

On High Order Tensor-based Diffusivity Profile Estimation

Mohammad Alipoor¹, Irene Y. H. Gu¹, Andrew J. H. Mehnert^{1,2}, Ylva Lilja³, Daniel Nilsson³

Abstract—Diffusion weighted magnetic resonance imaging (dMRI) is used to measure, *in vivo*, the self-diffusion of water molecules in biological tissues. High order tensors (HOTs) are used to model the apparent diffusion coefficient (ADC) profile at each voxel from the dMRI data. In this paper we propose: (i) A new method for estimating HOTs from dMRI data based on weighted least squares (WLS) optimization; and (ii) A new expression for computing the fractional anisotropy from a HOT that does not suffer from singularities and spurious zeros. We also present an empirical evaluation of the proposed method relative to the two existing methods based on both synthetic and real human brain dMRI data. The results show that the proposed method yields more accurate estimation than the competing methods.

I. INTRODUCTION

Non-Gaussian diffusion models have gained wide attention among researchers because of their ability to resolve complex fiber orientations such as fiber crossing, branching or kissing. Nevertheless because of the lack of a unified framework for estimating and interpreting these new models, the original diffusion tensor imaging (DTI), based on the 2nd order tensor, remains the most common clinically used model. In the case of 2nd order tensors (3×3 matrix) their estimation and interpretation using eigen-analysis is well established. From the eigenvalue-decomposition one obtains the main diffusion direction and can compute scalars that measure the anisotropy and the relative diffusion strength in different directions; e.g. fractional anisotropy (FA) and relative anisotropy (RA). One of the more promising alternatives to DTI for modeling complex fiber architectures in the brain is the HOT model introduced by Ozarslan and Mareci [1]. The estimation of HOTs from dMRI data, and their interpretation, remain active areas of research. The intrinsic properties of the ADC profile are antipodal symmetry and positivity [2]. Symmetry dictates using even order symmetric HOTs whilst positivity is taken into consideration during estimation. Other estimation considerations include applicability to any even order HOT, robustness to noise, and computational efficiency. Although each of these is important from a theoretical point of view, in practice one can assign different weightings to each.

To date two efficient methods for HOT-based ADC profile estimation have been proposed. The first, hereinafter called method 1, is that proposed in [3] for estimating symmetric

positive definite HOTs of any order using non-negative least-squares (NNLS) optimization. However we show that this approach yields large fitting errors in the noise-free case. The second, hereinafter called method 2, is the method proposed in [2] for estimating the positive semi-definite (PSD) tensor. The method stems from an exact closed-form linear least squares (LS) solution, testing whether the solution is PSD, and if not solving a non-linear least squares optimization problem. However [2] note that there is no practical difference observed between this solution and the LS solution, i.e. the closed-form LS solution is PSD itself, for the real human brain data they used. The question remains, when can we neglect the PSD constraint? An important observation about both methods 1 and 2 is that neither accounts for the Rician noise inherent in dMRI data. Herein we present a method that does. We also present an empirical evaluation of the method relative to the two existing methods based on both synthetic and real human brain dMRI data. In [2] a framework for the interpretation of HOTs is also presented based on the concept of Z-eigenvalues [4] (a generalization of eigen-decomposition to HOTs). An expression for computing the FA from a HOT is also given. A caveat however is that although the expression reduces to the usual definition for DTI (2nd order tensor), it yields singularities for HOTs with a single Z-eigenvalue and implausible values for HOTs with several equal Z-eigenvalues. Herein we proffer a new expression to resolve this problem. The remainder of this paper is organized as follows. In the next section we introduce notation and definitions for describing HOTs and briefly introduce methods 1 and 2 for estimating HOTs. Section III proposes a new method for HOT-based ADC profile estimation that is both more robust to noise and simple to implement. A new FA definition for HOTs is also given. In section IV, experimental results and comparisons with methods 1 and 2 are presented based on synthetic and real dMRI data. We also briefly highlight advantages and drawbacks of methods 1 and 2.

