Inverse Problems and Imaging
Volume 2, No. 2, 2008, 205–224

Web site: http://www.aimSciences.org

A NONSTANDARD SMOOTHING IN RECONSTRUCTION OF APPARENT DIFFUSION COEFFICIENT PROFILES FROM DIFFUSION WEIGHTED IMAGES

Yunmei Chen[1], Weihong Guo[2], Qingguo Zeng[1] and Yijun Liu[3]

[1] Department of Mathematics, University of Florida
Gainesville, FL 32611, USA

[2] Department of Mathematics, University of Alabama
Tuscaloosa, AL 35487, USA

[3] Department of Psychiatry & Neuroscience, University of Florida
Gainesville, FL 32653, USA

(Communicated by Jackie Shen (retired))

Abstract. We present a new variational framework for simultaneous smoothing and estimation of apparent diffusion coefficient (ADC) profiles from High Angular Resolution Diffusion-weighted MRI. The model approximates the ADC profiles at each voxel by a 4th order spherical harmonic series (SHS). The coefficients in SHS are obtained by solving a constrained minimization problem. The smoothing with feature preserved is achieved by minimizing a variable exponent, linear growth functional, and the data constraint is determined by the original Stejskal-Tanner equation. The antipodal symmetry and positiveness of the ADC are accommodated in the model. We use these coefficients and variance of the ADC profiles from its mean to classify the diffusion in each voxel as isotropic, anisotropic with single fiber orientation, or two fiber orientations. The proposed model has been applied to both simulated data and HARD MRI human brain data . The experiments demonstrated the effectiveness of our method in estimation and smoothing of ADC profiles and in enhancement of diffusion anisotropy. Further characterization of non-Gaussian diffusion based on the proposed model showed a consistency between our results and known neuroanatomy.

1. **Introduction.** Diffusion-weighted MRI (DW-MRI, shortened as DWI) adds to conventional MRI the capability of measuring the random motion of water molecules, referred as diffusion. Water in tissues contains a large number of fibers (such as cardiac muscle and brain white matter). The diffusion is the fastest along the direction that a fiber is pointing to, but the slowest in the direction perpendicular to it. Water diffuses isotropically in tissues that contain few fibers. DWI renders non-invasively such complex in vivo information about how water diffuses into intricate a 3D representation of tissues.

Conventional MRI can be used to investigate spatial relationship between different anatomical regions, but it is unable to infer the connectivity of these regions.

2000 *Mathematics Subject Classification.* Primary: 15A29, 94A08.

Key words and phrases. Reconstruction, smoothing, diffusion weighted images, spherical harmonic series.

The first author is supported by NSF CCF-0527967 and NIH R01 NS052831-01 A1.

DWI provides profound histological and anatomical information about tissue structure, composition, architecture, and organization. Changes in these tissue properties can often be correlated with processes that occur in development, degeneration, disease, and aging, so this method has become more and more widely applied ([1, 2, 3, 4]).

The diffusion of water molecules in tissues over a time interval with length t can be described by a probability density function (pdf) $p(\mathbf{r}, t)$ on the displacement $\mathbf{r}$. Since $p(\mathbf{r}, t)$ is largest in the directions of least hindrance to diffusion and smaller in other directions, the information about $p(\mathbf{r}, t)$ reveals fiber orientations and leads to meaningful inferences about the microstructure of tissues. Therefore, it can be used to characterize diffusion anisotropy and reconstruct fiber pathways.

The standard methodology employed in most DWI experiments is the Stejskal-tanner pulsed gradient spin echo method [5]. Two magnetic field gradient pulses of strength $\mathbf{G}$ and duration δ with a temporal separation of t between the onset of the pulses are applied to the simple spin-echo sequence. If the duration of the pulses δ is negligible comparing with t, the attenuation of the MR signal $s(\mathbf{q})$ with respect to the diffusion sensitizing gradient $\mathbf{q}$ measures the Fourier transformation (FT) of the average pdf $p(\mathbf{r}, t)$ on a spin displacement $\mathbf{r}$ over diffusion time t [6]. That is

$$s(\mathbf{q}) = s_0 \int p(\mathbf{r}, t) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}, \tag{1}$$

where $\mathbf{q} = (2\pi)^{-1}\gamma\delta G$, γ is gyromagnetic ratio of protons in water, and s_0 is the MR signal in the absence of any gradient. From (1) $p(\mathbf{r}, t)$ can be estimated by taking the inverse FT of $s(\mathbf{q})/s_0$. But this requires a large number of measurements of $s(\mathbf{q})$ over a wide range of $\mathbf{q}$ in order to perform a stable inverse FT. Recently, Tuch et al. [7] developed q-space imaging method to obtain high angular resolution diffusion (HARD) measurements. In [8] Wedeen et al. succeed in acquiring 512 measurements of $s(\mathbf{q})$ in each scan to perform a stable inverse FT. However, most common way to estimate $p(\mathbf{r}, t)$ is assuming it to be a Gaussian or a mixture of Gaussians.

One of the alternatives for character diffusion anisotropy is using the information of apparent diffusion coefficient (ADC) profiles. The ADC on each voxel in DWI is defined as a function $d(\theta, \phi)$ in the Stejskal-tanner equation:

$$s(\mathbf{q}) = s_0 e^{-bd(\theta,\phi)}, \tag{2}$$

i.e.

$$d(\theta, \phi) = -\frac{1}{b} log \frac{s(\mathbf{q})}{s_0}. \tag{3}$$

where (θ, ϕ) $(0 \le \theta < \pi,\ 0 \le \phi < 2\pi)$ represents the direction of $\mathbf{q}$ in spherical coordinates, the b-factor is defined as $b = 4\pi^2|\mathbf{q}|^2(t - \delta/3)$.

For Gaussian diffusion,

$$p(\mathbf{r}, t) = \frac{1}{\sqrt{(4\pi t)^3 |D|}} e^{\frac{-\mathbf{r}^T D^{-1}\mathbf{r}}{4t}},$$

where D is called the diffusion tensor. Inserting this to equation (1) it yields

$$s(\mathbf{q}) = s_0 e^{-b\mathbf{u}^T D\mathbf{u}}, \tag{4}$$

where $\mathbf{u} = \mathbf{q}/|\mathbf{q}|$. In this case

$$d(\mathbf{u}) = \mathbf{u}^T D\mathbf{u}.$$

The principle eigenvector of D indicates direction of the diffusion. The fractional anisotropy (FA) defined as

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^+(\lambda_3-\lambda_1)^2}{(\lambda_1+\lambda_2+\lambda_3)^2}}, \tag{5}$$

where λ_i $(i = 1, 2, 3)$ are the eigenvalues of D, has become the most widely used measure of diffusion anisotropy [4]. This is known as diffusion tensor imaging (DTI), and in particular useful for creating white matter fiber tracts ([9, 6, 10, 11]). However, it has been recognized that the single tensor model (4) is inappropriate for assessing multiple fiber tract orientations, when complex tissue structure is found within a voxel ([6, 12, 13, 14, 7, 8]).

One of the promising approaches for characterizing anisotropy of non-Gaussian diffusion is using SHS approximation of the ADC profiles estimated from HARD data. This technique was initiated by Frank [14], also studied by Alexander et al. [15], Chen et al. [16] and others. The basic idea of these work is to approximate a ADC profile $d(x, \theta, \phi)$ at each voxel x by its truncated SHS:

$$d(x,\theta,\phi) = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} A_{l,m}(x) Y_{l,m}(\theta,\phi),$$

where $Y_{l,m}(\theta, \phi) : S^2 \to \mathbb{C}$ are the spherical harmonics and $\mathbb{C}$ denotes the set of complex numbers. The odd-order terms in the SHS are set to be zero, since the HARD measurements are made by a series of 3-d rotation, $d(\theta, \phi)$ is real and antipodal symmetry.

The coefficients $A_{l,m}(x)$'s were estimated from HARD data by different methods, and used to characterize the diffusion anisotropy. In [14] the coefficients $A_{l,m}$ at each voxel were estimated from the measured HARD (log) signal acquired with $b = 3000 s/mm^2$ through the equation (3), i.e.

$$A_{l,m}(x) = \int_0^{2\pi}\int_0^{\pi} -\frac{1}{b} log\frac{s(\mathbf{q})}{s_0} Y^*_{l,m}(\theta,\phi) sin\theta d\theta d\phi, \tag{6}$$

where * denotes the complex conjugate. In [15] $A_{l,m}$ were estimated as the least-squares solutions of

$$-\frac{1}{b} log\frac{s(x,\theta,\phi)}{s_0(x)} = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} A_{l,m}(\mathbf{x}) Y_{l,m}(\theta,\phi). \tag{7}$$

Then, the voxels with the significant 4th order (l=4) components in SHS are classified as anisotropic with multi-fiber orientations (shortened as multi-fibers), while voxels with the significant 2nd order (l=2) but not the 4th order components are classified as anisotropic with single fiber orientation (shortened as one-fiber), which is equivalent to the DTI model (4). Voxels with the significant 0th order ($l = 0$) but not the 2nd and 4th order components are classified as isotropic. In [16] we presented a variational model with a non-negativity constraint for simultaneous smoothing and estimation of the ADC profile $d(x, \theta, \phi)$ from the noisy HARD measurements $s(\mathbf{q})$. This model minimizes an energy that includes a smoothing term and a data fidelity term.

