



Published in final edited form as:

Med Image Anal. 2014 October ; 18(7): 1143–1156. doi:10.1016/j.media.2014.06.003.

Multi-shell diffusion signal recovery from sparse measurements

Y. Rathi^a, O. Michailovich^b, F. Laun^d, K. Setsompop^e, P. E. Grant^c, and C-F Westin^a

^aBrigham and Women's Hospital, Harvard Medical School, Boston

^bUniversity of Waterloo, Canada

^cChildren's Hospital Boston, Harvard Medical School, Boston

^dGerman Cancer Research Center (DKFZ), Dept. of Medical Physics in Radiology, Heidelberg, Germany

^eMassachusetts General Hospital, Harvard Medical School, Boston

Abstract

For accurate estimation of the ensemble average diffusion propagator (EAP), traditional multi-shell diffusion imaging (MSDI) approaches require acquisition of diffusion signals for a range of b -values. However, this makes the acquisition time too long for several types of patients, making it difficult to use in a clinical setting. In this work, we propose a new method for the reconstruction of diffusion signals in the entire q -space from highly under-sampled sets of MSDI data, thus reducing the scan time significantly. In particular, to sparsely represent the diffusion signal over multiple q -shells, we propose a novel extension to the framework of spherical ridgelets by accurately modeling the monotonically decreasing radial component of the diffusion signal. Further, we enforce the reconstructed signal to have smooth spatial regularity in the brain, by minimizing the total variation (TV) norm. We combine these requirements into a novel cost function and derive an optimal solution using the Alternating Directions Method of Multipliers (ADMM) algorithm. We use a physical phantom data set with known fiber crossing angle of 45° to determine the optimal number of measurements (gradient directions and b -values) needed for accurate signal recovery. We compare our technique with a state-of-the-art sparse reconstruction method (i.e., the SHORE method of (Cheng et al., 2010)) in terms of angular error in estimating the crossing angle, incorrect number of peaks detected, normalized mean squared error in signal recovery as well as error in estimating the return-to-origin probability (RTOP). Finally, we also demonstrate the behavior of the proposed technique on human *in-vivo* data sets. Based on these experiments, we conclude that using the proposed algorithm, at least 60 measurements (spread over three b -value shells) are needed for proper recovery of MSDI data in the entire q -space.

© 2014 Elsevier B.V. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

diffusion MRI; compressed sensing; diffusion spectrum imaging; diffusion propagator; multi shell diffusion imaging; kurtosis

1. Introduction

Diffusion MRI (dMRI) is an imaging modality that is sensitive to the neural architecture and connectivity of the brain. Consequently, it is increasingly being used in clinical settings for investigating several brain disorders such as, Alzheimer's disease, stroke, schizophrenia, mild traumatic brain injury, etc. (Thomason and Thompson, 2011; Shenton et al., 2012). Apart from more traditional Diffusion Tensor Imaging (DTI), it is nowadays standard to use High Angular Resolution Diffusion Imaging (HARDI), which involves acquiring diffusion signals at a single b -value (single q -shell) in several gradient directions spread over the unit sphere in a quasi-uniform manner (Tuch et al., 2003; Asselmlal et al., 2011). While this protocol allows for resolving the complex angular structure of the neural fibers, it does not provide information about the radial signal decay, which is known to be sensitive to various anomalies of white matter (Cohen and Assaf, 2002).

To obtain accurate information about the neural architecture, diffusion spectrum imaging (DSI) was proposed by Wedeen et al. (2005). This dMRI technique involves acquiring multiple measurements over a Cartesian grid of points in the q -space, followed by application of discrete Fourier transform to obtain an estimate of the ensemble average propagator (EAP). Unfortunately, a large number of measurements required by DSI makes it impractical to use in clinical settings. Accordingly, to speed-up the acquisition of dMRI (and DSI) data, two complementary approaches have been proposed, namely: (i) the use of compressed sensing (CS) to reduce the number of measurements (Candès et al., 2006; Donoho, 2006), and (ii) the use of multi-slice acquisition sequences for faster data acquisition (Setsompop et al., 2011; Feinberg et al., 2010). This work focuses on methodology (i), i.e., CS-based reconstruction of diffusion signal from critically under-sampled measurements.

Several imaging and analysis schemes, which use fewer measurements than traditional DSI, have recently been proposed in the literature (Wu and Alexander, 2007; Jensen et al., 2005; Asselmlal et al., 2011; Merlet et al., 2012; Barmpoutis et al., 2008; Descoteaux et al., 2010; Zhang et al., 2012; Ye et al., 2011, 2012; Hosseinbor et al., 2012). Each of these techniques captures a different aspect of the underlying tissue organization, which is missed by HARDI. Traditional methods of EAP estimation that account for the non-monoexponential (radial) decay of diffusion signals, require a relatively large number of measurements at high b -values (greater than 3000 s/mm^2) (Assaf et al., 2004; Mulkern et al., 2001). Consequently, their associated scan times are deemed to be too long for non-cooperative patients, which is the main motivation for reducing the number of measurements in dMRI scans.

Although not new in application to MRI, CS-based methods of signal reconstruction has gained significant attention in the diffusion imaging community over the last few years. Several works have proposed CS-based algorithms for recovering HARDI, MSDI as well as

DSI data from undersampled (aka incomplete) measurements (Ye et al., 2011; Merlet et al., 2012; Landman et al., 2012; Gramfort et al., 2012; Duarte-Carvajalino et al., 2012; Freiman et al., 2013; Scherrer et al., 2013; Assemlal et al., 2011; Michailovich et al., 2011; Rathi et al., 2011). To this end, various types of signal representation bases have also been proposed, each having different sparsifying properties. For example, for HARDI data, spherical ridgelets were proposed in (Michailovich et al., 2008; Michailovich and Rathi, 2010), and for MSDI data, spherical polar Fourier (SPF) and its variants (SHORE) were proposed in (Assemlal et al., 2008; Ozarslan et al., 2008; Cheng et al., 2010; Merlet et al., 2012). In the case of the SHORE basis, to optimize the accuracy of signal reconstruction, one has to choose an appropriate scaling parameter, which could potentially be different for different types of tissue. To address this issue, (Merlet et al., 2012) used a dictionary learning technique to learn the scaling parameter and the appropriate polynomial to represent the radial decay term. On the other hand, in (Özarslan et al., 2013), this scaling parameter was adaptively obtained in a data driven fashion by computing the eigenvalues of a tensor at each voxel. However, at a fundamental level, both these methods extend the original SHORE basis to sparsely represent the diffusion data. In this work, we will compare our technique with the SHORE-based reconstruction (Cheng et al., 2010; Merlet and Deriche, 2013), where sparsity is enforced through the standard l_1 -norm minimization.

2. Our contributions

The framework of spherical ridgelets (SR) proposed in (Michailovich et al., 2011) was used to recover HARDI data on a single b-value shell from highly undersampled set of diffusion measurements. In this work, we propose a novel extension of this basis for recovering multi-shell diffusion data. Towards this end, we incorporate a novel radial decay term which is a monotonically decreasing function with its range bounded between 0 and 1. This property is quite desirable, since it is known that the values of normalized diffusion signals lie within this range (Clark and Le Bihan, 2000; Schwarcz et al., 2004; Mulkern et al., 2009). In this work, we use spherical ridgelets to perform CS-based reconstruction of MSDI signals over each of their associated b-value shells (q -shells), while using the radial decay term for representing the signal attenuation with increasing b -values. To obtain an optimal consensus solution that ensures spatially smooth signal recovery, we propose a novel computational framework based on the ADMM algorithm. We perform extensive testing of the proposed algorithm on a physical phantom data set and compare it with the SHORE-based method. We provide quantitative results in terms of the error in estimation of the orientation, incorrect number of peaks detected, normalized mean squared error (NMSE) in the estimation of the signal as well as NMSE in the estimation of the return-to-origin probability (RTOP). We also provide similar quantitative results on human *in-vivo* data set.

The primary aim of the algorithm presented in this work is the recovery of diffusion signal from sub-critically sampled measurements. Following this, any model or methodology (such as, multi-compartment models, kurtosis, diffusion propagator, free-water, etc.) can be used to compute diffusion measures or features (Özarslan et al., 2013). Thus, in this work, we do not focus on recovering model specific diffusion properties as they can be computed once an estimate of the diffusion signal in the entire q -space is available using the proposed method.

3. Background

3.1. Diffusion MRI

Under the narrow pulse assumption, the diffusion signal $S(\mathbf{q})$ in the q -space is related to the EAP $P(\mathbf{r})$ via the Fourier transform as given by (Stejskal and Tanner, 1965)

$$P(\mathbf{r}) = \int_{\mathbf{r} \in \mathbb{R}^3} E(\mathbf{q}) \exp(-i2\pi \mathbf{q} \cdot \mathbf{r}) d\mathbf{q},$$

where $E(\mathbf{q}) \triangleq S(\mathbf{q})/S(0) : \mathbb{R}^3 \rightarrow [0, 1]$ is the *normalized* diffusion signal, with $S(\mathbf{q})$ and $S(0)$ being the measured diffusion signal and its corresponding $b = 0$ value, respectively.

Alternatively, E can be written as a function of b -value and a unit vector $\mathbf{u} \in \mathbb{S}^2$, such that $E(b, \mathbf{u}) : \mathbb{R}^+ \times \mathbb{S}^2 \rightarrow [0, 1]$, where $b = \gamma^2 \delta^2 (\|\mathbf{g}\|^2 - \delta^2)$ s/mm^2 , with δ being the duration of the gradient pulse, τ is the mixing time (i.e., the time between the two diffusion-encoding gradients), γ is the gyromagnetic constant, and $\|\mathbf{g}\|$ denotes the Euclidean norm of the diffusion-encoding gradient \mathbf{g} . In the context of MSDI, the signal E is measured along N discrete orientations $\{\mathbf{u}_k\}_{k=1}^N$ for several different values of b . Thus, for each b value shell, the sampling points are spread over the unit sphere, thereby giving the measurements a multi-shell structure.

3.2. Compressed sensing

The theory of CS provides the mathematical foundation for accurate recovery of signals from their discrete measurements acquired at sub-critical (aka sub-Nyquist) rate (Candès et al., 2006; Donoho, 2006; Candès et al., 2011). The theory relies on two key concepts: *sparsity* and *incoherence*, although the latter requirement could be relaxed in certain cases (Candès et al., 2011). Sparsity implies that the signal of interest should have a sparse representation in some basis/frame $\Psi \in \mathbb{R}^{N \times M}$, which we term as the representation dictionary. The signal $E \in \mathbb{R}^N$ is said to admit a sparse representation in Ψ if its expansion coefficients contain only a small number of significant coefficients, i.e. if $E = \Psi \mathbf{c}$, then most of the elements of $\mathbf{c} \in \mathbb{R}^M$ are zero. If only K elements of \mathbf{c} are nonzero, then the signal E is said to be K -sparse in Ψ , where $K \ll M$.

The framework of CS also relies on a sensing or sampling basis Φ . In the context of diffusion MRI, since we have a single value $E(\mathbf{q})$ associated to each point \mathbf{q} in the q -space, we assume that the sampling basis Φ consists of rotated versions of a spherical Dirac delta function. Consequently, denoting by $\mathbf{s} \in \mathbb{R}^P$ a column vector of P discrete measurements of E , our goal is to recover the representation coefficients \mathbf{c} (and hence the corresponding diffusion signal) given the data vector

$$\mathbf{s} = \Phi \Psi \mathbf{c} + \eta, \quad (1)$$

where η is measurement noise and the basis Φ acts as a subsampling operator. CS theory asserts that, to reconstruct the full signal from its incomplete measurements \mathbf{s} , one can use a non-linear decoding scheme represented by the following ℓ_1 -norm minimization problem

$$\min \|\mathbf{c}\|_1, \text{ s.t. } \|\mathbf{s} - \Phi\Psi\mathbf{c}\|^2 \leq \varepsilon. \quad (2)$$

This minimization problem can be readily solved using a variety of optimization techniques (Beck and Teboulle, 2009; Asif and Romberg, 2009).

