SIAM J. IMAGING SCIENCES
Vol. 6, No. 3, pp. 1531–1552

Positive Semidefinite Generalized Diffusion Tensor Imaging via Quadratic Semidefinite Programming*

Yannan Chen†, Yuhong Dai‡, Deren Han§, and Wenyu Sun¶

Abstract. The positive definiteness of a diffusion tensor is important in magnetic resonance imaging because it reflects the phenomenon of water molecular diffusion in complicated biological tissue environments. To preserve this property, we represent it as an explicit positive semidefinite (PSD) matrix constraint and some linear matrix equalities. The objective function is the regularized linear least squares fitting for the log-linearized Stejskal–Tanner equation. The regularization term is the heuristic nuclear norm of the PSD matrix, since we expect it to be of low rank. In this way, we establish a convex quadratic semidefinite programming (SDP) model, whose global solution exists. The optimal solution could be solved by three efficient methods. While there are two state-of-the-art solvers—SDPT3 and QSDP—for the primal problem, we design a new augmented Lagrangian based alternating direction method (ADM) for the dual problem. Sensitivity analyses on the coefficients of the optimal diffusion tensor and the optimal objective function value with respect to noise-corrupted signals are presented. Experiments on synthetic data with multiple fibers show that the new method is robust to the Rician noise and outperforms several existing methods. Furthermore, when the fiber orientation distribution function is considered, the new method is competitive with the Q-ball imaging. Using the human brain data, we illustrate that the new method could capture the crossing of three nervous fiber bundles. Additionally, the new method generates positive definite generalized diffusion tensors in all voxels, while the unconstrained least squares fitting fails. Finally, we confirm that the ADM solver is more efficient than SDPT3 and QSDP for this special problem.

Key words. alternating direction method, generalized diffusion tensor, higher order tensor, magnetic resonance imaging, positive semidefinite Cartesian tensor, QSDP, SDPT3, semidefinite programming

AMS subject classifications. 58J65, 65H17, 65K05

DOI. 10.1137/110843526

1. Introduction.

Diffusion weighted magnetic resonance imaging (MRI) is a noninvasive medical tool for exploring some clinical information of biological tissues in vivo. This information includes tissue components, physical properties of tissue constituents, and tissue

*Received by the editors August 5, 2011; accepted for publication (in revised form) May 14, 2013; published electronically August 6, 2013. This work was supported by The National Natural Science Foundation of China 11071122, 11171159, and the Graduate Student Research and Innovation Program of Jiangsu Province of China CXZZ12_0384.

http://www.siam.org/journals/siims/6-3/84352.html

†School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing 210023, China, and College of Science, Nanjing Forestry University, Nanjing 210037, China (ynchen@njfu.edu.cn).

‡Institute of Computational Mathematics, Chinese Academy of Sciences (CAS), Beijing 100190, China (dyh@lsec.cc.ac.cn).

§Corresponding author. School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing 210023, China (handeren@njnu.edu.cn).

¶School of Mathematical Sciences, Key Laboratory for NSLSCS of Jiangsu Province, Nanjing Normal University, Nanjing 210023, China (wysun@njnu.edu.cn).

microstructure and its architectural organization. To obtain this information, we need to measure and quantize the diffuse behavior of water molecules, which is sensitive to the restrictive tissular environment.

Given a pulsed magnetic field gradient sequence with the gradient orientation $\mathbf{g}$ and a diffusion weighted b-value b, the Stejskal–Tanner equation for signal attenuation [28] is as follows:

$$S = S_0 \exp(-b\, D), \tag{1.1}$$

where S is a measured signal, S_0 is the zero-gradient signal, and D is the diffusion coefficient. After taking all gradient orientations $\mathbf{g}$ into account, we use the apparent diffusion coefficient (ADC) function to replace the diffusion coefficient. Here, the ADC function is approximately

$$D(\mathbf{g}) := \sum_{i_1=1}^{3} \sum_{i_2=1}^{3} \cdots \sum_{i_R=1}^{3} D_{i_1 i_2 \cdots i_R} g_{i_1} g_{i_2} \cdots g_{i_R}, \tag{1.2}$$

where $\mathbf{D}$ is an Rth order symmetric Cartesian tensor, $\mathbf{g} = (g_1, g_2, g_3)^\top$ is a unit vector, and R is an even number. Particularly, the widely used diffusion tensor imaging (DTI) uses a second order diffusion tensor (matrix, $R = 2$) [6, 5, 7]. That is to say, the water molecules' diffusion is assumed to have a Gaussian distribution in every local tissular environment. However, if the local environment of biological tissue is complex, such as crossing fiber configuration, the assumption of Gaussian diffusion is unsatisfactory; e.g., see Figure 1. Even if the two fibers meet at right angles, the orientations of these two fibers are unavailable. To deal with this difficulty, Özarslan and Mareci [19] propose using a higher order ($R \geq 4$) generalized diffusion tensor, which can better approximate the diffusivity profiles of the local tissular geometry.

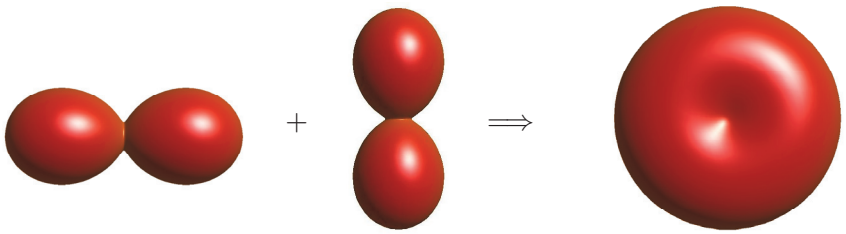

Figure 1. *DTI could capture the configuration of a single fiber, whose orientation corresponds to the direction of the maximum of ADC. See* (a) *and* (b). *However, in the presence of two-fiber crossing* (c), *the profile of ADC looks like a "doughnut" and a "pizza." So the orientations of crossing fibers are unavailable.*

Due to the physical inherence of the diffusion process, the ADC function should be positive in all gradient orientations. However, many numerical algorithms employ the following slightly relaxed form in practice:

$$D(\mathbf{g}) \geq 0 \qquad \forall\, \mathbf{g}. \tag{1.3}$$

If (1.3) holds, we call the (generalized) diffusion tensor $\mathbf{D}$ PSD. Many experts preserve the PSD constraint (1.3) by various decomposition methods. In DTI, the diffusion tensor is a PSD matrix. Wang et al. [36, 37] propose using the Cholesky factorization.

When a fourth order generalized diffusion tensor is used, the ADC function is a ternary quartic, which could be rewritten as

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} D_{ijkl}g_ig_jg_kg_l = u_2^\top \Upsilon_4 u_2,$$

where $u_2 := (g_1^2, g_2^2, g_3^2, g_1g_2, g_1g_3, g_2g_3)^\top$ is a basis of quadratic forms, and $\Upsilon_4 \in \mathcal{R}^{6\times 6}$ is symmetric and is called a Gram matrix. Powers and Reznick [22] point out that a Gram matrix Υ_4 has the following general formulation:

$$\begin{pmatrix}
D_{1111} & D_{1122}-\alpha_1 & D_{1133}-\alpha_2 & 2D_{1112} & 2D_{1113} & 2D_{1123}-\alpha_4 \\
D_{1122}-\alpha_1 & D_{2222} & D_{2233}-\alpha_3 & 2D_{1222} & 2D_{1223}-\alpha_5 & 2D_{2223} \\
D_{1133}-\alpha_2 & D_{2233}-\alpha_3 & D_{3333} & 2D_{1233}-\alpha_6 & 2D_{1333} & 2D_{2333} \\
2D_{1112} & 2D_{1222} & 2D_{1233}-\alpha_6 & 4D_{1122}+2\alpha_1 & 4D_{1123}+\alpha_4 & 4D_{1223}+\alpha_5 \\
2D_{1113} & 2D_{1223}-\alpha_5 & 2D_{1333} & 4D_{1123}+\alpha_4 & 4D_{1133}+2\alpha_2 & 4D_{1233}+\alpha_6 \\
2D_{1123}-\alpha_4 & 2D_{2223} & 2D_{2333} & 4D_{1223}+\alpha_5 & 4D_{1233}+\alpha_6 & 4D_{2233}+2\alpha_3
\end{pmatrix}.$$

There are 21 independent parameters: 15 are for the fourth order generalized diffusion tensor, and $\alpha_1, \alpha_2, \ldots, \alpha_6$ are free. So a fourth order generalized diffusion tensor may correspond to a lot of Gram matrices. On these Gram matrices, Hilbert established the following famous theorem in 1888 [27].

Theorem 1.1. *A fourth order generalized diffusion tensor is PSD if and only if there exists a corresponding* 6×6 *PSD Gram matrix with rank three.*

Based on Hilbert's theorem, two decomposition methods appear. Barmpoutis and colleagues [3, 2] propose decomposing a PSD Gram matrix as $\Upsilon_4 = C^\top C$, where $C \in \mathcal{R}^{3\times 6}$ is an upper triangular matrix. So there are 15 independent entries to be determined. Moreover, Barmpoutis et al. [2] give a regularized method for producing a positive definite generalized diffusion tensor. Ghosh, Deriche, and Moakher [14] give another decomposition. Let $V := \mathrm{diag}(u_2u_2^\top, u_2u_2^\top, u_2u_2^\top) \in \mathcal{R}^{18\times 18}$, let q_1, q_2, q_3 be the columns of $C^\top$, and let $x := (q_1^\top, q_2^\top, q_3^\top)^\top \in \mathcal{R}^{18}$. Then $u_2^\top C^\top C u_2 = x^\top V x$. So this method needs to determine 18 independent entries. Hereafter, the nonlinear least squares fitting for the Stejskal–Tanner equation is established and then solved by some well-understood nonlinear optimization algorithms. However, these unconstrained optimization problems are nonconvex and may suffer from local optimal solutions.

