Inverse Problems and Imaging
Volume 7, No. 2, 2013, 565–583
doi:10.3934/ipi.2013.7.565

# TOTAL VARIATION AND WAVELET REGULARIZATION OF ORIENTATION DISTRIBUTION FUNCTIONS IN DIFFUSION MRI

Yuyuan Ouyang and Yunmei Chen

Department of Mathematics
University of Florida
Gainesville, FL 32611, USA

Ying Wu

Center for Advanced Imaging
Evanston Hospital, 2650 Ridge Avenue
Evanston, IL 60201, USA

(Communicated by Hao-Min Zhou)

Abstract. We introduce a variational model and a numerical method for simultaneous ODF smoothing and reconstruction. The model uses the sparsity of MR images in finite difference domain and wavelet domain as the spatial regularization means in ODF's reconstruction. The model also incorporates angular regularization using Laplace-Beltrami operator on the unit sphere. A primal-dual scheme is applied to solve the model efficiently. The experimental results indicate that with spatial and angular regularization in the process of reconstruction, we can get better directional structures of reconstructed ODFs.

1. **Introduction.** Diffusion Weighted Magnetic Resonance Imaging (DW-MRI, or shortened as DWI) has been implemented widely as a non-invasive method to quantify water diffusion in tissues. Under the hypothesis that the preferred orientations of water diffusion will coincide with the fiber directions, DWI can determine the directionality of neuronal fiber bundles, that yield information on structural connections in brains [6, 18, 24, 25].

Water diffusion within tissue depends on the microstructure of the tissue. The average water diffusion probability density function (PDF) $P(\mathbf{r})$ at a specific voxel on a displacement $\mathbf{r}$ over an experiment diffusion time is related to the DWI measurements $S(\mathbf{q})$ [27] by a Fourier transform

$$S(\mathbf{q}) = S_0 \int_{\mathbb{R}^3} P(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}, \tag{1}$$

where $S(\mathbf{q})$ is the attenuation of the MR signal with respect to the diffusion sensitizing gradient $\mathbf{q}$, $S_0$ is the MRI signal in the absence of any gradient. The PDF $P(\mathbf{r})$ provides valuable information on the tissue microstructure. Since the water diffusion is more likely to happen at the direction of fiber tissue, the direction $\mathbf{r}$ of the maximum diffusion probability $P(\mathbf{r})$ will highly coincide with the direction of fiber tissue. However, for *in vivo* applications it is not feasible to reconstruct the diffusion PDF $P(\mathbf{r})$ from the MR signals $S(\mathbf{q})/S_0$ using the complex Fourier

2010 *Mathematics Subject Classification.* Primary: 92C55; Secondary: 15A29.
*Key words and phrases.* Diffusion magnetic resonance image, orientation distribution function, primal dual hybrid gradient method.

 

transform, since it requires a large number of measurements of $S(\mathbf{q})$ over a wide range of $\mathbf{q} \in \mathbb{R}^3$ in order to perform a stable inverse Fourier transform.

Diffusion Tensor Imaging (DTI) is a well-known classical MRI technique used to explore fiber tissue information in the brain. There have been a large amount of work on DTI that employs a second order, positive definite, symmetric diffusion tensor $D$ to represent the local tissue structure [4, 5, 6]. DTI implicitly assumes that the probability density function of the displacement of water diffusion is Gaussian with mean zero and covariance matrix $D$. The fractional anisotropy (FA) defined using the eigenvalues of D has become the most widely used measure of diffusion anisotropy in white matter. DTI has been shown to be a valuable tool in handling voxels with only one fiber, and studies have shown increasing clinical utility of DTI in the investigation of neuronal axon fiber integrity of white brain matter. However, it has been recognized that the single Gaussian model is inappropriate for assessing multiple fiber tract orientations, when complex tissue structure is found within a voxel [15, 29, 34].

In order to overcome these difficulties several approaches have been taken. Tuch et al. have proposed high angular resolution diffusion imaging (HARDI) method in which the acquisition makes the diffusion sensitizing gradients sample on the surface of a sphere [29, 30]. In [31] Tuch introduced Q-ball imaging (QBI), which is a HARDI technique, and used the orientation distribution function (ODF) to describe the orientational structure of fibre tissue. The local maxima of the ODFs implies the most probable fiber directions. In deterministic fiber tracing methods, such as streamline algorithms, the local maxima of the ODFs are assumed as the fiber directions. In statistical fiber tracing methods, such as Markov Chain Monte Carlo (MCMC) based algorithms, the ODFs can be used as the probability density functions of the fiber orientation.

The original definition of ODF is of the form

$$\Psi_1(\mathbf{u}) = \frac{1}{Z}\int_0^\infty P(r\mathbf{u})dr, \tag{2}$$

where $P(r\mathbf{u})$ is the same as in equation (1), $r = |\mathbf{r}|$, and $\mathbf{u} = \mathbf{r}/r$. With proper normalization constant $Z$, the ODF $\Psi_1(\mathbf{u})$ is a probability density function defined on a unit sphere. Tuch also showed in [31] that the ODF could be approximated directly from the raw HARDI signal $S(\mathbf{u})$ on a single unit sphere of q-space by the Funk-Radon transform (FRT) $\mathcal{G}$:

$$\Psi_1(\mathbf{u}) \approx \frac{1}{Z}\mathcal{G}[S](\mathbf{u})$$

where $\mathcal{G}[S](\mathbf{u})$ is defined as

$$\mathcal{G}[S](\mathbf{u}) = \int_{|\mathbf{w}|=1} \delta(\mathbf{u}^T\mathbf{w})S(\mathbf{w})d\mathbf{w},$$

and $\delta$ is the Dirac delta function.

In [1, 32] it is pointed out that, if we represent the orientation of unit vector $\mathbf{u}$ using spherical coordinate $(\theta, \phi)$, then

$$\begin{aligned}\int_{\mathbb{R}^3} P(\mathbf{r})d\mathbf{r} &= \int_0^\pi \int_0^{2\pi} \int_0^\infty P(r\mathbf{u})r^2\sin(\theta)drd\phi d\theta \\ &= \int_0^\pi \int_0^{2\pi} \left(\int_0^\infty P(r\mathbf{u})r^2dr\right)\sin(\theta)d\phi d\theta,\end{aligned}$$

and thus the marginal PDF on the unit sphere should be represented by

$$\Psi_2(\mathbf{u}) = \int_0^\infty P(r\mathbf{u})r^2 dr. \tag{3}$$

The definition of $\Psi_2(\mathbf{u})$ was actually proposed before in Wedeen et al. [35] as a "weighted radial summation". Comparing to $\Psi_2(\mathbf{u})$, in the definition of $\Psi_1$ the Jacobian factor $r^2$ is dropped, so $\Psi_1$ does not represent a true probability density function, and in practice the orientation information is blurred in the ODF estimation by $\Psi_1$. On the other hand, since $\Psi_2$ is a probability distribution function, it does not require the normalization factor $Z$ anymore.

Tristán-Vega et al. [32] use the property of Fourier transform and propose to estimate $\Psi_2(\mathbf{u})$ based on

$$\begin{aligned}\Psi_2(\mathbf{u}) &\approx C\mathcal{G}\left[\frac{\triangle S(\mathbf{q})}{S_0}\right](\mathbf{u}) \\ &= C\mathcal{G}\left[\frac{1}{q^2 S_0}\frac{\partial}{\partial q}\left(q^2\frac{\partial S(\mathbf{q})}{\partial q}\right) + \frac{1}{q^2 S_0}\triangle_b S(\mathbf{q})\right](\mathbf{u}),\end{aligned}$$

where $C$ is a constant, $q = |\mathbf{q}|$, and $\triangle_b$ is the Laplacian-Beltrami operator. Aganj et al. [1] showed that

$$\mathcal{G}\left[\frac{1}{q^2 S_0}\frac{\partial}{\partial q}\left(q^2\frac{\partial S(\mathbf{q})}{\partial q}\right)\right] \equiv -2\pi, \tag{4}$$

and developed a simple relationship between $\Psi_2(\mathbf{u})$ and the signal intensity on the unit sphere:

$$\Psi_2(\mathbf{u}) \approx \frac{1}{4\pi} + \frac{1}{16\pi^2}\mathcal{G}[\triangle_b \tilde{S}](\mathbf{u}), \tag{5}$$

where

$$\tilde{S}(\mathbf{q}) = \ln(-\ln(\frac{S(\mathbf{q})}{S_0})).$$

In equation (5), The ODF $\Psi_2(\mathbf{u})$ is estimated in each individual voxel, and no connection between the neighborhood points is considered. This can result in the error in ODF estimation when the data is noisy. There has been some work on the spatial regularization of the ODF results. H-E. Assemlal et al. [3] presented a variational framework for $\Psi_1(\mathbf{u})$. The model in his work is adaptable to the Rician distribution of MRI noise and able to use neighboring information by total variation (TV) based minimization. The similar methods have been proposed for the regularization of DTI [10, 23, 28] apparent diffusion coefficient (ADC) [9], and HARDI data [25]. However, there is still difficulty in incorporating TV based regularization into ODF estimation. One big problem is the computational complexity caused by non-differentiability of the TV norm. In many TV based smoothing algorithms a regularized TV norm was used to avoid non-smoothness problems, so that gradient descent methods can be applied. The drawback of using regularized TV norm is that it is sensitive to the regularization parameter, and takes longer time to get convergence.