Earlier works on CS focused on signal recovery when Ψ was assumed to be an orthonormal basis (Candès et al., 2006). In this case, high *incoherence* between the representation Ψ and sampling Φ bases was a necessary condition for a successful CS-based signal reconstruction. For the case when Ψ is chosen to be an overcomplete dictionary (as it is the case in the present study), the importance of the above condition was recently shown to be much less critical (Candès et al., 2011). As such, the ability of an overcomplete Ψ to provide sparse representation for the signals of interest *alone* can guarantee reliable signal recovery from incomplete measurements. However, in this scenario, the lower bound on the number of measurements required for signal recovery is still application dependent and has to be determined from realistic experimental validation studies. More importantly, this lower bound depends on the level of sparsity of the representation dictionary Ψ . Consequently, we will use an experimental setup to determine the minimal number of gradient directions (measurements) required for proper recovery of dMRI data in q -space.

3.3. Spherical Ridgelets

Spherical ridgelets (SR) form a basis for representing \mathcal{L}_2 functions defined on the unit sphere (Michailovich et al., 2008). Specifically, it was shown to provide sparse representation of diffusion signals over a single b -value shell. To avoid repetitions, we present only the principal concepts of SR design, while their detailed description can be found in (Michailovich et al., 2008).

Spherical ridgelets are constructed using the fundamental principles of wavelet theory. Specifically, let $x \in \mathbb{R}^+$ and $\rho \in (0, 1)$ be a positive scaling parameter. Further, let $\kappa(x) = \exp\{-\rho x(x+1)\}$ be a Gaussian function, which we subject to a series of dyadic scalings as shown below

$$\kappa_j(x) = \kappa(2^{-j}x) = \exp\left\{-\rho \frac{x}{2^j} \left(\frac{x}{2^j} + 1\right)\right\}, \quad (3)$$

where $j \in \mathbb{N} := \{-1, 0, 1, 2, \dots\}$. The corresponding spherical ridgelets with their energy spread around the great circle supported by \mathbf{v} is given by:

$$\psi_{j,\mathbf{v}}(\mathbf{u}) = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \lambda_n (\kappa_{j+1}(n) - \kappa_j(n)) P_n(\mathbf{u} \cdot \mathbf{v}), \quad (4)$$

where P_n denotes the Legendre polynomial of order n and $\kappa_{-1}(n) = 0, \forall n$ and

$$\lambda_n = \begin{cases} 2\pi(-1)^{n/2} \frac{1 \cdot 3 \cdots (n-1)}{2 \cdot 4 \cdots n}, & \text{if } n \text{ is even} \\ 0, & \text{if } n \text{ is odd.} \end{cases} \quad (5)$$

To obtain a finite overcomplete dictionary, we restrict the values of the resolution index j to a finite set $\{-1, 0, 1, \dots, J\}$, where J defines the highest level of “detectable” (high frequency) signal details. Additionally, the set of all possible *orientations* $\mathbf{v} \in \mathbb{S}^2$ of spherical ridgelets needs to be discretized as well. To find a proper discretization scheme, we first note that the construction in (4) suggests that the bandwidth of the spherical ridgelets (and therefore the dimensionality of the functional space they belong to) increases proportionally to 2^j . The space of spherical harmonics of degree n has a dimension of $(n+1)^2$. Similarly, we define the number of ridgelet orientations at resolution j to be equal to $M_j = (2^{j+1}m_0 + 1)^2$, with m_0 being the smallest spherical order resulting in $\kappa_0(m_0) \leq \varepsilon$ for some predefined $0 < \varepsilon \ll 1$ (e.g. $\varepsilon = 10^{-6}$). Consequently, for each j , a total of M_j orientations $\{\mathbf{v}_j^i\}_{i=1}^{M_j}$ are chosen so that the overcomplete SR dictionary is given by

$\Psi = \{\psi_{j,\mathbf{v}_j^i} | j = -1, 0, 1, \dots, J, i = 1, 2, \dots, M_j\}$. Note that, Ψ consists of a total of

$M = \sum_{j=-1}^J (2^{j+1}m_0 + 1)^2$ spherical ridgelets. To slightly simplify our notation, in what follows, the spherical ridgelets in Ψ will be indexed as $\psi_m(\mathbf{u})$, with $m = 1, 2, \dots, M$ being a combined index accounting for both different resolutions and orientations. In this and previous works (Michailovich et al., 2011), the number of spherical ridgelet orientations were predefined with $m_0 = 4$, resulting in $M_{-1} = 16$, $M_0 = 49$ and $M_1 = 169$ ridgelets spanning the resolution levels $j = -1$, $j = 0$ and $j = 1$, respectively. Thus, the total number of spherical ridgelets used in the reconstruction was equal to 234, which has been shown to provide robust estimates of the signal (Michailovich et al., 2011).

Given a sampling set of K diffusion-encoding orientations $\{\mathbf{u}_k\}_{k=1}^K$, one can use (4) to compute the values of the spherical ridgelets over the sampling set¹. The resulting values can be stored into a $K \times M$ matrix A defined as

$$A = \begin{bmatrix} \psi_1(\mathbf{u}_1) & \psi_2(\mathbf{u}_1) & \cdots & \psi_M(\mathbf{u}_1) \\ \psi_1(\mathbf{u}_2) & \psi_2(\mathbf{u}_2) & \cdots & \psi_M(\mathbf{u}_2) \\ \cdots & \cdots & \cdots & \cdots \\ \psi_1(\mathbf{u}_K) & \psi_2(\mathbf{u}_K) & \cdots & \psi_M(\mathbf{u}_K) \end{bmatrix}. \quad (6)$$

Then, for a given vector \mathbf{s} of diffusion measurements acquired at a fixed (yet arbitrary) spatial location, one can obtain a sparse solution vector \mathbf{c} by solving the l_1 minimization problem in (2), i.e.

$$\min \|\mathbf{c}\|_1, \text{ s.t. } \|\mathbf{Ac} - \mathbf{s}\|^2 \leq \varepsilon. \quad (7)$$

¹Since the definition in (4) involves an infinite summation, the latter needs to be truncated to render the computations practical. In practice, we truncate the summation to index n_{max} for which the magnitude of the summand drops below 10^{-9} .

4. Methods

In this work, it is assumed that the dMRI data is measured along the same K gradient directions $\{\mathbf{u}_k\}$ at several b -values. This allows for estimating the signal attenuation with increasing b -values along each gradient direction. In particular, we propose to model the radial signal decay of $E(\mathbf{q})$ ($E(\mathbf{q}) = S(\mathbf{q})/S(0)$), where the signal $S(\mathbf{q})$ is assumed to be normalized by the signal at $b = 0$ i.e. $S(0)$ using a monotonically decreasing function given by

$$f_{\alpha,\beta}(b) = (1+b^\alpha)^{-\beta}, \quad \alpha \geq 0, \quad \beta \geq 0, \quad (8)$$

which is related to the cumulative probability distribution function of the Burr distribution (Tadikamalla, 1980). A similar power-law based three-dimensional representation of q -space using Wishart and Gamma distributions has also been used in earlier works of (Jian and Vemuri, 2007; Scherrer et al., 2013). Both these works, represent the signal using probability distributions defined on the space of positive definite tensors. In the present work, however, we integrate the one-dimensional power-law decay function $f_{\alpha,\beta}$ with spherical ridgelets to characterize the diffusion signal in q -space. Note that the proposed function is bounded in the range $f_{\alpha,\beta}(b) \in [0, 1]$ and can model mono-exponential as well as multi-exponential decay for appropriate choice of α and β (Figure 1). This property of $f_{\alpha,\beta}(b)$ is in contrast to the formulations in (Rathi et al., 2011; Ozarslan et al., 2008; Cheng et al., 2010), where severe noise at high b -values can lead to “bumps” or non-monotonic signal values (which cannot arise from biological tissue of the brain). The proposed function, on the other hand, guarantees a monotonically decreasing estimate of the diffusion signal despite severe noise at high b -values. Additionally, this function has only two free parameters, which makes the estimation procedure more robust and stable. On the other hand, if an interleaved sampling scheme is used to measure data on several q -shells as in Ye et al. (2012), one would require at-least 2 shells to have the same set of gradient directions. Note that this function $f_{\alpha,\beta}(b)$ is a one-dimensional function (in b variable) and hence we need to estimate the free parameters α_k and β_k for each gradient direction \mathbf{u}_k . For better intuition and visualization, the b -value b is scaled by dividing the actual b -value by 1000 (so $b = 1$ (arbitrary units) corresponds to a b -value of 1000 s/mm^2 in Figure 1). Note that, changing the units of b in this manner does not affect the algorithm and is only meant for visualization purposes. Not scaling b will give the exact same results albeit with an appropriate scaling of α and β .

We use spherical ridgelets to estimate the diffusion signal at each b -value from undersampled measurements. Thus, given the measurements s_k^i corresponding to gradient direction \mathbf{u}_k and b -value b_i , we estimate the spherical ridgelet coefficients \mathbf{c}^i for each b_i . We should note that, since each b -value corresponds to different frequency content (with higher b -values accounting for higher frequency components of the signal), the dictionary of spherical ridgelets can be further adjusted for each spherical shell. Such an adjustment can be done by appropriately estimating the value of parameter ρ in (3) for each b_i , thereby resulting in a set of shell-specific ridgelet dictionaries A_i (more on this in section 4.4).

Consequently, a separate ℓ_1 -norm minimization problem (7) needs to be solved for each b_i to obtain coefficients \mathbf{c}^i .

Our goal is to estimate the diffusion signal in a continuous and consistent fashion in the entire q -space. However, the above formulation provides the signal estimate in the spherical domain using spherical ridgelets, and in the radial domain using the radial decay function (RDF) $f(b)$ for each gradient direction independently. In this case, the estimate using the SR basis ensures smooth signal in the spherical domain (at a given b -value), while the estimate using RDF ensures smooth signal decay with b -value at a given direction \mathbf{u}_k . Due to severe noise present in dMRI images, the solution using each of these methods will be different and to obtain a consistent estimate of the diffusion signal in the entire q -space, we need to attain a consensus between the estimates provided by both these methods.

4.1. Mathematical formulation

One way to obtain a consistent solution is using ADMM, which is a numerical scheme designed to solve composite optimization problems (Boyd et al., 2011). We propose the following cost function to obtain an estimate of the signal at each voxel:

$$\min_{\{\mathbf{c}^i, \alpha_k, \beta_k\}} \left\{ \sum_{i=1}^{N_b} \frac{1}{2} \|A^i \mathbf{c}^i - \mathbf{s}^i\|^2 + \lambda_1 \sum_{i=1}^{N_b} \|\mathbf{c}^i\|_1 + \lambda_2 \sum_{k=1}^K \|\mathbf{f}(\alpha_k, \beta_k) - \mathbf{s}_k\|^2 \right\}, \quad (9)$$

where $\mathbf{s}^i \in \mathbb{R}^K$ denotes a (column) vector of K measurements acquired at b -value b_i , while $\mathbf{s}_k \in \mathbb{R}^{N_b}$ is a (column) vector of N_b diffusion signals measured along direction \mathbf{u}_k over N_b different b -value shells². Also, the vector $\mathbf{f}(\alpha_k, \beta_k) \in \mathbb{R}^{N_b}$ in (9) represents the values of the radial model (8) corresponding to different q -shells and it is defined as

$$\mathbf{f}(\alpha_k, \beta_k) = [(1+b_1^{\alpha_k})^{-\beta_k}, (1+b_2^{\alpha_k})^{-\beta_k}, \dots, (1+b_{N_b}^{\alpha_k})^{-\beta_k}]^T.$$

We rewrite (9) in a more compact form by using the following notations. First, let $\mathbf{c} \in \mathbb{R}^{N_b M}$ be a vector obtained through column stacking the vectors of ridgelet coefficients \mathbf{c}_i

corresponding to each of the N_b b -shells used for data acquisition, i.e. $\mathbf{c} = [\mathbf{c}_1^T, \dots, \mathbf{c}_{N_b}^T]^T$.