In the context of general higher order cases, i.e., $R \geq 6$, checking whether a given ADC function is nonnegative is an NP-hard problem. Certainly, if the ADC function can be represented as a sum of squares (SOS) of several $R/2$th order homogeneous polynomials, the ADC function is obviously nonnegative [21]. However, its converse assertion does not hold when $R \geq 6$ [27]. The equivalence of the nonnegativity of the ADC function and its SOS representation is valid only in the cases of $R = 2$ and $R = 4$. Fortunately, using Lasserre's results [18], we can see that a nonnegative ADC function could be closely approximated by an SOS polynomial.

Barmpoutis and Vemuri [4] construct an SOS polynomial $\Psi(\mathbf{g}; \mathbf{D}) = \sum_j \theta_j [p(\mathbf{g}; c^{(j)})]^2$ to approximate the ADC function, where θ_j is a nonnegative weighted factor and $p(\mathbf{g}; c)$ is

a $R/2$th order homogeneous polynomial with c as its coefficient. The coefficient $c \in \mathcal{R}^Q$ is theoretically arbitrary, where $Q := (\frac{R/2+2}{2})$. However, they restrict $p(\mathbf{g}; c) := (\vartheta_1 g_1 + \vartheta_2 g_2 + \vartheta_3 g_3)^{R/2}$ with $\vartheta_1^2 + \vartheta_2^2 + \vartheta_3^2 = 1$ and take finite samples of the triple $(\vartheta_1, \vartheta_2, \vartheta_3)$ on the unit spherical surface. In this way, they preset some $R/2$th order homogeneous polynomials. The remainder of the nonnegative weighted factors θ_j are determined by the nonnegative linear least squares fitting for the linearized Stejskal–Tanner equation. The advantage of their method is that the model is convex. But some unnecessary noises in the special formulation of homogeneous polynomials and in the samples on the unit spherical surface are introduced.

Qi, Yu, and Wu [24] give a novel method that explains the PSD constraint (1.3) as preserving the minimal Z-eigenvalue [23] of the higher order generalized diffusion tensor nonnegativity. The minimal Z-eigenvalue of a generalized diffusion tensor is the minimal value of the corresponding ADC function. Due to the gradient orientation $\mathbf{g} \in \mathcal{R}^3$ belonging to the unit spherical surface, the minimal Z-eigenvalue is computable. Then, they establish a convex nonsmooth optimization model. Computationally, if the generalized diffusion tensor produced by the unconstrained least squares method happens to be PSD, their method does nothing.

The contributions of this paper are as follows.

1. We propose a convex optimization model with simple constraints to replace the decomposition methods that lead to nonconvex models. Here, we approximate the PSD constraint (1.3) via the SOS representation in a unified framework. An implicit PSD matrix is kept to express the coefficients of the SOS homogeneous polynomial, and a linear map links the coefficients between this PSD matrix and the generalized diffusion tensor. Then, the feasible domain, which is an image of the convex PSD matrix cone under a linear map, is convex.
2. Instead of restricting the rank of the Gram matrix to be three in the special case of $R = 4$ [3, 2, 14], we propose taking a low-rank PSD matrix, which is flexible. A novel nuclear norm minimization technique is employed to control the rank of the PSD matrix. Hence, the objective function here is the nuclear norm regularized least squares fitting for the log-linearized Stejskal–Tanner equation. Thus, it is a convex function. Hence, the new model is a convex constraint optimization problem, whose global solution exists.
3. We compare three solvers for the new model. Since the nuclear norm of a PSD matrix is its trace, we can rewrite the new model as a quadratic SDP problem. Then, two state-of-the-art solvers SDPT3 [32] and QSDP [30] are employable. Meanwhile, from the viewpoint of solving the equivalent dual problem, we establish an efficient augmented Lagrangian based ADM, whose subproblems have explicit solutions. This is the main advantage of the proposed algorithm over the state-of-the-art solvers, especially for the large-scale problems considered here.
4. Sensitivity analyses on the noise-corrupted signals are presented. Since Rician noise often exists in diffusion weighted MRI data, it is meaningful to analyze its effect on the diffusion tensor and ADC value. All the noise-corrupted signals in the new model are collected in the vector f. By our analysis, we find that, under a linear mapping, the operator of the optimal independent coefficients vector of the (generalized) diffusion tensor is monotone with respect to f. Furthermore, the optimal objective value function defined in f is a convex function and is bounded up by a quadratic function.

5. We report some preliminary experimental results to illustrate the validation of the convex quadratic SDP model. For the synthetic data, the new method provides a more accurate approximate solution than several existing methods when Rician noise arises. Then, we illustrate that the fiber orientations estimated by the new method are correct under different separation angles and different volume fractions. This result is competitive with the Q-ball imaging. For the human brain data, a fiber reconstruction of a region of multifiber crossing is established. We detect three different fiber orientations, which correspond to the corpus callosum, the corticospinal tracts, and the superior longitudinal fasciculus. Moreover, a comparison of SDPT3, QSDP, and ADM solvers are given for this special quadratic SDP problem.

The outline of this paper is as follows. We establish the quadratic SDP model for the generalized DTI in section 2. To solve this model, two state-of-the-art solvers SDPT3 and QSDP are introduced in section 3, and the ADM for the dual problem is proposed in section 4. Sensitivity analyses on noise effects is presented in section 5. Numerical experiments illustrating the validation of the new method are performed in section 6. The conclusion and remarks are drawn in section 7.

2. Convex quadratic SDP model. We first rewrite the ADC function (1.2) as a linear form,

$$D(\mathbf{g}) = \phi^\top w, \tag{2.1}$$

where $\phi := (g_1^R, g_1^{R-1}g_2, \ldots, g_3^R)^\top$ is a basis of Rth order homogeneous polynomials, the order R is even, $w := (d_{R,0,0}, d_{R-1,1,0}, \ldots, d_{0,0,R})^\top \in \mathcal{R}^P$ is the independent coefficients vector of the ADC function in this basis, and $P := (\frac{R+2}{2})$. As a consequence,

$$d_{i,j,R-i-j} := \frac{R!}{i!j!(R-i-j)!} D_{\underbrace{1\cdots1}_{i}\underbrace{2\cdots2}_{j}\underbrace{3\cdots3}_{R-i-j}}.$$

We then use an SOS polynomial to approximate the PSD constraint (1.3). Let $u := (g_1^{R/2}, g_1^{R/2-1}g_2, \ldots, g_3^{R/2})^\top$ be a basis of $R/2$th order homogeneous polynomials, which contains $Q = (\frac{R/2+2}{2})$ components. Under this basis, every PSD matrix $X \in \mathcal{S}^Q$ corresponds to an SOS homogeneous polynomial $u^\top Xu$, where $\mathcal{S}^Q$ is the $Q \times Q$ symmetric matrix space. Considering the sufficient and almost necessary condition, i.e., SOS, for the PSD constraint (1.3), we take

$$D(\mathbf{g}) = u^\top Xu.$$

Here, we say "almost necessary" for two reasons. First, SOS is necessary if $R = 2$ and $R = 4$ by Hilbert's theorem. The matrices X in proper bases are the diffusion tensor $\mathbf{D}$ and a Gram matrix Υ_4 if $R = 2$ and $R = 4$, respectively. Second, when $R \geq 6$, every nonnegative ADC function could be approximated by an SOS homogeneous polynomial as closely as possible. We rewrite $u^\top Xu = \langle uu^\top, X\rangle$, where the matrix inner product $\langle A, B\rangle := \text{trace}(A^\top B)$. Comparing the entries of the matrix $uu^\top$ and of the basis ϕ, we get the following interesting lemma.

Lemma 2.1. *Given bases ϕ and u of the Rth and $R/2$th order homogeneous polynomials, respectively, there exists a linear map $\mathcal{A} : \mathcal{S}^Q \mapsto \mathcal{R}^P$ such that*

$$\phi^\top \mathcal{A}(X) = \langle \mathcal{A}^*\phi, X\rangle = \langle uu^\top, X\rangle,$$

where

$$\mathcal{A}(X) := \begin{pmatrix} \langle A_1, X\rangle \\ \langle A_2, X\rangle \\ \vdots \\ \langle A_P, X\rangle \end{pmatrix}$$

for $X, A_1, A_2, \ldots, A_P \in \mathcal{S}^Q$, *and* $\mathcal{A}^*\phi := \sum_{p=1}^P A_p\phi_p$.

Furthermore, $\mathcal{A}\mathcal{A}^*$ *is a* $P \times P$ *positive definite diagonal matrix with diagonal entries*

$$[\mathcal{A}\mathcal{A}^*]_{pp} = \textit{the number of nonzeros of } A_p.$$

Proof. Since all the entries of ϕ and $uu^\top$ are Rth order homogeneous polynomials, we define the matrix A_p as the location of the entry ϕ_p in the matrix $uu^\top$, i.e.,

$$[A_p]_{ij} = \begin{cases} 1 & \text{if } [uu^\top]_{ij} = \phi_p, \\ 0 & \text{otherwise} \end{cases}$$

for $p = 1, \ldots, P$. Then, $uu^\top = \sum_{p=1}^P A_p\phi_p = \mathcal{A}^*\phi$, and $\langle uu^\top, X\rangle = \sum_{p=1}^P \phi_p\langle A_p, X\rangle = \phi^\top\mathcal{A}(X)$.

Moreover,

$$\langle A_p, A_q\rangle = \begin{cases} \text{the number of nonzeros of } A_p & \text{if } p = q, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, $\mathcal{A}\mathcal{A}^*$ is a positive definite diagonal matrix. ∎

By this lemma, if the ADC function is an SOS homogeneous polynomial, there exists a symmetric PSD matrix X such that

$$D(\mathbf{g}) = u^\top Xu = \langle uu^\top, X\rangle = \langle \mathcal{A}^*\phi, X\rangle = \phi^\top\mathcal{A}(X).$$

From its linear form (2.1) and the arbitrariness of the gradient orientation $\mathbf{g}$ as well as the basis ϕ, the coefficients vector w of the generalized diffusion tensor is forced to be

$$\mathcal{A}(X) = w \quad \text{and} \quad X \succeq 0, \tag{2.2}$$

where $X \succeq 0$ means that X is a symmetric and PSD matrix. The equality constraint here establishes the scalar relationship between the independent coefficients of the generalized diffusion tensor and the PSD matrix, which preserves the positive semidefiniteness of the generalized diffusion tensor. Hence, this SDP constraint is perfectly equivalent to the SOS condition.

