Information theoretic methods for diffusion-weighted MRI analysis*

Angelos Barmpoutis[1] and Baba C. Vemuri

Department of Computer Information Science & Engineering
University of Florida, Gainesville FL 32611, USA,
{abarmpou, vemuri}@cise.ufl.edu

Concepts from Information Theory have been used quite widely in Image Processing, Computer Vision and Medical Image Analysis for several decades now. Most widely used concepts are that of KL-divergence, minimum description length (MDL), etc. These concepts have been popularly employed for image registration, segmentation, classification etc. In this chapter we review several methods, mostly developed by our group at the Center for Vision, Graphics & Medical Imaging in the University of Florida, that glean concepts from Information Theory and apply them to achieve analysis of Diffusion-Weighted Magnetic Resonance (DW-MRI) data.

This relatively new MRI modality allows one to non-invasively infer axonal connectivity patterns in the central nervous system. The focus of this chapter is to review automated image analysis techniques that allow us to automatically segment the region of interest in the DWMRI image wherein one might want to track the axonal pathways and also methods to reconstruct complex local tissue geometries containing axonal fiber crossings. Implementation results illustrating the algorithm application to real DW-MRI data sets are depicted to demonstrate the effectiveness of the methods reviewed.

1 Introduction

Modern technological developments in image acquisition techniques have made it possible to capture images from various medical image modalities in high resolution. Magnetic Resonance Imaging (MRI) allows capturing of high contrast images of the soft human tissues. More specifically, Diffusion-Weighted MRI (DW-MRI) is the only non-invasive method for capturing the diffusivity of molecules of water in human tissue. The local diffusion properties usually change in different parts of the tissue being imaged due to changes encountered in anisotropy to water diffusion and these variations in anisotropy result in variations in signal attenuation which are captured in the acquired signal. By analyzing the local diffusion characteristics one can obtain information about the connectivity patterns prevalent say in the brain or the spinal cord, which motivates the development of appropriate methods for processing these datasets. The acquired DW-MRI signal is attenuated at locations of higher diffusivity and

* The research was in part funded by the grant NIH EB007082.

can be observed if they are along the direction of the diffusion sensitizing magnetic gradient. This attenuation is popularly approximated by the Stejkal-Tanner equation [1] as follows:

$$S/S_0 = e^{-bd(\mathbf{g})} \tag{1}$$

where S is the signal acquired by applying diffusion-weighting magnetic gradient field with direction $\mathbf{g}$ and (a weighting) b-value b, S_0 is the signal acquired without diffusion-weighting, and $d(\mathbf{g})$ is the so called diffusivity function. In the original form of the signal attenuation equation (by Stejskal and Tanner [1]) the diffusivity function was approximated by a constant $d(\mathbf{g}) = d$ representing the mean diffusivity. The advances in imaging techniques however made it possible to acquire several diffusion-weighted MR images S_i by applying different diffusion gradient directions $\mathbf{g}_i$. This allowed for the approximation of the diffusivity function with a second-order tensor $d(\mathbf{g}) = \mathbf{g}^T\mathbf{D}\mathbf{g}$, where $\mathbf{D}$ is a 3×3 symmetric and positive-definite matrix – see Basser et al., [2], for pioneering research in this direction.

The relation between the diffusion-weighted signal attenuation (eq. 1) and the diffusion propagator equation is given by the following Fourier integral expression

$$P(\mathbf{r}) = \int S(\mathbf{q})/S_0 e^{-2\pi i \mathbf{q}^T \mathbf{r}} d\mathbf{q} \tag{2}$$

where $\mathbf{q}$ is the reciprocal space vector, $S(\mathbf{q})$ is the DW-MRI signal value associated with vector $\mathbf{q}$, S_0 the zero gradient signal and $\mathbf{r}$ is the displacement vector [3]. Note that the direction of vector $\mathbf{q}$ in Eq. 2 is the same with that of $\mathbf{g}$ in Eq. 1 and that b is related to the magnitude of $\mathbf{q}$ with the expression $b = 4\pi q^2 t$, where t is the effective diffusion time.

By using the tensorial approximation of the diffusivity function, the Fourier transform in Eq. 2 can be computed analytically and is given by,

$$P(\mathbf{r}) = \frac{1}{\sqrt{(2\pi)^3 |2t\mathbf{D}|}} e^{-\frac{1}{4t}\mathbf{r}^T \mathbf{D}^{-1} \mathbf{r}}. \tag{3}$$

The orientation $\mathbf{r}$ that maximizes the displacement probability $P(\mathbf{r})$ corresponds to the orientation of maximum water molecule diffusion. In the diffusion tensor model case, it can be easily seen that Eq. 3 is maximized when the quantity $\mathbf{r}^T\mathbf{D}\mathbf{r}$ is also maximized, i.e. for vectors $\mathbf{r}$ which are parallel to the primary eigen-vector of matrix $\mathbf{D}$. Therefore, by computing the primary eigen-vector of $\mathbf{D}$ using the method of spectral decomposition one can easily estimate the orientation of maximum diffusion, which is one of the advantages of diffusion tensor imaging. Furthermore, by following the primary eigen-vectors one can trace the underlying fiber paths in the neural tissue [4]. This procedure is known as fiber tracking and it is a tool for obtaining information and analyzing the brain connectivity.

The estimation of a smooth field of diffusion tensors and its segmentation can be performed by using an information theoretic approach first introduced in literature by Wang et al., [5] and reviewed in Sec. 2. Other methods include the

geometric approach using the Riemannian metric of the space of positive-definite matrices [6–9], tensor spline regularization [10], the log-Euclidean framework [11] and the Geodesic-loxodromes method [12] for tensor interpolation.

The diffusion tensor model can be extended for cases of complex fiber structures, such as fiber crossings, using the multi-compartmental model also known as the mixture model given by:

$$S(\mathbf{q})/S_0 = \sum_i^N w_i S_i(\mathbf{q}) \tag{4}$$

where w_i are unknown mixing weights [13]. The drawback in this discrete mixture model is that its difficult to select a priori, the number N of mixing compartments. This problem can however be solved by using the continuous mixture model as was first proposed in Jian et al., [14] and later in Kumar et al., [15, 16], both of which can be expressed in the following unified de-convolution framework introduced in Jian and Vemuri [17],

$$S(\mathbf{q})/S_0 = \int f(X)K(\mathbf{q}|X)dX. \tag{5}$$

In Eq. 5 the signal response is parametrized using a kernel function $K(\mathbf{q}|X)$, where X is a set of parameters, $f(X)$ is a properly chosen mixing density function over the domain of X, and the integration is with respect to X.

In Sec. 3 we study two different versions of Eq. 5: a) setting X to be the diffusion tensor $\mathbf{D}$ and $f(\mathbf{D})$ to be a mixing density function on the space of 3×3 symmetric and positive-definite matrices [14], and b) setting X to be a 3-dimensional unit vector μ representing the orientation of maximum signal response and $f(\mu)$ to be a mixing density over the space of unit vectors (i.e. over the unit sphere) [15].