Similarly, let $\mathbf{s} \in \mathbb{R}^{N_b K}$ be a vector obtained through column stacking the associated HARDI signals \mathbf{s}^i , i.e. $\mathbf{s} = [\mathbf{s}_1^T, \dots, \mathbf{s}_{N_b}^T]^T$. Then, in the absence of model errors and measurement noises, the above vectors are related through $\mathbf{s} = \mathbf{A}\mathbf{c}$, where \mathbf{A} is an $N_b K \times N_b M$ block-diagonal matrix composed of the ridgelet matrices A^i as given by

²Note that, we have used superscripts to denote variables related to the spherical domain (ridgelets), and subscripts to index variables corresponding to the radial term.

$$\mathbf{A} = \begin{bmatrix} A^1 & 0 & \dots & 0 \\ 0 & A^2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & A^{N_b} \end{bmatrix}.$$

Subsequently, the minimization problem in (9) can be redefined as

$$\begin{aligned} \min_{\mathbf{c}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \left\{ \frac{1}{2} \|\mathbf{A}\mathbf{c} - \mathbf{s}\|^2 + \lambda_1 \|\mathbf{c}\|_1 + \lambda_2 \|\mathbf{f}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \mathbf{s}\|^2 \right\}, \\ \text{s.t. } \mathbf{A}\mathbf{c} - \mathbf{f}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0, \end{aligned} \quad (10)$$

where $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_K]$ and $\boldsymbol{\beta} = [\beta_1, \dots, \beta_K]$ are the vectors of parameters of our radial decay model corresponding to various diffusion-encoding directions \mathbf{u}_k , while $\mathbf{f}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ can be viewed as a concatenation of vectors $\mathbf{f}(\alpha_k, \beta_k)$, with $k = 1, 2, \dots, K$ *permuted* to comply with the order of coordinates in \mathbf{s} . Note that in (10) the original minimization (9) has been enhanced with an additional equality constraint that *couples* the variables \mathbf{c} , $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$, thereby aiming to produce a consensus estimate, for which the ridgelet and the radial decay models agree over the set of sampling points.

The optimization in (10) is performed at a *single* voxel, and therefore it ignores the correlation between the diffusion signal observed at neighboring voxels. To incorporate this prior knowledge, we add another term to our cost function as described in the next section.

4.2. Spatial regularization via total-variation (TV) minimization

Let $\Omega \in \mathbb{R}^3$ be a discrete (rectangular) volume of interest of size $|\Omega| = N_x \times N_y \times N_z$ over which the diffusion signal is observed. The reconstruction technique presented in the previous section estimates *both* the ridgelet coefficients and the radial model parameters at each voxel $\mathbf{r} \in \Omega$. Thus, introducing an explicit dependency on \mathbf{r} , (10) can be re-written as

$$\begin{aligned} \min_{\mathbf{c}(\mathbf{r}), \boldsymbol{\alpha}(\mathbf{r}), \boldsymbol{\beta}(\mathbf{r})} \left\{ \frac{1}{2} \|\mathbf{A}\mathbf{c}(\mathbf{r}) - \mathbf{s}(\mathbf{r})\|^2 + \lambda_1 \|\mathbf{c}(\mathbf{r})\|_1 + \lambda_2 \|\mathbf{f}(\boldsymbol{\alpha}(\mathbf{r}), \boldsymbol{\beta}(\mathbf{r})) - \mathbf{s}(\mathbf{r})\|^2 \right\}, \\ \text{s.t. } \mathbf{A}\mathbf{c}(\mathbf{r}) - \mathbf{f}(\boldsymbol{\alpha}(\mathbf{r}), \boldsymbol{\beta}(\mathbf{r})) = 0, \end{aligned} \quad (11)$$

to compute estimates of $\mathbf{c}(\mathbf{r})$, $\boldsymbol{\alpha}(\mathbf{r})$ and $\boldsymbol{\beta}(\mathbf{r})$ for each $\mathbf{r} \in \Omega$, independently. However, the diffusion signal is known to exhibit a certain degree of spatial-domain regularity, which can be exploited to further improve the accuracy and robustness of the estimation process. In particular, it is reasonable to assume the diffusion-encoded images to be of bounded variation which can be incorporated into our estimation procedure by minimizing the total-variation (TV) semi-norm of the estimated signal. We note that TV-based regularization has been used by several authors in the context of CS for data restoration from incomplete measurements (Osher et al., 2005; Yin et al., 2008; Ma et al., 2008; Michailovich et al., 2011).

The vector $\mathbf{Ac}(\mathbf{r}) \in \mathbb{R}^{Nb}$ represents the reconstructed MSDI signal at spatial location $\mathbf{r} \in \Omega$. When viewed as a function of the spatial coordinate \mathbf{r} , the k -th component of this vector forms a 3D image volume I_k , which approximates the diffusion-encoded image corresponding to the gradient direction \mathbf{u}_k (for a particular b-value). For a collection of image volumes $\mathcal{I} = \{I_k\}_{k=1}^K$, the TV-norm of \mathcal{I} can be written as a sum of individual TV-norms of each of the image volumes I_k (Blomgren and Chan, 1998)

$$\|\mathcal{I}\|_{TV} = \sum_k \|I_k\|_{TV} = \sum_k \sum_{\mathbf{t} \in \Omega} \left[\sum_{\mathbf{p} \in \mathcal{C}(\mathbf{t})} |I_k(\mathbf{t}) - I_k(\mathbf{p})|^2 \right]^{1/2}, \quad (12)$$

where $\mathcal{C}(\mathbf{t}) = \{(i-1, j, k), (i, j-1, k), (i, j, k-1)\}$.

Note that the optimization problem in (11) has been formulated for the diffusion signal observed at a single voxel, while the definition of the TV norm in (12) requires integration of image information over the whole domain Ω . Therefore, to incorporate the TV regularization into the proposed estimation framework, the optimization problem in (11) needs to be properly modified. To this end, we introduce the following notations. First, when viewed as a function of \mathbf{r} , the array of vectors $\mathbf{s}(\mathbf{r})$ can be considered to be a (discrete) vector field \mathbf{s} . We denote the set of all such vector fields by \mathcal{V} and endow it with the ℓ_2 norm defined in a standard manner as

$$\|\mathbf{s}\|_{\mathcal{V},2} = \left[\sum_{\mathbf{r} \in \Omega} \|\mathbf{s}(\mathbf{r})\|_2^2 \right]^{1/2} = \left[\sum_{\mathbf{r} \in \Omega} \sum_{k=1}^K |s_k(\mathbf{r})|^2 \right]^{1/2},$$

with $s_k(\mathbf{r})$ denoting the k -th component of $\mathbf{s}(\mathbf{r})$. Analogously, the array of vectors $\mathbf{c}(\mathbf{r})$ can also be considered to be a (discrete) vector field $\mathbf{c} \in \mathcal{U}$, with \mathcal{U} denoting the set of all such vector fields. Moreover, it is also possible to equip \mathcal{U} with the ℓ_1 norm of the form

$$\|\mathbf{c}\|_{\mathcal{U},1} = \sum_{\mathbf{r} \in \Omega} \|\mathbf{c}(\mathbf{r})\|_1 = \sum_{\mathbf{r} \in \Omega} \sum_{k=1}^{MN_b} |c_k(\mathbf{r})|,$$

where $c_k(\mathbf{r})$ is the k -th component of $\mathbf{c}(\mathbf{r})$.

Finally, let $\mathcal{A}: \mathcal{U} \rightarrow \mathcal{V}$ be a linear operator that, for every $\mathbf{r} \in \Omega$, maps $\mathbf{c}(\mathbf{r})$ to $\mathbf{Ac}(\mathbf{r})$, thereby suggesting a data formation model of the form $\mathbf{s} = \mathcal{A}\mathbf{c} + \eta$, with $\eta \in \mathcal{V}$ accounting for both measurement noise and modelling errors. Consequently, the above definitions can be used to rewrite our optimization problem as given by

$$\begin{aligned} \min_{\mathbf{c}, \alpha, \beta} \{ & \|\mathcal{A}\mathbf{c} - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_1 \|\mathbf{c}\|_{\mathcal{U},1} + \lambda_2 \|\mathcal{F}(\alpha, \beta) - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_3 \|\mathcal{A}\mathbf{c}\|_{TV} \}, \\ \text{s.t. } & \mathcal{A}\mathbf{c} - \mathcal{F}(\alpha, \beta) = 0, \end{aligned} \quad (13)$$

where \mathcal{F} aggregates the values of the radial decay function $f(\boldsymbol{\alpha}, \boldsymbol{\beta})(\mathbf{r})$ into a vector field (similar to that of \mathbf{s}). Evidently, this cost functional has two data fidelity terms (one for spherical and one for radial domain) and two regularization terms (sparsity in the SR domain and piecewise smoothness in the spatial domain). The regularization parameters $\lambda_1, \lambda_2, \lambda_3$ determine the relative importance given to the data fitting terms versus the sparsity and TV regularization terms. Next, we derive a computationally efficient and optimal solution for solving the cost function in (13) based on the ADMM (Boyd et al., 2011; Yin et al., 2008).

4.3. Optimization Algorithm

The minimization problem in (13) can be efficiently solved using the ADMM algorithm (Boyd et al., 2011; Yin et al., 2008). This optimization technique allows one to supersede a complex optimization problem by a sequence of easier problems, which often admit a closed form solution. Following the methodology developed in (Goldstein and Osher, 2009), we introduce an auxiliary optimization variable \mathbf{z} , leading to the following modified cost function

$$\min_{\mathbf{c}, \boldsymbol{\alpha}, \boldsymbol{\beta}} \{ \|\mathcal{A}\mathbf{c} - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_1 \|\mathbf{c}\|_{\mathcal{W},1} + \lambda_2 \|\mathcal{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_3 \|\mathbf{z}\|_{TV} \}, \quad (14)$$

$$\text{s.t. } \mathcal{A}\mathbf{c} - \mathcal{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0, \quad \mathcal{A}\mathbf{c} - \mathbf{z} = 0. \quad (15)$$

Then, starting from an arbitrary p_f^0 and p_z^0 at iteration $t = 0$, the above (constrained) minimization problem can be solved iteratively by the augmented Lagrange multiplier method, which updates the optimal values of \mathbf{c} , \mathbf{z} , $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$ according to

$$(\mathbf{c}^{t+1}, \boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}, \mathbf{z}^{t+1}) = \arg \min_{\mathbf{c}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{z}} \left\{ \frac{1}{2} \|\mathcal{A}\mathbf{c} - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_1 \|\mathbf{c}\|_{\mathcal{W},1} + \lambda_2 \|\mathcal{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \mathbf{s}\|_{\mathcal{V},2}^2 + \lambda_3 \|\mathbf{z}\|_{TV} + \frac{\delta_f}{2} \|\mathcal{A}\mathbf{c} - \mathcal{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) + p_f^t\|_{\mathcal{V},2}^2 + \frac{\delta_z}{2} \|\mathcal{A}\mathbf{c} - \mathbf{z} + p_z^t\|_{\mathcal{V},2}^2 \right\}, \quad (16)$$

followed by updating the Lagrange multipliers p_f and p_z as given by

$$p_z^{t+1} = p_z^t + (\mathcal{A}\mathbf{c}^{t+1} - \mathbf{z}^{t+1}) \quad (17)$$

$$p_f^{t+1} = p_f^t + (\mathcal{A}\mathbf{c}^{t+1} - \mathcal{F}(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1})) \quad (18)$$

for some $\delta_f > 0, \delta_z > 0$. Moreover, in the case of ADMM, the concurrent minimization over all the primary and auxiliary variables in (16) is replaced by sequential minimization with respect to each of the variables separately, which (after a few algebraic simplifications) results in the following set of update equations for \mathbf{c} , \mathbf{z} , $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$:

$$\mathbf{c}^{t+1} = \arg \min_{\mathbf{c}} \left\{ \left\| \mathcal{A} \mathbf{c} - \frac{\mathbf{s} + \delta_f (\mathcal{F}(\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t) - p_f^t) + \delta_z (\mathbf{z}^t - p_z^t)}{1 + \delta_f + \delta_z} \right\|_{\mathcal{V},2}^2 + \lambda_1 \|\mathbf{c}\|_{\mathcal{W},1} \right\} \quad (19)$$

$$(\boldsymbol{\alpha}^{t+1}, \boldsymbol{\beta}^{t+1}) = \arg \min_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \left\{ \frac{1}{2} \left\| \mathcal{F}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \frac{2\lambda_2 \mathbf{s} + \delta_f (\mathcal{A} \mathbf{c}^{t+1} - p_f^t)}{2\lambda_2 + \delta_f} \right\|_{\mathcal{V},2}^2 \right\} \quad (20)$$

$$\mathbf{z}^{t+1} = \arg \min_{\mathbf{z}} \left\{ \|\mathbf{z} - (\mathcal{A} \mathbf{c}^{t+1} - p_z^t)\|_{\mathcal{V},2}^2 + \frac{\lambda_3}{\delta_z} \|\mathbf{z}\|_{TV} \right\} \quad (21)$$

