WHY MULTIPLE B-VALUES ARE REQUIRED FOR MULTI-TENSOR MODELS. EVALUATION WITH A CONSTRAINED LOG-EUCLIDEAN MODEL.

Benoit Scherrer, Simon K. Warfield

Computational Radiology Laboratory, Department of Radiology
Children's Hospital, 300 Longwood Avenue, Boston, MA, 02115, USA

ABSTRACT

Multi-tensor models have been proposed to assess multiple fiber orientations but are known to be numerically challenging. We show that the estimation cannot be performed with a single-shell HARDI acquisition because the fitting procedure leads to an infinite number of solutions ; multiple-shell HARDI acquisitions are required. Additionally, we propose a new log-euclidean constrained two-tensor model capable of assessing crossing fibers configurations with a relative limited number of DW acquisitions. We provide numerical experiments with this model to verify experimentally the necessity of multiple-shell HARDI acquisitions schemes for multi-tensor models.

Index Terms— Diffusion MRI, two-tensor model, log-euclidean, b-value

1. INTRODUCTION

Diffusion tensor imaging (DTI) [1] has been widely used to describe the three-dimensional nature of anisotropic diffusion in the human brain from diffusion measurements in several directions. However, its underlying *homogeneous Gaussian diffusion* assumption is well known to be inappropriate for assessing multiple fibers orientations. High angular resolution diffusion imaging (HARDI) acquisition schemes combined with specific algorithms have been proposed to overcome this limitation. They all share the idea of introducing many gradient encoding directions with one gradient strength (single-shell HARDI, one b-value) or several (multiple-shell HARDI, multiple b-values). A wide number of HARDI approaches have been proposed which generally aim at estimating an approximate of the underlying fiber orientation distribution: diffusion spectrum imaging (DSI), Q-ball imaging (QBI), spherical decomposition, generalized diffusion tensor imaging (GDTI), ... If these approaches are promising they generally require a large amount of data to be acquired (typically more than 100 images, up to 500), limiting their use for practical clinical applications. Lower spatial resolution acquisitions can be considered to reduce the imaging time but at the expense of the complexity of each voxel, which may not be desirable. Additionally, non-parametric models such as DSI or QBI are limited by the need of the *narrow pulse approximation* and the need to truncate the Fourier representation, leading to quantization artifacts [2]. In contrast, parametric models require less imaging time as they describe a predetermined model of diffusion rather than an arbitrary one. Particularly, it is generally admitted that the fiber orientation distribution of an individual fiber bundle is well represented by a single Gaussian. We can expect voxels containing multiple fiber orientations to be well represented by a mixture of Gaussians. Following that idea, multi-tensor approaches [3, 4, 5, 6] model the diffusion signal by fitting a simple finite mixture of Gaussians (generally two). Having a very low number of parameters to estimate, only a limited number of diffusion images should be necessary. These models are however known to be numerically challenging, experiencing difficulties for their fitting.

In this paper we demonstrate that multi-tensor models cannot be properly estimated with a single-shell HARDI acquisition because the model parameters are collinear, leading to an infinite number of solutions. In contrast, multiple-shell HARDI acquisitions makes the fitting to be better determined, leading theoretically to a unique solution. Additionally we propose a new constrained two-tensor model in the log-euclidean framework, able to perform the estimation with a relative limited number of DW acquisitions. Numerical experiments with this model are provided to demonstrate the necessity of multiple-shell HARDI acquisitions schemes for multi-tensor models.

