is worth noting that differences between the algorithms in eigenvalues/eigenvectors at high SNR (>30) became negligible for robust approaches. However, all algorithms strongly demands on a priori noise correction, especially at high noise level (SNR < 5) (see, for example Fig. 4).

In the Example III we compared the performance of the LMS/LTS approaches in their applications to the *in vivo* experiments. Investigations of the brain microstructure demand the novel models describing the non-Gaussian diffusion in biological tissue. In particular, the biexponential model needs to apply very high *b*-values where the SNR becomes extremely low and might be corrupted by the noise or outliers. The application of the classical least squares frequently cannot provide successful results due to large error variations for the slow component of the diffusion tensors. In Fig. 5c we see that error maps related to the NLLS algorithm exhibit a few bright spots with high error values. This error variation has a decisive influence on the estimation of the slow components of the diffusion tensor (see Fig. 5b). However, we note that all methods improve their estimations with increasing SNR (see Fig. 6b).

A similar tendency can be seen in multiple gradient-direction experiments (see Fig. 7a). Visually, it is very difficult to find out the discrepancy in the colour coded FA maps. In fact, the histograms of the FA maps are almost the same (see Fig. 7b, left plot). However, the histograms of the error maps have characteristic differences. The NLLS algorithm has a wide distribution of the error while the histograms of the both robust algorithms, LTS and RESTORE, are much closer to zero axis.

The performance of the approaches examined here is summarised in Table 1. Each of the methods exhibits specific advantages and disadvantages, and should be evaluated with caution. The criterion of the positiveness of the evaluated diffusion tensors can be introduced in all methods such as NLLS, RESTORE and LMS/LTS. The main advantage of the robust approaches is a stability to the possible outliers which tend to rise up during the clinical or research scans due to the noise. In order to assess a computational speed of the applied algorithms, we estimated the averaged time required by each algorithm for the tensor evaluation in one single voxel in the Example II. Finally, the noise level has a decisive role in the diffusion tensor evaluation [8]. Our simulations demonstrate a different influence of the noise level on an accuracy of the applied approaches. The NLLS exhibits less sensitivity to the noise as opposite to the robust algorithms, RESTORE and LMS/LTS, see, for example, Figs. 3 and 4.

It is noteworthy that all regression algorithms are depended on the initial guess. This feature of the methods leads to the non-uniqueness of the obtained solutions. For the LMS or LTS algorithms the good initial guess is a critical point as well. In order to improve the evaluations of the robust approaches we suggest to use the results of the classical least squares as an input (see Appendix A). Another disadvantage of the robust approaches is a breakdown limit [12,20] which can lead to evaluation instabilities in the regression procedure. Thoughtful planning of the experiments may help to avoid this problem by supplying excessive information for the robust regression procedures (for example, enough *b*-values, additional acquisitions, etc.) [33].

The developed algorithm, LMS/LTS, exhibits a sufficient computational speed, numerical stability and precision, and it is rather insensitive to the presence of outliers. These features should allow one to implement the LMS/LTS approach as an important tool in the optimisation of data treatment from DTI protocols. One may suggest that a proper selection of the correction methods for post-processing will help to substantially increase the usefulness and validity of DTI for clinical and scientific research.

Table 1
A summary of properties of different diffusion-tensor reconstruction methods.

Feature	NLLS	RESTORE	LTS	LMS
Guaranteed positiveness	No	No*	Yes	Yes
Stability to the outliers	No	Yes	Yes	Yes
Computational speed (averaged time per voxel in s)	0.04 ± 0.04	0.51 ± 0.12	0.40 ± 0.09	0.12 ± 0.05
Stability to the noise level	High	Low	Medium	Low

* Original RESTORE [14] was realised without constraints. However, the algorithm can be easily modified with a demand of the positive definite tensors (see Appendix B).

Acknowledgement

I.I.M thanks Prof. Mia Hubert for valuable discussions and her help with LIBRA package. We thank Ezequiel A. Farrher for the assistance in DTI data acquisition and processing. Authors thank three anonymous referees for the substantial improvement of the manuscript.

Appendix A

In this appendix we provide the details of the robust algorithm based on the LMS/LTS approach. The flow diagram is presented in Fig. 8.