Recently, several methods have been developed to solve the TV denoising problem efficiently with exact (not approximated) TV norm. They include using dual formulation [7], variable splitting and continuation [33, 36], split Bregman [16], primal-dual formulation [37, 38, 14, 8], proximity gradients [26, 19] and various forms of operator splitting [20, 22]. Other alternatives to TV based regularizers are also considered for magnetic resonance image reconstruction. One of them is the use

of $L_1$ sparsity under a wavelet transform. It has been exploited that MR images are sparse both in the spatial finite differences domain and under wavelet transform [21]. These properties have been successfully applied in MR reconstructions in compressive sensing [21].

In this paper we focus on the joint estimation and regularization of the ODF $\Psi_2$. The purpose of this paper is to provide a framework that simultaneously estimate and smooth the ODFs from the HARDI data, and a fast robust numerical algorithm to get the model solutions. Inspired by the previous work on the regularization for $\Psi_1$, we apply the angular and spatial regularization framework on the ODF model for $\Psi_2$, which has not been implemented previously. Furthermore, we are the first ones that consider the combination of total variation and wavelet based regularization as the spatial regularization on ODF. We also adapt the primal-dual numerical algorithm for solving combined total variation and wavelet based regularization models in the estimation of the ODF. Moreover, unlike the work in [1, 11, 12, 13, 31, 32], where the estimation of ODF is done after the reconstruction of $S$ is performed, we introduce a direct estimation and smoothing model on ODFs in the hope to reduce the accumulation of estimation error in the calculation.

Experimental results and comparisons provided in this work indicate the efficiency of the proposed method.

## 2. Background.

The Q-Ball Imaging scheme for solving ODF $\Psi_1$ in [31] requires very high load of calculation on the Funk-Radon Transform (FRT). A simplifciation was provided by Descoteaux et al. [12], in which the HARDI data is represented by spherical harmonic series (SHS). By introducing SHS, the calculation of the FRT is much simpler. Descoteaux et al. implemented the SHS on the calculation of $\Psi_1$ [11, 12, 13]. The use of SHS is also applied by Aganj et al. and Vega et al. [1, 32] in the estimation of ODF $\Psi_2$.

### 2.1. Spherical harmonics series.

A spherical harmonic function, denoted as $Y_l^m(\theta,\phi)$, is of the form

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\cdot\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi}, \tag{6}$$

where $P_l^m$ is the associated Legendre polynomial. $l$ i s called the order of the spherical harmonic function, $m$ is the phase factor, and $m=-l,\ldots,0,\ldots,l$. The function $Y_l^m(\theta,\phi)$ is defined on unit sphere, and the set of all spherical harmonic functions is an orthonormal basis of complex functions defined on unit sphere.

For all real functions defined on unit sphere, the orthogonal set of real spherical harmonic basis is usually used. For $l$ even number, choose $k=0,2,4,\ldots,l$, $m=-k,\ldots,0,\ldots,k$, a modified spherical harmonic basis $Y_j$ can be defined by

$$Y_j = \begin{cases} \sqrt{2}\,\mathrm{Re}(Y_k^m) & ,\text{ if } -k\le m<0 \\ Y_k^0 & ,\text{ if } m=0 \\ \sqrt{2}\,\mathrm{Im}(Y_k^m) & ,\text{ if } 0<m\le k \end{cases}, \tag{7}$$

where the index $j$ follows $j=(k^2+k+2)/2+m$ , and we still say that the order of $Y_j$ is $k$ [12]).

For any real valued function $\Psi(\mathbf{u})$ on an unit sphere, it can be approximated by the modified spherical harmonic basis mentioned above. Denoting the point $\mathbf{u}$ on

the unit sphere by its polar angles $(\theta, \phi)$, we will have

$$\Psi(\theta,\phi) \approx \sum_{j=1}^{R} c_j Y_j(\theta,\phi), \tag{8}$$

where $c_j$'s are real numbers, $R = (l+1)(l+2)/2$.

One useful property of spherical harmonic functions $\{Y_j\}_{j=1}^{R}$ is that they are the eigenfunctions of the Laplace-Beltrami operator $\triangle_b$ on the unit sphere:

$$\triangle_b Y_j(\mathbf{u}) = -l_j(l_j+1)Y_j(\mathbf{u}), \tag{9}$$

where $l_j$ is the order of $Y_j(\mathbf{u})$. Descoteaux et al. [12] used this property on the angular regularization of ODF $\Psi_1$. Aganj et al. [1] applied this property in the process of solving the analytical solution to the equation (5) for $\Psi_2$.

2.2. **Spherical harmonics series approximation of Funk-Radon transform.** In diffusion MRI, at a fixed voxel, the achieved signal intensities $S(\mathbf{u})$ and its Funk-Radon transform $\mathcal{G}[S](\mathbf{u})$ are real valued functions defined on unit sphere, and thus they can be approximated by the real spherical harmonic basis. Descoteaux et al. [12] proved that if $S(\mathbf{u})$ can be approximated as

$$S(\mathbf{u}) = \sum_{j=1}^{R} c_j Y_j(\mathbf{u}), \tag{10}$$

then the Funk-Radon transform $\mathcal{G}[S](\mathbf{u})$ can be approximated by

$$\mathcal{G}[S](\mathbf{u}) = \sum_{j=1}^{R} 2\pi P_{l_j}(0) c_j Y_j(\mathbf{u}), \tag{11}$$

where $l_j$ is the order of the modified spherical harmonics function $Y_j$, and $P_{l_j}(0)$ is the Legendre polynomial of degree $l_j$ evaluated at 0, i.e.,

$$P_{l_j}(0) = \begin{cases} 0 & l_j \text{ odd}, \\ (-1)^{l_j/2} \frac{1\cdot 3\cdot 5\cdots(l_j-1)}{2\cdot 4\cdot 6\cdots l_j} & l_j \text{ even}. \end{cases} \tag{12}$$

Aganj et al. [1] showed that if the signal is represented using SHS, i.e.,

$$\tilde{S}(\mathbf{u}) = \ln(-\ln(\frac{S(\mathbf{u})}{S_0})) = \sum_{j=1}^{R} c_j Y_j(\mathbf{u}), \tag{13}$$

then

$$\Psi_2(\mathbf{u}) = \sum_{j=1}^{R} a_j Y_j(\mathbf{u}), \tag{14}$$

where

$$a_j = \begin{cases} \dfrac{1}{2\sqrt{\pi}}, & j = 1, \\ -\dfrac{1}{8\pi} P_{l_j}(0) l_j(l_j+1) c_j, & j > 1. \end{cases} \tag{15}$$

In equation (15), $a_1 = 1/2\sqrt{\pi}$ is due to the fact that $Y_1(\mathbf{u}) \equiv 1/2\sqrt{\pi}$. In fact, from equation (5),

$$\sum_{j=2}^{R} a_j Y_j(\mathbf{u}) = \frac{1}{16\pi^2} \mathcal{G}[\triangle_b \tilde{S}](\mathbf{u}).$$

Aganj et al. [1] propose to estimate $\Psi_2(\mathbf{u})$ in two steps: First estimate the SHS coefficients $c_j$'s of $S(\mathbf{u})$ by least squares from equation (13), and then calculate the SHS coefficients $a_j$'s of $\Psi_2(\mathbf{u})$ by the relationship in equation (15).