We argue that the PSD matrix X in (2.2) should be of low rank. On the one hand, as a special case, Hilbert's theorem says that the 6×6 PSD Gram matrix can take at most three ranks when $R = 4$. On the other hand, Jayachandra et al. [17] show that a fourth order generalized diffusion tensor could describe six different fiber orientations. In practice, the number of different fiber orientations in one voxel is usually less than six, which implies the degeneration of the rank. Furthermore, a higher order ADC function, which is a homogeneous

polynomial, is sensitive to noise. Hence, a low-rank restriction makes the PSD matrix stable. Since the rank of a matrix is a discrete nonconvex function [25], Fazel [13] introduces the heuristic nuclear norm, which is the best convex lower approximation of the rank function over the matrix's set $\{X \,|\, \|X\|_2 \le 1\}$. Here, $\|\cdot\|_2$ is the spectral norm that is the maximal singular value of a matrix, while the nuclear norm $\|\cdot\|_*$ is defined as the sum of all the singular values. In fact, the nuclear norm of a PSD matrix is indeed its trace, i.e.,

$$\|X\|_* = \langle X, I\rangle \qquad \text{if } X \succeq 0, \tag{2.3}$$

where I is an identity matrix. This is one part of the objective function.

The other part of the objective function is the least squares fitting for the log-linearized Stejskal–Tanner equation (1.1) with the ADC function (2.1):

$$\phi(\mathbf{g})^\top w + \frac{1}{b}\ln\left(\frac{S}{S_0}\right) = 0.$$

Given a set of magnetic field gradient orientations $\mathbf{g}^{(n)}$ and corresponding signals $S^{(n)}$ for $n = 1, 2, \ldots, N$, where $N \ge P$, define

$$\Phi := \begin{pmatrix} \phi(\mathbf{g}^{(1)})^\top \\ \phi(\mathbf{g}^{(2)})^\top \\ \vdots \\ \phi(\mathbf{g}^{(N)})^\top \end{pmatrix} \quad \text{and} \quad f := \begin{pmatrix} \frac{1}{b}\ln\left(\frac{S^{(1)}}{S_0}\right) \\ \frac{1}{b}\ln\left(\frac{S^{(2)}}{S_0}\right) \\ \vdots \\ \frac{1}{b}\ln\left(\frac{S^{(N)}}{S_0}\right) \end{pmatrix}. \tag{2.4}$$

The matrix $\Phi \in \mathcal{R}^{N\times P}$ is of column full rank due to the elaborate selection of the gradient orientations. After taking the SDP constraint (2.2) and the nuclear norm regularization (2.3) into account, the least squares fitting for the log-linearized Stejskal–Tanner equation is the convex quadratic SDP problem

$$\begin{cases} \min\limits_{w,X} & \dfrac{1}{2}\|\Phi w + f\|^2 + \mu\langle X, I\rangle \\ \text{s.t.} & \mathcal{A}(X) = w,\ X \succeq 0, \end{cases} \tag{2.5}$$

where μ is the weighted regularization parameter. This problem is well defined, and its global solution exists.

3. State-of-the-art solvers: SDPT3 and QSDP. For convex quadratic SDP problems, Toh, Tütüncü, and Todd developed two state-of-the-art solvers: SDPT3[1] [31, 35, 32] and QSDP[2] [29, 33, 30], which are infeasible primal-dual predictor-corrector path-following methods.

To use the SDPT3 solver, we model the quadratic term $\|\Phi w + f\|^2$ in the objective function of problem (2.5) by a linear equality constraint $h = \Phi w + f$, a second order cone constraint

[1]The MATLAB codes of SDPT3 are available from http://www.math.nus.edu.sg/~mattohkc/sdpt3.html.
[2]The MATLAB codes of QSDP are available from http://www.math.nus.edu.sg/~mattohkc/qsdp.html.

$(t;h) \in \mathcal{K}_{N+1} := \{(t,h) \,|\, t \geq \|h\|\}$, and an inequality $s \geq t^2$. Consequently, the objective function is converted to a linear function $\frac{1}{2}s + \langle X, \mu I\rangle$. Then, we find that the inequality $s \geq t^2$ is equal to $(\frac{s+1}{2})^2 \geq (\frac{s-1}{2})^2 + t^2$. Hence, the equivalent semidefinite-quadratic-linear programming of the model (2.5) is

$$\begin{cases} \min\limits_{X} & \dfrac{1}{2}s + \langle X, \mu I\rangle \\ \text{s.t.} & \mathcal{A}(X) = w,\ X \succeq 0, \\ & \Phi w + f = h,\ (t;h) \in \mathcal{K}_{N+1}, \\ & 2x = s + 1, \\ & 2y = s - 1,\ s \geq 0, \\ & z = t,\ (x;(y;z)) \in \mathcal{K}_3. \end{cases}$$

To use the QSDP solver, we substitute the equality constraint $\mathcal{A}(X) = w$ into the quadratic objective function and ignore the constant $\|f\|^2/2$. The problem (2.5) can be rewritten as

$$\begin{cases} \min\limits_{X} & \dfrac{1}{2}\langle X, \mathcal{A}^*\Phi^\top\Phi\mathcal{A}(X)\rangle + \langle X, \mathcal{A}^*\Phi^\top f + \mu I\rangle \\ \text{s.t.} & X \succeq 0. \end{cases}$$

Once the optimal matrix X has been obtained, the optimal vector w is computed by $w = \mathcal{A}(X)$.

4. A new solver: ADM. Corresponding to the convex quadratic SDP problem (2.5), which is called the primal problem, there is a dual problem

$$\begin{cases} \max\limits_{\lambda, Y} & -\dfrac{1}{2}\big(\Phi^\top f + \lambda\big)^\top\big(\Phi^\top\Phi\big)^{-1}\big(\Phi^\top f + \lambda\big) + \dfrac{1}{2}\|f\|^2 \\ \text{s.t.} & \mathcal{A}^*\lambda + Y = \mu I, \quad Y \succeq 0. \end{cases} \tag{4.1}$$

Since the primal problem (2.5) is convex and Slater's condition holds, i.e., there exist a positive definite matrix $X \in \mathcal{S}^Q$ and a vector $w = \mathcal{A}(X)$, strong duality holds [9]. Hence, the primal-dual solution of (2.5) can be obtained by solving the dual problem (4.1).

The augmented Lagrangian based ADM [16, 15, 38, 8] is a powerful tool for the equality constrained convex programming with separable variables if the involved subproblems are easy. Here, all of the subproblems have explicit solutions. We define the augmented Lagrangian function of the dual problem (4.1) as

$$\begin{aligned} \mathcal{L}(\lambda, Y, X) := {} & \frac{1}{2}(\Phi^\top f + \lambda)^\top(\Phi^\top\Phi)^{-1}(\Phi^\top f + \lambda) - \frac{1}{2}\|f\|^2 + \langle X, \mathcal{A}^*\lambda + Y - \mu I\rangle + \pi_{\mathcal{S}_+}(Y) \\ & + \frac{\beta}{2}\|\mathcal{A}^*\lambda + Y - \mu I\|_F^2, \end{aligned}$$

where β is a positive penalty parameter and $\pi_{\mathcal{S}_+}$ is the indicator function of the symmetrix PSD matrix cone $\mathcal{S}_+$:

$$\pi_{\mathcal{S}_+}(Y) := \begin{cases} 0 & \text{if } Y \in \mathcal{S}_+, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.2}$$

The matrix $X \in \mathcal{S}^Q$ is the multiplier associated to the equality constraint of the dual problem (4.1), which is also the matrix variable of the primal problem (2.5).

Given an initial iterate $(X^0, Y^0) \in \mathcal{S}^Q \times \mathcal{S}^Q$ such that $Y^0 \succeq 0$, for iteration $k = 0, 1, 2, \ldots$, we alternately update one variable while keeping the others.

- Update λ:

(4.3)
$$\begin{aligned}\lambda^{k+1} &= \arg\min_{\lambda}\left\{\mathcal{L}(\lambda, Y^k, X^k)\right\}\\ &= \arg\min_{\lambda}\left\{\frac{1}{2}(\Phi^\top f+\lambda)^\top(\Phi^\top\Phi)^{-1}(\Phi^\top f+\lambda)+\langle X^k, \mathcal{A}^*\lambda\rangle+\frac{\beta}{2}\|\mathcal{A}^*\lambda+Y^k-\mu I\|_F^2\right\}\\ &= \arg\min_{\lambda}\left\{\frac{1}{2}\lambda^\top\left[(\Phi^\top\Phi)^{-1}+\beta\mathcal{A}\mathcal{A}^*\right]\lambda+\lambda^\top\left[(\Phi^\top\Phi)^{-1}\Phi^\top f+\mathcal{A}(X^k)+\beta\mathcal{A}(Y^k-\mu I)\right]\right\}\\ &= -\left[(\Phi^\top\Phi)^{-1}+\beta\mathcal{A}\mathcal{A}^*\right]^{-1}\left[(\Phi^\top\Phi)^{-1}\Phi^\top f+\mathcal{A}(X^k+\beta Y^k-\beta\mu I)\right].\end{aligned}$$

Moreover, compute

$$w^{k+1} = -(\Phi^\top\Phi)^{-1}(\Phi^\top f+\lambda^{k+1}). \tag{4.4}$$

- Update Y:

$$\begin{aligned}Y^{k+1} &= \arg\min_{Y}\left\{\mathcal{L}(\lambda^{k+1}, Y, X^k)\right\}\\ &= \arg\min_{Y}\left\{\frac{\beta}{2}\|\mathcal{A}^*\lambda^{k+1}+Y-\mu I\|_F^2+\langle X^k, Y\rangle+\pi_{\mathcal{S}_+}(Y)\right\}\\ &= \arg\min_{Y}\left\{\frac{\beta}{2}\|Y+\mathcal{A}^*\lambda^{k+1}+\beta^{-1}X^k-\mu I\|_F^2\,\Big|\,Y\succeq 0\right\}\\ &= \mathcal{P}(-\mathcal{A}^*\lambda^{k+1}-\beta^{-1}X^k+\mu I),\end{aligned} \tag{4.5}$$

where $\mathcal{P}(\cdot)$ is the projection onto the PSD matrix cone.