In this paper we will present a slightly modified version of the model presented in [16], give a mathematical justification of the model: proof of the existence of a solution in the space of functions with bounded variation (BV). We will also discuss

some implementation issues on saving computational cost, since finding a model solution involves solving a large system. Moreover, we will give our algorithm for characterizing diffusion anisotropy, which uses not only the information from the $A_{l,m}(x)$'s, but also the variation of $d(\theta,\phi)$ about its mean at each voxel. Our experimental results showed the effectiveness of the proposed model in the estimation and enhancement of anisotropy of the ADC profile. The characterization of the diffusion anisotropy based on the reconstructed ADC profiles using the proposed method is consistent with the known fiber anatomy.

2. **Model description.** In this section we will present a variational framework for simultaneous smoothing and estimation of the ADC profile $d(x,\theta,\phi)$ from the noisy HARD measurements $s(\mathbf{q})$. This model minimizes an energy functional consisting of two parts: one is the cost for regularizing d, and the other is the cost for fitting d to the HARD measurements through the original Stejskal-Tanner equation (2). To explain the basic idea of our method, we focus our attention on the cases where there are at most two fibers passing through a single voxel. The same idea can be applied to the cases where there are more fibers within a voxel.

The challenge in regularizing d comes from two aspects. First, d is defined on $\Omega\times S^2$ rather than $\Omega\times\mathbb{R}^2$, hence, the derivatives for (θ,ϕ) should be along the sphere. Secondly, the regularized d has to preserve the antipodal symmetry property with respect to (θ,ϕ). Considering these facts we adopt the idea developed in [14, 15] that approximates d by its SHS consisting of only even order components up to order 4 , i.e.

$$d(x,\theta,\phi)=\sum_{l=0,2,4}\sum_{m=-l}^{l}A_{l,m}(\mathbf{x})Y_{l,m}(\theta,\phi). \tag{8}$$

The expression in (8) ensures the smoothness and antipodal symmetry property of $d(\mathbf{x},\theta,\phi)$ in terms of (θ,ϕ), this is easy to see from the definition of $Y_{l,m}(\theta,\phi)$. For the cases where possibly k fibers cross in a single voxel, the sum in (8) should be replaced by $\sum_{l=0,2,\ldots,2k}$.

Now the problem of regularization and estimation of $d(x,\theta,\phi)$ reduces to that for the 15 complex valued functions $A_{l,m}(x)$ ($l=0,2,4$ and $m=-l,\ldots,l$) in (8). Since $d(\theta,\phi)$ at each voxel is a real valued function, and $Y_{l,m}$ satisfies $Y_{l,-m}=(-1)^mY^*_{l,m}$, $A_{l,m}$ should be constrained by

$$A_{l,-m}=(-1)^mA^*_{l,m}.$$

This constraint reduces the number of the unknown coefficients $A_{l,m}$ in (8) to 15 real valued functions. They are

$$A_{l,0}(x),\ (l=0,2,4),\quad ReA_{l,m}(x),\quad ImA_{l,m}(x),\quad (l=2,4\ and\ m=1,\ldots,l). \tag{9}$$

By using (9), we can rewrite (8) as

$$\begin{aligned} d(x,\theta,\phi)=&\sum_{l=0,2,4}A_{l,0}(\mathbf{x})Y_{l,0}(\theta,\phi)\\ &+2\sum_{l=2,4}\sum_{m=1}^{l}(ReA_{l,m}(x)ReY_{l,m}(\theta,\phi)-ImA_{l,m}(x)ImY_{l,m}(\theta,\phi)), \end{aligned} \tag{10}$$

where ReF and ImF represent the real and imaginary part of a function F respectively. Now the problem of regularizing and estimating d reduces to smoothing and estimation of 15 functions in (9) simultaneously.

There are many choices of regularizing operators to smooth the 15 functions in (9). Total Variation (TV) based regularization, first proposed by Rudin, Osher and Fatemi[17], proved to be an invaluable tool for feature preserving smoothing. However, it sometimes causes a staircase effect making restored image blocky, and even containing 'false edges' [18, 19]. An improvement, that combines the TV based smoothing with isotropic smoothing, was given by Chambolle and Lions [19]. Their model minimizes the TV norm when the magnitude of the image gradient is larger, and the L^2 norm of the image gradient if it is smaller. However, this model is sensitive to the choice of the threshold which separates the TV based and isotropic smoothing. To further improve Chambolle and Lions' model and make the model having an ability to self adjust diffusion property, recently, certain nonstandard diffusion models based on minimizing $L^{p(x)}$ norm of image gradient have been developed [18, 20]. To recover an image u from an observed image I in [18] the diffusion was governed by minimizing

$$\min_u \int_\Omega |\nabla u|^{p(|\nabla u|)}$$

where $p(s)$ is monotonically decreasing function and $\lim_{s\to 0} p(s) = 2$, $\lim_{s\to\infty} p(s) = 1$. In [20] the diffusion was performed through minimizing

$$\min_u \int_\Omega \phi(x, Du)$$

where $\phi(x,r) := \frac{1}{p(x)}|r|^{p(x)}$ if $|r| \le a$, and $\phi(x,r) := |r| - \frac{ap(x)-a^{p(x)}}{p(x)}$, if $|r| > a$ for threshhold a. In this model $p(x)$ is chosen as

$$p(x) = p(|\nabla I|) = 1 + \frac{1}{1 + k|\nabla G_\sigma * I|^2},$$

where $k,\ \sigma > 0$ are parameters, G_σ is the Gaussian kernel. Both models in [18] and [20] are able to self adjust diffusion range from isotropic to TV-based depending on image gradient. At the locations with higher image gradients ($p = 1$), the diffusion is TV based and strictly tangential to the edges ([17, 21, 19]). In homogeneous regions the image gradients are very small ($p = 2$), the diffusion is essentially isotropic. At all other locations, the image gradient forces $1 < p < 2$, and the diffusion is between isotropic and total variation based and varies depending on the local properties of the image. This self adjusting ability enables these models to effectively preserve features when images are smoothed.

Applying the idea of minimizing functionals with variable exponent to the problem of regularizing the coefficients $A_{l,m}$'s in (8) , we propose to minimize

$$E_1(A_{l,m}) := \int_\Omega \sum_{l=0,2,4} \sum_{m=-l}^{l} \phi_{l,m}(x, DA_{l,m}), \tag{11}$$

where

$$\phi_{l,m}(x,r) := \begin{cases} \frac{1}{p_{l,m}(x)} |\frac{r}{M_{l,m}}|^{p_{l,m}(x)}, & |r| \le M_{l,m} \\ |\frac{r}{M_{l,m}}| - (1 - \frac{1}{p_{l,m}(x)}), & |r| > M_{l,m} \end{cases} \tag{12}$$

We would like to point out that for a function $A_{l,m} \in BV$, $DA_{l,m}$ is a measure, the definition of (11) is not obvious. This will be discussed in existence section 6 below.

In (12)

$$p_{l,m} = 1 + \frac{1}{1 + k|\nabla G_\sigma * a_{l,m}(x)|^2}, \tag{13}$$

and $a_{l,m}$ is the least-squares solution of

$$-\frac{1}{b} log \frac{s(x,\theta,\phi)}{s_0(x)} = \sum_{l=0,2,4} \sum_{m=-l}^{l} a_{l,m}(x) Y_{l,m}(\theta,\phi). \tag{14}$$

In real application, $M_{l,m}$ is picked as 90th percentile of r. Experimental results are not sensitive to this parameter.

Again we would like to point out that E_1 only needs to include 15 terms corresponding to the 15 real valued functions in (9). Here we write it in terms of $A_{l,m}$ in order to shorten the expression of the formula. If using $A_{l,m}$ instead of $a_{l,m}$ in (13) we may get better numerical results, since $p_{l,m}$ would depend on updated $A_{l,m}$ rather than the fixed $a_{l,m}$ in the iterations to minimize (11). However, it gives difficulty in the study of the existence of solutions.

Since $d(x,\theta,\phi)$ is related to the HARD measurements $s(x,\theta,\phi)$ and $s_0(x)$ through the Stejskal-Tanner equation (2), the estimation of the $A_{l,m}$'s is based on the original Stejskal-Tanner equation (2) rather than its (log) linearized form, that is,

$$E_2(A_{l,m}) := \frac{1}{2} \int_\Omega \int_0^{2\pi} \int_0^{\pi} |s(\mathbf{x},\theta,\phi) - s_0(x) e^{-bd(x,\theta,\phi)}|^2 sin\theta d\theta d\phi dx, \tag{15}$$

where d is determined in (8). As observed in [22] when the signal to noise ratio is low the linearized model gives different results.