2. THEORY

Diffusion signal modeling. The single tensor model consider that the local diffusion in a voxel can be modeled with a 3D Gaussian distribution, whose covariance matrix is proportional to the 3×3 diffusion tensor $\mathbf{D}$. The resulting diffusion-weighted signal S_k along a unit-gradient direction $\mathbf{g}_k$ is classically modeled as:

$$S_k(\mathbf{D}) = S_0 e^{-b_k \mathbf{g}_k^T \mathbf{D} \mathbf{g}_k},$$

where S_0 is the signal with no diffusion gradients applied and b_k is the b-value for the gradient direction k. In two-tensor models, or so called multi-compartment models [3], we consider that each voxel can be divided in a discrete number of homogeneous subregions in which the diffusion is Gaussian, i.e. fully described by a tensor, and that these subregions are in slow exchange. Let consider a model with two diffusion tensors $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2)$. The resulting diffusion-weighted signal S_k along a gradient direction $\mathbf{g}_k$ can be modeled as the finite mixture of Gaussians:

$$S_k(\mathbf{D}, \mathbf{f}) = S_0(f_1 e^{-b_k \mathbf{g}_k^T \mathbf{D}_1 \mathbf{g}_k} + f_2 e^{-b_k \mathbf{g}_k^T \mathbf{D}_2 \mathbf{g}_k}) \quad (1)$$

where $\mathbf{f} = (f_1, f_2)$ describes the volume *fractions* of each compartment ($f_j \in [0, 1]$) and verify $\sum_{i=0}^{2} f_i = 1$.

Why multiple b-values are required. We now demonstrate that the tensors $\mathbf{D}$ and the fractions $\mathbf{f}$ cannot be determined using a single shell HARDI acquisition. Let consider a constant b-value b, and let y_k be the measured signal for the direction k. $\mathbf{D}$ and $\mathbf{f}$ are generally estimated by a least-square approach by considering the K gradients directions:

$$\left(\widehat{\mathbf{D}}, \widehat{\mathbf{f}}\right) = \arg\min_{\mathbf{D}, \mathbf{f}} \sum_{k=1}^{K} [S_k(\mathbf{D}, \mathbf{f}) - y_k]^2 \quad (2)$$

If $(\widehat{\mathbf{D}}, \widehat{\mathbf{f}})$ is a solution of (2), then for any α, $\beta > 0$:

$$\begin{aligned} S_k\left(\widehat{\mathbf{D}}, \widehat{\mathbf{f}}\right) &= S_0\left(\frac{\alpha}{\alpha} \widehat{f}_1 e^{-b \mathbf{g}_k^T \widehat{\mathbf{D}}_1 \mathbf{g}_k} + \frac{\beta}{\beta} \widehat{f}_2 e^{-b \mathbf{g}_k^T \widehat{\mathbf{D}}_2 \mathbf{g}_k}\right) \\ &= S_0\left(\alpha \widehat{f}_1 e^{-b \mathbf{g}_k^T \widehat{\mathbf{D}}_1 \mathbf{g}_k - \log\alpha} + \beta \widehat{f}_2 e^{-b \mathbf{g}_k^T \widehat{\mathbf{D}}_2 \mathbf{g}_k - \log\beta}\right) \end{aligned}$$

with $\widehat{f}_2 = 1 - \widehat{f}_1$. Considering $\mathbf{g}_k^T \mathbf{g}_k = 1$, we have $\log\alpha = \mathbf{g}_k^T (\log\alpha\ \mathbf{I}_{3\times3})\, \mathbf{g}_k$ and:

$$S_k\left(\widehat{\mathbf{D}}, \widehat{\mathbf{f}}\right) = S_0 \left(\alpha \widehat{f}_1 e^{-b \mathbf{g}_k^T (\widehat{\mathbf{D}}_1 + \frac{\log\alpha}{b} \mathbf{I}_{3\times3}) \mathbf{g}_k} + \beta(1 - \widehat{f}_1) e^{-b \mathbf{g}_k^T (\widehat{\mathbf{D}}_2 + \frac{\log\beta}{b} \mathbf{I}_{3\times3}) \mathbf{g}_k}\right)$$

