Multi-fiber Reconstruction from Diffusion MRI Using Mixture of Wisharts and Sparse Deconvolution

Bing Jian and Baba C. Vemuri

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

Abstract. In this paper, we present a novel continuous mixture of diffusion tensors model for the diffusion-weighted MR signal attenuation. The relationship between the mixing distribution and the MR signal attenuation is shown to be given by the Laplace transform defined on the space of positive definite diffusion tensors. The mixing distribution when parameterized by a mixture of Wishart distributions (MOW) is shown to possess a closed form expression for its Laplace transform, called the Rigaut-type function, which provides an alternative to the Stejskal-Tanner model for the MR signal decay. Our model naturally leads to a deconvolution formulation for multi-fiber reconstruction. This deconvolution formulation requires the solution to an ill-conditioned linear system. We present several deconvolution methods and show that the nonnegative least squares method outperforms all others in achieving accurate and sparse solutions in the presence of noise. The performance of our multi-fiber reconstruction method using the MOW model is demonstrated on both synthetic and real data along with comparisons with state-of-the-art techniques.

1 Introduction

As the only noninvasive and in vivo imaging method available today which allows neural tissue architecture to be probed at a microscopic scale, diffusion-weighted magnetic resonance imaging (DW-MRI) provides unique clues to the microstructure of tissues and to changes associated with various physiological and pathological states. By producing quantitative data of water molecule motion that naturally occurs in brain tissues as part of the physical diffusion process, DW-MRI has also been used to map the fiber orientation in the brain white matter tracks. This valuable information can be further exploited for neuronal connectivity inference and brain developmental studies [1].

Assuming a displacement probability characterized by an oriented Gaussian probability distribution function, diffusion tensor MRI (DT-MRI) [2] provides a relatively simple way of quantifying diffusional anisotropy as well as predicting the local fiber direction within the tissue from multidirectional DW-MRI data. However, the major drawback of diffusion tensor MRI is that it can only reveal a single fiber orientation in

* This research was in part supported by NIH EB007082, NIH NS42075.We thank Dr. Evren Özarslan for many technical discussions. The MRI data were obtained at the Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) facility in the McKnight Brain Institute at the University of Florida.

N. Karssemeijer and B. Lelieveldt (Eds.): IPMI 2007, LNCS 4584, pp. 384–395, 2007.

each voxel and fails in voxels with orientational heterogeneity [3], which makes DT-MRI an inappropriate model for use in the presence of multiple fibers within a voxel.

This limitation of diffusion tensor model has prompted interest in the development of both improved image acquisition strategies and more sophisticated reconstruction methods. Both spherical harmonic expansion [4] and the equivalent higher order tensor model [5] have been used to represent the diffusivity profile based on the Stejskal-Tanner mono-exponential attenuation model. Knowing that the peaks of the diffusivity profile do not necessarily yield the orientations of the distinct fiber populations, a number of model-independent approaches attempt to transform the multi-directional signals into a probability function describing the probability of water molecular displacement. The q-ball imaging (QBI) method approximates the radial integral of the displacement probability distribution function (PDF) by the spherical Funk-Radon transform [6]. More recent studies have expressed QBI's Funk-Radon transform in a spherical harmonic basis [7, 8, 9]. Diffusion spectrum imaging (DSI) can measure the microscopic diffusion function directly based on the Fourier relation between the diffusion signal and the diffusion function, but is limited by the time-intensive q-space sampling burden [10]. The diffusion orientation transform (DOT) transforms the diffusivity profiles into probability profiles by explicitly expressing the Fourier relation in spherical coordinates and evaluating the radial part of the integral analytically [11].

Some multi-compartmental models have also been used to model the diffusion-attenuated MR signal using a finite mixture of Gaussians [3, 12, 13]. A continuous extension of the finite discrete mixture model is the spherical deconvolution method [14]. Compared to the multi-compartment models, the spherical deconvolution framework has two significant advantages. *First*, it is not required to specify the number of underlying fiber populations before deconvolution while this number has to be known in order to build the multi-compartment models. *Second*, the spherical deconvolution methods often yield a linear system which can be solved efficiently while the multi-compartment models usually involve the expensive nonlinear fitting. Recognizing these merits of the spherical deconvolution framework, recently many researchers have proposed a number of variants of spherical deconvolution approaches [7, 15, 16] with different choices of basis functions, deconvolution kernels and regularization schemes.

In this paper, we present a novel probabilistic model that significantly generalizes the traditional diffusion tensor model [2]. First, we assume that each voxel is associated with an underlying probability distribution defined on the space of diffusion tensors (the manifold of 3×3 positive-definite matrices). Conceptually, our model can be viewed as a natural extension of the multiple-compartment models [3, 13]. Moreover, this extension relates the continuous mixture model to MR signal attenuation through a Laplace transform defined for matrix-variate functions. It is worth noting that the Laplace transform can be evaluated in closed form for Wishart distributions and the resulting closed form leads to a Rigaut-type function which has been used in the past to explain the MR signal decay [17]. Our model naturally leads to a deconvolution formulation of the multi-fiber reconstruction problem, where the deconvolution kernel is the Laplace transform kernel and the basis functions are Wishart distributions. We develop an efficient and robust scheme for reconstructing the multiple fiber bundles using the proposed model and show several comparisons with other state-of-the-art methods.