- Update X:

$$X^{k+1} = X^k+\beta(\mathcal{A}^*\lambda^{k+1}+Y^{k+1}-\mu I). \tag{4.6}$$

Since the matrix Φ has column full rank and $\mathcal{A}\mathcal{A}^*$ is a positive definite diagonal matrix (see Lemma 2.1), the involved matrices in (4.3) are all invertible. Consequently, the equation (4.3) is well defined. On the other hand, from the optimal condition for the primal problem (2.5) and the fact that λ is the multiplier associated to the equality constraint, we see that $\Phi^\top(\Phi w+f)+\lambda = 0$. Thus, we can compute w^{k+1} from (4.4). As to the equality (4.6), we can see that $(\beta^{-1}X^{k+1}, -Y^{k+1})$ is the Moreau decomposition [26] of the matrix $\mathcal{A}^*\lambda^{k+1}+\beta^{-1}X^k-\mu I$ with respect to the PSD matrix cone in the sense of $\langle\beta^{-1}X^{k+1}, -Y^{k+1}\rangle = 0$, $\beta^{-1}X^{k+1}\succeq 0$, $Y^{k+1}\succeq 0$, and $\beta^{-1}X^{k+1}-Y^{k+1} = \mathcal{A}^*\lambda^{k+1}+\beta^{-1}X^k-\mu I$. According to iterates (4.5) and (4.6),

$$\begin{aligned}X^{k+1} &= \beta(\beta^{-1}X^k+\mathcal{A}^*\lambda^{k+1}+Y^{k+1}-\mu I)\\ &= \beta\big(\mathcal{A}^*\lambda^{k+1}+\beta^{-1}X^k-\mu I+\mathcal{P}(-\mathcal{A}^*\lambda^{k+1}-\beta^{-1}X^k+\mu I)\big)\\ &= \beta\mathcal{P}(\mathcal{A}^*\lambda^{k+1}+\beta^{-1}X^k-\mu I),\end{aligned} \tag{4.7}$$

which means that we can first, in an alternative way, compute the iterate X^{k+1} by (4.7), and then compute Y^{k+1} via $Y^{k+1} = \beta^{-1}X^{k+1} - (\mathcal{A}^*\lambda^{k+1} + \beta^{-1}X^k - \mu I)$.

We stop the algorithm if the duality gap, together with the primal infeasibility and the dual infeasibility, is sufficiently small. Here, the duality gap in each iteration is

$$\begin{aligned}\text{gap}^k &= \frac{1}{2}\|\Phi w^k + f\|^2 + \mu\langle X, I\rangle - \left(-\frac{1}{2}(\Phi^\top f + \lambda^k)^\top(\Phi^\top\Phi)^{-1}(\Phi^\top f + \lambda^k) + \frac{1}{2}\|f\|^2\right)\\ &= \frac{1}{2}\|\Phi w^k\|^2 + f^\top\Phi w^k + \frac{1}{2}{w^k}^\top(\Phi^\top\Phi)w^k + \langle X, \mu I\rangle\\ &= {w^k}^\top(\Phi^\top\Phi w^k + \Phi^\top f) + \langle X, \mu I\rangle\\ &= -{w^k}^\top\lambda^k + \langle X, \mu I\rangle,\end{aligned}$$

where the second and the fourth equalities hold because of (4.4).

5. Sensitivity analysis.

In diffusion weighted MRI data, noises may appear in measured signals including S_0 and $S^{(n)}$ for $n = 1, 2, \ldots, N$. These noise-corrupted signals constitute the vector f in (2.4). So it is important to analyze the influences on the optimal solution and the optimal value of the new model (2.5) as the vector f changes. For convenience, we denote the noise-corrupted vector f as $\widetilde{f}$ in the remainder of this section.

Theorem 5.1. *Consider the parameterized problem*

$$w(f) := \arg\min_{w,X}\left\{\frac{1}{2}\|\Phi w + f\|^2 + \mu\langle X, I\rangle \;\middle|\; \mathcal{A}(X) = w, X \succeq 0\right\}, \tag{5.1}$$

where $w(f)$ represents the optimal diffusion tensor coefficients corresponding to the input f. Then, for any vectors f and its noise corruption $\widetilde{f}$, we have

$$[w(f) - w(\widetilde{f})]\Phi^\top(f - \widetilde{f}) \le 0.$$

Proof. Let $(w(f), X(f))$ and $(w(\widetilde{f}), X(\widetilde{f}))$ be the optimal solutions of the parameterized problem (5.1) with inputs f and $\widetilde{f}$, respectively. Then, we have

$$\frac{1}{2}\|\Phi w(f) + f\|^2 + \mu\langle X(f), I\rangle \le \frac{1}{2}\|\Phi w(\widetilde{f}) + f\|^2 + \mu\langle X(\widetilde{f}), I\rangle$$

and

$$\frac{1}{2}\|\Phi w(\widetilde{f}) + \widetilde{f}\|^2 + \mu\langle X(\widetilde{f}), I\rangle \le \frac{1}{2}\|\Phi w(f) + \widetilde{f}\|^2 + \mu\langle X(f), I\rangle.$$

After adding the above two inequalities and rearranging terms, we get

$$w(f)^\top\Phi^\top f + w(\widetilde{f})^\top\Phi^\top\widetilde{f} \le w(\widetilde{f})^\top\Phi^\top f + w(f)^\top\Phi^\top\widetilde{f},$$

which completes the proof. ∎

We now analyze the property of the optimal objective function value as a function of the input f. We have the following result.

Theorem 5.2. *We define the optimal objective function value as a function of the input f:*

$$\mathbb{P}(f) := \min_{w,X} \left\{ \frac{1}{2}\|\Phi w + f\|^2 + \mu\langle X, I\rangle \;\middle|\; \mathcal{A}(X) = w, X \succeq 0 \right\}. \tag{5.2}$$

Then, $\mathbb{P}(f)$ is a convex function.

Moreover, there exists a vector $\lambda(f)$ such that for every noise-corrupted $\widetilde{f}$,

$$0 \le \mathbb{P}(\widetilde{f}) - \mathbb{P}(f) + \lambda(f)^\top(\widetilde{f} - f) \le \frac{1}{2}\|\widetilde{f} - f\|^2. \tag{5.3}$$

Proof. To establish the convexity of $\mathbb{P}(f)$, we introduce a linear equality constraint $\Phi w + f = h$ and a second order cone $\mathcal{K} := \{(t;h) \,|\, t \ge \|h\|\}$. Then, the quadratic term in the objective function of problem (5.2) could be rewritten as

$$\begin{aligned}
\mathbb{P}(f) &= \min_{(t;h),w,X} \left\{ \frac{t^2}{2} + \mu\langle X, I\rangle \;\middle|\; \Phi w + f = h, (t,h) \in \mathcal{K}, \mathcal{A}(X) = w, X \succeq 0 \right\} \\
&= \min_{(t;h),X} \left\{ \frac{t^2}{2} + \mu\langle X, I\rangle \;\middle|\; \Phi\mathcal{A}(X) + f = h, X \succeq 0, (t,h) \in \mathcal{K} \right\} \\
&= \min_{(t;h),X} \left\{ \frac{t^2}{2} + \mu\langle X, I\rangle + \pi_{\mathcal{S}_+}(X) + \pi_{\mathcal{K}}(t,h) \;\middle|\; \Phi\mathcal{A}(X) + f = h \right\}, \qquad (5.4)
\end{aligned}$$

where $\pi_{\mathcal{S}_+}(\cdot)$ and $\pi_{\mathcal{K}}(\cdot)$ are indicator functions (4.2) of the convex sets $\mathcal{S}_+$ and $\mathcal{K}$, respectively. Then, the parameterized Lagrangian of the optimization problem (5.4) is

$$L_f(t,h,X,\lambda) := \frac{t^2}{2} + \mu\langle X, I\rangle + \pi_{\mathcal{S}_+}(X) + \pi_{\mathcal{K}}(t,h) - \lambda^\top(\Phi\mathcal{A}(X) + f - h),$$

where λ is the multiplier corresponding to the equality constraint.

Let $(t(f), h(f), X(f), \lambda(f))$ and $(t(\widetilde{f}), h(\widetilde{f}), X(\widetilde{f}), \lambda(\widetilde{f}))$ be the optimal solutions of the optimization problem with inputs f and $\widetilde{f}$, respectively. Then,

$$\begin{aligned}
\mathbb{P}(f) &= L_f(t(f), h(f), X(f), \lambda(f)) \\
&\le L_f(t(\widetilde{f}), h(\widetilde{f}), X(\widetilde{f}), \lambda(f)) \\
&= L_{\widetilde{f}}(t(\widetilde{f}), h(\widetilde{f}), X(\widetilde{f}), \lambda(f)) + \lambda(f)^\top(\widetilde{f} - f) \\
&\le L_{\widetilde{f}}(t(\widetilde{f}), h(\widetilde{f}), X(\widetilde{f}), \lambda(\widetilde{f})) + \lambda(f)^\top(\widetilde{f} - f) \\
&= \mathbb{P}(\widetilde{f}) + \lambda(f)^\top(\widetilde{f} - f). \qquad (5.5)
\end{aligned}$$

So $\mathbb{P}(f)$ is a convex function, and the left inequality of (5.3) holds.