We can show that for $\beta = \frac{1-\alpha\widehat{f}_1}{1-\widehat{f}_1}$ and $\alpha \in]0, 1[$ we have (1) $\alpha\widehat{f}_1 + \beta(1 - \widehat{f}_1) = 1$, (2) $\alpha\widehat{f}_1 \leq 1$, (3) $\beta(1 - \widehat{f}_1) \leq 1$ and (4) $\beta > 0$. We can see that when all images are acquired with a single b-value, then if $(\widehat{f}_1, 1 - \widehat{f}_1)$ and $(\widehat{\mathbf{D}}_1, \widehat{\mathbf{D}}_2)$ is a solution of (2), for *any* $0 < \alpha < 1$, $(\alpha\widehat{f}_1, 1 - \alpha\widehat{f}_1)$ and $\left(\widehat{\mathbf{D}}_1 + \frac{\log\alpha}{b} \mathbf{I}_{3\times3}, \widehat{\mathbf{D}}_2 + \frac{\log(\frac{1-\alpha\widehat{f}_1}{1-\widehat{f}_1})}{b} \mathbf{I}_{3\times3}\right)$ is a solution of (2) as well : there is an infinite number of solutions. Additionally, non-degenerate tensors are obtained for $e^{-b\lambda_1^{\min}} < \alpha < \frac{1-(1-\widehat{f}_1)e^{-b\lambda_2^{\min}}}{\widehat{f}_1}$, $\lambda_1^{\min}$ and $\lambda_2^{\min}$ being respectively the minimum eigenvalues of $\widehat{\mathbf{D}}_1$ and $\widehat{\mathbf{D}}_2$. A single b-value conflates the tensor size indicated by the magnitude of its eigenvalues and the partial volume fractions. In contrast, the use of multiple b-values enables a unique solution to be found and disambiguates the estimation of $\mathbf{f}$ and $\mathbf{D}$. This allows measurements of the partial volume occupancy of each tensor in addition to the tensor estimation.

Constrained two-tensor model in a log-euclidean framework. We now propose an original two-tensor estimation framework to evaluate experimentally the need of multiple b-values for the estimation of two-tensor models. Different (multi-) tensor estimation schemes have been proposed in the literature. Particularly, care must be taken to ensure the positive-definitive property of the $\mathbf{D}_j$ and then avoid degenerate tensors with null or negative eigenvalues. Solutions include the Cholesky parameterization of the diffusion tensor, or Bayesian approaches with priors on the eigenvalues [6, 5]. An other elegant approach is to consider symmetric positive definite matrixes as elements of a Riemannian manifold with a particular affine-invariant metric [7], with which null and negative eigenvalues are at an infinite distance. Such a metric provides excellent theoretical properties but at a extremely high computational cost. The log-euclidean approach is a computationally efficient close approximation and has been successfully applied to one-tensor estimation [8]. To our knowledge it has never been used in multi-tensor models.
Additionally, following the model of [4], we introduce a geometrical constraint in the modeling: $\mathbf{D}_1$ and $\mathbf{D}_2$ are constrained to lie in the plane defined by the two biggest eigenvalues of the one-tensor solution, reducing the number of free parameters. However, contrary to [4], we do not constrain the principal diffusivity magnitude to be the same in both tracts. More precisely, if $\mathbf{D}^{1\mathrm{T}}$ is the one-tensor solution, *i.e.*:

$$\mathbf{D}^{1\mathrm{T}} = \mathbf{V}^T \Lambda \mathbf{V} \text{ with } \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \lambda_1 \geq \lambda_2 \geq \lambda_3,$$

then the j^{th} tensor can be expressed as:

$$\mathbf{D}_j = \mathbf{V}^T \widetilde{\mathbf{D}}_j \mathbf{V} \quad \text{with} \quad \widetilde{\mathbf{D}}_j = \begin{pmatrix} \widetilde{D}_{j,a} & \widetilde{D}_{j,b} & 0 \\ \widetilde{D}_{j,b} & \widetilde{D}_{j,c} & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

