Published in final edited form as:
IEEE Trans Med Imaging. 2011 March ; 30(3): 642–654. doi:10.1109/TMI.2010.2090539.

Minimization of Imaging Gradient Effects in Diffusion Tensor Imaging

Alpay Özcan [Senior Member, IEEE]
Biomedical MR Laboratory, Mallinckrodt Institute of Radiology, Washington University in Saint Louis, School of Medicine, St. Louis, MO 63110 USA (ozcan@zach.wustl.edu)

Abstract

A new, sample independent optimization criterion for minimizing the effect of the imaging gradients, including the directional awareness they create, is defined for diffusion tensor imaging (DTI) experiments. The DTI linear algebraic framework is expanded to a normed space to design optimal diffusion gradient schemes (DGS) in an integral fashion without separating the magnitude and direction of the DGS vectors. The feasible space of DGS vectors, for which the estimation equations are determinate, thus a hard constraint for the optimization, is parametrized. Newly generated optimal DGSs demonstrate on an isotropic sample and an ex-vivo baboon brain that the optimization goals are reached as well as a significant increase in estimation performance.

Index Terms

Diffusion tensor imaging (DTI); imaging gradients; optimization

I. INTRODUCTION

The diffusion tensor imaging (DTI) is an important technique that is being used extensively in clinical and research environments. Its importance stems from its capability of inferring the tissue microstructure by representing the measurements of the spin motion using magnetic resonance imaging (MRI). With this information, it is possible to monitor different conditions such as multiple sclerosis [1], [2], cancer [3], Alzheimer's disease [4], and spinal cord injury [5]. DTI is also the basis of fiber tracking algorithms [6], [7] that found a wide variety of application areas [8].

The origin of DTI is the adaptation of the work of Torrey [9], which resulted in Bloch–Torrey or modified Bloch equations, to the pulsed-gradient spin-echo (PGSE) experiment by Stejskal and Tanner [10]. In essence, the measurements in [10] rely on a pair of strong magnetic field gradients that sensitize the signal to the motion of the spins. The gradients attenuate the nuclear magnetic resonance (NMR) signal's peak as a function of the newly added gradients.

With the invention of MRI creating the ability of applying the magnetic field gradients as three dimensional vectors, the spin motion, which was only described in NMR experiments as a scalar quantity, is calculated as a three dimensional construct. The modeling of NMR diffusion experiments was directly adapted to MRI using a tensor algebraic framework in [11], [12]. The diffusion tensor estimated by DTI is a symmetric tensor of rank two and thus is a quadratic form that is represented by a symmetric matrix D. In simplest terms, DTI presents the general tendency of the incoherent spin motion with ellipsoids.

Whether seen as matrices or quadratic forms, these objects are elements of a vector space. In [13] the DTI estimation equations were written appropriately using linear algebra. The imaging gradients were not included in the formulation, in a sense parallel to the adaptation from NMR experiments that do not possess any gradients other than the diffusion encoding ones.

Considering that physically any magnetic field gradient affects the signal according to the motion of the spin packets and that the magnitude of the imaging gradients is inversely proportional to the field of view (FOV), it is imperative that a complete DTI model includes the effect of the imaging gradients, especially for the experiments with small FOV. The inclusion was presented for the tensor algebraic framework in [12]. The same task is accomplished in a much less tedious fashion with the minimalist linear algebraic framework introduced in [14], [15]. The framework is more utilitarian than the tensor algebraic one: it provided a set of necessary conditions for the design of the DTI experiments that guarantee the uniqueness of solutions for linear and nonlinear estimation [16]–[18]. Moreover, it reveals modeling incongruities that are exposed in detail in [14], [15], [19] and are also summarized in Appendix A–C. Most importantly, it created the basis of this manuscript's foundation: the normed space that allows the description of a sample independent, geometric optimization problem to design more efficient DTI experiments and the parametrization of the feasible space where the search for the optimal diffusion gradient vectors scheme takes place.

The design of DTI experiments consists of finding a diffusion gradient scheme (DGS) in the 3D space i.e., a set of at least six vectors that define the diffusion encoding magnetic field gradient magnitude and direction.

In a scenario where the effect of the imaging gradients is neglected, experiments conducted with an isotropic sample using a given DGS and any other scheme obtained by rotating the original DGS as a group should give identical results. Clearly, the imaging gradients introduce a directional awareness. The DTI parameters (e.g., the eigenvalues) obtained from a given DGS and its rotated version will not be the same due to the additional encoding created by the imaging gradients in different strengths for each scheme. The situation is more complicated for anisotropic samples, specifically for tissues with a geometric structure constraining the motion of the spins where the directional awareness is highly enhanced. Under these circumstances, if, for example, the sample orientation changes for any reason, e.g., during a longitudinal time study, the affected eigenvalues and eigenvectors will cause inconsistencies for the fiber tracking algorithms [20]. Moreover, depending on their alignment with respect to the imaging gradients, fine structure might become undetectable (see Fig. 4) in a detrimental manner, e.g., for tractography.

The essence of the optimization process is to design new DGSs with their cones (formed by their vectors, see Fig. 1 in Section II) oriented in such a manner that minimizes the effect of the imaging gradients. With this in hand, the optimal DGSs allow ignoring the imaging gradients in the estimation equations with higher fidelity and consequently addressing the modeling incongruities.

In Section II-A, the vector and operator norms that are used to define the optimization criterion for this purpose are presented. It is natural that the optimized DGS have vectors with different magnitudes favoring the directions that minimize and balance the effect of the imaging gradients. The factorization of the switching times and DGS vectors for the DTI estimation equations described in [14] and [15] provides that capability (see (29) and the discussion below it). In brief, with the new framework the DGS vectors are *not constrained to the unit sphere* as it is traditionally done in the existing DTI experiment designs. In consequence, the optimization presented in this manuscript is integral: there is no separate optimization of the vector magnitudes and directions as its done in [21]–[25].

However, the choice of the DGS vectors is not arbitrary. First of all, the newly designed DGS must make the coefficient matrix of the estimation equation full rank. Otherwise, the DTI experiment will collect redundant data leading to a set of underdetermined equations [14], [16] regardless of whether least squares or nonlinear estimation is used [17]. The optimization is therefore constrained to the feasible DGS space where the rank condition is guaranteed. The constraint is implemented by utilizing the linear parametrization of Section II-B [19]. Second, as there is a natural tendency in the optimization to enhance the diffusion sensitizing gradients to make the imaging gradients negligible, the DGS vectors must be prevented from breaching the hardware limits. The objective function in Section II-D includes a hard constraint that limits the magnitude of the magnetic field gradients.

Most important of all, the optimization problem presented in Section II is *sample independent*, quite different and novel than the existing optimization criterions [21]–[25]. It is founded on bounding the difference between the eigenvalues of diffusion matrix D calculated with and without including the imaging gradients in the estimation equations. Whereas the eigenvalues are uniquely dependent on the characteristics of the sample, the bound is solely dependent on coefficient matrices of the DTI estimation equations that are completely independent of the sample. As the eigenvalues are the building blocks of the anisotropy indices due their invariance under similarity transformations (e.g., rotations), bounding their difference to minimize their directional awareness is an effective choice for providing consistency. The advantage of the optimal DGSs is demonstrated with the fractional anisotropy index on a fixed baboon brain in Fig. 4 of Section IV.

As the coefficient matrices are functions of the pulse sequence parameters (e.g., the echo time) and imaging gradients defined by the slice FOV, orientation, position etc., the optimization problem possesses a different solution for different parameter values. Therefore each time a new experiment is conducted, the optimal DGS must be computed on-the-fly. The implementation is provided in Section III-B.

The optimization generates new DGSs by using the existing DGSs in the literature as the starting points for the numerical optimization routine. The newly generated DGSs are tested with the experimental setups of Section III that consist of a water tube and a fixed baboon brain. In Section IV, the results are analyzed and compared with the results from the nonoptimal schemes.

II. DEFINITION OF THE OPTIMIZATION PROBLEM

The linear algebraic framework derived in [14] and [15], which is summarized in Appendix A, formulates the governing DTI estimation equations by deriving explicitly the coefficient matrix V of the equations as shown in (31). The matrix is the sum of three matrices corresponding respectively to the diffusion gradients, imaging gradients and the cross terms between them

$$V=b_tV_g+V_I+V_{C(g)}. \quad (1)$$

The matrix V_I from (31) is for imaging gradients and is constant with equal rows. $V_D \doteq b_t V_g$ is conveniently factorized as a function of the gradient switching times, b_t from (33) and the matrix V_g defined in (30) as a function of the DGS denoted by g. $V_{C(g)}$ given in (32) represents the cross terms between the imaging and diffusion gradients. Since the imaging gradients are fixed for the experiment, V_C is also a function of g.

In the case where the imaging gradients are neglected the equations take the form

$$\gamma^2 V_D d=\gamma^2 b_t V_g d=p \quad (2)$$

where the entries of p are the logarithms of the measurements shown in (37), $d \in \mathbb{R}^6$ represents the six elements of the (symmetric) diffusion quadratic form D [see (17)] and γ is the gyromagnetic ratio.

A. Normed Space and the Error Bound

The central idea is to treat (1) as a matrix perturbation of V_D in (2): $V = V_D + (V_I + V_C)$ and use the following theorem.

Theorem 1 ([26]): Let $A \in \mathbb{R}^{m\times m}$ be nonsingular and let $\tilde{A} = A + E$ be a perturbation of A. For $p \in \mathbb{R}^m$ let $A\,d = p$ and let $\left\| \cdot \right\|$ be an operator norm. If there is a vector $\tilde{d}$ such that $\tilde{A}\,\tilde{d} = p$, then

$$\frac{\| d-\tilde{d} \|}{\| \tilde{d} \|} \le \| A^{-1}E \| . \quad (3)$$

The theorem provides an upper bound on the relative error between the entries of the solutions $\tilde{d}$ from (1), i.e., V, and d from V_D as in (2) using any vector norm and the corresponding operator norm

$$\frac{\| d-\tilde{d} \|}{\| \tilde{d} \|} \le \| V_D^{-1}(V_I+V_C) \| . \quad (4)$$

In this work, the Frobenius norm F is adopted due to the following property for a symmetric matrix D with eigenvalues $\lambda = [\lambda_1, \lambda_2, \lambda_3]$

$$F(D)=\left[\sum_i^3 \lambda_i^2\right]^{1/2} = \| \lambda \|_2 . \quad (5)$$

$F(D)$ is computed as a R-weighted norm for $d \in \mathbb{R}^6$

$$\begin{aligned} F(D) &= \| d \|_R \doteq \sqrt{d^T R\, d} \\ &= \sqrt{d_1^2+d_2^2+d_3^2+2(d_4^2+d_5^2+d_6^2)} \end{aligned}$$

with

$$R=\begin{bmatrix} I_3 & 0 \\ 0 & 2I_3 \end{bmatrix}$$

where I_3 is 3×3 identity matrix. It can be shown that the compatible operator Frobenius norm is given by

$$\begin{aligned} \| V \|_R &= \| R^{1/2} V R^{-1/2} \|_2 \\ &= \sqrt{r_\sigma(R^{-1/2} V^T R\, V R^{-1/2})} \end{aligned} \quad (6)$$

with r_σ denoting the spectral radius of a matrix. The Frobenius norm also possesses the following important property:

Theorem 2 ([26], p. 205): Let d and $\tilde{d} \in \mathbb{R}^6$ represent two symmetric $n \times n$ matrices with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_n$, respectively. Then

$$\left[\sum_{i}^{n}(\lambda_i - \tilde{\lambda}_i)^2\right]^{1/2} \leq \| d - \tilde{d} \|_R .$$

The (4), (5), and Theorem 2 are now combined to place a bound on the relative error of eigenvalues

$$\frac{\left[\sum_{i}^{3}(\lambda_i - \tilde{\lambda}_i)^2\right]^{1/2}}{\left[\sum_{i}^{3}\tilde{\lambda}_i^2\right]^{1/2}} \leq \| V_D^{-1}(V_I + V_C) \|_R . \quad (7)$$

The goal of the optimization algorithms is to find a DGS that will minimize the right-hand side of (7). This is a geometric search for schemes that define cones that are optimally “stretched” and “oriented” with respect to the fixed imaging gradient vectors (see Fig. 1).