2 Theory

By generalizing the discrete Gaussian mixture model to the continuous case, we postulate that at each voxel there is an underlying probability measure associated with the manifold of $n \times n$ symmetric positive-definite matrices, $\mathcal{P}_n$ (by default $\mathcal{P}_3$). Let $f(\mathsf{D})$ be its density function with respect to some carrier measure $\mathrm{d}\mathsf{D}$ on $\mathcal{P}_n$. Then the diffusion weighted MR signal $S(\boldsymbol{q})$ can be modeled as:

$$S(\boldsymbol{q})/S_0 = \int_{\mathcal{P}_n} f(\mathsf{D}) \, \exp[-b\boldsymbol{g}^T \mathsf{D} \boldsymbol{g}] \, \mathrm{d}\mathsf{D} \, , \tag{1}$$

where S_0 is the signal in the absence of diffusion weighting gradient, $\boldsymbol{q}$ encodes the magnitude (G) and direction ($\boldsymbol{g}$) of the diffusion sensitizing gradients, and b is the diffusion weighting factor depending on the strength as well as the effective time of diffusion. Note that Eq. (1) implies a continuous form of mixture model with $f(\mathsf{D})$ being a mixing density over the components in the mixture. Clearly, our model simplifies to the diffusion tensor model when the underlying probability measure is the Dirac measure.

Since $b\,\boldsymbol{g}^T \mathsf{D} \boldsymbol{g}$ in Eq.(1) can be replaced by $\mathrm{trace}(\mathsf{BD})$ where $\mathsf{B} = b\,\boldsymbol{g}\boldsymbol{g}^T$, the equation (1) can be expressed as the Laplace transform (matrix variable case) [18]:

$$S(\boldsymbol{q})/S_0 = \int_{\mathcal{P}_n} \exp(-\mathrm{trace}(\mathsf{BD})) \, f(\mathsf{D}) \mathrm{d}\mathsf{D} = (\mathscr{L}_f)(\mathsf{B}) \, , \tag{2}$$

where $\mathscr{L}_f$ denotes the Laplace transform of a function f which takes its argument as symmetric positive definite matrices from $\mathcal{P}_n$.

This expression naturally leads to an inverse problem: recovering of a distribution defined on $\mathcal{P}_n$ that best explains the observed diffusion signal $S(\boldsymbol{q})$. This is an ill-posed problem and in general is intractable without prior knowledge of the probabilistic structure. In conventional DT-MRI, the diffusion tensor is usually estimated by solving a linear or nonlinear least squares problem, which amounts to applying the maximum likelihood estimator. Instead our approach views the diffusion tensor as random variable (matrix) belonging to some known distribution family, which allows us to model the uncertainty in the diffusion tensor estimation. Note that in DT-MRI, the diffusion tensor can be interpreted as the concentration matrix (inverse of the covariance matrix) of the Gaussian distribution in the q-space. It is a common practice to put a Wishart distribution (see definition below) prior, on the concentration matrix in multivariate analysis. Moreover, in the case of a Wishart distribution, a closed form expression for the Laplace transform exists and leads to a Rigaut-type asymptotic fractal law for the MR signal decay behavior which has been observed in the past (see explanation below).

Definition 1. *[18] For $\Sigma \in \mathcal{P}_n$ and for p in $\left(\frac{n-1}{2}, \infty\right)$, the Wishart distribution $\gamma_{p,\Sigma}$ with scale parameter Σ and shape parameter p is defined as* [1]

$$\mathrm{d}\gamma_{p,\Sigma}(\mathsf{Y}) = \Gamma_n(p)^{-1} \, |\mathsf{Y}|^{p-(n+1)/2} \, |\Sigma|^{-p} \, e^{-\mathrm{trace}(\Sigma^{-1}\mathsf{Y})} \, \mathrm{d}\mathsf{Y}, \tag{3}$$

where Γ_n is the multivariate gamma function and $|\cdot|$ is the matrix determinant.

[1] Note that the correspondence between this definition and the conventional Wishart distribution $W_n(p, \Sigma)$ is given simply by $\gamma_{p/2,2\Sigma} = W_n(p, \Sigma)$.