Thus, the composite cost function in (13) can be solved by iteratively solving (19), (20) and (21) followed by updating the Lagrange multipliers p_z^t and p_f^t as given by (17) and (18), respectively. A typical stopping criteria is to check if p_f^t and p_z^t have “stopped changing”, i.e. $\|p_f^{t+1} - p_f^t\|_{\mathcal{V},2}^2 \leq \varepsilon_1$ and $\|p_z^{t+1} - p_z^t\|_{\mathcal{V},2}^2 \leq \varepsilon_2$, for some user defined choice of $\varepsilon_1, \varepsilon_2$. The ADMM algorithm is guaranteed to converge to a global minimum of the cost function as shown in (Goldstein and Osher, 2009; Deng and Yin, 2012)

Some important points to note from the above update equations are: (i) The iterative update for \mathbf{c}^t in (19) is separable in the spatial coordinate, i.e. the l_1 -minimization can be performed at each voxel independently. (ii) The update equation (20) is separable in both the spatial coordinate and gradient directions, i.e. the least squares fitting of parameters α, β can be carried out for each gradient direction and at each spatial location \mathbf{r} independently (iii) The update equation for \mathbf{z} (21) is separable in the gradient directions and b-values, i.e., solving the TV de-noising problem in (21) amounts to solving a total of KN_b scalar TV de-noising problems on 3-D image volumes corresponding to K different directions of diffusion encoding and N_b q -shells. It is important to note that the separable structure of the update equations (21)–(23) suggests significant numerical advantages through the use of parallel computing.

Intuitively, an optimal solution to (13) is expected to attain a consensus between sparse ridgelet approximation in the spherical domain, a monotone signal decay in the radial domain, and piecewise smoothness of 3-D diffusion volumes in the spatial domain. Thus, the optimal solution would be the one which satisfies the regularity imposed for each of the three terms (sparsity in spherical, monotonic decay in radial and minimum TV-norm in the spatial domain). The update equations for p_f^t and p_z^t exemplify these points where the “error” between the solution from different domains is fed-back into the system, eventually converging to zero.

Finally, we note that, if one sets $\lambda_3 = 0$ in (13), it would lead to solving the cost function without the TV-norm. In such a scenario, we do not have a variable \mathbf{z} and the minimization proceeds by iteratively updating \mathbf{c}^t and $\boldsymbol{\alpha}^t, \boldsymbol{\beta}^t$, with $\delta_z = 0$. Thus, the above update equations

are a solution to the most generalized problem and specific update equations can be obtained by setting the appropriate parameters to zero.

4.4. Practical considerations

While the overall minimization procedure is described in the previous section, there are a few parameters that need to be set for a successful execution of the proposed framework. Specifically, we first describe a way to set the parameter ρ in the construction of the spherical ridgelet basis (3). While, this parameter was set to 0.5 in our earlier works (Michailovich et al., 2008, 2011), here, we use a line search minimization procedure to set the value of ρ_i while constructing the basis A^i for each value of b_i . Given a representative sample of the white matter diffusion signal for a given b -value, we find the optimal ρ that minimizes the error in fitting the signal. The representative sample could be obtained from *in vivo* scans, from phantom data set or using a synthetic model of the data. In this work, we used the data from a single fiber voxel from a phantom data set that closely approximates the diffusion signal *in vivo*.

In our experimental study, we set $\lambda_1 = \lambda_2 = \lambda_3 = 1$, thereby assigning for the terms corresponding to spherical (SR basis), radial (radial decay function) and TV-norm (spatial smoothness). Although the ADMM algorithm is guaranteed to converge for an arbitrary choice of parameters $\delta_f > 0$ and $\delta_z > 0$, in our numerical experiments, the fastest convergence was observed for $\delta_f = \delta_z = 0.5$. In this work, we used the homotopy based algorithm of (Asif and Romberg, 2009) to solve the l_1 -minimization problem in (19), although other choices are also available (Becker et al., 2009; Beck and Teboulle, 2009; Candes et al., 2008). To solve the TV-minimization, we used the fixed point algorithm of (Chambolle, 2004) and utilized the nonlinear least squares fitting routine 'lsqnonlin' from the MATLAB (Mathworks Inc.) toolbox to carry out the minimization in (20).

5. Experiments

To evaluate the proposed algorithm, we constructed a physical phantom with diffusion properties similar to that of human brain tissue. The phantom consisted of a spherical spindle wound with $15\mu\text{m}$ polyfil fibers to obtain a 45° crossing angle. A detailed description of how the phantom was made is given in (Moussavi-Biugui et al., 2011). The phantom was scanned using a Siemens 3T scanner at a spatial resolution of $2\text{mm} \times 2\text{mm} \times 7\text{mm}$, so that the crossing region lay in the center of the axial slice. We used a larger slice thickness to ensure that all crossing fibers contributed to the diffusion signal (the crossing region is a few millimeters deep as seen in Figure 2). The average fractional anisotropy (FA) for the single fiber region at a b -value of 1000 was 0.78. This is very close to the FA value estimated (0.79) in *in-vivo* human brain as reported by several authors.

The main goal of this work has been to show that the diffusion signal in the entire q -space can be recovered using a sub-critical (incomplete) set of measurements. We should point out that once the signal is recovered, any type of diffusion model can be used to analyze the resulting data. For example, the signal could be used to compute the orientation distribution function (ODF), the EAP, return-to-origin probability, kurtosis, fast and slow diffusing

compartments, free-water compartment, or any of the novel measures proposed in (Özarslan et al., 2013).

To perform a quantitative comparison of our signal reconstruction technique, we needed a “gold standard” data set. To obtain one, we acquired 10 separate scans (repetitions) of the phantom with the following acquisition parameters : b -values of {1000, 2000, 3000, 4000, 5000} s/mm^2 , 81 gradient directions per shell (a total of 405 measurements) and lines $b = 0$ image. The 10 scans were then averaged to obtain the “gold standard” data with high SNR (signal-to-noise ratio).

In order to quantify the effect of noise as well as to ensure that we do not introduce bias in the signal reconstruction process by subsampling the high-angular resolution data (subsampling could potentially smooth the data, which does not occur in realistic scenarios), we acquired the following data sets to test the signal reconstruction accuracy of our algorithm: For each of the following number of gradient directions $K = \{16, 20, 24, 26, 30, 36, 42, 60\}$, we acquired MSDI data over a total of $5q$ -shells corresponding to $b = \{1000, 2000, 3000, 4000, 5000\} s/mm^2$. Thus, the first data sample consisted of 16 gradient directions at each of the b -values (a total of 80 measurements). Further, 5 repetitions were acquired for each of these data samples to test the effect of noise on signal reconstruction quality. Signal reconstruction was performed using each of these 5 repetitions separately, and error metrics were computed from the same. The average SNR for each b -value shell was $b = \{15.5, 10.60, 8.02, 7.1, 6.7\}$ and the overall average over all gradient directions and b -values was 9.5 ($SNR = s/\sigma$, where s is the mean signal and σ is the standard deviation). The average SNR was estimated by computing the mean of the SNR estimated for each gradient direction and at each voxel location. Given that the SNR is quite low, the results obtained in this work show the effectiveness of the proposed method in the presence of high noise.

5.1. Comparison metrics

To test the signal reconstruction ability of the proposed algorithm, we used a number of quantitative metrics which are described below.

- 1. Normalized mean squared error (NMSE) in signal reconstruction:** Let $\mathbf{s}(\mathbf{r})$ denote the reconstructed signal at location \mathbf{r} , and $\mathbf{s}_g(\mathbf{r})$ represent the “gold standard” signal obtained as described earlier. Then, the NMSE error N_s in signal reconstruction can be computed as

$$N_s = \frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} \frac{\|\mathbf{s}(\mathbf{r}) - \mathbf{s}_g(\mathbf{r})\|^2}{\|\mathbf{s}_g(\mathbf{r})\|^2}, \quad (22)$$

where $|\Omega|$ is the number of voxels in the region of interest.

- 2. Angular error:** The peak of the fiber orientation distribution function (fODF) indicates the most likely orientation of the fibers at each voxel location. The peaks of the ODF are often used by deterministic tractography methods to map the anatomical connectivity of the brain. Thus, accurate estimation of the fODF peaks

is imperative for proper tracing and study of white matter fiber bundles. For the case of the phantom data set, the ground truth in the crossing region is known (45°), whereas for the in-vivo case, the ground truth was computed from the fully sampled data set. For the phantom data, our main goal was to visualize if any particular sampling scheme (data sets) overestimated or underestimated the angle between the two fiber bundles and as such, we computed the deviation of the estimated fiber orientation from the ground truth of 45° . Consequently, we report the estimated angle between the fiber bundles in the crossing region of the phantom data.

Since ground truth is not known for in-vivo data, the angular error AE over a pre-defined region of interest Ω can be computed using

$$AE = \frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} \sum_{h=1}^{n_p} |\arccos(\mathbf{v}_g^h(\mathbf{r}) \cdot \mathbf{v}^h(\mathbf{r}))|, \quad (23)$$

where \mathbf{v}_g is the “gold standard” orientation and \mathbf{v} is the estimated orientation. At each voxel \mathbf{r} , we expect one or more significant orientation peaks of the fODF given by n_p (number of peaks), and we sum over all peaks to compute the final error at each voxel.

3. **Percentage of Incorrect Peaks:** Computation of the angular error as defined in (23) is feasible only if the number of detected fODF peaks are the same in both, the test data and the gold standard data. However, in many cases, the recovered signal may miss a peak or generate spurious peaks that are false positives. Thus, knowledge of the percentage of incorrect peaks in addition to the angular error can provide the proper perspective on the accuracy of fODF estimated from the recovered signal. We compute the fraction of incorrect peaks using the following expression

$$I_p = \frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} \mathcal{B}(n_{pg} - n_p)(\mathbf{r}), \quad (24)$$

where \mathcal{B} is a binary function whose value is 1 if $n_{pg} = n_p$ and 0 otherwise (n_{pg} is the number of peaks at a voxel in the gold standard data and n_p is the number of peaks in the recovered data). In other words, we compute the percentage of voxels with incorrect number of peaks and denote it by I_p .

4. **NMSE error in estimation of return-to-origin probability P_o :** Another quantity that can be directly computed from the signal is the return-to-origin probability or zero displacement probability P_o (Wu et al., 2008). Mathematically, P_o can be defined as $P_o = \int S(\mathbf{q}) d\mathbf{q}$, where $S(\mathbf{q})$ is the diffusion signal at q -value \mathbf{q} . In practice, however, we restrict the integrand to some maximum q -value (or b -value) and perform numerical summation of the signal values over this domain. NMSE error (N_p) in the estimation of P_o is computed as follows:

$$N_p = \frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} \frac{\|P_o(\mathbf{r}) - P_{og}(\mathbf{r})\|^2}{\|P_{og}(\mathbf{r})\|^2}, \quad (25)$$

where P_{og} is the return-to-origin probability of the gold standard. A significant error in the computation of P_o would imply that the radial decay is not accurately estimated by a given algorithm.

We use all of these metrics to quantify the reconstruction quality of the data using the proposed method as well as the l_1 3D-SHORE method (Cheng et al., 2010). An important point to note is that, we compute N_s and N_p at all points at which the gold standard data is available. Thus, essentially, for sparse methods, we compute the error in extrapolating the data from incomplete samples and lower b -value data to higher b -value data (i.e., the entire sampled q -space).