The left-hand side of the inequality (7) is unknown before the measurements and reflects uniquely the sample’s diffusion properties. In contrast, the right-hand side is solely determined by the magnetic field gradient vectors and their activation times. Therefore, the upper bound is completely independent of measurement values and properties of the diffusion at any given location including the eigenvectors and eigenvalues. The bound guarantees the generation of optimal schemes that provide impartial experimental results for two different samples imaged with the same pulse sequence parameters. It is an extremely sound choice for an objective function that does not possess any sample dependence (tissue properties, orientation, location in the sample, etc.) of the existing optimization criterions [21]–[25].

The advantage of a formulation that separates the timings and the strengths of the gradients (as $V_D = b_t V_g$) presented in [14], [15] now becomes clear. When the gradient pulses are approximated by boxcar functions, b_t defined in [14] [different than the definition of *b*-value used in the literature, see (33)] is a factor for the on and off times of the gradient pulses and is independent of their strength. This prevents the restriction of optimizing the *b*-value and the direction of the diffusion gradient vectors separately, so it is more integral and sound. Each of the DGS vectors (i.e., each acquisition) will have its own *b*-value defined by the product of its magnitude and b_t. The search for an optimal solution, then, transforms into a true vectorial optimization geometrically balancing the effect of the imaging gradients. This is different and much more general than restricting the gradients to the unit sphere in order to use a single *b*-value for all of them.

B. Parametrization of the Feasible Space

Regardless of how the estimation is computed, with or without the incorporation of the imaging gradients, using linear or nonlinear methods [17], any DGS, g, must create a full rank V_g, (thus a full rank V_D) [16] in concert with the conditions of Theorem 1. In addition, the full version of Theorem 1 in [26] asserts that provided V_D is nonsingular, if $\| V_D^{-1}(V_I + V_C) \|$ is small, the sum of all three matrices, V, will also be nonsingular. Therefore the optimization problem must respect the full rank condition for V_g.

The constraint is satisfied by using the following parametrization of the set of admissible diffusion gradient vectors. It can be shown that when the DGS

$$g=\begin{bmatrix} g_{1x} & g_{1y} & g_{1z} \\ \vdots & \vdots & \vdots \\ g_{mx} & g_{my} & g_{mz} \end{bmatrix} \in \mathbb{R}^{m\times 3}$$

is transformed to $\bar{g} = g\,P$ with the 3 × 3 matrix P, V_g transforms via a matrix M, which is a function of P, to

$$V_{\bar{g}}=V_{(gP)}=V_g M. \quad (8)$$

The calculations are routine and use the linearity of h following from its definition (16):

$$h(G_{iD}\,P+G_I)=h(G_{iD})\,P+h(G_I). \quad (9)$$

The transformation obviously leaves V_I intact since the imaging gradients are not modified. V_g and V_C change because they are functions of the diffusion gradients. M is given in (12), shown at the bottom of the page, with the property that det(M) = (det(P))^4. This implies that det($V_{\bar{g}}$) = det(V_g) (det(P))^4. Hence, $V_{\bar{g}}$ will be nonsingular if and only if P is nonsingular. Thus, P is the parameter that describes the feasible gradient schemes. The original DGS, g, acts like a “pivot” and new admissible schemes are generated by changing P in the set of non-singular matrices. All of the schemes that can be obtained from a given pivot form a congruence class, with each member characterized by an element of the set of nonsingular matrices.

The feasible DGS set is characterized by the parametrization of the set of nonsingular matrices via polar decomposition: $P = U\,Q$. Here U is an orthogonal matrix (“angle”) and Q is a positive definite matrix (“magnitude”) [27]. In addition, Q is parametrized by $q \in \mathbb{R}^6$ using the Cholesky decomposition

$$Q=\widehat{Q}^T\,\widehat{Q}=\begin{bmatrix} q_1 & q_4 & q_6 \\ 0 & q_2 & q_5 \\ 0 & 0 & q_3 \end{bmatrix}^T \begin{bmatrix} q_1 & q_4 & q_6 \\ 0 & q_2 & q_5 \\ 0 & 0 & q_3 \end{bmatrix}.$$

In this work, the orthogonal part of the parametrization is confined to the set of rotation matrices using the Euler angles (ψ, θ, ϕ) ∈ [0, 2π] × [0, π] × [0, 2π] = Ω

$$U=\begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \times \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10)$$

However, the entirety of the feasible DGS space is not reachable starting from a single good DGS using the linear parametrization of (8). There are gradient schemes such that $\bar{g} = g\,P$ does not hold

$$\tilde{g}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \dot{g}=\begin{bmatrix} \frac{(1+\sqrt{5})}{2} & 1 & 0 \\ \frac{(1+\sqrt{5}}{2} & -1 & 0 \\ 0 & 1 & \frac{(1+\sqrt{5})}{2} \\ 0 & \frac{(-1-\sqrt{5})}{2} & -1 \\ -1 & 0 & \frac{(1+\sqrt{5})}{2} \\ 1 & 0 & \frac{(1+\sqrt{5})}{2} \end{bmatrix}. \tag{11}$$

$$M=\begin{bmatrix} P_{11}^2 & P_{12}^2 & P_{13}^2 & 2P_{11}P_{12} & 2P_{12}P_{13} & 2P_{11}P_{13} \\ P_{21}^2 & P_{22}^2 & P_{23}^2 & 2P_{21}P_{22} & 2P_{22}P_{23} & 2P_{21}P_{23} \\ P_{31}^2 & P_{32}^2 & P_{33}^2 & 2P_{31}P_{32} & 2P_{32}P_{33} & 2P_{31}P_{33} \\ P_{11}P_{21} & P_{12}P_{22} & P_{13}P_{23} & P_{12}P_{21}+P_{11}P_{22} & P_{13}P_{22}+P_{12}P_{23} & P_{13}P_{21}+P_{11}P_{23} \\ P_{21}P_{31} & P_{22}P_{32} & P_{23}P_{33} & P_{22}P_{31}+P_{21}P_{32} & P_{23}P_{32}+P_{22}P_{33} & P_{23}P_{31}+P_{21}P_{33} \\ P_{11}P_{31} & P_{12}P_{32} & P_{13}P_{33} & P_{12}P_{31}+P_{11}P_{32} & P_{13}P_{32}+P_{12}P_{33} & P_{13}P_{31}+P_{11}P_{33} \end{bmatrix}. \tag{12}$$

Although $M=V_{\dot{g}}^{-1}V_{\tilde{g}}$ exists, there is no P that provides M as in (12), shown at the bottom of the page. Consequently, the search is forcefully restricted to the congruence classes represented by different pivots. Since the classes are disjoint and disconnected, there is no way to find a global minimizer starting from a single pivot scheme with the linear parametrization. The strategy for covering the feasible DGS space is described in Section III-B. The investigation of more complicated parametrizations that will enable the discovery of a global optimum is the subject of future research.

C. Hardware Constraints

Intuitively speaking, if V_D has large entries, the effect of V_I and V_C will be reduced. Clearly, without any boundaries on the hardware limits, the solution of the minimization problems tends towards strong diffusion gradients in order to make V_D dominant. The main hardware constraint is the upper limit for the strength of the diffusion gradients, $G_{\max}$. The value cannot be higher than the maximum value imposed by the system. Alternatively, the user can set it at a lower level for satisfactory amplifier performance. The constraint, which is incorporated in the objective function below (13), guarantees that the limit is not violated in the search for the optimal DGSs.

D. Objective Function

At the heart of each optimization problem lays the choice of an adequate objective function. The analysis of the previous sections provided the foundation for including the minimization of the estimation differences in the objective function. The R-norm condition number of V_D is added for handling the perturbation of the measurements (e.g., due to noise) and for increasing the robustness and the numerical stability, which is also valid for nonlinear estimation

$$10 \parallel (b_t\, V_{(gP)})^{-1}(V_I + V_{C(gP)})\parallel_R + b_t^2 \parallel V_{(gP)}\parallel_R \parallel V_{(gP)}^{-1}\parallel_R + 100\left|\max_{i=1,\ldots,m} \parallel g_i P\parallel_1 - G_{\max}\right| . \quad (13)$$

The last portion defines the hardware constraint that confines the diffusion gradients to the cube $[-G_{\max}, G_{\max}]^3$. The largest weight, 100, was assigned to it to guarantee that the optimal DGS vectors have the largest allowed magnitudes to maximally attenuate the effect of the imaging gradients, which assists reaching the main goal with the second largest weight, 10. The choice of the weights was accomplished by trial and error while monitoring the magnitude of the DGS vectors after each trial. The degree of freedom in the weight assignment provides application-specific flexibility. The process relies both on the problem at hand, theoretical and experimental results, as well as on the experience of the designer. The investigation of weight determination methods is left as a future research topic.

The search is performed over the variables $(\psi, \theta, \phi, q) \in \Omega \times \mathbb{R}^6$. that parametrize P as in Section II-B.

III. EXPERIMENTAL SETUP, OPTIMIZATION, AND ANALYSIS

Although the main purpose of the DTI method is the discovery of microstructure, first and foremost, the estimation procedures must work properly for the simplest case of diffusion with known characteristics: an isotropic sample. For that reason, a polypropylene centrifuge tube by FisherBrand (Cat. No. 05-539-6) filled with tap water at room temperature, with an inner diameter at the slice of 2.7 cm was chosen as the phantom.

A fixed baboon brain immersed in 4% paraformaldehyde was used for testing the optimization procedures on biological tissue. The primate was prematurely delivered at 125th day and sacrificed at 59th day after delivery. All animal husbandry, handling, and procedures were performed at the Southwest Foundation for Biomedical Research, San Antonio, Texas. Animal handling and ethics were approved to conform to American Association for Accreditation of Laboratory Animal Care (AAALAC) guidelines. Further details of the sample preparation are described in [28].

A. Experimental Setup

The experiments were carried out on a 4.7T MR scanner (Varian NMR Systems, Palo Alto, CA) with a gradient system of bore size of 15 cm, maximum gradient strength of 45 gauss/cm and rise time of 0.2 ms using a quadrature birdcage coil (Varian NMR Systems, Palo Alto, CA) with 63 mm inner diameter size. DTI data were obtained using the standard spin-echo multislice sequence with in-house modifications that store all of the relevant parameters, including the timing and amplitudes of all of the crusher gradients. The images were 128 × 128 pixels with a field of view $64 \times 64\ \text{mm}^2$. The slice thickness was 1 mm for the water sample and 0.5 mm for the baboon brain. The repetition time $T_R = 1$ s, echo times $T_E = 35$ ms (water) $T_E = 30.5$ ms (baboon brain), diffusion pulse separation $\Delta = 18$ ms and diffusion pulse duration $\delta = 6$ ms were used.