Other methods for multi-fiber reconstructions include higher-order tensor models [18, 19], and their equivalent spherical tensor expansion [20], diffusion orientation transform for computing the displacement probability profiles from given diffusivity profiles [21], the estimation of orientation distribution function in q-ball imaging [22–24], spherical de-convolution [25–28], diffusion spectrum magnetic resonance imaging [29]. Recently, Jian and Vemuri [17] developed the unified de-convolution framework (Eq. 5) for multi-fiber reconstruction within a voxel from diffusion weighted MRI and posed several of the existing methods in this framework that facilitated easy comparison and showed superior performance of their continuous mixture of Wisharts model [14].

The performance of methods reviewed in this chapter are demonstrated using synthetic and real diffusion-weighted MRI data. The goal of applying various information theoretic concepts to analyze the real data is to unravel the underlying fiber geometry making explicit the connectivity patterns in various regions of the neural tissue. From a clinical point of view, to date, several methods for analyzing the diffusivity have been used in monitoring encephalopathy, sclerosis, ischemia and other brain disorders [30, 31]. We hope that some of the methods

reviewed here will be of use in analyzing these neurological disorders in the near future. The rest of the chapter is organized as follows. In Sec. 2 we review the information theoretic method from [5] for regularizing and segmenting diffusion tensor fields. Information geometric methods for multi-fiber reconstruction are reviewed in Sec. 3. Each section is accompanied by an experimental results subsection demonstrating the performance of the presented methods. Finally in Sec. 4 we conclude.

2 Information theoretic methods for processing DTI

In this section we review an information theoretic dissimilarity measure for the space of positive definite-matrices. A dissimilarity measure between tensors is needed in several tensor-valued image analysis methods, such as in regularization, interpolation and segmentation of tensor fields.

2.1 An information theoretic dissimilarity measure

It is well known that Brownian motion of the water molecules in the soft tissue is a Gaussian process, and therefore it is known that the diffusion propagator is given by the Gaussian probability expressed in Eq. 3. The diffusion tensor is defined as the covariance matrix in this probability density function; hence a natural dissimilarity measure between diffusion tensors can be defined by employing a divergence between the corresponding Gaussian probabilities. Wang and Vemuri defined a dissimilarity measure between two given tensors $\mathbf{D}_1$ and $\mathbf{D}_2$ by using the J-divergence as follows

$$d(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{J(P(\mathbf{q}|\mathbf{D}_1), P(\mathbf{q}|\mathbf{D}_2))} = \sqrt{\frac{KL(P_1, P_2) + KL(P_2, P_1)}{2}} \qquad (6)$$

where $P_1 = P(\mathbf{q}|\mathbf{D}_1)$, $P_2 = P(\mathbf{q}|\mathbf{D}_2)$ and KL is the well known KL-divergence [32] given by $KL(p_1, p_2) = \int p_1(\mathbf{q}) log(\frac{p_1(\mathbf{q})}{p_2(\mathbf{q})}) d\mathbf{q}$. By substituting $P(\mathbf{q}|\mathbf{D}_1)$ and $P(\mathbf{q}|\mathbf{D}_2))$ in Eq. 6 by the Gaussian probability in Eq. 3, the following closed form can be derived

$$d(\mathbf{D}_1, \mathbf{D}_2) = \frac{1}{2}\sqrt{trace(\mathbf{D}_1^{-1}\mathbf{D}_2 + \mathbf{D}_2^{-1}\mathbf{D}_1) - 2n} \qquad (7)$$

where n is the dimension of the matrix, i.e. in the diffusion tensor case $n = 3$. The proof of Eq. 7 was first published in [5]. A useful property of the dissimilarity measure defined above is the invariance to affine transformations. In other words, $d(\mathbf{D}_1, \mathbf{D}_2) = d(\mathbf{A}\mathbf{D}_1\mathbf{A}^T, \mathbf{A}\mathbf{D}_2\mathbf{A}^T)$ where $\mathbf{A}$ is an affine transformation matrix.

The dissimilarity measure given by Eq. 7 cannot be considered as a distance between tensors since it fails to satisfy the triangle inequality. However, it can be shown that for infinitesimally close tensors $\mathbf{D}_1$ and $\mathbf{D}_2$, Eq. 7 approximates the squared Riemannian geodesic distance between them (which was used in

[6]). Therefore, in this case we can say that Eq. 7 approximates the Rao distance (which is also the Riemannian distance) between the two nearby Gaussian distributions $P_1(\mathbf{q})$ and $P_2(\mathbf{q})$.

The above dissimilarity measure can be employed for defining the average tensor $\mathbf{D}_\mu$ in a tensor field as the minimizer of $\int d^2(\mathbf{D}_\mu, \mathbf{D}(\mathbf{x}))d\mathbf{x}$, i.e. $\mathbf{D}_\mu$ is the tensor representing the mean of the elements of the field. Here the vector $\mathbf{x}$ denotes the lattice index in the tensor field. By substituting Eq. 7 into the above integral, it was shown in [5] that the average tensor is given by

$$\mathbf{D}_\mu = \sqrt{\mathbf{B}^{-1}}\sqrt{\sqrt{\mathbf{B}}\mathbf{A}\sqrt{\mathbf{B}}}\sqrt{\mathbf{B}^{-1}} \tag{8}$$

where $\mathbf{A} = \int_R \mathbf{D}(\mathbf{x})d\mathbf{x}$ and $\mathbf{B} = \int_R \mathbf{D}^{-1}(\mathbf{x})d\mathbf{x}$

Here we note that Eq. 8 is valid also for computing the average tensor $\mathbf{D}_\mu$ defined as the mid-point in the geodesic between two tensors $\mathbf{D}_1$ and $\mathbf{D}_2$ using the Riemmanian metric of $\mathcal{P}_n$ [6, 7, 9] and setting $\mathbf{A} = \mathbf{D}_1$ and $\mathbf{B} = \mathbf{D}_2^{-1}$. The symbol $\mathcal{P}_n$ denotes the space of $n \times n$ symmetric positive-definite matrices. However, in the Riemmanian framework there is no analytic formula for computing the average of more than two tensors (e.g. tensor field), and an iterative optimization method has been used instead, which significantly increases the execution time of the diffusion tensor processing algorithms.

In the next section, we review how the formulas presented here can be employed in a level-set framework for segmenting diffusion tensor fields.

2.2 Application to DTI segmentation

The segmentation of a DTI field can be performed by minimizing the following Mumford-Shah energy function

$$E(\mathbf{D}, C) = E_{dist} + E_{reg} + E_{arc} = \alpha \int_\Omega d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}_0(\mathbf{x}))d\mathbf{x} + \int_{\Omega/C} \nabla\mathbf{D}(\mathbf{x})d\mathbf{x} + \beta|C| \tag{9}$$

where $\mathbf{D}_0(\mathbf{x})$ is the given noisy tensor field, $\mathbf{D}(\mathbf{x})$ is the approximated (fitted) tensor field, Ω is the domain of the field (i.e. $\Re^2$ or $\Re^3$ for 2D or 3D fields respectively), C is the boundary between the segmented regions, and α and β are constant factors. The first term in Eq. 9 measures the dissimilarity between the fitted tensor field and the original noisy field, the second term measures the variation (or the lack of smoothness) within the segmented regions, and the last term measures the arc length of the segmentation curve C.