The problem is then recast into a 2D minimization problem: we look to estimate two 2D tensors $\widetilde{\mathbf{D}}_j$ subsequently rotated by $\mathbf{V}$, leading to the estimation of only three parameters per tensor. We parameterize each tensor $\widetilde{\mathbf{D}}_j$ by its logarithm: $\widetilde{\mathbf{L}} = (\widetilde{\mathbf{L}}_1, \widetilde{\mathbf{L}}_2) = (\log(\widetilde{\mathbf{D}}_1), \log(\widetilde{\mathbf{D}}_2))$. The predicted signal for a gradient direction k can then be modeled as (replacing $S_k(\widetilde{\mathbf{L}}, \mathbf{f})$ by S_k for simplification concerns):

$$S_k = S_0 \sum_{j=1}^{2} f_j e^{-b_k (\mathbf{V}\mathbf{g}_k)^T \exp(\widetilde{\mathbf{L}}_j)(\mathbf{V}\mathbf{g}_k)}$$

To ensure properly bounded and normalized fractions, we parameterize them trough a *softmax* transformation [3]: $f_j(\eta) = \frac{\exp\eta_j}{\sum_l \exp\eta_l}$, with $\eta = (\eta_l \in I\!\!R, l = 1, \ldots, 2)$. The estimation of $\widetilde{\mathbf{L}}$ and η is performed via a least-square criteria (see Equation 2) using a conjugate gradient method, the Fletcher-Reeves-Polak-Ribiere algorithm. It raises the problem of the initial position. Considering the one tensor solution $\mathbf{D}^{1\mathrm{T}}$, we propose to initialize $\widetilde{\mathbf{D}}_1^{(0)}$ and $\widetilde{\mathbf{D}}_2^{(0)}$ (and consequently $\widetilde{\mathbf{L}}_1^{(0)}$ and

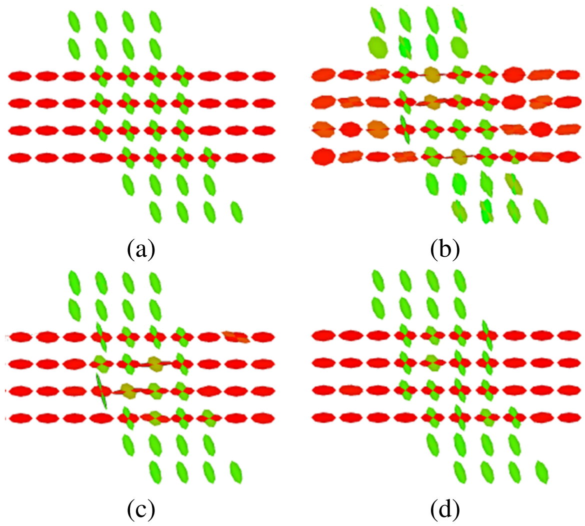

Fig. 1. Generated phantom of two fiber bundles crossing at 70° (a). Estimation from different simulated acquisitions: **(b):** one shell, D_1=90 directions, b_1=1000s/mm^2 ; **(c):** two shells, D_j=$(30, 15)$, b_j=$(1000, 7000)$; **(d):** two shells, D_j=$(45, 45)$, b_j=$(1000, 7000)$. Remark: double-tensors outside of the crossing region in (b,c,d) may appear as a single tensor because they are very similar.

$\widetilde{\mathbf{L}}_2^{(0)}$) according to a rotation of angle ϕ in the plane formed by (λ_1, λ_2) composed with a shrink of λ_2:

$$\widetilde{\mathbf{D}}_j^{(0)} = \begin{pmatrix} cos(\pm\phi) & \frac{1}{4}sin(\pm\phi) & 0 \\ -sin(\pm\phi) & \frac{1}{4}cos(\pm\phi) & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