II. RELATED WORK

This section briefly reviews the basics of HOT-based ADC profile estimation both for the sake of completeness and to define notation. Readers are referred to [3], [2] for more detail. The Stejskal-Tanner equation for dMRI signal attenuation is [5]:

$$-\frac{1}{b} \ln \left(\frac{S}{S_0} \right) = d(\mathbf{g}) \quad (1)$$

where $d(\mathbf{g})$ is the diffusivity function, S is the measured signal when the diffusion sensitizing gradient is applied in the direction \mathbf{g} , S_0 is the observed signal in the absence of

¹M. Alipoor, I. Y. H. Gu and A. Mehnert are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden alipoor@chalmers.se

²A. Mehnert is also with MedTech West, Sahlgrenska University Hospital, Gothenburg, Sweden

³I. Lilja and D. Nilsson are with the Institute of Neuroscience and Physiology, Sahlgrenska University Hospital, Gothenburg, Sweden

such a gradient, and b is the diffusion weighting taken to be constant over all measurements. The diffusivity function $d(\mathbf{g})$ is modeled using even order symmetric tensors as follows:

$$d(\mathbf{g}) = \sum_{i=0}^m \sum_{j=0}^{m-i} t_{ij} g_1^i g_2^j g_3^{m-i-j} \quad (2)$$

where $\mathbf{g} = [g_1 \ g_2 \ g_3]^T$ and the t_{ij} denote $n = (m+1)(m+2)/2$ distinct entries of the m -th order tensor. In [2] the diffusivity function is expressed as the inner product

$$d(\mathbf{g}) = \mathbf{t}^T \hat{\mathbf{g}} \quad (3)$$

where $\hat{\mathbf{g}} = [g_3^m \ g_2 g_3^{m-1} \ g_2^2 g_3^{m-2} \ \dots \ g_1^m]^T$ and $\mathbf{t} = [t_{00} \ t_{01} \ \dots \ t_{m0}]$. Note that both vectors \mathbf{t} and $\hat{\mathbf{g}}$ are vectors in \mathbb{R}^n and $d(\mathbf{g}, \mathbf{t}) = d(\mathbf{g})$ is used for simplification. Given measurements in $N > n$ different directions \mathbf{g}_k , uniformly distributed over the unit sphere Ω , the HOT estimation problem can then be formulated as [2]:

$$\min \sum_{k=1}^N (d_k - \hat{d}_k)^2, \quad \text{s.t. } d(\mathbf{g}) \geq 0 \quad (4)$$

where $d_k = d(\mathbf{g}_k)$ are the ADC values predicted by the model, and $\hat{d}_k = -b^{-1} \ln(S_k/S_0)$ are the measured ADC values. Let \mathbf{G} be an $N \times n$ matrix defined $\mathbf{G} = [\hat{\mathbf{g}}_1 \ \hat{\mathbf{g}}_2 \ \dots \ \hat{\mathbf{g}}_N]^T$. For relatively big N it is assumed that \mathbf{G} has row rank n . Let \mathbf{y} be a column vector of size N , whose elements are the measured ADC values; i.e. $\mathbf{y} = -\frac{1}{b} [\ln(S_1/S_0) \ \dots \ \ln(S_N/S_0)]^T$. Equation (4) can then be written

$$\min_{\mathbf{t} \in \mathbb{R}^n} (\mathbf{y} - \mathbf{G}\mathbf{t})^T (\mathbf{y} - \mathbf{G}\mathbf{t}), \quad \text{s.t. } d(\mathbf{g}) \geq 0 \quad (5)$$

In method 2 this optimization problem is solved as follows. A solution is initially sought for the problem without constraints. This turns out to be the well-known LS solution $\mathbf{t}_{\text{opt}} = \mathbf{G}^+ \mathbf{y} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{y}$. If the LS solution satisfies $d(\mathbf{g}) \geq 0, \forall \mathbf{g} \in \Omega$, then it is the final solution. If not then we have to solve another optimization problem [2]:

$$\min_{\mathbf{t} \in \mathbb{R}^n} (\mathbf{y} - \mathbf{G}\mathbf{t})^T (\mathbf{y} - \mathbf{G}\mathbf{t}), \quad \text{s.t. } \lambda_{\min}(\mathbf{T}) = 0 \quad (6)$$

where $\lambda_{\min}(\mathbf{T})$ is the smallest Z-eigenvalue of \mathbf{T} (m -th order tensor). In method 2 the optimization problem is solved by the gradient descent method and an algorithm for Z-eigen decomposition of HOTs is also proffered. Z-eigenvectors correspond to the critical points of the spherical function defined in equation (2). They represent the dominant diffusion directions. The corresponding Z-eigenvalues are used to interpret the estimated HOT; e.g. to compute the FA. In method 1 the HOT-based ADC profile estimation problem is formulated as a NNLS optimization and solved using an iterative method. Thus it does not necessarily converge to the optimal (LS) solution.