3. **Model description.** In this section we present our model that is able to simultaneously estimate and smooth the ODF $\Psi_2(\mathbf{u})$, where the smoothing is performed with respect to both the spatial variable $x$ and the angular variable $(\theta, \phi)$.

The data fidelity term in our model is based on equation (14). Instead of voxel-by-voxel least square fitting on $\tilde{S}(\mathbf{u})$ as in [1], we start by assuming that

$$\Psi_2(x, \mathbf{u}) = \sum_{j=1}^{R} a_j(x) Y_j(\mathbf{u}), \ x \in \Omega, \tag{16}$$

where $\Omega$ is the image domain. The goal is simultaneously estimating and regularizing $\Psi_2(x, \mathbf{u})$ from the data $S(x, \mathbf{u})$ and $S_0(x)$. By the linear expansion described in equation (16), this problem reduce to the estimation and regularization of the coefficients $a_j(x)$, where $j = 1, 2, \ldots, R$, $x \in \Omega$. Moreover, from equation (15) we already have $a_1(x) \equiv 1/\sqrt{2\pi}, \forall x \in \Omega$.

Our model consists of four terms: a least squares energy, an angular regularization energy, a total variation regularization energy, and a wavelet $L_1$ sparsity regularization energy.

3.1. **Least squares energy.** We first present a least squares type energy for the estimation of the coefficients $a_j(x)$ in this subsection as the data fitting term in our energy functional. By using the relation of equations (13) and (16), from (14) and (15) we have

$$\tilde{S}(x, \mathbf{u}) = c_1(x) Y_1(\mathbf{u}) - \sum_{j=2}^{R} a_j(x) \cdot 8\pi [P_{l_j}(0) l_j (l_j + 1)]^{-1} Y_j(\mathbf{u}), \tag{17}$$

where $c_1(x)$ is the coefficient of $Y_1(\mathbf{u})$ in the SHS representation of $S(x, \mathbf{u})$. In fact, from the orthogonality of the real SHS, we have $Y_1(\mathbf{u}) \equiv 1/\sqrt{2\pi}$ and

$$\int_{\partial B^1} Y_j(\mathbf{u}) d\mathbf{u} = 0, \ \forall j > 1,$$

therefore

$$c_1(x) Y_1(\mathbf{u}) \equiv \frac{1}{4\pi} \int_{\partial B^1} \tilde{S}(x, \mathbf{u}) d\mathbf{u},$$

and

$$\left( \frac{1}{4\pi} \int_{\partial B^1} \tilde{S}(x, \mathbf{u}) d\mathbf{u} \right) - \tilde{S}(x, \mathbf{u}) = \sum_{j=2}^{R} a_j(x) \cdot 8\pi [P_{l_j}(0) l_j (l_j + 1)]^{-1} Y_j(\mathbf{u}). \tag{18}$$

Now if we let

$$F(x, \mathbf{u}) = \left( \frac{1}{4\pi} \int_{\partial B^1} \tilde{S}(x, \mathbf{u}) d\mathbf{u} \right) - \tilde{S}(x, \mathbf{u}),$$

and let

$$\widetilde{Y}_j(\mathbf{u}) = 8\pi [P_{l_j}(0) l_j (l_j + 1)]^{-1} Y_j(\mathbf{u}),$$

then from equation (18) we have

$$F(x, \mathbf{u}) = \sum_{j=2}^{R} a_j(x) \widetilde{Y}_j(\mathbf{u}). \tag{19}$$

where $F(x,\mathbf{u})$ can be calculated directly from the signal data $S(x,\mathbf{u})$. Therefore we define the least squares energy as

$$\begin{aligned} & E_1(a_2,\ldots,a_R) \\ = \ & \frac{1}{2}\int_\Omega\int_{\partial B^1}(F(x,\mathbf{u})-\sum_{j=2}^{R}a_j(x)\widetilde{Y}_j(\mathbf{u}))^2 d\mathbf{u}dx, \end{aligned} \tag{20}$$

where $\partial B^1$ denotes the unit sphere.

Although the DWI data has Rician noise, we use the least squares fidelity for simplicity, since the primal-dual optimization schemes are well studied especially for least squares fidelity terms. For Rician noise, it is also possible to use a likelihood based fidelity term, and then use a general primal-dual scheme to solve the problem.

3.2. **Angular regularization.** M. Descoteaux et al. [12] proposed a angular regularization on the signal $S(x,\mathbf{u})$ by minimizing the following term:

$$\int_{\partial B^1}\triangle_b^2 S(x,\mathbf{u})d\mathbf{u}, \tag{21}$$

where $\triangle_b$ denotes the Laplace-Beltrami operator on unit sphere. This angular regularization can reduce the affection by noise in ODF estimation, especially when using higher order spherical harmonics in the representation of the ODF. Given the property in equation (9), it is very easy to evaluate the Laplace-Beltrami operator acting on $S$ in equation (21). Inspired by [12] we apply the Laplace-Beltrami operator $\triangle_b$ on $\Psi_2(x,\mathbf{u})$. Then, we define

$$\begin{aligned} & E_2(a_2,\ldots,a_R) \\ = \ & \frac{1}{2}\int_{\partial B^1}\left(\triangle_b\Psi_2(x,\mathbf{u})\right)^2 d\mathbf{u} \\ = \ & \frac{1}{2}\int_{\partial B^1}\left(\sum_{j=1}^{R}a_j(x)\triangle_b Y_j(\mathbf{u})\right)^2 d\mathbf{u} \\ = \ & \frac{1}{2}\sum_{j=2}^{R}a_j^2(x)l_j^2(l_j+1)^2, \end{aligned} \tag{22}$$

where $\triangle_b Y_1(\mathbf{u})=0$ since $Y_1(\mathbf{u})\equiv 1/\sqrt{2\pi}$.

There are two advantages by including the angular regularization. First, since the weights $l_j^2(l_j+1)^2$ are larger for high order coefficients, the angular regularization tends to suppress the value of $a_j^2(x)$'s when $j$ is large, which helps in reducing the fake ODF maxima caused by the noise. This effect is well studied in [12]. Second, the energy functional $E_2$ is strongly convex with respects to the coefficients $a_2(x),\ldots,a_R(x)$, which will provide faster convergence and better robustness for the numerical scheme that solves our model.

There is a slight difference between our energy functions $E_1, E_2$ and the energy functions in the work of Descoteaux et al.. In [12, 13], where the least-squares fit and angular smoothing are applied on $S(x,\mathbf{u})$, the ODF $\Psi_1(x,\mathbf{u})$ is calculated using the smoothed $S(x,\mathbf{u})$ through the coefficients in the SH representations of $S$. In this paper the estimation and smoothing are applied directly to the ODFs $\Psi_2(x,\mathbf{u})$.

3.3. **Spatial regularization.** Minimizing the following energy functional

$$E_3(a_2,\dots,a_R)=\sum_{j=2}^{R}\int_\Omega |\nabla_x a_j(x)|\,dx \tag{23}$$

is the total variation (TV) based regularization, which is a technique used widely in MRI reconstruction. The idea of applying TV on the spherical harmonic representation for diffusion MRI is from [3, 9].

In our model we consider one other spatial regularization energy functional

$$E_4(a_2,\dots,a_R)=\sum_{j=2}^{R}\int_\Lambda |W[a_j](y)|\,dy, \tag{24}$$

where $W:\Omega\to\Lambda$ is a wavelet transform operator. This is a sparsity constraint for images $a_j$'s on domain $\Lambda$.

The purpose of spatial regularization is to enhance the images of $a_j(x)$'s, and remove image noises by regularizing the sparsity of $a_j(x)$'s in finite difference domain and wavelet domain. The combination of total variation and wavelet based regularization has been proven to be very effective in MRI, since most MRI images have been shown to be sparse in both the finite difference domain and wavelet domain [21]. For TV based image restoration, the restored image is often sharper in edge, but with possible staircase effects, while by wavelet based image restoration, the restored image is smoother. In [21], Lustig et al. introduced this regularization technique in multi-channel fast MRI reconstruction. In fact, our model can be seen as an extension of the method by Lustig et al. on diffusion MRI with different least squares energy, which represents the correlation between the multiple channels. Other studies on the total variation and wavelet regularization technique are also in [17, 22].

