HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2021 May 27.

Published in final edited form as:

Neuroimage. 2014 November 01; 101: 750–764. doi:10.1016/j.neuroimage.2014.07.062.

Non-Negative Spherical Deconvolution (NNSD) for Estimation of Fiber Orientation Distribution Function in Single-/Multi-Shell Diffusion MRI

Jian Cheng[a,*], **Rachid Deriche**[b], **Tianzi Jiang**[c], **Dinggang Shen**[a], **Pew-Thian Yap**[a,*]

[a]Department of Radiology and BRIC, The University of North Carolina at Chapel Hill, USA

[b]Athena Project-Team, INRIA Sophia Antipolis – Méditerranée, France

[c]Center for Computational Medicine, LIAMA, Institute of Automation, Chinese Academy of Sciences, China

Abstract

Spherical Deconvolution (SD) is commonly used for estimating fiber Orientation Distribution Functions (fODFs) from diffusion-weighted signals. Existing SD methods can be classified into two categories: 1) Continuous Representation based SD (CR-SD), where typically Spherical Harmonic (SH) representation is used for convenient analytical solutions, and 2) Discrete Representation based SD (DR-SD), where the signal profile is represented by a discrete set of basis functions uniformly oriented on the unit sphere. A feasible fODF should be non-negative and should integrate to unity throughout the unit sphere $\mathbb{S}^2$. However, to our knowledge, most existing SH-based SD methods enforce non-negativity only on discretized points and not the whole continuum of $\mathbb{S}^2$. Maximum Entropy SD (MESD) and Cartesian Tensor Fiber Orientation Distributions (CT-FOD) are the only SD methods that ensure non-negativity throughout the unit sphere. They are however computational intensive and are susceptible to errors caused by numerical spherical integration. Existing SD methods are also known to overestimate the number of fiber directions, especially in regions with low anisotropy. DR-SD introduces additional error in peak detection owing to the angular discretization of the unit sphere. This paper proposes a SD framework, called non-negative SD (NNSD), to overcome all the limitations above. NNSD is significantly less susceptible to the false-positive peaks, uses SH representation for efficient analytical spherical deconvolution, and allows accurate peak detection throughout the whole unit sphere. We further show that NNSD and most existing SD methods can be extended to work on multi-shell data by introducing a three-dimensional fiber response function. We evaluated NNSD in comparison with Constrained SD (CSD), a quadratic programming variant of CSD, MESD, and an $L1$-norm regularized non-negative least-squares DR-SD. Experiments on synthetic and real single-/multi-shell data indicate that NNSD improves estimation performance in terms of mean

*Corresponding author: jian_cheng@med.unc.edu (Jian Cheng), ptyap@med.unc.edu (Pew-Thian Yap).

difference of angles, peak detection consistency, and anisotropy contrast between isotropic and anisotropic regions.

Keywords

Spherical Deconvolution; Diffusion MRI; Fiber Orientation Distribution Function; Non-Negativity Constraint; Spherical Harmonics

1. Introduction

Diffusion MRI (dMRI) non-invasively reveals the microstructure of white matter by capturing the diffusion patterns of water molecules. The most widely used dMRI approach, Diffusion Tensor Imaging (DTI), captures one fiber direction per voxel and is incapable of describing complex diffusion processes due to its Gaussian diffusion assumption (Johansen-Berg and Behrens, 2009). In view of this, many High Angular Resolution Diffusion Imaging (HARDI) (Tuch et al., 2002) methods, have been developed in recent years to characterize non-Gaussian diffusion and compute quantities such as the Ensemble Average Propagator (EAP) (Wedeen et al., 2005; Descoteaux et al., 2010; Cheng et al., 2010b; Özarslan et al., 2009; Cheng et al., 2012), diffusion Orientation Distribution Function (dODF) (Tuch, 2004; Hess et al., 2006; Descoteaux et al., 2007; Aganj et al., 2010; Cheng et al., 2010a, 2012), and fiber Orientation Distribution Function (fODF) (Tournier et al., 2004, 2007; Alexander, 2005; Jian and Vemuri, 2007; Dell'Acqua et al., 2007, 2010; Landman et al., 2012; Weldeselassie et al., 2012).

Spherical Deconvolution (SD) has been shown to be effective for estimating the fODF by assuming that the measured diffusion-weighted signal can be obtained via spherically convolving a latent fODF with a fiber response function estimated from voxels known to be traversed by a single fascicle (Tournier et al., 2004, 2007; Jian and Vemuri, 2007; Johansen-Berg and Behrens, 2009). The fODF can hence be recovered via an inverse problem by deconvolving the signal with the estimated fiber response function. The local peaks (maxima) of the fODF give the corresponding fiber directions. SD methods can be classified into two categories, 1) Continuous Representation based SD (CR-SD), which is normally based on the Spherical Harmonic (SH) basis (Tournier et al., 2004, 2007; Anderson, 2005), and 2) Discrete Representation based SD (DR-SD), which is based on a discrete mixture of rotated versions of the fiber response function (Jian and Vemuri, 2007; Dell'Acqua et al., 2007, 2010; Landman et al., 2012).

Existing SD methods in both continuous and discrete representation categories share some common limitations. First, they often result in false-positive fiber directions (Tournier et al., 2004, 2007; Alexander, 2005; Johansen-Berg and Behrens, 2009; Jian and Vemuri, 2007; Landman et al., 2012; Weldeselassie et al., 2012), especially in low-anisotropy gray matter and cerebrospinal fluid (CSF) regions. Second, they normally fall short in ensuring that the estimated fODF is a proper probability density function, because non-negativity and unit integral throughout the unit sphere are not explicitly enforced. Most SD methods, including the popular Constrained SD (CSD) (Tournier et al., 2007) and all DR-SD methods (Jian and Vemuri, 2007; Dell'Acqua et al., 2007, 2010; Landman et al., 2012), consider non-negativity

only on discretized points but not the whole continuum of the unit sphere $\mathbb{S}^2$. To the best of our knowledge, Maximum Entropy SD (MESD) (Alexander, 2005) and Cartesian Tensor Fiber Orientation Distributions (CT-FOD) (Weldeselassie et al., 2010, 2012) are the only existing methods that ensure non-negativity throughout $\mathbb{S}^2$. However, they are computationally inefficient and rely on the error-prone process of numerical spherical integration. Ad-hoc normalization is also employed in these methods to obtain fODFs with unit integral. Some methods estimate continuously non-negative dODFs (Schwab et al., 2012; Cheng et al., 2012; Krajsek and Scharr, 2012) and EAPs (Cheng et al., 2012) using eigenvalue distribution of spherical functions and square root representation. But to our knowledge, none of these methods has been proposed to estimate continuously non-negative fODF in a SD framework. Third, for estimation of the fODF with reasonable accuracy, DR-SD methods (Jian and Vemuri, 2007; Dell'Acqua et al., 2007, 2010; Landman et al., 2012) require a significant amount of rotated fiber response functions along directions that are distributed densely on the unit sphere, significantly increasing the dimensionality and the time cost of the optimization problem. Furthermore, since for DR-SD methods (Jian and Vemuri, 2007; Dell'Acqua et al., 2007, 2010; Landman et al., 2012) the local peaks (maxima) of the fODF are detected from discretized points on the unit sphere, the angular resolution is limited.

To our knowledge, existing SD methods deal only with single-shell data (i.e., single b-value) and do not consider the radial component of diffusion. With advances in dMRI, multi-shell data are increasingly available (e.g., Human Connectome Project (HCP)[1] (Sotiropoulos et al., 2013)). For example, the HCP Q1 data[2] come with three b-values (b = 1000/2000/3000s/mm^2). Recent estimation methods such as the ball-stick model (Jbabdi et al., 2012), Q-Ball Imaging (Aganj et al., 2010), and other multi-shell HARDI methods (Assemlal et al., 2011; Cheng et al., 2010b,a; Descoteaux et al., 2010; Özarslan et al., 2009) demonstrated that ODF, EAP as well as fiber directions can be estimated with greater accuracy from multi-shell data compared with single-shell data. However, there is currently no existing work on how to perform SD on multi-shell data.

In this paper, we propose a method called Non-Negative Spherical Deconvolution (NNSD) to estimate fODF from both single- and multi-shell data. The main contributions of this paper are summarized as follows:

- NNSD is the first SH-based SD method to guarantee non-negativity throughout $\mathbb{S}^2$, not only on discretized points on the unit sphere, as in CSD (Tournier et al., 2007). In NNSD, non-negativity is achieved by representing the square root of the fODF as a linear combination of SH basis functions. Compared with non-SH methods like MESD and CT-FOD, which also guarantee non-negativity on the whole $\mathbb{S}^2$, NNSD is significantly faster due to the use of closed-form expression for spherical convolution. Compared with CSD which suppresses negative values in discrete samples using iteratively re-weighted regularization, the non-

[1] http://www.humanconnectome.org/data/
[2] http://www.humanconnectome.org/documentation/Q1/imaging-protocols.html#dMRI

negativity constraint in NNSD is built into its fODF representation and hence NNSD works well even without any regularization.

- In addition to the non-negativity constraint, NNSD reduces spurious peaks by implementing Riemannian gradient descent with an adaptive stopping condition. As a result, the anisotropy values of the fODFs estimated by NNSD in gray and white matter regions exhibit large contrast. Existing SD methods result in many false-positive peaks and hence high anisotropy in regions that are less anisotropic.
- Compared with traditional single-shell methods, we show that multi-shell data can be used for fODF estimation with greater robustness.
- Performance evaluation using real data is difficult due to the lack of ground truth. We propose in this paper a measure called Peak Consistency (PC) for quantitative fODF evaluation without exact knowledge of the ground truth.

Part of this work has been reported in our conference paper (Cheng et al., 2013b,a). Herein, we provide additional examples, results, derivations, and insights that are not part of this conference publication.

The rest of the paper is organized as follows. SD theory and algorithms are reviewed in Section 2. Section 2.1 provides an overview of existing SD methods using single-/multi-shell data, i.e. CSD (Tournier et al., 2007), a variant of CSD based on quadratic programming, MESD (Alexander, 2005), and DR-SD via $L1$ regularized non-negative least-squares fitting (L1-NNLS) (Jian and Vemuri, 2007; Landman et al., 2012). Section 2.2 describes NNSD and the associated Riemannian gradient descent algorithm. Two stopping strategies for Riemannian gradient descent are discussed. Section 3 furnishes evaluation details, including fiber response function estimation (Section 3.1), peak detection (Section 3.2), synthetic data generation (Section 3.3), and peak consistency evaluation (Section 3.4). In Section 4, NNSD is evaluated in comparison with the methods discussed in Section 2.1. Section 5 provides additional discussions on various aspects of NNSD. Section 6 concludes this paper.

2. Theory

2.1. Spherical Deconvolution Revisited

In this section, we describe CSD, MESD, and L1-NNLS, which were originally proposed for single-shell data, and generalize them for multiple-shell data. We also proposed a new implementation of CSD using quadratic programming.

Constrained SD (CSD).—SD (Tournier et al., 2004; Anderson, 2005) methods assume that the measured signal in each voxel is the product of convolving a latent fODF with an axisymmetric fiber response function. For $\mathbf{u} \in \mathbb{S}^2$, the fODF is represented as

$$\Phi(\mathbf{u}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{lm} Y_l^m(\mathbf{u}), \tag{1}$$

and the axisymmetric 3D fiber response function along the z-axis is represented as

$$H(q\mathbf{u} \mid (0,0,1)) = \sum_{l=0}^{L} h_l(q) Y_l^0(\mathbf{u}), \tag{2}$$

where $Y_l^m(\mathbf{u})$ is the l order m degree real SH basis function (Johansen-Berg and Behrens, 2009; Descoteaux et al., 2007) defined in spherical coordinates as

$$Y_l^m(\theta,\phi) = \begin{cases} \sqrt{2}\mathrm{Re}\left(y_l^{|m|}(\theta,\phi)\right) & \text{if } -l \le m < 0, \\ y_l^m(\theta,\phi) & \text{if } m = 0, \\ \sqrt{2}\mathrm{Im}(y_l^m(\theta,\phi)) & \text{if } \quad 0 < m \le l, \end{cases} \tag{3}$$

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

Re(·) and Im(·) denote respectively the real and imaginary parts of a complex number. $P_l^m(\cdot)$ is the associated Legendre polynomial. Since the fODF is antipodal symmetric, only even l's are used. Based on the property of spherical convolution and spherical harmonics (Tournier et al., 2004; Anderson, 2005), the diffusion signal can be represented in closed form as the convolution of the fiber response function with the fODF (Tournier et al., 2004) using

$$E(q\mathbf{u}) = \int_{\mathbb{S}^2} H(q\mathbf{u} \mid \mathbf{r}) \Phi(\mathbf{r}) \mathrm{d}\mathbf{r} = \sum_{l=0}^{L} \sum_{m=-l}^{l} \sqrt{\frac{4\pi}{2l+1}} f_{lm} h_l(q) Y_l^m(\mathbf{u})\,. \tag{5}$$