We now turn to problem (5.2) and define $w(f) := \mathcal{A}(X(f))$ and $w(\widetilde{f}) := \mathcal{A}(X(\widetilde{f}))$. Then,

$$\begin{aligned}
\mathbb{P}(\widetilde{f}) &= \frac{1}{2}\|\Phi w(\widetilde{f}) + \widetilde{f}\|^2 + \mu\langle X(\widetilde{f}), I\rangle \\
&\le \frac{1}{2}\|\Phi w(f) + \widetilde{f}\|^2 + \mu\langle X(f), I\rangle \\
&= \frac{1}{2}\|\Phi w(f) + f\|^2 + \mu\langle X(f), I\rangle + \frac{1}{2}\|\widetilde{f} - f\|^2 + (\Phi w(f) + f)^\top(\widetilde{f} - f) \\
&= \mathbb{P}(f) + \frac{1}{2}\|\widetilde{f} - f\|^2 + (\Phi w(f) + f)^\top(\widetilde{f} - f). \qquad (5.6)
\end{aligned}$$

It remains to prove $\lambda(f) = -(\Phi w(f) + f)$. Adding inequalities (5.5) and (5.6) and rearranging terms yield

$$\frac{1}{2}\|\widetilde{f} - f\|^2 + (\lambda(f) + \Phi w(f) + f)^\top (\widetilde{f} - f) \geq 0.$$

Multiplying $\|\widetilde{f} - f\|^{-1}$ on both sides of the above inequality and letting $\widetilde{f}$ get close to f, we get

$$(\lambda(f) + \Phi w(f) + f)^\top e_f \geq 0,$$

where

$$e_f = \lim_{\|\widetilde{f}-f\|\to 0} \frac{\widetilde{f} - f}{\|\widetilde{f} - f\|}$$

is a unit vector with arbitrary orientations. Hence, $\lambda(f) + \Phi w(f) + f = 0$. ∎

6. Experimental results. To show the validation of the new quadratic SDP model for the estimation of higher order generalized diffusion tensors, we apply it to synthetic and human brain data. Before we start, it is important to choose a proper regularization parameter for the quadratic SDP model.

6.1. Choosing the regularization parameter μ. The determination of the optimal regularization parameter μ for the convex quadratic SDP problem (2.5) is a challenging problem. Here, we aim to choose a proper μ which makes a balance between the least squares fitting and the nuclear norm regularization. Let w_0 be the solution for the basic unconstrained least squares model, i.e.,

$$w_0 := \arg\min_w \left\{ \frac{1}{2}\|\Phi w + f\|^2 \right\},$$

which is easy to obtain. Then, two terms $\|\Phi w_0 + f\|^2/2$ and $\|w_0\|_1$ should have the same order of magnitude as $\|\Phi w^* + f\|^2/2$ and $\langle X^*, I\rangle$, respectively, where (w^*, X^*) denotes the optimal solution of the quadratic SDP model (2.5). So, we take

$$\mu = \kappa \cdot \frac{\|\Phi w_0 + f\|^2}{2\|w_0\|_1},$$

where κ is a fixed scalar factor.

By numerical experiences, we list the "best" scalar factor κ corresponding to each order R of the generalized diffusion tensor in Table 1. We find that as the order R of the generalized diffusion tensor increases, the regularization parameter μ should increase too.

Table 1

The "best" scalar factor for the regularization parameter μ according to R.

R	2	4	6
κ	0.01	1	10

6.2. Synthetic experiments. First, we examine the accuracy of the generalized diffusion tensors estimated by the quadratic SDP model (2.5). The synthetic data used here is generated by the multitensor model [1, 10]

$$S^{(n)} = S_0 \sum_{k=1}^{K} v_k \exp\left(-b\, \mathbf{g}^{(n)\top} \mathbf{D}_k \mathbf{g}^{(n)}\right).$$

The setting of parameters is as follows: the zero-gradient signal $S_0 = 1$, the diffusion weighted b-value $b = 1500s/mm^2$, each volume fraction $v_k = K^{-1}$, eigenvalues of each 3×3 diffusion tensor $\mathbf{D}_k$ are $(0.2, 0.2, 1.7) \times 10^{-3} mm^2/s$, and gradient orientations $\mathbf{g}^{(n)}$ for $n = 1, \ldots, 81$ on the unit semisphere come from third order tessellation of the icosahedron. Once the synthetic signals are obtained, we compute the coefficient matrix Φ and the vector f by (2.4). Then, we define the ground truth solution as

$$w_{groundTruth} := -(\Phi^\top \Phi)^{-1} \Phi^\top f.$$

Hereafter, the original signal S is corrupted by the Rician noise [10]. Given a signal-to-noise ratio (SNR), we produce two Gaussian noises $n_r, n_i \sim \mathcal{N}(0, \sigma^2)$, where $S_0/\sigma = \text{SNR}$. Then, the Rician noise-corrupted signal $\widetilde{S}$ is given by

$$\widetilde{S} = \sqrt{(S + n_r)^2 + n_i^2}.$$

To show the robustness of the quadratic SDP model, we compare it with three existing methods.

♢ (Unconst. LS) The unconstrained least squares method [19].

♢ (Decomposition) The nonconvex decomposition method proposed by Ghosh, Deriche, and Moakher [14], which fits the log-linearized Stejskal–Tanner equation. In this experiment, the nonlinear least squares problem involved is solved by the "`lsqcurvefit`" function in MATLAB. However, this method works only for estimating fourth order diffusion tensors.

♢ (fanDTasia) A unified discrete convex model proposed by Barmpoutis and Vemuri [4].

♠ (New method) The quadratic SDP model established in section 2.

For each method, we perform 1000 tests for each Rician noise level and show the means of differences between the estimated diffusion tensors and the ground truth solution.

Now, we consider the case of two-fiber crossing with a separation angle of 60°. The numerical results are illustrated in Figure 2, when fourth and sixth order generalized diffusion tensors are exploited. Obviously, the convex quadratic SDP model outperforms the unconstrained least squares method. So the positive definiteness constrained for the generalized diffusion tensor is necessary and meaningful. However, the nonconvex decomposition method performs poorly since it may fall into local optima.

Compared with the discrete convex nonnegative linear least squares model fanDTasia, the continuous convex quadratic SDP model usually produces a solution with higher accuracy for lower SNR data. So it is promising and inspiring. Particularly, when SNR varies between 30 and 60, we find it is hard to distinguish the performance of fanDTasia from the unconstrained least squares method for sixth order generalized diffusion tensor estimations. However, distinct

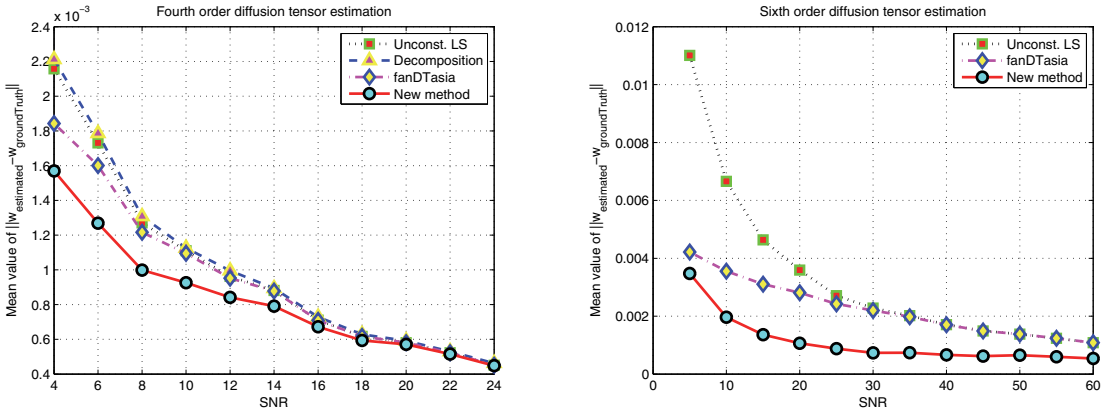

Figure 2. *The performance of generalized diffusion tensor estimations produced by four kinds of methods applied to the two-fiber crossing data.*

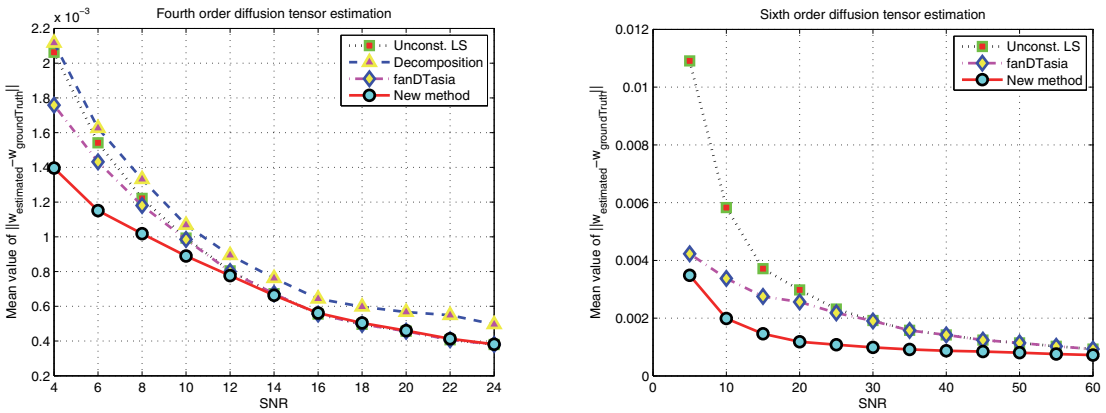

Figure 3. *The performance of generalized diffusion tensor estimations produced by four kinds of methods applied to the three-fiber crossing data.*

improvement is made by the convex quadratic SDP method. We think that the main reason for this improvement is due to the successful application of the nuclear norm regularization.

In Figure 3, similar results are observed for the three-fiber crossing, where two fibers are separated by an angle of 60° and the third fiber is perpendicular to them.