Finally, to simultaneously regularize and estimate the ADC $d(x,\theta,\phi)$, our model minimizes the energy function

$$E(A_{l,m}) := \lambda E_1(A_{l,m}) + E_2(A_{l,m}), \tag{16}$$

with respect to $A_{l,m}$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) in the space of $BV(\Omega)$, (in fact, only 15 functions in (9) are needed), and subject to the constraint:

$$d(x,\theta,\phi) \geq 0. \tag{17}$$

In (16),(17), $s(x,\theta,\phi)$ and $s_0(\mathbf{x})$ are the noisy HARD measurements (real valued), $d(x,\theta,\phi)$ is the SHS given in (8), $\Omega \subset \mathbb{R}^3$ is the image domain, $\lambda > 0$ is a parameter which could be different for different A_{lm}. E_1 and E_2 are given in (11) and (15), respectively.

Before we derive the Euler-Lagrange equations for our model (16),(17), we would like to point out that if the measurements satisfy the condition $s(x,\theta,\phi) \leq s_0(x)$, the solution of (16) meets the constraint (17) automatically. Therefore we can treat our model as an unconstrained minimization. This is given in the following lemma.

Lemma: *Under the assumption that*

$$s(x,\theta,\phi) \leq s_0(x), \quad for\ all\ \ \mathbf{x} \in \Omega,\ 0 \leq \theta < \pi,\ 0 \leq \phi < 2\pi \tag{18}$$

the minimizer of (16) always satisfies the constraint (17).

Proof. Let $A_{l,m}(x)$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) be the minimizer of (16) in $BV(\Omega)$, and $d(x,\theta,\phi)$ be the function defined in (8) associated with these optimal

$A_{l,m}(x)$'s. Given $x \in \Omega$, if $d(\mathbf{x},\theta,\phi) < 0$ for some $0 \le \theta < \pi,\ 0 \le \phi < 2\pi$, then define $\hat{d}(x,\theta,\phi) = 0$, otherwise, define $\hat{d}(x,\theta,\phi) := d(x,\theta,\phi)$.

Correspondingly,

$$\hat{A}_{l,m}(x) := \int_0^{2\pi}\int_0^{\pi} \hat{d}(x,\theta,\phi)Y_{l,m}(\theta,\phi)sin\theta d\theta d\phi.$$

Then, using the orthonormality of the spherical harmonics and the definition of d, we have

$$\hat{A}_{l,m}(x) = \begin{cases} A_{l,m}(x), & if\ d(x,\theta,\phi) \ge 0, \forall\ 0 \le \theta < \pi,\ 0 \le \phi < 2\pi \\ 0, & if\ d(x,\theta,\phi) < 0 \end{cases} \tag{19}$$

This implies that

$$\phi_{l,m}(x, D\hat{A}_{l,m}) \le \phi_{l,m}(x, DA_{l,m}),$$

hence

$$E_1(\hat{A}_{l,m}) \le E_1(A_{l,m}).$$

Moreover, it is easy to obtain

$$E_2(\hat{d}) \le E_2(d),$$

if (18) holds. From the last two inequality above, we obtain that $E(\hat{d}) \le E(d)$. This contradicts to the fact that d minimizes energy functional (16).

Now we give the evolution equations associated with the Euler-Lagrange (EL) equations for (16): for $l = 0, 2, 4$ and $m = -l, \ldots, l$,

$$\frac{\partial A_{l,m}}{\partial t} = \lambda div(\partial_r \phi_{l,m})(x, DA_{l,m}) - b\int_0^{2\pi}\int_0^{\pi} s_0 e^{-bd}(s - s_0 e^{-bd})Y_{l,m} sin\theta d\theta d\phi, \tag{20}$$

with the initial and boundary conditions:

$$A_{l,m} = a_{l,m}, \qquad on\ \ \Omega \times \{t = 0\},$$

$$(\partial_r \phi_{l,m})(x, DA_{l,m}) \cdot n = 0 \qquad on\ \ \partial\Omega \times \Re^+.$$

In the above EL equation n is the unit outward normal to the boundary of Ω, and $\partial_r \phi_{l,m}(x,r)$ is a continuously differentiable function in r, and

$$\partial_r \phi_{l,m}(x,r) := \begin{cases} \frac{1}{M_{l,m}^p(x)}|r|^{p(x)-2}r, & |r| \le M_{l,m} \\ \frac{1}{M_{l,m}}|r|^{-1}r, & |r| > M_{l,m} \end{cases} \tag{21}$$

$\partial_r \phi_{l,m}$ can also be written as

$$\partial_r \phi_{l,m}(x,r) := \frac{1}{M_{l,m}^q(x)}|r|^{q(x)-2}r, \tag{22}$$

where $q(x) = p(x)$ if $|r| \le M_{l,m}$, and $q(x) = 1$ if $|r| > M_{l,m}$. □

3. **Characterization of anisotropy.** In [14] the $|A_{l,m}(x)|$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) in the truncated SHS (8) are used to characterize the diffusion anisotropy at each voxel x. Our experimental results, however, indicate this information alone is insufficient to separate isotropic diffusion, one-fiber diffusion, and multi-fiber diffusion within a voxel. We propose to combine the information from $|A_{l,m}|$ with the variances of $d(\phi,\theta)$ about its mean value to characterize the diffusion anisotropy. We outline our algorithm as follows:

(1). If

$$R_0 := \frac{|A_{0,0}|}{\sum_{l=0,2,4}\sum_{m=-l}^{l}|A_{l,m}|}, \tag{23}$$

is large, or the variance of $d(\theta,\phi)$ about its mean is small, the diffusion at such voxels is classified as isotropic.

(2). For the remaining voxels, if

$$R_2 := \frac{\sum_{m=-2}^{m=2}|A_{2,m}|}{\sum_{l=2,4}\sum_{m=-l}^{l}|A_{l,m}|} \tag{24}$$

is large, the diffusion at such voxels is characterized as one-fiber diffusion. Figure 3 D) presents an intensity-coded image of R_2 in a brain slice through the external capsule, an important structure of the human white matter. In Figure 3D) those voxels of a high intensity (bright regions on the image) are characterized as one-fiber diffusion.

(3). For each uncharacterized voxel after the above two steps, search the directions (θ,ϕ), where $d(\theta,\phi)$ attains its local maxima. Note, $d(\theta,\phi)$ is antipodal symmetric, i.e., $d(\theta,\phi) = d(\pi-\theta,\phi+\pi)$, we mod out this symmetry when count the number of local maxima. Then we compute the weights for the local maxima (say we have 3 local maxima):

$$W_i := \frac{d(\theta_i,\phi_i) - d_{min}}{\sum_{i=1}^{3} d(\theta_i,\phi_i) - 3d_{min}},$$

where (θ_i,ϕ_i) $(i=1,2,3)$ are the directions in which d attains its local maxima. If one of the weights is significant, it is considered as one fiber diffusion. If two weights are similar but much larger than the third one, it is viewed as two-fiber diffusion, if all three weights are similar, d can be considered either three-fiber diffusion or isotropic diffusion. In our experiment we restrict ourselves to the cases where we only distinguish isotropic, one-fiber or two-fiber diffusions. Under this restriction if three weights are similar, we include this voxel in the class of isotropic diffusion. Figure 5A) shows our classification of isotropic diffusion (dark region), one-fiber diffusion (gray region), and two-fiber diffusion (bright region) in the same slice as in Figure 3.

4. **Numerical implementation issues.** To efficiently solve the Euler-Lagrange equations (20), we use Additive Operator Splitting(AOS) algorithm for the diffusion operator (see [23, 24]). By using this algorithm, the computational and storage costs are linear in the number of voxels, and the computational efficiency can be increased by a factor of 10 under realistic accuracy requirements([23]). The algorithm is ready to be modified to a parallel version.

To avoid the complicated notation, we use X to represent any $A_{l,m}$ in the Euler-Lagrange equations, and write the algorithm for only one of the equations (20) in the system, since each equation has the same structure as others.

We use semi-implicit finite difference scheme:

$$\frac{X_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) + \lambda div\left(\frac{\nabla X_{i,j}^{(n+1)}}{M^{q_{ij}}|\nabla X_{i,j}^{(n)}|^{2-q_{ij}}}\right)$$

$$= f(X_{i,j}^{(n)}) + \frac{-\lambda ln M}{M^{q_{ij}}} \frac{\nabla q_{ij} \cdot \nabla X_{i,j}^{(n+1)}}{|\nabla X_{i,j}^{(n)}|^{2-q_{ij}}} + \frac{\lambda}{M^{q_{ij}}} div\left(\frac{\nabla X_{i,j}^{(n+1)}}{|\nabla X_{i,j}^{(n)}|^{2-q_{ij}}}\right) \tag{25}$$

Here X can be replaced by one of $A_{l,m}$'s with $l = 0, 2, 4,\ m = -l, \cdots, l$, and f is a function of results from last iteration,namely, f is a function of all $A_{l,m}^{(n)}$'s. $q(x) = p(x)$ if $|\nabla X| \leq M$, and $q(x) = 1$ if $|\nabla X| > M$ for some fixed constant M, which was chosen based on initial value of X, so M might be different for different $A_{l,m}$'s.