Note that in the original formulation (Tournier et al., 2004), signal measurements are performed on a single-shell in the **q**-space[3], fixing q at a certain value. In the current work, we let q vary and define the fiber response function in $\mathbb{R}^3$, hence allowing the coefficients $\{h_l(q)\}$ to vary across shells. This generalization allows us to extend CSD to work with multi-shell data. Based on Eq. (1) and Eq. (2), the SH coefficient vector $\boldsymbol{f} = (f_{00}, \ldots, f_{LL})^{\mathrm{T}}$ of the fODF can be estimated via

$$\min_{\boldsymbol{f}} \|\mathbf{M}\boldsymbol{f} - \mathbf{E}\|_2^2, \tag{6}$$

where $\mathbf{M} = \left[\sqrt{\frac{4\pi}{2l+1}} h_l(q_i) Y_l^m(\mathbf{u}_i)\right]$ is an $N \times \frac{(L+1)(L+2)}{2}$ basis matrix that is the element-wise product between the SH basis matrix $\left[Y_l^m(\mathbf{u}_i)\right]$ and the radial basis matrix $\left[\sqrt{\frac{4\pi}{2l+1}} h_l(q_i)\right]$, and $\mathbf{E} = (E_1, \ldots, E_N)^{\mathrm{T}}$ is the measured signal vector, including samples from single shell or multiple shells. The least-squares solution is $(\mathbf{M}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{E}$ (Tournier et al., 2004). Due to the smoothness of $H(q\mathbf{u})$, $h_l(q)$ decreases to zero as l increases, causing instability in estimating f_{lm} when l is large. Therefore, Tournier et al. (2004) proposed a low-pass Filtered SD (FSD) by down-weighting the high-order SH coefficients. However, FSD is known to result in many false-positive peaks because fODF non-negativity is not considered. In view

[3] $\mathbf{q} = q\mathbf{u}$ is a wave vector in **q**-space.

of this, Constrained SD (CSD) (Tournier et al., 2007) attempts to iteratively suppress the negative values by a discrete reconstruction of each fODF estimate, i.e.

$$\boldsymbol{f}^{(k+1)} = \underset{\boldsymbol{f}}{\operatorname{argmin}} \parallel \mathbf{M}\boldsymbol{f} - \mathbf{E} \parallel_2^2 + \lambda_{\mathrm{CSD}}^2 \parallel \mathbf{L}^{(k)}\boldsymbol{f} \parallel_2^2, \tag{7}$$

where $\mathbf{L}_{i,j}^{(k)} = \mathbf{P}_{i,j}$ if $\left(\mathbf{P}\boldsymbol{f}^{(k)}\right)_i < \tau$, $\mathbf{L}_{i,j}^{(k)} = 0$ if $\left(\mathbf{P}\boldsymbol{f}^{(k)}\right)_i > \tau$, and $\mathbf{P}$ is a matrix for reconstructing the fODF at discretized points on $\mathbb{S}^2$. $\boldsymbol{\tau}$ is a threshold normally chosen as 0.1 (Tournier et al., 2007). Note that non-negativity is only imposed on a set of discretized points, not throughout the whole unit sphere. Although CSD reduces significantly the negative values compared with FSD (Tournier et al., 2004), a significant amount of negative values still exist, even at points where non-negativity is imposed. An implementation of CSD for single shell data is available via MRtrix[4] (Tournier et al., 2012).

Maximum Entropy SD (MESD).—MESD is a CR-SD method that involves solving a non-linear least-squares problem using the Levenberg-Marquardt algorithm (Alexander, 2005). Based on the maximum entropy principle, the fODF is represented as an exponential function of the mixture of fiber response functions, i.e.,

$$\min_{\{\lambda_j\}_{i=0}^{K}} \sum_{i=1}^{N} \left(\int_{\mathbb{S}^2} H(\mathbf{q}_i \mid \mathbf{r}) \boldsymbol{\Phi}\left(\mathbf{r} \mid \{\lambda_j\}_{j=0}^{K}\right) \mathrm{d}\mathbf{r} - E_i \right)^2, \tag{8}$$

$$\boldsymbol{\Phi}\left(\mathbf{r} \mid \{\lambda_j\}_{j=0}^{K}\right) = \exp\left(\lambda_0 + \sum_{j=1}^{K} \lambda_j H(\mathbf{q}_j \mid \mathbf{r})\right), \tag{9}$$

where N is the number of diffusion-weighted samples. The exponential representation naturally ensures non-negativity on $\mathbb{S}^2$. However, without any closed-form expression for spherical convolution as in Eq. (5), the spherical integration in Eq. (8) needs to be approximated numerically. Based on the method of Lagrange multipliers and the maximum entropy principle, K should ideally be equal to N and the sampling points $\{\mathbf{q}_j\}_{j=1}^{K}$ should correspond to the N sampling points of the diffusion-weighted signal (Alexander, 2005). However, as a trade-off between accuracy and speed, only $K < N$ evenly distributed points are used[5]. After estimating $\{\lambda_j\}_{j=0}^{K}$, ad-hoc normalization of the fODF needs to be performed for unit integral. CT-FOD (Weldeselassie et al., 2010, 2012), which achieves non-negativity by representing the fODF using a sum of squares of homogeneous polynomials, share the same limitations as MESD by relying on numerical spherical integral and ad-hoc normalization.

L1-NNLS.—DR-SD uses a discrete representation of the fODF: $\{\boldsymbol{w}_i = \Phi(\mathbf{u}_i)\}$ (Jian and Vemuri, 2007; Landman et al., 2012; Yap and Shen, 2012).

[4]MRtrix: http://www.brain.org.au/software/mrtrix/
[5]Camino: http://cmic.cs.ucl.ac.uk/camino//index.php?n=Tutorials.MultifibreReconstruction#toc5

The SD problem in this case is typically formulated as

$$\min_{\boldsymbol{w}} \| \mathbf{A}\boldsymbol{w} - \mathbf{E} \|_2^2 + \lambda_{L1} \| \boldsymbol{w} \|_1 , \quad \text{s.t.} \quad \boldsymbol{w} \geq 0, \tag{10}$$

where $\mathbf{A}$ is a matrix with columns containing rotated versions of the fiber response function, and λ_{L1} is the sparsity tuning parameter. We refer to this method as $L1$ regularized Non-Negative Least Squares (L1-NNLS). The problem in Eq. (10) can be solved using efficient algorithms, such as those reported in (Lee et al., 2006; Kim et al., 2007). When $\lambda_{L1} = 0$, Eq. (10) becomes the Non-Negative Least Squares (NNLS) (Landman et al., 2012; Jian and Vemuri, 2007; Dell'Acqua et al., 2007). The Richardson-Lucy SD (Dell'Acqua et al., 2007), another DR-SD method, was shown to converge to the solution of NNLS in the case of Gaussian noise (De Pierro, 1993). The sparsity of the solution increases as λ_{L1} increases. If λ_{L1} is large enough, all elements in $\boldsymbol{w}$ are zero. Landman et al. (2012) proposed to determine λ_{L1} based on the so-called breakdown parameter $\lambda_{L1}^* = \|2\mathbf{A}^{\mathrm{T}}\mathbf{E}\|_\infty$, which is the minimal λ_{L1} that causes $\boldsymbol{w} = \mathbf{0}$ (Kim et al., 2007). Experimentally, Landman et al. (2012) set $\lambda_{L1} = 0.1\lambda_{L1}^*$. We followed this approach for experiments performed in this paper. Since the fODF in L1-NNLS is represented using discretized points, peak detection is restricted to these points instead of the whole $\mathbb{S}^2$, limiting the angular resolution.

QP-CSD.—We propose a quadratic programming implementation of CSD (QP-CSD) for imposing non-negativity in a manner that is more similar to NNLS, i.e.,

$$\min_{\boldsymbol{f}} \| \mathbf{M}\boldsymbol{f} - \mathbf{E} \|_2^2 , \quad \text{s.t.} \quad f_{00} = \frac{1}{\sqrt{4\pi}}, \quad \mathbf{P}\boldsymbol{f} \geq 0, \tag{11}$$

where $f_{00} = \frac{1}{\sqrt{4\pi}}$ ensures unit integral

$$\int_{\mathbb{S}^2} \boldsymbol{\Phi}(\mathbf{u})\mathrm{d}\mathbf{u} = \int_{\mathbb{S}^2} \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{lm} Y_l^m(\mathbf{u})\mathrm{d}\mathbf{u} = 1 . \tag{12}$$

QP-CSD, unlike CSD, produces the global solution and guarantees non-negativity on fODF points specified by matrix $\mathbf{P}$. Compared with L1-NNLS, QP-CSD works in a low dimension space thanks to the SH representation and allows peak detection to be performed on the continuous sphere.

2.2. Non-Negative Spherical Deconvolution (NNSD)

SH-based SD methods are able to leverage the nice property of having a closed-form solution to spherical deconvolution. But currently they fall short in enforcing non-negativity throughout the unit sphere. To address this issue, we propose a new approach called Non-Negative Spherical Deconvolution (NNSD), which we will describe next.

2.2.1. Square Root Representation of the fODF—

The square root representation has been proposed for the dODF (Cheng et al., 2009) and the EAP (Cheng et al., 2011), and it has been used for non-negative dODF and EAPs estimation (Cheng et al., 2012). In this

work, we propose to utilize the square root representation for non-negative fODF estimation by letting

$$\begin{aligned}\Phi(\mathbf{u} \mid \boldsymbol{c}) &= \left(\sum_{l=0}^{L}\sum_{m=-l}^{l} c_{lm} Y_l^m(\mathbf{u})\right)^2 \\ &= \sum_{\alpha=0}^{2L}\sum_{\beta=-\alpha}^{\alpha} \underbrace{\left(\sum_{l,m}^{L}\sum_{l',m'}^{L} c_{lm}c_{l'm'}Q_{ll'\alpha}^{mm'\beta}\right)}_{f_{\alpha\beta}} Y_\alpha^\beta(\mathbf{u}),\end{aligned} \tag{13}$$

where

$$Q_{ll'\alpha}^{mm'\beta} = \int_{\mathbb{S}^2} Y_l^m(\mathbf{u}) Y_{l'}^{m'}(\mathbf{u}) Y_\alpha^\beta(\mathbf{u}) \mathrm{d}\mathbf{u}\,. \tag{14}$$

The notation $\sum_{l,m}^{L}$ is a shorthand for $\sum_{l=0}^{L}\sum_{m=-l}^{l}$. $Q_{ll'\alpha}^{mm'\beta}$ is a constant resulting from the integration of three real SH functions (Cheng et al., 2012), which can be calculated from the Wigner 3-j symbol and Eq. (3). Note that the sum over α is up to $2L$, because when $\alpha > 2L$, $Q_{ll'\alpha}^{mm'\beta} = 0$, based on the property of the Wigner 3-j symbol[6].

Note that 1) a SH representation of $\sqrt{\Phi(\mathbf{u})}$ with maximum order L corresponds to a SH representation of $\Phi(\mathbf{u})$ with maximum order $2L$. Therefore compared with Eq. (1), the square root representation Eq. (13) is more compact in representing high angular resolution information. This is also contributive to reducing the Gibbs ringing artifact associated with the truncation of the SH series (Raffelt et al., 2012); 2) the fODF is naturally non-negative throughout $\mathbb{S}^2$ due to the square root representation.

