

Denoising and Visualization of HARDI Data

Tim McGraw, Evren Özarslan, Baba Vemuri, *Fellow, IEEE*, Yunmei Chen, and Thomas Mareci,

Abstract—Despite its apparent success, Diffusion Tensor MRI (DT-MRI) has significant shortcomings when the tissue of interest has a complicated structure. This is due to the relatively simple tensor model that assumes a unidirectional—if not isotropic—local structure. As a more viable alternative Tuch et. al. have proposed to do the data acquisition such that the diffusion sensitizing gradient directions sample the surface of a sphere. In this high angular resolution diffusion imaging (HARDI) method, one does not have to be restricted to the tensor model and instead, it is possible to calculate diffusion coefficients independently along many directions. This imaging technique can reveal white matter fiber crossings which would not be apparent in DT images.

In this paper, we present a novel variational formulation for restoring the HARDI data and visualizing the fibers from this restored data. This formulation involves smoothing signal measurements over the spherical domain and across the 3D image lattice. The smoothing on the spheres at each lattice point is achieved using first and second order smoothness constraints, and across the lattice via a total variation norm based scheme. For the smoothing problem on the sphere, we use the finite element method (FEM). Unlike the reported work on spherical harmonic basis expansion of the diffusivity function on the sphere, the FEM basis functions have local support and are therefore more stable with respect to local perturbations caused by noise in the data. To visualize the fiber paths, the probability values for water molecules to move a particular distance along different orientations were calculated using a Laplace series expansion of these probabilities. We compute the Shannon as well as the Renyi entropies of this distribution to characterize the anisotropy of diffusion, with higher entropy corresponding to lower anisotropy. These local distributions are also used to compute a vector quantity called expected direction. Surfaces rendered using colors corresponding to expected direction reveal anisotropy and fiber direction in the imaged tissue. Further, examples are presented to depict the performance of the HARDI data restoration and visualization schemes on synthetic data, rat brain and spinal cord data sets respectively.

Index Terms—high angular resolution diffusion image, diffusion tensor magnetic resonance imaging

I. INTRODUCTION

Fundamental advances in understanding living biological systems require detailed knowledge of structural and functional organization. This is particularly important in the nervous system where anatomical connections determine the information pathways and how this information is processed. In addition, understanding fundamental structural relationships is essential to the development and application of therapies to treat pathological conditions (e.g. disease or injury).

Observing the directional dependence of water diffusion can allow us to infer structural information about the surrounding tissue. White matter fiber bundles present a barrier to diffusion, causing relatively high diffusivity along the fiber direction, and lower diffusivity across the fiber.

Most imaging methods give only an anatomically isolated representation of living tissue because the images do not contain connectivity information. Such information would allow the identification and correlation of system elements responding during function. For example in brain trauma, the relationship between anatomy and behavior will only become apparent when we are able to resolve the afferent nerve fiber pathways transmitting the sensation from a stimulus to the brain or the efferent pathways transmitting impulses from the brain area controlling behavior.

MR measurements can be made sensitive to the translational diffusion of water molecules by the utilization of magnetic field gradients [1]. In general, the signal acquired depends on the strength and the direction of these diffusion sensitizing gradients. Repeated measurements of water diffusion in tissue with varying gradient directions provide a means to quantify the level of anisotropy as well as to determine the local fiber orientation within the tissue. In a series of publications, Basser and colleagues [2], [3], [4] have formulated a new imaging modality called “diffusion tensor MRI (DT-MRI)” that employs a second order, positive definite, symmetric diffusion tensor to represent the local tissue structure. They have proposed several rotationally invariant scalar indices that quantify different aspects of water diffusion observed in tissue, similar to different “stains” used in histological studies [5]. Under the hypothesis that the direction along which the diffusion coefficient is largest will yield the local fiber orientation, one can determine the directionality of neuronal fiber bundles. This fact has been exploited to generate fiber-tract maps that yield information on structural connections in human [4], [6], [7], [8] as well as rat brains [9], [10], [11], [12], [13] and spinal cords [14].

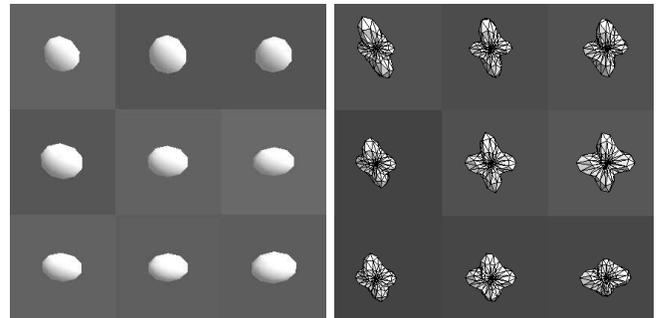


Fig. 1. Orientational heterogeneity in DT-MRI (left), and HARDI (right).

Despite its apparent success, DT-MRI has significant shortcomings when the tissue of interest has a complicated geometry. This is due to the relatively simple tensor model that assumes a unidirectional—if not isotropic—local structure. In the case of orientational heterogeneity, DT-MRI technique is

likely to yield incorrect fiber directions, and artificially low anisotropy values. This is due to violation of the assumption of Gaussian probability model characterizing the diffusion implicit in DTI. In order to overcome these difficulties several approaches have been taken. Q-space imaging, a technique commonly used to examine porous structures [15], has been suggested as a possible solution [16]. However this scheme requires strong gradient strengths and long acquisition times [17], or significant reduction in the resolution of the images. As a more viable alternative Tuch et. al. have proposed to do the acquisition such that the diffusion sensitizing gradients sample the surface of a sphere [18], [19]. In this high angular resolution diffusion imaging (HARDI) method, one does not have to be restricted to the tensor model and instead, it is possible to calculate diffusion coefficients independently along many directions. This method *does not* require more powerful hardware systems than that required by DT-MRI. Several groups have already performed HARDI acquisitions in clinical settings and have reported 43 to 126 different diffusion weighted images acquired in 20 to 40 minutes of total scanning time [20], [19], [21] indicating the feasibility of the high angular resolution scheme as a clinical diagnostic tool. In Figure 1, we present a matrix of voxels from a rat brain data set showing, renderings of DTI-based estimates of orientation and HARDI-based orientation estimates. The orientation heterogeneity is evident from the HARDI-based renderings at each voxel since HARDI measurements can resolve multiple dominant directions of anisotropy at a voxel, a feature which is lacking in the DTI. Since the HARDI data acquisition is very nascent, not many techniques of processing the HARDI data have been reported in literature. In the following, we will briefly review the few very recently reported techniques of HARDI data denoising, which must be done prior to further analysis or visualization.

A. Restoration

Processing of HARDI data sets has received increased attention lately and a few researchers have reported their results in literature. The use of spherical harmonic expansions have been quite popular in this context since the HARDI data primarily consists of scalar signal measurements on a sphere located at each lattice point on a 3D image grid. Tuch [18], [19] developed the HARDI acquisition and processing and later Frank [20] showed that it is possible to use the spherical harmonics expansion of the HARDI data to characterize the local geometry of the diffusivity profiles. In his work however, there is no discussion of denoising/restoring the HARDI signal measurements, which is essential for subsequent processing and interpretation. Chen et al. [22] find a regularized spherical harmonic expansion by solving a constrained minimization problem. However the expansion is a truncated spherical harmonic expansion of order four, and hence the solution can at best only represent two fiber directions within a voxel. In [21], Jansons and Alexander described a new statistic which was called persistent angular structure that was obtained from the samples of a 3D function, in this case the displacement

of water molecules in each direction. The goal in their work was to resolve voxels containing one or more fibers. However, there was no discussion on how to restore the noisy HARDI data prior to resolution of the fiber paths.

In contrast to HARDI denoising, DT-MRI denoising has been more popular and numerous techniques exist in literature. Most of the techniques in literature have employed vector-valued (See [23]) smoothing to tensor field data by using spectral decomposition of the tensors and smoothing the eigen vectors and eigen values using techniques introduced in [24]. Recent work by Tschumperlé et al., [25] deals with smoothing the eigenvector field of the diffusion tensors computed from the raw echo intensity image data. Several other methods have been developed for restoring the DT-MRI data sets but most of them use existing restoration schemes for scalar or vector-valued functions from image processing literature. More recently, matrix valued function restoration was introduced in [26] and applied to the restoration of noisy diffusion tensor fields. An interesting alternative to the variational principle approach was taken by Weickert and Brox [27] wherein, they developed an anisotropic diffusion filter that achieves an “edge” preserving smoothing of the positive semi-definite tensor-valued image. All of the matrix-valued smoothing methods share one commonality i.e., they use a linearized Stejskal-Tanner equation as the data acquisition model which does not accurately reflect the physics of the data acquisition. In recently reported work, by Wang et al., [28], [29], an alternative approach which overcomes this weakness by directly estimating a smooth positive semidefinite tensor field from the raw data using the actual nonlinear mono-exponential model characterized by the Stejskal-Tanner equation [1] was reported.

B. Tractography

Water in the brain preferentially diffuses along white matter fibers. By tracking the direction of fastest diffusion, as measured by MRI, non-invasive fiber tracking of the brain can be accomplished. In the context of DTI data, fiber tracks estimated in reported literature were obtained by repeatedly stepping in the direction of fastest diffusion. The direction along which the diffusion is dominant corresponds to the direction of eigenvector corresponding to the largest eigenvalue of the tensor \mathbf{D} . This approach was taken by many researchers [30], [9], [31], [32], [23], [33], [14], [13]. Most of these techniques do incorporate some regularization in their stream line estimation schemes in order to generate the fiber pathways. Techniques that are quite distinct from the idea of stream line generation have also been reported in literature. Batchelor et. al., [34] reported a fiber tract mapping scheme wherein, they produce a map indicating the probability of a fiber passing through each location in the field. However, no discussion on how to estimate the actual fibers was described. An alternative approach based on sequential importance sampling and regularization techniques was proposed in [35], which allowed paths to originate from a single location and branch out and produced a probability distribution of the paths. O’Donnell et. al., [36] describe ways to estimate the

connectivity from the given tensor field. One approach they suggested was to estimate the geodesics in the locally warped space where the warping is derived from the local tensor. Behrens et al. [37] estimate the local probability distribution of fiber directions at each voxel and proceed to compute global connectivity from these local PDFs. Unlike previous work, the model parameters are described using distributions, and the PDF at each voxel are estimated using Markov Chain Monte Carlo techniques.