The Wishart distribution $\gamma_{p,\Sigma}$ is known to have the closed-form Laplace transform:

$$\int e^{-\mathrm{trace}(\Theta \mathsf{Y})}\, \mathrm{d}\gamma_{p,\Sigma}(\mathsf{Y}) = (1+\mathrm{trace}(\Theta\Sigma))^{-p} \quad \text{where } (\Theta+\Sigma^{-1}) \in \mathcal{P}_n. \tag{4}$$

Let f in (2) be the density function of $\gamma_{p,\Sigma}$ with the expected value $\hat{\mathsf{D}} = p\Sigma$. We have

$$S(\boldsymbol{q}) = S_0\,(1+(b\,\boldsymbol{g}^T\hat{\mathsf{D}}\boldsymbol{g})/p)^{-p}\,. \tag{5}$$

This is a familiar Rigaut-type asymptotic fractal expression [19] implying a signal decay characterized by a power-law which is the expected asymptotic behavior for the MR signal attenuation in porous media. Note that although this form of a signal attenuation curve had been phenomenologically fitted to the diffusion-weighted MR data before [17], until now, there was no rigorous derivation of the Rigaut-type expression used to explain the MR signal behavior as a function of b-value. Therefore, this derivation may be useful in understanding the apparent fractal-like behavior of the neural tissue in diffusion-weighted MR experiments. Also note when p tends to infinity, we have $S(\boldsymbol{q}) \rightarrow S_0\,\exp(-b\boldsymbol{g}^T\hat{\mathsf{D}}\boldsymbol{g})$, which implies that the mono-exponential model can be viewed as a limiting case ($p \rightarrow \infty$) of our model.

The single Wishart distribution model has a drawback in that it can not resolve the intra-voxel orientational heterogeneity due to the single mode of the Wishart distribution. Hence it is natural to use a discrete mixture of Wishart distribution model where the mixing distribution in Eq.(2) is expressed as $\mathrm{d}F = \sum_{i=1}^{N} w_i \mathrm{d}\gamma_{p_i,\Sigma_i}$. In order to make the problem tractable, several simplifying assumptions are made as follows. First, in this model (p_i, Σ_i) are treated as the basis and will be fixed as described below. This leaves us with the weights, $\boldsymbol{w} = (w_i)$, as the unknowns to be estimated. Note that the number of components in mixture, N, only reflects the resolution of the discretization and should not be interpreted as the expected number of fiber bundles. We assume that all the p_i take the same value, $p_i = p = 2$, based on the analogy between the Eq.(5) and Debye-Porod law of diffraction [20] in 3D space. Since the fibers have an approximately cylindrical geometry, it is reasonable to assume that the two smaller eigenvalues of diffusion tensors are equal. In practice, we fix the eigenvalues of $\mathsf{D}_i = p\Sigma_i$ to specified values $(\lambda_1, \lambda_2, \lambda_3) = (1.5, 0.4, 0.4)\mu^2/ms$ consistent with the values commonly observed in the white-matter tracts [3]. This rotational symmetry leads to a tessellation where N unit vectors evenly distributed on the unit sphere are chosen as the principal directions of Σ_i. For K measurements with $\boldsymbol{q}_j$, the signal model equation:

$$S(\boldsymbol{q}) = S_0 \sum_{i=1}^{N} w_i (1+\mathrm{trace}(\mathsf{B}\Sigma_i))^{-p} \tag{6}$$

leads to a linear system $\mathsf{A}\boldsymbol{w} = \boldsymbol{s}$, where $\boldsymbol{s} = (S(\boldsymbol{q})/S_0)$ contains the normalized measurements, A is the matrix with $A_{ji} = (1+\mathrm{trace}(\mathsf{B}_j\Sigma_i))^{-p}$, and $\boldsymbol{w} = (w_i)$ is the weight vector to be estimated.

Like many existing reconstruction methods, our method can be cast into a unified convolution framework as follows:

$$S(\boldsymbol{q})/S_0 = \int_{\mathcal{M}} R(\boldsymbol{q}, x) f(x) \mathrm{d}x \tag{7}$$

In Eq. (7), the signal is expressed as the convolution of a probability density function and a kernel function. The integration is over a manifold $\mathcal{M}$ whose elements x contain the useful information like orientation and anisotropy. The convolution kernel, $R(\boldsymbol{q}, x) : \mathbb{R}^3 \times \mathcal{M} \mapsto \mathbb{R}$ represents the response derived from a single fiber. In order to handle the intra-voxel orientational heterogeneity, the volume fractions are represented by a continuous function $f(x) : \mathcal{M} \mapsto \mathbb{R}$. The deconvolution problem is to estimate the $f(x)$ given the specified $R(\boldsymbol{q}, x)$ and measurements $S(\boldsymbol{q})/S_0$. In literature, $f(x)$ is usually expressed as a linear combination of N basis functions: $f(x) = \sum w_j f_j(x)$. The choices of convolution kernels and basis functions are related to the underlying manifold $\mathcal{M}$. A simple example is to set $\mathcal{M}$ to the unit sphere, which leads to the spherical deconvolution problem [15]. Though involving the manifold of diffusion tensors, several other approaches still reduce to the sphere deconvolution problem since only rotationally symmetric tensors are considered. [14, 13, 7].

The types of basis functions include radial basis functions [15], spherical harmonics [14, 7]. Like in [14, 13, 7], our method uses the standard diffusion tensor kernel. However, it is the Wishart basis function that distinguishes our method from these related methods. It is worth noting that the Wishart basis reduces to the Dirac function on $\mathcal{P}_n$ when $p = \infty$ and thus leading to the tensor basis function method introduced in [13]. The fiber orientation estimated using the continuous axially symmetric tensors (FORECAST) method proposed in [7] also resembles (very closely) our method with the basis function being chosen as the spherical harmonics.