Based on Eq. (2), Eq. (5) and Eq. (13), the diffusion signal can be represented using convolution as

$$\begin{aligned}E(q\mathbf{u}) &= \sum_{\alpha=0}^{2L}\sum_{\beta=-\alpha}^{\alpha} \sqrt{\frac{4\pi}{2\alpha+1}} f_{\alpha\beta} h_\alpha(q) Y_\alpha^\beta(\mathbf{u}) \\ &= \sum_{\alpha=0}^{2L}\sum_{\beta=-\alpha}^{\alpha}\sum_{l,m}^{L}\sum_{l',m'}^{L} \sqrt{\frac{4\pi}{2\alpha+1}} c_{lm}c_{l'm'}Q_{ll'\alpha}^{mm'\beta} h_\alpha(q) Y_\alpha^\beta(\mathbf{u}) \\ &= \boldsymbol{c}^{\mathrm{T}}\mathbf{K}(q\mathbf{u})\boldsymbol{c},\end{aligned} \tag{15}$$

where for any fixed $q\mathbf{u}$, $\mathbf{K}(q\mathbf{u})$ is a square matrix of dimension $(L+1)(L+2)/2$ with elements

$$K_{ll'}^{mm'}(q\mathbf{u}) = \sum_{\alpha=0}^{2L}\sum_{\beta=-\alpha}^{\alpha} \sqrt{\frac{4\pi}{2\alpha+1}} Q_{ll'\alpha}^{mm'\beta} h_\alpha(q) Y_\alpha^\beta(\mathbf{u})\,. \tag{16}$$

2.2.2. Riemannian Gradient Descent—Based on Eq. (15), we propose to estimate $\boldsymbol{c}$ from the measured $\{E_i\}_{i=1}^{N}$ by minimizing

[6] http://mathworld.wolfram.com/Wigner3j-Symbol.html

$$J(\boldsymbol{c}) = \frac{1}{2}\sum_{i=1}^{N}\left(\boldsymbol{c}^{\mathrm{T}}\mathbf{K}(q_i\mathbf{u}_i)\boldsymbol{c} - E_i\right)^2 + \frac{1}{2}\boldsymbol{c}^{\mathrm{T}}\Lambda\boldsymbol{c}, \quad \text{s.t.} \quad \| \boldsymbol{c} \|_2 = 1 \tag{17}$$

where Λ is a diagonal matrix with elements $\Lambda_{lm} = \lambda_{\mathrm{NNSD}} l^2(l+1)^2$ for Laplace-Beltrami regularization (Descoteaux et al., 2007), penalizing the high-order coefficients. The constraint $\|\boldsymbol{c}\|_2 = 1$ is a result of the unit integral, i.e. $\int_{\mathbb{S}^2}\boldsymbol{\Phi}(\mathbf{u})\mathrm{d}\mathbf{u} = 1$, and the orthogonality of the SH basis:

$$\int_{\mathbb{S}^2}\boldsymbol{\Phi}(\mathbf{u})\mathrm{d}\mathbf{u} = \sum_{l,m}^{L}\sum_{l',m'}^{L} c_{lm}c_{l'm'}\int_{\mathbb{S}^2} Y_l^m(\mathbf{u})Y_{l'}^{m'}(\mathbf{u})\mathrm{d}\mathbf{u} = \| \boldsymbol{c} \|_2^2 = 1, \tag{18}$$

which is equivalent to $f_{00} = \frac{1}{\sqrt{4\pi}}$, as in Eq. (11). The influence of the unit integral constraint on NNSD will be discussed in Section 5.3.

We propose a Riemannian gradient descent method on the sphere $\|\boldsymbol{c}\|_2 = 1$ to minimize Eq. (17). The Euclidean gradient of $J(\boldsymbol{c})$ is

$$\frac{\partial J(\boldsymbol{c})}{\partial \boldsymbol{c}} = \sum_{i=1}^{N} 2\left(\boldsymbol{c}^{\mathrm{T}}\mathbf{K}(q_i\mathbf{u}_i)\boldsymbol{c} - E_i\right)\mathbf{K}(q_i\mathbf{u}_i)\boldsymbol{c} + \Lambda\boldsymbol{c} \tag{19}$$

and the Riemannian gradient is the projection of Euclidean gradient onto the tangent space of $\boldsymbol{c}$, i.e.,

$$\nabla J(\boldsymbol{c}) = \frac{\partial J(\boldsymbol{c})}{\partial \boldsymbol{c}} - \left(\boldsymbol{c}^{\mathrm{T}}\frac{\partial J(\boldsymbol{c})}{\partial \boldsymbol{c}}\right)\boldsymbol{c}. \tag{20}$$

The Riemannian gradient descent associated with $J(\boldsymbol{c})$ is

$$\boldsymbol{c}^{(k+1)} = \exp_{\boldsymbol{c}^{(k)}}\left(-\mathrm{d}t\frac{\nabla J(\boldsymbol{c})}{\| \nabla J(\boldsymbol{c}) \|_2}\right), \tag{21}$$

$$\mathrm{Exp}_{\boldsymbol{c}}(\boldsymbol{v}) = \boldsymbol{c}\cos \| \boldsymbol{v} \|_2 + \frac{\boldsymbol{v}}{\| \boldsymbol{v} \|_2}\sin \| \boldsymbol{v} \|_2, \tag{22}$$

where $\boldsymbol{c}^{(k)}$ is the estimated $\boldsymbol{c}$ in the k-th step, $\mathrm{d}t$ is the step size chosen from inexact line search in $(0, \mathrm{d}t_0]$, and $\mathrm{Exp}_{\boldsymbol{c}}(\boldsymbol{v})$ is the exponential map on the sphere (Cheng et al., 2009). See Algorithm 1 for a summary of the whole process. We use the isotropic fODF with $\boldsymbol{c}^{(0)} = (1, 0, \ldots, 0)^{\mathrm{T}}$ for initialization and choose $\mathrm{d}t_0 = 0.1$ based on experimental observations and the fact that $\|\boldsymbol{c}\|_2 = 1$. We found that NNSD is robust to noise and $\lambda_{\mathrm{NNSD}} = 0$ works well in the experiments. The stopping condition is described next.

2.2.3. Adaptive Stopping Condition—

The stopping condition for Algorithm 1 is

$$\tilde{J}(k) = \frac{J\left(\boldsymbol{c}^{(k-1)}\right) - J\left(\boldsymbol{c}^{(k)}\right)}{J\left(\boldsymbol{c}^{(k-1)}\right)} < \delta\,. \tag{23}$$

Parameter δ should be chosen as a small number such that the algorithm converges to a local minimum of Eq. (17). Experimentally $\delta = 10^{-2}$ works well in most cases. If δ is further reduced beyond 10^{-2}, the fODFs in regions with high anisotropy will increase in sharpness, but the fODFs in regions with low anisotropy will however suffer from spurious peaks due to over-fitting. Therefore, we propose to adaptively stop the algorithm if $\mathrm{GFA}(\boldsymbol{c}^{(k)}) < T$ and $\tilde{J}(k) < \delta_0$ or if $\mathrm{GFA}(\boldsymbol{c}^{(k)}) \geq T$ and $\tilde{J}(k) < 0.01\delta_0$, where $\delta_0 = 10^{-2}$ and

$$\mathrm{GFA}(\boldsymbol{c}) = \sqrt{1 - \frac{c_{00}^2}{\| \boldsymbol{c} \|_2^2}} = \sqrt{1 - c_{00}^2} \tag{24}$$

is the Generalized Fractional Anisotropy (GFA) (Tuch, 2004) of the square root of the fODF $\sqrt{\Phi(\mathbf{u} \mid \boldsymbol{c})}$, and $T \in [0, 1]$ is a threshold depending on the noise level. We call the NNSD implementation with this adaptive stopping condition ASC-NNSD. Since gradient descent is initialized using the isotropic fODF, $\mathrm{GFA}(\boldsymbol{c}^{(k)}) < T$ is satisfied for all voxels during the early stage of optimization. As k increases, $\tilde{J}(k)$ decreases such that for a certain step k_0 we have $\tilde{J}(k_0) < \delta_0$. If for a voxel $\tilde{J}(k_0) < \delta_0$ and $\mathrm{GFA}\left(\boldsymbol{c}^{(k_0)}\right) < T$, we consider the voxel insufficiently anisotropic to proceed further and hence stop the algorithm early at step k_0. If $\tilde{J}(k_0) < \delta_0$ and $\mathrm{GFA}\left(\boldsymbol{c}^{(k_0)}\right) \geq T$, we consider the voxel anisotropic and further refine the solution by proceeding with gradient descent. When $T = 0$, ASC-NNSD becomes NNSD with $\delta = 0.01\delta_0 = 10^{-4}$; when $T = 1$, ASC-NNSD becomes NNSD with $\delta = \delta_0 = 10^{-2}$. Both NNSD and ASC-NNSD converge quickly, normally in a dozen of steps. Based on our C++ implementation, if $L = 6$, 1000 fODFs can be estimated within 5 seconds using an ordinary laptop, significantly faster than MESD as implemented in Camino (Alexander, 2005).

Algorithm 1: Non-Negative Spherical Deconvolution (NNSD)

Input: Signal measurements $\{E_i\}_{i=1}^N$.
Output: SH coefficient vector $\boldsymbol{c}$ and the fODF $\Phi(\mathbf{u}|\boldsymbol{c})$.
begin
  initialization: $\boldsymbol{c}^{(0)} = (1, 0, ..., 0)^{\mathrm{T}}$; // isotropic fODF
  repeat
    calculate $\boldsymbol{v} = \nabla J(\boldsymbol{c}^{(k)})$ as in Eq. (20);
    if $\|\boldsymbol{v}\| < \epsilon$ **then** break;
    choose step size $\mathrm{d}t \in (0, \mathrm{d}t_0]$ via line search;
    $\boldsymbol{c}^{(k+1)} = \mathrm{Exp}_{\boldsymbol{c}^{(k)}}\left(-\mathrm{d}t \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}\right)$, $k \leftarrow k + 1$;
  until *stopping condition is met*;
  compute $\Phi(\mathbf{u}|\boldsymbol{c})$ as in Eq. (13);
end

3. Methods

3.1. Estimation of Fiber Response Function

SD methods require an estimate of the fiber response function, which is essentially the diffusion signal profile of a single-directional highly-coherent fiber bundle. Anderson (2005) computed a SH representation of the fiber response function using the diffusion signal generated from a tensor model. Tournier et al. (2007) proposed to first fit the tensor model to

the data of voxels with high Fractional Anisotropy (FA) and then rotate the diffusion-weighted measurements such that the principal directions coincide with the z-axis. From the rotated measurements, the SH coefficients of the fiber response function are computed. Although these two methods described above were originally proposed for single-shell data, they can be generalized for multi-shell data using two different approaches:

- First fit diffusion models, such as DTI (Basser et al., 1994) or Spherical Polar Fourier Imaging (Assemlal et al., 2009; Cheng et al., 2010b), to high-anisotropy diffusion signal $E(\mathbf{q})$ in $\mathbb{R}^3$. Then, reconstruct the signal in each shell and compute the SH representation (Anderson, 2005; Cheng et al., 2010b).
- If a sufficiently large number of measurements are collected for each shell, the fiber response function can be estimated independently for each shell using the approach described in (Tournier et al., 2007).

Both approaches were used in the experiments (see Section 4.1).

3.2. Peak Detection

At each voxel, fiber orientations are given by the peaks or local maxima of each fODF. For L1-NNLS, the peaks are determined based on the local maxima of 321 points evenly distributed on a hemisphere, generated by subdividing the faces of an icosahedron. We discard local maxima with low fODF value and consider only peaks with fODF values greater than the mean of the minimal and maximal fODF values. We assume that the fiber crossing angle is no smaller than $\theta_\Delta = 15°$ and within this angle only one peak with the greatest fODF value is retained. The choice $\theta_\Delta = 15°$ is motivated by the fact the angle is approximately twice the minimal angular separation (i.e., 7.9°) between the 321 points.

For all CR-SD methods, including NNSD, CSD, QP-CSD, and MESD, we first determine candidate local maxima from the same 321 points. Then, using the continuous representation associated with these methods, a gradient ascent method, also implemented in MRtrix (Tournier et al., 2012), is performed to give a set of refined peaks.

3.3. Synthetic Data Simulation and Evaluation

For evaluation, synthetic data were generated by a mixture of tensor models:

$$E(q\mathbf{u}) = \frac{S(\mathbf{u})}{S(0)} = \sum_{k=1}^{K} p_k \exp\left(-q^2 \mathbf{u}^{\mathrm{T}} \mathbf{D}_k \mathbf{u}\right), \quad b = q^2, \tag{25}$$

where $\mathbf{D}_k$ is a 3×3 matrix representing a diffusion tensor and $K = 1, 2$ corresponds to the single-direction case and the crossing case, respectively. The baseline signal is $S(0) = 1$ and b is the diffusion weighting (Johansen-Berg and Behrens, 2009). For the crossing case, the tensors share the same eigenvalues (λ_1, λ_2, λ_2) and are equally weighted with $p_1 = p_2 = 0.5$. We generated Rician noise corrupted signal by

$$E_i = \sqrt{\left(E(q_i \mathbf{u}_i) + s_1\right)^2 + s_2^2} \tag{26}$$

where s_1 and s_2 are the Gaussian distributed noise with zero mean and standard deviation σ. The signal-to-noise ratio (SNR) is defined as $\frac{1}{\sigma}$ (Descoteaux et al., 2007). The methods used for comparison might not necessarily detect the correct number of peaks (2 in this case). Therefore, we report the *success ratio*, which tells us the proportion of trials that correctly detect the number of peaks. The Mean Difference of Angles (MDA) (Descoteaux et al., 2007) is then computed only for the successful trials.

3.4. Real Data Evaluation via Peak Consistency

Real data evaluation is challenging due to the lack of ground truth. To make comparison possible, we gauge the correctness of the fiber directions at each voxel based on the agreement of different datasets and methods. We assume that if a certain direction is widely agreed upon, then there is a high possibility that the direction reflects the true direction. Based on this assumption, we propose here a novel measure, called Peak Consistency (PC), to evaluate peak estimation accuracy. Assuming we have N datasets (e.g., resulting from different sampling schemes) pertaining to the same subject, then M methods will yield MN fODFs at each voxel, resulting in more than MN peaks (at least one peak for each fODF).

Two peaks are said to be consistent if the angle between them is no more than threshold θ_C. Our choice of angular threshold $\theta_C = 5°$ is based on the empirical observation that, for moderate Rician noise, the peak detection error of the algorithms is around 2°. This angular threshold gives a good tolerance to cover the variability caused by noise.