In the context of HARDI processing, *suggestions* for fiber path computation have been reported in [20], [21], [38]. Frank [20] described a spherical harmonic transform representation of the HARDI data and pointed out that, due to the antipodal symmetry on the sphere, only the even terms in the spherical harmonic transform contributed toward the representation. Any non-zero odd terms would be due to artifacts in imaging such as noise. The fiber pathways were not computed in his reported work, however it was suggested that one could estimate the direction of the crossing fibers by using a multi-tensor model and finding the principal diffusion directions of this multitensor expansion. The approach of using HARDI data and then reverting to a multi-tensor approach seems somewhat awkward. The approach described by Özarlan et al. [38] expresses the diffusivity function, a function defined on the sphere, as a generalized (higher rank) Cartesian tensor and then estimates the probability distribution of water molecule displacement over all directions using the FFT (fast Fourier transform) of the signal measurements either on the sphere or an interpolated Cartesian grid. These distribution profiles are then sharpened to display the possible orientation of the fibers with complex local geometries. *In this paper, we will not be computing the fiber tracts but will present ways of visualizing the probability maps of water diffusion via color displays of measures of anisotropy.*

C. Anisotropy

In DTI there are many scalar measures which characterize the anisotropy of the diffusion phenomenon. Most useful are the rotationally invariant measures, those independent of the laboratory reference frame, such as volume ratio, rational anisotropy and fractional anisotropy [5]. These can be computed in terms of the eigenvalues of the diffusion tensor. These scalar indices alone have been useful in clinical studies [39], so it is important to find analogous measures of anisotropy for HARDI. The measures for HARDI should overcome the weaknesses of the tensor model in regards to regions of crossing fibers. In the case of orientational heterogeneity, DT-MRI technique yields incorrect fiber directions, and artificially low anisotropy values [40].

Frank [20] used the coefficients of the spherical harmonic expansion of the apparent diffusion coefficient to quantify anisotropy. The 0th, 2nd and 4th order coefficients describe isotropic diffusion, single-fiber diffusion and two-fiber diffusion respectively. Voxels can be classified using the relative magnitudes of these coefficients. In [38] Özarlan et al. use

a high rank tensor to describe the diffusion process and generalize the rank-2 tensor trace to higher rank tensors. The variance of the diffusivity is then expressed in terms of the generalized trace, and a transfer function maps this value to the $[0, 1]$ range. This quantity is called "generalized anisotropy" (GA). It was shown to be useful in detecting anisotropy at fiber crossings which are not detectable using FA.

D. Overview of Our Modeling Scheme

Our goal here is to denoise the HARDI data, develop a fast computation scheme for estimating the molecular displacement distribution, and compute a novel index of anisotropy from these distributions along with a visualization scheme for the same. We present a novel and effective variational formulation that will directly estimate a smooth signal $S(\theta, \phi)$ and the probability distribution of the water molecule displacement over all directions $p(\theta, \phi)$, given the noisy measurement

$$\hat{S}(\theta, \phi) = S_0 \exp(-bd(\theta, \phi)) + \eta(\theta, \phi) \quad (1)$$

where \hat{S} is the signal measurement taken on a sphere of constant gradient magnitude over all (θ, ϕ) , b is the diffusion weighting factor, $d(\theta, \phi)$ is the diffusivity as a function of the direction expressed by the elevation and azimuth angles on the sphere and $\eta(\theta, \phi)$ is the Gaussian noise. The variational principle involves smoothing S values over the sphere and across the 3D image lattice. The key factor that complicates this problem is that the domain of the data at each voxel in the voxel lattice is a sphere. One may use the level-set techniques developed by Tang et al., [41] to achieve this smoothing however, when data sets are large, it becomes computationally impractical to apply the level-set technique at each voxel independently to restore these scalar-valued measurements on the sphere. We arrive at a computationally efficient solution to this problem by using the finite element method (FEM) on the sphere and choosing local basis functions for the data restoration. Unlike the reported work on spherical harmonic basis expansion of the diffusivity function on the sphere [42], [38], [22], the FEM basis functions have local support and are more stable to perturbations due to noise in the data.

From the denoised data we will compute a probability, $p_i(\theta, \phi)$, of molecular diffusion over a sphere of directions. The Shannon entropy as well as Renyi entropy of this distribution will then be used to quantify anisotropy.

II. METHODS

The HARDI processing proceeds by acquiring diffusion weighted images with many diffusion encoding gradient directions, effectively sampling a spherical shell of the q-space. It is desired that this sampling be uniform, or nearly so. The gradient direction for each image is usually chosen to correspond to the vertices of an icosahedron which has been repeatedly subdivided. Since the process of diffusion is known to be symmetric, we need only sample one hemisphere of the q-space. In the case of our data, we consider 81 or 46

gradient directions. The gradient directions correspond to the vertices of the surface shown in Figure (2). In addition a low b-value (small diffusion encoding gradient magnitude) image is acquired in order to estimate S_0 .

The random process of diffusion of water molecules is described by the diffusion displacement PDF $p_t(\mathbf{r})$. This is the probability that a given molecule has a diffusion displacement of \mathbf{r} after time t . The relation between the measured image, and the diffusion displacement PDF is given by Callaghan in [15] as

$$p_t(\mathbf{r}) = \int \frac{S(\mathbf{q})}{S_0} \exp(-2\pi i \mathbf{q} \cdot \mathbf{r}) d\mathbf{q} \quad (2)$$

where S_0 is the image acquired with no diffusion encoding gradient applied. This is simply the Fourier transform of $\frac{S(\mathbf{q})}{S_0}$. It is the modes of $p_t(\mathbf{r})$ that are taken to be the underlying fiber directions.

III. VARIATIONAL PRINCIPLE FORMULATION

We propose a membrane-spline deformation energy minimization for smoothing the measured image $\hat{\mathbf{S}}(\mathbf{x}, \theta, \phi)$. The variational principle for estimating a smooth $\mathbf{S}(\mathbf{x}, \theta, \phi)$ is given by

$$\begin{aligned} \min_{\mathbf{S}} E(\mathbf{S}) = & \mu \int_{\Omega} \int_{S^2} |\mathbf{S}(\mathbf{x}, \theta, \phi) - \hat{\mathbf{S}}(\mathbf{x}, \theta, \phi)|^2 dS d\mathbf{x} \\ & + \int_{S^2} \|\nabla_{(\theta, \phi)} \mathbf{S}\|^2 dS + \int_{\Omega} g(\mathbf{x}) \|\nabla_x \mathbf{S}\| d\mathbf{x} \end{aligned} \quad (3)$$

where Ω is the domain of the image lattice and S^2 is the sphere on which the signal measurements are specified at each voxel. The first term of Equation (3) is a data fidelity term which makes the solution to be close to the given data. The degree of data fidelity can be controlled by the input parameter μ . The second term is a regularization constraint enforcing smoothness of the data over the spherical domain at each voxel. The third term is another regularization term which causes the solution to be smooth over the spatial domain (the 3D voxel lattice).

A. Finite Element smoothing of $S(\theta, \phi)$

We will consider a deformation energy functional which is a weighted combination of thin-plate spline energy and membrane spline energy, that is commonly used in computer vision literature for smoothing scalar-valued data in \mathfrak{R}^3 (see McInerney and Terzopoulos [43]).

The diffusion-encoding gradient directions are taken as the vertices of a subdivided icosahedron, to achieve a nearly uniform sampling of spherical direction. We map this piecewise linear approximation of a sphere to the plane by taking the spherical coordinates (θ, ϕ) of the imaging gradient direction as the planar global coordinates (u, v) . The gradient directions, and their embedding in the plane are shown in Figures (2) and (3).

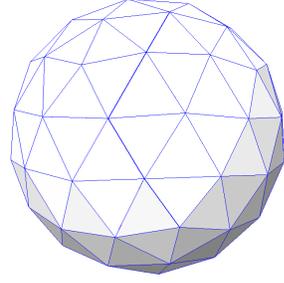


Fig. 2. HARD gradient directions (θ, ϕ) correspond to the vertices of a subdivided icosahedron.

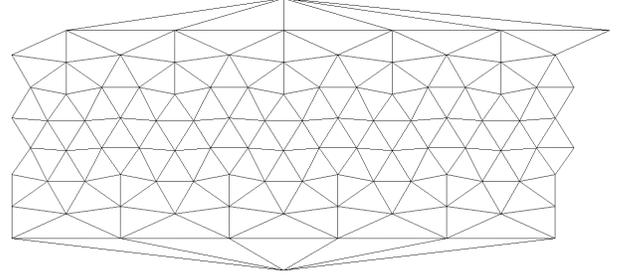


Fig. 3. 2 dimensional (u, v) domain for FEM-based smoothing of HARDI data.

Note that, the data can be seen as height values defined on the sphere at the coordinates specified parametrically by (u, v) . The smoothing will be applied to these height measurements $z(u, v)$ using the following smoothing functional.

$$\mathcal{E}_p = \int \int_{\Omega} (\alpha(|z_u|^2 + |z_v|^2) + \beta(|z_{uu}|^2 + 2|z_{uv}|^2 + |z_{vv}|^2)) dudv \quad (4)$$

The weight on the membrane term is α and the weight on the thin-plate term is β . Once we have computed a smooth $z(u, v)$, the result will then be mapped back to the image on the sphere, $S(\theta, \phi)$.

The data energy due to virtual work of the data forces is

$$\mathcal{E}_d = - \int \int_{\Omega} z(u, v) f(u, v) dudv \quad (5)$$

The restoration at each voxel is formulated at the energy minimization

$$\min_{\mathbf{S}} \mathcal{E}(\mathbf{S}) = \mathcal{E}_p(\mathbf{S}) + \mathcal{E}_d(\mathbf{S}) \quad (6)$$

We use polynomial shape functions, N_i as a basis for the data over the domain of spherical directions.

$$z(u, v) = \sum_{i=1}^n q_i N_i(u, v) = \mathbf{N} \mathbf{q} \quad (7)$$

Where the \mathbf{N} is a $(1 \times n)$ row vector, and \mathbf{q} is a column vector of nodal variables.