3 Stable, Sparse and Positive Deconvolution

The deconvolution problem can be formulated in a general form of as,

$$\mathsf{A}\boldsymbol{w} = \boldsymbol{s} + \eta, \tag{8}$$

where $\boldsymbol{s}$ contains K measurements $S(\boldsymbol{q})/S_0$, the $K \times N$ matrix $\mathsf{A} = \{A_{ij}\}$ is given by $A_{ij} = \int_M R(\boldsymbol{q}_i, x) f_j(x) \mathrm{d}x$ and η represents certain noise model. Note that the integral to compute the entries of A may have an analytical solution as in our model and others that use the tensor kernel [14, 13, 7], or needs to be numerically approximated as in [15]. But, once the response kernel $R(\boldsymbol{q}, x)$ and the basis function are specified, the matrix A can be fully computed (or approximated) and only $\boldsymbol{w}$, a column vector containing K unknown coefficients, remains to be estimated.

Under the assumption that the measurement errors η are i.i.d. and normally distributed, the maximum likelihood estimate of $\boldsymbol{w}$ naturally leads to the L_2 norm as a measure of goodness of the fit. Without inequality constraints, the corresponding quadratic programming (QP) problem minimizing the residual sum of squares

$$(P_1) \qquad \min \|\mathsf{A}\boldsymbol{w} - \boldsymbol{s}\|^2 \tag{9}$$

can be efficiently solved by solving a linear system using for instance, direct methods when the size of the linear system in Eq. (9) is not large as in our application. The solution in the least squares sense is given by $\boldsymbol{w} = \mathsf{A}^+\boldsymbol{s}$ where $\mathsf{A}^+ = (\mathsf{A}^T\mathsf{A})^{-1}\mathsf{A}^T$ is the pseudoinverse of the A. The advantage of applying the pseudoinverse is in its light computational burden since the matrix A is identical in each voxel and its pseudoinverse only

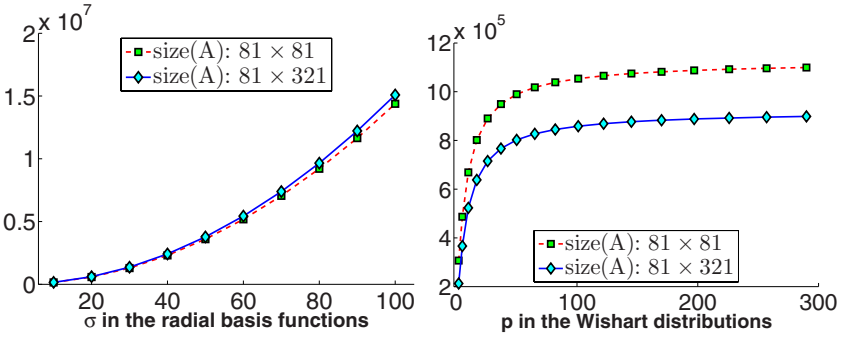

Fig. 1. Left plot shows the case where the A matrices are constructed from the radial basis function and the tensor kernel model as in [15]. Right plot shows the case with a standard diffusion tensor kernel weighted by a mixture of Wisharts. Both assume 81 diffusion gradient directions and two tessellation schemes (81 and 321 directions) are considered for each model.

needs to be computed once. However, this simplicity and efficiency comes at the cost of higher susceptibility to noise, due to the fact that the matrix A is usually ill-conditioned in our application as illustrated in Fig. 1. Efforts at reliable multi-fiber reconstruction in the presence of noise have included low-pass filtering [14], the maximum entropy principle [15] and Tikhonov regularization [16]. In the Tikhonov regularization framework, the problem in Eq. (8) can be formulated as:

$$(P_2) \qquad \min \|\mathsf{A}\boldsymbol{w} - \boldsymbol{s}\|^2 + \alpha \|\mathsf{T}\boldsymbol{w}\|^2 \tag{10}$$

where α is a regularization parameter and T is a regularization operator. In order to penalize the magnitude of the estimates, (P_2) in (10) with T being the identity operator I is often used and yields the relation: $\boldsymbol{w} = (\mathsf{A}^T\mathsf{A} + \alpha\mathsf{I})^{-1}\mathsf{A}^T\boldsymbol{s}$. Recently, a Damped Singular Value Decomposition was used to regularize the fiber orientation distribution [21] where the damping factor α is determined by minimizing the Generalized Cross Validation (GCV) criterion, which provides a simple and objective method, though not really optimal, for choosing the regularization parameter.