In the case that the fitted field $\mathbf{D}(\mathbf{x})$ is chosen to be a piecewise constant model, the second term in Eq. 9 becomes $E_{reg} = 0$, while the first term is given by

$$E_{dist} = \sum_{\mathbf{x}\in R} d^2(\mathbf{D}_1, \mathbf{D}_0(\mathbf{x})) + \sum_{\mathbf{x}\in R^c} d^2(\mathbf{D}_2, \mathbf{D}_0(\mathbf{x})) \tag{10}$$

where $\mathbf{D}_1$ is the mean tensor in the region R enclosed by the curve C and $\mathbf{D}_2$ is the mean tensor in the region R^c outside the segmentation curve. Note that

since the given tensor field is on a discrete lattice, the integral of Eq. 9 has been replaced by a summation in Eq. 10.

Instead of the piecewise constant model used above, a piecewise continuous model may be chosen to approximate the tensor field. In this case the dissimilarity term in Eq. 9 is given by the following expression

$$E_{dist} = \alpha \sum_{\mathbf{x} \in R} d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}_0(\mathbf{x})) + \alpha \sum_{\mathbf{x} \in R^c} d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}_0(\mathbf{x})) \tag{11}$$

and the second term in Eq. 9 becomes

$$E_{reg} = \sum_{(\mathbf{x},\mathbf{y}) \in R, \mathbf{y} \in N_{\mathbf{x}}} d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})) + \sum_{(\mathbf{x},\mathbf{y}) \in R^c, \mathbf{y} \in N_{\mathbf{x}}} d^2(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y})). \tag{12}$$

where each term measures the lack of smoothness within the segmented regions and $N_{\mathbf{x}}$ is a neighborhood centered at location $\mathbf{x}$.

In both piecewise constant and piecewise continuous models, the distance function $d(,)$ can be set to the dissimilarity measure defined in Eq. 7 since the latter approximates the geodesic distance between two nearby elements in the space $\mathcal{P}_n$ of positive definite matrices. Furthermore, the simple closed-form expression of Eq. 7 and 8 make it possible to produce analytic update formulas for minimizing the variational principle, which is one of the main advantages of this information theoretic dissimilarity measure over the $\mathcal{P}_n$ Riemannian metric.

The curve evolution equations for the above segmentation framework are reported along with their derivations in [5]. There in also lies a detailed discussion about the implementation of the algorithm and the numerical techniques used to solve the flow equations. The above methods are demonstrated in the next section using simulated synthetic and real diffusion-weighted MR data sets.

2.3 Experimental Results

In this section we present experimental results obtained by applying the DTI segmentation method reviewed earlier to synthetic and real DW-MRI data sets. The synthetic tensor field (shown in Fig. 1) is of size 32×32 and consists of the following regions: (1) a ring with principal eigenvectors tangent to circles centered in the lower left corner of the image; (2) two triangular regions with horizontal principal eigenvectors; and (3) two triangular regions with vertical principal eigenvectors. All three regions have distinct piecewise constant fractional anisotropy. Figure 1 (upper left plate) shows a plot of the primary eigenvectors.

We applied the piecewise continuous DTI segmentation method presented in Sec. 2.2 for segmenting the circular region in the synthetic data set. The challenge in segmenting this data set is due to the smooth transitions between the regions, especially at the locations shown in the upper central plate of Fig. 1. In the method we used the values $\alpha = -2.5$ (advection) and the $\beta = 0.01$ (smoothing). The segmentation boundaries were initialized as shown in the lower left plate of the same figure. After the execution of the method the boundaries converged

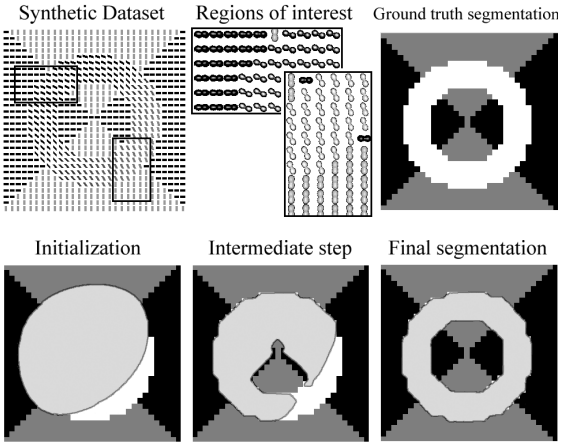

Fig. 1. Application of the piecewise continuous DTI segmentation method in a synthetic dataset.

as shown in the lower right plate, accurately segmenting the ring region. This demonstrates the effectiveness of the presented method.

Finally, in order to illustrate the performance of the segmentation framework on real data sets, we applied the method to a DTI data set from an excised rat hippocampus (shown in Fig. 2). The original DW-MRI data set contained 22 images acquired using a pulsed gradient spin echo pulse sequence, with 21 different diffusion gradients and approximate b value of 1250 s/mm^2. From this data set we estimated a DTI field using the variational formulation proposed in [33, 34]. Then, we segmented the hippocampal region that consists of fimbria, stratum oriens and dorsal commisure using the method reviewed in this section. The tensor field in this region is smoothly varying and highly anisotropic. For the segmentation of this field we used the same value for α as before, but we increased the regularization parameter $\beta = 2.0$ in order to enforce smoothing of the segmentation curve and therefore avoid producing results with 'bumps' and other artifacts due to the presence of noise in the data set. Figure 2 shows, in 2D and 3D views, part of the principal eigenvector field and the segmentation result. By observing the figures we can see that smooth segmentation boundaries have been created successfully enclosing the aforementioned hippocampal regions.

For more experimental results testing the performance of the method under various noise conditions the reader is referred to [5].

3 Multi-fiber reconstruction

In the information theoretic framework discussed in the previous section, a second order tensor (Diffusion Tensor) was employed to approximate the diffusivity

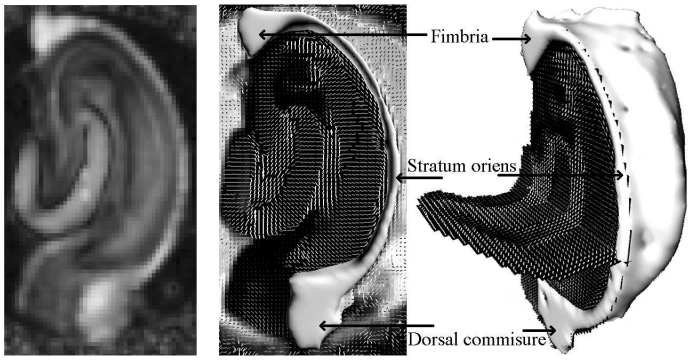

Fig. 2. Segmentation of the fimbria-stratum oriens-dorsal commisure region in a real rat hippocampus dataset. Left: Fractional anisotropy, Center and Right: 2D and 3D view of the segmented region.

function in the Stejskal-Tanner model of DW-MRI attenuation (Eq. 1). However, the second-order tensorial approximation fails to represent complex local geometries of the tissue, such as fiber crossings [21, 4]. Several methods have been proposed for multi-fiber reconstruction, and they can be categorized into model-free and model-based methods as was mentioned in Sec. 1. Jian and Vemuri [17] have shown that many of the model-based methods can be regarded as special cases of the unified framework formulation presented by them and given by Eq. 5.