For voxel $\boldsymbol{x}$, we denote $p_{i,k}^{\boldsymbol{x}}$ as the k-th peak from the i-th fODF ($i \in [1 : MN]$, $[a : b]$ denotes $\{a, a+1, \ldots, b\}$). Since $\theta_C < \theta_\Delta$, for peak $p_{i,k}^{\boldsymbol{x}}$, at most one consistent peak can be found from peaks $\left\{p_{j,l}^{\boldsymbol{x}}\right\}_{l=1}^{K_j^{\boldsymbol{x}}}$ for a given $j \in [1 : MN]$, $j \neq i$, where $K_j^{\boldsymbol{x}}$ denotes the number of peaks for the j-th fODF at voxel $\boldsymbol{x}$. The maximum number of possible matches for all $j \neq i$ is hence $MN - 1$. For each peak, we define the Single-Peak Consistency (SPC)

$$\mathrm{SPC}(i, k, \boldsymbol{x}) = \frac{|\Gamma(i, k, \boldsymbol{x})|}{MN - 1}, \tag{27}$$

where $\Gamma(i, k, \boldsymbol{x}) = \{j \mid j \neq i, j \in [1 : NM], l \in \left[1 : K_j^{\boldsymbol{x}}\right], d\left(p_{i,k}^{\boldsymbol{x}}, p_{j,l}^{\boldsymbol{x}}\right) \leq \theta_C\}$, $d\left(p_{j,l}^{\boldsymbol{x}}, p_{i,k}^{\boldsymbol{x}}\right)$ is the angle between the two peaks, and $|\cdot|$ denotes the number of elements in a set. SPC = 0 indicates that no peak is consistent with the reference peak $p_{i,k}^{\boldsymbol{x}}$. SPC = 1 indicates that the reference peak $p_{i,k}^{\boldsymbol{x}}$ is fully consistent with all other $MN - 1$ peaks. For concurrent consistency evaluation of multiple peaks, we define the Multi-Peak Consistency (MPC) as

$$\mathrm{MPC}(i, \boldsymbol{x}) = \frac{\left|\cap_{k \in \left[1 : K_i^{\boldsymbol{x}}\right]} \Gamma(i, k, \boldsymbol{x})\right|}{MN - 1}. \tag{28}$$

See Fig. 1 for a visual summary of the single- and multi-peak consistency measures. Given a field of fODFs with voxels $\{\boldsymbol{x} | \boldsymbol{x} \in \Omega\}$, we define for a threshold $\mathrm{PC}_T \in (0, 1)$ (0.5 in the current work) the overall single- and multi-peak consistency measures as

$$\overline{\mathrm{SPC}}(i) = \frac{\sum_{\boldsymbol{x}\in\Omega} |\{k \mid \mathrm{SPC}(i,k,\boldsymbol{x}) \geq \mathrm{PC_T}, k \in [1{:}K_i^{\boldsymbol{x}}]\}|}{\sum_{\boldsymbol{x}\in\Omega} K_i^{\boldsymbol{x}}} \tag{29}$$

and

$$\overline{\mathrm{MPC}}(i) = \frac{\sum_{\boldsymbol{x}\in\Omega} |\{\boldsymbol{x} \mid MPC(i,\boldsymbol{x}) \geq \mathrm{PC_T}\}|}{|\Omega|}. \tag{30}$$

4. Experiments

4.1. Parameters

We compared NNSD and ASC-NNSD with CSD, QP-CSD, MESD, and L1-NNLS, using both synthetic and real data. For CSD, MRtrix (Tournier et al., 2012) was used in all experiments involving single-shell data (Tournier et al., 2007), and an in-house implementation of CSD was used in experiments involving multi-shell data. The default parameters specified in (Tournier et al., 2007) and 321 orientations evenly distributed in a hemisphere were used to generate $\mathbf{L}^{(k)}$ in Eq. (7) for CSD (Tournier et al., 2007) and $\mathbf{P}$ in Eq. (11) for QP-CSD. Accelerated MESD with $K = 16$ in Eq. (8) was used as suggested in Camino. Following (Landman et al., 2012), we set $\lambda_{L1} = 0.1\lambda_{L1}^{*}$ for L1-NNLS. To demonstrate the robustness of NNSD, we intentionally set $\lambda_{\mathrm{NNSD}} = 0$ to switch off Laplace-Beltrami regularization and set $T = 0.5$ for ASC-NNSD. For the synthetic data experiments, the fiber response function for all methods was set according to the signal from the tensor model with eigenvalues [1.7, 0.2, 0.2] × 10^{-3} mm^2/s. For the real data experiments, the response function was estimated by MRtrix (Tournier et al., 2012) using voxels with FA greater than 0.7.

4.2. Synthetic Data

Evaluation with SNR=1000.—We would like to investigate how the different methods behave when the anisotropy of the signal changes and when there is a mismatch between the fiber response function and the data. Diffusion signal was first generated from the mixture of tensor models in Eq. (25) with $b = 1500 s/mm^2$, crossing angle 60°, mean diffusivity 0.7 × 10^{-3} mm^2/s, and different FA values (Johansen-Berg and Behrens, 2009) by varying tensor eigenvalues (λ_1, λ_2, $\lambda_3 = \lambda_2$). Fig. 2 (SNR=1000) shows the estimated fODFs when λ_1 increases from 0.7 × 10^{-3} mm^2/s (left with FA= 0) to 1.7 × 10^{-3} mm^2/s (right with FA= 0.87). For all methods, the fiber response function was the tensor model with eigenvalues [1.7, 0.2, 0.2] × 10^{-3} mm^2/s. Therefore, the response function matches the data on the far right but not on the left. The results indicate that CSD, QP-CSD and L1-NNLS result in many spurious peaks for data with low anisotropy, even when the SNR is high. NNSD and ASC-NNSD work well when the signal is anisotropic or less so. The fODFs estimated by NNSD with $\delta = 10^{-2}$ are a little smoother than the fODFs estimated by NNSD with $\delta = 10^{-4}$. Note that for isotropic signal, the fODFs given by NNSD and ASC-NNSD are close to being isotropic.

Evaluation with SNR=30.—The estimation results when the noise level is increased are shown in Fig. 2 (SNR=30). Despite the noise, both NNSD and ASC-NNSD obtain significantly smoother results for signal that is more isotropic when compared with other methods. ASC-NNSD works best among all methods in obtaining sharper fODFs for anisotropic signal and smoother fODFs for isotropic signal. The results given by ASC-NNSD can be seen as a combined outcome of NNSD with $\delta = 10^{-2}$ for the less anisotropic signal and NNSD with $\delta = 10^{-4}$ for the anisotropic signal.

Evaluation of Non-Negativity and Anisotropy.—1000 Rician-noise realizations of isotropic signal with eigenvalues $[0.7, 0.7, 0.7] \times 10^{-3}$ mm^2/s and anisotropic signal with eigenvalues $[1.7, 0.2, 0.2] \times 10^{-3}$ mm^2/s were generated at SNR=15 and 30 for $b = 1500$s/mm^2 and 60 evenly distributed directions. From the estimated fODFs, we recorded the GFA values (Tuch, 2004) and the proportion of directions (among 5121 points uniformly distributed on the unit sphere) with negative fODF values. We ignored negative values close to zero and only took into consideration negative values with absolute values larger than 1% of the maximal fODF value. The top subfigure in Fig. 3 indicates that CSD is still susceptible to negative values despite its non-negativity constraint. QP-CSD results in a significantly lower number of negative values than CSD because it enforces a stricter non-negativity requirement than CSD. ASC-NNSD and MESD guarantee non-negativity throughout $\mathbb{S}^2$. Note that the fODFs estimated by L1-NNLS are non-negative on the 321 points used during optimization; however, since fODF values on the other points are unknown, we cannot compute the proportion of negative values for L1-NNLS. The bottom subfigure indicates the GFA values of the fODFs given by CSD, QP-CSD, MESD and L1-NNLS are high even for isotropic signal, indicating overestimation of the number of fiber directions. Only ASC-NNSD yields significant contrast between anisotropic and isotropic signals.

Evaluation of Peak Accuracy – Single Shell.—Noise-corrupted signals were generated from a two-tensor model with $b = 1500$s/mm^2, eigenvalues $[1.7, 0.2, 0.2] \times 10^{-3}$ mm^2/s, different crossing angles in the range of [30°, 90°], and at SNR=10 and 20. For peaks detected from the fODFs estimated by each method, we recorded the success ratio and the MDA as described in Section 3.3. The results are shown in Fig. 4 for both SNR=10 and SNR=20. When SNR=10, MESD, CSD, and NNSD with $L = 10$ give low success ratios for small crossing angles, but high success ratios for the large crossing angles. With $L = 6$, when the crossing angle is larger than 55°, ASC-NNSD generally gives higher success ratio than CSD and lowest MDA among all methods. When SNR=20, MESD, CSD, and NNSD with $L = 10$ give higher success ratios for both small and large crossing angles. CSD, QP-CSD NNSD with $L = 10$ give better performance than their counterparts with $L = 6$. This implies that, if SNR is high, CSD and NNSD benefit from using a higher order representation. It is interesting to see that the success ratio of L1-NNLS is high for small crossing angle but low for large crossing angle. This indicates that L1-NNLS has high sensitivity, but low specificity due to false-positive peaks. Partially due to limited angular resolution, L1-NNLS yields larger MDA values than CSD and NNSD. The MDA performance of MESD can potentially be improved by increasing parameter K.

Evaluation of Peak Accuracy – Multiple Shells.—We performed an evaluation similar to the previous section using multi-shell data. Since the results reported in the previous section indicate that CSD and NNSD generally give better performance than QP-CSD, MESD, and L1-NNLS, we focus our comparison in this section on CSD and NNSD. We chose $L = 8$, which is good a trade-off between angular resolution and robustness. This is also suggested in MRtrix (Tournier et al., 2012). The data were generated by sampling from a mixture of tensor models (Caruyer et al., 2011) using $b = 1500\text{s/mm}^2$, 3000s/mm^2 in 30 staggered directions per shell. Fig. 5 shows the success ratios and MDA values given by CSD and NNSD using samples respectively from the outer $b = 3000\text{s/mm}^2$ shell (30 samples) and from two shells (60 samples) concurrently. When SNR=10, ASC-NNSD and CSD give similar success ratios, but ASC-NNSD generally gives lower MDA than CSD for both single- and two-shell cases. For both ASC-NNSD and CSD, the two-shell data generally result in higher success ratios and lower MDA values than single-shell data. When SNR=20, ASC-NNSD gives better performance than CSD both in terms of success ratio and MDA. For both ASC-NNSD and CSD, the two-shell data result in success ratios similar to the single-shell data with $b = 3000\text{s/mm}^2$, but lower MDA values.

4.3. Real Data

4.3.1. Single-Shell Data—

Evaluation was performed using real human data with $b = 2000\text{s/mm}^2$, 120 gradient directions, 2mm isotropic resolution, and TR/TE=12400/116ms. We set $L = 8$ for CSD, QP-CSD, NNSD, and ASC-NNSD. MRtrix was used for the estimation of the fiber response function from voxels with FA > 0.7 (Tournier et al., 2004, 2007). The results are shown in Fig. 6. NNSD with $\delta = 10^{-4}$ and ASC-NNSD with $T = 0.5$ give similar results, implying that the result is insensitive to T for this data. For each method, the GFA map computed from the estimated fODFs is displayed as the background image. Consistent with the results in Fig. 3, the results in Fig. 6 for CSD, QP-CSD, MESD, and L1-NNLS show a significant amount of false-positive peaks, which are especially evident in the regions with low anisotropy, as indicated by the high GFA values in gray matter and cerebrospinal fluid regions. NNSD and ASC-NNSD dramatically reduced the spurious peaks of the estimated fODFs. See for example the region marked by the yellow circles in Fig. 6. Compared with CSD, NNSD and ASC-NNSD yield sharper fODFs in anisotropic regions and more isotropic fODFs in isotropic regions. This gives high GFA contrast between these regions. Note that existing work (Tournier et al., 2004, 2007; Jian and Vemuri, 2007; Landman et al., 2012; Weldeselassie et al., 2012) attempts to avoid such low GFA contrast by computing FA or GFA using some other models, such as the tensor model.

4.3.2. High Spatial Resolution Multi-Shell HCP Data—

Evaluation was also performed using a high-resolution dataset (1.25mm in all dimensions) obtained from the Human Connectome Project (HCP)[7]. fODF estimation from this dataset is challenging because the SNR is low due to the small voxel size. This dataset was acquired using three shells, with 90 staggered directions per shell, and at $b = 1000, 2000, 3000\text{s/mm}^2$.