III. PROPOSED WORK

In attempting to minimize the error between the HOT model and the observed ADC values one must consider that the intensity S in magnetic resonance intensity images is corrupted by Rician distributed noise [6]. The measurements of diffusion signal magnitude can thus be modeled as [7]:

$$S(\mathbf{g}_i) = |A(\mathbf{g}_i) + c|, \quad i = 1, \dots, N \quad (7)$$

where c is Rician distributed random noise with standard deviation σ , $A(\mathbf{g}_i) = \sum_{j=1}^u \frac{S_0}{u} \exp(-b \mathbf{g}_i^T D_j \mathbf{g}_i)$ is the ideal signal (without noise) from a voxel containing u fiber bundles and D_j is the 2nd order diffusion tensor for the j -th fiber. Note that SNR is measured as the baseline signal divided by the standard deviation of noise [8] (which is $1/\sigma$ in our case). Taking the noise into account would improve estimation accuracy. Neither method 1 nor method 2 provide a mechanism to account for noise during the optimization/estimation step. This produces a large estimation error when we deal with real world problems ($SNR \leq 50$ [3], [2]). We propose a new HOT estimation method which maintains all of the advantages of methods 1 and 2 but includes a mechanism to mitigate noise. First we eliminate the PSD constraint (based on the conclusion from [2] and our experiments) and then we add an $N \times N$ weighting matrix \mathbf{W} to the cost function in (5) as follows:

$$\min_{\mathbf{t} \in \mathbb{R}^n} (\mathbf{y} - \mathbf{G}\mathbf{t})^T \mathbf{W} (\mathbf{y} - \mathbf{G}\mathbf{t}) \quad (8)$$

where \mathbf{W} is a symmetric positive definite matrix [9]. For instance it can be a diagonal matrix with positive diagonal elements. The rationale for introducing the weighting to the cost function is to reduce the contribution of those measurements deemed not to be reliable. The problem in (8) is known as weighted least squares (WLS) optimization and has the following closed-form solution:

$$\mathbf{t}_{\text{opt}} = (\mathbf{G}^T \mathbf{W} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{W} \mathbf{y} \quad (9)$$

In principle the weighting matrix is defined in terms of the statistics of the interfering noise. We propose to set \mathbf{W} as $\text{diag}(w_{11}, \dots, w_{ii}, \dots, w_{NN})$ where:

$$w_{ii} = \frac{\|\mathbf{t}\|_2}{\|\mathbf{t} - \mathbf{t}_i\|_2} \quad (10)$$

where \mathbf{t} is the LS solution considering all measured data and \mathbf{t}_i is the LS solution considering all measured data excluding the measurement in the gradient direction \mathbf{g}_i . We normalize the diagonal elements to sum to one. Exclusion of noise-affected measurements leads to a large deviation from the LS solution. Thus outliers and noisy measurements are assigned low weights. On the other hand, inliers are expected to achieve large weights. Therefore WLS is expected to outperform the LS solution (especially in the presence of noise) because it prevents over-fitting to the observed noisy signal. We note that [10] also uses WLS for 2nd order diffusion tensor estimation but point out two significant differences. Firstly we take a different approach to obtain the weighting matrix. Secondly we define a criterion (based on SNR level) to evaluate whether the PSD constraint is necessary.