Second, using the generalized diffusion tensor estimation, we study the fiber orientation distribution function (ODF), whose maximums approximate orientations of fiber bundles in each voxel. The flowchart of this process is shown in Figure 4. Given an ADC function $D(\mathbf{g})$, we compute the trustable signal function

$$\overline{S}(\mathbf{g}) = \exp(-bD(\mathbf{g})).$$

Then, we approach this trustable signal function by even order real spherical harmonics (SH)

$$\sum_j c_j Y_j(\mathbf{g}) = \overline{S}(\mathbf{g}),$$

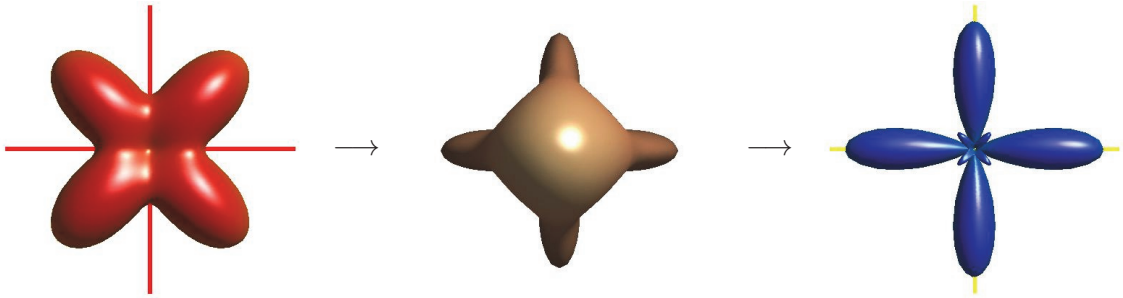

Figure 4. *The flowchart of estimating the fiber ODF from a generalized diffusion tensor is illustrated here. We first compute trustable signals from the generalized diffusion tensor. Then, the Q-ball imaging and the spherical deconvolution are employed to generate the fiber ODF. Additionally, the red lines in the ADC subfigure (left) show true fiber orientations, and the yellow lines in the ODF subfigure (right) show estimated fiber orientations.*

where $Y_j(\cdot)$ is the real SH basis with an order l_j, and c_j is the corresponding SH coefficients. Since the nuclear norm regularization is considered in the generalized diffusion tensor estimation, we prefer to estimate these SH coefficients from the trustable signal without any regularization. Once these SH coefficients are obtained, the fiber ODF gets the following analytical formulation:

$$\Psi(\mathbf{g}) = \sum_j \frac{2\pi P_{l_j}(0) c_j}{f_j} Y_j(\mathbf{g}),$$

where P_{l_j} is the Legendre polynomial of a degree l_j, and the sharpening coefficient

$$f_j = 2\pi \int_{-1}^{1} P_{l_j}(t) R(t)\, \mathrm{d}t$$

comes from the spherical deconvolution with a proper diffusion ODF kernel $R(t)$. For more details on computing the analytical fiber ODF from signals, we refer the reader to works on the Q-ball imaging [34, 11, 12].

By this flowchart, we show in Figure 5 the fiber ODF profiles estimated by the quadratic SDP model (2.5) under multiple fiber orientations, multiple orders of generalized diffusion tensors, and multiple diffusion weighted b-values, where a moderate SNR = 35 and 100 tests are performed. The best performance is observed when we take sixth order generalized diffusion tensors and use diffusion weighted b-value $3000s/mm^2$. In the remainder of this section, we always use these settings.

Next, we compare the quadratic SDP model (2.5) with the Q-ball imaging method of Descoteaux et al. [12] that directly estimates fiber ODF from raw signals. We take two-fiber crossing, for example.

Figure 6 shows fiber ODF profiles corresponding to multiple separation angles between 90° and 40°. When the separation angle is large, both methods report correct fiber orientations. However, when the separation angle decreases, the estimated fiber ODF trends to oversmooth, which may fail to capture the profiles of fiber crossing. Particularly, when the separation angle

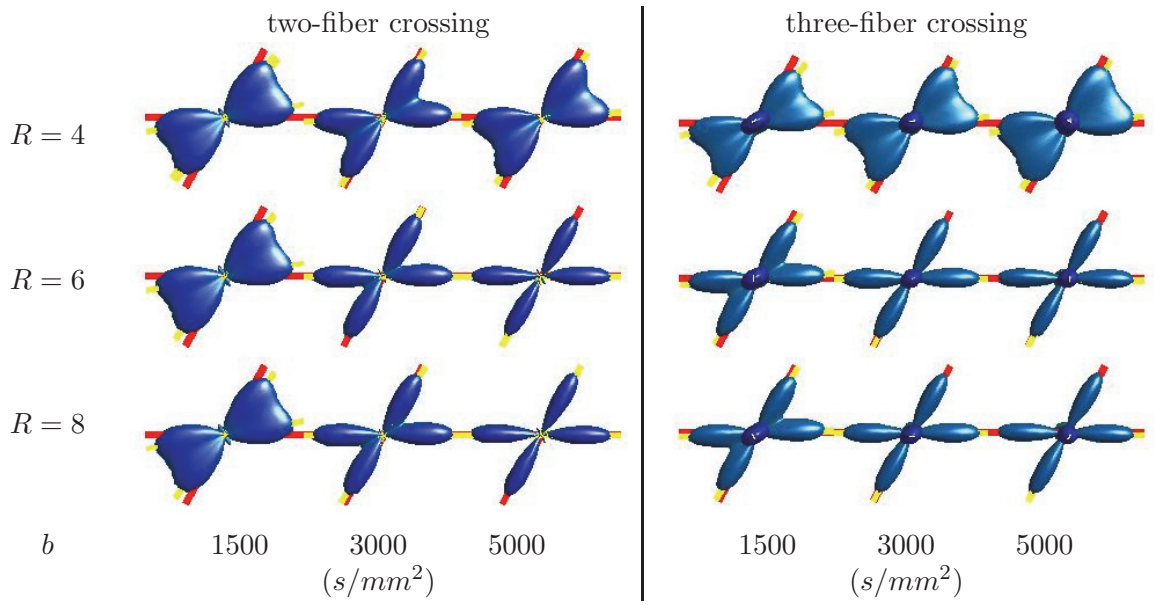

Figure 5. *The fiber ODFs estimated by the quadratic SDP model with multiple fiber orientations, diffusion weighted b-values, and orders of the generalized diffusion tensor.*

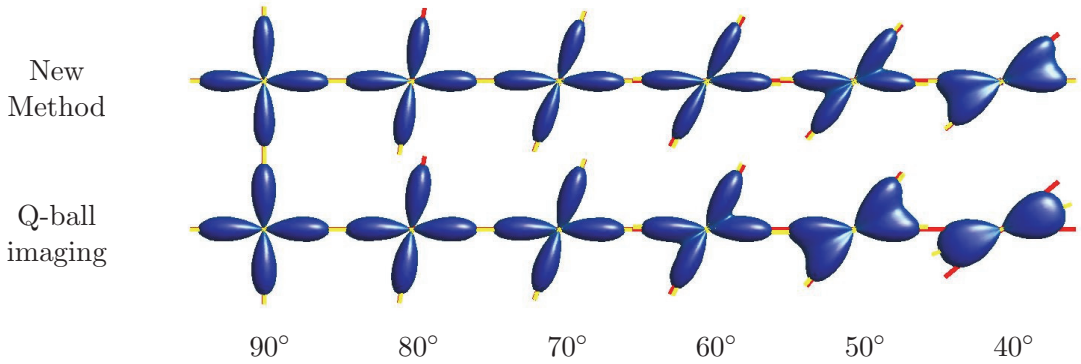

Figure 6. *Comparisons of the new quadratic SDP method with the Q-ball imaging [12] in the context of varying separation angles.*

is 40°, the quadratic SDP model (2.5) successfully detects two-fiber crossing, while the direct Q-ball imaging fails.

Figure 7 reports estimated fiber ODF profiles under the impact of volume fractions, where a fixed separation angle 60° is used. Obviously, the quadratic SDP model is competitive with the direct Q-ball imaging method.

6.3. Human brain study. The human brain data used here is a high angular resolution diffusion imaging dataset, where the diffusion weighted b-value is $3000 s/mm^2$, the number of gradient orientations is 200, and the size of each voxel is $1.875 \times 1.875 \times 2\, mm^3$. For the convenience of illustrating experimental results on higher order generalized diffusion tensors,

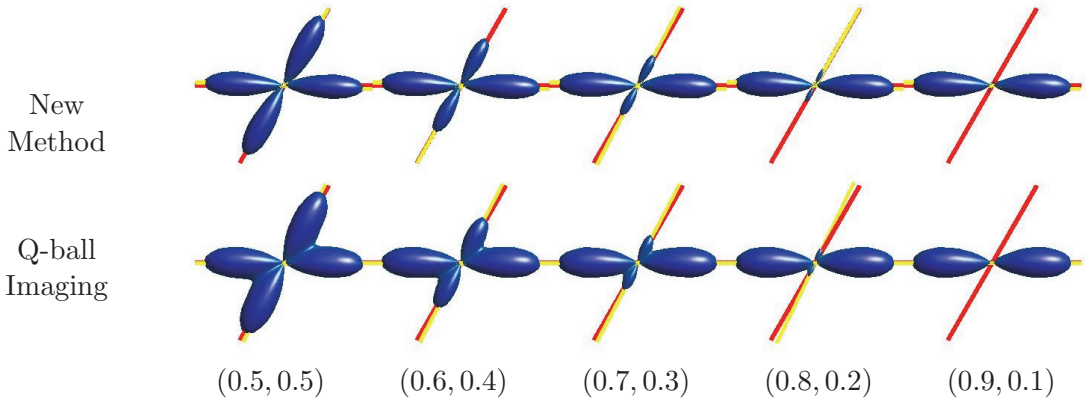

Figure 7. *Comparisons of the new quadratic SDP method with the Q-ball imaging* [12] *in the context of varying volume fractions.*

we employ two rotationally invariant scalar measures [20]. The first is the generalized mean diffusivity (MD),

$$\mathrm{MD} := \frac{1}{4\pi}\int_{\Omega} D(\mathbf{g})\mathrm{d}\mathbf{g},$$

where Ω is the unit spherical surface. The second is the generalized anisotropy (GA), which is based on the generalized variance V of diffusivity,

$$V = \frac{1}{9}\left(\frac{1}{\mathrm{MD}^2}\cdot\frac{1}{4\pi}\int_{\Omega}[D(\mathbf{g})]^2\mathrm{d}\mathbf{g} - 1\right).$$

Then, we scale the range of the function V into $[0, 1)$ and obtain

$$\mathrm{GA} := 1 - \frac{1}{1 + (250V)^{\varepsilon(V)}},$$

where

$$\varepsilon(V) := 1 + \frac{1}{1 + 5000V}.$$

Here, GA getting close to one means that the diffusion is anisotropic. The MD and GA for generalized diffusion tensors are the generalizations of often used MD and fractional anisotropy (FA) for second order diffusion tensors.