Inspired by those work, we combine both the total variation and wavelet regularization for analyzing diffusion MRI.

3.4. **Proposed model.** In this section we present our model. Define the following energy functional:

$$\begin{aligned} & E(a_2,\dots,a_R)\\ = \;& E_1(a_2,\dots,a_R)+\lambda E_2(a_2,\dots,a_R)\\ = \;& +\eta E_3(a_2,\dots,a_R)+\mu E_4(a_2,\dots,a_R), \end{aligned}$$

where $E_1$ is the least square energy defined in equation (20), $E_2$ is the energy for angular regularization defined in (22), $E_3$ is the TV regularization energy in (23), and $E_4$ is the wavelet $L_1$ sparsity regularization in (24). The parameters $\lambda,\eta,\mu$ are the balancing weights for angular regularization, TV based regularization, and wavelet based regularization. Our model estimates the coefficients $a_2,\dots,a_R$ by minimizing the energy functional $E(a_2,\dots,a_R)$.

3.5. **Discrete form of our model.** In this section we provide the discrete form of our model. Let $\mathbf{F}\in\mathbb{R}^{M\times N}$ be the matrix representing the discrete form of the MR signal information $\mathbf{F}(x,\mathbf{u})$. $M$ is the total number of the sensitizing gradient applied to get the data in Q-ball imaging, and $N$ is the total number of voxels in image domain $\Omega$. Let $\mathbf{A}\in\mathbb{R}^{(R-1)\times N}$ be the matrix of the discrete form of the $a_j(x)$'s, and $\mathbf{B}\in\mathbb{R}^{M\times(R-1)}$ be the matrix for real spherical harmonic basis functions $\widetilde{Y}_j(\mathbf{u})$.

We can rewrite the least square energy function $E_1$ as

$$E_1(\mathbf{A}) = \|\mathbf{BA} - \mathbf{F}\|_F^2,$$

where $\|\cdot\|_F$ is the Frobenius norm.

Also, write $\mathbf{A} = (\mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_R)^T$. Then, for each $j = 2, 3, \ldots, R$, $\mathbf{A}_j \in \mathbb{R}^{1\times N}$ is the row vector for the image defined by function $a_j(x)$, $x \in \Omega$. Under this notation our model can be written as minimizing the following energy function $E(\mathbf{A})$:

$$\begin{aligned} E(\mathbf{A}) = \quad & \frac{1}{2}\|\mathbf{BA} - \mathbf{F}\|_F^2 + \frac{\lambda}{2}\|\mathbf{LA}\|_F^2 \\ & +\eta \sum_{j=2}^{R} \|\mathbf{A}_j^T\|_{TV} + \mu \sum_{j=2}^{R} \|W(\mathbf{A}_j^T)\|_1, \end{aligned} \tag{25}$$

where $\mathbf{L}$ is a row vector with $\mathbf{L}_j = l_j(l_j + 1)$, $W$ denotes the discrete wavelet transform operator, and $\|\cdot\|_{TV}$ is the total variation of a image.

4. **Numerical scheme.** To minimize the energy function in (25), we adapt the modified primal-dual hybrid gradient algorithm proposed by E. Esser, X, Zhang and T. Chan [14], with slightly modification to cope with the wavelet regularization term. The primal-dual scheme is also equivalent to a special case of the primal-dual algorithms discussed in [8].

4.1. **Primal-dual formulation.** In [37] a primal-dual hybrid gradient (PDHG) scheme was developed on linear inversion problems with only TV regularization. Now we extend their scheme to the problem consisting of more regularization terms:

$$\min_{x\in\mathbb{R}^n} H(x) + \eta \sum_{i=1}^{n} \|D_i x\|_2 + \mu\|Wx\|_1. \tag{26}$$

Here $H(x)$ is a closed proper convex function, $D_i \in \mathbb{R}^{2\times n}, \Psi_1 \in \mathbb{R}^{n\times n}$ are linear operators acting on $x$. For the norms $\|\cdot\|_2$ and $\|\cdot\|_1$, we have

$$\begin{aligned} \eta\|D_i x\|_2 &= \max_{p_i\in\mathbb{R}^2, \|p_i\|_2\le 1} < \eta D_i x, p_i > \\ &= \max_{p_i\in\mathbb{R}^2, \|p_i\|_2\le 1} < x, \eta D_i^T p_i >, \\ \mu\|Wx\|_1 &= \max_{q\in\mathbb{R}^n, \|q\|_\infty\le 1} < \mu W x, q > \\ &= \max_{q\in\mathbb{R}^n, \|q\|_\infty\le 1} < x, \mu W^T q > . \end{aligned}$$

Therefore if we let

$$D = \begin{pmatrix} \eta D_1 \\ \eta D_2 \\ \vdots \\ \eta D_n \\ \mu W \end{pmatrix} \in \mathbb{R}^{3n\times n}, \; p = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \\ q \end{pmatrix} \in \mathbb{R}^{3n},$$

then

$$\eta \sum_{i=1}^{n} \|D_i x\|_2 + \mu\|\Psi_1 x\|_1 = \max_{p\in X} < Dx, p >,$$

where

$$\begin{aligned} X = \{p \in \mathbb{R}^{3n} : \quad & \|(p_{2i-1}, p_{2i})^T\|_2 \le 1, \forall i = 1, \cdots, n, \\ & \|(p_{2n+1}, p_{2n+2}, \cdots, p_{3n})^T\|_\infty \le 1\}. \end{aligned}$$

Then the minimization problem in equation (26) becomes a min-max problem:

$$\min_{x\in\mathbb{R}^n}\max_{p\in X}\ H(x)+<Dx,p>.$$

By [14, 37], the PDHG scheme is as follows:

(Dual Step)

$$\begin{aligned} p^{k+1} &= \operatorname*{argmax}_{p\in X} <Dx^k,p> -\frac{1}{2\tau_k}\|p-p^k\|_2^2 \\ &= \operatorname*{argmin}_{p\in X} <-Dx^k,p> +\frac{1}{2\tau_k}\|p-p^k\|_2^2 \\ &= \operatorname*{argmin}_{p\in X} \frac{1}{2\tau_k}\|p-(p^k+\tau_k Dx^k)\|_2^2 \\ &= \pi_X(p^k+\tau_k Dx^k), \end{aligned}$$

(Primal Step)

$$\begin{aligned} x^{k+1} &= \operatorname*{argmin}_{x\in\mathbb{R}^n} <x,D^Tp^{k+1}> +H(x) \\ &\quad +\frac{1}{2\theta_k}\|x-x^k\|_2^2, \end{aligned}$$

where $\pi_X$ denotes the projection onto space $X$, $\tau_k$ and $\theta_k$ are stepsizes.

A modified PDHG scheme is proposed in [14], by modifying the iteration of $p^{k+1}$ in dual step to

$$\begin{aligned} p^{k+1} &= \operatorname*{argmin}_{p\in X} <-Dy^k,p> +\frac{1}{2\tau_k}\|p-p^k\|_2^2 \\ &= \pi_X(p^k+\tau_k Dy^k), \end{aligned}$$

where

$$y^k=(1+\frac{\theta_k}{\theta_{k-1}})x^k-\frac{\theta_k}{\theta_{k-1}}x^{k-1}.$$

In fact, if $\{\theta_k\}$ is a constant sequence, $y^k=2x^k-x^{k-1}$. In this case, the modified PDHG algorithm is also a special case of the primal-dual algorithms studied in [8]. The convergence analysis is discussed in [14, 8].

4.2. **Primal-dual scheme for our model.** The variable $x$ in the minimization problem (26) is a vector in $\mathbb{R}^n$. In our proposed model, the variable $\mathbf{A}$ is a matrix. However, since the Frobenius norm is an entry-wise matrix norm, we can easily adapt our model to a vector form.