We choose $\phi = \frac{\lambda_2}{\lambda_1}\frac{\pi}{4}$ so that when $\lambda_1 = \lambda_2$ the initial $\widetilde{\mathbf{D}}_j$'s describe two tensors whose principal diffusivities are perpendicular. In contrast, a one tensor solution with $\lambda_1 >> \lambda_2$ is likely to represent an individual fiber bundle, and the initial $\widetilde{\mathbf{D}}_j$'s are two tensors with almost parallel principal diffusivities. η is initialized to $\eta = (1, 1)$, leading to $\mathbf{f}(\eta) = (0.5, 0.5)$. Finally, the differentiation in the log-euclidean framework needed by the conjugate gradient algorithm gives:

$$\frac{\partial}{\partial \eta_j}(\cdot) = 2S_0 \sum_{k=1}^{K} (S_k - y_k) \left(\sum_{l=1}^{2} \frac{\partial f_l(\eta)}{\partial \eta_j} e^{-b_k \widetilde{\mathbf{g}_k}^T \exp(\widetilde{\mathbf{L}}_l) \widetilde{\mathbf{g}_k}} \right)$$

$$\nabla_{\widetilde{\mathbf{L}}_j}(\cdot) = -2S_0 f_j(\eta) \sum_{k=1}^{K} (S_k - y_k) b_k e^{-b_k \widetilde{\mathbf{g}_k}^T \exp(\widetilde{\mathbf{L}}_j) \widetilde{\mathbf{g}_k}}$$
$$\partial_{\widetilde{\mathbf{g}_k}\widetilde{\mathbf{g}_k}^T} \exp(-\widetilde{\mathbf{L}}_j),$$

with $\widetilde{\mathbf{g}_k} = \mathbf{V}\mathbf{g}_k$ (see [8] for practical implementation of $\partial_{\widetilde{\mathbf{g}_k}\widetilde{\mathbf{g}_k}^T} \exp(-\widetilde{\mathbf{L}}_j)$). After convergence, the final tensors are obtained by $\widehat{\mathbf{D}}_j = \mathbf{V}^T \exp(\widehat{\widetilde{\mathbf{L}}}_j)\, \mathbf{V}$.

3. EVALUATION

Simulations. We experimentally evaluated the need of multiple-shell HARDI acquisitions with our log-euclidean constrained two-tensor approach. The evaluation was currently performed with simulations. We generated a phantom representing two fiber bundles crossing with an angle of 70° (see Fig. 1.a). The typical tensor profile representing an individual fiber bundle was estimated from a real DWI acquisition by averaging the eigenvalues of voxels with highest FA, leading to $(\lambda_1, \lambda_2, \lambda_3) = (5.119 \times 10^{-4}, 2.178 \times 10^{-4}, 1.071 \times 10^{-4})$. In the crossing region, the fractions $\mathbf{f}$ were set to $(f_1, f_2) = (0.7, 0.3)$. We simulated the diffusion-weighted signal (see Equation (1)), corrupted by a Rician-noise (from two $\mathcal{N}(0, \sigma = 1)$), for different acquisition schemes. We used a b-value of $b_1 = 1000$s/mm^2 for the first shell, which is a generally admitted optimal b-value in the adult human brain. Different b-values for the other shells were considered when using multiple-shell schemes.