1) *Element Matrices:* The domain is subdivided into elements, each with their own local shape functions. For each element j , we have,

$$z(u, v) = \mathbf{N}^j(u, v) \mathbf{q}^j \quad (8)$$

for $(u, v) \in \Omega_j$. The local energy function for each element is given by

$$\mathcal{E}_p^j = \int \int_{\Omega_j} (\alpha |z_u^j|^2 + \alpha |z_v^j|^2 + \beta |z_{uu}^j|^2 + 2\beta |z_{uv}^j|^2 + \beta |z_{vv}^j|^2) dudv \quad (9)$$

Then, the global energy is the sum of the energies of each element,

$$\mathcal{E}_p = \sum_j \mathcal{E}_p^j \quad (10)$$

The element strain vector (given by Dhatt and Touzot [44]) is,

$$\boldsymbol{\varepsilon}^j = \begin{bmatrix} z_u^j \\ z_v^j \\ z_{uu}^j \\ z_{uv}^j \\ z_{vv}^j \end{bmatrix} \quad (11)$$

$$\boldsymbol{\varepsilon}^j = \begin{bmatrix} (N_1)_u & \dots & (N_n)_u \\ (N_1)_v & \dots & (N_n)_v \\ (N_1)_{uu} & \dots & (N_n)_{uu} \\ (N_1)_{uv} & \dots & (N_n)_{uv} \\ (N_1)_{vv} & \dots & (N_n)_{vv} \end{bmatrix} \mathbf{q}^j = \mathbf{B} \mathbf{q}^j \quad (12)$$

where we have defined \mathbf{B} as the $(n \times 5)$ matrix of derivatives of the nodal basis functions. We can then write the element strain energy as:

$$\mathcal{E}_p^j = \int \int_{\Omega_j} \boldsymbol{\varepsilon}^{jT} \mathbf{D} \boldsymbol{\varepsilon}^j dudv \quad (13)$$

where we define

$$\mathbf{D} = \begin{bmatrix} \alpha & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 2\beta & 0 \\ 0 & 0 & 0 & 0 & \beta \end{bmatrix} \quad (14)$$

Since \mathbf{q}^j is constant over each element we can derive the element stiffness matrix in terms of \mathbf{D} and \mathbf{B} as follows:

$$\mathcal{E}_p^j = \int \int_{\Omega_j} \mathbf{q}^{jT} \mathbf{B}^{jT} \mathbf{D} \mathbf{B}^j \mathbf{q}^j dudv = \mathbf{q}^{jT} \mathbf{K}^j \mathbf{q}^j. \quad (15)$$

We will model the data constraint as springs pulling $z(u, v)$ toward some given value $z_0(u, v)$. The force at each point will obey $f = k(z - z_0)$, where k is the spring constant.

$$\mathcal{E}_d^j = - \int \int_{\Omega_j} \mathbf{N}^j \mathbf{q}^j k (\mathbf{N}^j \mathbf{q}^j - z_0) dudv \quad (16)$$

$$\mathcal{E}_d^j = -k \mathbf{q}^{jT} \int \int_{\Omega_j} \mathbf{N}^{jT} \mathbf{N}^j \mathbf{q}^j dudv + k \mathbf{q}^{jT} \int \int_{\Omega_j} \mathbf{N}^{jT} z_0 dudv \quad (17)$$

\mathbf{F}_k^j and \mathbf{f}^j are defined such that the first term of Equation (17) is $\mathbf{q}^{jT} \mathbf{F}_k^j \mathbf{q}^j$ and the second term is $\mathbf{q}^{jT} \mathbf{f}^j$. We can then balance deformation energy and data energy by solving the following linear system:

$$(\mathbf{K}^j + \mathbf{F}_k^j) \mathbf{q}^j = \mathbf{f}^j. \quad (18)$$

TABLE I
GAUSS-RADAU WEIGHTS

i	r_i	w_i	ξ_i	a_i
1	0.0469100770	0.1184634425	0.0398098571	0.1007941926
2	0.2307653449	0.2393143353	0.1980134179	0.2084506672
3	0.5	0.2844444444	0.4379748102	0.2604633916
4	0.7692346551	0.2393143353	0.6954642734	0.2426935942
5	0.9530899230	0.1184634425	0.9014649142	0.1598203766

2) *Local Element Coordinates:* We now present the coordinate system for the local elements. For local elements, triangular elements are used with a barycentric coordinate system (γ, ξ, η) so that each coordinate is in $[0, 1]$ and $\gamma = 1 - \xi - \eta$.

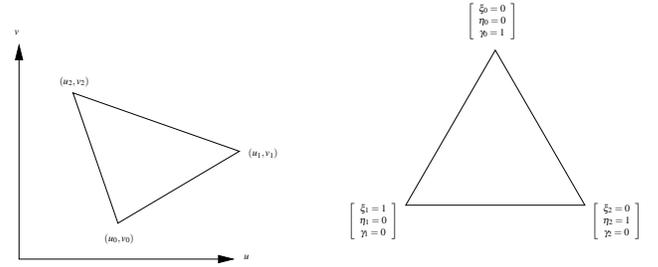


Fig. 4. Mapping to barycentric coordinates

$$\begin{aligned} u(\xi, \eta) &= (1 - \xi - \eta)u_0 + \xi u_1 + \eta u_2 \\ v(\xi, \eta) &= (1 - \xi - \eta)v_0 + \xi v_1 + \eta v_2 \end{aligned} \quad (19)$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_1 - u_0 & u_2 - u_0 \\ v_1 - v_0 & v_2 - v_0 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (20)$$

$$\begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial \xi} & \frac{\partial u}{\partial \eta} \\ \frac{\partial v}{\partial \xi} & \frac{\partial v}{\partial \eta} \end{bmatrix} \begin{bmatrix} d\xi \\ d\eta \end{bmatrix} = \mathbf{J} \begin{bmatrix} d\xi \\ d\eta \end{bmatrix} \quad (21)$$

Integrals over the (u, v) domain can be converted to integrals over the local (ξ, η) domain in the following way:

$$\int \int_{\Omega_j} f(u, v) dudv = \int \int_{\Omega_j} f(u(\xi, \eta), v(\xi, \eta)) \det(\mathbf{J}) d\xi d\eta \quad (22)$$

Using Gauss-Radau quadrature rules, we can approximate the integral in Equation (22) by

$$\sum_{i=1}^5 \sum_{j=1}^5 w_i w_j f(u(\xi_j, \eta_{i,j}), v(\xi_j, \eta_{i,j})) \det(\mathbf{J}) \quad (23)$$

where $\eta_{i,j} = r_i(1 - s_j)$, $w_{j,j} = a_j(1 - \xi_j)$, ξ_j , and w_i are given in Table I. Derivatives over (u, v) become

$$\begin{aligned} \frac{\partial N}{\partial u} &= \frac{\partial N}{\partial \xi} \frac{\partial \xi}{\partial u} + \frac{\partial N}{\partial \eta} \frac{\partial \eta}{\partial u} \\ \frac{\partial N}{\partial v} &= \frac{\partial N}{\partial \xi} \frac{\partial \xi}{\partial v} + \frac{\partial N}{\partial \eta} \frac{\partial \eta}{\partial v} \end{aligned} \quad (24)$$

The partial derivatives of ξ and η with respect to u and v can be computed by inverting the Jacobian

$$\begin{bmatrix} d\xi \\ d\eta \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi}{\partial u} & \frac{\partial \xi}{\partial v} \\ \frac{\partial \eta}{\partial u} & \frac{\partial \eta}{\partial v} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} = \mathbf{J}^{-1} \begin{bmatrix} du \\ dv \end{bmatrix} \quad (25)$$

$$\mathbf{J}^{-1} = \frac{1}{\det(\mathbf{J})} \begin{bmatrix} v_2 - v_0 & -(u_2 - u_0) \\ -(v_1 - v_0) & u_1 - u_0 \end{bmatrix} \quad (26)$$

We use the fifth order element shape functions given by Dhatt and Touzot [44]. *This element guarantees C^1 (surface normal) continuity across triangles.* The basis functions are given by:

$$\begin{aligned} N_1 &= \lambda^2(10\lambda - 15\lambda^2 + 6\lambda^3 + 30\xi\eta(\xi + \eta)) \\ N_2 &= \xi\lambda^2(3 - 2\lambda - 3\xi^2 + 6\xi\eta) \\ N_3 &= \eta\lambda^2(3 - 2\lambda - 3\eta^2 + 6\xi\eta) \\ N_4 &= \frac{1}{2}\xi^2\lambda^2(1 - \xi + 2\eta) \\ N_5 &= \xi\eta\lambda^2 \\ N_6 &= \frac{1}{2}\eta^2\lambda^2(1 + 2\xi - \eta) \\ N_7 &= \xi^2(10\xi - 15\xi^2 + 6\xi^3 + 15\eta^2\lambda) \\ N_8 &= \frac{1}{2}\xi^2(-8\xi + 14\xi^2 - 6\xi^3 - 15\eta^2\lambda) \\ N_9 &= \frac{1}{2}\xi^2\eta(6 - 4\xi - 3\eta - 3\eta^2 + 3\xi\eta) \\ N_{10} &= \frac{1}{4}\xi^2(2\xi(1 - \xi)^2 + 5\eta^2\lambda) \\ N_{11} &= \frac{1}{2}\xi^2\eta(-2 + 2\xi + \eta + \eta^2 - \xi\eta) \\ N_{12} &= \frac{1}{4}\xi^2\eta^2\lambda + \frac{1}{2}\xi^3\eta^2 \\ N_{13} &= \eta^2(10\eta - 15\eta^2 + 6\eta^3 + 15\xi^2\lambda) \\ N_{14} &= \frac{1}{2}\xi\eta^2(6 - 3\xi - 4\eta - 3\xi^2 + 3\xi\eta) \\ N_{15} &= \frac{1}{2}\eta^2(-8\eta + 14\eta^2 - 6\eta^3 - 15\xi^2\lambda) \\ N_{16} &= \frac{1}{4}\xi^2\eta^2\lambda + \frac{1}{2}\xi^2\eta^2 \\ N_{17} &= \frac{1}{2}\xi\eta^2(-2 + \xi + 2\eta + \xi^2 - \xi\eta) \\ N_{18} &= \frac{1}{4}\eta^2(2\eta(1 - \eta)^2 + 5\xi^2\lambda) \end{aligned} \quad (27)$$

The quintic shape functions have nodal variables which can be written in terms of local or global coordinates as,

$$\mathbf{q}_{\xi,\eta} = \begin{bmatrix} z \\ z\xi \\ z\eta \\ z\xi\xi \\ z\xi\eta \\ z\eta\eta \end{bmatrix}, \mathbf{q}_{u,v} = \begin{bmatrix} z \\ z_u \\ z_v \\ z_{uu} \\ z_{uv} \\ z_{vv} \end{bmatrix} \quad (28)$$

which we can relate to each other by

$$\mathbf{q}_{u,v} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi_u & \eta_u & 0 & 0 & 0 \\ 0 & \xi_v & \eta_v & 0 & 0 & 0 \\ 0 & 0 & 0 & \xi_u^2 & 2\xi_u\eta_u & \eta_u^2 \\ 0 & 0 & 0 & \xi_u\xi_v & (\xi_u\eta_v + \eta_u\xi_v) & \eta_u\eta_v \\ 0 & 0 & 0 & \xi_v^2 & 2\xi_v\eta_v & \eta_v^2 \end{bmatrix} \mathbf{q}_{\xi,\eta} \quad (29)$$

3) *Global Matrices:* We now wish to construct global matrices so that the energy balance over the entire FEM mesh is given by the linear system

$$\mathbf{K}\mathbf{q} = \mathbf{f} \quad (30)$$

where \mathbf{K} is a $(6n \times 6n)$ matrix since we have 6 variables per node.