In this section we discuss two special cases in this unified framework by setting the mixing density $f(X)$ in Eq. 5 to be (a) a mixture of Wishart distributions, and (b) a mixture of von Mises-Fisher distributions. The motivation for studying these two methods is that the first one approximates the DW-MRI signal attenuation by using statistics on the space of diffusion tensors while the second one solves the same problem by following a more general approach for approximating any function on a spherical domain.

3.1 The Mixture of Wisharts model

The Wishart distribution of positive-definite matrices is a generalization of the gamma distribution (to non-integer degrees of freedom) and of the chi-square distribution (to multiple dimensions) and its probability density function [35] in the case of (3×3) matrices is given by

$$f_w(\mathbf{D}|\mathbf{\Sigma}, p) = \frac{|\mathbf{D}|^{p-2} exp(-trace(\mathbf{\Sigma}^{-1}\mathbf{D}))}{2^{2p}|\mathbf{\Sigma}|^p \Gamma_3(p)} \tag{13}$$

where $\mathbf{D}$ is the matrix-valued random variable, $\mathbf{\Sigma}$ is the scale parameter (both $\mathbf{D}$ and $\mathbf{\Sigma}$ are positive definite matrices), and p is a scalar that controls the shape of the distribution.

The DW-MRI signal attenuation model of discrete mixture of Gaussians can be extended to a continuous mixture model in this unified framework [14] given by 5 and setting the parameter vector X to be a 3×3 positive-definite matrix $\mathbf{D}$, the mixing density $f(X)$ to be the Wishart distribution (Eq. 13) and the signal Kernel $K(\mathbf{q}|X)$ to be the Stejskal-Tanner model (Eq. 1). After these substitutions, we obtain a continuous mixture model that is given by

$$S/S_0 = \int f_w(\mathbf{D}) e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} d\mathbf{D} = \int f_w(\mathbf{D}) e^{-trace(\mathbf{BD})} d\mathbf{D} \tag{14}$$

where $\mathbf{B}$ is a 3×3 matrix defined as $b\mathbf{g}\mathbf{g}^T$ and the integration is over the space of positive-definite matrices. Note that the exponent in the right side of Eq. 14 is written in the equivalent form of the trace of $\mathbf{BD}$. In this form, Eq. 14 can be seen as the Laplace transform (in the case of matrices) of the Wishart distribution [14]. This Laplace transform integral can be computed analytically as it has been shown in [14], giving the following expression

$$S/S_0 = |\mathbf{I} + \mathbf{B}\mathbf{\Sigma}|^{-p} = |1 + b\mathbf{g}^T \mathbf{\Sigma} \mathbf{g}|^{-p}. \tag{15}$$

Equation 15 is a novel model for the DW-MRI signal attenuation distinct from the commonly used Eq. 1. The latter was however shown to be a limiting case of Eq. 15 by setting $\mathbf{\Sigma} = \mathbf{D}/p$ and $p \to \infty$. Thus, the model in equation 15 is a generalization of the 35 year old and popular Stejskal-Tanner model for MR signal decay. Another interesting observation is that it is also a generalization of the multi-compartmental (bi-Gaussian etc.) models [22] popular in literature.

The obtained model is still incapable of approximating complex local geometries of the tissue due to the fact that the mixing density was limited to the case of a single Wishart distribution. This problem can be solved by setting the mixing density to be a mixture of N Wishart distributions. Note however that, we still have a continuous approximation and not a discrete approximation. In this case, Eq. 14 becomes

$$S/S_0 = \int \sum_{i=1}^{N} w_i f_w(\mathbf{D}|\mathbf{\Sigma}_i, p) e^{-b\mathbf{g}^T \mathbf{D} \mathbf{g}} d\mathbf{D} = \sum_{i=1}^{N} w_i |1 + b\mathbf{g}^T \mathbf{\Sigma}_i \mathbf{g}|^{-p} \tag{16}$$

where w_i are the mixing weights and $\mathbf{\Sigma}_i$ are the corresponding scaling parameters of the Wishart distributions. For simplicity the shape parameter p was taken to be the same in all mixing components.

Here we should note that although Eq. 16 can be seen as a discrete mixture of functions in the form $|1 + b\mathbf{g}^T \mathbf{\Sigma}_i \mathbf{g}|^{-p}$, it was derived as a continuous mixture of signal attenuations modeled by the Stejskal-Tanner Eq. 1. Hence, the number of mixing components N should not be interpreted as the number of underlying distinct fiber populations but as the resolution of the mixing density $f(\mathbf{D})$. Theoretically, any distribution of positive-definite matrices can be arbitrarily approximated by taking an appropriately large number N in the mixture of Wisharts.

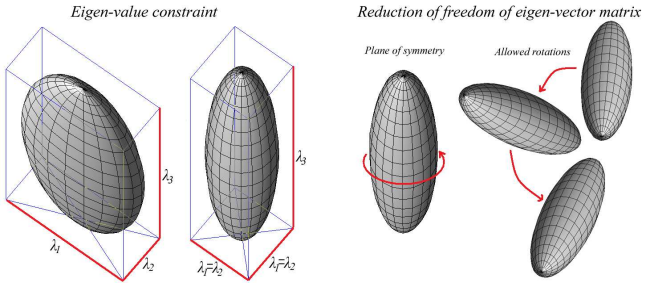

Fig. 3. Illustration of the dimensionality reduction of 3×3 symmetric positive-definite matrices, shown here as Gaussian ellipsoids. Left: We assume the two smallest eigen-values λ_1 and λ_2 to be equal due to the cylindrical geometry of fibers. Right: Reduction of the space of rotations due to symmetry.

The set of the positive-definite matrices $\mathbf{\Sigma}_i$ must be constructed in such a way that the full space of 3×3 positive-definite matrices is appropriately sampled. However, the space of symmetric positive-definite matrices is a 6-dimensional space, hence it cannot be efficiently sampled by a computationally feasible manner. In order to overcome this issue, further assumptions about the fiber geometry must be made in order to reduce the above space to those matrices $\mathbf{\Sigma}_i$ which are practically meaningful in our particular application.

If we express the matrix $\mathbf{\Sigma}$ using the spectral decomposition as $\mathbf{v\Lambda v}^T$ (where $\mathbf{v}$ is an orthogonal eigen-vector matrix and $\mathbf{\Lambda}$ is a diagonal matrix of eigen-values), 3 out of the 6 degrees of freedom are in the three eigenvalues λ_1, λ_2, λ_3 and the other 3 degrees are in the orthogonal eigen-vector matrix. By considering the approximately cylindrical geometry of the fibers, we can reduce our solution to those matrices $\mathbf{\Sigma}$ whose two smallest eigen-values are equal [14], i.e. elimination of 1 degree of freedom. Furthermore, due to the rotational symmetry of those matrices along the plane defined by the two smallest eigenvectors, the dimensionality of the orthogonal eigen-vector matrices is reduced to 2 [14], i.e. elimination of 1 additional degree of freedom. The above assumptions (illustrated in Fig. 3.1) produce a 4-dimensional space whose sampling is practically more feasible and is employed in the experiments presented in Sec. 3.4.

In the next section, we review another special case of the unified de-convolution framework, by using the von Mises-Fisher distribution instead of the Wishart.