5.2. Results: Phantom data

Our goal is to determine the performance of the proposed method while varying the number of MSDI measurements. Specifically, the community is still debating whether to use more b -value shells or use more gradient directions per shell. Thus, for example, if the scan time allows to acquire only 60 measurements, should we use three b -values with 20 directions each or alternatively use two b -values with 30 directions each? Further, which set of b -values to use is also not clear. The purpose of the extensive validation we perform in this work is an attempt to answer these questions using the proposed algorithm.

From the phantom data set acquired above, we create the following multi-shell test data sets to validate our method: For each of the following number of gradient directions $K = \{16, 20, 24, 26, 30, 36, 42, 60, 81\}$, we generated test data with the following b -values: (i) $b = \{1000, 3000\}$ (2 q -shells), (ii) $b = \{1000, 2000, 3000\}$ (3 q -shells), (iii) $b = \{1000, 2000, 3000, 4000\}$ (4 b -shells), (iv) $b = \{1000, 2000, 3000, 4000, 5000\}$ (5 b -shells). This corresponded to a total of 36 data sets with differing levels of sparsity. Note that, for each of these 36 test sets, we had 5 acquisitions (a total of 180 test sets) which we used to compute the average error metrics described in the previous section.

5.2.1. Signal Reconstruction Error—Figure 3 shows the fODF computed from the “gold standard” phantom data (81 gradient directions per shell, 5 b -value shells, and 10 repetitions) using the proposed method (spherical ridgelets with radial decay) and using the SHORE basis. For the proposed method, we used the solid angle formulation of (Tristán-Vega and Westin, 2011) to estimate the fODF. In the case of the SHORE basis, we set the scaling parameter $\psi = 500$ after running several experiments (this choice of ψ gave the best results). Note that, this choice of the parameter was optimal for the region that contained the fibers, but not for the noisy isotropic region surrounding the fiber crossing, and hence no fODFs are seen in this region in Figure 3(b). In particular, in this region, the SHORE algorithm did not converge. However, our region of interest is really the single fiber and the crossing fiber region and hence we compute all our error metrics only in this region.

Figure 4 shows the values of NMSE obtained with each of the tested methods for all b -values in use. An important point to note here is that, we extrapolate the signal reconstructed using the incomplete data to all points in the q -space at which the gold standard data is available. For example, for the case of $K = 16$ and two b -values $b = \{1000, 3000\}$, we reconstruct the data at each of the 5 q -shells and 81 gradient directions per shell that is available for the gold standard data. Thus, NMSE is computed over a wide range of q -space at which one can hope to measure the signal with descent SNR.

Figure 4 shows that the NMSE in signal reconstruction N_s using the SR basis is less than 2% for all data sets. Further, the NMSE for $K = 20$ and $b = \{1000, 2000, 3000\}$ is 1.8%, whereas for the full data set with all directions and b -values it is around 1.2%, which we believe is not substantially different. On the other hand, the errors using the SHORE basis is less than 5% only if all b -values are used (yellow line in Figure 4b), indicating that it is not particularly suitable for extrapolating data in the radial q -space domain.

5.2.2. Angular Error—Next, we show the angular error between the principal diffusion directions obtained from the peaks of the fODFs. A maxima of the fODF is considered a valid peak, if its value is at-least 20% of the highest peak in that fODF. The fODF peaks were extracted from a spherical tessellation consisting of 2562 points sampled on the sphere.

The results in Figure 5(a) show that the angular error using the SR basis is never more than 5° even with $K = 16$ and two b -values (a total of 32 measurements), while it is about 2° for $K = 20$ and three b -values. With increasing number of measurements, the error is within 1° . A slight increase in error (0.42°) for $K = 26$ compared to $K = 24$ is partly due to noise in the data as well as the fact that for $K = 26$ the measurements were not quite uniformly spread. In general though, the error decreases with increasing number of measurements. However, note that the angle between the fibers is underestimated when using two b -values, and overestimated with three or more b -values. In the case of the SHORE basis, the angular error is much higher with at-least a 6° error in all cases.

5.2.3. Incorrect Peak Detection—We should note that, the above angular errors were computed only at voxels within the region of crossing fibers, where two fODF peaks were detected. However, it is also important to know the percentage of voxels for which an incorrect number of peaks were found. For the crossing region in the phantom data, we computed the average number of voxels (normalized) with incorrect peaks (averaged over the crossing region and the 5 repetitions).

From Figures 5(a) and 6(a), it becomes clear, that even though the angular error is very low for $K = 24$ and two b -values (green line), in 10% of the voxels the fiber crossing is not detected. However, with $K = 20$ and three b -values, the algorithm performs reasonably well with about $I_p = 5\%$ and angular error of 2° . On the other hand, the SHORE basis does not find the crossing fibers in a majority of the voxels unless at-least four b -values are used with $K = 42$. Note that, in this case as well, about 15% of the voxels show only one peak, when there exists two peaks in the gold standard data. Thus, the proposed method significantly outperforms the SHORE-based reconstruction.

Figure 7 shows the computed fODFs from signal recovered using the proposed method for two different test sets ($K = 20$, 3 b -values, and $K = 24$, 3 b -values), as well as using the SHORE basis ($K = 20$, 3 b -values, and $K = 42$ and 4 b -values). Note that, using the proposed basis, the voxels where we do not find the second peak are at the interface of single and crossing fiber regions (voxels which are circled). Further, these voxels have a very small component that corresponds to the second peak. On the other hand, using the SHORE basis and with $K = 20$ and 3 b -values, most of the voxels do not show the crossing fiber region. However, a descent signal recovery with the SHORE basis can be seen with $K = 42$ and 4 b -values (a total of 168 measurements).

5.2.4. Error in P_o —Figure 8 shows the NMSE in the estimation of P_o , the return-to-origin probability. The proposed method estimates this quantity to a very high degree of accuracy for all data sets (error is less than 0.1%). The best case scenario for the SHORE basis is an error of about 1% when four b -values were used.

5.3. Results: In vivo

To test the proposed method *in vivo*, we scanned a young adult (age 25 years) in a 3T Siemens scanner with the following acquisition parameters: (i) spatial resolution: $2.5 \text{ mm} \times 2.5 \text{ mm} \times 2.5 \text{ mm}$, (ii) TE = 128 ms, TR = 10.4s, (iii) 60 gradient directions for each of the following four b -values $b = \{900, 2000, 3600, 5600\} \text{ s/mm}^2$ and 10 $b = 0$ acquisitions (a total of 250 acquisitions). To obtain a sparse set of samples, we subsampled this data set by fitting an order 10 spherical harmonic basis (without any regularization), which can characterize the high-frequency noise in the data and then evaluating the values at the desired location for each b -value. Figure 9 shows the estimated fODF's using the proposed method with all measurements (gold standard) and with $K = 20$ directions and three b -values $b = \{900, 2000, 3600\}$. The return-to-origin probability P_o is also shown in the background for each of the fODF glyphs.

For better visualization, a zoomed-in version of the glyphs where a 3-way fiber crossing is visible is shown in Figure 10. This is the region in the centrum semiovale, where the corpus callosum, the cortico-spinal tract and the superior-longitudinal fasciculus (SLF) cross each other. Note the similarity of the fODF's in the crossing areas, using the sparsely sampled data ($K = 20$ and three b -values) and the gold standard.

On a quantitative note, for the white matter region shown in Figure 9, the average angular error of the proposed algorithm with respect to the gold standard was 3.6° . The error in estimation of the signal (NMSE - N_s) was 0.008, whereas the NMSE in estimation of P_o was 0.0001.

Figure 11 shows the fODFs estimated using the SHORE basis for the same region shown in Figure 9. Note that, the SHORE basis fails to detect the crossing at a few voxel locations even using the gold standard data. On the other hand, the fODF from the sub-critically sampled data misses a lot of secondary peaks.

We also computed the return-to-origin probability P_o in the entire brain using the gold standard as well as the sub-critically sampled data ($K = 20$, three b -values). Figure 12 shows

an axial slice of the computed P_o as well as a difference image scaled to better visualize the areas of differences (NMSE N_p). The average NMSE N_p computed over the entire brain region was 1.63×10^{-4} , which indicates the accuracy of our estimation approach using incomplete measurements.

Finally, in Figure 13, we show three representative slices in the brain where we computed the NMSE error in signal estimation (N_s). The error was computed between the gold standard and sub-critically sampled data with 60 measurements ($K = 20$, three b -values). The average NMSE over the entire brain was 0.0068, however for better visualization we scaled the error values so that a value of 100 corresponded to an error of 0.1. Note that the signal estimation error is relatively large mostly in the areas near the skull, and significantly lower in the gray and white matter areas.

In Figure 14, we computed the mean kurtosis at each voxel using data with 60 measurements, by computing the kurtosis along each gradient direction followed by taking the mean as is shown in Jensen et al. (2005). Kurtosis measures the amount of non-Gaussian diffusion, which is highest in tightly bound fiber bundles like the corpus callosum (blue regions in the figure) and is close to zero in the CSF areas (red) where diffusion is almost unrestricted.

6. Discussion

In this work, we proposed a novel framework for recovering dMRI data in the entire q -space from very few measurements. In particular, the proposed method is designed to use diffusion samples acquired at multiple q -shells by extending the spherical ridgelets basis to be used within a MSDI framework. We proposed a new monotonically decreasing radial decay function to be used within the estimation framework, which allows accurate modeling of single and multi-exponential attenuation of diffusion signal as a function of b -value. A novel application of the ADMM algorithm along with a total-variation based spatial regularity term was used for consistent and spatially regular recovery of the diffusion signal from a sub-critically sampled diffusion measurements.

Through a comprehensive analysis of a physical phantom data set, and from in-vivo experiments, it is evident that at-least 60 measurements are required for a satisfactory recovery of the dMRI data that is equivalent to a fully sampled data set with 240 measurements. Some of the conclusions one can draw from the experimental analysis done earlier are: (i) For the same number of measurements, spreading the acquisition over three q -shells is slightly better than over two q -shells. For example, from Figures 4, 5 and 6, it is clear that 60 measurements with three b -values (20 measurements per shell) is a slightly better than spreading them over two b -values (30 measurements per shell). In making this choice, we preferred to have fewer incorrect peaks and sacrifice a little of the angular accuracy (about $1 - 2^\circ$). (ii) Depending on the amount of time one is willing to spend on a dMRI scan, it is evident that at-least 60 measurements are needed for a good recovery of the dMRI signal (in the entire q -space, and not just in the lower b -value range) using the proposed method. However, as expected, the more the measurements, the lower the error (although, the gain in accuracy with more measurements is not substantial). (iii) For the

phantom and in-vivo data, the measurements estimated beyond a b -value of 3000 was quite accurate, indicating that acquisition of up-to a maximum b -value of 3000 – 3500 is probably enough for recovering information in the remaining q -space³.

We should note that, we used one particular method to recover the fODFs in this work. However, there exist a lot of different formulations which could be used to obtain the *fiber* ODF's, the EAP as well as other measures of diffusion. These methods could be used to analyze the data once the full diffusion signal is recovered from sparse measurements using the method proposed in this work. One of the limitations of the present study is that we do not know how accurately the proposed method preserves long range fiber connectivity. A detailed investigation of this aspect as well as its effect on network connectivity measures will form part of our future work.

Acknowledgments

This work has been supported by NIH grants: R01MH097979 (Rathi), R01MH074794 (Westin), P41RR013218, P41EB015902 (Kikinis, Core PI: Westin), and Swedish research grant VR 2012-3682 (Westin).