We start with non-linear least squares fitting of the raw data. The NLLS approach is implemented using Levenberg–Marquardt minimisation algorithm [32]. Next step is an estimation of the fitting criterion. The fitting criterion can be formulated through different definitions: the evaluated value has to be inside of the confidential interval, the fitting estimations have to satisfy to a given threshold, any kind of χ^2 analysis, etc. The best fitting criterion has to be chosen taking into account the details of the problem (biexponential model, kurtosis imaging, Q-ball, etc.) in order to avoid an erroneous treatment. If the criterion is satisfied the results are stored in tensor maps; otherwise we have to choose: what kind of the problem is to be solved? In the case of the multiple *b*-values we apply the LMS algorithm based on Eq. (10). In the case of the multiple gradient directions we apply the LTS algorithm based on Eq. (5). Since the problem of non-positive definite tensors is essential in both cases we use the Cholesky decomposition that guarantee the feasible evaluations. The realisation of the Cholesky parametrisation is done in Ref. [11].

Note that the minimisation problem of Eq. (10) in the LMS method is solved using the Levenberg–Marquardt algorithm as well. An application of the LTS algorithm has several steps. Using the NLLS estimation as an initial guess we provided the iterative steps in a following way:

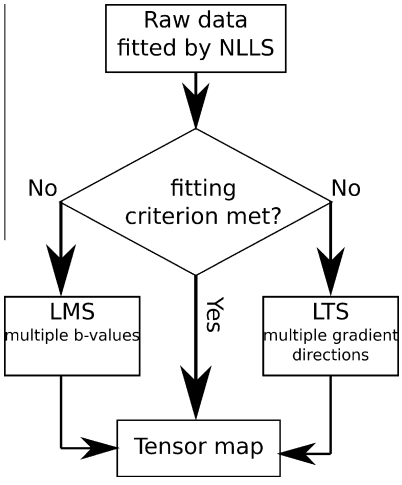

Fig. 8. Flow diagram of the LMS/LTS robust algorithm.

1. Initial guess obtained from NLLS
2. Evaluation of all residuals
3. Creation of truncated residual set using given *h*-factor from Eq. (5)
4. Application of NLLS to the truncated residuals using the Levenberg–Marquardt algorithm
5. Test fitting criterion: if the criterion is satisfied, stop the iterations, else go to step 2

The fitting criterion in the LTS algorithm was restricted by a number of iteration steps equal to, approximately, 10–20 iterations.

Moreover, we have to consider an intermediate case when determination of the area of the application of the robust statistics according to the above criteria is not obvious: does a data set under evaluation belong to the case of multiple *b*-values or to the case of multiple gradient directions? The typical example refers to diffusion kurtosis imaging. The solution might be compromised in a following way: if the number of the *b*-values is notably larger than, or comparable to, the number of the encoding gradients we suggest the LMS algorithm should be applied; in all other cases – one should use the LTS algorithm.

Appendix B

The constrained RESTORE algorithm can be obtained by replacing the non-linear least squares fitting with weightings in the diagram flow of Ref. [14] with the constrained NLLS algorithm.

References

[1] H. Johansen-Berg, T.E.J. Behrens, Diffusion MRI. From Quantitative Measurement to In vivo Neuroanatomy, Elsevier, London, Burlington, San Diego, 2009. p. 490.
[2] C. Beaulieu, The basis of the anisotropic water diffusion in the nervous system: a technical review, NMR Biomed. 15 (2002) 435–455.
[3] D. Le Bihan, The wet mind: water and functional neuroimaging, Phys. Med. Biol. 52 (2007) R57–R90.
[4] P.J. Basser, J. Mattiello, D. Le Bihan, MR diffusion tensor and imaging, Biophys. J. 66 (1994) 259–267.
[5] P.J. Basser, D. Le Bihan, J. Mattiello, Estimation of the effective self-diffusion tensor from NMR spin echo, J. Magn. Reson. B 103 (1994) 247–254.
[6] C. Pierpaoli, P.J. Basser, Toward a quantitative assessment of diffusion anisotropy, Magn. Reson. Med. 36 (1996) 893–906.
[7] Q. Dong, R.C. Welsh, T.L. Chenevert, R.C. Carlos, P. Maly-Sundgren, D.M. Gomez-Hassan, S.K. Mukherji, Clinical applications of diffusion tensor imaging, J. Magn. Reson. Imaging 19 (2004) 6–19.
[8] C.G. Koay, L.C. Chang, C. Pierpaoli, P.J. Basser, Error propagation framework for diffusion tensor imaging via diffusion tensor representations, IEEE Trans. Med. Imaging 26 (2007) 1017–1034.
[9] C.G. Koay, J.D. Carew, A.L. Alexander, P.J. Basser, M.E. Meyerand, Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging, Magn. Reson. Med. 55 (2006) 930–936.
[10] C.G. Koay, L.C. Chang, J.D. Carew, C. Pierpaoli, P.J. Basser, A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging, J. Magn. Reson. 182 (2006) 115–125.
[11] C.G. Koay, Least squares approaches to diffusion tensor estimation, in: D.K. Jones (Ed.), Diffusion MRI: Theory, Methods, and Applications, Oxford University Press, Oxford, 2011, p. 272.
[12] P.J. Rousseeuw, Tutorial to robust statistics, J. Chemom. 5 (1991) 1–20.
[13] J.F. Mangin, C. Poupon, C. Clark, D. Le Bihan, I. Bloch, Distortion correction and robust estimation for MR diffusion imaging, Med. Imaging Anal. 6 (2002) 191–198.