3.2 The Mixture of von Mises-Fisher model

The von Mises-Fisher (vMF) distribution is the special case in 3 dimensions of the von Mises distribution of unit vectors. The vMF distribution is the analogous of the Gaussian distribution on the space of 3-dimensional unit vectors and it has the following probability density function

$$f_{vMF}(\mathbf{x}|\mu,\kappa) = \frac{\kappa}{4\pi sinh\kappa} e^{\kappa\mu^T\mathbf{x}} \tag{17}$$

where $\mathbf{x}$ and μ are unit vectors, μ is the center of the distribution and κ is a positive scalar that controls the concentration of the probability. Figure 3.2 shows examples of the von Mises-Fisher distribution for different concentration values.

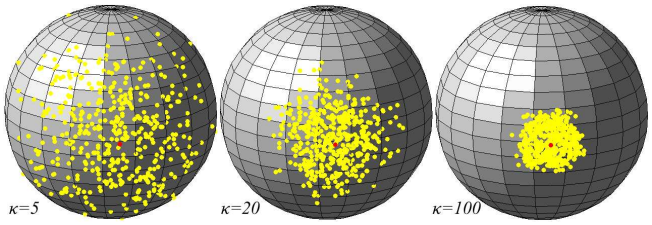

Fig. 4. Samples drawn from von Mises-Fisher distributions using different concentrations κ.

In the area of diffusion-weighted MR imaging, the vMF distribution has been used by McGraw et al. [36, 37] for approximating orientation distribution functions (ODF) and more recently mixtures of hyperspherical vMF have been used for high angular resolution DW-MRI approximation [38]. Here we should note that the magnitude of the DW-MR signal as well as the estimated ODF, the displacement probability and the diffusivity function are all antipodally symmetric, i.e. $f(x) = f(-x)$. Since the vMF distribution function is not antipodally symmetric, a symmetrised vMF expression was employed in a mixture of vMF distributions [36] in order to model the ODF at each voxel of a DW-MRI data set as shown in the following equation.

$$\sum_{i=1}^{N} w_i \left[f_{vMF}(\mathbf{x}|\mu_i, \kappa_i) + f_{vMF}(-\mathbf{x}|\mu_i, \kappa_i) \right] / 2 \tag{18}$$

where w_i are the unknown mixing weights, which are non-negative and sum up to 1 in order the obtained mixture of vMFs (Eq. 18) to be also a probability density function. N is the number components and is assumed to be predefined, which is the main drawback of this model.

The above discrete mixture of vMFs can be extended using a continuous mixture of vMFs [15] by following similar reasoning with the formulation of the continuous mixture of Wisarts distributions, discussed in the previous section. In general, a continuous mixture of vMFs can be used to approximate any spherical function. The diffusion-weighted MR measurements when acquired over a single sphere of the $\mathbf{q}$-space, i.e. with constant b-value and varying diffusion gradient orientation $\mathbf{g}_i$, can be approximated by a spherical function model such as the

continuous mixture of vMFs given by the expression

$$S/S_0 = \int f(\mu) \left[f_{vMF}(\mathbf{g}|\mu, 1) + f_{vMF}(-\mathbf{g}|\mu, 1) \right] d\mu = \int f(\mu) \frac{cosh(\mu^T \mathbf{g})}{4\pi sinh(1)} d\mu \tag{19}$$

where $f(\mu)$ is a mixing probability density function and the integration is over the unit sphere, i.e. the space of unit vectors μ. Since $f(\mu)$ is also a probability in the space of unit vectors, it can be modeled by the vMF density function (Eq. 17), or in the more general case by a mixture of vMFs. In this case, the MR signal attenuation is expressed by the following continuous mixture model

$$S/S_0 = \int \sum_{i=1}^{N} w_i f_{vMF}(\mu|\mathbf{v}_i, \kappa) \frac{cosh(\mu^T \mathbf{g})}{4\pi sinh(1)} d\mu \tag{20}$$

where N is the resolution of the mixing density, i.e. the number of unit vectors $\mathbf{v}_i$ spanning the unit hemi-sphere. Note that the same concentration parameter κ was used in all the components of the mixture for simplicity.

By substituting Eq. 17 into Eq. 20 and taking the summation out of the integral, we obtain a sum of integrals in the form of Laplace transforms. These integrals can be computed analytically, as it was shown in [16, 15] obtaining finally the model

$$S/S_0 = \sum_{i=1}^{N} \frac{w_i \kappa}{4\pi sinh(1) sinh(\kappa)} \left[\frac{sinh(\parallel \kappa \mathbf{v}_i - \mathbf{g} \parallel)}{\parallel \kappa \mathbf{v}_i - \mathbf{g} \parallel} + \frac{sinh(\parallel \kappa \mathbf{v}_i + \mathbf{g} \parallel)}{\parallel \kappa \mathbf{v}_i + \mathbf{g} \parallel} \right]. \tag{21}$$

Theoretically, any spherical function $S(\mathbf{g})$ can be arbitrarily accurately approximated by Eq. 21 by using an appropriately large number N for the resolution of mixing density. As it was pointed out also in Sec. 3.1, although Eq. 21 is expressed in the form of a discrete mixture, the approximation is still a continuous mixture of symmetrized vMF distributions since it was derived from Eq. 20.

3.3 Estimation from DW-MRI data

In this section, we study methods for fitting to DW-MRI data a mixture model such as the mixture of Wisharts or the mixture of von Mises-Fisher distributions discussed in Sec. 3.1 and 3.2 respectively. In both cases, the goal of the fitting procedure is to estimate the unknown mixing weights w_i such that the squared distance between the given data and the model is minimized.

Having acquired a set of M diffusion-weighted MR images, the goal is to fit either one of the mixture models described in the previous sections. At each voxel of the acquired images a spherical function modeled by the selected mixture model is fitted to the data by minimizing the following sum of squares energy

$$E(w_1, ..., w_N) = \sum_{j=1}^{M} \left(S_j/S_0 - \sum_{i=1}^{N} w_i S(b, \mathbf{g}_j|\mathbf{X}_i) \right)^2 \tag{22}$$

where S_j are the M acquired diffusion-weighted images associated with b-value b and magnetic gradient direction $\mathbf{g}_j$, and S_0 is the acquired image without diffusion weighting. The expression $\sum_{i=1}^{N} w_i S(b, \mathbf{g}_j|\mathbf{X}_i)$ is the general form of Eq. 16 and Eq. 21 for the case of the mixture of Wisharts and the mixture of von Mises-Fisher respectively, where $\mathbf{X}_i$ denotes the parameters of each model, i.e. $\mathbf{\Sigma}_i$ and $\mathbf{v}_i$ respectively. The energy function is minimized with respect to the unknown mixing weights w_i, $i = 1...N$.