We will consider the simple case of 2 elements. Expanding the element Equation(18) in terms of nodal variables for element 0, we get

$$\begin{bmatrix} \mathbf{K}_{0,0}^0 & \mathbf{K}_{0,1}^0 & \mathbf{K}_{0,2}^0 \\ \mathbf{K}_{1,0}^0 & \mathbf{K}_{1,1}^0 & \mathbf{K}_{1,2}^0 \\ \mathbf{K}_{2,0}^0 & \mathbf{K}_{2,1}^0 & \mathbf{K}_{2,2}^0 \end{bmatrix} \begin{bmatrix} \mathbf{q}_0^0 \\ \mathbf{q}_1^0 \\ \mathbf{q}_2^0 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_0^0 \\ \mathbf{f}_1^0 \\ \mathbf{f}_2^0 \end{bmatrix} \quad (31)$$

and likewise for element 1

$$\begin{bmatrix} \mathbf{K}_{3,3}^1 & \mathbf{K}_{3,2}^1 & \mathbf{K}_{3,1}^1 \\ \mathbf{K}_{2,3}^1 & \mathbf{K}_{2,2}^1 & \mathbf{K}_{2,1}^1 \\ \mathbf{K}_{1,3}^1 & \mathbf{K}_{1,2}^1 & \mathbf{K}_{1,1}^1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_3^1 \\ \mathbf{q}_2^1 \\ \mathbf{q}_1^1 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_3^1 \\ \mathbf{f}_2^1 \\ \mathbf{f}_1^1 \end{bmatrix} \quad (32)$$

where each \mathbf{q}_i^j is a (6×1) column vector of nodal variables. We expand each \mathbf{K}^j to be $(6n \times 6n)$ by inserting rows and columns of zeros corresponding to each node of the mesh. Also expand \mathbf{f}^j to $(6n \times 1)$. The global \mathbf{K} and \mathbf{q} are obtained by summing the expanded matrices from each element in the mesh. For our 2 element example we have

$$\begin{bmatrix} \mathbf{K}_{0,0}^0 & \mathbf{K}_{0,1}^0 & \mathbf{K}_{0,2}^0 & 0 \\ \mathbf{K}_{1,0}^0 & \mathbf{K}_{1,1}^0 + \mathbf{K}_{1,1}^1 & \mathbf{K}_{1,2}^0 + \mathbf{K}_{1,2}^1 & \mathbf{K}_{1,3}^1 \\ \mathbf{K}_{2,0}^0 & \mathbf{K}_{2,1}^0 + \mathbf{K}_{2,1}^1 & \mathbf{K}_{2,2}^0 + \mathbf{K}_{2,2}^1 & \mathbf{K}_{2,3}^1 \\ 0 & \mathbf{K}_{3,1}^1 & \mathbf{K}_{3,2}^1 & \mathbf{K}_{3,3}^1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_0^0 \\ \mathbf{q}_1^0 \\ \mathbf{q}_2^0 \\ \mathbf{q}_3^1 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_0^0 \\ \mathbf{f}_1^0 + \mathbf{f}_1^1 \\ \mathbf{f}_2^0 + \mathbf{f}_2^1 \\ \mathbf{f}_3^1 \end{bmatrix} \quad (33)$$

The global linear system Equation (30) is symmetric, and has a sparse banded structure with 18 nonzero diagonal bands. An efficient solution to \mathbf{q} is obtained via Cholesky factorization.

(27) B. Spatial smoothing of $S(x)$

Smoothing the raw vector-valued data, $S(x)$, is posed as a variational principle involving a first order smoothness constraint on the solution to the smoothing problem. Note that the data at each voxel is a large set of S measurements over a sphere of directions and can be assembled into a vector after the smoothing on the spherical coordinate domain has been accomplished. We propose a weighted TV-norm minimization for smoothing this vector-valued image \mathbf{S} . The variational principle for estimating a smooth $\mathbf{S}(\mathbf{x})$ is given by

$$\min_{\mathbf{S}} \mathcal{E}(\mathbf{S}) = \int_{\Omega} (g(\mathbf{x}) \sum_{i=1}^m \|\nabla S_i\| + \frac{\mu}{2} \sum_{i=1}^m (S_i - \hat{S}_i)^2) d\mathbf{x} \quad (34)$$

where, Ω is the image domain and μ is a regularization factor. The first term here is the regularization constraint on the solution to have a certain degree of smoothness. The second term in the variational principle makes the solution faithful to the data to a certain degree. We have used the coupling function $g(\mathbf{x}) = 1/(1 + GA(\mathbf{x}))$ for smoothing HARDI, where GA is the generalized anisotropy defined in Özarslan et al., [40] and is computed from the variance of normalized diffusivity. For a more detailed discussion on GA , we refer the reader to [40]. This selection criterion preserves locations of dominant GA while smoothing the rest of the data. Note that since we are interested in regions of dominant anisotropy, it is apt to choose such a selective term.

Here we have used a different TV-norm than the one used by Blomgren and Chan [45]. The $TV_{n,m}$ norm is an L_2 norm of the vector of $TV_{n,1}$ norms ($\int_{\Omega} \|\nabla S_i(x)\|^2 dx$) for each channel. We use the L_1 norm instead, which is known to have better discontinuity preservation properties.

The gradient descent form of the above minimization is given by

$$\begin{aligned} \frac{\partial S_i}{\partial t} &= \text{div} \left(\frac{g \nabla S_i}{\|\nabla S_i\|} \right) - \mu(S_i - \hat{S}_i) \quad i = 1, \dots, m \\ \frac{\partial S_i}{\partial n} \Big|_{\partial \Omega \times \mathbb{R}^+} &= 0 \quad \text{and} \quad \mathbf{S}(\mathbf{x}, t = 0) = \hat{\mathbf{S}}(\mathbf{x}) \end{aligned} \quad (35)$$

The use of a modified TV-norm in equation (34) results in a looser coupling between channels than when using the $TV_{n,m}$ norm. This reduces the numerical complexity of Equation (35) and makes solution for large data sets feasible.

The gradient descent of the vector-valued image smoothing using the $TV_{n,m}$ -norm $TV_{n,m}(\mathbf{S}(\mathbf{x})) = \sqrt{\sum_{i=1}^m [TV_{n,1}(S_i)]^2}$ presented in [45] is given by,

$$\begin{aligned} \frac{\partial S_i(x, t)}{\partial t} &= \frac{TV_{n,1}(S_i)}{TV_{n,m}(\mathbf{S})} \nabla \cdot \left(\frac{\nabla S_i}{\|\nabla S_i\|} \right) \\ \mathbf{S}(x, 0) &= \mathbf{S}_0(x). \end{aligned} \quad (36)$$

Note that the $TV_{n,m}$ norm appears in the gradient descent solution of the vector-valued minimization problem. Considering that our data sets consist of up to 82 images, corresponding to (magnetic field) gradient directions, calculating the $TV_{n,m}$ norm by numerically integrating over the 3-dimensional data set at each step of an iterative process would be prohibitively expensive. In contrast using our modified TV-norm suggested earlier leads to a more efficient solution.

1) Fixed-Point Lagged-Diffusivity: Since the m Equations(35) are coupled only through the function g , we can drop the subscript on S with no ambiguity (later the subscript will refer to spatial coordinates.) In this section we will discuss the numerical solution for each channel, S , of the vector-valued image \mathbf{S} . Equation (35) is nonlinear due to the presence of $\|\nabla S_i\|$ in the denominator of the first term. We linearize Equation (35) by using the method of ‘‘lagged-diffusivity’’ presented by Chan and Mulet [46]. By considering $\|\nabla S\|$ to be a constant for each iteration, and using the value from the

previous iteration we can instead solve

$$-\frac{1}{\|\nabla S^t\|} (\nabla g \cdot \nabla S^t + g \nabla^2 S^{t+1}) + \mu(S^{t+1} - S^0) = 0 \quad (37)$$

Here the superscript denotes iteration number. First, rewrite Equation (37) with all of the S^{t+1} terms on the left-hand side

$$-\nabla^2 S^{t+1} + \frac{\mu \|\nabla S^t\|}{g} S^{t+1} = \frac{\mu \|\nabla S^t\| S^0 + \nabla g \cdot \nabla S^t}{g} \quad (38)$$

2) Discretized Equations: To write Equation (38) as a linear system ($\mathbf{A}\mathbf{S}^{t+1} = \mathbf{f}^t$), discretize the Laplacian and gradient terms. Using central differences for the Laplacian we have

$$\begin{aligned} \nabla^2 S^{t+1} &= S_{x-1,y,z}^{t+1} + S_{x,y-1,z}^{t+1} + S_{x,y,z-1}^{t+1} \\ &- 6S_{x,y,z}^{t+1} + S_{x+1,y,z}^{t+1} + S_{x,y+1,z}^{t+1} + S_{x,y,z+1}^{t+1} \end{aligned} \quad (39)$$

Define the standard central differences to be

$$\begin{aligned} \Delta_x S &= \frac{1}{2}(S_{x+1,y,z} - S_{x-1,y,z}) \\ \Delta_y S &= \frac{1}{2}(S_{x,y+1,z} - S_{x,y-1,z}) \\ \Delta_z S &= \frac{1}{2}(S_{x,y,z+1} - S_{x,y,z-1}) \end{aligned} \quad (40)$$

We can rewrite Equation (38) in discrete form using the definitions in Equation (40)

$$\begin{aligned} &-S_{x-1,y,z} - S_{x,y-1,z} - S_{x,y,z-1} \\ &+ (6 + \frac{\mu \sqrt{(\Delta_x S^t)^2 + (\Delta_y S^t)^2 + (\Delta_z S^t)^2}}{g}) S_{x,y,z} \\ &- S_{x+1,y,z} - S_{x,y+1,z} - S_{x,y,z+1} \\ &= \frac{1}{g} (\mu S^0 \sqrt{(\Delta_x S^t)^2 + (\Delta_y S^t)^2 + (\Delta_z S^t)^2} \\ &+ \Delta_x g \Delta_x S^t + \Delta_y g \Delta_y S^t + \Delta_z g \Delta_z S^t) \end{aligned} \quad (41)$$

This results in a sparse banded linear system with 7 nonzero coefficients per row.

$$\begin{bmatrix} 6 + \frac{\mu \|\nabla S^t\|_0}{g_0} & -1 & \dots & -1 & \dots & -1 & \dots \\ -1 & 6 + \frac{\mu \|\nabla S^t\|_1}{g_1} & & -1 & \dots & -1 & \dots \\ 0 & & 6 + \frac{\mu \|\nabla S^t\|_2}{g_2} & & -1 & \dots & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} S_0^{t+1} \\ S_1^{t+1} \\ \vdots \\ S_n^{t+1} \end{bmatrix} = \begin{bmatrix} f_0^t \\ f_1^t \\ \vdots \\ f_n^t \end{bmatrix} \quad (42)$$

where the right-hand side of Equation (41) has been replaced with f_n^t . The matrix in Equation (42) is symmetric and diagonally dominant. We have successfully used conjugate gradient descent to solve this system.