Center-symmetric diffusion gradient schemes with 12 diffusion gradient vectors were used to obtain data. Initial nonoptimal DGSs were constructed by appending to the six vector gradient schemes [29] their central symmetric part: Tetrahedral (geometrically, the scheme is build on the four tetrahedral vectors of [30] by adding two more elements and is different from the scheme presented in [30]), Cond6, Jones noniso (without the last vector) renamed as Cond* because it yields to a V_g with a good condition number, Jones ($N = 6$), Muthupallai, (which is equal to Icosahedron (ICOSA6) scheme from [31]), Downhill Simplex Minimization (DSM) and Dual Gradient. Also a modified version of the

Muthupallai scheme, denoted by MUTM is included in the experiments. The DGSs are provided in Table V with the modified entries of MUTM in bold. Maximum diffusion gradient strengths of g_{diff} = 12 gauss/cm (water) and g_{diff} = 21 gauss/cm (baboon brain) were used. With boxcar approximation, at maximum diffusion gradient, the value of the scalar coefficient for the water and baboon brain were respectively: $\gamma^2 b_t\, g_{\text{diff}}^2 = 593.61\ \text{s/mm}^2$ and $\gamma^2 b_t\, g_{\text{diff}}^2 = 1817.94\ \text{s/mm}^2$. The *ex vivo* tissue experiments were carried out consecutively after leaving the sample in the scanner for approximately 12 h to reach a stable temperature.

B. Implementation of the Optimization

Optimal gradient schemes were generated using each of the diffusion gradient schemes listed in Section III-A, which are explicitly given in Table V, with six vectors as the pivot DGS g in the objective function of (13). Standard Matlab Optimization Toolbox (Mathworks, Natick, MA) routine *fmincon* was used to solve the optimization problems.

The initial condition supplied to the numerical routines plays a crucial role in the convergence of the algorithm. In this work, for each pivot, several initial conditions obtained by changing Euler angles in multiples of $\pi/4$ rad, equivalent to rotating the pivots, were evaluated. The "magnitude" of the pivot was left untouched having in mind from Section II-D that the vectorial approach deals with the strengths of the gradients automatically. After the algorithm converged for each initial condition of every pivot, which takes between 2.5 to 3.1 s with the computer hardware described in Section III-C, the best of the congruence class optimal solutions was selected as listed in Tables VI and VII.

It should be emphasized that the optimized schemes are not universal, they are specific to the parameters of the experiments: FOV, slice thickness, timing of the gradients, etc. For a new experiment, the optimization algorithm must be run to find the optimal solution corresponding to the new parameters.

Practically, the optimization is implemented in real time as described in the flowchart of Fig. 2. In the beginning, all of the pulse sequence parameters are obtained from the MR console based on the physical size, orientation and location of the slices, and the pulse sequence times (e.g., T_E), including the timing of the imaging gradients. V_I is calculated using these data. If a connected parametrization of the feasible set is discovered in future, the optimization will be executed using it. However, the current linear parametrization describes the feasible space as the union of disjoint and disconnected congruence classes. The optimization routines must run separately for each of the congruence classes. The larger the number of representative pivots (i.e., the number of classes), the larger will be the covered portion of the feasible space. For this reason, the congruence classes of interest are defined by a list of representative pivots. As the convergence of numerical optimization routines depends crucially on the initial conditions, different initial conditions are obtained by multiplying the representative pivot with an invertible matrix before running the optimization algorithms. Once the list of initial conditions is completed in a congruence class, the optimal solution that results in the minimum cost among all is chosen as the optimizer for that class. Afterwards, the next pivot in the database is subjected to the same procedure. The loop is repeated until the cost is found for each congruence class. Finally, the optimal solution with minimal cost between all of these is selected. The optimal DGS is provided as input to the MR scanner and the DTI experiment is executed.

The search can be sped up by keeping in a database the optimal solutions corresponding to sets of pulse sequences and imaging parameters. In the beginning of the experiment, after the operator chooses all of the parameters, the optimization starts from the initial condition equal to the optimal solution that corresponds to the parameters closest to those of the current experiment. If the parameter space is sampled densely enough, numerical routines will quickly convergence to the optimal solution.

C. Analysis of the Experimental Results

In-house Mathematica (Wolfram Research, Champaign, IL) code was used to compute components of V as described in [14], using the parameter values written to the console computer's hard disk by in-house additions to the pulse sequence. In-house Matlab (Mathworks, Natick, MA) programs were utilized for the estimation of d at each pixel, for graphical representation and maps of related results on a two quad core 2.3 GHz Intel Xeon cpu and 8 GB memory Dell Precision Workstation 490 running Windows XP 64-bit operating system. Standard Matlab Image Processing Toolbox routines, Sobel edge detection and morphological reconstruction were used to detect the signal region of the phantom in non-diffusion weighted images for each gradient scheme. The edges were removed to obtain a region free of susceptibility artifacts and the intersection of all regions was taken to obtain the circular area with 2022 pixels for the water sample and 6080 pixels for the baboon brain.

Among several estimation methods such as weighted and non-weighted least squares, total fitting [32] and nonlinear optimization (constrained and unconstrained), nonweighted least squares is chosen because of its speed and simplicity. The nonlinear estimation is investigated in [17]. In the computation of V_C and V_I the phase-encoding gradient value of 0 was selected as described in [14].

IV. ANALYSIS RESULTS

Table I summarizes the numerical values of the cost function (13) at different stages of the optimization procedure: The pivot DGSs are the ones listed in Section III-A, the initial condition is found by rotating the pivots as explained in Section III-B and the last row is the optimal solution. For Muthupallai, DSM, Dual Gradient and Tetrahedral, the original DGS as the initial condition has resulted in the optimal cost as shown in the appropriate columns of Table I. The optimal scheme initiated from Jones6 had the lowest cost for both of the samples.

Due to different imaging parameters, the optimization yielded different DGSs for the water sample and baboon brain. The center symmetric counterparts of the DGSs listed in Tables VI and VII are appended to them to obtain the NoCroT schemes (see [14], [15], the Appendix A–C). The new twelve gradient DGSs are used in the experiments respectively on both samples.

There are three levels for the inclusion of the imaging gradients in the estimation equations: "all gradients" (V_{DC}) as in (39), NoCroT as in (42) and "diffusion gradients only" as in (2). Since NoCroT becomes equivalent to neglecting all of the imaging gradients in the case of center-symmetric DGSs [see (42)], the estimation using the experimental data was carried out with two methods: "all gradients" and NoCroT (equivalently V_D).

The main goal of the optimization defined in (7) is the minimization of the differences between the eigenvalues ($\Delta\lambda_i = \lambda_i - \tilde{\lambda}_i$) estimated by excluding and including the imaging gradients (λ_i and $\tilde{\lambda}_i$) as in (2) and as in (40), respectively. Table II shows that the eigenvalue difference obtained from the nonoptimal DGSs ($(\Delta\lambda_i)_{\mathrm{Non}}$) is higher than the ones from the optimal DGSs ($(\Delta\lambda_i)_{\mathrm{Opt}}$), for the water sample and the overall baboon brain including individually the region of interests (ROIs) displayed in Fig. 3. This observation holds regardless of the ROIs that are specifically chosen to be physically near each other and to depict different levels of microstructure: cerebrospinal fluid (CSF) is similar to an isotropic sample, white matter (WM) possess structure and corpus callosum (CC) is highly organized. In short, the optimization goal has been reached.

Aside from the specific goal, it is also important to judge the overall performance of the newly designed DGSs. As DTI is a model matching method, the most relevant performance measure is the model matching error, the difference between the model S of (19) and the m measurement values $\hat{S}_i$

$$\|\chi\|_2^2=\frac{1}{m}\sum_{i=1}^{m}(S_i-\widehat{S}_i)^2. \quad (14)$$

$\bar{\chi}$ denotes the mean of the pixel residuals [14] from the signal region. Table III shows that the model matching error is reduced by using the optimized DGS for both samples, including different structures of the baboon brain. For example, in the isotropic sample the error is reduced by 46.96% for the dual gradient scheme, by 28.97%, 14.15%, and 16.46% in CSF, WM, and CC respectively for the Cond* scheme. These results are highly significant. But despite the success, the model matching error for “all gradients” method is larger compared to the less “complete” model NoCroT.

The eigenvalues obtained from the optimal DGSs are smaller compared to their non-optimal counterparts. For most of the optimal DGSs the standard deviation of the eigenvalues is also lower up to 7% (with the exception of Cond6 and Modified Muthupallai with 43% and 55%, respectively). In other words, the precision for the optimal DGSs is increased.

In addition, in [15] and [14] the negative eigenvalues of the water sample and CC estimated with “all gradients” for Tetrahedron DGS completely disappear for the optimal version. The negative eigenvalues are physically meaningless because their existence violates the Fick’s law of diffusion. In the case of the overall baboon brain sample, all of the nonoptimal DGSs, except DSM and Jones6, exhibit negative eigenvalues. The number of the negative eigenvalues are either eliminated (Dual Gradient and Muthupallai) or reduced by up to 70% with the optimal DGSs. In the CSF, nonoptimal Cond6 and Tetrahedron estimate negative eigenvalues whose numbers are reduced by up to 80% with the optimal schemes and in the WM they are completely eliminated. These observations demonstrate that the newly obtained DGSs are more robust than the DGSs in the literature.

One of the issues that motivated this research was to overcome the indeterminateness of the phase-encoding gradient values in the estimation equations, which is briefly summarized in Appendix A–C. As first noted in [14] and [15], the eigenvectors of the water sample tend to align themselves according to the value of the phase-encode gradient used in the coefficient matrices (V_I and V_C) for the nonoptimal schemes. The observation persists for nonlinear estimation [17] as well. The experimental results of the water sample show that the bias exists also for the optimal schemes for both least squares and nonlinear estimation [17]. It can only be concluded that the problem originates from a fundamental level of modeling that is beyond the scope of this manuscript.

Since the only physically known fact is the isotropic nature of the water sample, the only way to assess the accuracy performance is to investigate the anisotropy index fractional anisotropy (FA). Table IV summarizes the percent change of the mean value of FA ($\overline{FA}$) between the nonoptimal and optimal schemes. For the optimal DGSs, $\overline{FA}$ is lowered (with the exception of Cond*) for “all gradients” estimation up to 57.54%. However, this consistency is not observed for NoCroT: Cond6 and Modified Muthupallai report lowering of $\overline{FA}$ (up to 23.65%) but Dual Gradient and Tetrahedron DGSs show a significant increase (up to 48.3%) while the other schemes have small increases (up to 6.8%). Like the model matching error, $\overline{FA}$ estimated by V is higher than the one from NoCroT. This is an indication of the deficiency of a complete DTI model [14].

For the *ex vivo* sample, the true FA values are unknown but the FA maps of Fig. 4 present a great deal of information about the anatomical structure with known characteristics. The optimal DGS elucidates the subcortical U-fibers, which are not visible with the nonoptimal DGS. Also, the separation of the adjacent gyri is much better with the optimal scheme. Fig. 5 shows that the minimization of the eigenvalue difference, which is the main aim of the optimization criterion, reflects well on the fractional anisotropy. The relative FA difference for optimal DGS is lower and more homogenous over the baboon brain than the one from the nonoptimal version.

V. CONCLUSION AND FUTURE DIRECTIONS

As it is summarized in the Appendix A–C, [14] and [15] show that the DTI model equations cannot be resolutely written when the imaging gradients are needed to be included to obtain a complete model. The following, which is realized in [14], is one of the motivators of this work:

> The decrease of the accuracy for the inclusion of all of the gradients validates the existence of modeling incongruities in Diffusion Tensor Imaging model.