Equation 22 can be rewritten in the form of an over determined linear system $\mathbf{Aw} = \mathbf{b}$, where $\mathbf{A}$ is an $M \times N$ matrix whose elements are $A_{j,i} = S(b, \mathbf{g}_j|\mathbf{X}_i)$, $\mathbf{w}$ is a N-dimensional vector of unknowns w_i, and $\mathbf{b}$ is a N-dimensional vector whose elements are the acquired signal attenuations $b_j = S_j/S_0$. This over determined linear system can be solved in a least square sense, whose solution will correspond to the solution obtained by minimizing Eq. 22. In the case of the mixture of Wisharts model, the elements of matrix $\mathbf{A}$ are given by

$$A_{j,i} = |1 + b\mathbf{g}_j^T \mathbf{\Sigma}_i \mathbf{g}_j|^{-p} \tag{23}$$

while in the case of the mixture of von Mises-Fisher model, the elements are

$$A_{j,i} = \frac{\kappa}{4\pi sinh(1) sinh(\kappa)} \left[\frac{sinh(\parallel \kappa\mathbf{v}_i - \mathbf{g}_j \parallel)}{\parallel \kappa\mathbf{v}_i - \mathbf{g}_j \parallel} + \frac{sinh(\parallel \kappa\mathbf{v}_i + \mathbf{g}_j \parallel)}{\parallel \kappa\mathbf{v}_i + \mathbf{g}_j \parallel} \right]. \tag{24}$$

Different methods for solving the obtained linear system have been compared extensively by Jian and Vemuri in [39]. According to the results presented in [39] the best results are obtained by using the NNLS algorithm. Here, we should note that theoretically the weights w_i are non negative and they sum up to 1, since they were introduced as the mixing components in a probability distribution function (see Sec. 3.1 and 3.2). However, due to inaccuracies in measuring the signal attenuation ratio S_j/S_0, the sum of the estimated weights may not be 1, although each w_i is greater than or equal to zero. In this case the weights can be normalized by dividing the vector $\mathbf{w}$ as well as the vector $\mathbf{b}$ with the normalizing factor Σw_i.

After having fitted the mixture model to the data, the reconstructed signal $S(\mathbf{g})$ at each voxel can be plotted as a spherical function, i.e. over the space of unit vectors $\mathbf{g}$. However, in order to understand the estimated diffusivity pattern which corresponds to the underlying fiber structure, the displacement probability function $P(\mathbf{r})$ must be computed, whose peaks corresponds to the orientation of distinct fiber distributions. The displacement probability can be estimated by evaluating the Fourier integral given by Eq. 2.

In both mixture model cases this integral cannot be evaluated analytically. In the case of the mixture of Wisharts the integral can be approximated by $\Sigma w_i P_i(\mathbf{r})$, where $P_i(\mathbf{r})$ is given by Eq. 3. The error introduced by this approximation decreases with increasing parameter values for p in the Wishart distribution and becomes zero when $p \to \infty$. Jian et al. [14] have used the value $p = 2$ based on the analogy between Eq. 15 and the Debye-Porod law of diffraction [40]

for porous media, and it has been shown that the accuracy of the approximated displacement probability in estimating fiber orientations is higher than that of other existing techniques.

In the case of the mixture of von Mises-Fisher distributions, the Fourier integral cannot be computed analytically either, and no approximation formula has been reported to date. An efficient way to estimate the displacement probability from this model is to use the general method presented in [41] for approximating the probability from a set of DW-MR acquired images. In our particular case of the von Mises-Fisher model, the recovered signal attenuation can be computed from the approximated mixture model by evaluating Eq. 21 for a large set of unit vectors $\mathbf{g}_i$. The set of vectors can be constructed by tessellating the icosahedron on the unit hemi-sphere. After having computed the $S_i = S(\mathbf{g}_i)/S_0$ the set of S_i can be considered as a new high angular resolution DW-MR data set and used by the algorithm in [41] for estimating the displacement probability as a 4^{th}-order Cartesian tensor.

Finally, after having estimated the displacement probability $P(\mathbf{r})$ it can be plotted as a spherical function over displacement vectors $\mathbf{r}$ of same magnitude. The orientations $\mathbf{r}$ that correspond to the peaks of the probability function can be computed by following the maxima of this spherical function. In order to compute all the peaks of the possibly multi-lobed displacement probability, the gradient ascent is initialized in multiple different orientations $\mathbf{r}$. The obtained orientations of maximum water molecule displacement probability can be further used by a fiber tracking method for computing complicated fiber structures such as crossing and splaying fibers [42, 43].

In the next section several examples of multi-fiber reconstruction from synthetic as well as real diffusion-weighted MR data are shown. A table of comparison between the discussed mixture models is also presented and it is supported by quantitative experimental results.

3.4 Experimental results

In this section we present experimental results obtained by applying the multi-fiber reconstruction methods presented previously, using synthetic and real DW-MRI datasets. The synthetic data set was produced by simulating the signal response on a fiber of cylindrical geometry using the realistic model in [44].

First, we compared the basis derived from the MoW and MovMF models (Eq. 23 and 24 respectively) with the simulated signal response, in order to demonstrate the ability of the basis derived from the MoW model in approximating the true DW-MRI signal. In this experiment we first defined a fiber orientation $\mathbf{v}$ and simulated the DW-MRI signal for various diffusion magnetic gradient directions $\mathbf{g}_i$. Then, we evaluated the basis functions given by Eq. 23 and Eq. 24 respectively for various $\mathbf{g}_i$. The primary eigenvector of $\mathbf{\Sigma}$ in Eq. 23 was taken to be parallel to the fiber orientation $\mathbf{v}$, and we used the parameters $p = 2$ and $\kappa = 10$ for the MoW and MovMF models respectively. The plots of the three functions are shown in Fig. 5.

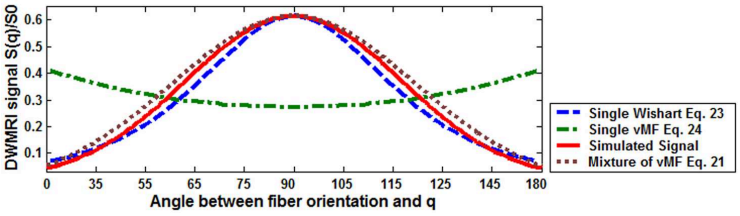

Fig. 5. Plots of the DW-MRI signal response for a single fiber when using the MoW and MovMF basis function-based models respectively.

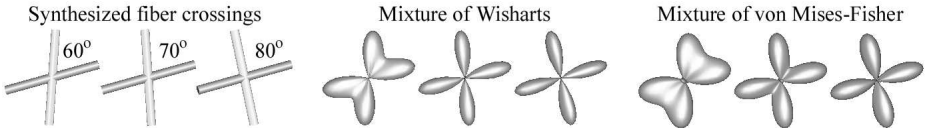

Fig. 6. Multi-fiber reconstruction example. Left: orientations of the synthetic fiber crossings. Center and Right: The corresponding displacement probabilities estimated by the MoW and the MovMF models.

By observing Fig. 5 we can see that the MoW model better approximates the simulated DW-MRI signal in comparison to the MovMF model-based approximation. Therefore, a single fiber response can be approximated by employing only one basis in the MoW model (i.e. only one non-zero w_i in Eq. 16), while on the other hand an appropriate mixture of the MovMF basis must be employed (i.e. several non-zero w_i in Eq. 17). This experimentally validates the MoW model as a better suited approximation over the MovMF model which is simply a general set of basis for approximating spherical functions but not necessarily well suited to represent the MR-signal decay. As an interesting side, note that the MovMF model has been also used in approximating facial apparent bidirectional reflectance distribution functions [16], which are also spherical functions.