References

- Asif, MS.; Romberg, JK. Dynamic updating for l1 minimization. 2009. CoRR abs/0903.1443
- Assaf Y, Freidlin R, Rohde G, Basser P. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*. 2004; 52 (5):965–978. [PubMed: 15508168]
- Assemlal H, Tschumperlé D, Brun L. Efficient computation of pdf-based characteristics from diffusion mr signal. *MICCAI*. 2008; 2008:70–78. [PubMed: 18982591]
- Assemlal HE, Tschumperlé D, Brun L, Siddiqi K. Recent advances in diffusion mri modeling: Angular and radial reconstruction. *Medical image analysis*. 2011; 15 (4):369–396. [PubMed: 21397549]
- Barpoutis A, Vemuri B, Forder J. Fast displacement probability profile approximation from hardi using 4th-order tensors. *ISBI*. 2008:911–914.
- Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*. 2009; 2 (1):183–202.
- Becker, S.; Bobin, J.; Candes, E. NESTA: A fast and accurate first-order method for sparse recovery. 2009. Arxiv preprint arXiv 904
- Blomgren P, Chan TF. Color tv: total variation methods for restoration of vector-valued images. *Image Processing, IEEE Transactions on*. 1998; 7 (3):304–309.
- Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*. 2011; 3 (1):1–122.
- Candès E, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*. 2006; 52 (2):489–509.
- Candes E, Wakin M, Boyd S. Enhancing sparsity by reweighted L1 minimization. *Journal of Fourier Analysis and Applications*. 2008; 14 (5):877–905.
- Candes EJ, Eldar YC, Needell D, Randall P. Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*. 2011; 31 (1):59–73.
- Chambolle A. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*. 2004; 20 (1–2):89–97.

³Assuming that the narrow pulse approximation is satisfied and one does not see the diffraction pattern in clinical *in vivo* scans

- Cheng J, Ghosh A, Jiang T, Deriche R. Model-Free and Analytical EAP Reconstruction via Spherical Polar Fourier MRI. *MICCAI*. 2010:590–597. [PubMed: 20879279]
- Clark CA, Le Bihan D. Water diffusion compartmentation and anisotropy at high b values in the human brain. *MRM*. 2000; 44 (6):852–859.
- Cohen Y, Assaf Y. High b-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues—a technical review. *NMR in Biomedicine*. 2002; 15 (7–8):516–542. [PubMed: 12489099]
- Deng W, Yin W. On the global and linear convergence of the generalized alternating direction method of multipliers. Tech rep, DTIC Document. 2012
- Descoteaux M, Deriche R, Bihan D, Mangin J, Poupon C. Multiple q-shell diffusion propagator imaging. *Medical Image Analysis*. 2010
- Donoho D. Compressed sensing. *IEEE Transactions on Information Theory*. 2006; 52 (4):1289–1306.
- Duarte-Carvajalino, JM.; Lenglet, C.; Ugurbil, K.; Moeller, S.; Carin, L.; Sapiro, G. A framework for multi-task bayesian compressive sensing of dw-mri. *Proceedings of the CDMRI MICCAI workshop; Nice, France*. 2012. p. 1-13.
- Feinberg D, Moeller S, Smith S, Auerbach E, Ramanna S, Glasser M, Miller K, Ugurbil K, Yacoub E. Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging. *PLoS One*. 2010; 5(12)
- Freiman, M.; Afacan, O.; Mulkern, RV.; Warfield, SK. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*. Springer; 2013. Improved multi b-value diffusion-weighted mri of the body by simultaneous model estimation and image reconstruction (smeir); p. 1-8.
- Goldstein T, Osher S. The split bregman method for l_1 -regularized problems. *SIAM Journal on Imaging Sciences*. 2009; 2 (2):323–343.
- Gramfort A, Poupon C, Descoteaux M. Sparse DSI: Learning DSI structure for denoising and fast imaging. *MICCAI*. 2012:288–296. [PubMed: 23286060]
- Hosseinbor AP, Chung MK, Wu Y-C, Alexander AL. Bessel fourier orientation reconstruction (bfor): An analytical diffusion propagator reconstruction for hybrid diffusion imaging and computation of q-space indices. *NeuroImage*. 2012
- Jensen J, Helpert J, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*. 2005; 53 (6):1432–1440. [PubMed: 15906300]
- Jian B, Vemuri B. Multi-fiber Reconstruction from Diffusion MRI Using Mixture of Wisharts and Sparse Deconvolution. *Lecture Notes in Computer Science*. 2007; 4584:384.
- Landman BA, Bogovic JA, Wan H, ElShahaby FEZ, Bazin PL, Prince JL. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI. *NeuroImage*. 2012; 53:21752186.
- Ma, S.; Yin, W.; Zhang, Y.; Chakraborty, A. An efficient algorithm for compressed mr imaging using total variation and wavelets. *Computer Vision and Pattern Recognition, 2008. CVPR 2008; IEEE Conference on. IEEE; 2008*. p. 1-8.
- Merlet S, Caruyer E, Deriche R. Parametric dictionary learning for modeling eap and odf in diffusion mri. *MICCAI*. 2012:10–17. [PubMed: 23286108]
- Merlet SL, Deriche R. Continuous diffusion signal, eap and odf estimation via compressive sensing in diffusion mri. *Medical image analysis*. 2013; 17 (5):556–572. [PubMed: 23602920]
- Michailovich O, Rathi Y. Fast and accurate reconstruction of HARDI data using compressed sensing. *MICCAI*. 2010; 6361:606–613.
- Michailovich O, Rathi Y, Dolui S. Spatially regularized compressed sensing for high angular resolution diffusion imaging. *TMI*. 2011; 30 (5):1100–1115.
- Michailovich O, Rathi Y, Shenton ME. On approximation of orientation distributions by means of spherical ridgelets. *Int Symp on Biomedical Imaging*. 2008:939–942.
- Moussavi-Biugui A, Stieltjes B, Fritzsche K, Semmler W, Laun FB. Novel spherical phantoms for q-ball imaging under in vivo conditions. *Magnetic Resonance in Medicine*. 2011; 65 (1):190–194. [PubMed: 20740652]

- Mulkern RV, Haker SJ, Maier SE. On high b-value diffusion imaging in the human brain: ruminations and experimental insights. *Magnetic resonance imaging*. 2009; 27 (8):1151–1162. [PubMed: 19520535]
- Mulkern RV, Vajapeyam S, Robertson RL, Caruso PA, Rivkin MJ, Maier SE. Biexponential apparent diffusion coefficient parametrization in adult vs newborn brain. *Magnetic resonance imaging*. 2001; 19 (5):659–668. [PubMed: 11672624]
- Osher S, Burger M, Goldfarb D, Xu J, Yin W. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling & Simulation*. 2005; 4 (2):460–489.
- Özarslan E, Koay C, Basser P. Simple harmonic oscillator based estimation and reconstruction for one-dimensional q-space mr. *ISMRM*. 2008; 16:35.
- Özarslan E, Koay CG, Shepherd TM, Komlosh ME, rfanoöglu MO, Pierpaoli C, Basser PJ. Mean apparent propagator (map) mri: A novel diffusion imaging method for mapping tissue microstructure. *NeuroImage*. 2013
- Rathi Y, Michailovich O, Setsompop K, Bouix S, Shenton M, Westin C-F. Sparse multi-shell diffusion imaging. *MICCAI*. 2011:58–65. [PubMed: 21995013]
- Scherrer, B.; Schwartzman, A.; Taquet, M.; Prabhu, SP.; Sahin, M.; Akhondi-Asl, A.; Warfield, SK. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer; 2013. Characterizing the distribution of anisotropic micro-structural environments with diffusion-weighted imaging (diamond); p. 518-526.
- Schwarcz A, Bogner P, Meric P, Correze JL, Berente Z, Pál J, Gallyas F, Doczi T, Gillet B, Beloeil JC. The existence of bi-exponential signal decay in magnetic resonance diffusion-weighted imaging appears to be independent of compartmentalization. *Magnetic resonance in medicine*. 2004; 51 (2):278–285. [PubMed: 14755652]
- Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magnetic Resonance in Medicine*. 2011; 67 (5):1210–1224. [PubMed: 21858868]
- Shenton M, Hamoda H, Schneiderman J, Bouix S, Pasternak O, Rathi Y, Vu MA, Purohit M, Helmer K, Koerte I, et al. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. *Brain imaging and behavior*. 2012; 6 (2):137–192. [PubMed: 22438191]
- Stejskal E, Tanner J. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*. 1965; 42 (1):288.
- Tadikamalla PR. A look at the burr and related distributions. *International Statistical Review/Revue Internationale de Statistique*. 1980:337–344.
- Thomason ME, Thompson PM. Diffusion imaging, white matter, and psychopathology. *Annual review of clinical psychology*. 2011; 7:63–85.
- Tristán-Vega, A.; Westin, C-F. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*. Springer; 2011. Probabilistic odf estimation from reduced hardi data with sparse regularization; p. 182-190.
- Tuch D, Reese T, Wiegell M, Wedeen V. Diffusion MRI of complex neural architecture. *Neuron*. 2003; 40:885–895. [PubMed: 14659088]
- Wedeen V, Hagmann P, Tseng W, Reese T, Weisskoff R. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*. 2005; 54 (6): 1377–1386. [PubMed: 16247738]
- Wu Y, Alexander A. Hybrid diffusion imaging. *NeuroImage*. 2007; 36 (3):617–629. [PubMed: 17481920]
- Wu YC, Field AS, Alexander AL. Computation of diffusion function measures. *Medical Imaging, IEEE Transactions on*. 2008; 27 (6):858–865.
- Ye W, Portnoy S, Entezari A, Blackband SJ, Vemuri BC. An efficient interlaced multi-shell sampling scheme for reconstruction of diffusion propagators. *Medical Imaging, IEEE Transactions on*. 2012; 31 (5):1043–1050.
- Ye, W.; Portnoy, S.; Entezari, A.; Vemuri, BC.; Blackband, SJ. Box spline based 3d tomographic reconstruction of diffusion propagators from mri data. *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE; 2011. p. 397-400.

- Yin W, Osher S, Goldfarb D, Darbon J, Bregman J. Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing. *SIAM J Imaging Sci.* 2008; 1 (1):143–168.
- Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage.* 2012

Highlights of the paper

1. We present a novel method for reconstruction of the diffusion MRI signal from sparse measurements
2. A novel function to model the bi-exponential decay of the signal in the radial q -space.
3. Novel algorithmic derivation for spatially smooth signal reconstruction.
4. Extensive experiments done on phantom and in-vivo data.
5. Our results show very low error in the recovered signal using at-least 60 samples.

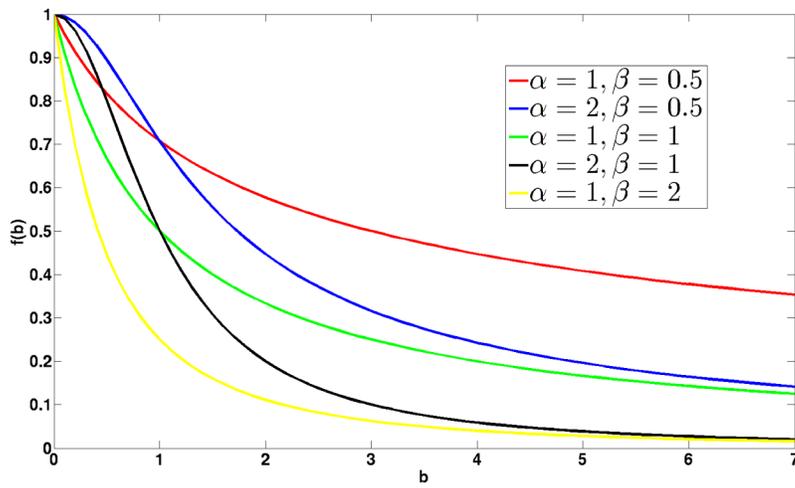


Figure 1. Radial decay function $f(b)$ that we use to model signal attenuation with increasing b -values. Various curves for different parameter settings α, β are shown.