A. An Alternative Expression for FA

We propose the following expression as a measure of fractional anisotropy derived from a HOT having v Z-eigenvalues:

$$FA_{\star} = \frac{\lambda_{\max}}{v\lambda_{\text{mean}}} \quad (11)$$

where λ_{\max} and λ_{mean} are the maximum and mean respectively of the Z-eigenvalues. The equation defined in [2] for computing FA, $FA_{\text{PSDT}} = \sqrt{\frac{v}{v-1} \frac{\sum_{i=1}^v (\lambda_i - \lambda_{\text{mean}})^2}{\sum_{i=1}^v \lambda_i^2}}$, can yield values approaching infinity for $v = 1$, or zero for two equally important dominant directions. In the context of 2nd order tensors, having 3 equal eigenvalues implies isotropic diffusion ($FA = 0$, as occurs for cerebrospinal fluid). However in the case of a HOT it may simply reflect a complex anisotropic environment (e.g. 2 crossing fiber bundles). Consider the set of eigenvalues $\{\alpha, \alpha + \beta\}$ where α and β are real. It can be seen that $\lim_{\beta \rightarrow 0} FA_{\star} = \lim_{\alpha \rightarrow \infty} FA_{\star} = 0.5$ and $\lim_{\beta \rightarrow 0} FA_{\text{PSDT}} = \lim_{\alpha \rightarrow \infty} FA_{\text{PSDT}} = 0$ and so our proposed definition gives more plausible FA values for HOTs. Although our definition does not reduce to the regular FA expression for the 2nd order model, Fig 3 shows that it conveys the same information.

IV. EXPERIMENTS AND RESULTS

This section describes the simulation setup, test results and evaluation of the proposed method. Further, we also compare the performance of the proposed method with the two existing state-of-the-art HOT-based diffusion profile estimation methods; i.e. methods 1 and 2. Synthetic data were generated using the model given by (7). The following set up was used: $N = 81$, $b = 1500 \text{sec}/(\text{mm}^2)$, $D_1 = \text{diag}(17, 1, 1) \times 10^{-4}(\text{mm}^2)/\text{sec}$, $D_2 = \text{diag}(1, 17, 1) \times 10^{-4}(\text{mm}^2)/\text{sec}$, $u = 2$. We used $m = 4$ and the same criterion for computing the estimation error as in [3]:

$$\text{Error} = \frac{\sum_{i=1}^N |d_{\text{true}}(\mathbf{g}_i) - d(\mathbf{g}_i)|}{\sum_{i=1}^N |d_{\text{true}}(\mathbf{g}_i)|} \quad (12)$$

First we generated four synthetic ADC profiles simulating two fibers crossing in different directions (first row of Fig. 1). Then we added noise ($\text{SNR}=12.5$) and applied the proposed method to reconstruct the ADC profile. Results (averaged over 100 iterations) show that the underlying ADC profile and main diffusion directions are reconstructed successfully (second row of Fig. 1). A quantitative comparison of methods 1 and 2 with our proposed method for ADC profile estimation on synthetic data is shown in Fig. 2. Reconstruction error is averaged over 100 realizations of Rician noise. The graph shows that our proposed (WLS) method performs the same as method 2 for high SNR measurements. However for the typical level of noise $\text{SNR} \simeq 12.5$ [8], the performance of the proposed method is significantly better than that of method 1 and slightly better than that of method 2. Fig. 3 shows that our proposed definition for FA provides a more plausible measure of anisotropy both on a real healthy human brain (first row) and realistic phantom dMRI data [11] (second row). The real data was acquired in 32 directions with a 1.5T scanner using $b = 800 \text{sec}/(\text{mm}^2)$. Voxels in which FA_{PSDT} (left column) incorrectly goes to zero or infinity are marked with white circles and triangles

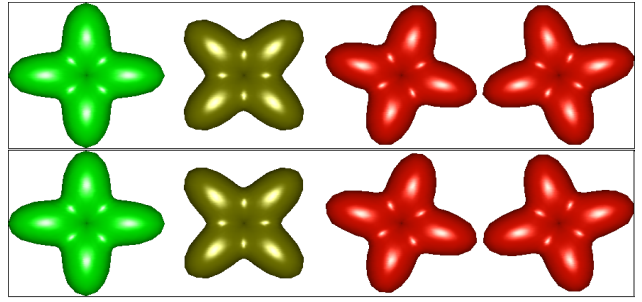


Fig. 1. WLS-based ADC profile reconstruction in four synthetic voxels: First row is the ground truth and the second row is the reconstructed ADC profile when $\text{SNR}=12.5$. The plots were produced using the same software as [12].