The profound challenge is to overcome DTI's theoretical inability to provide the account of the imaging gradients for a complete model, using the best available theoretical information within the model itself. Motivated by the lack of an exact description, this manuscript created a new and robust method for designing appropriate DGSs that minimize the effect of the imaging gradients in the DTI equations by positioning and orienting the cone defined by the diffusion gradient vectors so that the imaging gradients appear not be in place. The goal, which was initiated by the usage of the center symmetric DGSs in [14] and [15], also encompasses the elimination of the directional awareness created by the imaging gradients.

Although the optimization objective-bringing closer the eigenvalues estimated with and without the inclusion of the imaging gradients-was achieved with the bonus of model matching error reduction for the optimal DGSs, for the water sample the more complete model, "all gradients," V, still reports higher model matching error and higher anisotropy than the calculations that ignore the imaging gradients, i.e., NoCroT. In other words despite quantitative improvements with the optimized DGSs, qualitatively the behavior of the DTI model remains the same whether optimal or nonoptimal DGSs are used.

The model matching error and water sample's anisotropy is consistently lowered for estimation with V throughout the optimal DGSs. This a benefit of the optimization procedure for "all gradients." But the increase in $\overline{FA}$ for the majority of the optimal DGSs in NoCroT despite the consistent reduction of model matching error and attained optimization objective, is a reflection of modeling incongruities of the DTI method. Moreover, these observations are also valid for nonlinear estimation [17], fortifying the last remark. Therefore, the DTI model is not powerful enough to capture adequately the essence of the diffusion-weighted MRI signal.

Despite these deficiencies, the simplicity of DTI method makes it still a useful protocol. The optimization method introduced in this manuscript improved the quality of the information obtained by DTI: The superior anatomical resolution provided by the optimal DGS originates from the elimination of the directional awareness and establishes a stronger foundation for the fiber tracking algorithms (see Fig. 4). Also, the optimal DGS minimized the effect of the modeling incongruity by bringing closer the results obtained with and without the imaging gradients for the eigenvalues and consequently for the FA map (see Fig. 5).

In brief, the optimization problem furnished valuable tools to maximize the usefulness of the DTI method.

In this work, the optimization problem was posed for DGSs with six vectors mainly because Theorem 1 and 2 are valid for square matrices. The addition of the six center symmetric counterparts is for taking advantage of NoCroT method. The twelve vector DGSs force the usage of least squares estimation rather than matrix inversion. For acquisitions with arbitrary number of DGS vectors ($m > 6$), a different and more complicated perturbation theorem [33] concerning the least squares estimation comes into play. Further investigations of this subject are left for future research with important considerations for clinical practice. For example, the adoption of the manuscript’s optimization framework can significantly influence the results of [34] and [35] where the organization and acquisition order of DGSs are arranged to minimize the impact of partial scans, e.g., due to patient motion, for salvaging efficiently the results that would otherwise be lost.

Acknowledgments

This work was supported in part by the Washington University Small Animal Imaging Resource, a National Cancer Institute funded Small Animal Imaging Resource Program facility (U24-CA83060) and in part by the National Institutes of Health/National Institute of Neurological Disorders and Stroke (NIH/NINDS) grant Biomarkers and Pathogenesis of MS (P01-NS059560).

APPENDIX A

BRIEF DERIVATION OF THE DTI ESTIMATION EQUATIONS IN THE LINEAR ALGEBRAIC FRAMEWORK

In this section, the derivation of the coefficient matrix V that is presented in full details in [14] is concisely reviewed.

A) The Estimation Equation

The solution of Stejskal and Tanner for the NMR pulsed-gradient spin-echo (PGSE) experiments [10] has been directly translated to MR-DTI experiments in [11] and [12]. The translation in [11], (3), is written here in its most general form by replacing the spin-echo image S_0 with S_p that is related to the proton density

$$S(G,t)=S_p\exp\left(-\gamma^2\int_0^t h(G,\zeta)\,D\,[h(G,\zeta)]^T d\zeta\right) \quad (15)$$

where $G(\cdot) = [G_x(\cdot)\ G_y(\cdot)\ G_z(\cdot)]$ denotes the time course of the magnetic field gradient vector, γ is the gyromagnetic ratio, S denotes the signal intensity at each pixel. The initial time, 0, is taken in this manuscript as the time “immediately following the 90° pulse” [36]. The time t is chosen to be T_E, the echo time. The analytic form of the vector valued function h for the PGSE imaging experiment is derived by incorporating the unit step function for the sign change created by the π-pulse

$$h(G,\zeta)=\int_0^\zeta G(\xi)d\xi-2u(\zeta-\tau)\int_0^\tau G(\xi)d\xi \quad (16)$$

where τ is the time of the π-pulse and $u(\cdot)$ denotes the unit step function.

By the appearance of D as a quadratic from acting on h in (15), D is a symmetric 3 × 3 matrix that is represented as a vector $d \in \mathbb{R}^6$

$$d = [d_1 \ \cdots \ d_6]^T \mapsto D = \begin{bmatrix} d_1 & d_4 & d_6 \\ d_4 & d_2 & d_5 \\ d_6 & d_5 & d_3 \end{bmatrix}. \quad (17)$$

For a fixed $h \in \mathbb{R}^3$, the argument of the exponential function in (15) becomes a linear function of D and therefore it is expressed as the product of a row vector υ_i and d

$$\upsilon_i\, d = \int_0^t h(G_i, \zeta) D [h(G_i, \zeta)]^T d\zeta \quad (18)$$

where the subscript i corresponds to the acquisition with the i^{th} diffusion gradient vector. Equation (15) transforms to

$$S(\upsilon_i\, d) \doteq S_p \exp(-\gamma^2\, \upsilon_i\, d) = S_i \quad (19)$$

and after the application the logarithm, the linear estimation equation is obtained

$$\gamma^2 \upsilon_i d = -\ln\left(\frac{S_i}{S_p}\right) = \ln S_p - \ln S_i. \quad (20)$$

Here, the substraction on the right is preferred to emphasize the reduction of propagation of errors in numerical implementation. Stacking the row vectors υ_i to form a matrix V and using m, measurements to form a vector

$$p_r = [\ln(S_p) - \ln(S_1), \ \ldots, \ \ln(S_p) - \ln(S_m)]^T \quad (21)$$

the linear equations for the estimation of the diffusion matrix are written as

$$\gamma^2\, V\, d = p_r. \quad (22)$$

This expression is the *true governing equation* for DTI experiments and consequently the optimization problem is posed based on the coefficient matrix V.

Ideally, in order to obtain an impartial observation, S_p ought to be obtained from an imaging experiment other than DTI. Nevertheless, S_p can be estimated by adding another (minimum seventh) acquisition

$$\begin{aligned} \begin{bmatrix} \upsilon_1 & -1 \\ \vdots & \vdots \\ \upsilon_7 & -1 \end{bmatrix} \begin{bmatrix} d \\ \frac{\ln(S_p)}{\gamma^2} \end{bmatrix} &= [V \ -\mathbb{1}] \begin{bmatrix} d \\ \frac{\ln(S_p)}{\gamma^2} \end{bmatrix} \\ &= -\frac{1}{\gamma^2} \begin{bmatrix} \ln(S_1) \\ \vdots \\ \ln(S_7) \end{bmatrix}. \end{aligned} \quad (23)$$

The augmented coefficient matrix uses V rather than V_D that is used by Papadakis *et al.* [32] in the method called *total fitting*.

However, in practice, the image obtained without any diffusion gradients, S_0, is used to replace S_p as described in the Appendix A–C.

B) Components of V

The magnetic gradient vector G, is the sum of two components corresponding to the diffusion sensitizing gradients and to the imaging gradients

$$G_i=G_{iD}+G_I.$$

Using this decomposition and the linearity of h from (16)

$$h(G_D+G_I)=h(G_D)+h(G_I)$$

V is formulated in (31) with three components corresponding to diffusion gradients V_D, imaging gradients V_I, and the cross terms between them, V_C. V is calculated explicitly by calculating each of the components as follows.

Although the vector G_{iD} changes at each acquisition, G_I defined by the MR slice parameters (slice thickness, field of view) is fixed. Without loss of generality, the coordinate frame is taken as the frame defined by the read-out ($x = ro$), phase-encode ($y = pe$), and slice select ($z = ss$) axes

$$G_I(\xi)=[\beta_{ro}(\xi)\beta_{pe}(\xi)\beta_{ss}(\xi)] \quad (24)$$

where each β describes the on-off times and the strength of each gradient channel as shown in Fig. 6. Similarly, G_{iD} is written as a product of a scalar function and of a vector (not necessarily of unit norm)

$$G_{iD}(\xi)=\beta_D(\xi)g_i. \quad (25)$$

The calculation in (16) is systematically carried for different gradient types using the scalar valued integral

$$\mu_*(\zeta)=\int_0^{\zeta}\beta_*(\xi)d\xi-2u(\zeta-\tau)\int_0^{\tau}\beta_*(\xi)d\xi \quad (26)$$

to obtain from (16)

$$h(G_*,\zeta)=\mu_*(\zeta)g_*=\mu_*(\zeta)\,[g_{*x}g_{*y}g_{*z}]. \quad (27)$$

The asterisk can be any of the labels *ro, pe, ss* for imaging or D for diffusion gradients.

The calculation of V_I in (31) only requires one row vector

$$\int_0^t [\mu_{ro}^2\mu_{pe}^2\mu_{ss}^2 2\mu_{ro}\mu_{pe} 2\mu_{pe}\mu_{ss} 2\mu_{ro}\mu_{ss}]\,d\zeta. \quad (28)$$

On the other hand V_D is calculated as the product of a scalar and a matrix

$$V_D=b_t\,V_g \quad (29)$$

where V_g is the $m \times 6$ matrix

$$\begin{bmatrix} g_{1x}^2 & g_{1y}^2 & g_{1z}^2 & 2g_{1x}g_{1y} & 2g_{1y}g_{1z} & 2g_{1x}g_{1z} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{mx}^2 & g_{my}^2 & g_{mz}^2 & 2g_{mx}g_{my} & 2g_{my}g_{mz} & 2g_{mx}g_{mz} \end{bmatrix} \quad (30)$$

$$\begin{aligned} V= & \int_0^t \begin{bmatrix} h_x^2(G_{1D}) & \cdots & 2h_x(G_{1D})h_z(G_{1D}) \\ \vdots & & \vdots \\ h_x^2(G_{mD}) & \cdots & 2h_x(G_{mD})h_z(G_{mD}) \end{bmatrix} d\zeta + \int_0^t \begin{bmatrix} h_x^2(G_I) & \cdots & 2h_x(G_I)h_z(G_I) \\ \vdots & & \vdots \\ h_x^2(G_I) & \cdots & 2h_x(G_I)h_z(G_I) \end{bmatrix} d\zeta + 2\int_0^t \begin{bmatrix} h_x(G_{1D})h_x(G_I) & \cdots & h_x(G_{1D})h_z(G_I)+ \\ \vdots & & \vdots \\ h_x(G_{mD})h_x(G_I) & \cdots & h_x(G_{mD})h_z(G_I)+ \end{bmatrix} \\ & = V_D+V_I+V_C=b_t V_g+V_I+V_{C(g)}. \end{aligned}$$

$$V_C=2\int_0^t \mu_D \begin{bmatrix} \mu_{ro}g_{1x} & \mu_{pe}g_{1y} & \mu_{ss}g_{1z} & (\mu_{pe}g_{1x}+\mu_{ro}g_{1y}) & (\mu_{pe}g_{1y}+\mu_{ss}g_{1z}) & (\mu_{ss}g_{1x}+\mu_{ro}g_{1z}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{ro}g_{mx} & \mu_{pe}g_{my} & \mu_{ss}g_{mz} & (\mu_{pe}g_{mx}+\mu_{ro}g_{my}) & (\mu_{pe}g_{my}+\mu_{ss}g_{mz}) & (\mu_{ss}g_{mx}+\mu_{ro}g_{mz}) \end{bmatrix} d\zeta. \quad (32)$$

(see (31) and (32) at the bottom of the page) and b_t is calculated using (26) and (27):

$$\begin{aligned} b_t & =\int_0^t \left(\int_0^\zeta \beta_D(\xi)\,d\xi - 2u(\zeta-\tau)\int_0^\tau \beta_D(\xi)d\xi\right)^2 d\zeta \\ & =\int_0^t \mu_D^2(\zeta)d\zeta. \end{aligned} \quad (33)$$

For example, the scalar factor of rectangular diffusion gradient pulses is $b_t = \delta^2 (\Delta - \delta/3)$ and for trapezoidal pulses it is $\delta^2(\Delta - \delta/3) - \delta\, t_{\text{rise}}^2/6 + t_{\text{rise}}^3/30$ where δ is the length of the pulses, Δ is the time between them, and t_{rise} is the time for the gradients to reach a specified value.