In order to demonstrate the ability of the methods in resolving fiber orientations in the presence of fiber crossings, we simulated the DW-MRI signal [44] for the cases where two fibers are crossing each other and form an angle of 60^o, 70^o and 80^o degrees respectively (shown in the left plate of Fig. 6). The simulated signal, consisting of 81 measurements (at each crossing location) corresponding to different gradient directions, was approximated by the MoW and MovMF methods. The estimated displacement probabilities for the two methods are shown in the central and right plates of Fig. 6 respectively. In the MoW case the resulting spherical function plots have sharper lobes than those in the MovMF model. The sharpness of the plots in the MoW model is due to the ability of the model to better approximate the true DW-MRI signal and the existence of an analytic

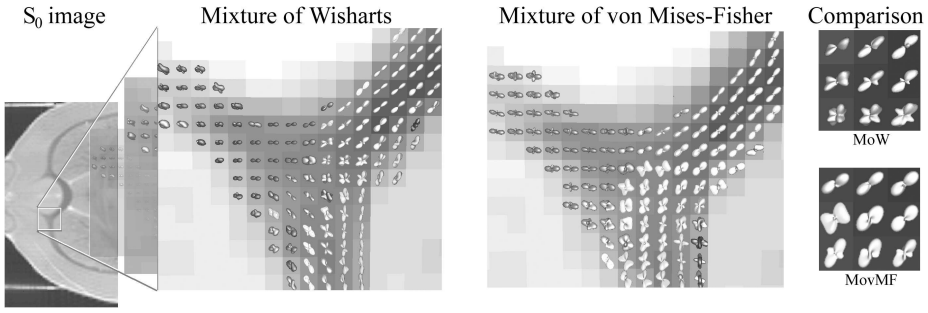

Fig. 7. Application of the MoW and MovMF methods to a rat brain data set. The depicted ROI shows fibers from the cingulum and corpus callosum crossing over.

form for approximating the displacement probability, which improves the accuracy of computation. A detailed comparison of the accuracy of each method in estimating fiber orientations can be found in [15].

Further, we applied the methods to a real data set from an excised perfusion-fixed rat brain. The DW-MRI data set was acquired using a pulsed gradient spin echo pulse sequence with 32 diffusion gradients and $b \simeq 1250\ s/mm^2$. The region of interest depicted in Fig. 7 contains intersecting fibers from cingulum and corpus callosum. By observing the computed displacement probabilities, as anticipated, we can see that both methods estimated crossings at voxels located in the center of this region.

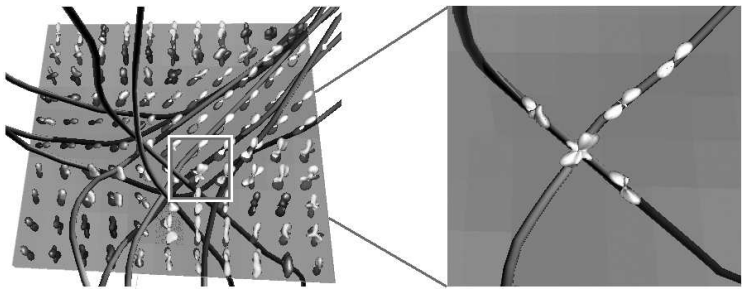

Fig. 8. Fiber tracking example by following the intra-voxel fiber orientations estimated by the MoW method.

After having estimated the water molecule displacement probabilities as shown in Fig. 7, one can perform fiber tracking by applying a fiber tracking method to the field of probabilities [4, 43]. Figure 8 shows some of the estimated

fibers from the cingulum and the corpus callosum crossing over. A fiber crossing is more clearly depicted in the right plate of the same figure, with the estimated displacement probability plots superimposed.

	Wishart model	von Mises-Fisher model
Model specialized in:	DW-MRI	spherical functions
Integration space:	$\mathcal{P}_n$	S_2
Pre-defined shape parameter:	p	κ
The space of mixing densities:	not spanned completely	spanned uniformly
Signal attenuation is modeled:	very accurately	poorly

Table 1. Properties of the two models discussed in Sec. 3.1 and 3.2 [14, 15].

Finally we present in Table 1 a summary of the properties of the two methods studied in this section. First, the MoW model was defined (in Eq. 14) as a continuous mixture of Gaussians, where each Gaussian is in the form of the DW-MRI signal attenuation defined in Eq. 1. This produces a model (in Eq. 23) which is natural for DW-MRI, while on the other hand, the MovMF model is a general basis (Eq. 24) for approximating any spherical function. Furthermore, the basis (Eq. 23) derived in the MoW model, approximates closely the true DW-MRI signal response obtained by using the realistic simulation model in [44]. However, as it was expected, the basis (Eq. 23) derived in the MovMF case fails to approximate the signal response since it is a general spherical function basis and not specifically tailored to approximate the MR signal response from single fibers.

4 Conclusions

In this chapter we reviewed several information theoretic methods for DTI and DW-MRI processing and analysis. In the case of DTI, we reviewed an information theoretic dissimilarity measure between two tensors and then employed it in a tensor field segmentation framework. The main advantages of this dissimilarity measure over other existing metrics is that it has an analytic form and it also approximates the Riemannian geodesic distance between two nearby tensors [5]. Additionally, unlike most other non-Euclidean measures, it provides a closed form expression for computing the mean tensor of a set of tensors and hence is very useful as a computationally efficient tensor interpolation technique.

Furthermore, we studied two methods for multi-fiber reconstruction by modeling the DW-MRI signal as a continuous mixture of basis. In the first method the mixing density was set to be mixture of Wishart distributions, while the von Mises-Fisher distribution of unit vectors was employed in the other method. The mixture of Wisharts model is a natural choice for modeling the MR signal response in the presence of multiple fibers in a voxel since this model constitutes

a continuous mixture of responses from single fibers. In contrast, the continuous mixture of vMFs is a natural choice for expressing any multi-lobed spherical functions which may or may not necessarily have anything to do with MR signal responses obtained in the presence of multiple fibers in a voxel. For more details on the mixture of Wishart's and the mixture of vMFs model we refer the reader to [14, 39, 17] and [15] respectively.

5 Acknowledgments

Authors thank Drs. Paul R. Carney and Thomas H. Mareci respectively for providing the rat brain data set and Drs. T. M. Shepherd and Evren Özarslan respectively for providing the rat hippocampus data set.