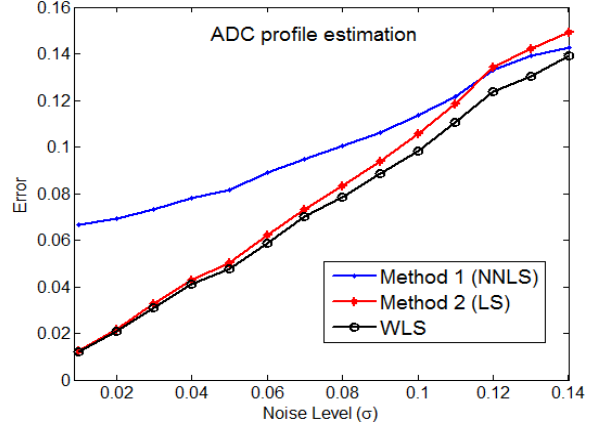


Fig. 2. Comparison of the HOT estimation error for our WLS method and methods 1 and 2.

respectively. It can be seen that the FA map generated with the new definition: (i) does not suffer from singularities or spurious zeros; (ii) has the same dynamic range as that of [2]; (iii) is as informative as that of [2] (see also Fig. 4); and (iv) can compute more feasible FA values for HOTs with an arbitrary number of Z-eigenvalues (see also Fig. 5).

A. Discussion

In contrast to the results reported in Fig. 2 of [2], our results show that negative Z-eigenvalues are most likely to happen as SNR reduces (LS solution is not PSD anymore). This means that if we acquire dMRI data with high enough SNR, the PSD constraint is guaranteed. According to our experiments, for synthetic data of $\text{SNR} \geq 8.3$, the LS solution is PSD. This can explain the observation in [2] that no negative eigenvalues were observed in their experiments on a human brain dataset. Both [2], [3] evaluate the performance of their respective methods based on signal reconstruction error (defined in (12)). However, the major direction of diffusion obtained from a Z-eigen-decomposition of the estimated tensor could be another criterion. An interesting observation, not noted in [2], is that the additional effort associated with obtaining the PSDT solution does not significantly change either the accuracy of signal reconstruction or the estimated major diffusion direction (see Fig. 1 and tables 1 and 2 of [2]). This is confirmed in our experiments and forms the rationale for elimination of the PSD constraint.

The proposed Z-eigenvalue computation algorithm in [2] can be used for Z-eigen decomposition of HOTs regardless



Fig. 3. Spurious FA changes are resolved using our FA definition. FA_{PSDT} (left column) and FA_* (right column) maps are shown for an axial slice of a human brain (first row) and realistic phantom data (second row). The first row shows that FA_* carries the same information as FA_{PSDT} . The second row shows that FA_* is more informative than FA_{PSDT} . Voxels for which the regular FA incorrectly goes to zero or infinity are marked with white circles and triangles, respectively.

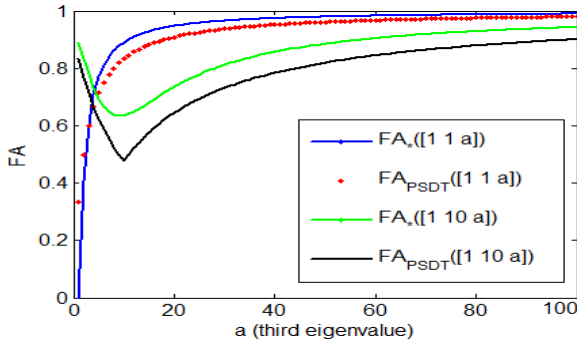


Fig. 4. FA changes for two sets of eigenvalues: $\{1, 1, a\}$ and $\{1, 10, a\}$. FA_* carries the same information as FA_{PSDT} .

of the applied estimation method. However it has several shortcomings. In some cases it may lead to an infinite number of eigenvalues. For example when all entries of a 4-th order tensor are zero except $t_{31} = t_{40} = 1$.