For simplicity of formulas, we define:

$$\triangle_-^x X_{i,j} = X_{i,j} - X_{i-1,j},\ \triangle_+^x X_{i,j} = X_{i+1,j} - X_{i,j},\ \triangle^x X_{i,j} = X_{i+1,j} - X_{i-1,j}$$

$$\triangle_+^y X_{i,j} = X_{i,j+1} - X_{i,j},\ \triangle_-^y X_{i,j} = X_{i,j} - X_{i,j-1},\ \triangle^y X_{i,j} = X_{i,j+1} - X_{i,j-1}$$

Adopting a discretization of the divergence operator from [25], one can write (25) as:

$$\frac{X_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) - \frac{\lambda ln M}{M^{q_{ij}}} \frac{[\triangle^x q_{ij},\ \triangle^y q_{ij}]}{2h} \cdot \frac{\left[\triangle^x X_{i,j}^{(n+1)}, \triangle^y X_{i,j}^{(n+1)}\right]/(2h)}{\left(\frac{(\triangle^x X_{i,j}^{(n)})^2}{(2h)^2} + \frac{(\triangle^y X_{i,j}^{(n)})^2}{(2h)^2}\right)^{\frac{2-q_{ij}}{2}}}$$

$$+ \frac{\lambda}{M^{q_{ij}} h^2} \cdot$$

$$\left[\triangle_-^x \left(\frac{\triangle_+^x X_{i,j}^{(n+1)}}{\left(\frac{(\triangle_+^x X_{i,j}^{(n)})^2}{h^2} + \frac{(\triangle^y X_{i,j}^{(n)})^2}{(2h)^2}\right)^{\frac{2-q_{ij}}{2}}}\right) + \triangle_-^y \left(\frac{\triangle_+^y X_{i,j}^{(n+1)}}{\left(\frac{(\triangle_+^y X_{i,j}^{(n)})^2}{h^2} + \frac{(\triangle^x X_{i,j}^{(n)})^2}{(2h)^2}\right)^{\frac{2-q_{ij}}{2}}}\right)\right]$$

$$= f(X_{i,j}^{(n)}) + (C_{i,j} - G_{i,j}) X_{i-1,j}^{(n+1)} - (C_{i,j} + D_{i,j}) X_{i,j}^{(n+1)} + (D_{i,j} + G_{i,j}) X_{i+1,j}^{(n+1)} +$$

$$+ \ (E_{i,j} - H_{i,j}) X_{i,j-1}^{(n+1)} - (E_{i,j} + F_{i,j}) X_{i,j}^{(n+1)} + (F_{i,j} + H_{i,j}) X_{i,j+1}^{(n+1)} \tag{26}$$

Where C, D, E and F are from divergence operation, while G and H are generated by dot product, in detail:

$$C_{i,j} = \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i,j}^{(n)} - X_{i-1,j}^{(n)})^2}{h^2} + \frac{(X_{i-1,j+1}^{(n)} - X_{i-1,j-1}^{(n)})^2}{(2h)^2}\right]^{\frac{q_{i-1,j}-2}{2}}$$

$$D_{i,j} = \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i,j}^{(n)})^2}{h^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2}\right]^{\frac{q_{ij}-2}{2}}$$

$$E_{i,j} = \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j-1}^{(n)} - X_{i-1,j-1}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j}^{(n)} - X_{i,j-1}^{(n)})^2}{(h)^2}\right]^{\frac{q_{i,j-1}-2}{2}}$$

$$F_{i,j} = \frac{\lambda}{M^{q_{ij}} h^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j}^{(n)})^2}{(h)^2}\right]^{\frac{q_{ij}-2}{2}}$$

$$G_{i,j} = -\frac{\lambda\, ln M (q_{i+1,j} - q_{i-1,j})}{M^{q_{ij}} (2h)^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2}\right]^{\frac{q_{ij}-2}{2}}$$

$$H_{i,j} = -\frac{\lambda\, ln M (q_{i,j+1} - q_{i,j-1})}{(2h)^2} \left[\frac{(X_{i+1,j}^{(n)} - X_{i-1,j}^{(n)})^2}{(2h)^2} + \frac{(X_{i,j+1}^{(n)} - X_{i,j-1}^{(n)})^2}{(2h)^2}\right]^{\frac{q_{ij}-2}{2}}$$

Solving (26) would involve matrix inverse operation, which would become more and more complicated and dramatically expensive as dimension increases if we solve

it directly. Instead, here we use Additive Operator Splitting (AOS) algorithm, which allows us to reformat system (26) into following system:

$$\frac{\bar{X}_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) +$$

$$2\left[(C_{i,j} - G_{i,j})\bar{X}_{i-1,j}^{(n+1)} - (C_{i,j} + D_{i,j})\bar{X}_{i,j}^{(n+1)} + (D_{i,j} + G_{i,j})\bar{X}_{i+1,j}^{(n+1)}\right] \tag{27}$$

$$\frac{\bar{\bar{X}}_{i,j}^{(n+1)} - X_{i,j}^{(n)}}{\tau} = f(X_{i,j}^{(n)}) +$$

$$2\left[(E_{i,j} - H_{i,j})\bar{\bar{X}}_{i,j-1}^{(n+1)} - (E_{i,j} + F_{i,j})\bar{\bar{X}}_{i,j}^{(n+1)} + (F_{i,j} + H_{i,j})\bar{\bar{X}}_{i,j+1}^{(n+1)}\right] \tag{28}$$

and

$$X_{i,j}^{(n+1)} = \frac{\bar{X}_{i,j}^{(n+1)} + \bar{\bar{X}}_{i,j}^{(n+1)}}{2}$$

To accommodate the boundary condition $\frac{\partial X}{\partial n} = 0$ for the $M \times N$ matrix X, one needs to have:

$$X_{1,j}^{(n+1)} = X_{2,j}^{(n+1)}, X_{M-1,j}^{(n+1)} = X_{M,j}^{(n+1)}$$
$$X_{i,1}^{(n+1)} = X_{i,2}^{(n+1)}, X_{i,N-1}^{(n+1)} = X_{i,N}^{(n+1)}$$

Then (27) and (28) correspond to linear systems in matrix-vector notation:

$$A_1\underline{\bar{X}}^{(n+1)} = \tau f(\underline{\bar{X}}^{(n)}) + \underline{\bar{X}}^{(n)}$$
$$A_2\underline{\bar{\bar{X}}}^{(n+1)} = \tau f(\underline{\bar{\bar{X}}}^{(n)}) + \underline{\bar{\bar{X}}}^{(n)}$$

where $\underline{\bar{X}}$ and $\underline{\bar{\bar{X}}}$ are $(M-2)(N-2) \times 1$ vectors formed by columns and transpose of rows of the original matrix X respectively, both A_1 and A_2 are $(M-2)(N-2) \times (M-2)(N-2)$ matrices, specifically, A_1 is a tri-diagonal matrix that repeats a $(M-2) \times (M-2)$ tri-diagonal matrix $(N-2)^2$ times diagonally, and A_2 is a tri-diagonal matrix that repeats a $(N-2) \times (N-2)$ tri-diagonal matrix $(M-2)^2$ times. They are defined as:

$A_1 = I - 2\tau\cdot$

$$\begin{bmatrix} -D_{2,2} - G_{2,2} & D_{2,2} + G_{2,2} & 0 & \cdot & \cdot \\ C_{3,2} - G_{3,2} & -C_{3,2} - D_{3,2} & D_{3,2} + G_{3,2} & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & D_{M-2,N-1} + G_{M-2,N-1} \\ \cdot & \cdot & 0 & C_{M-1,N-1} - G_{M-1,N-1} & -C_{M-1,N-1} + G_{M-1,N-1} \end{bmatrix}$$

$A_2 = I - 2\tau\cdot$

$$\begin{bmatrix} -F_{2,2} - H_{2,2} & F_{2,2} + H_{2,2} & 0 & \cdot & \cdot \\ E_{2,3} - H_{2,3} & -E_{2,3} - F_{2,3} & F_{2,3} + H_{2,3} & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & F_{M-1,N-2} + H_{M-1,N-2} \\ \cdot & \cdot & 0 & E_{M-1,N-1} - H_{M-1,N-1} & -E_{M-1,N-1} + H_{M-1,N-1} \end{bmatrix}$$

$$\underline{\bar{X}} = [X_{2,2}\ X_{3,2} \cdots X_{M-1,2}\ X_{2,3} \cdots X_{M-2,N-1}\ X_{M-1,N-1}]^T$$
$$\underline{\bar{\bar{X}}} = [X_{2,2}\ X_{2,3} \cdots X_{2,N-1}\ X_{3,2} \cdots X_{M-1,N-2}\ X_{M-1,N-1}]^T$$

Since both A_1 and A_2 are tri-diagonal matrices, one can get their inverses efficiently by using Thomas Algorithm([26]).