The solution of Equation (42) represents one fixed-point iteration. This iteration is continued until $|S^t - S^{t+1}| < c$, where c is a small constant.

C. Computing Probabilities

The probability in Equation 2 can now be evaluated by computing the quantity $S(q)/S_0$ and performing the FFT. Since we know that the signal, S , decays exponentially from the origin of q -space (where $S(0) = S_0$) we can interpolate signal values for arbitrary q . We resample from spherical coordinates to cartesian and perform the FFT on the resampled data. The result is a probability of water molecule displacement over a small time constant. We are interested in only the direction of

water displacement, so we integrate out the radial component of $p_t(\mathbf{r})$ to get $p_t(\theta, \phi)$. This is commonly referred to as the diffusion orientation distribution function (ODF). Computing the ODF with this method is computationally expensive since it requires a 3D FFT at each voxel, and then a numerical integration for each direction. In this section we present an alternative method which makes computing the ODF for large datasets extremely efficient.

The Fourier transform that relates the signal attenuation to the water displacement probability (Eq. 2) can be written in spherical coordinates. This is a consequence of the pointwise convergent expansion of the plane wave in spherical coordinates [47] given by

$$e^{\pm 2\pi i \mathbf{q} \cdot \mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l (\pm i)^l j_l(2\pi q r) Y_{lm}(\theta, \phi) Y_{lm}^*(\theta_r, \phi_r) \quad (43)$$

where j_l is the l -th order spherical Bessel function, Y_{lm} is the spherical harmonics and θ_r and ϕ_r are the spherical coordinates depicting the direction of the vector \mathbf{r} . Inserting this expression into Eq. 2 and setting $r = r_0$, we get

$$p_t(r_0, \theta_r, \phi_r) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (-i)^l Y_{lm}(\theta_r, \phi_r) \int_S Y_{lm}(\theta, \phi) Y_{lm}^*(\theta, \phi) I_l(\theta, \phi) dS \quad (44)$$

where $I_l(\theta, \phi)$ are evaluated from the radial part of the integral in q -space. If we assume that along each direction in q -space the signal values attenuate according to Equation (1), then the radial integrals yield

$$I_l(\theta, \phi) = \frac{r_0^l \Gamma(\frac{l+3}{2})}{2^{l+3} \pi^{3/2} (D(\theta, \phi) t)^{(l+3)/2} \Gamma(l+3/2)} {}_1F_1\left(\frac{l+3}{2}; l + \frac{3}{2}; -\frac{r_0^2}{4D(\theta, \phi)t}\right) \quad (45)$$

where ${}_1F_1$ is the confluent hypergeometric function [48].

Note that the function $p_t(r_0, \theta_r, \phi_r)$ is the probability of the water molecule to move from the origin to a point r_0 away from the origin along the direction specified by the spherical coordinates θ_r and ϕ_r , i.e., we will be interested in the probability values on a sphere of radius r_0 . In order to estimate the probability on the surface of a sphere of radius r_0 , we go back to Eqs. 44 and 45, and expand $I_l(\theta, \phi)$ in Laplace series, i.e.,

$$I_l(\theta, \phi) = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} \alpha_{ll'm'} Y_{l'm'}(\theta, \phi) \quad (46)$$

where

$$\alpha_{ll'm'} = \int_S Y_{l'm'}(\theta, \phi) Y_{lm}^*(\theta, \phi) dS \quad (47)$$

Comparing the integration over the sphere in Eq. 44 with the expression in Eq. 47, it is easy to see that

$$p_t(r_0, \theta_r, \phi_r) = \sum_{l=0}^{\infty} \sum_{m=-l}^l (-i)^l \alpha_{llm} Y_{lm}(\theta_r, \phi_r) \quad (48)$$

which is just a Laplace series expansion of $p_t(r_0, \theta, \phi)$. Note that coefficients of this Laplace series for some l value come from the l -th order spherical harmonic transform of $I_l(\theta, \phi)$.

In summary, the estimation of the angular probability distribution at a distance r_0 away from the origin involves the following steps:

- Given HARDI data including N_G images acquired with gradient directions along different orientations, calculate the diffusivity $D(\theta, \phi)$ along each direction.
- Calculate $I_l(\theta, \phi)$ using Eq. 45.
- Calculate α_{llm} , the l -th order spherical harmonic transform of $I_l(\theta, \phi)$ for each l .
- Evaluate Eq. 48, which is just

$$p_t(r_0, \theta_r, \phi_r) = \frac{\alpha_{000}}{\sqrt{4\pi}} - \sum_{m=-2}^2 \alpha_{22m} Y_{2m}(\theta_r, \phi_r) + \sum_{m=-4}^4 \alpha_{44m} Y_{4m}(\theta_r, \phi_r) - \sum_{m=-6}^6 \alpha_{66m} Y_{6m}(\theta_r, \phi_r) \pm \dots \quad (49)$$

Note that only even order l terms are retained which is a consequence of the antipodal symmetry of $D(\theta, \phi)$. This scheme provides a fast way to calculate the orientation profiles. In our implementation we have evaluated the series given in Eq. 49 up to $l = 6$ terms, and r_0 was set to $17.5 \mu m$.

D. Visualizing Probabilities

In diffusion tensor image processing, the scalar quantity known as fractional anisotropy (FA) is often considered a useful quantity to visualize. An example of an FA image is shown in Figure (5). The FA value at each voxel can be computed from the eigenvalues of the tensor at that voxel. The values of FA range from 0, for completely isotropic diffusion, to 1, for completely anisotropic diffusion. The bright regions of Figure (5) correspond to high FA values. In rank-2 tensor images FA can indicate the presence of a single fiber direction within a voxel.

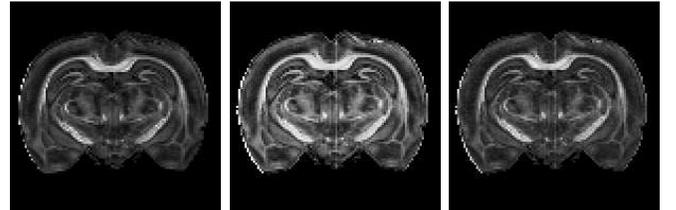


Fig. 5. FA image (left), generalized anisotropy (center), Shannon anisotropy (HA) (right), from coronal slice of rat brain.

Since we have a distribution at each voxel, we can compute the Shannon entropy value at each voxel as given by,

$$H(p) = - \sum_{i=1}^n p(\theta_i, \phi_i) \log p(\theta_i, \phi_i). \quad (50)$$

Considering the entropy of several example distributions, we can get a feel for the interpretation of entropy in the context of HARDI. Entropy attains its maximum value for a uniform distribution. In our case, this corresponds to isotropic diffusion. The entropy of a Gaussian distribution decreases as the variance decreases. A voxel with this distribution has oriented diffusion greatest in the direction of the mode of the distribution, implying the presence of fibrous structure. Tuch [49] defined a scalar measure of anisotropy called "normalized entropy", which was also based on Shannon entropy.

The anisotropy images we present are mapped using Equation (51) such that the black color is high entropy (isotropic diffusion) and higher intensity grey colors represent low entropy (anisotropic diffusion). We denote the anisotropy measure computed from Shannon entropy as HA .

$$HA(p) = (1.0 + \frac{H(p)}{\log \frac{1}{n}}) \quad (51)$$

This allows the image to be interpreted in the same way as a fractional anisotropy image where the white color corresponds to white matter, grey corresponds to grey matter and black corresponds to cerebrospinal fluid.

The HA index is not a generalization of FA, and their values cannot be compared in a meaningful way. In order to highlight the difference in anisotropy measures for HARDI and DTI, a Gaussian ODF was computed from the tensor data, and the Shannon entropy of these distributions was computed. The anisotropy values for the Gaussian ODFs can then be compared to the anisotropy values for the true ODFs. The difference image is shown in Figure (6). It can be seen that there is a structure to the difference image. *This difference is more pronounced in regions where the tensor model predicts diffusion is more isotropic than it actually is.*



Fig. 6. Shannon anisotropy difference between Gaussian ODF and general ODF.

Shannon entropy is not the only measure of uncertainty in a random variable. Another definition of entropy, called the Rényi entropy can be seen as a generalization of Shannon entropy, [50]. Rényi entropy of order α is given by

$$H_\alpha(p) = \frac{1}{1-\alpha} \log \left(\sum_{i=1}^n p(\theta_i, \phi_i)^\alpha \right) \quad (52)$$

The parameter, α , of this entropy formulation has several interesting properties:

- $\lim_{\alpha \rightarrow 1} H_\alpha(p) = H(p)$ (Shannon entropy)
- $H_0(p) =$ number of nonempty bins in the histogram of p .
- $H_\infty(p) = -\log(\max_i(p_i))$ (depends only on the mode of p)

One may interpret the order, α , as a parameter which changes the shape of the distribution, p . For $\alpha > 1$, small values of $p(x)$ will shrink closer to zero. As α increases, these small probabilities approach zero more quickly. For high α , events with high probability influence the entropy more. This can be seen as controlling the contrast between white and gray matter. We can formulate an anisotropy index based on Rényi entropies just as we did for Shannon entropy. Anisotropy images computed from Rényi entropies of different orders are presented in Figure (7).

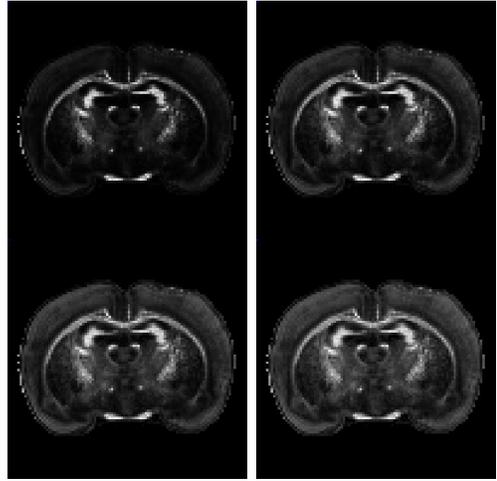


Fig. 7. Anisotropy computed from H_2 (top-left), H_5 (top-right), H_{10} (bottom-left), H_{20} (bottom-right).