In practice, the number of diffusion MR image acquisition sequence, K, is rarely greater than 100. On the other hand, a high resolution tessellation with $N > 100$ is usually taken to obtain an accurate reconstruction. This under-determined linear system has infinite solutions in the least squares sense and usually produces the $\boldsymbol{w}$ with many negative-valued components which are not physically meaningful. Another issue related to this configuration is the sparsity constraint. Since the number of significant spikes in $\boldsymbol{w}$ is indicative of the number of maxima in the displacement probability surfaces, $\boldsymbol{w}$ is expected to have a sparse support. Recently a series of significant research articles have been published by Candés and collaborators (see [22] and references therein) on a theory of signal recovery from highly incomplete information. The central result relevant to us here, states that a sparse vector $\boldsymbol{w} \in \mathbb{R}^N$ can be recovered from a small number of linear measurements $\boldsymbol{s} = \mathsf{A}\boldsymbol{w} \in \mathbb{R}^K$, $K \ll N$ (or $\boldsymbol{s} = \mathsf{A}\boldsymbol{w} + \eta$ when there is measurement noise) by solving a convex program. Among the several problems

they have discussed, we are particularly interested in the following two problems: (1) Min-L_1 with equality constraints:

$$(P_3) \quad \min \|\boldsymbol{w}\|_1 \quad \text{subject to} \quad \mathsf{A}\boldsymbol{w} = \boldsymbol{s} \tag{11}$$

and (2) Min-L_1 with quadratic constraints:

$$(P_4) \quad \min \|\boldsymbol{w}\|_1 \quad \text{subject to} \quad \|\mathsf{A}\boldsymbol{w} - \boldsymbol{s}\|_2 \leq \epsilon \tag{12}$$

where ϵ is a user specified parameter. Both problems (P_3) and (P_4) find the vector with smallest L_1 norm ($\|\boldsymbol{w}\|_1 = \sum_i |w_i|$) that best explains the observation $\boldsymbol{s}$. (P_3) can be recast as an linear-programming (LP) problem while (P_4) can be recast as a second order cone programming (SOCP) problem (see [22, 23] and references therein for details). We will report the results of implementation of these methods for the sake of comparisons in the next section.

However, (P_3) and (P_4) do not explicitly enforce the nonnegative constraints. The straight forward solution is to incorporate a nonnegative constraint while minimizing the least-squares criterion:

$$(P_5) \quad \min \|\mathsf{A}\boldsymbol{w} - \boldsymbol{s}\|^2 \quad \text{subject to} \quad \boldsymbol{w} \geq 0. \tag{13}$$

This non-negative least squares (NNLS) minimization is precisely a quadratic programming problem: Find the minimum point of a concave quadratic function in a linearly bounded convex feasible hyperspace. The most used algorithm for NNLS was developed in [24, Ch. 23], which treats the linear inequality constraints using an *active set strategy*. Though the sparsity constraint is not explicitly imposed, the active set strategy tends to find the sparse solution quickly if there exists such one. Additionally, unlike other iterative methods mentioned above, this algorithm requires no arbitrary cutoff parameter and hence the output is not susceptible to mis-tuning of the input parameters. More comparisons of these methods on simulation data are shown in Section 4.

After $\boldsymbol{w}$ is estimated, the displacement probabilities can be approximated by the Fourier transform $P(\boldsymbol{r}) = \int (S(\boldsymbol{q})/S_0) \exp(-i\boldsymbol{q} \cdot \boldsymbol{r}) \, \mathrm{d}\boldsymbol{q}$ where $\boldsymbol{r}$ is the displacement vector. Assuming a continuous diffusion tensor model (1) with mixing distribution $F(\mathsf{D}) = \sum_{i=1}^{N} w_i \mathrm{d}\gamma_{p_i,\Sigma_i}$, we have

$$P(\boldsymbol{r}) = \int_{R^3} \int_{\mathcal{P}_n} e^{-\boldsymbol{q}^T \mathsf{D} \boldsymbol{q} t} \mathrm{d}F(\mathsf{D}) e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{q} \approx \sum_{i=1}^{N} \frac{w_i}{\sqrt{(4\pi t)^3 |\hat{\mathsf{D}}_i|}} \exp\left(\frac{-\boldsymbol{r}^T \hat{\mathsf{D}}_i^{-1} \boldsymbol{r}}{4t}\right) \tag{14}$$

where $\hat{\mathsf{D}}_i = p\Sigma_i$ are the expected values of γ_{p,Σ_i}. Note that the end result is expressed as a mixture of oriented Gaussians. Due to its good analytic properties, many of the quantities produced by other methods including the radial integral of $P(\boldsymbol{r})$ in QBI [6] and the integral of $P(\boldsymbol{r})\boldsymbol{r}^2$ in DSI [10] are analytically computable using our technique.