Note that b_t involves only the switching times of the gradients, and the b-value given in the literature is equal to the product of $\gamma^2 b_t \parallel g_i \|_2^2$. The factorization of V_D in (29) with a pure function of switching times, b_t, and V_g that is a function of the diffusion gradient vectors is the key factor for the definition and the implementation of the geometric optimization problem (13) in Section II-D.

The cross terms matrix, $V_{C(g)}$ is calculated as a function of the diffusion gradient scheme g as shown in (32). Finally, (22) is written explicitly as

$$\gamma^2(b_t\, V_g+V_I+V_{C(g)})d=p_r. \quad (34)$$

C) Estimation With the Spin-Echo Image S_0

The reference image obtained without diffusion gradients, S_0, is obtained by turning off the diffusion gradients resulting in the row vector υ_I of (28) in for the calculation of (18)

$$S_0=S_p \exp(-\gamma^2 \upsilon_I d). \quad (35)$$

Since S_p is not directly available, the usual approach is to replace it by writing (19) as

$$S(\upsilon_i\, d) \doteq S_0 \exp(-\gamma^2 (\upsilon_i - \upsilon_I)\, d)=S_i. \quad (36)$$

Applying the logarithm to this equation yields

$$\gamma^2 (\upsilon_i - \upsilon_I) d = -\ln\left(\frac{S_i}{S_0}\right) = \ln S_0 - \ln S_i. \quad (37)$$

With

$$p = [\ln(S_0) - \ln(S_1), \ldots, \ln(S_0) - \ln(S_m)]^T \quad (38)$$

the linear equations are written as

$$\gamma^2 (V - V_I) d = \gamma^2 V_{DC} d = p. \quad (39)$$

D) No Cross Terms Estimation: NoCroT

The main difference between the diffusion-weighted NMR and MRI experiments is that in the former the estimation is based on a single point-the peak of the signal-whereas the DTI model uses the pixel intensity (15) and (19) that originates from the entirety of the *k*-space lines.

The read-out and the slice select gradients, and therefore β_{ro} and β_{ss} in (24), have fixed values in Fig. 6 and in all of the computations of Appendix A–B. However, the phase-encode gradient, thus β_{pe}, changes at each *k*-space line. It is, therefore, not clear which value of the phase-encoding gradient to utilize in the estimation equations. When the equations are solved for different phase-encode gradient strengths, the eigenvectors align with the direction dictated by those strengths [14] regardless of whether least squares or nonlinear estimation [17] is utilized. Taking the strength to be equal to zero is equivalent to completely neglecting the phase-encode gradient which defeats the purpose of deriving a complete model.

Moreover, because the imaging gradients do not possess null moments, the integral of μ_{ro}, μ_{pe} and μ_{ss} do not form a plateau around T_E in (15). In consequence [14], each of the *k*-space sample points collected during the readout acquisition period is differently diffusion sensitized. In other words, $V_{C(g)}$ in (32) and V_I in (28) and (31) are in fact not well defined.

The resolution of this indeterminateness proposed in [14], [15] relies on the properties of the coefficient matrices originating from (30) and (32): when the sign of the DGS is changed, V_D remains the same, because $V_g = V_{(-g)}$ while the cross terms matrix changes sign, $V_{C(-g)} = -V_{C(g)}$. Experiments performed with center-symmetric DGSs, i.e., with $(g, -g) = (g_1, \ldots, g_{m/2}, -g_1, \ldots, -g_{m/2})$ result in

$$\gamma^2 \begin{bmatrix} b_t V_g + V_{C(g)} \\ b_t V_{(-g)} + V_{C(-g)} \end{bmatrix} d = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}. \quad (40)$$

The sum of the equations in (40) yield

$$\begin{aligned} 2\gamma^2 b_t V_g d &= \begin{bmatrix} 2\ln S_0 - \ln S_1 - \ln S_{(m/2+1)} \\ \vdots \\ 2\ln S_0 - \ln S_{m/2} - \ln S_m \end{bmatrix} \\ &= p_1 + p_2 (\text{NoCroT}). \end{aligned} \quad (41)$$

This is equivalent to solving the following set of inconsistent equations:

$$\gamma^2 \begin{bmatrix} V_D \\ V_D \end{bmatrix} d = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \quad (42)$$

by least squares estimation. Hence, the center-symmetric gradient schemes guarantee that the NoCroT portion of the complete set of DTI estimation equations is equivalent to completely neglecting the imaging gradients from the equations.

The cross terms matrix, V_C, does not appear in (41) or (42), hence the name "No Cross Terms:" NoCroT is adopted. The central assumption behind the NoCroT strategy is that irrespective of the true nature of the indeterminateness, the sign change in the DGS causes a sign change in V_C. The correct way to choose the center-symmetric diffusion gradients is to first select 6 vectors that will ensure the full rank condition on V_g [16], [18] and then add their center-symmetric counterparts to obtain $V_{(g,-g)}$ with full rank.

APPENDIX B

SUMMARY OF THE RESULTS

In this section, Tables II–IV are presented.

APPENDIX C

GRADIENT SCHEMES USED IN THE EXPERIMENTS

In this section, Tables V–VII are presented.

REFERENCES

1. Song S-K, Sun S-W, Ramsbottom MJ, Chang C, Russell J, Cross AH. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. NeuroImage. 2002 Nov.vol. 17(no. 3):1429–1436. [PubMed: 12414282]
2. Budde MD, Kim JH, Liang H-F, Schmidt RE, Russell JH, Cross AH, Song S-K. Toward accurate diagnosis of white matter pathology using diffusion tensor imaging. Magn. Reson. Med. 2007 Apr.vol. 57(no. 4):688–695. [PubMed: 17390365]
3. Xu J, Humphrey PA, Kibel AS, Snyder AZ, Narra VR, Ackerman JJ, Song S-K. Magnetic resonance diffusion characteristics of histologically defined prostate cancer in humans. Magn. Reson. Med. 2009 Apr.vol. 61(no. 4):842–850. [PubMed: 19215051]
4. Mielke M, Kozauer N, Chan K, George M, Toroney J, Zerrate M, Bandeen-Roche K, Wang M-C, Zijl Pv, Pekar J, Mori S, Lyketso C, Albert M. Regionally-specific diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. NeuroImage. 2009 May; vol. 46(no. 1):47–55. [PubMed: 19457371]
5. Kim JH, Loy DN, Liang H-F, Trinkaus K, Schmidt RE, Song S-K. Noninvasive diffusion tensor imaging of evolving white matter pathology in a mouse model of acute spinal cord injury. Magn. Reson. Med. 2007 Aug.vol. 58(no. 2):253–260. [PubMed: 17654597]
6. Conturo TE, Lori NF, Cull TS, Akbudak E, Snyder AZ, Shimony JS, McKinstry RC, Burton H, Raichle ME. Tracking neuronal fiber pathways in the living human brain. PNAS. 1999 Aug.vol. 96(no. 18):10422–10427. [PubMed: 10468624]
7. Xue R, van Zijl PC, Crain BJ, Solaiyappan M, Mori S. In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging. Magn. Reson. Med. 1999; vol. 42(no. 6): 1123–1127. [PubMed: 10571934]
8. Mori S, van Zijl PCM. Fiber tracking: Principles and strategies—A technical review. NMR Biomed. 2002 Nov.vol. 15(no. 7–8):468–480. [PubMed: 12489096]
9. Torrey H. Bloch equations with diffusion terms. Phys. Rev. 1956 Nov.vol. 104(no. 3):563–565.

10. Stejskal EO, Tanner J. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field. J. Chem. Phys. 1965 Jan.vol. 42(no. 1):288–292.
11. Basser PJ, Matiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. J. Magn. Reson. Series B. 1994; vol. 103(no. 3):247–254.
12. Matiello J, Basser PJ, LeBihan D. Analytical expressions for the B matrix in NMR diffusion imaging and spectroscopy. J. Magn. Reson. Series A. 1994; vol. 108(no. 2):131–141.
13. Papadakis NA, Xing DC, Huang L-H, Hall LD, Carpenter TA. A comparative study of acquisition schemes for diffusion tensor imaging using MRI. J. Magn. Reson. 1999; vol. 137(no. 1):67–82. [PubMed: 10053134]
14. Özcan A. Characterization of imaging gradients in diffusion tensor imaging. J. Magn. Reson. 2010 Nov.vol. 207(no. 1):24–33.
15. Özcan, A. Theoretical and experimental analysis of imaging gradients in DTI; Proc. 31st Annu. Int. Conf. IEEE EMB Soc.; Sep. 2009; Minneapolis, MN. p. 2703-2706.
16. Özcan A. Mathematical necessary conditions for the selection of gradient vectors in DTI. J. Magn. Reson. 2005 Feb.vol. 172(no. 2):238–241.
17. Özcan A. Noise, nonlinear estimation and gradient vector selection in DTI. Magn. Reson. Imag. 2010 Nov.vol. 28(no. 9):1335–1343.
18. Rathore, R. Necessary and sufficient conditions for the admissibility of DTI gradient vectors; Proc. Int. Soc. Mag. Reson. Med.; May 2007; Berlin, Germany. p. 1509-1509.
19. Özcan, A. Decoupling of imaging and diffusion gradients in DTI; Proc. 31st Annu. Int. Conf. IEEE EMB Soc.; Sep. 2009; Minneapolis, MN. p. 2707-2710.
20. Landman BA, Farrell JA, Jones CK, Smith SA, Prince JL, Mori S. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. NeuroImage. 2007 Jul.vol. 36(no. 4):1123–1138. [PubMed: 17532649]
21. Xing D, Papadakis NG, Huang CL-H, Lee VM, Carpenter TA, Hall LD. Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. Magn. Reson. Imag. 1997; vol. 15(no. 7):771–784.
22. Bastin ME, Armitage PA, Marshall I. A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging. Magn. Reson. Imag. 1998; vol. 16(no. 7):773–785.
23. Armitage PA, Bastin ME. Utilizing the diffusion-to-noise ratio to optimize magnetic resonance diffusion tensor acquisition strategies for improving measurements of diffusion anisotropy. Magn. Reson. Med. 2001; vol. 45(no. 6):1056–1065. [PubMed: 11378884]
24. Kingsley PB, Monahan WG. Selection of the optimum b factor for diffusion-weighted magnetic resonance imaging assesment of ischemic stroke. Magn. Reson. Med. 2004; vol. 51(no.5):996–1001. [PubMed: 15122682]
25. Alexander DC, Barker GJ. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. Neuroimage. 2005; vol. 27(no. 2):357–367. [PubMed: 15921931]
26. Stewart, GW.; Sun, J. Matrix Perturbation Theory. First ed.. Academic Press; 1990.
27. Gallier, J. Geometric Methods and Applications: For Computer Science and Engineering. 1st ed.. New York: Springer Verlag; 2000.
28. Kroenke CD, Bretthorst GL, Inder TE, Neil JJ. Diffusion MR imaging characteristics of the developing primate brain. NeuroImage. 2005 May; vol. 25(no. 4):1205–1213. [PubMed: 15850738]
29. Skare S, Hedehus M, Moseley ME, Li T-Q. Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. J. Magn. Reson. 2000; vol. 147(no. 2): 340–352. [PubMed: 11097823]
30. Conturo TE, McKinstry RC, Akbudak E, Robinson BH. Encoding of anisotropic diffusion with tetrahedral gradients: A general mathematical diffusion formalism and experimental results. Magn. Reson. Med. 1996; vol. 35(no. 3):399–412. [PubMed: 8699953]
31. Hasan KM, Parker DL, Alexander AL. Comparison of gradient encoding schemes for diffusion-tensor MRI. J. Magn. Reson. Imag. 2001; vol. 13(no. 5):769–780.