In statistical physics, a quantity called structural entropy has useful physical interpretation [51]. This entropy is computed as the difference $H_1 - H_2$. It is unclear whether there is any such physical meaning for this quantity in the context of diffusion, but it does provide a means of applying a transfer function to entropy images which has a firm information theoretic footing. We present entropy difference images for several values of α in Figure (8).

The scalar entropy value has no directional information however. In DTI, there is precedence for using color values to represent direction. Color FA images are a mapping of the principal diffusion direction (the dominant eigenvector of the tensor) to a hue, and the fractional anisotropy value to an intensity. These images are useful for distinguishing between adjacent anisotropic regions which differ in direction. Since we may have high diffusion coefficients in several directions, we integrate over the sphere to determine a representative color for each voxel. We denote this value as ED , for "expected direction".

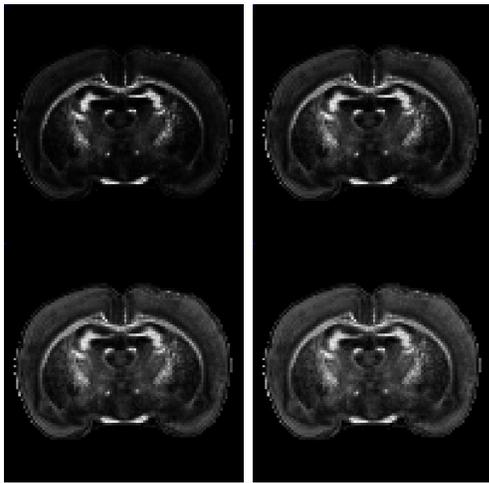


Fig. 8. Rényi entropy differences $H_1 - H_2$ (top-left), $H_1 - H_5$ (top-right), $H_1 - H_{10}$ (bottom-left), $H_1 - H_{20}$ (bottom-right).

$$\mathbf{ED} = \sum_{i=1}^n \begin{bmatrix} |\cos \theta_i \sin \phi_i| \\ |\sin \theta_i \sin \phi_i| \\ |\cos \phi_i| \end{bmatrix} (p(\theta_i, \phi_i) - p_{min}) \quad (53)$$

The resulting vector is interpreted as red, green, and blue color components. The directions with the highest probability of diffusion will have their corresponding color contribute most to the resulting color. Figure (10) depicts the colors computed for a synthetic ODF field directed in a curve. Figure (11) shows the colors computed for ODFs oriented in various directions, both with and without crossings.

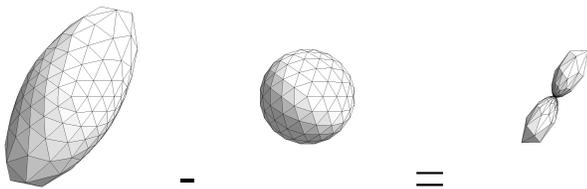


Fig. 9. Original ODF (left), Minimum probability sphere (center), and sharpened ODF (right).

To enhance the visual impact of the ODF we apply a sharpening transform to the ODF by subtracting a uniform distribution (sphere) from each distribution, as shown in Figure (9). The radius of the sphere is the minimum of the probability over all directions. By performing this operation the direction of maximum probability becomes more apparent. This causes regions of isotropic diffusion to have lower intensity in the color expected direction images.

IV. EXPERIMENTAL RESULTS

The denoising and rendering techniques described in the previous section were first applied to a synthetic HARDI dataset. This dataset was generated using the technique described by Özarlan et al. in [40]. The dataset was designed to depict a region of curving fibers, a region of straight fibers, and

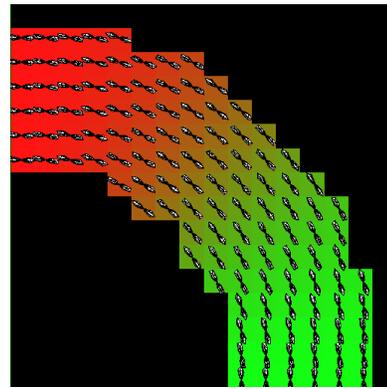


Fig. 10. ED colors for synthetic distribution field.

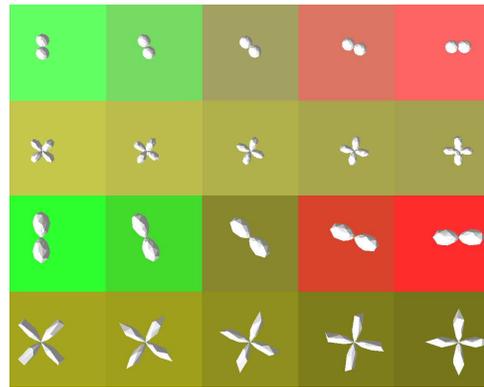


Fig. 11. ED colors for synthetic distribution field.

a crossing between the two. Representative sharpened ODF profile are shown overlaid on the HA image on the left side of Figure (12). The expected direction visualization shown on the right side of Figure (12).

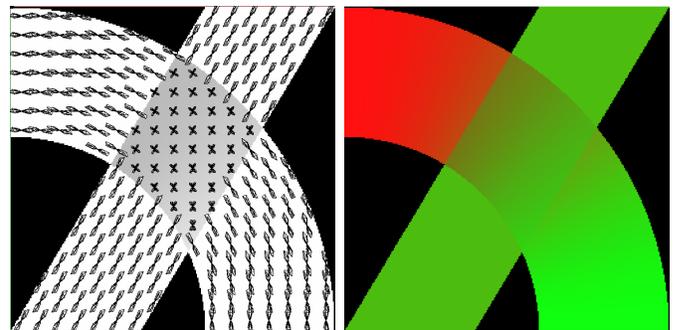


Fig. 12. ODF profiles overlaid on HA image (left), and ED (right) of the synthetic dataset.

A small sample of the synthetic data, taken from near the crossing region, is shown in Figure (13). The synthetic data was corrupted with Gaussian noise of mean zero, and variance $\sigma^2 = 0.005$. The noisy data is shown in Figure (14). The same voxels are shown after smoothing over the spherical manifold in Figure (15), after smoothing over the cartesian image domain in Figure (16) and after both techniques have been used in Figure (17). The right-hand side of each figure shows the sharpened ODF computed from the S values depicted on

the left-hand side.

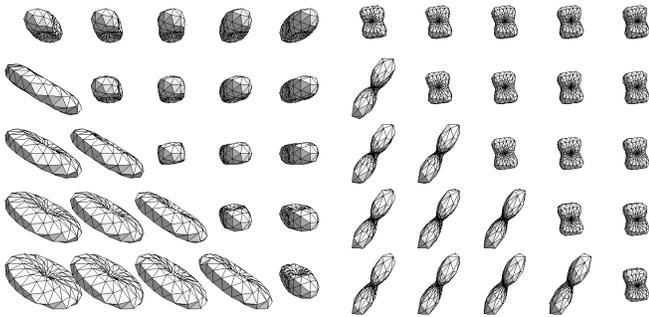


Fig. 13. Synthetic S (left), resulting ODF (right).

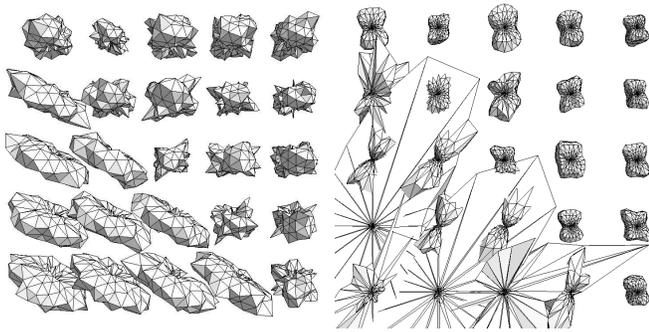


Fig. 14. Synthetic S with noise added (left), resulting ODF (right).

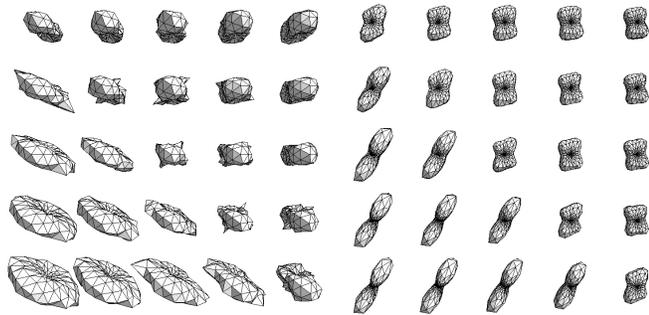


Fig. 15. Manifold (FEM) smoothing results for S (left), resulting ODF (right).

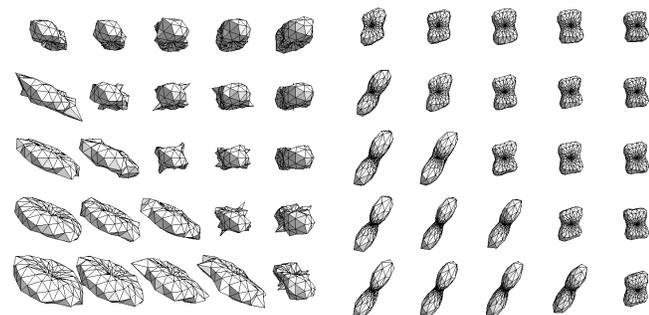


Fig. 16. Lattice (TV) smoothing results for S (left), resulting ODF (right).

From Figure (14), it can be seen that the noise has a large influence on the smoothness of the ODF. As expected from the

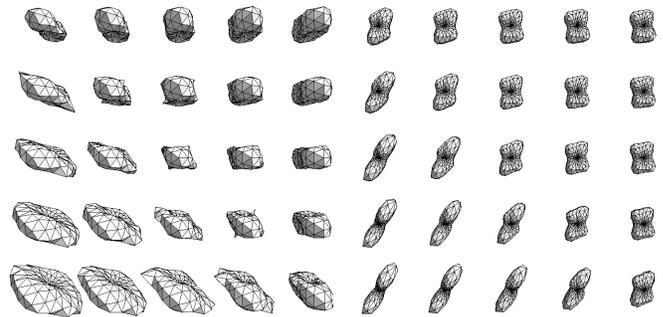


Fig. 17. Manifold and lattice smoothed S (left), resulting ODF (right).