Results. Fig. 1.b shows that with a single b-value, even a high angular resolution acquisition (D_1=90 directions) do not provide a satisfying result. In comparison, an acquisition with a total of only 45 directions with one low and one high b-value provides better results (Fig. 1.c), which are improved by considering 45 directions with two b-values each (Fig. 1.d). We then quantitatively evaluated the accuracy of both the tensors estimates and the fractions estimates depending on the value of b_2 for three different two-shells HARDI schemes (Fig. 2.a and 2.b). Each two-tensor $\widehat{\mathbf{D}}$ was compared to the ground-truth $\mathbf{D}^{\text{gt}}$ in term of average minimum log-euclidean distance (AMD): $\Delta_{\text{err}} = \frac{1}{2}\sum_{j=1}^{2} \min_k || \log(\widehat{\mathbf{D}}_j) - \log(\mathbf{D}_k^{\text{gt}})||$. Fig. 2.a reports the mean Δ_{err} over all tensors of the phantom. The fractions were compared to the ground-truth by considering their average absolute difference (AAD) over the crossing region (Fig. 2.b). Figs. 2.a and 2.b show that the combination of a low and a high b-value helps to stabilize the estimation of both the tensors and the fractions. Additionally, HARDI schemes such as (D_1=60,D_2=30) or (D_1=45,D_2=45) should be preferred to (D_1=75,D_2=15). Figs. 2.c and 2.d show that for a same number of directions, multiple-shells HARDI schemes always provides better results than a single-shell scheme, which is consistent with our demonstration in Section 2. For a very low number of directions (45), the tensor are in average surprisingly quite well estimated compared to the fractions. It is likely due to a good estimation outside of the crossing region but less good inside (see also Fig. 1.c).

4. DISCUSSION

In this paper, we (1) have shown theoretically that multi-tensor approaches require at least two b-value acquisitions for their estimations and (2) have proposed a log-euclidean constrained two-tensor model to assess multiple fiber orientations from a relative limited number of DW acquisitions.

In the literature, a number of approaches fit a multi-tensor model with only one b-value [3, 6, 4]. Other approaches use several b-values [5] but, to our knowledge, do not describe the theoretical reasons to do so. We have shown that with only

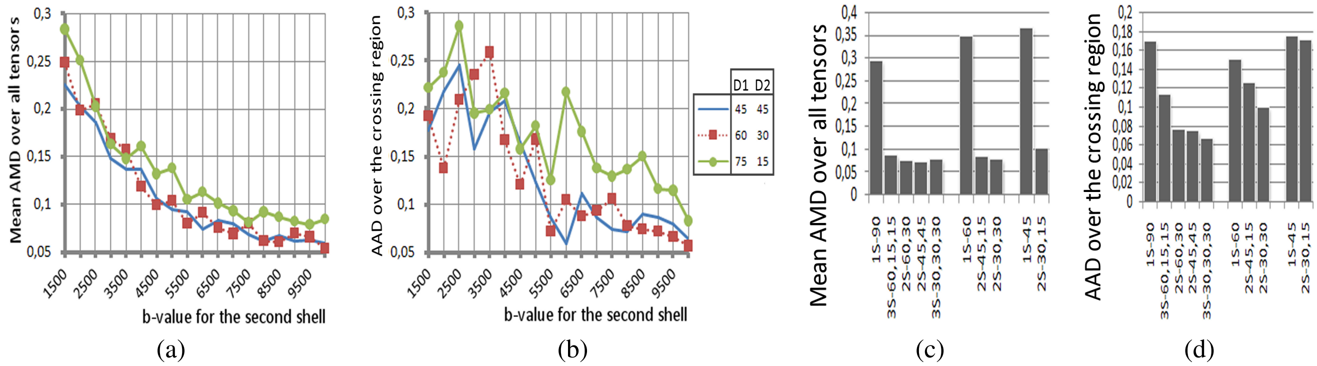

Fig. 2. Evaluation of the tensors (a) and the fractions (b) estimated from simulated two-shells acquisitions in function of b_2 ($b_1 = 1000\text{s/mm}^2$) for three different acquisition schemes of 90 directions each (D_j={(45, 45), (60, 30), (75, 15)}). Fig (c) and (d): Evaluation with different simulated acquisitions composed of a total of 90, 60 or 45 directions (**1S-90:** 1 shell, b_1=1000, D_1=90 ; **2S-60,15,15:** 3 shells, b_j=(1000, 4000, 7000), D_j=(60, 15, 15) ; **2S-60,30:** 2 shells, b_j=(1000, 7000), D_j=(60, 30) ; **2S-45,45:** 2 shells, b_j=(1000, 7000), D_j=(45, 45) ; **3S-30,30,30:** 3 shells, b_j=(1000, 4000, 7000), D_j=(30, 30, 30) ; **1S-60:** 1 shell, b_1=1000, D_1=60 ; **2S-45,15:** 2 shells, b_1=(1000, 7000), D_j=(45, 15) ; **2S-30,30:** 2 shells, b_1=(1000, 7000), D_j=(30, 30) ; **1S-45:** 1 shell, b_1=1000, D_1=45 ; **2S-30,15:** 2 shells, b_1=(1000, 7000), D_j=(30, 15)) .