[14] L.C. Chang, D.K. Jones, C. Pierpaoli, RESTORE: robust estimation of tensors by outlier rejection, Magn. Reson. Med. 53 (2005) 1088–1095.
[15] S. Geman, D.E. McClure, Statistical methods for tomographic image reconstruction, Bull. Int. Stat. Inst. 52 (1987) 5–21.
[16] L. Walker, L.C. Chang, C.G. Koay, N. Sharma, L. Cohen, R. Verma, C. Pierpaoli, Effects of physiological noise in population analysis of diffusion tensor MRI data, NeuroImage 54 (2010) 1168–1177.
[17] T. Roy, Robust non-linear regression analysis, J. Chemom. 9 (1995) 451–457.
[18] B.A. Landman, P.L. Bazin, S.A. Smith, J.L. Prince, Robust estimation of spatially variable noise fields, Magn. Reson. Med. 62 (2009) 500–509.
[19] B.A. Landman, P.L. Bazin, J.L. Prince, Estimation and application of spatially variable noise fields in diffusion tensor imaging, Magn. Reson. Imaging 27 (2009) 741–751.
[20] P.J. Rousseeuw, Least median of squares regression, J. Am. Stat. Ass. 79 (1984) 871–880.
[21] L.M. Li, An algorithm for computing exact least-trimmed squares estimate of simple linear regression with constraints, Comp. Stat. Data Anal. 48 (2005) 717–734.
[22] A.W. Anderson, Theoretical analysis of the effects of noise on diffusion tensor imaging, Magn. Reson. Med. 46 (2001) 1174–1188.
[23] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge UK, 1985.
[24] J.C. Pinheiro, D.G. Bates, Unconstrained parametrizations for variance–covariance matrices, Stat. Comput. 6 (1996) 289–296.
[25] C.A. Clark, M. Hedehus, M.E. Moseley, In vivo mapping of the fast and slow diffusion tensors in human brain, Magn. Reson. Med. 47 (2002) 623–628.
[26] J.H. Jensen, J.A. Helpern, A. Ramani, H. Lu, K. Kaczynski, Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging, Magn. Reson. Med. 53 (2005) 1432–1440.
[27] F. Grinberg, E. Farrher, J. Kaffanke, A.M. Oros-Peusquens, N.J. Shah, Non-Gaussian diffusion in human brain tissue at high b-factors as examined by a combined diffusion kurtosis and biexponential diffusion tensor analysis, NeuroImage 57 (2011) 1087–1102.
[28] D.S. Tuch, Q-ball imaging, Magn. Reson. Med. 52 (2004) 1358–1372.
[29] D.S. Tuch, T.G. Reese, M.R. Wiegell, N. Makris, J.W. Belliveau, V.J. Wedeen, High angular resolution diffusion imaging reveals introvoxel white matter fiber heterogeneity, Magn. Reson. Med. 48 (2002) 577–582.
[30] B. Jeurissen, A. Leemans, D.K. Jones, J.D. Tournier, J. Sijbers, Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution, Hum. Brain Mapping 32 (2011) 461–479.
[31] S. Verboven, M. Hubert, LIBRA: a MATLAB library for robust analysis, Chemom. Intell. Lab. Syst. 75 (2005) 127–136.
[32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipies, third ed., Cambridge University Press, Cambridge, UK, 2007.
[33] L.C. Chang, L. Walker, B. Behseta, C. Pierpaoli, Informed RESTORE for removal of physiological noise artifacts in low redundancy DTI data, Proc. Int. Soc. Mag. Reson. Med. 19 (2011) 3898.