[7] http://www.humanconnectome.org/documentation/Q1/

First, we performed a single-shell evaluation ($b = 2000\text{s/mm}^2$) using NNSD with $\delta = 10^{-4}$, ASC-NNSD with $T = 0.5$, CSD, QP-CSD, MESD, and L1-NNLS. $L = 8$ was used in all SH-based methods. The results shown in Fig. 7 demonstrate the effectiveness of NNSD and ASC-NNSD. In isotropic regions, only NNSD and ASC-NNSD give isotropic fODFs, while other methods result in a significant amount of false-positive peaks. See for example the region marked by the yellow circles. In anisotropic regions, NNSD and ASC-NNSD yield the sharpest fODFs. In particular, the fODFs estimated by CSD are not as sharp as those estimated by NNSD, ASC-NNSD, and MESD. This can be seen in the region marked by the yellow squares. Spurious peaks can be observed for L1-NNLS, as indicated by the high anisotropy values, even in regions with isotropic diffusion.

Second, we included the two other shells ($b = 1000, 3000\text{s/mm}^2$) in the estimation, totaling up to 90×3 samples. The results are shown in Fig. 8. For ASC-NNSD, the $b = 1000\text{s/mm}^2$ data give fODFs that are not as sharp as the results given by the two other b-values. The $b = 3000\text{s/mm}^2$ data give fODFs that are the sharpest but with more false-positive peaks. The $b = 2000\text{s/mm}^2$ data seem to yield more moderate results than among all b-values.

Third, we detected the peaks from the fODFs estimated using the 4 sampling schemes (the individual shells and all three shells) and 5 methods[8](ASC-NNSD, CSD, QP-CSD, MESD, and L1-NNLS) and then computed the corresponding peak consistency values defined in Section 3.4. Note that fODF peaks detected in isotropic regions are not reliable and are hence not included in the evaluation. Only voxels with GFA greater than 0.5, estimated based on ASC-NNSD using all three shells, were used for consistency evaluation. The detected peaks, color-coded with SPC values, are shown in Fig. 9, indicating that 1) peak consistency is high anisotropic and single directional regions; 2) peak consistency is low in fiber crossing regions; and 3) the lowest peak consistency occurs at boundaries between white matter and gray matter or cerebrospinal fluid regions. It can also be observed that for ASC-NNSD, CSD, and QP-CSD, the multi-shell data result in higher SPC than single-shell data. For both single- and multi-shell data, ASC-NNSD yields higher SPC values than other methods. These observations are confirmed by the overall peak consistency values ($\overline{\text{SPC}}$ and $\overline{\text{MPC}}$) shown in Table 1.

5. Discussion

5.1. Multi-Shell SD

Existing fODF estimation methods such as CSD only can be applied to single-shell data. We have generalized SD to work with multi-shell data by defining the fiber response function in $\mathbb{R}^3$. Data with low diffusion weighting have high SNR but low angular resolution, whereas data with high diffusion weighting data have low SNR but high angular resolution. Our method takes advantage of the whole spectrum of the data. Evaluation based on both synthetic and real data, as shown in Figs. 5, 6, 8, 9, and Table 1, indicate that multi-shell data improve peak estimation accuracy compared with single-shell data.

[8]Multi-shell versions of MESD and L1-NNLS were not implemented. For peak consistency evaluation, we therefore have $4 \times 5 - 2 = 18$ fODFs per voxel.

It should be noted that for the HCP data, the multi-shell SD methods use all 90 × 3 diffusion-weighted images for estimation. This is 3 times the number of samples used by single-shell SD methods. It will be interesting to fix the total number of samples and evaluate the advantages of the two different sampling schemes, e.g., a multi-shell sampling scheme with 90 samples per shell against a single-shell scheme with 270 samples. However, choosing a good sampling scheme (Cheng et al., 2014) with proper consideration of the tradeoff between angular resolution and SNR is still an open problem beyond the scope of this paper. The proposed multi-shell SD framework can be used as a tool to evaluate different sampling schemes in the future.

5.2. Non-Local NNSD (NL-NNSD)

In practice, denoising methods such as non-local mean (Buades et al., 2005; Descoteaux et al., 2008; Yap et al., 2014) and spatial regularization (Goh et al., 2009) can significantly improve SNR and hence fODF estimation accuracy. Our initial implementation of a non-locally regularized (Buades et al., 2005) version of NNSD, called non-local NNSD (NL-NNSD) (Cheng et al., 2013a), supports this idea. In the ISBI 2013 HARDI Reconstruction Challenge[9], which considered DTI, HARDI, and DSI-like sampling schemes and SNR=10,20,30, NL-NNSD was ranked the best technique in terms of local fiber orientation estimation for all sampling categories and all SNRs[10].

5.3. Unit Integral Constraint

Unlike NNSD, existing SD methods, such as FSD, CSD, MESD and L1-NNLS, do not explicitly incorporate the unit integral constraint. In Appendix A, we show that the solutions given by FSD in Eq. (6) with and without the constraint differ only by the isotropic component, and under some conditions the results given by NNSD with and without the constraint also differ only by the isotropic component. The constraint is beneficial for fODF estimation as explained below:

1. As a probability distribution function of orientations, fODF should naturally have unit integral.
2. The constraint helps mitigate the representation error inherent in SD-based fODF estimation methods. Based on Eq. (5) and Eq. (15), the spherical means of $E(q\mathbf{u})$ and $H(q\mathbf{u})$ are respectively

$$\bar{E}(q) = \frac{1}{4\pi}\int_{\mathbb{S}^2} E(q\mathbf{u})\mathrm{d}\mathbf{u} = f_{00}h_0(q), \quad \bar{H}(q) = \frac{1}{\sqrt{4\pi}}h_0(q)\,. \tag{31}$$

The value of f_{00} should ideally be adaptive to the values of these means in order to minimize the representation error. For example, if the diffusion signal and fiber response function have the same spherical mean, then one should set $f_{00} = \frac{1}{\sqrt{4\pi}}$ to avoid representation error. This requirement is obvious for $q = 0$, where we expect $\bar{E}(0) = \bar{H}(0) = 1$ and $h_0(0) = \sqrt{4\pi}$, especially when the fiber

[9] http://hardi.epfl.ch/static/events/2013_ISBI/
[10] http://hardi.epfl.ch/static/events/2013_ISBI/_static/talk_Max.pdf

response function is based on the tensor model. If the spherical means differ, our analysis in Appendix A shows that under some conditions setting $f_{00} = \frac{1}{\sqrt{4\pi}}$ will result in an fODF with the same anisotropic component.

3. Due to the constraint, the solution space of SD methods with the constraint is one dimension smaller than the solution space of SD methods without this constraint. This reduces the feasible solutions to a smaller space, effectively reducing the complexity of the associated non-convex optimization, and focuses the optimization on the anisotropic components of the fODF.

NNSD without the constraint can be solved by traditional gradient descent with inexact line search. The stopping condition Eq. (23) and its adaptive version can be used to terminate the algorithm. Our experiments indicate that $\delta = 10^{-4}$ gives the best results. Fig. 10 shows the fODF results given by NNSD without the constraint using the synthetic data shown in Fig. 2. The glyphs are normalized by $\|\boldsymbol{c}\| = 1$ for visualization. Compared with the results given by NNSD in Fig. 2 with $\delta = 10^{-4}$, the fODFs given by NNSD without the constraint are generally similar, but not as sharp. It should be noted that when δ varies, the fODFs given by NNSD without the constraint vary more dramatically than NNSD with the constraint, indicating that the latter is more robust to the choice of δ.

6. Conclusion

The contribution of this paper is threefold. First, by introducing a $\mathbb{R}^3$ fiber response function, we have generalized the existing single-shell SD framework to cater to multi-shell data. The experimental results indicate that, compared with single-shell data, multi-shell data give lower MDA, higher success ratios and more consistent peaks. Second, we have proposed a novel SD method, called Non-Negative SD (NNSD), which utilizes a square root representation to impose non-negativity. Comparisons with existing SD methods, including CSD, MESD, and L1-NNLS, demonstrate the advantages of NNSD: 1) NNSD guarantees non-negativity with unit integral of the fODF throughout the whole $\mathbb{S}^2$; 2) NNSD significantly reduces false-positive peaks and yields high GFA contrast between isotropic and anisotropic regions; 3) Due to the SH representation, NNSD is efficient and allows accurate peak detection on $\mathbb{S}^2$. Third, we have proposed a new metric called peak consistency for evaluation of the different methods in the absence of ground truth. Using this metric, we demonstrated that NNSD gives better estimates of peak orientations in real data. Improvements in fODF estimation obtained with NNSD will be helpful for applications such as fiber tractography (Yap et al., 2011b,c), fiber clustering (Wang et al., 2013), and brain connectomics (Yap et al., 2010, 2011a; Wee et al., 2010, 2012; Shi et al., 2012).

Acknowledgment

This work was supported in part by a UNC BRIC-Radiology startup fund and NIH grants (EB006733, EB008374, EB009634, AG041721, MH100217, AG042599). Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Appendix A.: fODF estimation with and without the unit integral constraint

Since for each shell the matrix $[Y_l^m(\mathbf{u}_i)]$ has orthonormal columns, Eq. (6), in continuous form, is equivalent to

$$\min_{\boldsymbol{f}} \sum_{s=1}^{S} \int_{\mathbb{S}^2} \left(\sum_{l,m}^{L} \sqrt{\frac{4\pi}{2l+1}} f_{lm} h_l(q_s) Y_l^m(\mathbf{u}) - E(q_s\mathbf{u}) \right)^2 \mathrm{d}\mathbf{u} \tag{A.1}$$

The minimization is quadratic with respect to f_{lm} and the optimal coefficients can be found at the point where the derivative vanishes as

$$f_{lm}^* = \sqrt{\frac{2l+1}{4\pi}} \frac{\sum_{s=1}^{S} h_l(q_s) e_{lm}(q_s)}{\sum_{s=1}^{S} h_l(q_s)^2}, \tag{A.2}$$

where $e_{lm}(q_s)$ is the SH coefficients of $E(q_s\mathbf{u})$. With the unit integral constraint, we have $f_{00}^* = \frac{1}{\sqrt{4\pi}}$, but this does not alter f_{lm}^* for $l > 0$. Thus the difference between the fODFs with and without the constraint is only the isotropic component.

The NNSD formulation can be seen as an extension of FSD because based on Eq. (15) we can write $E(q_i\mathbf{u}_i) = \boldsymbol{c}^{\mathrm{T}}\mathbf{K}(q_i\mathbf{u}_i)\boldsymbol{c} = \mathbf{M}(q_i\mathbf{u}_i)\boldsymbol{f}$, where $\mathbf{M}(q_i\mathbf{u}_i)$ is the row in basis matrix $\mathbf{M}$ corresponding to the sample $q_i\mathbf{u}_i$. Without taking unit integral constraint into consideration, the unconstrained version of NNSD solves the following problem:

$$\min_{\boldsymbol{c}} \sum_{s=1}^{S} \int_{\mathbb{S}^2} \left(\sum_{\alpha,\beta}^{2L} \sqrt{\frac{4\pi}{2\alpha+1}} f_{\alpha\beta} h_\alpha(q_s) Y_\alpha^\beta(\mathbf{u}) - E(q_s\mathbf{u}) \right)^2 \mathrm{d}\mathbf{u} \tag{A.3a}$$

$$\text{s.t.} \quad f_{\alpha\beta} = \sum_{l,m}^{L} \sum_{l',m'}^{L} c_{lm} c_{l'm'} Q_{ll'\alpha}^{mm'\beta}. \tag{A.3b}$$

Note that by setting $\alpha = \beta = 0$ in Eq. (A.3b) we have $f_{00} = \frac{\| \boldsymbol{c} \|^2}{\sqrt{4\pi}}$. Without loss of generality, we let, for optimal solution $\boldsymbol{c}^*$, $\frac{1}{\sqrt{4\pi}} \|\boldsymbol{c}^*\|^2 = f_{00}^* = \frac{1}{\sqrt{4\pi}} + \Delta$. The problem given by Eq. (A.3a) and Eq. (A.3b) then becomes

$$\min_{\boldsymbol{c}} \sum_{s=1}^{S} \int_{\mathbb{S}^2} \left(\sum_{\alpha,\beta}^{2L} \sqrt{\frac{4\pi}{2\alpha+1}} f_{\alpha\beta} h_\alpha(q_s) Y_\alpha^\beta(\mathbf{u}) - E(q_s\mathbf{u}) \right)^2 \mathrm{d}\mathbf{u} \tag{A.4a}$$

$$\text{s.t.} \quad f_{\alpha\beta} = \sum_{l,m}^{L} \sum_{l',m'}^{L} c_{lm} c_{l'm'} Q_{ll'\alpha}^{mm'\beta}, \quad f_{00} = \frac{1}{\sqrt{4\pi}} + \Delta. \tag{A.4b}$$