Assume $\mathbf{A}=\{a_{i,j}\}_{(R-1)\times N}=(\mathbf{A}_2,\mathbf{A}_3,\ldots,\mathbf{A}_R)^T$, and $\mathbf{F}=\{s_{i,j}\}_{M\times N}$. Here each $\mathbf{A}_i$ can be treated as a vector form of a 2D image. Now let $x\in\mathbb{R}^{(R-1)N\times 1}$ and $s\in\mathbb{R}^{MN\times 1}$ be the vector form of $\mathbf{A}$ and $\mathbf{F}$ by using dictionary order, i.e.,

$$\begin{aligned} x&=(a_{2,1},\ldots,a_{R,1},a_{2,2},\ldots,a_{R,2},\ldots,a_{2,N},\ldots,a_{R,N})^T,\\ s&=(s_{1,1},\ldots,s_{M,1},s_{1,2},\ldots,s_{R,2},\ldots,s_{1,N},\ldots,s_{M,N})^T. \end{aligned}$$

Then we have

$$\frac{1}{2}\|\mathbf{BA}-\mathbf{F}\|_F^2+\frac{\lambda}{2}\|\mathbf{LA}\|_F^2=\frac{1}{2}\|\mathbf{B}'x-s\|_2^2+\frac{\lambda}{2}\|\mathbf{L}'x\|_2^2,$$

where $\mathbf{B}'=\operatorname{diag}(\mathbf{B},\mathbf{B},\ldots,\mathbf{B})$, $\mathbf{L}'=\operatorname{diag}(\mathbf{L},\mathbf{L},\ldots,\mathbf{L})$, $\mathbf{B}',\mathbf{L}'\in\mathbb{R}^{(R-1)N\times(R-1)N}$. Now define $D_i\in\mathbb{R}^{2\times(R-1)N}$ to be the discrete form of gradient operator acting

on $x$ at voxel $i$, define $\Psi_1$ to be the operator that performs 2D discrete wavelet transform on each $\mathbf{A}_i$, and let

$$H(x) = \frac{1}{2}\|\mathbf{B}'x - s\|_2^2 + \frac{\lambda}{2}\|\mathbf{L}'x\|_2^2.$$

Then the minimization problem

$$\min_{x\in\mathbb{R}^{(R-1)N}} H(x) + \eta \sum_{i=1}^{(R-1)N} \|D_i x\|_2 + \mu\|Wx\|_1$$

will be equivalent to our model in (25).

To be consistent in notation, we write the primal-dual scheme here using our original notation with matrices.

Define $\mathbf{D} : \mathbb{R}^{(R-1)\times N} \to \mathbb{R}^{(R-1)\times N\times 2}$ to be the discrete form of gradient operator, and $\mathbf{W} : \mathbb{R}^{(R-1)\times N} \to \mathbb{R}^{(R-1)\times N}$ to be the discrete wavelet transform operator:

$$\mathbf{DA} = \mathbf{D}(\mathbf{A}_2, \ldots, \mathbf{A}_N)^T = (D\mathbf{A}_2^T, \ldots, D\mathbf{A}_N^T),$$
$$\mathbf{WA} = \mathbf{W}(\mathbf{A}_2, \ldots, \mathbf{A}_N)^T = (W\mathbf{A}_2^T, \ldots, W\mathbf{A}_N^T),$$

Notice that in the above $\mathbf{A}_i$ are actually the vector form of a 2D image. Thus $D$ is a 2D gradient operator, and $W$ is a 2D wavelet transform operator.

In dual step, we have

$$\begin{aligned} \mathbf{P}^{k+1} &= \pi_{\mathbf{X_1}}(\mathbf{P}^k + \tau_k\eta\mathbf{D}\mathbf{Y}^k), \\ \mathbf{Q}^{k+1} &= \pi_{\mathbf{X_2}}(\mathbf{Q}^k + \tau_k\mu\mathbf{W}\mathbf{Y}^k), \end{aligned} \tag{27}$$

where $\mathbf{P}^k \in \mathbb{R}^{(R-1)\times N\times 2}$, $\mathbf{Y}^k, \mathbf{Q}^k \in \mathbb{R}^{(R-1)\times N}$, and

$$\begin{aligned} \mathbf{X}_1 = \ & \{(\mathbf{P}_{i,j}) \in \mathbb{R}^{(R-1)\times N\times 2} : \|\mathbf{P}_{i,j}\|_2 \le 1, \\ & \forall i = 2, \ldots, R, \forall j = 1, \ldots, N\}, \\ \mathbf{X}_2 = \ & \{(\mathbf{Q}_{i,j}) \in \mathbb{R}^{(R-1)\times N} : |\mathbf{Q}_{i,j}| \le 1, \\ & \forall i = 2, \ldots, R, \forall j = 1, \ldots, N\}. \end{aligned}$$

For any $\mathbf{P} = (\mathbf{P}_{i,j}) \in \mathbb{R}^{(R-1)\times N\times 2}$, $\mathbf{Q} = (\mathbf{Q}_{i,j}) \in \mathbb{R}^{(R-1)\times N}$, we can actually write the projections component-wise as follows:

$$\begin{aligned} (\pi_{\mathbf{X_1}}(\mathbf{P}))_{i,j} &= \frac{\mathbf{P}_{i,j}}{\max(\|\mathbf{P}_{i,j}\|_2, 1)}, \\ (\pi_{\mathbf{X_2}}(\mathbf{Q}))_{i,j} &= \frac{\mathbf{Q}_{i,j}}{\max(|\mathbf{Q}_{i,j}|, 1).} \end{aligned}$$

In primal step, the optimal condition for $\mathbf{A}^{k+1}$ is

$$\begin{aligned} &\eta\mathbf{D}^T\mathbf{P}^{k+1} + \mu\mathbf{W}^T\mathbf{Q}^{k+1} + \mathbf{B}^T(\mathbf{B}\mathbf{A}^{k+1} - \mathbf{F}) \\ &+\lambda\mathbf{L}^2\mathbf{A}^{k+1} + \frac{1}{\theta_k}(\mathbf{A}^{k+1} - \mathbf{A}^k) = 0, \end{aligned}$$

Thus, we can write

$$\begin{aligned} \mathbf{A}^{k+1} &= \left(\theta_k(\mathbf{B}^T\mathbf{B} + \lambda\mathbf{L}^\mathbf{2}) + \mathbf{I}\right)^{-1}(\mathbf{A}^k \\ &\quad -\theta_k\eta\mathbf{D}^T\mathbf{P}^{k+1} - \theta_k\mu\mathbf{W}^T\mathbf{Q}^{k+1} \\ &\quad +\theta_k\mathbf{B}^T\mathbf{F}), \\ \mathbf{Y}^{k+1} &= (1 + \frac{\theta_k}{\theta_{k-1}})\mathbf{A}^{k+1} - \frac{\theta_k}{\theta_{k-1}}\mathbf{A}^k. \end{aligned} \tag{28}$$

Here $\mathbf{D}^T : \mathbb{R}^{(R-1)\times N\times 2} \to \mathbb{R}^{(R-1)\times N}$ is in fact the discrete form of negative divergence operator.

Finally, we write our primal-dual scheme as follows:

**Algorithm 1** Primal-Dual Scheme for solving (25)

$\mathbf{A}^0 \leftarrow 0, \mathbf{P}^0 \leftarrow 0, \mathbf{Q}^0 \leftarrow 0, \mathbf{Y}^0 \leftarrow 0$
**repeat**
  Iterate $\mathbf{P}^{k+1}$, $\mathbf{Q}^{k+1}$ by (27)
  Iterate $\mathbf{A}^{k+1}$, $\mathbf{Y}^{k+1}$ by (28)
**until** convergence condition is met

5. **Experimental results.** To verify the effectiveness of the proposed model and numerical algorothms in this section we provide our experimental results, and compare with the algorithms in [1], which incorporate least squares estimation and angular regularization over signal information $F(x, \mathbf{u})$ defined in equation (19). For the 2D wavelet transforms, we use a level 2 Daubechies-6 wavelet transform. To perform the transform in our program we use the Rice Wavelet Toolbox (RWT).

5.1. **Synthetic results.** The aim of this experiment is to examine the accuracy and robustness to noise of the proposed model in the reconstruction of ODFs. We generate the diffusion weighted signal $S$ using a bi-Gaussian model. For each gradient direction $\mathbf{u}$, we generate the signal intensity by

$$S(\mathbf{u}) = \frac{1}{2}\exp(-b\mathbf{u}^T D_1 \mathbf{u}) + \frac{1}{2}\exp(-b\mathbf{u}^T D_2 \mathbf{u}),$$

where $D_1$ and $D_2$ are diffusion tensor profiles with eigenvalues $[1700, 300, 300]\times 10^{-6}$, and $b = 3000$. The eigenvectors of the tensor profiles $D_1$ and $D_2$ are chosen to simulate a system of two crossing fiber bundles in a domain of 32x32 voxels. The region of the simulated fiber crossings is shown in figure 1. 55 gradient directions are used, and Rician random noise is added to the signal with different signal-to-noise ratio (SNR): 15, 20, 25 and 30.