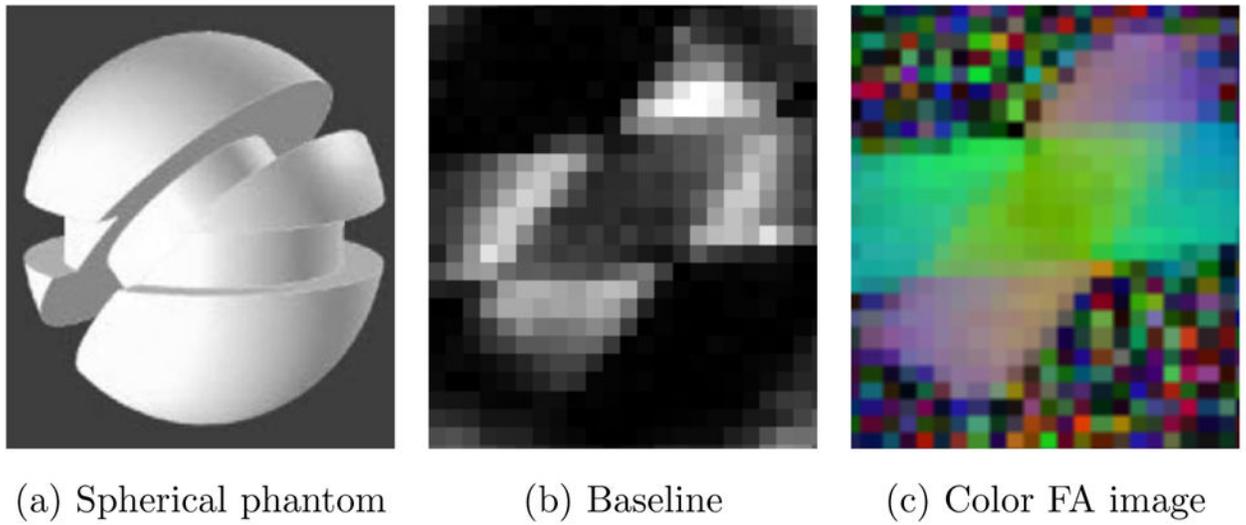


Figure 2.
The spherical phantom with a 45° crossing angle is shown along with the baseline and color coded FA images.

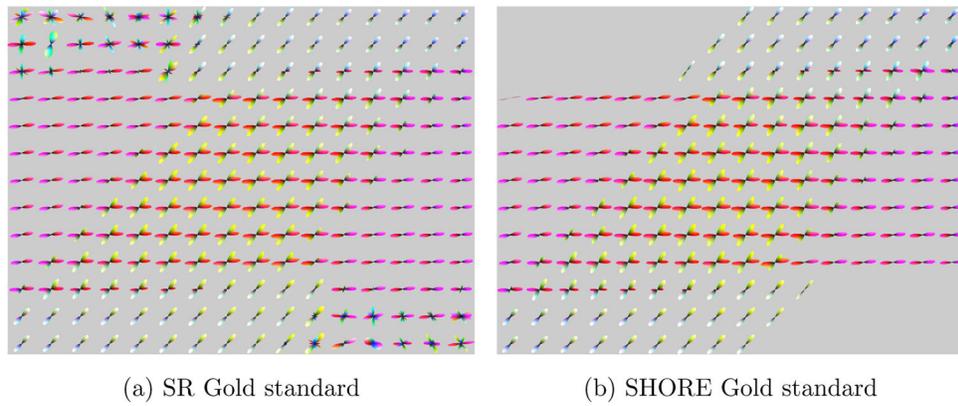


Figure 3. fODF computed from the “gold standard” data using the proposed method (SR basis) and the SHORE basis.

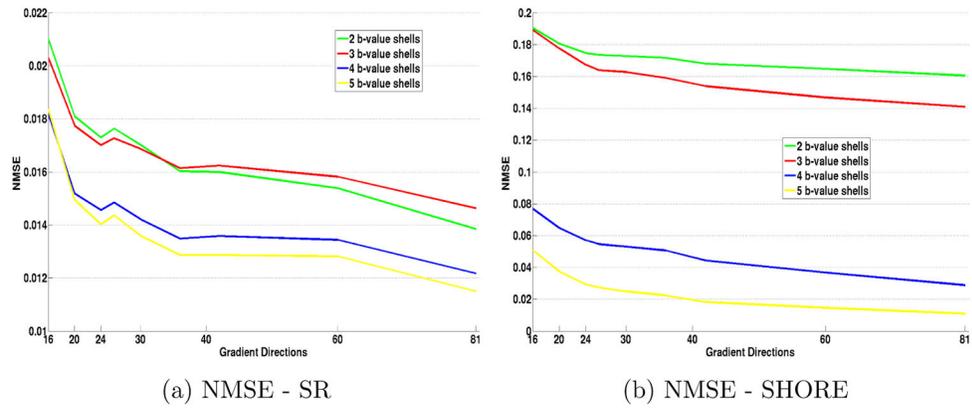
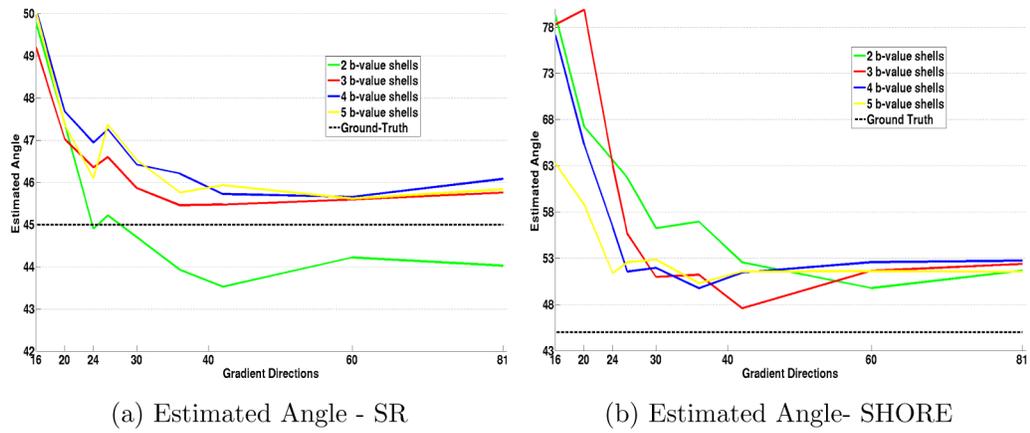


Figure 4.

NMSE in signal reconstruction N_s computed using the proposed SR basis (a) and the SHORE basis (b). x-axis is the number of gradient directions K per b -value, while the colored plots represent the errors for different number of b -values used. Notice that, for the SR basis, the signal reconstruction error is significantly lower (by a factor of 10) compared to the SHORE basis.



(a) Estimated Angle - SR

(b) Estimated Angle- SHORE

Figure 5.

Estimated fODF angle (in degrees) between the fiber bundles using the proposed SR basis (a) and the SHORE basis (b). x-axis is the number of gradient directions K per b-value shell, while the colored plots are the errors for different number of b-value shells used. Note that, the y-axis in both plots has a very different range.

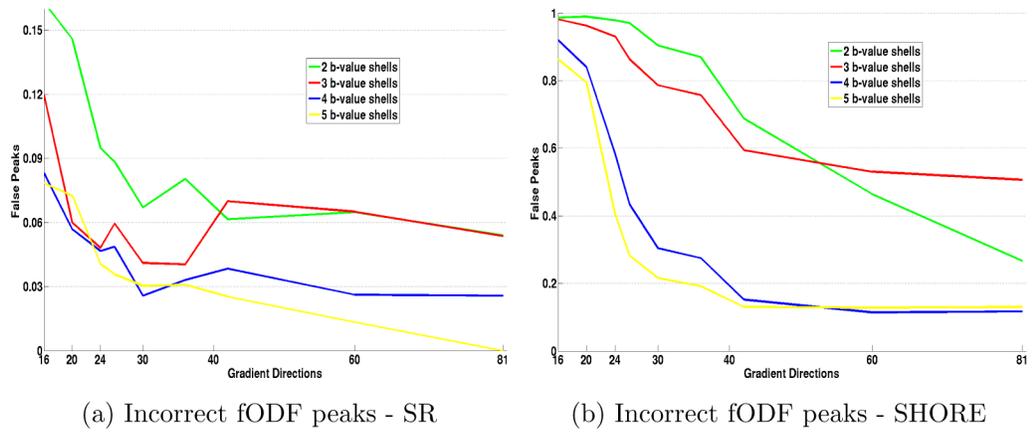


Figure 6. Incorrect number of fiber peaks I_p detected (normalized to 1) using the proposed SR basis (a) and the SHORE basis (b). Note that, the y-axis has a very different range in both the plots.

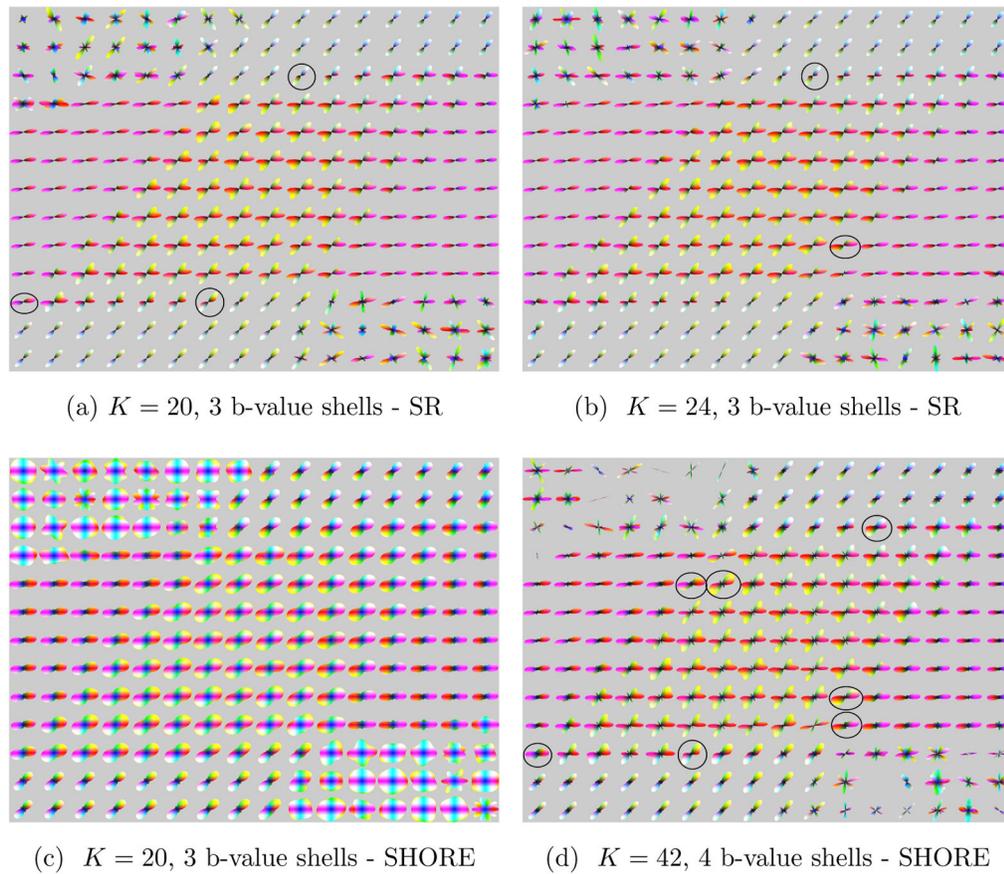


Figure 7.

Computed fODF's from dMRI data recovered using the proposed method and from a total of 60 measurements ($K = 20$, $b = \{1000, 2000, 3000\}$) and 72 measurements ($K = 24$, $b = \{1000, 2000, 3000\}$). Also shown is the fODF field recovered using the SHORE basis with $K = 20$ and three b -values (where most voxels do not show a crossing) as well as $K = 42$ and four b -values. Some of the voxels, where the the second peak was not found are circled.