32. Papadakis NA, Murrils CD, Hall LDC, Huang L-H, Carpenter TA. Minimal gradient encoding for robust estimation of diffusion anisotropy. Magn. Reson. Imag. 2000; vol. 18(no. 6):671–679.
33. Golub, GH.; Loan, CFV. Matrix Computations. 3rd ed.. Baltimore, MD: Johns Hopkins Univ. Press; 1996.
34. Dubois J, Poupon C, Lethimonnier F, Bihan DL. Optimized diffusion gradient orientation schemes for corrupted clinical DTI data sets. Magn. Reson. Mater. Phys., Biolo. Med. 2006; vol. 19(no. 3): 134–143.
35. Cook PA, Symms M, Boulby PA, Alexander DC. Optimal acquisition orders of diffusion-weighted MRI measurements. J. Magn. Resona. Imag. May; 2007 vol. 25(no. 5):1051–1058.
36. Stejskal EO. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic restricted diffusion and flow. J. Chem. Phys. 1965 Nov.vol. 43(no. 10):3597–3603.

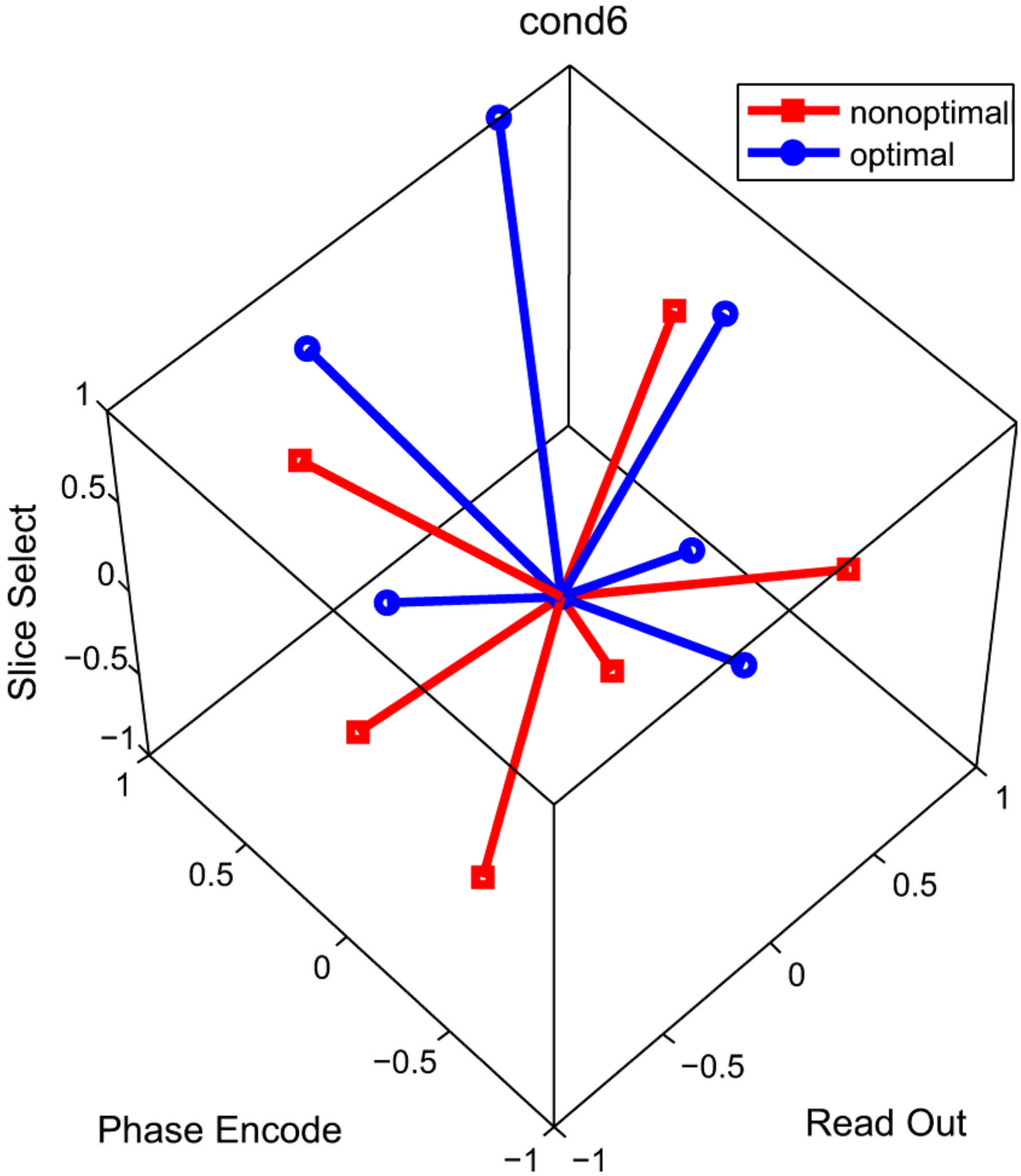

Fig. 1.

Cond6 DGS and its optimized version (see Tables V and VI). Notice that the cones are different but the schemes are not completely antipodal to each other. In the optimization problem the gradients are confined to the cube $[-1, 1]^3$ to respect the hardware limitations.

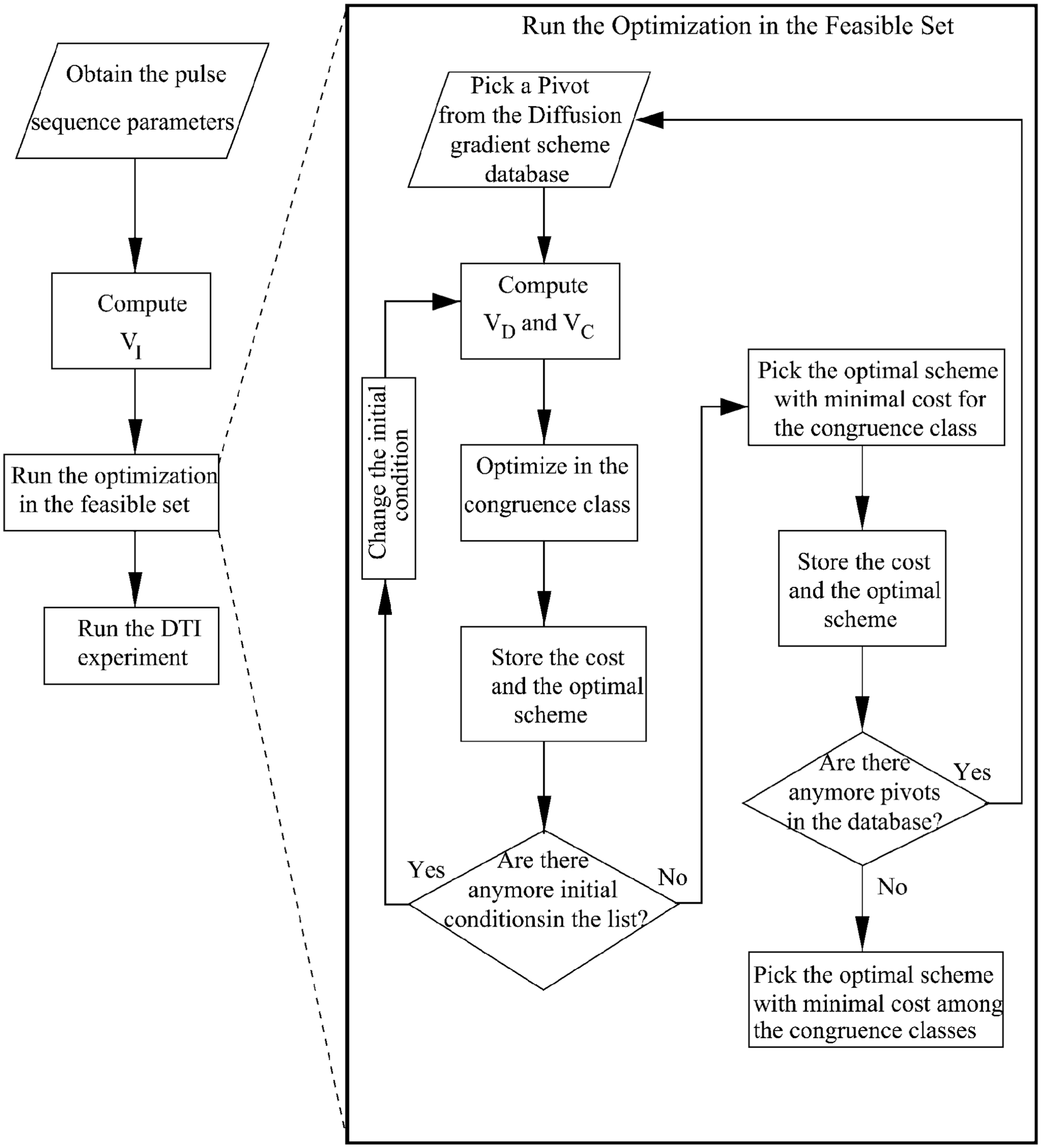

Fig. 2.

Flowchart for implementing the optimization procedure in real time during DTI experiments. On the left is the overall setup. On the right is description of the algorithm inside the box "Run the Optimization" for the case of current linear parametrization of the feasible space that can only describe disconnected congruence classes. If the feasible space can be parametrized in a connected fashion, the search in disconnected sets will not be necessary.