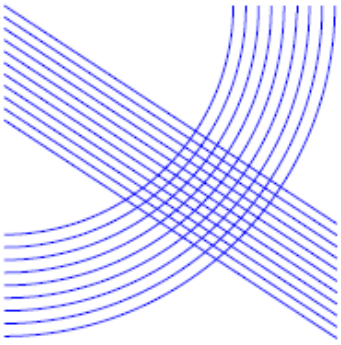

FIGURE 1. The simulated region of fiber crossings.

To examine the accuracy of the proposed model on the directional structures of the reconstructed ODFs, we estimate the regularized ODF, and then compare the fiber directions with the true values. By comparison, we also apply the scheme by Aganj et al. [1]. For each voxel, the estimated fiber directions are assumed to be the local maxima that surpasses a certain threshold (we use 0.5 here) of the estimated ODF, as suggested in [12]. The true fiber directions are assumed to be

the eigenvectors corresponding to the largest eigenvalues of the tensor profiles $D_1$ and $D_2$ in the bi-Gaussian model. We calculated the degree of angular difference in the fiber directions between the estimated ODFs and true ODFs. The estimation error is represented as the root mean square error (RMSE) of the angular differences in fiber directions: for estimated fiber directions $\{d_i^e\}_{i=1}^N$ and true fiber directions $\{d_i^t\}_{i=1}^N$,

$$RMSE = \sqrt{\frac{\sum_{i=1}^N g(d_i^e, d_i^t)^2}{N}}, \tag{29}$$

where for any vectors $d^e$ and $d^t$, $g(d^e, d^t)$ denotes the angle between them (in degrees). The parameters are optimized for the best fiber direction estimation. The results are presented in table 1. We can see that the RMSE is effectively reduced by our model, and the combination of TV smoothing and Wavelet smoothing provides the best estimation.

| RMSE / SNR / Model | 15 | 20 | 25 | 30 |
|---|---|---|---|---|
| $\Psi_2$ by equations (13)-(15) (Aganj) | 5.87 | 4.70 | 3.89 | 3.37 |
| $\Psi_2$ by our model, $\mu = 0$ (TV regularization only) | 1.68 | 1.43 | 1.42 | 1.41 |
| $\Psi_2$ by our model, with TV and wavelet regularization | 1.33 | 1.32 | 1.30 | 1.24 |

Table 1. Comparison of RMSE resulting from three models under SNR = 15, 20, 25, 30 respectively. The angular difference is measured in degrees.

To show how the regularization parameters affect the performance of the proposed model, we apply the model to two datasets, while varying one parameter and fixing the other two. In figure 2, from the left to the right, the graphs are the RMSE of fiber direction estimation results from the proposed model, with varying TV regularization parameter $\eta$ ($\mu$ and $\lambda$ are fixed), varying $\mu$ ($\eta$ and $\lambda$ are fixed), and varying $\lambda$ ($\eta$ and $\mu$ are fixed), respectively. We can see from the first column that when $\mu$ and $\lambda$ are fixed, the RMSE decreases significantly when the TV regularization parameter $\eta$ varies from 0 to 0.7 (first column), while the RMSE decreases slightly when the wavelet regularization parameter $\mu$ varies from 0 to 0.3 ($\eta$ and $\lambda$ fixed, second column). This shows that the TV regularization is dominant within spatial regularization. On the other hand, the RMSE decreases greatly when the angular regularization parameter $\lambda$ varies from 0 to 0.001 ($\eta$ and $\mu$ fixed, third column), while barely changed when $\lambda$ varies from 0.001 to 0.008. This shows that the angular regularization is important, while insensitive to the choice of $\lambda$ when $\lambda > 0.001$.

Figure 2 provides a guideline of choosing the regularization parameters. Since the angular regularization is insensitive to the choice of $\lambda$ when $\lambda > 0.001$, we can choose a constant small $\lambda$ for most diffusion MRI problems. Furthermore, since the TV regularization is dominant among the spatial regularization, we can set the wavelet regularization parameter $\mu = 0$ first, and fine tune the TV regularization

parameter $\eta$. After we get a desirable range of $\eta$, we can start tuning $\mu$ to suppress the staircase effect. In most of our synthetic and practical experiments, we find the above guideline useful.

Next we show that the proposed model provides more accurate estimation of spherical harmonic coefficients $a_2(x), a_3(x), \ldots, a_R(x)$ of ODFs. In figure 3 the spherical harmonic coefficients $a_2(x), a_3(x), \ldots, a_R(x)$ of ODF estimation is show as 2D images. The estimation is performed on the synthetic data with SNR 20. We can see from figure 3 that if only TV regularization is used, a staircase effect can be observed. On the other hand, the estimation result by implementation of both TV and wavelet regulariztion is effectively improved.

To quantify the performance of the proposed model under different noise level, we compare the estimated SHS coefficients of the ODFs and compare them with the ground truth. We calculate a set of spherical harmonic coefficients $a_1^t(x)$, $a_2^t(x), \ldots, a_R^t(x)$ on the synthetic data with no noise, and use these coefficients as ground truth. Based on the ground truth, we compare the performance of different models by comparing the sum of squares of the deviation (SSD) of the estimated coefficients $\{a_i(x)\}_{i=1}^R$ with the ground truth $\{a_i^t(x)\}_{i=1}^R$. The SSD is defined as

$$SSD = \sum_{i=1}^{R} (a_i(x) - a_i^t(x))^2. \tag{30}$$

The comparison by SSD is listed in table 2. From table 2 we can clearly see that the combination of both TV and wavelet regularization provides the best estimation of spherical harmonic coefficients.

| SSD / SNR / Model | 15 | 20 | 25 | 30 |
|---|---|---|---|---|
| $\Psi_2$ by equations (13)-(15) (Aganj) | 15.40 | 11.30 | 9.99 | 8.49 |
| $\Psi_2$ by our model, $\mu = 0$ (TV regularization only) | 9.72 | 7.05 | 6.28 | 5.99 |
| $\Psi_2$ by our model, with TV and wavelet regularization | 7.18 | 6.62 | 6.04 | 5.90 |

TABLE 2. Comparison of SSD resulting from three models under SNR = 15, 20, 25, 30 respectively.

The comparison of computational times is shown in table 3. Our codes are written in MATLAB and run on a Linux (version 2.6.38) computer with 2.67 GHz Intel i5 CPU and 8GB memory. We can see that although our model requires more computational time, but due to the efficient numerical scheme, the total computational load is still reasonable for the dataset with 45x32x32 spherical harmonic coefficients. We also perform the ODF estimation on one larger domain of 64x64 voxels, and from table 4 we can see that the computational load is still reasonable for solving a set of 45x64x64 spherical harmonic coefficients.

5.2. **Real data.** We apply the proposed model in a set of real experimental data. The DWI data is obtained on a SIMENS 3.0 Telsa scanner, with repetition time (TR)=9835ms, echo time (TE) =96ms, (FOV)=170.1 mm x 204.8 mm, $b$ = 1000,