TABLE II

ERROR BETWEEN GROUND-TRUTH AND RESTORED SYNTHETIC DATA

Method	$\mu(d(\hat{p}, p))$	$\sigma^2(d(\hat{p}, p))$
$p = \text{No Restoration}$	1.0409	0.0173
$p = \text{FEM Restoration}$	0.9088	0.0125
$p = \text{TV Restoration}$	0.7420	0.0119
$p = \text{FEM + TV Restoration}$	0.6576	0.0139

variational formulation, the spikes of noise present in the raw data have been smoothed while preserving the overall shape the S profile. This smoothness is propagated to the computed ODF profiles.

We can compare the resulting distributions with the ground-truth by using the square root of J-divergence (symmetric KL-divergence) as a measure. This divergence is defined as

$$d(p, q) = \sqrt{J(p, q)} \quad (54)$$

where

$$J(p, q) = \frac{1}{2} \sum_{i=1}^n p(\theta_i, \phi_i) \log \frac{p(\theta_i, \phi_i)}{q(\theta_i, \phi_i)} + q(\theta_i, \phi_i) \log \frac{q(\theta_i, \phi_i)}{p(\theta_i, \phi_i)} \quad (55)$$

In Table (II) we compare the distances, $d(\hat{p}, p)$, between the original synthetic data, (\hat{p}) , and the unrestored data, the data restored only using the FEM method, the data restored using only the TV-norm minimization, and the data restored using both techniques. For each technique, the mean distance $\mu(d(\hat{p}, p))$, as well as the variance $\sigma^2(d(\hat{p}, p))$, between the distributions in corresponding voxels is presented.

As shown in Table (II), the TV restoration had superior performance to the FEM technique, both in terms of the mean error and variance of the error. The combination of techniques had a lower mean error than either the FEM or TV restoration, however the variance of the error was higher than that with either technique alone.

The denoising algorithm was applied to a dataset consisting of 47 diffusion weighted images of a rat spinal cord. Axial slices of one such image, before and after denoising are shown in Figure (18).

The ringing artifacts visible near the sample boundary in the raw DWI in Figure (18) have been noticeably decreased. Note that the edges in the image have been well preserved. From the expected direction images in Figure (19), it is clear that

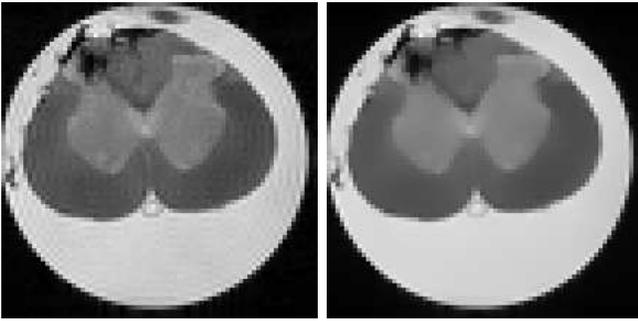


Fig. 18. Original diffusion-weighted image (left), and denoised (right) from spinal cord data.

the white matter fiber tracts are predominantly in the axial direction, which is represented by the blue color.

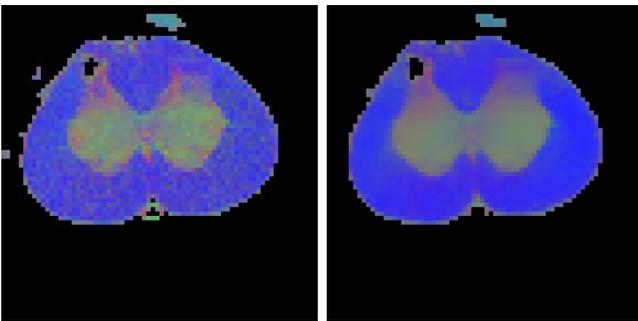


Fig. 19. Original expected direction image (left), and denoised (right) from spinal cord data.

Figures (20) and (21) show the HA and expected direction visualization results respectively for various coronal slices of a rat brain dataset. The chosen slices show the corpus callosum, an anatomical structure known to have long-range white matter tracts.

The HA images show a visible distinction between grey and white matter. In addition, the distinction between fiber directions is evident in the expected direction images. Red corresponds to the left- right direction and green corresponds to up-down and blue is in-out of the page.

V. CONCLUSION

In this paper, we presented a new variational formulation for restoring HARDI data, an FEM technique for implementing the restoration, a novel and efficient technique for computing the water molecule diffusion probability over a sphere of directions, and two novel techniques for visualizing the water molecule diffusion probability fields (computed from this restored data). To the best of our knowledge, there are no reported results on efficient computation and visualization of these probability fields. Our formulation of the HARDI restoration involves two types of smoothness constraints. The first is smoothness over the spherical domain of acquisition directions, and the second is smoothness between neighboring voxels in the cartesian domain. The smoothing technique is

capable of preserving discontinuities in the data. This was demonstrated on synthetic and real anatomical data. By using J-divergence as a measure of distance between distribution, we were able to show quantitatively that the combination of restoration techniques performs better than either technique alone.

A statistical property of the ODF, namely entropy, was shown to be a useful indicator of anisotropy computed from HARDI data. Novel results showing how expected direction images computed from the ODF could be used to visualize diffusion direction and anisotropy in HARDI data were presented. The effect of the restoration on these measures was shown to improve the clarity of the images.

ACKNOWLEDGEMENT

This research was supported in part by the grant NIH-NS42075 and by Siemens Corporate Research (Princeton, NJ). We wish to thank Sara Berens and Robert Yeziarski for providing the spinal cord sample and Ron Hayes for the brain.

REFERENCES

- [1] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient," *J. Chem. Phys.*, vol. 42, no. 1, pp. 288–292, 1965.
- [2] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.*, vol. 66, pp. 259–267, 1994.
- [3] P. J. Basser, J. Mattiello, and D. LeBihan, "Estimate of the effective self-diffusion tensor from the NMR spin echo," *J. Magn. Reson.*, vol. B, no. 103, pp. 247–254, 1994.
- [4] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 44, pp. 625–632, 2000.
- [5] P. J. Basser, "New histological and physiological stains derived from diffusion-tensor MR images," *Ann. N Y Acad. Sci.*, vol. 820, pp. 123–138, 1997.
- [6] D. K. Jones, A. Simmons, S. C. R. Williams, and M. A. Horsfield, "Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI," *Magn. Reson. Med.*, vol. 42, pp. 37–41, 1999.
- [7] N. Makris, A. J. Worth, A. G. Sorensen, G. M. Papadimitriou, O. Wu, T. G. Reese, V. J. Wedeen, and *et al*, "Morphometry of in vivo human white matter association pathways with diffusion-weighted magnetic resonance imaging," *Ann. Neurol.*, vol. 42, pp. 951–962, 1999.
- [8] T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle, "Tracking neuronal fiber pathways in the living human brain," *Proc. Natl. Acad. Sci.*, vol. 96, pp. 10422–10427, 1999.
- [9] S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, pp. 265–269, 1999.
- [10] R. Xue, P. C. M. van Zijl, B. J. Crain, H. Solaiyappan, and S. Mori, "In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging," *Magn. Reson. Med.*, vol. 42, pp. 1123–1127, 1999.
- [11] B. C. Vemuri, Y. Chen, M. Rao, Z. Wang, T. McGraw, T. Mareci, S. J. Blackband, and P. Reier, "Automatic fiber tractography from DTI and its validation," in *IEEE Intl. Symp. on Biomedical Imaging*, 2002, pp. 501–504.