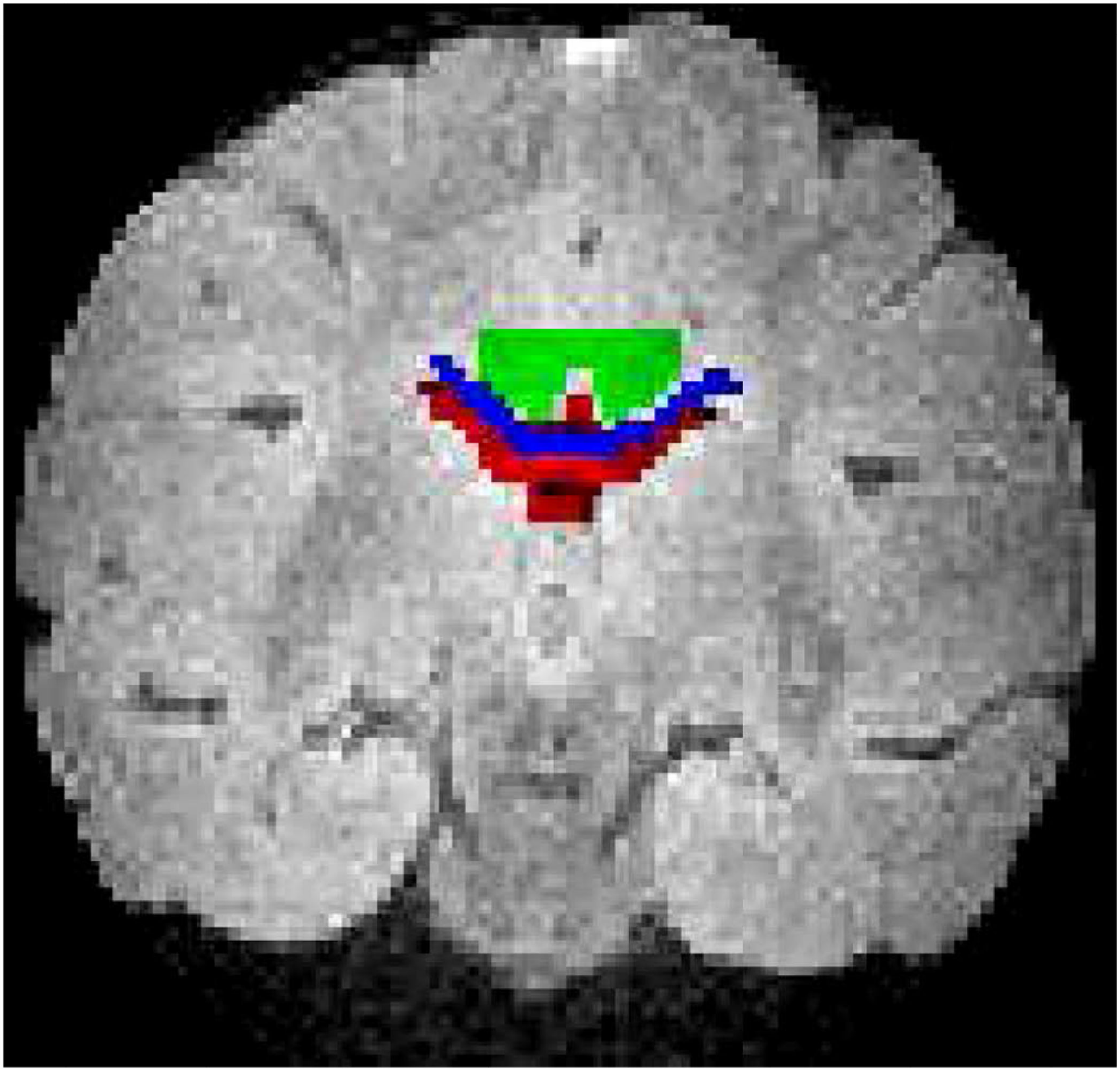

Fig. 3.

The T_2-weighted image of the baboon brain and the regions of interest used in the analysis: CSF in red ($n_{\text{roi}} = 82$), WM in green ($n_{\text{roi}} = 86$), and the CC in blue ($n_{\text{roi}} = 53$).

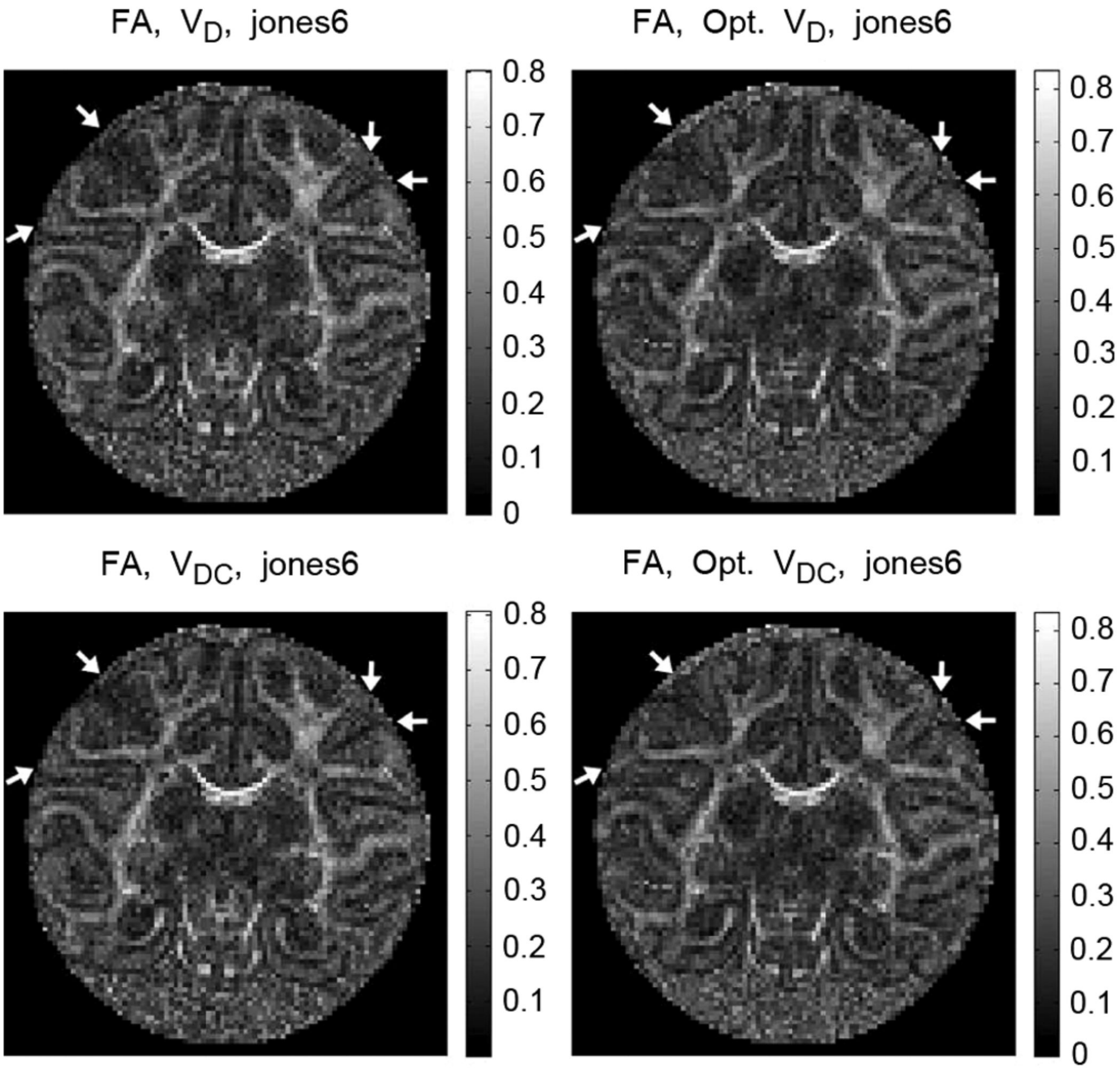

Fig. 4.

The fractional anisotropy maps obtained using nonoptimal and optimal versions of Jones6 DGS. On the right column, the arrows point to the structures that become visible when the optimal DGS is utilized. The structures on the left side of each image are the subcortical U-fibers that are not identifiable with the nonoptimal Jones6 DGS. The U-fiber, couple of pixels away from the left top arrow in its direction, is not distinguishable in nonoptimal schemes. The optimal DGS provides also better separability of the adjacent gyri as shown on the right side of the images with the arrows.

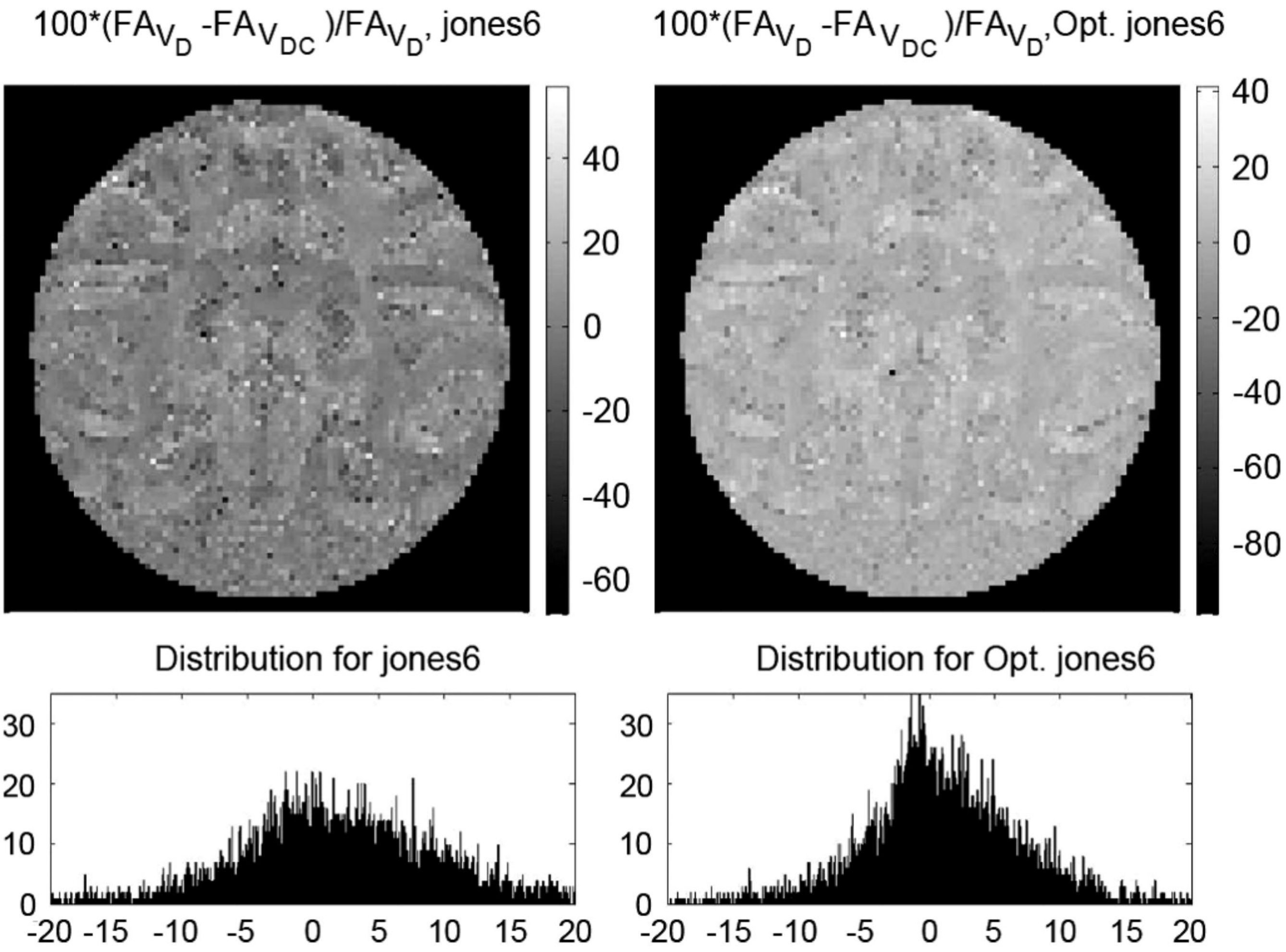

Fig. 5.