In method 1 the HOT-based ADC profile estimation problem, formulated as a NNLS optimization, is solved using an iterative method. Consequently it does not necessarily converge to the LS solution. Our experiments show that it produces a large error in low-noise cases (see Fig. 2). In addition it does not address the issue of HOT interpretation.

V. CONCLUSIONS

A novel WLS-based method for voxel-wise HOT-based ADC profile estimation was presented and empirically evaluated relative to the two existing state-of-the-art methods of [3] and [2] (which utilize NNLS and LS optimization respectively). The results show that the proposed method yields more accurate estimation. The superior performance stems from the fact that in contrast to the existing methods it accounts for the Rician distributed noise inherent in dMRI

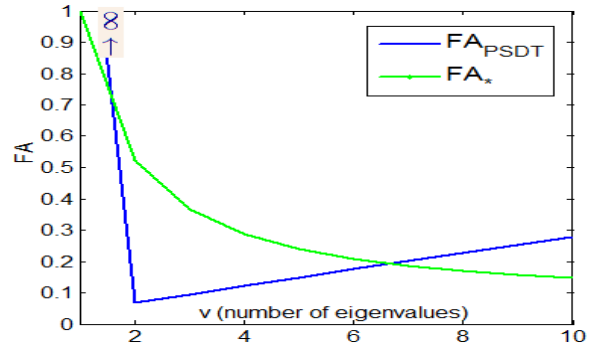


Fig. 5. FA changes with respect to v for a tensor having v , almost equal, Z-eigenvalues $\{1, 1.1, \dots, 1.1\lambda_{v-1}\}$. The FA value should approach zero as $v \rightarrow \infty$. This is the case for FA_* but not FA_{PSDT} .

data. Also the results show that the NNLS method does not converge to the optimal solution for low-noise cases (leads to large estimation error). In contrast our proposed framework is able to mitigate noise but still converge to the optimal solution in the noise-free case. Finally the results show that negative Z-eigenvalues are more likely to occur in low SNR cases. Consequently the elimination of the PSD constraint does not adversely affect HOT estimation accuracy for high SNR measurements. Several shortcomings of the Z-eigen analysis of HOTs were noted and remain open research problems. Finally a new FA definition for HOTs was presented that does not suffer from singularities or spurious zeros while having the same information content as generalized conventional map.

REFERENCES

- [1] E. Ozarslan and T. H. Mareci, "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging," *Magn Reson Med*, vol. 50, no. 5, pp. 955–965, 2003.
- [2] L. Qi, G. Yu, and E. Wu, "Higher order positive semidefinite diffusion tensor imaging," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 416–433, 2010.
- [3] A. Barmpoutis and B. C. Vemuri, "A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints," *Proc IEEE Int Symp Biomed Imaging*, pp. 1385–1388, 2010.
- [4] L. Qi, "The spectral theory of tensors (rough version)," *arXiv preprint arXiv:1201.3424*, 2012.
- [5] E. Stejskal and J. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *Journal of Chemical Physics*, vol. 42, no. 1, pp. 288–292, Jan. 1965.
- [6] H. Gudbjartsson and S. Patz, "The rician distribution of noisy MRI data," *Magn Reson Med*, vol. 34, no. 6, pp. 910–914, Dec. 1995.
- [7] K. Jansons and D. Alexander, "Persistent angular structure: new insights from diffusion magnetic resonance imaging data," *Inverse Problems*, vol. 19, pp. 1031–1046, 2003.
- [8] F. J. et. al, "Detection of crossing white matter fibers with high-order tensors and rank-k decompositions," in *Proc 22nd Int Conf Information processing in medical imaging*, ser. IPMI'11, 2011, pp. 538–549.
- [9] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, ser. Prentice Hall Signal Processing Series. PTR Prentice-Hall, 1993.
- [10] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo," *J. Magn. Reson.*, vol. Series B, 103, pp. 247–254, 1994.
- [11] P. F. et. al, "Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom," *NeuroImage*, vol. 56, no. 1, pp. 220–34, May 2011.
- [12] A. B. et. al, "Regularized positive-definite fourth-order tensor field estimation from DW-MRI," *NeuroImage*, vol. 45, no. 1 sup.1, pp. 153–162, March 2009.