If $\Delta = 0$, then Eq. (A.4a) is just the NNSD with unit integral in Eq. (17). If $\Delta \neq 0$, we let $f^1_{00} = \frac{1}{\sqrt{4\pi}}$, and $f^1_{\alpha\beta} = f_{\alpha\beta}$ for $\alpha > 0$. Then Eq. (A.4a) is equivalent to

$$\begin{aligned}&\sum_{s=1}^{S}\int_{\mathbb{S}^2}\left(\Delta h_0(q_s)+\sum_{\alpha,\beta}^{2L}\sqrt{\frac{4\pi}{2\alpha+1}}f^1_{\alpha\beta}h_\alpha(q_s)Y^\beta_\alpha(\mathbf{u})-E(q_s\mathbf{u})\right)^2\mathrm{d}\mathbf{u}\\&=\sum_{s=1}^{S}\int_{\mathbb{S}^2}\left(\sum_{\alpha,\beta}^{2L}\sqrt{\frac{4\pi}{2\alpha+1}}f^1_{\alpha\beta}h_\alpha(q_s)Y^\beta_\alpha(\mathbf{u})-E(q_s\mathbf{u})\right)^2\mathrm{d}\mathbf{u}\\&+2\sum_{s=1}^{S}\Delta h_0(q_s)\sqrt{4\pi}(h_0(q_s)-e_{00}(q_s))+\sum_{s=1}^{S}4\pi\Delta^2h_0(q_s)^2.\end{aligned}\tag{A.5}$$

Since $e_{00}(q_s)$, $\{h_\alpha(q_s)\}$ and Δ are known constants, the problem given by Eq. (A.4a) and Eq. (A.4b) is equivalent to

$$\min_{\boldsymbol{c}}\sum_{s=1}^{S}\int_{\mathbb{S}^2}\left(\sum_{\alpha,\beta}^{2L}\sqrt{\frac{4\pi}{2\alpha+1}}f^1_{\alpha\beta}h_\alpha(q_s)Y^\beta_\alpha(\mathbf{u})-E(q_s\mathbf{u})\right)^2\mathrm{d}\mathbf{u}\tag{A.6a}$$

$$\text{s.t.}\quad f^1_{\alpha\beta}=\sum_{l,m}^{L}\sum_{l',m'}^{L}c_{lm}c_{l'm'}Q^{mm'\beta}_{ll'\alpha},\quad f^1_{00}=\frac{1}{\sqrt{4\pi}}\tag{A.6b}$$

which is exactly the NNSD with unit integral constraint. The conditions for this equivalence are (a) $\min_{\mathbf{u}\in\mathbb{S}^2}\sum_{lm}f^*_{lm}Y^m_l(\mathbf{u})\geq\frac{\Delta}{\sqrt{4\pi}}$ and (b) L is large enough. Condition (a) is to ensure that the fODF as represented by $\left\{f^1_{lm}\right\}$ is nonnegative; this can be satisfied if $\Delta \leq 0$. Condition (b) is to ensure representation of this nonnegative fODF using $\{c^*_{lm}\}$ without error. With these two conditions satisfied, the difference between the fODF solutions given by NNSD with and without the unit integral constraint is an isotropic component. In practice, since obtaining the solutions involves non-convex optimization, local minima might cause differences even in the anisotropic component.

References

Aganj I, Lenglet C, Sapiro G, Yacoub E, Ugurbil K, Harel N, 2010. Reconstruction of the orientation distribution function in single and multiple shell q-ball imaging within constant solid angle. Magnetic Resonance in Medicine 2, 554–566.

Alexander D, 2005. Maximum entropy spherical deconvolution for diffusion MRI, in: Information Processing in Medical Imaging, Springer. pp. 27–57.

Anderson A, 2005. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. Magnetic Resonance in Medicine 54, 1194–1206. [PubMed: 16161109]

Assemlal H, Tschumperlé D, Brun L, Siddiqi K, 2011. Recent advances in diffusion MRI modeling: Angular and radial reconstruction. Medical Image Analysis 15, 369–396. [PubMed: 21397549]

Assemlal HE, Tschumperlé D, Brun L, 2009. Efficient and robust computation of PDF features from diffusion MR signal. Medical Image Analysis 13, 715–729. [PubMed: 19665917]

Basser PJ, Mattiello J, LeBihan D, 1994. MR diffusion tensor spectroscopy and imaging. Biophysical Journal 66, 259–267. [PubMed: 8130344]

Buades A, Coll B, Morel JM, 2005. A review of image denoising algorithms, with a new one. Multiscale Modeling and Simulation 4, 490–530.

Caruyer E, Cheng J, Lenglet C, Sapiro G, Jiang T, Deriche R, 2011. Optimal Design of Multiple Q-shells experiments for Diffusion MRI, in: Computational Diffusion MRI - MICCAI Workshop.

Cheng J, Deriche R, Jiang T, Shen D, Yap PT, 2013a. Non-Local Non-Negative Spherical Deconvolution for Single and Multiple Shell Diffusion MRI, in: HARDI Reconstruction Challenge, International Symposium on Biomedical Imaging (ISBI).

Cheng J, Deriche R, Jiang T, Shen D, Yap PT, 2013b. Non-Negative Spherical Deconvolution (NNSD) for Fiber Orientation Distribution Function Estimation, in: Computational Diffusion MRI - MICCAI Workshop.

Cheng J, Ghosh A, Deriche R, Jiang T, 2010a. Model-Free, Regularized, Fast, and Robust Analytical Orientation Distribution Function Estimation, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 648–656. [PubMed: 20879286]

Cheng J, Ghosh A, Jiang T, Deriche R, 2009. A Riemannian Framework for Orientation Distribution Function Computing, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI, pp. 911–918. [PubMed: 20426075]

Cheng J, Ghosh A, Jiang T, Deriche R, 2010b. Model-free and Analytical EAP Reconstruction via Spherical Polar Fourier Diffusion MRI, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI, pp. 590–597. [PubMed: 20879279]

Cheng J, Ghosh A, Jiang T, Deriche R, 2011. Diffeomorphism Invariant Riemannian Framework for Ensemble Average Propagator Computing, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI, Springer Berlin / Heidelberg. pp. 98–106.

Cheng J, Jiang T, Deriche R, 2012. Nonnegative Definite EAP and ODF Estimation via a Unified Multi-Shell HARDI Reconstruction, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI, Springer Berlin / Heidelberg. pp. 98–106.

Cheng J, Shen D, Yap PT, 2014. Designing single- and multiple-shell sampling schemes for diffusion MRI using spherical code, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI).

De Pierro AR, 1993. On the relation between the ISRA and the EM algorithm for positron emission tomography. Medical Imaging, IEEE Transactions on 12, 328–333.

Dell'Acqua F, Rizzo G, Scifo P, Clarke RA, Scotti G, Fazio F, 2007. A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. Biomedical Engineering, IEEE Transactions on 54, 462–472.

Dell'Acqua F, Scifo P, Rizzo G, Catani M, Simmons A, Scotti G, Fazio F, 2010. A modified damped Richardson–Lucy algorithm to reduce isotropic background effects in spherical deconvolution. Neuroimage 49, 1446–1458. [PubMed: 19781650]

Descoteaux M, Angelino E, Fitzgibbons S, Deriche R, 2007. Regularized, Fast and Robust Analytical Q-ball Imaging. Magnetic Resonance in Medicine 58, 497–510. [PubMed: 17763358]

Descoteaux M, Deriche R, Bihan D, Mangin J, Poupon C, 2010. Multiple q-Shell Diffusion Propagator Imaging. Medical Image Analysis 15, 603–621. [PubMed: 20685153]

Descoteaux M, Wiest-Daessle N, Prima S, Barillot C, Deriche R, 2008. Impact of Rician Adapted Non-local Means Filtering on HARDI, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI).

Goh A, Lenglet C, Thompson P, Vidal R, 2009. Estimating orientation distribution functions with probability density constraints and spatial regularity, in: MICCAI, Springer. pp. 877–885.

Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB, 2006. Q-Ball Reconstruction of Multimodal Fiber Orientations Using The Spherical Harmonic Basis. Magnetic Resonance In Medicine 56, 104–117. [PubMed: 16755539]

Jbabdi S, Sotiropoulos SN, Savio AM, Granã M, Behrens TE, 2012. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. Magnetic Resonance in Medicine 68, 1846–1855. [PubMed: 22334356]

Jian B, Vemuri BC, 2007. A Unified Computational Framework for Deconvolution to Reconstruct Multiple Fibers from Diffusion Weighted MRI. IEEE Transactions on Medical Imaging 26, 1464–1471. [PubMed: 18041262]

Johansen-Berg H, Behrens TE, 2009. Diffusion MRI: From quantitative measurement to In vivo neuroanatomy. Elsevier.

Kim SJ, Koh K, Lustig M, Boyd S, Gorinevsky D, 2007. An Interior-Point Method for Large-Scale l1-Regularized Least Squares. Selected Topics in Signal Processing, IEEE Journal of 1, 606–617. doi:10.1109/JSTSP.2007.910971.

Krajsek K, Scharr H, 2012. A Riemannian approach for estimating orientation distribution function (ODF) images from high-angular resolution diffusion imaging (HARDI), in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, IEEE. pp. 1019–1026.

Landman B, Bogovic J, Wan H, El Zahraa E, Bazin P, Prince J, 2012. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI. NeuroImage 59, 2175. [PubMed: 22019877]

Lee H, Battle A, Raina R, Ng A, 2006. Efficient sparse coding algorithms, in: Advances in neural information processing systems, pp. 801–808.

Özarslan E, Koay C, Shepherd T, Blackband S, Basser P, 2009. Simple harmonic oscillator based reconstruction and estimation for three-dimensional q-space MRI, in: International Society for Magnetic Resonance in Medicine (ISMRM), p. 1396.

Raffelt D, Tournier JD, Crozier S, Connelly A, Salvado O, 2012. Reorientation of Fiber Orientation Distributions using Apodized Point Spread Functions. Magnetic Resonance in Medicine 67, 844–855. [PubMed: 22183751]

Schwab E, Afsari B, Vidal R, 2012. Estimation of Non-negative ODFs Using the Eigenvalue Distribution of Spherical Functions, in: Ayache N, Delingette H, Golland P, Mori K (Eds.), Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer Berlin Heidelberg. pp. 322–330.

Shi F, Yap PT, Gao W, Lin W, Gilmore JH, Shen D, 2012. Altered structural connectivity in neonates at genetic risk for schizophrenia: A combined study using morphological and white matter networks. NeuroImage 62, 1622–1633. [PubMed: 22613620]

Sotiropoulos SN, Jbabdi S, Xu J, Andersson JL, Moeller S, Auerbach EJ, Glasser MF, Hernandez M, Sapiro G, Jenkinson M, et al., 2013. Advances in diffusion MRI acquisition and processing in the Human Connectome Project. NeuroImage 80, 125–43. [PubMed: 23702418]

Tournier JD, Calamante F, Connelly A, 2007. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. NeuroImage 35, 1459–1472. [PubMed: 17379540]

Tournier JD, Calamante F, Connelly A, 2012. MRtrix: Diffusion tractography in crossing fiber regions. International Journal of Imaging Systems and Technology 22, 53–66.

Tournier JD, Calamante F, Gadian D, Connelly A, 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. NeuroImage 23, 1176–1185. [PubMed: 15528117]

Tuch DS, 2004. Q-ball imaging. Magnetic Resonance in Medicine 52, 1358–1372. [PubMed: 15562495]

Tuch DS, Reese TG, Wiegell MR, Makris N, Belliveau JW, Wedeen VJ, 2002. High Angular Resolution Diffusion Imaging Reveals Intravoxel White Matter Fiber Heterogeneity. Magnetic Resonance in Medicine 48, 577–582. [PubMed: 12353272]

Wang Q, Yap PT, Wu G, Shen D, 2013. Application of neuroanatomical features to tractography clustering. Human Brain Mapping 34, 2089–2102. [PubMed: 22461221]

Wedeen VJ, Hagmann P, Tseng WYI, Reese TG, Weisskoff RM, 2005. Mapping Complex Tissue Architecture With Diffusion Spectrum Magnetic Resonance Imaging. Magnetic Resonance In Medicine 54, 1377–1386. [PubMed: 16247738]

Wee CY, Yap PT, Li W, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D, 2010. Enriched white matter connectivity networks for accurate identification of MCI patients. NeuroImage 54, 1812–1822. [PubMed: 20970508]

Wee CY, Yap PT, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D, 2012. Identification of MCI individuals using structural and functional connectivity networks. NeuroImage 59, 2045–2056. [PubMed: 22019883]

Weldeselassie Y, Barmpoutis A, Atkins M, 2010. Symmetric Positive-Definite Cartesian Tensor Orientation Distribution Functions (CT-ODF), in: Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2010.

Weldeselassie YT, Barmpoutis A, Stella Atkins M, 2012. Symmetric positive semi-definite Cartesian Tensor fiber orientation distributions (CT-FOD). Medical Image Analysis 16, 1121–1129. [PubMed: 22890050]

Yap PT, An H, Chen Y, Shen D, 2014. Uncertainty estimation in diffusion MRI using the non-local bootstrap. IEEE Transactions on Medical Imaging, to appear.