5. **Validation and application to diffusion weighted images.** In this section we present our experimental results on the application of the proposed model (16)-(17) to simulated data and a set of HARD MRI data from the human brain.

5.1. **Analysis of simulated data.** The aim of our experiment on the simulated data is to test whether our model can efficiently reconstruct a regularized ADC profile from the noisy HARD measurements. We simulated an ADC profile on a 2D lattice of size 8×4. The volume consists of two homogeneous regions, values of S_0 and all the $A_{l,m}$'s were shown in table 1.

In Figure 1 we displayed the true, noisy, and recovered ADC profiles $d(x, \theta, \phi)$ for the synthetic data with size 8×4. The ADC profile $d(x, \theta, \phi)$ was computed by (10) based on these simulated data, and the corresponding $s_{true}(x, \theta, \phi)$ was constructed via (2) with $b = 1000 s/mm^2$. Then the noisy HARD MRI signal $s(x, \theta, \phi)$ was generated by adding a zero mean Gaussian noise with standard deviation $\sigma = 0.15$. Figure 1B) shows the ADC profile d computed by (10), where the coefficients of the SHS are the least-squares solutions of (7) with noisy s.

Table 1: List of S_0 and $A_{l,m}$'s for two regions

Region	1	2
S_0	414	547
$A_{0,0}$	5.21×10^{-3}	1.43×10^{-2}
$A_{2,0}$	-1.17×10^{-3}	-2.68×10^{-3}
$ReA_{2,1}$	-4.37×10^{-5}	0
$ReA_{2,2}$	1.43×10^{-3}	0
$ImA_{2,1}$	3.64×10^{-5}	0
$ImA_{2,2}$	3.28×10^{-5}	0
$A_{4,0}$	-3.15×10^{-5}	8.4×10^{-6}
$ReA_{4,1}$	-1.56×10^{-4}	0
$ReA_{4,2}$	1.02×10^{-4}	0
$ReA_{4,3}$	6.30×10^{-5}	0
$ReA_{4,4}$	-8.54×10^{-5}	-1.73×10^{-3}
$ImA_{4,1}$	-8.01×10^{-5}	0
$ImA_{4,2}$	$0.9961.55 \times 10^{-4}$	0
$ImA_{4,3}$	1.41×10^{-5}	0
$ImA_{4,4}$	3.63×10^{-5}	0

We then applied our model (16)-(17) to the noisy $s(x, \theta, \phi)$ to test the effectiveness of the model, with $\lambda_{0,0} = 4, \lambda_{2,m} = 40 (m = -2, \ldots, 2), \lambda_{4,m} = 60 (m = -4, \ldots, 4)$. By solving the system of equations (20) in 2.5 seconds on computer with PIV 2.8GHZ CPU and 2G RAM using Matlab script code, we obtained 15 reconstructed functions as in (9). Using these $A_{l,m}$ (the solutions of (20)) we computed $d(x, \theta, \phi)$ via (10). The reconstructed $d(x, \theta, \phi)$ is shown in Figure 1C). Comparing these three figures, it is clear that the noisy measurements s have changed Figure 1A), the original shapes of d, into Figure 1B). After applying our model (16)-(17) to

reconstruct the ADC profiles, the shapes of d in Figure 1A) were recovered, as shown in Figure 1C). These simulated results demonstrate that our model is effective in simultaneously regularizing and recovering ADC profiles.

5.2. **Analysis of human MRI data.** The second test is to reconstruct and characterize ADC profiles $d(x, \theta, \phi)$ from human HARD MRI data.

The raw DWI data, usually contains a certain level of noise, were obtained on a GE 3.0 Tesla scanner using a single shot spin-echo EPI sequence. The scanning parameters for the DWI acquisition are: repetition time (TR)=1000ms, echo time (TE) =85ms, the field of view (FOV)=220 mm x 220 mm. 24 axial sections covering the entire brain with the slice thickness=3.8 mm and the intersection gap=1.2 mm. The diffusion-sensitizing gradient encoding is applied in fifty-five directions (selected for the HARD MRI acquisition) with $b = 1000s/mm^2$. Thus, a total of fifty-six diffusion-weighted images, with a matrix size of 256 x 256, were obtained for each slice section. We applied model (16) to these data to compute the ADC profiles in the entire brain volume. By solving a system of equations (20) we obtained all the coefficients $A_{l,m}$'s in (9), and determined $d(x, \theta, \phi)$ using (10).

Then, we used these $A_{l,m}(x)$ to calculate R_0 and R_2 defined in (23) and (24) respectively, as well as the variance $\sigma(x)$ of $d(x, \theta, \phi)$ about its mean: $\sigma(x) = \int_0^{\pi}\int_0^{2\pi}(d(x, \theta, \phi) - \sum_{i=1}^{55} d(\mathbf{x}, \theta_i, \phi_i)/55)^2 d\theta d\phi$. Based on results from the HARD MRI data of this particular patient, we characterized the diffusion anisotropy according to the following procedure. If $R_0(x) > 0.856$, or $\sigma(\mathbf{x}) < 19.65$ the diffusion at x is classified as isotropic. For the remaining voxels if $R_2(x) > 0.75$, the diffusion at such voxels is considered as one-fiber diffusion. For uncharaterized voxels from these two steps we further classified them by the principles stated in the section 3. The selection of the thresholds mentioned above for R_0, R_2 and σ involves experts' input and large sample experiments. Experimental results definitely depend on these thresholds, but not sensitively.

Figure 2 presents $A_{2,0}(x)$, one of the coefficients in (10), for the particular slice in the volume. The images $A_{2,0}(x)$ in Figure 2A) and 2B) are estimated by using (14) and solving (16), respectively.

Figure 3 compares FA and three $R_2(x)$'s with $A_{l,m}(\mathbf{x})$'s obtained from three different models for the same slice as shown in Figure2. Figure 3A) displays the FA image obtained by using advanced system software from GE. The $A_{l,m}(\mathbf{x})$'s used to obtain $R_2(\mathbf{x})$ in Figure 3B) are directly computed from (14). Those used to obtain $R_2(\mathbf{x})$ in Figures. 3C) and 3D) are the least-squares solutions of (7) and the solutions of (16), respectively. In Figures. 3C) and 3D) the voxels with high levels of intensities (red, yellow, yellow-light blue) are characterized as one-fiber diffusion.

Although the FA image in Figure 3A) is obtained based on a conventional DTI model (4), it is still comparable with the R_2 map, since single tensor diffusion characterized by SHS representation from the HARD images agrees with that characterized by the DTI model. However, in DTI a voxel with a low intensity of FA indicates isotropic diffusion, while using our algorithm, multi-fibers diffusion may occur at the location with the low value of R_2.

The ability to characterize anisotropic diffusion is enhanced by this algorithm, as shown in Figures. 3A)-3D). Figure 3B) indicates again that the estimates of $A_{l,m}$ directly from the *log* signals usually are not good. Even the least-squares solution of (7) is not always effective. This can be seen by comparing the anatomic region inside the red square of Figures. 3C) and 3D), which are zoomed in Figures.

4A) and 4B), respectively. There is a dark broken line showing on the map of the external capsule (arrow to the right on Figure 4A), this same region was recovered by the proposed model and characterized by the third step in our algorithm as two-fiber anisotropic diffusion (arrow to the right in Figure 4B). (The model solutions reduced the value of R_0, increased the values of R_1 slightly, and made the 3rd step in our characterization to be applied). Our results also show the connection in a cortical associative tract (arrow to the left in Figures. 4B), however, this connection was not mapped out on Figure 3C) or the zoomed image in Figure 4A). In fact this connection was not mapped out on Figures 3A)-B) either. All these mapped connections are consistent with known neuroanatomy. Combined together, our results indicate that our proposed model for joint recovery and smoothing of the ADC profiles has an advantage over existing models for enhancing the ability to characterize diffusion anisotropy.

Figure 5A) shows a partition of isotropic, one-fiber, and two-fiber diffusion for the same slice used in Figure 4. The two-fiber, one-fiber, and isotropic diffusion regions were further characterized by the white, gray, and black regions, respectively. The region inside the white square in Figure 5A), which is the same one squared in Figures. 3C) and 3D), is zoomed in Figure 4C). It is seen that the two arrayed voxels in Figure 4B) are classified as two-fiber diffusion. The characterization of the anisotropy on the voxels and their neighborhoods is consistent with the known fiber anatomy.

Figure 5 B) represents the shapes of $d(x, \theta, \phi)$ at three particular voxels (upper, middle and lower rows). The d in all three voxels is computed using (10). However, the $A_{l,m}(\mathbf{x})$ used in computing d on the left column are the least-squares solutions of (7), while in the right column they are the solutions of the proposed model (16). The first and second rows show two voxels that can be characterized as isotropic diffusion before denoising, but as two-fiber diffusion after applying model (16). These two voxels are the same voxels as in Figure 4 directed by arrows. The lower row of Figure 5 B) shows the one-fiber diffusion was enhanced after applying our model.