References

1. Stejskal, E.O., Tanner, J.E.: Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. J. Chem. Phys. **42 (1)** (1965) 288–292
2. Basser, P., Mattiello, J., Lebihan, D.: Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo. J. Magn. Reson. B **103** (1994) 247–254
3. Callaghan, P.T.: Principles of Nuclear Magnetic Resonance Microscopy. Clarendon Press, Oxford (1991)
4. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using dt-mri data. Magnetic Resonance in Medicine **44**(4) (2000) 625–632
5. Wang, Z., Vemuri, B.C.: DTI segmentation using an information theoretic tensor dissimilarity measure. IEEE Transactions on Medical Imaging **24**(10) (2005) 1267–1277
6. Lenglet, C., Rousson, M., Deriche, R., Faugeras, O.: Statistics on the manifold of Multivariate Normal Distributions: Theory and Applications to Diffusion Tensor MRI processing. J. Math. Imaging Vis. **25** (2006) 423–444
7. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. International Journal of Computer Vision **65** (2005)
8. Pennec, X.: Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. IEEE Workshop on Nonlinear Signal and Image Processing (1999)
9. Fletcher, P., Joshi, S.: Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. Proc. of CVAMIA (2004) 87–98
10. Barmpoutis, A., Vemuri, B., Shepherd, T., Forder, J.: Tensor splines for interpolation and approximation of DT-MRI with applications to segmentation of isolated rat hippocampi. Medical Imaging, IEEE Transactions on **26**(11) (Nov. 2007) 1537–1546
11. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Fast and Simple Calculus on Tensors in the Log-Euclidean Framework. In: Proceedings of MICCAI. LNCS (2005) 259–267
12. Kindlmann, G., Estepar, R., Niethammer, M., Haker, S., Westin, C.F.: Geodesic-loxodromes for diffusion tensor interpolation and difference measurement. MICCAI **10** (2007) 1–9
13. Tuch, D.S., Reese, T.G., Wiegell, M.R., Wedeen, V.J.: Diffusion MRI of complex neural architecture. Neuron (40) (2003) 885–895

14. Jian, B., Vemuri, B.C., Özarslan, E., Carney, P.R., Mareci, T.H.: A novel tensor distribution model for the diffusion-weighted MR signal. NeuroImage **37**(1) (2007) 164–176
15. Kumar, R., Barmpoutis, A., Vemuri, B.C., Carney, P.R., Mareci, T.H.: Multi-fiber reconstruction from DW-MRI using a continuous mixture of von mises-fisher distributions. Computer Vision and Pattern Recognition Workshops, 2008. CVPR Workshops 2008. IEEE Computer Society Conference on (June 2008) 1–8
16. Barmpoutis, A., Kumar, R., Vemuri, B.C., Banerjee, A.: Beyond the Lambertian assumption: A generative model for apparent BRDF fields of faces using anti-symmetric tensor splines. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on (June 2008) 1–6
17. Jian, B., Vemuri, B.C.: A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. TMI **26**(11) (2007) 1464–1471
18. Ozarslan, E., Mareci, T.H.: Generalized diffusion tensor imaging and analytical relationships between DTI and HARDI. MRM **50**(5) (Nov 2003) 955–965
19. Barmpoutis, A., Jian, B., Vemuri, B.C., Shepherd, T.M.: Symmetric positive 4th order tensors and their estimation from diffusion weighted mri. In Karssemeijer, N., Lelieveldt, B.P.F., eds.: IPMI. Volume 4584 of Lecture Notes in Computer Science., Springer (2007) 308–319
20. Frank, L.R.: Characterization of anisotropy in high angular resolution diffusion-weighted MRI. Magn Reson Med **47**(6) (2002) 1083–1099
21. Özarslan, E., Shepherd, T.M., Vemuri, B.C., Blackband, S.J., Mareci, T.H.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). NeuroImage (2006)
22. Tuch, D.S.: Q-ball imaging. Magn. Reson. Med. **52**(6) (2004) 1358–1372
23. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast and robust analytical q-ball imaging. MRM **58**(3) (2007) 497–510
24. Wassermann, D., Descoteaux, M., Deriche, R.: Diffusion maps clustering for magnetic resonance q-ball imaging segmentation. Journal of Biomedical Imaging **8**(3) (2008) 1–12
25. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from DW-MRI data using spherical deconvolution. NeuroImage **23**(3) (2004) 1176–1185
26. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution.
27. Tournier, J.D., Yeh, C.H., Calamante, F., Cho, K.H., Connelly, A., Lin, C.P.: Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. NeuroImage **42**(2) (2008) 617–625
28. Alexander, D.C.: Maximum entropy spherical deconvolution for diffusion mri. Inf. Process Med. Imaging (2005) 76–87
29. Wedeen, V., Wang, R.P., Schmahmann, J.D., Benner, T., Tseng, W.Y., Dai, G., Pandya, D.N., Hagmann, P., D'Arceuil, H., de Crespigny, A.J.: Diffusion spectrum magnetic resonance imaging (dsi) tractography of crossing fibers. Neuroimage **41**(4) (2008) 1267–1277
30. Hasan, K.M., Gupta, R.K., Santos, R.M., Wolinsky, J.S., Narayana, P.A.: Diffusion tensor fractional anisotropy of the normal-appearing seven segments of the corpus callosum in healthy adults and relapsing-remitting multiple sclerosis patients. Journal of Magnetic Resonance Imaging **21**(6) (2005) 735–743

31. van Gelderen, P., de Vleeschouwer, M.H.M., DesPres, D., Pekar, J., van Zijl, P.C.M., Moonen, C.T.W.: Water diffusion and acute stroke. Magnetic Resonance in Medicine **31**(2) (1994) 154–163
32. Cover, T.M., Thomas, J.A.: Elements of Information Theory. John Wiley and Sons Inc. (2001)
33. Wang, Z., Vemuri, B., Chen, Y., Mareci, T.: A Constrained Variational Principle for Direct Estimation and Smoothing of the Diffusion Tensor Field from DWI. IPMI (2003) 660–671
34. Wang, Z., Vemuri, B., Chen, Y., Mareci, T.: A Constrained Variational Principle for Direct Estimation and Smoothing of the Diffusion Tensor Field from complex DWI. TMI **23**(8) (2004) 930–939
35. Letac, G., Massam, H.: Quadratic and inverse regressions for Wishart distributions. Ann. Stat. **2**(26) (1998) 573595
36. McGraw, T., Vemuri, B.C., Yezierski, R., Mareci, T.: von Mises-Fisher mixture model of the diffusion ODF. In: ISBI. (2006) 65–68
37. McGraw, T., Vemuri, B.C., Yezierski, R., Mareci, T.: Segmentation of high angular resolution diffusion MRI modeled as a field of von Mises-Fisher mixtures. In: ECCV. (2006) 463–475
38. Bhalerao, A., Westin, C.F.
39. Jian, B., Vemuri, B.C.: Multi-fiber reconstruction from diffusion MRI using mixture of Wisharts and sparse deconvolution. In: IPMI. (2007) 384–395
40. Sen, P., Hrlimann, M., de Swiet, T.: DebyePorod law of diffraction for diffusion in porous media. Phys. Rev. **51**(1) (1995) 601–604
41. Barmpoutis, A., Vemuri, B.C., Forder, J.R.: Fast displacement probability profile approximation from hardi using 4th-order tensors. In Proceedings of ISBI08: IEEE International Symposium on Biomedical Imaging (14 - 17 May 2008) 911–914
42. Barmpoutis, A., Vemuri, B.C., Howland, D., Forder, J.R.: Extracting tractosemas from a displacement probability field for tractography in DW-MRI. In LNCS, (Springer) Proceedings of MICCAI08: Int. Conf. on Medical Image Computing and Computer Assisted Intervention (6-10 September 2008) 9–16
43. Deriche, R., Descoteaux, M.: Splitting tracking through crossing fibers: Multidirectional q-ball tracking. ISBI (2007) 756–759
44. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. J. Magn. Reson. **A (117)** (1995) 94–97