4 Experimental Results

4.1 Numerical Simulations

Prior to performing the experiments on real diffusion MRI data, we first test the performance of the methods described in the previous section on the HARDI simulations of 1-,2- and 3-fiber geometries with known fiber orientations as shown in Fig. 2. The diffusion MR signals were realistically simulated by using the formulas from the cylindrical boundary restricted diffusion model in [25] with the same parameter settings as in [11]. In order to compare the performance of the five deconvolution methods described in Section 3, we first apply all of them on the noiseless 1-fiber HARDI simulation data and the results of $\boldsymbol{w}$ obtained from these methods are plotted in Fig. 3. We observe that the least squares solution to (P_1) contains a large portion of negative weights and has relatively large magnitude. A zeroth-order Tikhonov regularization (P_2) is able to reduce the magnitude significantly but does not help achieve the sparsity and non-negativity. By minimizing the L_1 norm with equality constraints, (P_3) yields relative sparse solution but the magnitude and the negative values are not controlled. The result produced by (P_4) has better sparsity and non-negativity. Evidently, the best result is obtained by solving (P_5) using NNLS. Among the 321 components, there are only two nonzero and significant spikes which both lie in the neighborhood of true fiber orientation $(30^\circ, 90^\circ)$. It is important to note that (1) the true fiber orientations do not necessarily occur at the maximum of the discrete $\boldsymbol{w}$ vector; and (2) although all of these different results for $\boldsymbol{w}$ actually lead to a very good approximation of the true displacement probability function $P(\boldsymbol{r})$ after taking the Fourier transforms, a sparse positive representation of $\boldsymbol{w}$ obviously offers a great advantage in setting the initial guess in the optimization procedure used to find the fiber orientations by estimating the extrema of $P(\boldsymbol{r})$. Considering the additional computational overhead for solving $(P_3), (P_4)$ and (P_5) due to the iterative optimization, (P_5) is slightly slower than (P_3) (LP) but significantly faster than (P_4) (SOCP). It only takes MATLAB's built-in *lsqnonneg* around 20-200ms to solve a problem of size 81×321. Fig. 4 further shows the results of using NNLS on the noisy simulated data. Clearly, NNLS is still able to produce quite accurate solutions which also are sparse representations.

Fig. 2. HARDI simulations of 1-, 2- and 3-fibers ($b = 1500s/mm^2$) visualized in QBI ODF surfaces using [7, Eq.(21)]. Orientation configurations: azimuthal angles: $\phi_1 = 30^\circ, \phi_2 = \{20^\circ, 100^\circ\}, \phi_3 = \{20^\circ, 75^\circ, 135^\circ\}$; polar angles are all 90°.

Finally, as a conclusion to our experiments on the simulated data, we compare the proposed method mixture of Wisharts (MOW) model with two model-free methods, namely, the Q-ball ODF [6] and the DOT [11]. In order to provide a quantitative comparison, all the resulting $P(\boldsymbol{r})$ surfaces were represented by spherical harmonics coefficients up to order $l = 6$. As before, the Q-ball ODF is computed using the formula in [7, Eq.(21)]. First, to gain a global assessment of these methods in terms of stability,

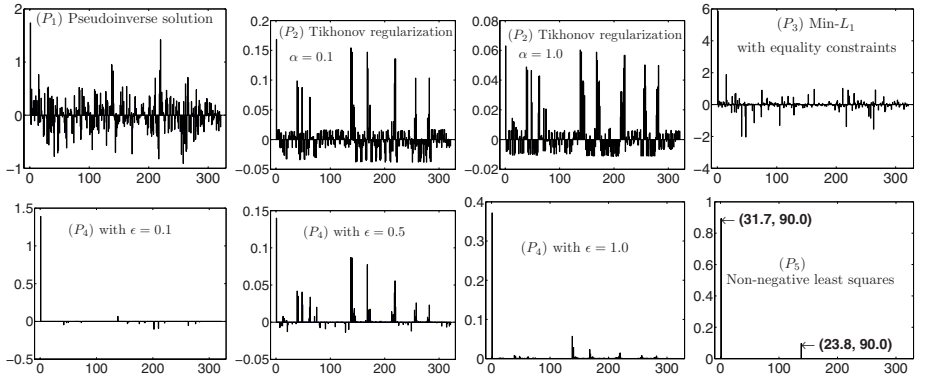

Fig. 3. Deconvolution-based computation of $\boldsymbol{w}$ in the 1-fiber HARDI simulation. The matrix A is of size 81×321 and is built from the Wishart model but with $p = \infty$. The Min-L_1 algorithms are solved using the package developed in [23]. The NNLS solver is the MATLAB built-in *lsqnonneg* function based on Lawson and Hanson's algorithm [24].

we calculated the similarity between each noisy $P(\boldsymbol{r})$ and the corresponding noiseless $P(\boldsymbol{r})$ using the angular correlation coefficient formula given in [7, Eq.(71)]. The angular correlation ranges from 0 to 1 where 1 implies identical probability profiles. Then, we estimated the fiber orientations of each system by numerically finding the maxima of the probability surfaces with a Quasi-Newton algorithm and computed the deviation angles between the estimated and the true fiber orientations. Figure 5 shows the mean and standard deviation of the angular correlation coefficients, and error angles, respectively, for the two-fiber simulation. Note that among the three methods examined, only MOW results in small error angles and high correlation coefficients in presence of relatively large noise. This trend also holds for the 1-fiber and the 3-fiber simulations. This can be explained by noting that NNLS is able to locate the sparse spikes quite accurately even in the presence of large noise.