Solving $A_{l,m}$'s of size $15 \times 109 \times 86 \times 8$ from 4-D data of size $55 \times 109 \times 86 \times 8$ takes 46.2 seconds for each iteration on computer with PIV 2.8GHZ CPU and 2G RAM in Matlab script code.

6. **An existence theorem for the model.** In this section we will discuss the existence of a solution to our minimization problem (16) using the idea developed in [20].

Recall that for a function $u \in BV(\Omega)$,

$$Du = \nabla u \cdot \mathcal{L}^n + D^s u$$

is a Radon measure, where ∇u is the density of the absolutely continuous part of Du with respect to the n-dimensional Lebesgue measure $\mathcal{L}^n$, and $D^s u$ is the singular part. To minimize (16) over the functions in $BV(\Omega)$, we first need to give a precise definition for E_1.

Definition: For $A_{l,m} \in BV(\Omega)$, define

$$\int_\Omega \phi_{l,m}(x, DA_{l,m}) := \int_\Omega \phi_{l,m}(x, \nabla A_{l,m})dx + \int_\Omega |D^s A_{l,m}|$$

where $\phi_{l,m}$ is defined as in (12), and $\int_\Omega |D^s A_{l,m}|$ is the total variation norm of $A_{l,m}$.

Then, energy functional (16) is defined as

$$E(A_{l,m}) = \lambda \int_\Omega \sum_{l=0,2,4} \sum_{m=-l}^{l} \phi_{l,m}(x, \nabla A_{l,m}) + \lambda \int_\Omega |D^s A_{l,m}|$$

$$+ \frac{1}{2} \int_\Omega \int_0^{2\pi} \int_0^{\pi} |s(x,\theta,\phi) - s_0(x) e^{-bd(x,\theta,\phi)}|^2 sin\theta d\theta d\phi dx. \tag{29}$$

In the discussion of existence, without loss of generality, we set the parameter $\lambda = 1$ in (16) and threshold $M_{l,m} = 1$ in (12) to reduce the complexity in the formulation.

Next we will show lower semi-continuity of the energy functional (16) in L^1, i.e. if for each l, m ($l = 0, 2, 4$ and $m = -l, \ldots, l$), as $k \to \infty$,

$$A^k_{l,m} \to A^0_{l,m} \quad in \quad L^1(\Omega),$$

then

$$E(A^0_{l,m}) \leq liminf_{k\to\infty} E(A^k_{l,m}), \tag{30}$$

To prove this we need the following lemma:

Lemma: *Let*

$$\phi(x, r) := \begin{cases} \frac{1}{p(x)}|r|^{p(x)}, & |r| \leq 1 \\ |r| - (1 - \frac{1}{p(x)}), & |r| > 1 \end{cases} \tag{31}$$

For $u \in BV(\Omega)$ denote

$$\Phi(u) := \int_\Omega \phi(x, Du),$$

and

$$\widetilde{\Phi}(u) := \sup_{\substack{\psi \in C^1_0(\Omega, \mathbb{R}^n) \\ |\psi| \leq 1}} \int_\Omega -udiv\psi - \frac{p(x)-1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx.$$

Then,

$$\Phi(u) = \widetilde{\Phi}(u) \tag{32}$$

Furthermore, Φ is lower semi-continuous on $L^1(\Omega)$, i.e. if $u_j, u \in BV(\Omega)$ satisfy $u_j \to u$ weakly in $L^1(\Omega)$ as $j \to \infty$ then

$$\Phi(u) \leq \liminf_{j\to\infty} \Phi(u_j).$$

Proof. First note that for each $\psi \in C^1_0(\Omega, \mathbb{R}^n)$, the map

$$u \longmapsto \int_\Omega -udiv\psi - \frac{p(x)-1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx$$

is continuous and affine on $L^1(\Omega)$. Therefore, $\widetilde{\Phi}$ is convex and lower semi-continuous on $L^1(\Omega)$ and the domain of $\widetilde{\Phi}$, $\{u \,|\, \widetilde{\Phi}(u) < \infty\}$, is precisely $BV(\Omega)$.

Next we show (32). For $u \in BV(\Omega)$, we have that for each $\psi \in C^1_0(\Omega, \mathbb{R}^n)$,

$$-\int_\Omega udiv\psi dx = \int_\Omega \nabla u \cdot \psi dx + \int_\Omega D^s u \cdot \psi$$

and so

$$\widetilde{\Phi}(u) = \sup_{\substack{\psi \in C^1_0(\Omega, R^n) \\ |\psi| \leq 1}} \int_\Omega \nabla u \cdot \psi - \frac{p(x)-1}{p(x)} |\psi|^{\frac{p(x)}{p(x)-1}} dx + \int_\Omega D^s u \cdot \psi$$

Since the measures dx and $D^s u$ are mutually singular, by a standard argument we can have

$$\widetilde{\Phi}(u) = \sup_{\substack{\psi\in C_0^1(\Omega,R^n)\\ |\psi|\le 1}} \int_\Omega (\nabla u\cdot\psi - \frac{p(x)-1}{p(x)}|\psi|^{\frac{p(x)}{p(x)-1}})dx + \int_\Omega |D^s u|.$$

To prove (32) it only remains to show that

$$\int_\Omega \phi(x,\nabla u)dx = \sup_{\substack{\psi\in C_0^1(\Omega,R^n)\\ |\psi|\le 1}} \int_\Omega (\nabla u\cdot\psi - \frac{p(x)-1}{p(x)}|\psi|^{\frac{p(x)}{p(x)-1}})dx. \tag{33}$$

Since any $\rho\in L^\infty(\Omega,R^n)$ can be approximated in measure by $\psi\in C_0^1(\Omega,R^n)$, we have that

$$\begin{aligned}&\sup_{\substack{\psi\in C_0^1(\Omega,\mathbb{R}^n)\\ |\psi|\le 1}} \int_\Omega \nabla u\cdot\psi - \frac{p(x)-1}{p(x)}|\psi|^{\frac{p(x)}{p(x)-1}}dx\\ &= \sup_{\substack{\rho\in L^\infty(\Omega,R^n)\\ |\rho|\le 1}} \int_\Omega \nabla u\cdot\rho - \frac{p(x)-1}{p(x)}|\rho|^{\frac{p(x)}{p(x)-1}}dx.\end{aligned} \tag{34}$$

Choosing $\rho(x) = 1_{\{|\nabla u|\le 1\}}|\nabla u|^{p(x)-1}\frac{\nabla u}{|\nabla u|} + 1_{\{|\nabla u|>1\}}\frac{\nabla u}{|\nabla u|}$, where χ_E is the indicator function on E, we see that the right hand side of (34) is

$$\begin{aligned}&\ge \int_\Omega \frac{1}{p(x)}|\nabla u|^{p(x)}1_{\{|\nabla u|\le 1\}} + \left[|\nabla u| - \frac{p(x)-1}{p(x)}\right]1_{\{|\nabla u|>1\}}dx = \int_\Omega \phi(x,\nabla u)dx.\end{aligned} \tag{35}$$

To show the opposite inequality, we argue as follows. For any $\rho\in L^\infty(\Omega,R^n)$, since $p(x)>1$ we have that for almost all x,

$$\nabla u(x)\cdot\rho(x) \le \frac{1}{p(x)}|\nabla u|^{p(x)} + \frac{p(x)-1}{p(x)}|\rho(x)|^{\frac{p(x)}{p(x)-1}}$$

In particular, if $|\nabla u|\le 1$,

$$\nabla u(x)\cdot\rho(x) - \frac{p(x)-1}{p(x)}|\rho(x)|^{\frac{p(x)}{p(x)-1}} \le \frac{1}{p(x)}|\nabla u|^{p(x)} \tag{36}$$

If $|\nabla u|>1$, noticing $p(x)>1$ and $|\rho|\le 1$ for almost all x we have that

$$\nabla u\cdot\rho = |\nabla u|\frac{\nabla u}{|\nabla u|}\cdot\rho \le |\nabla u|\left[\frac{1}{p(x)} + \frac{p(x)-1}{p(x)}|\rho|^{\frac{p(x)}{p(x)-1}}\right]$$

and so

$$\nabla u\cdot\rho - \frac{p(x)-1}{p(x)}|\rho|^{\frac{p(x)}{p(x)-1}} \le \frac{1}{p(x)}|\nabla u| + (|\nabla u|-1)\frac{p(x)-1}{p(x)}|\rho|^{\frac{p(x)}{p(x)-1}} \le |\nabla u| - \frac{p(x)-1}{p(x)} \tag{37}$$

Combining, (34), (35), (36), and (37), we have (33), and hence for all $u\in BV(\Omega)$, $\widetilde{\Phi}(u) = \Phi(u)$.

Note that $\phi(x,r) = \phi_{l,m}(x,r)$ if $p(x) = p_{l,m}(x)$. A direct consequence of this lemma is that we have that E_1 in (11) is weakly lower semi-continuous in L^1 topology on $BV(\Omega)$ norm.