Yap PT, Fan Y, Chen Y, Gilmore J, Lin W, Shen D, 2011a. Development trends of white matter connectivity in the first years of life. PLoS ONE 6, e24678. [PubMed: 21966364]

Yap PT, Gilmore J, Lin W, Shen D, 2011b. PopTract: Population-based tractography. IEEE Transactions on Medical Imaging 30, 1829–1840. [PubMed: 21571607]

Yap PT, Gilmore JH, Lin W, Shen D, 2011c. Longitudinal tractography with application to neuronal fiber trajectory reconstruction in neonates, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 66–73.

Yap PT, Shen D, 2012. Spatial Transformation of DWI Data Using Non-Negative Sparse Representation. IEEE Transaction on Medical Imaging 31, 2035–2049.

Yap PT, Wu G, Shen D, 2010. Human brain connectomics: Networks, techniques and applications. IEEE Signal Processing Magazine 27, 131–134.

Highlights

1. Improved fODF estimation with non-negative and unit integral constraints.
2. Significantly increased anisotropy contrast between WM and GM/CSF regions.
3. Improvements over state-of-the-art methods, such as constrained spherical deconvolution.

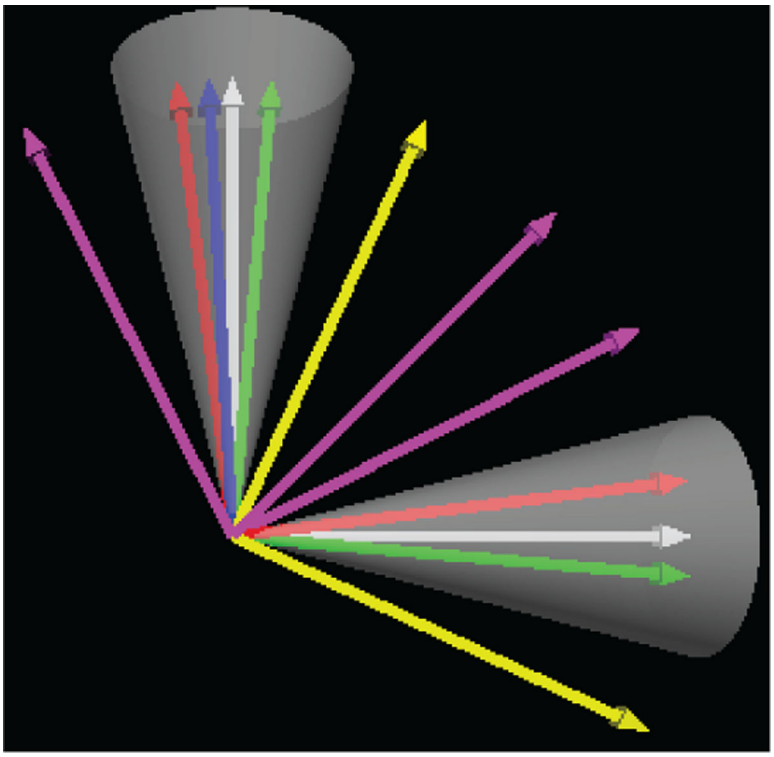


Figure 1: Peak Consistency.

The peaks detected from each fODF are indicated by identically colored arrows. The cones cover peak orientations subtending an angle θ_C with the white arrows. The white peaks are said to be consistent with the peaks within these cones. In this case, SPC = 3/5 and SPC = 2/5, respectively for the vertical and horizontal white peaks, and MPC = 2/5.

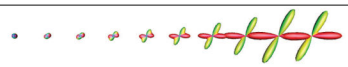
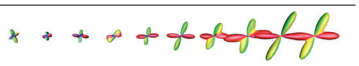
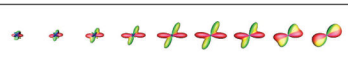
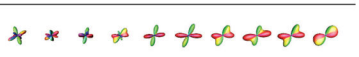
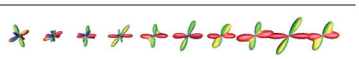
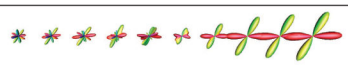
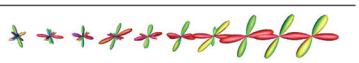
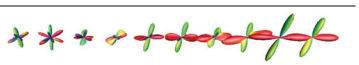
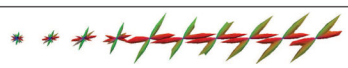
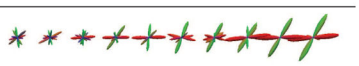
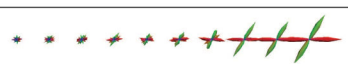
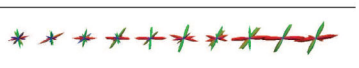
	SNR=1000	SNR=30
NNSD $L = 6, \delta = 10^{-2}$		
NNSD $L = 6, \delta = 10^{-4}$		
ASC-NNSD $L = 6, T = 0.5$		
QP-CSD $L = 6$		
QP-CSD $L = 10$		
CSD $L = 6$		
CSD $L = 10$		
MESD $K = 16$		
L1-NNLS $\lambda = 0.1\lambda^*_{L1}$		
L1-NNLS $\lambda_{L1} = 0$		

Figure 2: Synthetic Data.

fODF estimation results given by different methods under different parameter settings for synthetic data with varying anisotropy.

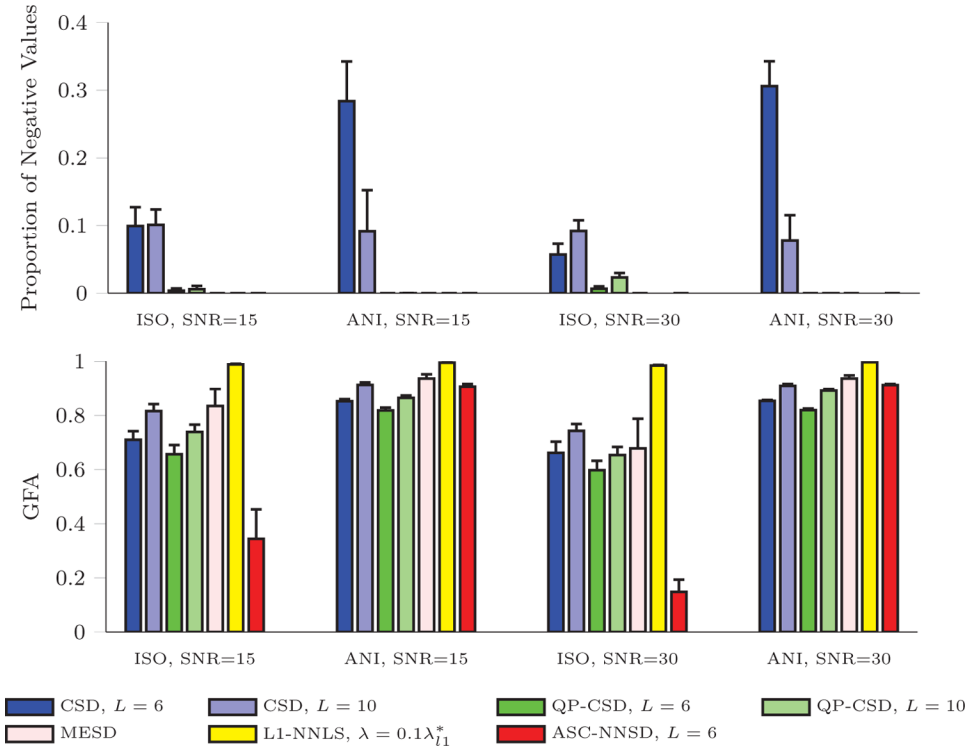


Figure 3: Nonnegativity and Anisotropy.

Proportions of negative values and GFA values of fODFs estimated from isotropic (ISO) and anisotropic (ANI) data. The error bars indicate standard deviations.

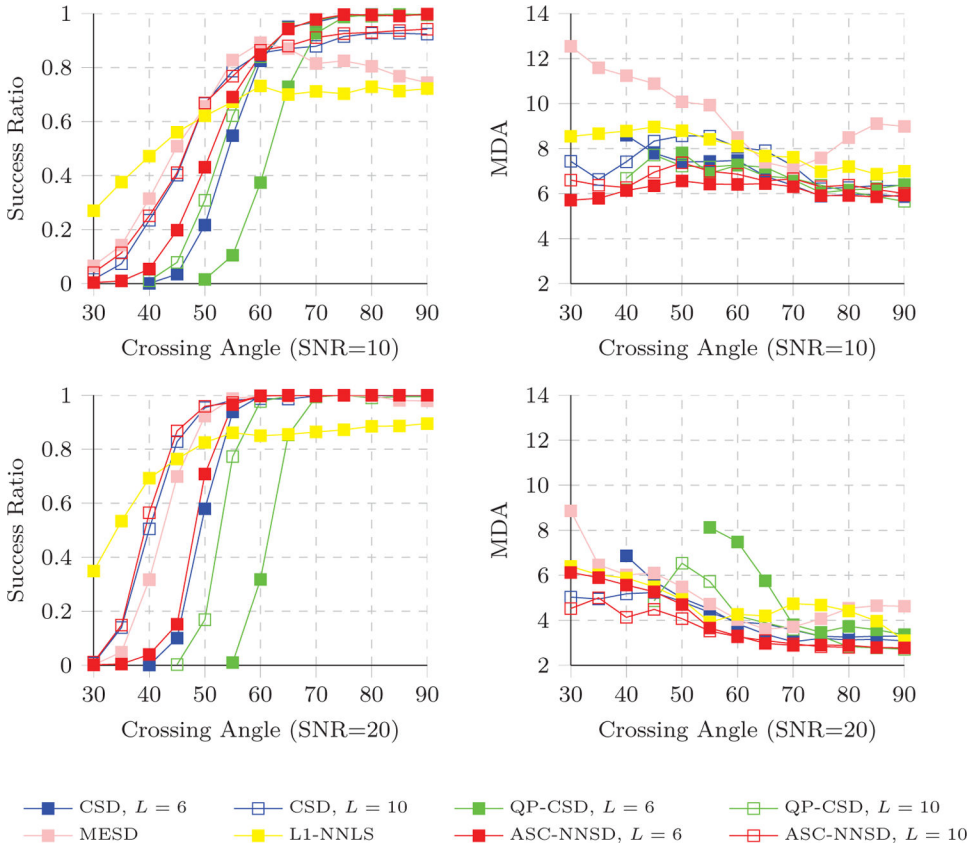


Figure 4: Peak Accuracy – Single Shell.

Success ratios and MDA values for different crossing angles at SNR=10 and SNR=20.

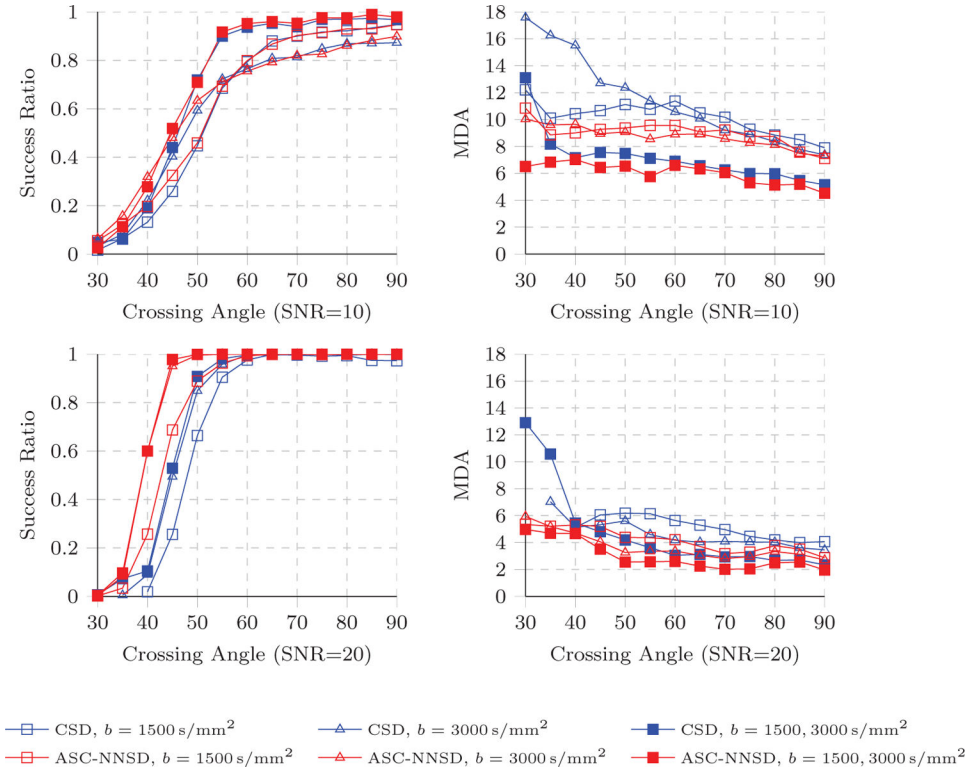


Figure 5: Peak Accuracy – Multiple Shells.