one b-value, the magnitude of the estimated tensors eigenvalues and the estimated fractions are melded, which is not a desirable property: in this situation, a fiber bundle with a uniform $\mathbf{D}_1$ across its entire length may appear to grow and shrink as it passes through voxels and experiences different partial volume effects. The introduction of constraint in the modeling cannot resolve the ambiguity because the problem comes from the collinearity in the parameters. Only the use of multiple-shell HARDI acquisitions allows to disambiguate the estimation of the tensors and the fractions. This was experimentally verified with simulations by using our proposed log-euclidean constrained two-tensor model. It appears that combining low and high b-values provides better results, possibly due to numerical reasons. Such situations probably reduce the number of local minima during the model fitting. We show that two tensors configurations can be estimated from a relative low number of DW images with this model, and thus clinically compatible scan times may be reached. An interesting refinement could be to add an isotropic term to the mixture of Gaussian to account for an isotropic diffusion compartment [5, 6]. Future works should concern a fully characterization of our new constrained log-euclidean two-tensor model, including (1) the acquisition of real human brain multiple-shell HARDI data and (2) an investigation of the tradeoff between the quality of the tensor fitting, the imaging time, the noise level, and the angular resolution detection of crossing fibers. Model selection may also be considered to choose between the one-tensor and the two-tensor solutions for each voxel.

5. REFERENCES

[1] P J Basser, J Mattiello, and D LeBihan, “Estimation of the effective self-diffusion tensor from the NMR spin echo,” *J Magn Reson B*, vol. 103, no. 3, pp. 247–254, 1994.

[2] E J Canales-Rodríguez, L Melie-García, and Y Iturria-Medina, “Mathematical description of q-space in spherical coordinates: exact q-ball imaging,” *Magn Reson Med*, vol. 61, no. 6, pp. 1350–1367, 2009.

[3] D S Tuch, T G Reese, M R Wiegell, N Makris, J W Belliveau, and V J Wedeen, “High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity,” *MRM*, vol. 48, no. 4, pp. 577–582, 2002.

[4] S Peled, O Friman, F Jolesz, and C F Westin, “Geometrically constrained two-tensor model for crossing tracts in DWI,” *Magn Reson Imaging*, vol. 24, no. 9, pp. 1263–1270, 2006.

[5] L Melie-García, E J C-Rodríguez, Y Alemán-Gómez, C P Lin, Y Iturria-Medina, and P A Valdés-Hernández, “A Bayesian framework to identify principal intravoxel diffusion profiles based on diffusion-weighted MR imaging,” *Neuroimage*, vol. 42, no. 2, pp. 750–770, 2008.

[6] T E Behrens, M W Woolrich, M Jenkinson, H Johansen-Berg, R G Nunes, S Clare, P M Matthews, J M Brady, and S M Smith, “Characterization and propagation of uncertainty in diffusion-weighted MR imaging,” *Magn Reson Med*, vol. 50, no. 5, pp. 1077–1088, 2003.

[7] Xavier Pennec, Pierre Fillard, and Nicholas Ayache, “A Riemannian framework for tensor computing,” *Int. J. of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.

[8] P Fillard, X Pennec, V Arsigny, and N Ayache, “Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics,” *IEEE Trans Med Imaging*, vol. 26, no. 11, pp. 1472–1482, 2007.