Furthermore, we can show that E_2 in (11) is lower semi-continuous on $L^1(\Omega)$. Indeed, when

$$A^k_{l,m} \to A^0_{l,m}, \quad in \quad L^1(\Omega), \quad as \quad k \to \infty,$$

for all $l = 0, 2, 4$ and $m = -l, \ldots, l$), $A^k_{l,m} \to A^0_{l,m}$ *a.e.* on Ω. Then, if $s(x, \theta, \phi) \in L^2(\Omega \times S^2)$ and $s_0(x) \in L^2(\Omega)$, by the Dominated Convergence Theorem we have

$$E_2(A^k_{l,m}) \to E_2(A^0_{l,m}).$$

Therefore, $E = E_1 + E_2$ is lower semi-continuous in L^1 topology on $BV(\Omega)$ norm, and (30) holds.

Now we can prove our existence results. □

Theorem: *Let Ω be a bounded open set of $\mathbb{R}^n$. Assume that $s(x, \theta, \phi) \in L^2(\Omega \times S^2)$ and $s_0(x) \in L^2(\Omega)$. Then, there exists a solution consisting of functions $A^0_{l,m}$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) to the minimization problem (16) over the space of $BV(\Omega)$.*

Proof. Let $A^k_{l,m}$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) be the minimizing sequences of (16) in $BV(\Omega)$. Then for each (l, m) the sequence $A^k_{l,m}$ is bounded in $BV(\Omega)$. From the compactness of $BV(\Omega)$ there exist subsequences of $A^k_{l,m}$ (still denoted by $A^k_{l,m}$) and functions $A^0_{l,m} \in BV(\Omega)$ satisfying

$$A^k_{l,m} \to A^0_{l,m} \quad strongly \quad in \quad L^1(\Omega).$$

By the lower semi-continuity of the energy functional on $L^1(\Omega)$ (see (30)), we have

$$E(A^0_{l,m}) \leq liminf_{k\to\infty} E(A^k_{l,m}) \leq inf_{A_{l,m} \in BV(\Omega)} E(A_{l,m}).$$

Hence, all these $A^0_{l,m}$ ($l = 0, 2, 4$ and $m = -l, \ldots, l$) together form a solution to the minimization problem (16). □

7. **Conclusion.** A novel variational framework was introduced for simultaneous smoothing and estimation of ADC profiles in the form of truncated SHS based on HARD MRI. Features of this model included minimizing a nonstandard growth function with nonlinear data fitting. Moreover, the constraints on the positivity and antipodal symmetry properties of d was also accommodated in the model. We also demonstrated our algorithm for using the variance of d from its mean and the coefficients of its truncated SHS approximation to characterize diffusion anisotropy.

Our experiments on both synthetic data and human HARD MRI data showed the effectiveness of the proposed model in the estimation of ADC profiles and the enhancement of the characterization of diffusion anisotropy. The characterization of non-Gaussian diffusion from the proposed method was consistent with known neuroanatomy.

The choice of the current parameters, however, may affect the results. Our choice was made based on the principle that classification for one-fiber diffusion from the model solution should agree with a priori knowledge of the fiber connections.

Acknowledgments. The first author is partially supported by NSF CCF-0527967 and NIH R01 NS052831-01 A1. The authors would like to thank M. Rao and H. Zhang from Department of Mathematics, University of Florida for their valuable discussion.

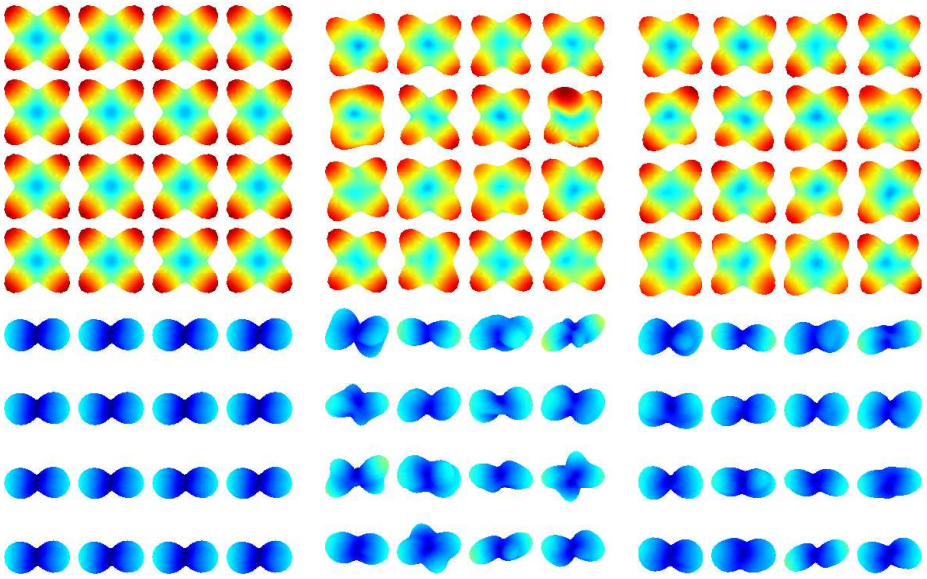


FIGURE 1. Comparing shapes of d. A) True d. B) The d generated by (10), with $A_{l,m}$'s the least square solution of (14) with the noisy measurement s. C) Recovered d by applying model (16).

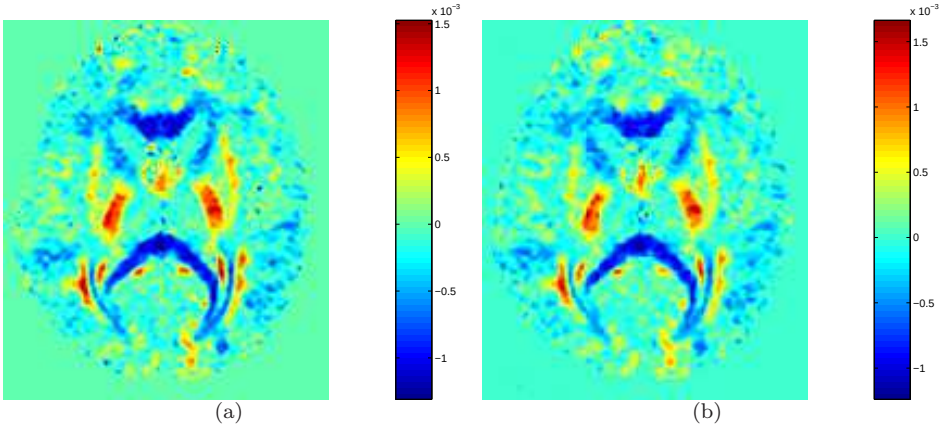


FIGURE 2. Comparison of A_{20}. A) A_{20} computed from (14). B) A_{20} obtained from model(16).

REFERENCES

[1] D. LeBihan and P. J. Basser, *Molecular diffusion and nuclear magnetic resonance*, Diffusion and perfusion magnetic resonance imaging, 1995.

[2] M. E. Moseley, Y. Cohen, J. Mintorovitch, J. L. Chileuitt, D. Norman and P. Weinstein, *Evidence of anisotropic self-diffusion in cat brain,* Proc. of the 8th ISMRM, (1989), 136–136.

[3] M. E. Moseley, J. Kucharczyk, H. S. Asgari and D. Norman, *Anisotropy in diffusion weighted MRI*, Magn. Reson. Med., **19** (1991), 321–326.