Success ratios and MDA values for different crossing angles at SNR=10 and SNR=20.

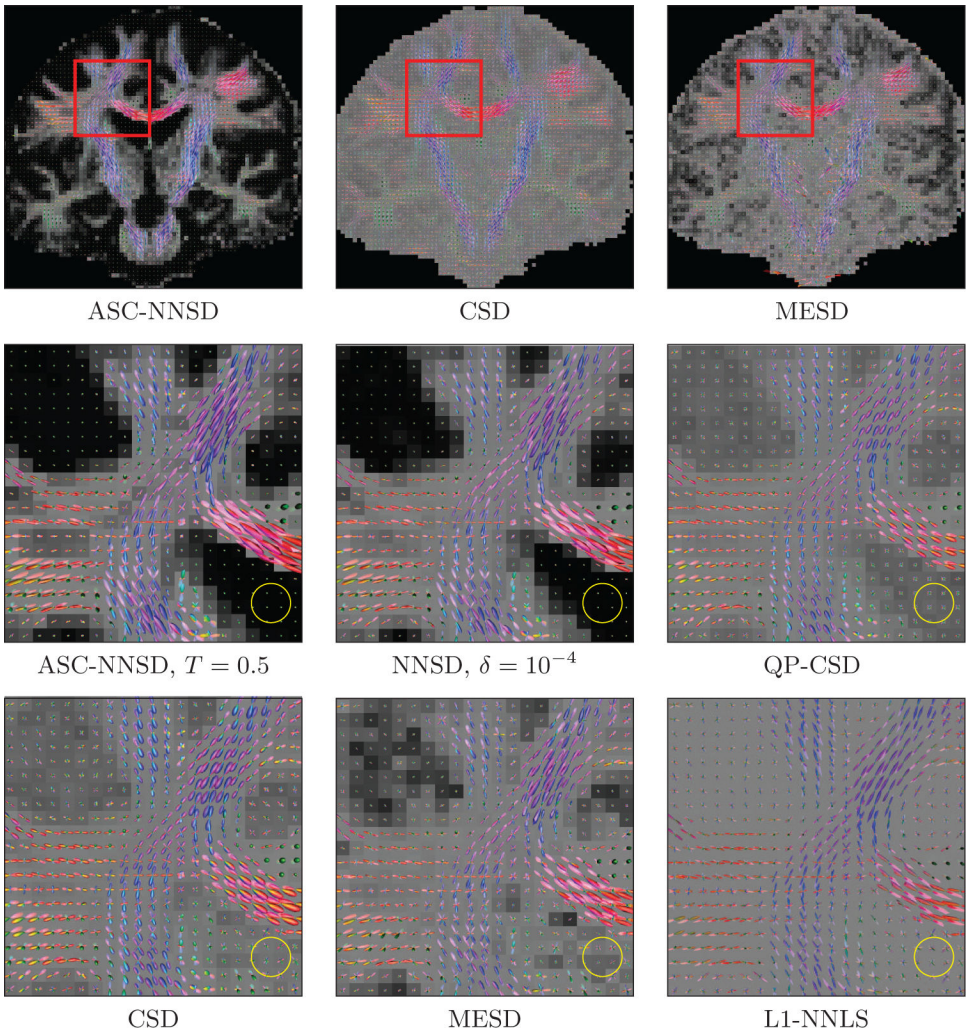


Figure 6: Real Data – Single Shell.

Coronal views and close-up views of the fODF fields generated by various methods. The background is the GFA map.

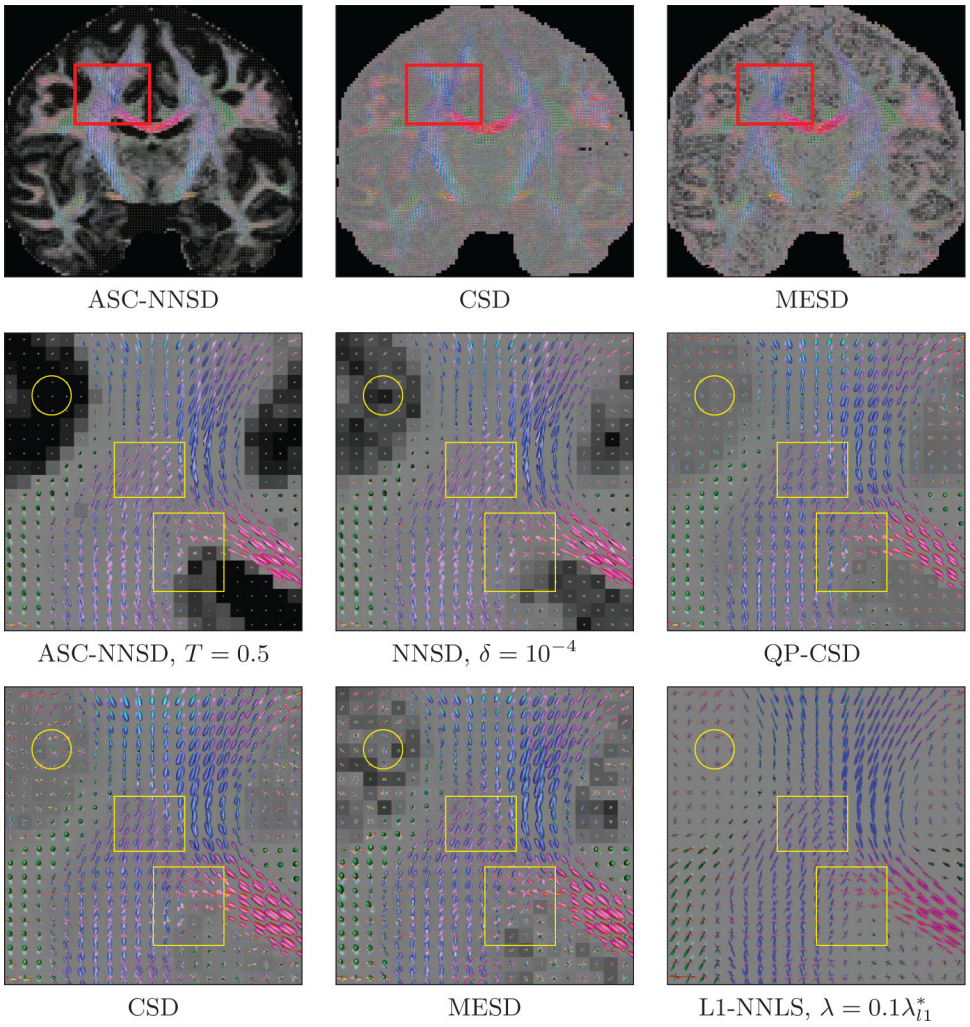


Figure 7: HCP Data – Single Shell.

Coronal views and close-up views of the fODF fields generated by various methods from single shell data with $b = 2000\text{s/mm}^2$. The background is the GFA map.

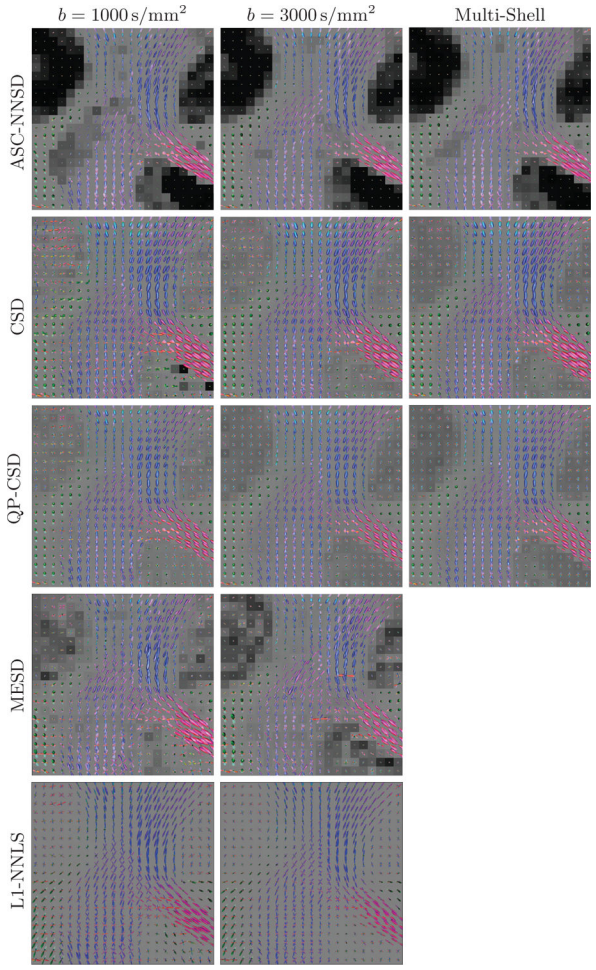


Figure 8: Single Shell vs. Multiple Shells.

Estimation results for different methods using data from each shell (90 samples) and concurrently from the three shells (90 × 3 samples). See Fig. 7 for the $b = 2000\mathrm{s/mm^2}$ results.

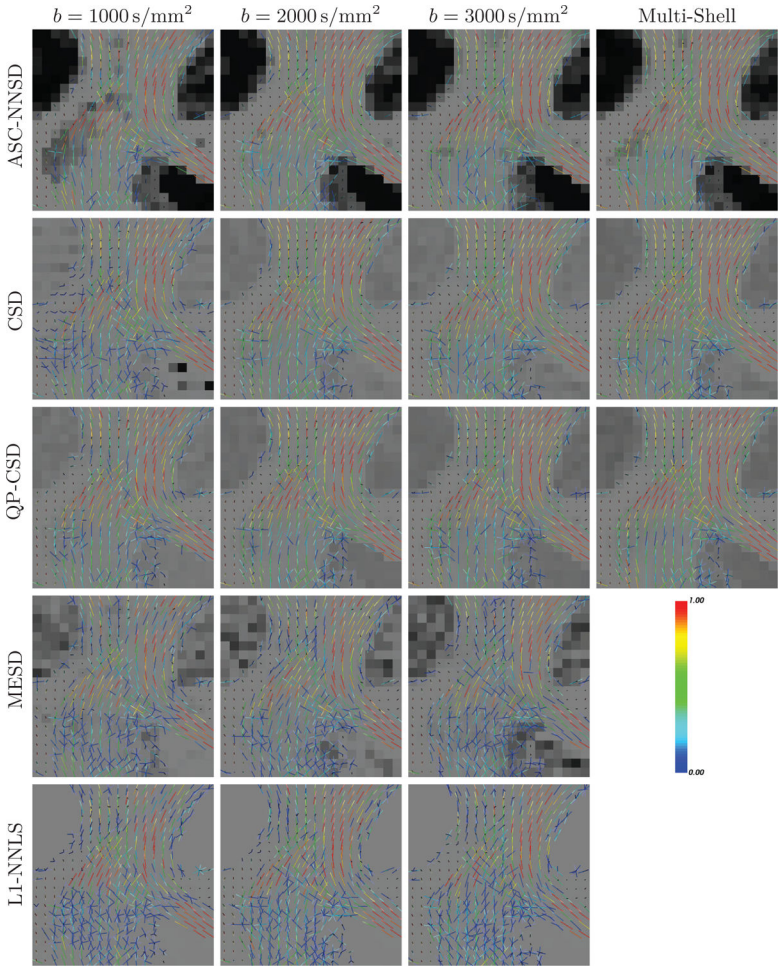


Figure 9: Single Peak Consistency.

Single Peak Consistency (SPC) for different methods and different sampling schemes. Cool and warm colors indicate low and high SPC, respectively.

	SNR=1000	SNR=30
$L = 6,\ \delta = 10^{-3}$		
$L = 6,\ \delta = 10^{-4}$		
$L = 6,\ \delta = 10^{-5}$		

Figure 10: NNSD without unit integral constraint.

Estimation results for NNSD without unit integral constraint for mixture of tensor models with varying anisotropy.

Table 1:

$\overline{\text{SPC}}$ and $\overline{\text{MPC}}$ values for fODF fields estimated by different methods for different sampling schemes.

	b = 1000 s/mm^2		b = 2000 s/mm^2		b = 3000 s/mm^3		Multi-Shell	
	$\overline{\text{SPC}}$	$\overline{\text{MPC}}$	$\overline{\text{SPC}}$	$\overline{\text{MPC}}$	$\overline{\text{SPC}}$	$\overline{\text{MPC}}$	$\overline{\text{SPC}}$	$\overline{\text{MPC}}$
ASC-NNSD	**0.2411**	**0.2503**	**0.2839**	**0.2931**	**0.2857**	**0.2995**	**0.3845**	**0.3745**
CSD	0.1533	0.1870	0.2245	0.2569	0.2359	0.2793	0.2779	0.3236
QP-CSD	0.1811	0.2319	0.2158	0.2720	0.1938	0.2593	0.2597	0.3218
MESD	0.1347	0.1752	0.1399	0.1815	0.1600	0.2102		
L1-NNLS	0.1080	0.1571	0.1148	0.1650	0.1030	0.1653		