- [12] T. McGraw, B. Vemuri, Z. Wang, Y. Chen, M. Rao, and T. Mareci, "Line integral convolution for visualization of fiber tract maps from DTI," in *Fifth Intl. Conf. on MICCAI*, Tokyo, Japan, 2002, pp. 615–622.
- [13] T. McGraw, B. C. Vemuri, Y. Chen, M. Rao, and T. Mareci, "DT-MRI denoising and neuronal fiber tracking," *Medical Image Analysis*, vol. 8, no. 2, pp. 95–111, June 2004.
- [14] B. C. Vemuri, Y. Chen, M. Rao, T. McGraw, Z. Wang, and T. Mareci, "Fiber tract mapping from diffusion tensor MRI," in *VLSM '01: Proceedings of the IEEE Workshop on Variational and Level Set Methods (VLSM'01)*. Vancouver, Canada: IEEE Computer Society, 2001, p. 81.
- [15] P. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*. Oxford: Clarendon Press, 1991.
- [16] V. J. Wedeen, T. G. Reese, D. S. Tuch, M. R. Weigel, J. G. Dou, R. Weiskoff, and D. Chessler, "Mapping fiber orientation spectra in cerebral white matter with fourier transform diffusion MRI," in *Proc. of the 8th Annual Meeting of ISMRM*, Denver, CO, 2000, p. 82.
- [17] P. J. Basser, "Relationships between diffusion tensor and q-space MRI," *Magn. Reson. Med.*, vol. 47, pp. 392–397, 2002.
- [18] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen, "High angular resolution diffusion imaging of the human brain," in *Proc. of the 7th Annual Meeting of ISMRM*, Philadelphia, PA, 1999, p. 321.
- [19] D. S. Tuch, T. G. Reese, M. R. Wiegell, and V. J. Wedeen, "Diffusion MRI of complex neural architecture," *Neuron*, vol. 40, pp. 885–895, 2003.
- [20] L. R. Frank, "Characterization of anisotropy in high angular resolution diffusion-weighted MRI," *Magn Reson Med*, vol. 47, no. 6, pp. 1083–1099, 2002.
- [21] K. M. Jansons and D. C. Alexander, "Persistent angular structure: new insights from diffusion magnetic resonance imaging data," *Inverse problems*, vol. 19, pp. 1031–1046, 2003.
- [22] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B. C. Vemuri, and Y. Liu, "Estimation, smoothing, and characterization of apparent diffusion coefficient profiles from high angular resolution dwi," in *CVPR (1)*, 2004, pp. 588–593.
- [23] O. Coulon, D. C. Alexander, and S. R. Arridge, "A regularization scheme for diffusion tensor magnetic resonance images," in *Proc. of the 17th Intl. Conf. on Information Processing in Medical Imaging*, Davis, CA, 2001, pp. 92–105.
- [24] T. F. Chan and J. Shen, "Restoration of non-flat image features: models and algorithms," *SIAM J. Appl. Math.*, vol. 61, no. 4, pp. 1338–1361, 2000.
- [25] D. Tschumperlé and R. Deriche, "Variational frameworks for DT-MRI estimation, regularization and visualization," *ICCV*, pp. 116–121, 2003.
- [26] C. Chef d'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, "Regularizing flows for constrained matrix-valued images," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 147–162, 2004. [Online]. Available: <ftp://ftp-sop.inria.fr/odyssee/Publications/2004/chefdhotel-tschumperle-et-al:04.pdf>
- [27] J. Weickert and T. Brox, "Diffusion and regularization of vector- and matrix-valued images," *Contemporary Mathematics*, vol. 313, pp. 251–268, 2002.
- [28] F. Wang, B. C. Vemuri, M. Rao, and Y. Chen, "A new and robust information theoretic measure and its application to image alignment," in *Proc. of the 18th Intl. Conf. on Information Processing in Medical Imaging*, 2003, pp. 388–400.
- [29] —, "Cumulative residual entropy, a new measure of information and its application to image alignment," *ICCV*, pp. 548–553, 2003.
- [30] T. E. Conturo, R. C. McKinstry, E. Akbudak, and B. H. Robinson, "Encoding of anisotropic diffusion with tetrahedral gradients: A general mathematical diffusion formalism and experimental results," *Magn. Reson. Med.*, vol. 35, pp. 399–412, 1996.
- [31] C. F. Westin, S. E. Maier, B. Khidir, P. Everett, F. A. Jolesz, and R. Kikinis, "Image processing for diffusion tensor magnetic resonance imaging," in *Second Intl. Conf. on MICCAI*, 1999, pp. 441–452.
- [32] C. Poupon, C. A. Clark, V. Frouin, J. Régis, I. Bloch, D. L. Bihan, and J.-F. Mangin, "Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles," *NeuroImage*, vol. 12, no. 2, pp. 184–195, 2000.
- [33] G. J. M. Parker, C. A. M. Wheeler-Kinghtshott, and G. J. Baker, "Distributed anatomical brain connectivity derived from diffusion tensor imaging," in *Proc. of the 17th Intl. Conf. on Information Processing in Medical Imaging*, Davis, CA, 2001, pp. 106–120.
- [34] P. G. Batchelor, D. L. G. Hill, F. Calamante, and D. Atkinson, "Study of connectivity in the brain using the full diffusion tensor from MRI," in *Proceedings of the 17th Intl. Conf. on Information Processing in Medical Imaging*, Davis, CA, 2001, pp. 121–133.
- [35] M. Björnemo, A. Brun, R. Kikinis, and C. F. Westin, "Regularized stochastic white matter tractography using diffusion tensor MRI," in *Fifth Intl. Conf. on MICCAI*, vol. 1, 2002, pp. 435–442.
- [36] L. O'Donnell, S. Haker, and C.-F. Westin, "New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space," in *Fifth Intl. Conf. on MICCAI*, Tokyo, Japan, 2002, pp. 459–466.
- [37] T. Behrens, H. Johansen-Berg, M. Woolrich, S. Smith, C. Wheeler-Kingshott, P. Boulby, G. Barker, E. Sillery, K. Sheehan, O. Ciccarelli, A. Thompson, J. Brady, and P. Matthews, "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging," *Nature Neuroscience*, vol. 6, no. 7, pp. 750–757, July 2003.
- [38] E. Özarslan and T. H. Mareci, "Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging," *Magn Reson Med*, vol. 50, pp. 955–965, 2003.
- [39] P. Sundgren, Q. Dong, D. Gómez-Hassan, S. Mukherji, P. Maly, and R. Welsh, "Diffusion tensor imaging of the brain: review of clinical applications," *Neuroradiology*, vol. 46, pp. 339–350, May 2004.
- [40] E. Özarslan, B. C. Vemuri, and T. H. Mareci, "Generalized scalar measures for diffusion MRI using trace, variance and entropy," *Magn Reson Med*, vol. 53, no. 4, pp. 866–876, 2005.
- [41] B. Tang, G. Sapiro, and V. Caselles, "Diffusion of general data on non-flat manifolds via harmonic maps theory: The direction diffusion case," *IJCV*, vol. 36, no. 2, pp. 149–161, 2000.
- [42] L. R. Frank, "Anisotropy in high angular resolution diffusion-weighted MRI," *Magn Reson Med*, vol. 45, no. 6, pp. 935–939, 2001.
- [43] T. McInerney and D. Terzopoulos, "A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis," *Comp. Med. Imag. Graph.*, vol. 19, no. 1, pp. 69–83, 1995. [Online]. Available: <citeseer.ist.psu.edu/mcinerney95dynamic.html>
- [44] G. Dhatt and G. Touzot, *The finite element method displayed*. New York: Wiley, 1984.
- [45] P. Blomgren and T. F. Chan, "Color TV: total variation methods for restoration of vector-valued images," *IEEE Transaction on Image Processing*, vol. 7, no. 3, pp. 304–309, 1998.
- [46] T. Chan and P. Mulet, "On the convergence of the lagged diffusivity fixed point method in total variation image restoration," *SIAM Journal on Numerical Analysis*, vol. 36, no. 2, pp. 354–367, 1999.
- [47] F. Schwabl, *Quantum Mechanics*. Berlin: Springer-Verlag, 1989.
- [48] G. Arfken and H. Weber, *Mathematical Methods for Physicists*. London: Academic Press, 1995.
- [49] D. S. Tuch, "Q-ball imaging," *Magn Reson Med.*, vol. 52, no. 6, pp. 1358–1372, 2004.
- [50] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & sons, 1991.
- [51] K. Zyczkowski, "Renyi extrapolation of shannon entropy," *Open Syst. Inf. Dyn.*, vol. 10, pp. 297–310, 2003.

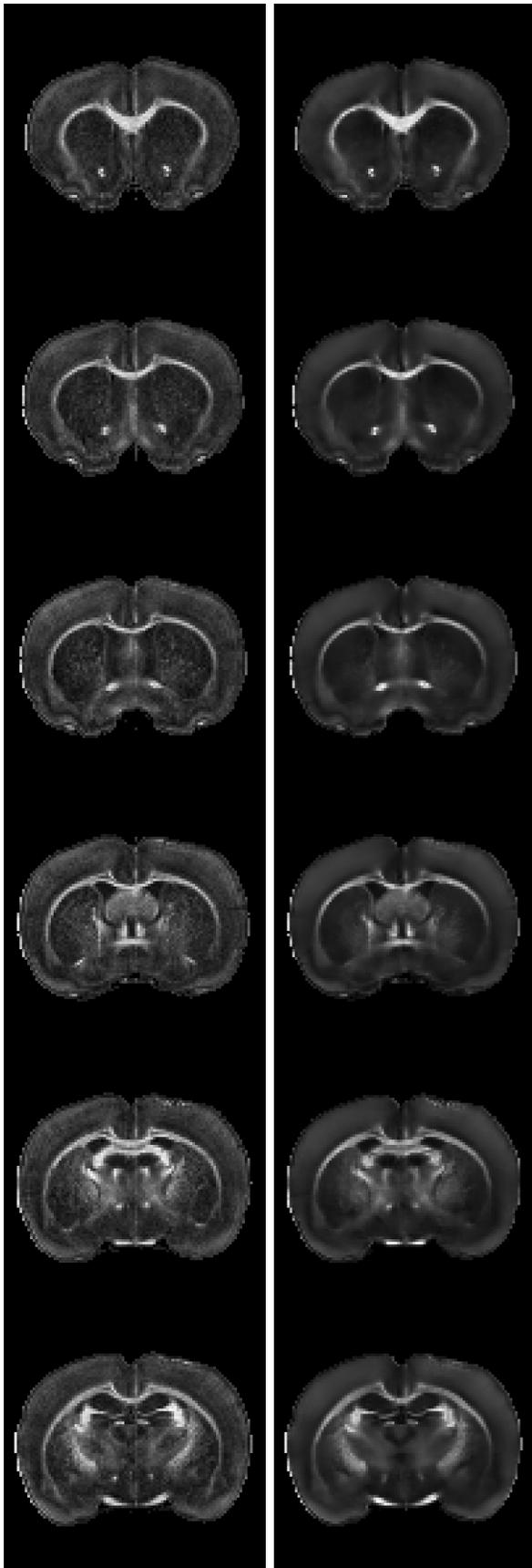


Fig. 20. Original (left), and denoised (right) HA images for coronal slices of rat brain.

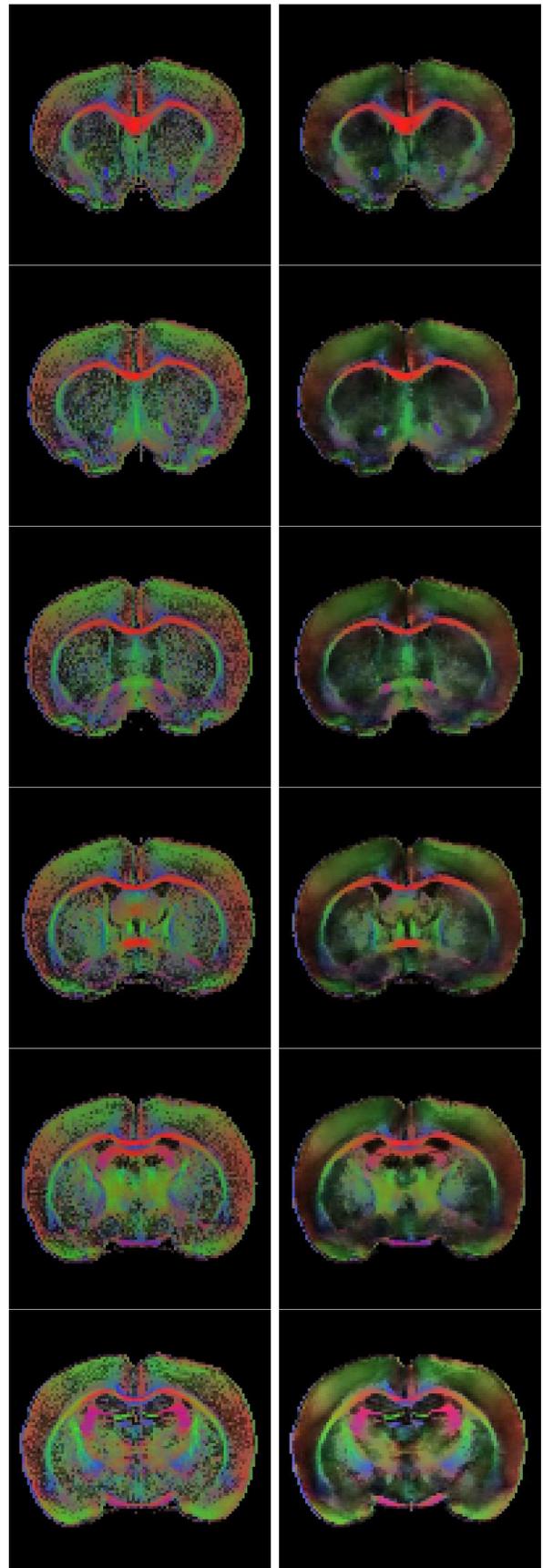


Fig. 21. Original (left), and denoised (right) ED images for coronal slices of rat brain.