4.2 Real Data Experiments

The rat optic chiasm provides an excellent "platform" to experimentally validate our approach due to its distinct myelinated structure with both parallel and descussating (crossing) optic nerve fibers. Decussating fibers carry information from the temporal visual fields to the geniculate body. A HARDI data set was acquired from a perfusion-fixed excised rat optic chiasm at 14.1T using a Bruker Avance imaging system with a diffusion-weighted spin echo pulse sequence. DW-MRI data were collected using 46 directions with a b-value of $1250 s/mm^2$ and a single image with $b \approx 0 s/mm^2$. Echo time and repetition time were $23ms$ and $0.5s$ respectively; Δ and δ values were set to $12.4ms$ and $1.2ms$ respectively; bandwidth was set to $35kHz$; signal average was 10; a matrix size of $128 \times 128 \times 5$ and a resolution of $33.6 \times 33.6 \times 200 \mu m^3$ was used. The optic chiasm images were signal averaged to $67.2 \times 67.2 \times 200 \mu m^3$ resolution prior to computation of the water molecule displacement probability field.

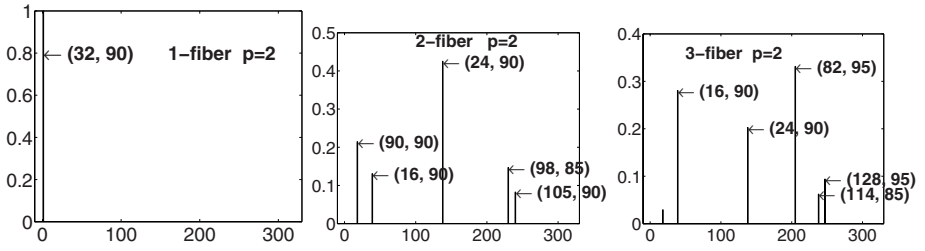

Fig. 4. Deconvolution-based computation of $\boldsymbol{w}$ using NNLS on simulated data in presence of Riccian noise with $\sigma = .06$. The matrix A is constructed by using the Wishart model with $p = 2$ and the tessellation of size $N = 321$. From left to right are the 1-fiber, 2-fiber and 3-fiber simulations, respectively. The spikes in each resulting $\boldsymbol{w}$ are marked with the corresponding azimuthal and polar angles in degrees. Note all the spikes detected are close to the ground truth orientations specified in Figure 2.

Three methods are used to generate the displacement probability functions for the optic chiasm image. The results on a region of interest are shown in Figure 6. The corresponding S_0 image is also shown in the upper left corner as a reference. As seen from the figure, the fiber-crossings in the optic chiasm region is not identifiable in Figure 6 (c). Note that both the DOT method and the MOW method are able to demonstrate the distinct fiber orientations in the central region of the optic chiasm where ipsilateral myelinated axons from the two optic nerves cross and form the contralateral optic tracts. However, it is evident from the figure that compared to all other solutions, the MOW scheme yields significantly sharper displacement probability surfaces. This is particularly borne out in the central location of the figure labeled, the optic chiasm. One of the reasons for the blurred appearance of these probability surfaces in the QBI and DOT models is due to the fact that neither of them yield the actual displacement probability surfaces but a corrupted $P(\boldsymbol{r})$ where the corrupting factor is a a zeroth order Bessel function in the QBI method and a function that does not have an analytic form in the

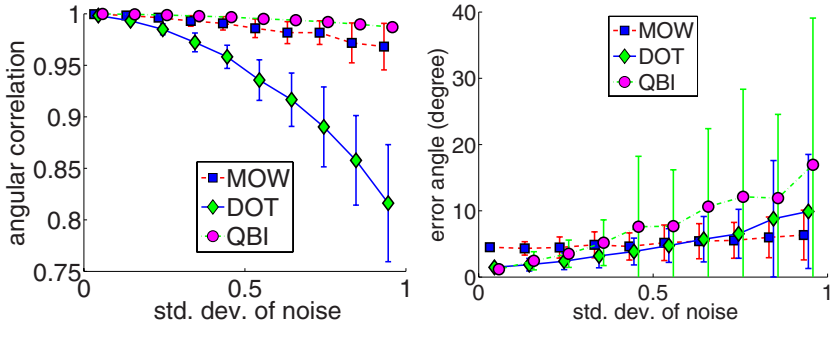

Fig. 5. Mean and standard deviation of (a) angular correlation coefficient and (b) error angles for the two-fiber simulation. The displayed values for error angles are averaged over the two fibers.

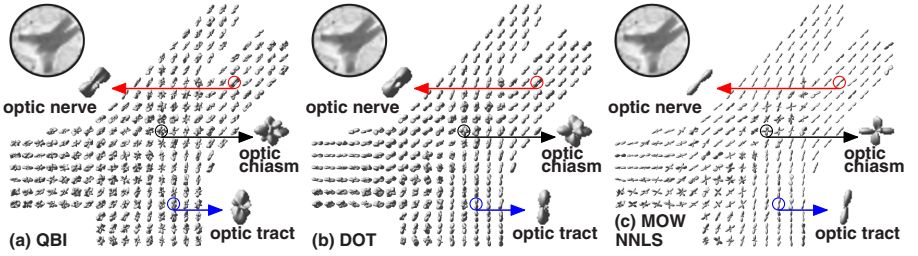

Fig. 6. Probability surfaces computed from a rat optic chiasm image using three methods. Note the decussation of myelinated axons from the two optic nerves at the center of the optic chiasm.

case of DOT. This corruption effects the accuracy of the reconstructed fiber orientations as evidenced in the simulated data case where ground truth was known.