As an application of the quadratic SDP model, we report the fiber reconstruction of the corticospinal tract meeting the corpus callosum. We take a coronal slice of a healthy human brain, whose FA map is shown in part (a) of Figure 8. The fiber ODFs corresponding to the region marked by a magenta box are illustrated in part (b) of Figure 8. In its center, we highlight a typical fiber ODF (marked by a red circle) in part (c). Obviously, there are three nervous fiber bundles detected in this voxel. The left-to-right directional fiber bundle is the corpus callosum, which connects the left and right cerebral hemispheres and facilitates interhemispheric communication. The up-to-down directional fiber bundle is the corticospinal tracts, which control voluntary movement of skeletal muscles. The front-to-back directional

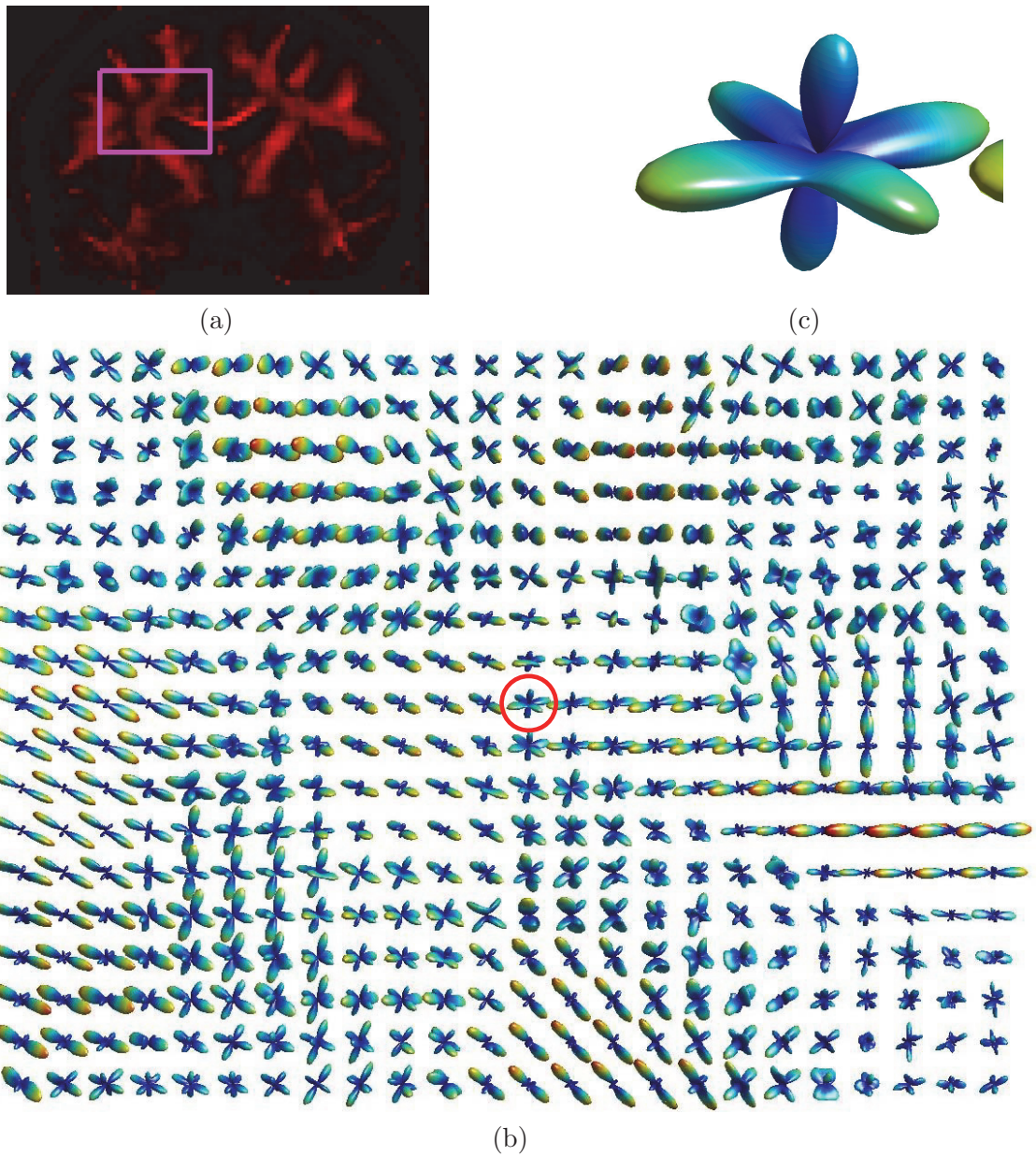

Figure 8. *A coronal slice of a healthy human brain is analyzed by the quadratic SDP model. Part* (a) *is the FA map. Part* (b) *shows the fiber ODFs of the region marked by the magenta box. A typical fiber ODF* (c) *highlighted by the red circle illustrates the crossing of three fibers.*

fiber bundle is the superior longitudinal fasciculus, which connects the front and the back of the cerebrum.

Another experiment is performed to validate the positive definiteness of the generalized diffusion tensor estimated by the convex quadratic SDP model. First, we use the unconstrained least squares method to estimate generalized diffusion tensors with order two, four, and six.

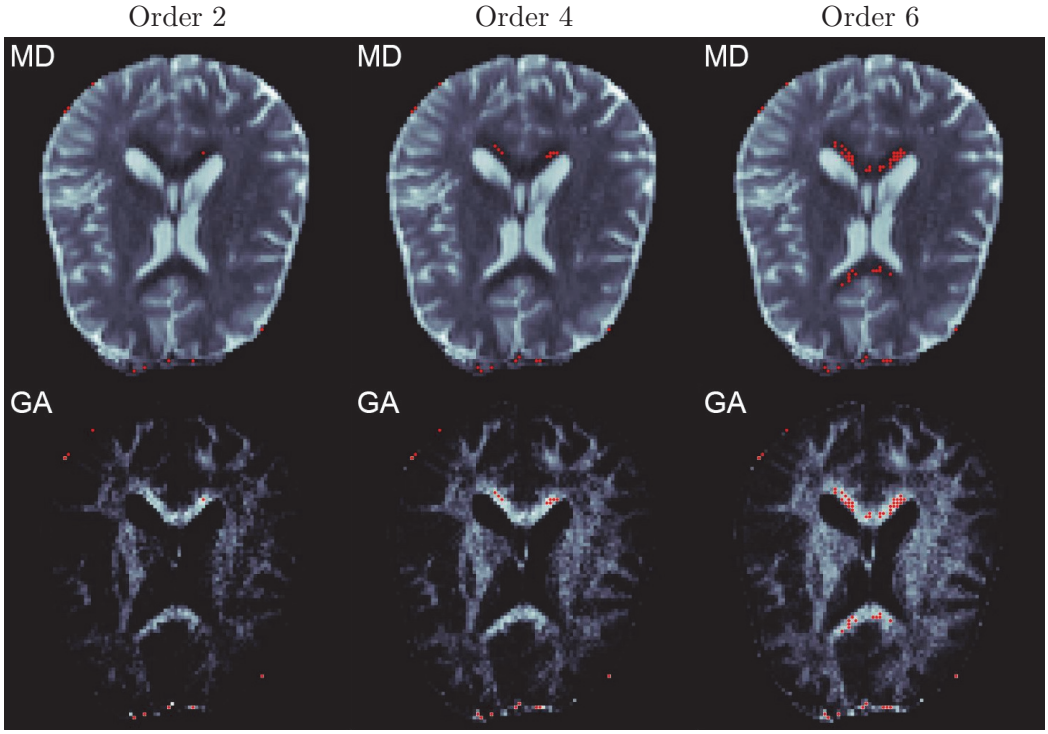

Figure 9. *MD and GA images of second, fourth, and sixth order diffusion tensors estimated by the unconstrained least squares method.*

These MD and GA maps are shown in Figure 9. In each voxel, we check whether the positive diffusivity is violated in any gradient orientation. If negative diffusivity is detected, we point out the corresponding voxel by red points. Except for the obviously noise voxel in the boundary of the cerebrum, the unconstrained least squares method suffers from 1, 8, and 42 meaningless red-pointed voxels for second, fourth, and sixth order diffusion tensor estimations, respectively. So the PSD constraint (1.3) is important for diffusion weighted MRI.

The MD and GA maps corresponding to the new convex quadratic SDP estimator are illustrated in Figure 10. Negative diffusion is not detected in any gradient orientation in each voxel. So we conclude that the new convex quadratic SDP estimator has the ability to produce meaningful generalized diffusion tensors.

6.4. Comparison of three solvers. The convex quadratic SDP model could be solved by three methods: SDPT3 and QSDP for the primal problem and ADM for the dual problem. Here, the ADM is accelerated by two skills: adaptively updating the penalty parameter and using a step size to update the multiplier [16, 15, 38].

We apply these three solvers to deal with the human brain data used in Figure 10. Their mean computational times for each voxel are listed in Table 2.[3] Obviously, compared with

[3]The computational time of QSDP applied to diffusion tensor estimation with second order is longer than that with fourth order because 0.4% of the problems reach the maximum number of iterations in the former case.