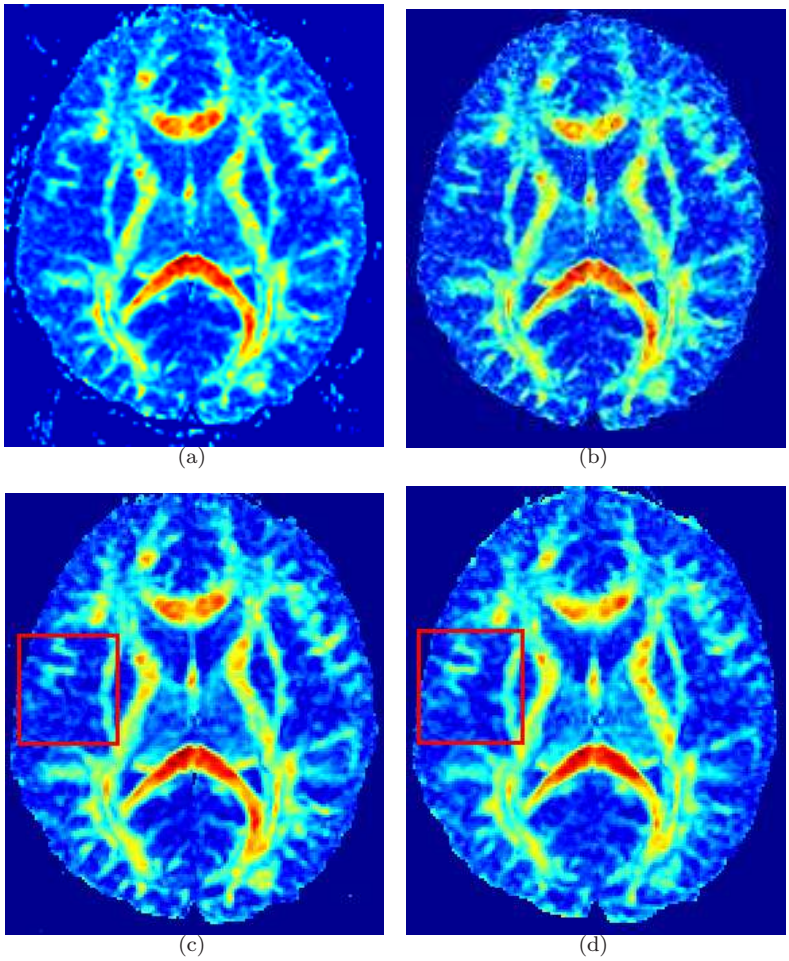
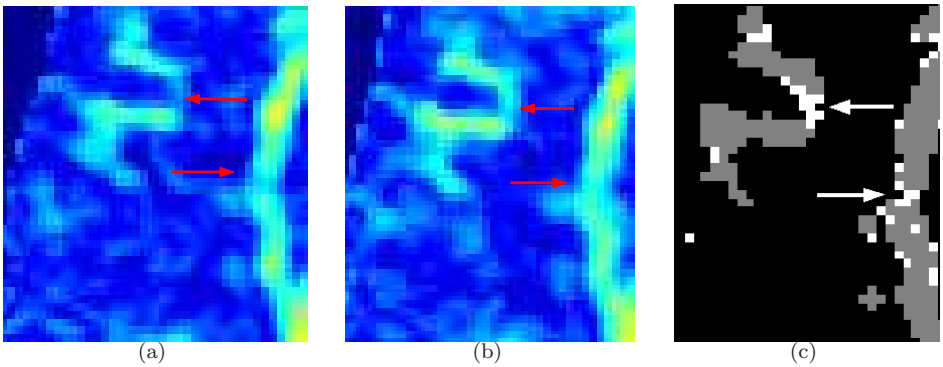


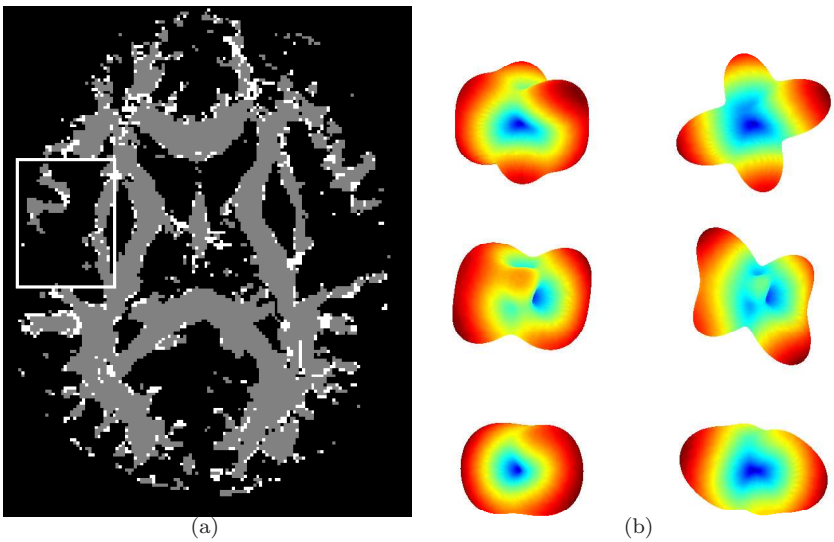
FIGURE 3. Images of FA and R_2. A) FA from GE software. B)-D) R_2 with the $A_{l,m}$'s as the solutions of (14), least-squares solutions of (7), and model solutions, respectively.

[4] P. J. Basser and C. Pierpaoli, *Microstructual and physiological features of tissues elucidated by quantitative diffusion tensor MRI*, Magn. Reson. Med., **111(B)** 1996, 209–219.
[5] E. O. Stejskal and J. E. Tanner, *Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient*, Chem. Phys., **42** (1965), 288–292.
[6] P. J. Basser, J. Mattiello and D. LeBihan, *MR diffusion tensor spectroscopy and imaging,* Biophys, **66** 1994, 259–267.
[7] D. S. Tuch, R. M. Weisskoff, J. W. Belliveau and V. J. Wedeen, *High angular resolution diffusion imaging of the human brain*, Proc. of the 7th ISMRM, (1999), 321–321.
[8] V. J. Wedeen, T. G. Reese, D. S. Tuch, M. R. Weigel, J.-G. Dou, R. M. Weisskoff and D. Chesler, *Mapping fiber orientation spectra in cerebral white matter with fourier transform diffusion MRI*, Proc. of the 8th ISMRM, (2000), 82–82.
[9] P. J. Basser, J. Mattiello and D. Lebihan, *Estimation of the effective self-diffusion tensor from the NMR*, Spin Echo. J. Magn. Reson., series B, **103** (1994), 247–254.
[10] T. L. Chenevert, J. A. Brunberg and J. G. Pipe, *Anisotropic diffusion in human white matter: demonstration with MR techniques in vivo*, Radiology, **177** (1990), 401–405.



(a) (b) (c)

Figure 4. Zoomed FA and A_{20}. A)-B) Enlarged portions inside the red squares in Figures. 3 C) and 3D), respectively. C) Enlarged portions inside the white squares in Figure 5A).



(a) (b)

Figure 5. Classification of voxels based on d. A) Classification: white, gray, and black voxels are identified as two-fiber, one-fiber, and isotropic diffusion respectively. B) Shapes of $d(x, \theta, \phi)$ at three particular points (upper, middle and lower rows). The d is computed via (10). $A_{l,m}(\mathbf{x})$ used in (10) in the left columns are the least-squares solutions of (7), while in the right column are the solutions from our model.

[11] E. W. Hsu and S. Mori, *Analytical expression for the NMR apparent diffusion coefficients in an anisotropy system and a simplified method for determing fiber orientation,* Magn. Reson. Med., **34** (1995), 194–200.

[12] L. Frank, *Characterization of anisotropy in high angular resolution diffusion weighted mri,* in "Proc. of 9th ISMRM," Glasgow, Scotland, (2001).

[13] A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda and D. L. Parker, *Analysis of partial volume effects in diffusion-tensor MRI*, Magn. Reson. Med., **45** (2001), 770–780.
[14] L. Frank, *Anisotropy in high angular resolution diffusion-weighted MRI,* Magn. Reson. Med., **45** (2001), 935–939.
[15] D. C. Alexander, G. J. Barker and S. R. Arridge, *Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data,* Magn. Reson. Med., **48** (2002), 331–340.
[16] Y. Chen, W. Guo, Q. Zeng, X. Yan, F. Huang, H. Zhang, G. He, B. Vemuri and Y. Liu, *Estimation, smoothing, and charaterization of apparent diffusion coefficient profiles from high angular resolution DWI,* Proc. of CVPR, (2004), 588–593.
[17] L. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithm,* Physica D, **60** (1992), 259–268.
[18] P. Blomgren, T. Chan, P. Mulet and C. K. Wong, *Total variation image restoration: Numerical methods and extensions,* Proceeding of IEEE Int'l Conference on Image Processing, **3** (1997), 384–387.
[19] A. Chambolle and P-L.Lions, *Image recovery via total variation minimization and related problems,* Numerische Mathematik, **76** (1997), 167–188.
[20] Y. Chen, S. Levine and M. Rao, *Variable exponent, linear growth functionals in image restoration,* SIAM Journal of AMath., **66** (2006), 1383–1406.
[21] P. Blomgren and T. Chan, *Color TV: total variation methods for restoration of vector-valued images,* IEEE Trans. on Image Processing, **7** (1998), 304–309.
[22] Z. Wang, B. C. Vemuri, Y. Chen and T. Mareci, *A constrained variational principle for direct estimation and smoothing of the tensor field from complex DWI*, IEEE TMI, **23** (2004), 930–939.
[23] J. Weickert, B. Romeny and M. Viergever, *Efficient and reliable schemes for nonlinear diffusion filtering,* IEEE Trans. on Img. Proc., **7** (1998), 398–410.
[24] T. Lu, P. Neittaanm and X. Tai, *A parallel splitting up method and its application to Navier-Stokes equations,* Applied Mathematics Letters, **4** (1991), 25–29.
[25] T. F. Chan and L. A. Vese, *Active contours without edges,* IEEE Trans. Image Processing, **10** (2001), 266–277.
[26] S. D. Conte and C. DeBoor, "Elementary Numerical Analysis," McGraw-Hill, New York, 1972.

Received May 2007; revised December 2007.

E-mail address: yun@math.ufl.edu
E-mail address: wguo@as.ua.edu
E-mail address: qingguo@math.ufl.edu
E-mail address: yijunliu@psychiatry.ufl.edu