The relative difference between the fractional anisotropy maps obtained using “all gradients,” V, and “diffusion gradients only,” V_D (equivalently NoCroT) for the non-optimal and optimal versions of Jones6 DGS: $100 \times (FA_{V_D} - FA_V)/FA_{V_D}$. The distribution of the difference is highly concentrated around 0 on the second row right column for the optimal DGS, whereas for the nonoptimal scheme it is much more dispersed as shown on the left column. For the optimal DGS, the FA values obtained with and without the imaging gradients in the calculations are homogenously closer than the ones from the nonoptimal DGS.

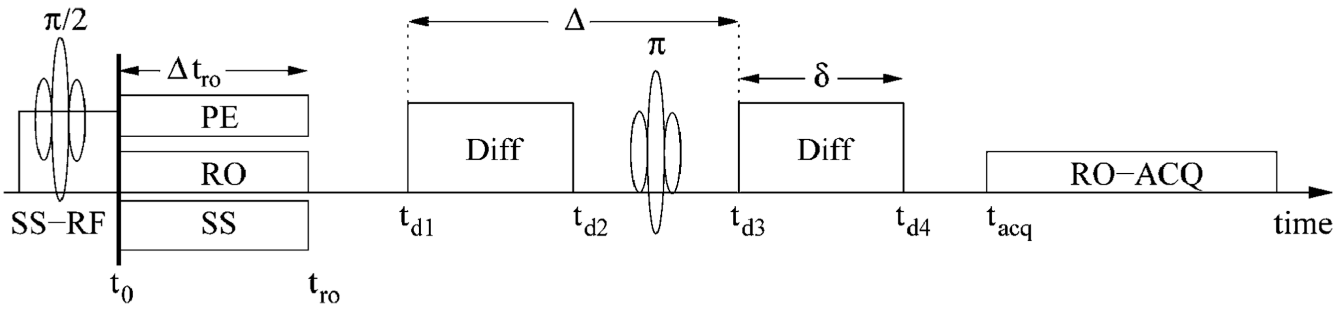

Fig. 6.

The pulsed-gradient spin-echo pulse sequence and the definition of the variables used in the calculations. RO is for read-out, PE for phase-encode, SS is the slice select gradient and ACQ stands for the digital to analog conversion acquisition.

TABLE I

Value of the Cost Function for the Pivot, Initial, and Optimized Schemes

Water Sample

	cond6	cond*	dsm	dualgr	mutm	jones6	muthup	tetra
Pivot	50.18	5.921	13.81	34.48	35.13	5.085	19.81	45.33
Initial	40.64	20.64	13.81	34.48	21.71	12.96	19.81	45.33
Opt.	25.35	5.177	4.546	4.348	9.437	4.241	4.41	12.53
Baboon Brain								
Pivot	37.51	5.052	12.69	33.26	30.7	3.757	18.6	42.14
Initial	33.39	19.63	10.05	4.011	27.93	7.146	6.53	12.13
Optimal	18.89	4.35	3.445	3.571	7.671	3.259	3.513	10.59

TABLE II

Percent Change Between Non-Optimal and Optimal Differences Between the Eigenvalues, $\Delta\lambda_i = \lambda_i - \tilde{\lambda}_i$, From V_D and V_{DC}

$100 \times ((\Delta\lambda_i)_{Non} - (\Delta\lambda_i)_{Opt})/(\Delta\lambda_i)_{Non}$

	cond6	**cond***	**dsm**	**dualgr**	**mutm**	**jones6**	**muthup**	**tetra**
Water	68.71	10.91	−209	79.99	85.86	47.03	88.31	48.08
	44.99	−104.6	−18.45	112.3	10.01	95.58	166.5	47.2
	17.04	−54.24	29.26	74.41	44.19	38.64	39.11	46.36
Brain	−142.1	83.65	31.83	−12.44	110.2	64.35	2.794	73.97
	−80.93	−70.82	39.07	−32.53	51.83	69.1	20.44	149.6
	80.14	315.5	47.04	62.08	99.2	11.91	37.97	75.11
CSF	197.8	328.1	58.2	79.18	107	122.2	44.08	62.39
	−35.2	−311.4	86.51	−72.34	208.3	81.64	5.018	146.3
	105	−1727	−15.63	31.15	79.35	21.34	36.51	70.98
WM	143.8	67.5	60.17	471.2	98.76	55.42	−29.67	74.64
	−93.53	−4.371	28.73	6.909	38.75	36.69	17.83	138.6
	60.67	226.7	1.596	41.84	240.4	76.01	31.51	70.36
CC	−545.3	−129.2	90.69	224.3	96.49	74.79	−0.2668	101.4
	125.8	−24.55	792.5	273.1	184.7	3.054	848.7	105.1
	145	−320.8	0.7862	68.16	49.22	43.91	34.76	67.36

TABLE III

Percent Change in the Model Matching Error

$100 \times (\bar{\chi}_{Non} - \bar{\chi}_{Opt})/\bar{\chi}_{Non}$

		cond6	cond*	dsm	dualgr	mutm	jones6	muthup	tetra
Water	V_{DC}	30.75	37.36	9.271	46.96	−8.979	12.41	17.63	52.36
	V_D	31	35	9.728	46.55	6.978	12.31	18.1	62.5
Brain	V_{DC}	22.83	17.85	6.936	1.703	17.23	4.257	2.016	9.192
	V_D	15.22	11.33	4.174	1.029	11.21	2.979	1.469	9.625
CSF	V_{DC}	24.26	28.97	14.53	3.14	21.82	14.54	7.437	2.633
	V_D	18.19	14.34	8.016	3.238	10.74	8.744	4.828	−0.5225
WM	V_{DC}	19.41	14.15	5.055	1.402	14.32	4.797	2.109	9.191
	V_D	12.1	3.751	−1.351	3.467	10.79	4.617	4.719	7.749
CC	V_{DC}	15.13	1646	7.221	−6.688	6.165	1.941	−2.423	1.661
	V_D	4.222	8.603	1.285	−6.669	−7.782	−0.3637	−5.418	2.659

TABLE IV

Percent Change in the Anisotropy Indices of the Water Sample

$$100 \times (\overline{FA}_{Non} - \overline{FA}_{Opt})/\overline{FA}_{Non}$$

	cond6	**cond***	**dsm**	**dualgr**	**mutm**	**jones6**	**muthup**	**tetra**
V_{DC}	32.17	−20.97	12.65	39.45	57.54	19.1	21.28	27.54
V_D	23.65	−6.573	−2.922	−42.2	20.11	−6.549	−6.801	−48.3

TABLE V

Non-Optimal DGSs [14], [15], [29]

cond6			**cond***		
0.755	0.26	0.602	1	0	0
−0.479	0.711	0.515	0	1	0
−0.394	−0.63	0.669	0	0	1
−0.616	−0.262	−0.743	0.707	0.707	0
0.558	−0.741	0.375	0	0.707	0.707
−0.954	−0.067	0.292	0.707	0	0.707
dsm			dualgr		
0.91	0.416	0	0.707	0.707	0
0	0.91	0.416	0.707	0	0.707
0.416	0	0.91	0	0.707	0.707
0.91	−0.416	0	0.707	−0.707	0
0	0.91	−0.416	0.707	0	−0.707
−0.416	0	0.91	0	0.707	−0.707
mutm			jones6		
0.851	0.526	0	1	0	0
0.851	−0.526	0	0.446	0.895	0
0	**0.526**	**0.851**	0.447	0.275	−0.851
0	**−0.851**	−0.526	0.448	−0.723	−0.525
−0.526	0	0.851	0.447	−0.724	0.526
0.526	0	0.851	−0.449	−0.277	−0.85
muthup			tetra		
0.851	0.526	0	0.577	0.577	0.577
0	0.851	0.526	−0.577	−0.577	0.577
0.526	0	0.851	0.577	−0.577	−0.577
0.851	−0.526	0	−0.577	0.577	−0.577
0	0.851	−0.526	0.707	0.707	0

cond6			cond*		
−0.526	0	0.851	0.707	0	0.707

TABLE VI

Optimal DGSs Used With the Water Tube.

Opt. cond6			**Opt. cond***		
−0.384	−0.513	−0.845	−0.5	1	−0.543
0.519	−1	−0.573	−0.828	−0.904	−0.541
0.859	0.96	−0.768	0.959	−0.135	−0.871
0.188	0.502	1	−0.939	0.0682	−0.767
−0.0932	1	−0.534	0.0928	−0.734	−0.999
0.996	0.223	−0.23	0.325	0.612	−1
Opt. dsm			Opt. dualgr		
−1	0.461	0.00318	−1	0.995	0.0199
−0.0034	0.995	0.446	−1	0.0501	0.981
−0.464	0.047	0.999	0	0.996	1
−1	−0.432	0.0116	−1	−0.946	−0.0193
0.00325	0.958	−0.465	−1	0	−0.98
0.45	0.0337	0.993	0	0.945	−0.961
Opt. mutm			Opt. jones6		
−0.314	1	0.898	−0.923	0.144	0.749
−0.727	−0.755	1	−0.295	1	0.39
0.807	0.11	0.362	−0.988	0.479	−0.295
−0.705	−0.946	−0.173	−0.885	−0.613	−0.109
0.922	−0.843	−0.173	−0.129	−0.771	0.689
0.278	−0.692	1	−0.232	−0.228	−1
Opt. muthup			Opt. tetra		
−0.993	0.609	0.0362	−0.727	0.822	0.82
−0.0265	0.978	0.639	0.906	−0.811	0.993
−0.627	0.0379	1	−0.976	−0.887	−1
−0.975	−0.578	0.00126	0.796	0.876	−0.813
−0.00332	0.943	−0.583	−1	1	−0.106

Opt. cond6			Opt. cond*		
0.59	0.0187	0.977	−0.933	−0.0333	1

TABLE VII

Optimal DGSs Used With the Baboon Brain

Opt. cond6			**Opt. cond***		
−0.0455	0.825	−0.872	−0.619	1	−0.556
0.935	0.309	−0.999	−0.795	−0.974	−0.528
0.681	1	1	0.985	−0.177	−0.858
−0.172	−0.977	0.845	−1	0.0185	−0.767
−0.304	0.852	0.812	0.135	−0.814	−0.981
0.944	0.208	0.464	0.259	0.582	−1
Opt. dsm			Opt. dualgr		
−0.107	−0.931	−0.753	−0.727	0.13	−0.853
0.938	−0.737	−0.061	0.21	0.48	−1
0.669	0.0439	−1	0.0455	−0.505	−1
−0.616	−0.137	−0.952	0.165	0.985	0
0.175	−1	0.497	−0.772	0.635	0.147
0.999	0.532	−0.22	−0.937	−0.35	0.147
Opt. mutm			Opt. jones6		
−0.974	−0.865	−0.679	−1	0.13	−0.587
−1	0.959	−0.336	−0.392	−0.777	−0.761
0.395	−0.142	−0.851	−0.969	−0.631	0.166
−0.257	1	0.697	−0.824	0.466	0.499
0.992	0.741	−0.366	−0.158	1	−0.218
−0.228	0.799	−0.993	−0.106	−0.231	0.999
Opt. muthup			Opt. tetra		
−1	0.262	−0.395	0.655	−0.5	−0.3
−0.347	−0.294	−1	0.365	1	0.0692
−0.17	0.731	−0.808	−0.907	−0.165	−0.686
−0.376	0.908	0.256	−0.113	−0.336	0.916
−0.663	−0.751	−0.054	0.177	−0.918	−0.226

Opt. cond6			Opt. cond*		
0.681	0.00759	−0.722	0.47	−0.1	−0.744