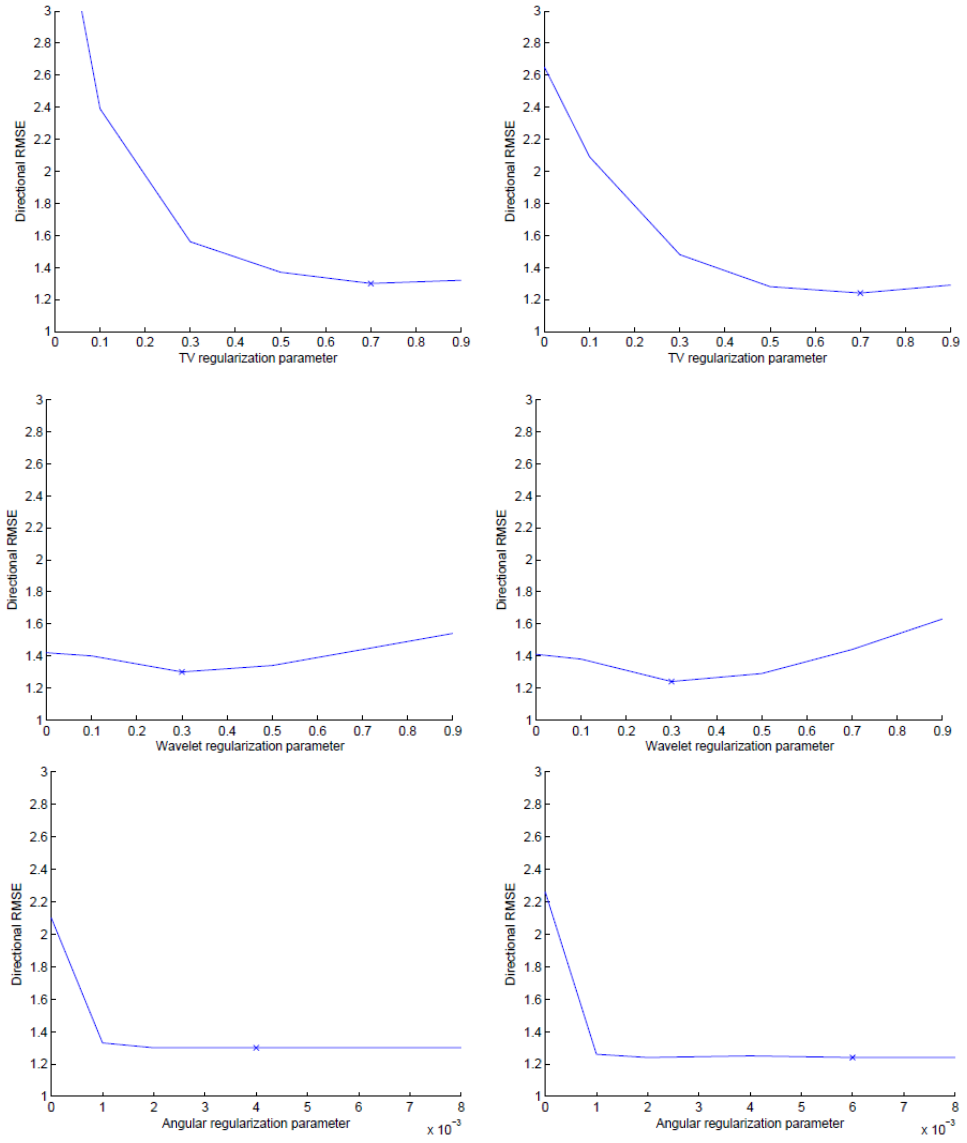


FIGURE 2. The performance of the proposed model while varying one parameter and fixing the other two. The x-axis denotes the choices of parameters, and the y-axis denotes the directional RMSE defined in equation (29). From top to bottom: RMSE under different choices of TV regularization parameter $\eta$ (with fixed $\mu$ and $\lambda$), wavelet regularization parameter $\mu$ (with fixed $\eta$ and $\lambda$) and angular regularization parameter $\lambda$ (with fixed $\eta$ and $\mu$), respectively. The left column is the performance on the dataset with SNR=25 (the best choice of parameters are $\eta = 0.7$, $\mu = 0.3$ and $\lambda = 0.004$). The right column is the performance on the dataset with SNR=30 (the best choice of parameters are $\eta = 0.7$, $\mu = 0.3$ and $\lambda = 0.006$).

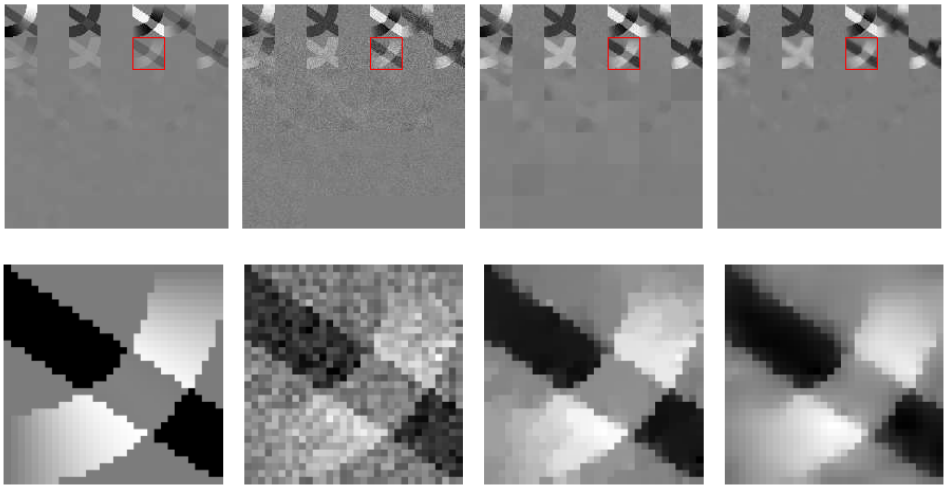

FIGURE 3. The image of spherical harmonic coefficients $a_2(x), a_3(x), \ldots, a_R(x)$ of ODF $\Psi_2$, estimated from a synthetic dataset with SNR 20. In the first row are the images of coefficients from different models. The first column is the ground truth, the second column is the result by Aganj et al. in [1], the third column is our model with only TV used in spatial regularization, and the last column is our model with both TV and wavelet in spatial regularization. For each image of coefficients, the images $a_2(x), a_3(x), \ldots, a_R(x)$ are arranged from top to bottom and left to right order with $a_2(x)$ at the top left corner. The second row is the zoomed in image of $a_{13}(x)$ (the region inside the red box in the first row), where the staircase effect of TV regularization can be observed.

| Computational Time / SNR / Model | 15 | 20 | 25 | 30 |
|---|---|---|---|---|
| $\Psi_2$ by equations (13)-(15) (Aganj) | 0.27 | 0.26 | 0.27 | 0.30 |
| $\Psi_2$ by our model, $\mu = 0$ (TV regularization only) | 0.87 | 0.77 | 0.76 | 0.96 |
| $\Psi_2$ by our model, with TV and wavelet regularization | 1.43 | 1.48 | 1.70 | 1.66 |

TABLE 3. Comparison of computational time (in seconds) resulting from three models under SNR = 15, 20, 25, 30 respectively. The dataset is on a domain of 32x32 voxels.

$M = 30$. The smoothing parameters are $R = 15$, $\lambda = 0.006$, $\eta = \mu = 0.2$. The region of interest (ROI) is shown in figure 4, and the estimated ODF results is presented in figure 5. From figure 5 we can see that by the proposed model the

| Computational Time / SNR / Model | 15 | 20 | 25 | 30 |
|---|---|---|---|---|
| $\Psi_2$ by equations (13)-(15) (Aganj) | 1.03 | 1.06 | 1.06 | 1.05 |
| $\Psi_2$ by our model, $\mu = 0$ (TV regularization only) | 3.58 | 3.67 | 3.55 | 3.35 |
| $\Psi_2$ by our model, with TV and wavelet regularization | 5.13 | 5.29 | 5.37 | 5.16 |

TABLE 4. Comparison of computational time (in seconds) resulting from three models under SNR = 15, 20, 25, 30 respectively. The dataset is on a domain of 64x64 voxels.

noise on fiber directions is effectively reduced, and a clear track of fiber directions can be seen.

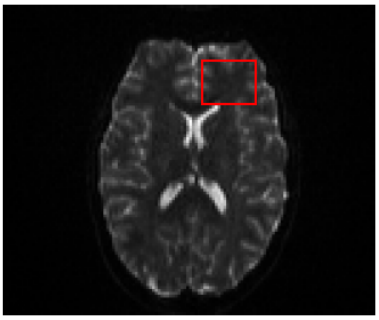

FIGURE 4. The region of interest in real data.