5 Conclusions

In this paper, we present a novel mathematical model which relates the diffusion MR signals and probability distributions for positive definite matrix-valued random variables through Laplace transforms. We further show that the closed form expression for the Laplace transform of Wishart distributions leads to Rigaut-type asymptotic fractal law for the MR signal decay behavior which has been observed experimentally in the past [17]. Moreover in this case, the traditional diffusion tensor model is the limiting case of the expected signal attenuation. We further develop a spherical deconvolution method for resolving multiple fiber orientations using the mixture of Wisharts (MOW) model. To address the numerical issues and sparsity constraints raised in solving this deconvolution problem, we investigate a number of deconvolution techniques and demonstrate that the classic non-negative least squares (NNLS) algorithm developed in [24] is most suitable for our deconvolution problem in achieving sparseness and robustness. Experimental results on both synthetic and real data sets have also shown that the proposed MOW model combined with NNLS deconvolution provides better overall performance than other state-of-the-art techniques for multi-fiber reconstruction.

References

[1] Bihan, D.L.: Looking into the functional architecture of the brain with diffusion MRI. Nat Rev Neurosci. 4, 469–480 (2003)
[2] Basser, P.J., Mattiello, J., Bihan, D.L.: MR diffusion tensor spectroscopy and imaging. Biophys. J. 66, 259–267 (1994)
[3] Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. Magn. Reson. Med. 48, 577–582 (2002)
[4] Frank, L.: Characterization of anisotropy in high angular resolution diffusion weighted MRI. Magn. Reson. Med. 47, 1083–1099 (2002)

[5] Özarslan, E., Mareci, T.H.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. Magn. Reson. Med. 50, 955–965 (2003)
[6] Tuch, D.S.: Q-ball imaging. Magn. Reson. Med. 52, 1358–1372 (2004)
[7] Anderson, A.W.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. Magn. Reson. Med. 54, 1194–1206 (2005)
[8] Hess, C.P., Mukherjee, P., Han, E.T., Xu, D., Vigneron, D.B.: Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magn. Reson. Med. 56, 104–117 (2006)
[9] Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: A fast and robust ODF estimation algorithm in q-ball imaging. In: ISBI 2006, pp. 81–84 (2006)
[10] Wedeen, V.J., Hagmann, P., Tseng, W.Y.I., Reese, T.G., Weisskoff, R.M.: Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn. Reson. Med. 54, 1377–1386 (2005)
[11] 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 31, 1086–1103 (2006)
[12] Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Brady, J., Smith, S.: Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088 (2003)
[13] Ramirez-Manzanares, A., Rivera, M., Vemuri, B.C., Mareci, T.H.: Basis functions for estimating intra-voxel structure in DW-MRI. In: Proc. IEEE Medical Imaging Conference, Roma, Italy 2004, pp. 4207–4211 (2004)
[14] Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage 23, 1176–1185 (2004)
[15] Alexander, D.C.: Maximum entropy spherical deconvolution for diffusion MRI. In: IPMI 2005, pp. 76–87 (2005)
[16] Tournier, J.D., Calamente, F., Connelly, A.: Improved characterisation of crossing fibres: spherical deconvolution combined with Tikhonov regularization. In: ISMRM 2006 (2006)
[17] Köpf, M., Metzler, R., Haferkamp, O., Nonnenmacher, T.F.: NMR studies of anomalous diffusion in biological tissues: Experimental observation of Lévy stable processes. Fractals in Biology and Medicine. 2, 354–364 (1998)
[18] Mathai, A.M.: Jacobians of Matrix Transformations and Functions of Matrix Argument. World Scientific, Singapore (1997)
[19] Rigaut, J.P.: An empirical formulation relating boundary lengths to resolution in specimens showing 'non-ideally fractal dimensions. J Microsc 133, 41–54 (1984)
[20] Sen, P.N., Hürlimann, M.D., de Swiet, T.M.: Debye-Porod law of diffraction for diffusion in porous media. Phys Rev B 51, 601–604 (1995)
[21] Sakaie, K.E., Lowe, M.J.: An objective method for regularization of fiber orientation distribution derived from diffusion-weighed MRI. NeuroImage 34, 169–176 (2007)
[22] Candès, E.J., Romberg, J.K., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Info. Theory 52, 489–509 (2006)
[23] Candés, E., Romberg, J.: l_1-MAGIC (2006) //www.l1-magic.org
[24] Lawson, C., Hanson, R.J.: Solving Least Squares Problems. Prentice-Hall, Englewood Cliffs (1974)
[25] Söderman, O., Jönsson, B.: Restricted diffusion in cylindirical geometry. J. Magn. Reson. B 117, 94–97 (1995)