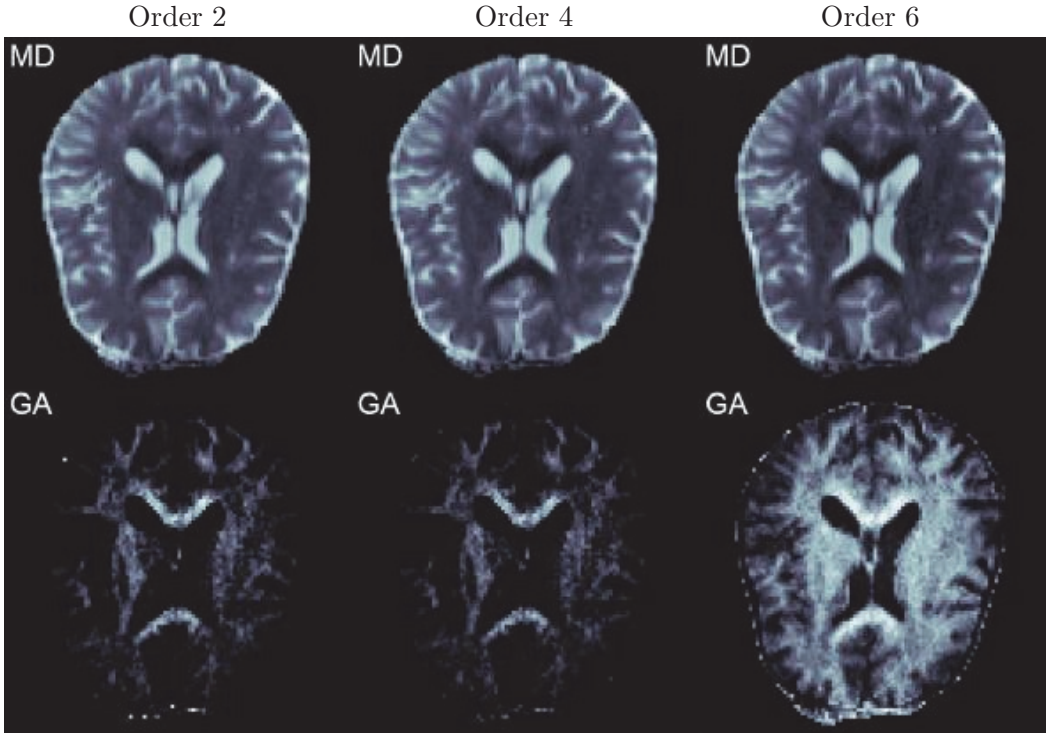

Figure 10. *MD and GA images of second, fourth, and sixth order diffusion tensors estimated by the convex quadratic SDP model.*

Table 2
The mean computational time (ms) of SDPT3, QSDP, and ADM solvers applied to one voxel.

R	SDPT3	QSDP	ADM
2	874	133	0.62
4	1027	130	7.2
6	1335	201	16.5

SDPT3 and QSDP solvers, the new solver ADM saves at least 90% CPU time. So ADM is very efficient for this special quadratic SDP model, which derives from the diffusion weighted MRI.

7. Conclusion. A novel quadratic SDP model was established for estimating higher order generalized diffusion tensors for the diffusion weighted MRI. This model benefits from a PSD constraint and a nuclear norm regularization. The positive semidefiniteness ensures that the resulting generalized diffusion tensor is physically significant, and the regularization makes the estimated solution robust to Rician noise. Sensitivity analysis on noise-corrupted signals was presented. For solving the new quadratic SDP model, two existing solvers, i.e., SDPT3 and QSDP, were introduced. In addition, we also proposed an ADM and compared it with SDPT3 and QSDP. Our numerical experiments show that the first order solver ADM, which exploits the special structure of the new model successfully, is efficient.

A well-selected regularization parameter μ could improve the quadratic SDP model (2.5). Although the strategy described in section 6.1 seems to work well, the problem of how to determine the best regularization parameter is still an open qustion.

Acknowledgments. The authors are grateful to Associate Editor René Vidal and the three anonymous referees for helping us improve the original manuscript.

REFERENCES

[1] D. C. Alexander, G. J. Barker, and S. R. Arridge, *Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data*, Magn. Reson. Med., 48 (2002), pp. 331–340.

[2] A. Barmpoutis, M. S. Hwang, D. Howland, J. R. Forder, and B. C. Vemuri, *Regularized positive-definite fourth order tensor field estimation from DW-MRI*, NeuroImage, 45 (2009), pp. 153–162.

[3] A. Barmpoutis, B. Jian, B. C. Vemuri, and T. M. Shepherd, *Symmetric positive 4th order tensors & their estimation from diffusion weighted MRI*, in Information Processing in Medical Imaging, Lecture Notes in Comput. Sci. 4584, N. Karssemeijer and B. Lelieveldt, eds., Springer-Verlag, Berlin, Heidelberg, 2007, pp. 308–319.

[4] A. Barmpoutis and B. C. Vemuri, *A unified framework for estimating diffusion tensors of any order with symmetric positive-definite constraints*, in Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2010, pp. 1385–1388.

[5] P. J. Basser and D. K. Jones, *Diffusion-tensor MRI: Theory, experimental design and data analysis—A technical review*, NMR Biomed., 15 (2002), pp. 456–467.

[6] P. J. Basser, J. Mattiello, and D. LeBihan, *Estimation of the effective self-diffusion tensor from the NMR spin echo*, J. Magn. Reson., 103 (1994), pp. 247–254.

[7] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, *Diffusion tensor imaging: Concepts and applications*, J. Magn. Reson. Imaging, 13 (2001), pp. 534–546.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn., 3 (2011), pp. 1–122.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[10] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, *Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications*, Magn. Reson. Med., 56 (2006), pp. 395–410.

[11] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche, *Regularized, fast, and robust analytical Q-ball imaging*, Magn. Reson. Med., 58 (2007), pp. 497–510.

[12] M. Descoteaux, R. Deriche, T. R. Knösche, and A. Anwander, *Deterministic and probabilistic tractography based on complex fibre orientation distributions*, IEEE Trans. Med. Imag., 28 (2009), pp. 269–286.

[13] M. Fazel, *Matrix Rank Minimization with Applications*, Ph.D. thesis, Department of Electrical Engineering, Stanford University, Stanford, CA, 2002.

[14] A. Ghosh, R. Deriche, and M. Moakher, *Ternary quartic approach for positive 4th order diffusion tensors revisited*, in Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009, pp. 618–621.

[15] B. He, L.-Z. Liao, D. Han, and H. Yang, *A new inexact alternating directions method for monotone variational inequalities*, Math. Program., 92 (2002), pp. 103–118.

[16] B. S. He, H. Yang, and S. L. Wang, *Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities*, J. Optim. Theory Appl., 106 (2000), pp. 337–356.

[17] M. R. Jayachandra, N. Rehbein, C. Herweh, and S. Heiland, *Fiber tracking of human brain using fourth-order tensor and high angular resolution diffusion imaging*, Magn. Reson. Med., 60 (2008), pp. 1207–1217.

[18] J. B. Lasserre, *A sum of squares approximation of nonnegative polynomials*, SIAM Rev., 49 (2007), pp. 651–669.
[19] E. Özarslan and T. H. Mareci, *Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging*, Magn. Reson. Med., 50 (2003), pp. 955–965.
[20] E. Özarslan, B. C. Vemuri, and T. H. Mareci, *Generalized scalar measures for diffusion MRI using trace, variance, and entropy*, Magn. Reson. Med., 53 (2005), pp. 866–876.
[21] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.
[22] V. Powers and B. Reznick, *Notes towards a constructive proof of Hilbert's theorem on ternary quartics*, in Quadratic Forms and Their Applications (Dublin 1999), Contemp. Math. 272, E. Bayer-Fluckiger, D. Lewis, and A. Ranicki, eds., AMS, Providence, RI, 2000, pp. 209–227.
[23] L. Qi, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302–1324.
[24] L. Qi, G. Yu, and E. X. Wu, *Higher order positive semidefinite diffusion tensor imaging*, SIAM J. Imaging Sci., 3 (2010), pp. 416–433.
[25] B. Recht, M. Fazel, and P. A. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev., 52 (2010), pp. 471–501.
[26] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[27] W. Rudin, *Sums of squares of polynomials*, Amer. Math. Monthly, 107 (2000), pp. 813–821.
[28] E. O. Stejskal and J. E. Tanner, *Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient*, J. Chem. Phys., 42 (1965), pp. 288–292.
[29] K.-C. Toh, *An inexact primal-dual path following algorithm for convex quadratic SDP*, Math. Program., 112 (2008), pp. 221–254.
[30] K.-C. Toh, *User Guide for QSDP-0—A MATLAB Software Package for Convex Quadratic Semidefinite Programming*, Technical report, Department of Mathematics, National University of Singapore, Singapore, 2010.
[31] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3—A MATLAB software package for semidefinite programming, Version* 1.3, Optim. Methods Softw., 11 (1999), pp. 545–581.
[32] K.-C. Toh, R. H. Tütüncü, and M. J. Todd, *On the Implementation and Usage of SDPT3—A MATLAB Software Package for Semidefinite-Quadratic-Linear Programming, Version* 4.0, Technical report, Department of Mathematics, National University of Singapore, 2006.
[33] K.-C. Toh, R. H. Tütüncü, and M. J. Todd, *Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems*, Pac. J. Optim., 3 (2007), pp. 135–164.
[34] D. S. Tuch, *Q-ball imaging*, Magn. Reson. Med., 52 (2004), pp. 1358–1372.
[35] R. H. Tütüncü, K.-C. Toh, and M. J. Todd, *Solving semidefinite-quadratic-linear programs using SDPT3*, Math. Program., 95 (2003), pp. 189–217.
[36] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci, *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI*, in Information Processing in Medical Imaging, Lecture Notes in Comput. Sci. 2732, C. Taylor and J. Noble, eds., Springer-Verlag, Berlin, Heidelberg, 2003, pp. 660–671.
[37] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci, *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI*, IEEE Trans. Med. Imag., 23 (2004), pp. 930–939.
[38] Z. Wen, D. Goldfarb, and W. Yin, *Alternating direction augmented Lagrangian methods for semidefinite programming*, Math. Program. Comput., 2 (2010), pp. 203–230.