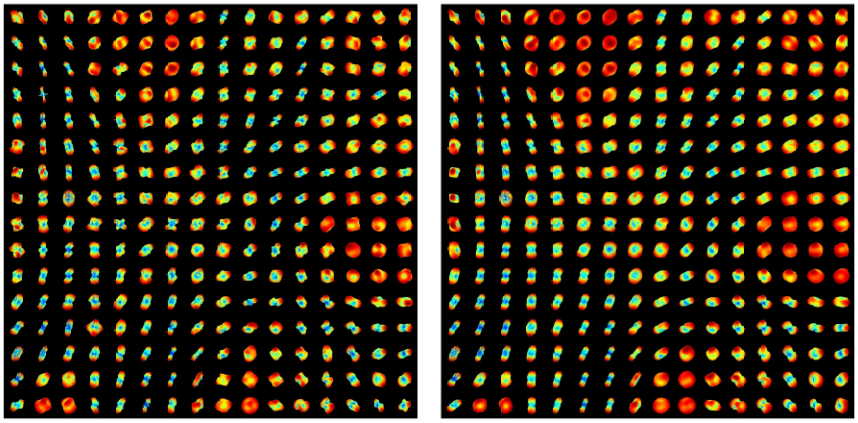

FIGURE 5. Real data. At the left is the result with no spatial smoothing, and at the right is the smoothed ODF by our framework.

6. **Conclusion.** We propose a model for regularization in the estimation of ODFs. The model performs simultaneous angular and spatial regularization to ODFs fields. The angular regularization is by using Laplace-Beltrami operator. For the spatial regularization, we use total variation and wavelet transform. The implemented numerical method is recently developed and very fast. We demonstrate the drawback of only angular regularization and the advantage of incorporating both angular and spatial regularization in our synthetic experiments. With our model we can achieve better orientational information for the reconstructed ODF fields.

7. **Acknowledgments.** This research was partly supported by National Science Foundation grant 1115568. The authors would like to thank the anonymous referees for their helpful suggestions.

## REFERENCES

[1] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil and N. Harel, *Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle*, Magnetic Resonance in Medicine, **64** (2010), 554–566.

[2] K. Arrow, L, Hurwicz, H. Uzawa and H. Chenery, *Studies in linear and non-linear programming*, Stanford Mathematical Studies in the Social Sciences, **II**, Stanford University Press, Stanford, Calif., (1958).

[3] H. Assemlal, D. Tschumperlé and L. Brun, *Fiber tracking on HARDI data using robust ODF fields*, in "IEEE International Conference on Image Processing," Citeseer, (2007), 344–351.

[4] P. Basser, J. Mattiello and D. Lebihan, *Estimation of the effective self-diffusion tensor from the NMR spin echo*, Journal of Magnetic Resonance, Series B, **103** (1994), 247–247.

[5] P. Basser and C. Pierpaoli, *Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI*, Journal of Magnetic Resonance, Series B, **111** (1996), 209–219.

[6] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda and A. Aldroubi, *In vivo fiber tractography using DT-MRI data*, Magnetic Resonance in Medicine, **44** (2000), 625–632.

[7] A. Chambolle, *An algorithm for total variation minimization and applications*, Journal of Mathematical Imaging and Vision, **20** (2004), 89–97.

[8] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, **40** (2011), 120–145.

[9] Y. Chen, W. Guo, Q. Zeng and Y. Liu, *A nonstandard smoothing in reconstruction of apparent diffusion coefficient profiles from diffusion weighted images*, Inverse Problems and Imaging, **2** (2008), 205–224.

[10] O. Christiansen, T. Lee, J. Lie, U. Sinha and T. Chan, *Total variation regularization of matrix-valued images*, International Journal of Biomedical Imaging, **2007** (2007).

[11] M. Descoteaux, E. Angelino, S. Fitzgibbons and R. Deriche, *Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications*, Magnetic Resonance in Medicine, **56** (2006), 395–410.

[12] M. Descoteaux, E. Angelino, S. Fitzgibbons and R. Deriche, *Regularized, fast, and robust analytical Q-ball imaging*, Magnetic Resonance in Medicine, **58** (2007), 497–510.

[13] M. Descoteaux, R. Deriche, T. Knösche and A. Anwander, *Deterministic and probabilistic tractography based on complex fibre orientation distributions*, IEEE transactions on medical imaging, **28** (2009), 269–286.

[14] E. Esser, X. Zhang and T. Chan, *A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science*, SIAM Journal on Imaging Sciences, **3** (2010), 1015–1046.

[15] L. Frank, *Anisotropy in high angular resolution diffusion-weighted MRI*, Magnetic Resonance in Medicine, **45** (2001), 935–939.

[16] T. Goldstein and S. Osher, *The split Bregman method for L1 regularized problems*, SIAM Journal on Imaging Sciences, **2** (2009), 323–343.

[17] L. He, T.-C. Chang, S. Osher, T. Fang and P. Speier, *MR image reconstruction by using the iterative renement method and nonlinear inverse scale space methods*, UCLA CAM Reports 06-35, (2006).

[18] D. Jones, A. Simmons, S. Williams and M. Horsfield, *Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI*, Magnetic Resonance in Medicine, **42** (1999), 37–41.
[19] Q. Li, C. A. Micchelli, L. Shen and Y. Xu, *A proximity algorithm accelerated by Gauss-Seidel iterations for L1/TV denoising models*, Inverse Problems, **28** (2012).
[20] P. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM Journal on Numerical Analysis, **16** (1979), 964–979.
[21] M. Lustig, D. Donoho and J. Pauly, *Sparse MRI: The application of compressed sensing for rapid MR imaging*, Magnetic Resonance in Medicine, **58** (2007), 1182–1195.
[22] S. Ma, W. Yin, Y. Zhang and A. Chakraborty, *An efficient algorithm for compressed MR imaging using total variation and wavelets*, in "IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008)," (2008).
[23] A. Ramirez-Manzanares and M. Rivera, *Basis tensor decomposition for restoring intra-voxel structure and stochastic walks for inferring brain connectivity in DT-MRI*, International Journal of Computer Vision, **69** (2006), 77–92.
[24] T. McGraw, B. Vemuri, Y. Chen, M. Rao and T. Mareci, *DT-MRI denoising and neuronal fiber tracking*, Medical Image Analysis, **8** (2004), 95–111.
[25] T. McGraw, B. Vemuri, E. Ozarslan, Y. Chen and T. Mareci, *Variational denoising of diffusion weighted MRI*, Inverse Problems and Imaging, **3** (2009), 625–648.
[26] C. A. Micchelli, L. Shen and Y. Xu, *Proximity Algorithms for Image Models: Denoising*, Inverse Problems, **27** (2011).
[27] E. Stejskal and J. Tanner, *Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient*, The Journal of Chemical Physics, **42** (1965), 288.
[28] D. Tschumperlé and R. Deriche, *Variational frameworks for DT-MRI estimation, regularization and visualization*, in "Ninth IEEE International Conference on Computer Vision," (2003), 116–121.
[29] D. Tuch, R. Weisskoff, J. Belliveau and V. Wedeen, *High angular resolution diffusion imaging of the human brain*, in "Proceedings of the 7th Annual Meeting of ISMRM," (1999), 321–321.
[30] D. Tuch, T. Reese, M. Wiegell and V. J. Wedeen, *Diffusion MRI of complex neural architecture*, Neuron, **40** (2003), 885–895.
[31] D. Tuch, *Q-ball imaging*, Magnetic Resonance in Medicine, **52** (2004), 1358–1372.
[32] A. Tristán-Vega, C. Westin and S. Aja-Fernández, *Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging*, NeuroImage, **47** (2009), 638–650.
[33] Y. Wang, J. Yang, W. Yin and Y. Zhang, *A new alternating minimization algorithm for total variation image reconstruction*, SIAM Journal on Imaging Sciences, **1** (2008), 248–272.
[34] V. Wedeen, T. Reese, D. Tuch, M. Weigel, J. Dou, R. Weiskoff and D. Chessler, *Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI*, in "Proc. Intl. Sot. Mag. Reson. Med.," **8** (2000), 82–82.
[35] V. Wedeen, P. Hagmann, W. Tseng, T. Reese and R. Weisskoff, *Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging*, Magnetic Resonance in Medicine, **54** (2005), 1377–1386.
[36] J. Yang, Y. Zhang and W. Yin, *A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data*, IEEE Journal of Selected Topics in Signal Processing, **4** (2010), 288–297.
[37] M. Zhu and T. Chan, *An efficient primal-dual hybrid gradient algorithm for total variation image restoration*, UCLA CAM Report 08–34, (2008).
[38] M. Zhu, S. Wright and T. Chan, *Duality-based algorithms for total-variation-regularized image restoration*, Computational Optimization and Applications, **47** (2010), 377–400.

Received October 2011; revised August 2012.

*E-mail address:* ouyang@ufl.edu
*E-mail address:* yun@ufl.edu
*E-mail address:* ywu@northshore.org