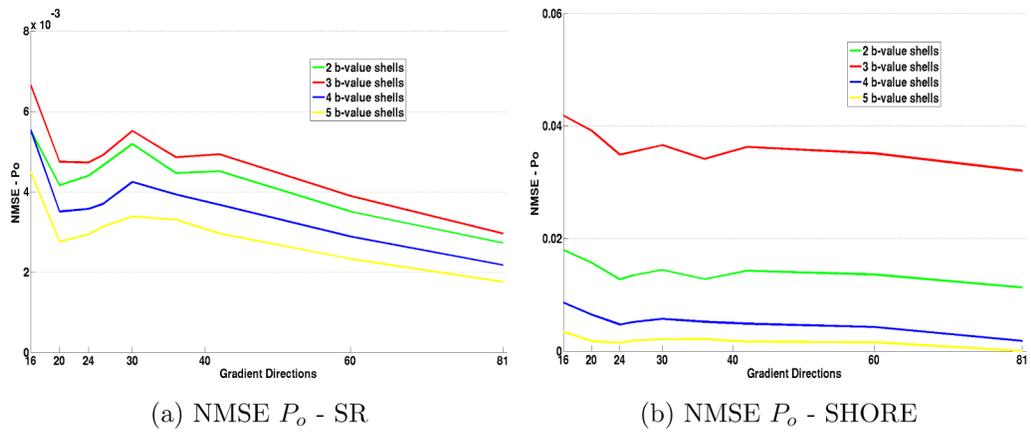
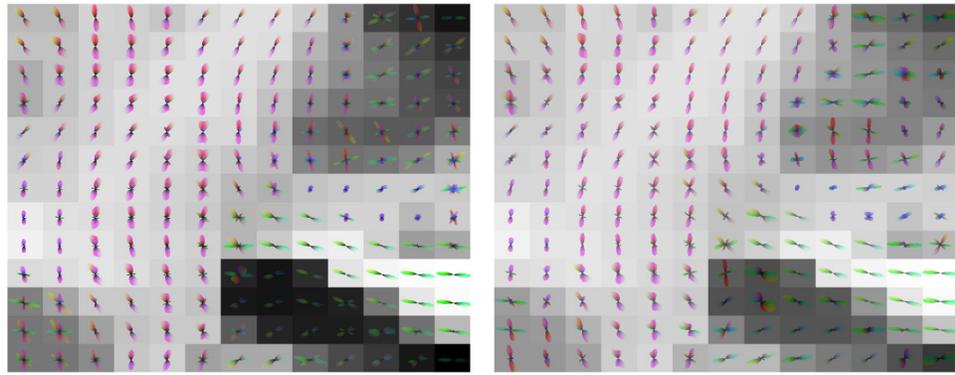
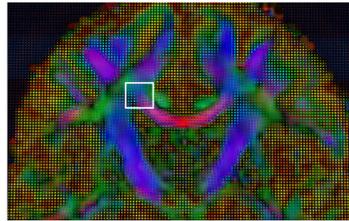


Figure 8. NMSE error in return-to-origin probability using the proposed method and SHORE. Note that, for the proposed method, the error is less than 0.1%, but much higher for the SHORE basis.



(a) Gold standard - SR

(b) $K = 20$, 3 b-value shells - SR

(c) Color FA image

Figure 9.

fODF glyphs are shown for the rectangular region in the color FA image. Return-to-origin probability P_o was used as the background for the glyphs, with (a) being the gold standard computed from all the 240 measurements, while (b) was obtained with 60 measurements.

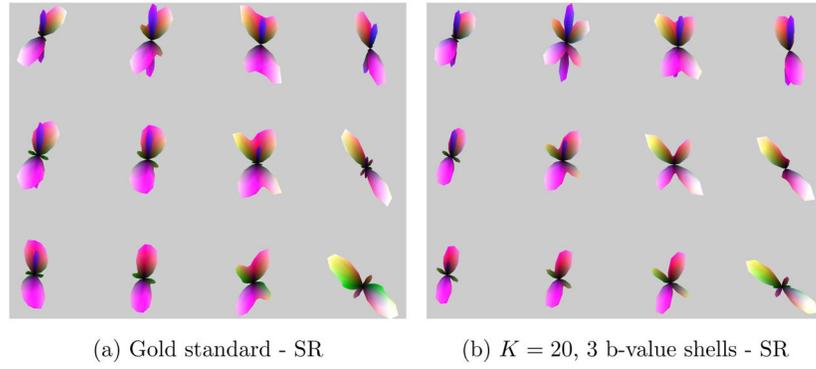


Figure 10. Two and three fiber crossing of the fODFs estimated in the centrum semiovale region using (a) the gold standard, and (b) sub-critically sampled data (60 measurements).

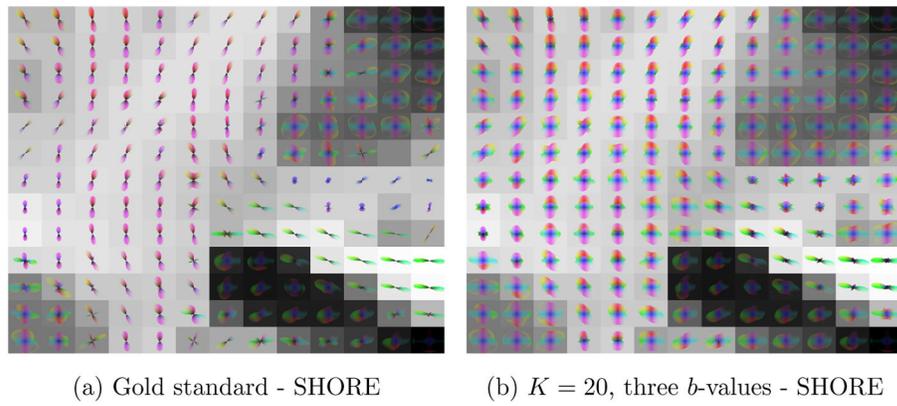
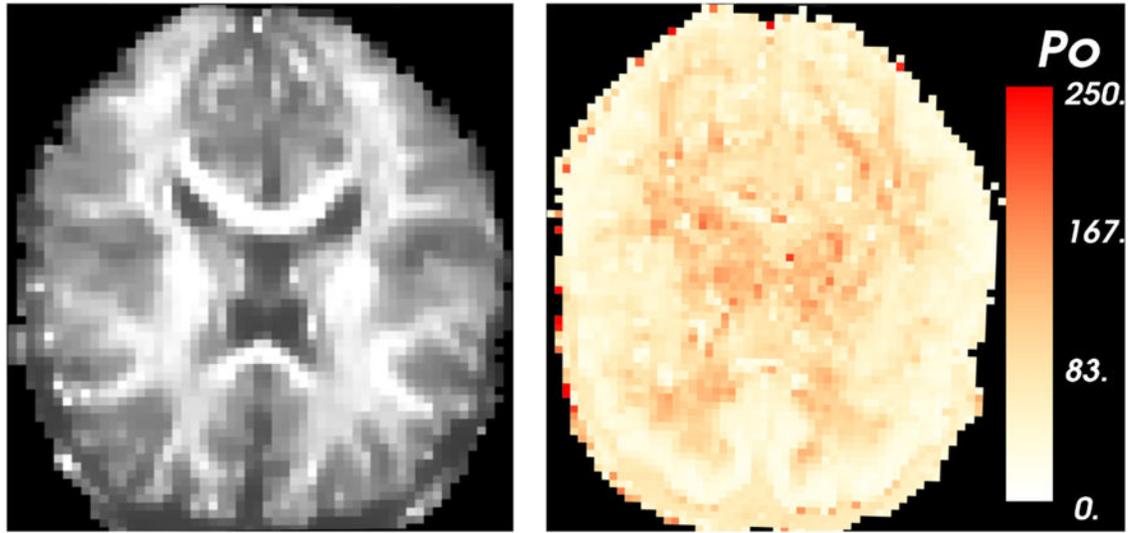


Figure 11. fODF glyphs estimated using the SHORE basis with the gold standard and sub-critically sampled data.



(a) Axial slice P_o

(b) Difference image

Figure 12.

(a) shows an axial slice of the estimated return-to-origin probability P_o and (b) shows the NMSE error in estimation of P_o . The image in (b) was scaled for better visualization so that the highest error of 0.01 was scaled to 250.

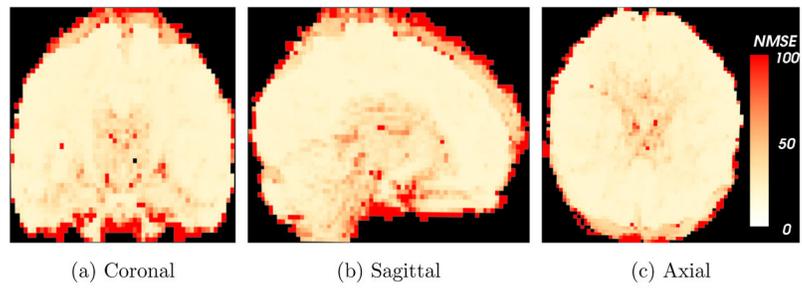
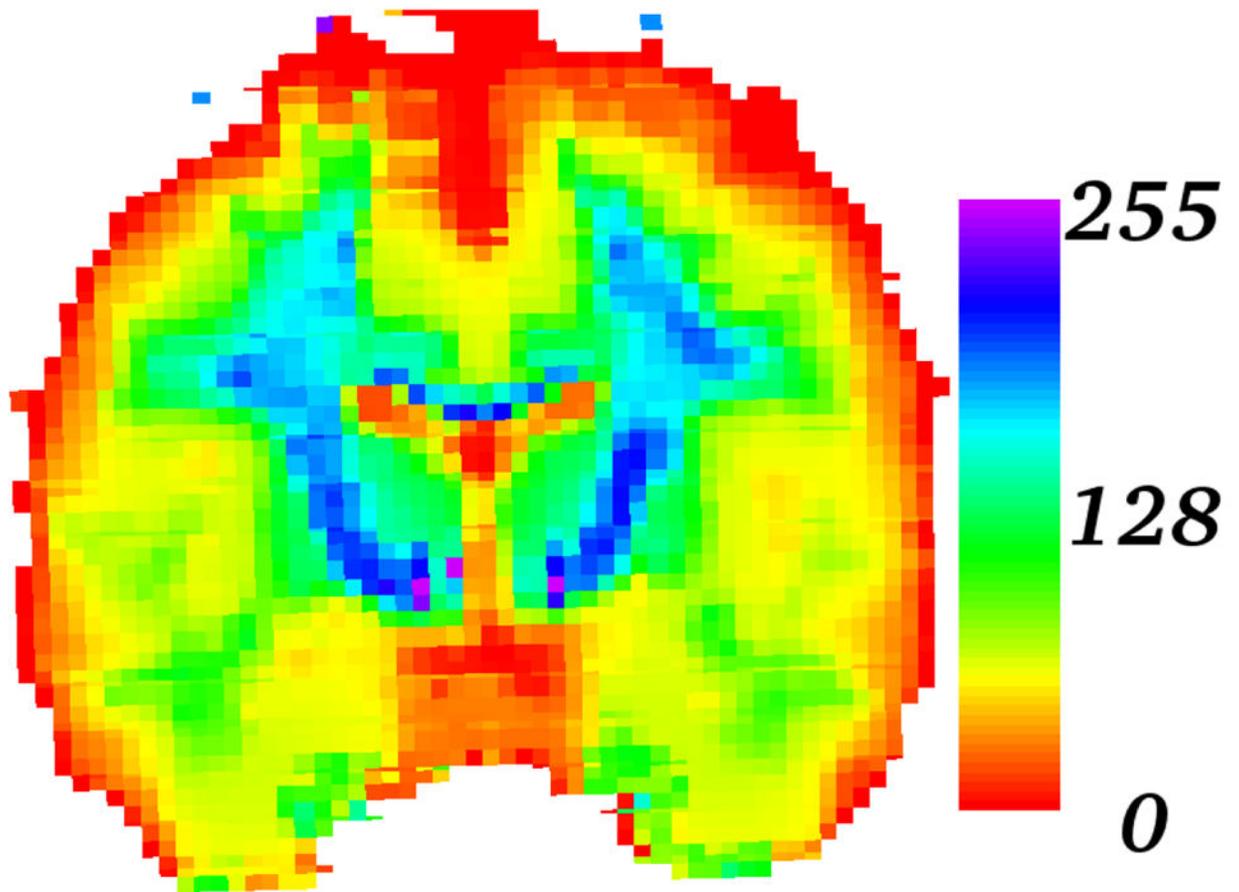


Figure 13.

Three representative slices showing the NMSE error in signal estimation using the sub-critically sampled data with $K = 20$ and three b -values. The error images were scaled for better visualization so that the highest error of $N_s = 0.1$ was scaled to 100.



(a) Mean kurtosis

Figure 14.

Figure shows a coronal slice with mean kurtosis computed at each voxel. Highest kurtosis is seen in the tightly bundled white matter region (blue regions), while it decreases in gray matter (yellow) and CSF areas (red). In this figure, the kurtosis values were scaled to a range of 0 to 255 for better visualization.