Estimation of the CSA-ODF using Bayesian Compressed Sensing of Multi-shell HARDI

Julio M. Duarte-Carvajalino[a], **Christophe Lenglet**[a], **Junqian Xu**[a,b], **Essa Yacoub**[a], **Kamil Ugurbil**[a], **Steen Moeller**[a], **Lawrence Carin**[c], and **Guillermo Sapiro**[c,d,*]

[a]Radiology – CMRR, University of Minnesota, Minneapolis MN, USA

[b]Icahn School of Medicine at Mount Sinai, New York, NY, USA

[c]Electrical and Computer Engineering, Duke University, Durham, NC, USA

[d]Biomedical Engineering, and Computer Science, Duke University, Durham, NC, USA

Abstract

Purpose—Diffusion MRI provides important information about the brain white matter structures and has opened new avenues for neuroscience and translational research. However, acquisition time needed for advanced applications can still be a challenge in clinical settings. There is consequently a need to accelerate diffusion MRI acquisitions.

Methods—A multi-task Bayesian compressive sensing (MT-BCS) framework is proposed to directly estimate the constant solid angle orientation distribution function (CSA-ODF) from under-sampled (i.e., accelerated image acquisition) multi-shell high angular resolution diffusion imaging (HARDI) datasets, and accurately recover HARDI data at higher resolution in q-space. The proposed MT-BCS approach exploits the spatial redundancy of the data by modeling the statistical relationships within groups (clusters) of diffusion signal. This framework also provides uncertainty estimates of the computed CSA-ODF and diffusion signal, directly computed from the compressive measurements. Experiments validating the proposed framework are performed using realistic multi-shell synthetic images and *in-vivo* multi-shell high angular resolution HARDI datasets.

Results—Results indicate a practical reduction in the number of required diffusion volumes (q-space samples) by at least a factor of four to estimate the CSA-ODF from multi-shell data.

Conclusion—This work presents, for the first time, a multi-task Bayesian compressive sensing approach to simultaneously estimate the full posterior of the CSA-ODF and diffusion-weighted volumes from multi-shell HARDI acquisitions. It demonstrates improvement of the quality of acquired datasets via CS de-noising, and accurate estimation of the CSA-ODF, as well as enables a reduction in the acquisition time by a factor of two to four, especially when "staggered" q-space sampling schemes are used. The proposed MT-BCS framework can naturally be combined with parallel MR imaging to further accelerate HARDI acquisitions.

*Corresponding author. Tel.: 919-660-5252; fax: 919-660-5293; Guillermo.sapiro@duke.edu.

Keywords

Bayesian Compressed Sensing (BCS); Constant Solid Angle (CSA); Dirichlet process; High Angular Resolution Diffusion Imaging (HARDI); Multi-task; Orientation Distribution Function (ODF)

INTRODUCTION

Diffusion magnetic resonance imaging (MRI) can characterize the anisotropic thermal motion of water molecules in biological tissues, such as brain white matter. This information can be used to estimate the local orientation of white matter fiber bundles. Diffusion tensor imaging (DTI) (1) was introduced to describe the three-dimensional nature of anisotropic diffusion in biological tissues. However, the diffusion tensor model is inadequate when two or more fiber bundles cross (at the acquisition resolution). It is estimated that between one third and two-thirds of imaging voxels in the human brain's white matter contain multiple fiber crossings (2). High angular resolution diffusion imaging (HARDI) overcomes limitations of DTI for characterizing complex tissue geometries such as fiber crossings, by measuring the diffusion-weighted signal along tens to hundreds of directions (q-space). The high angular resolution of HARDI allows estimating the orientation distribution function (ODF), which is the average probability of diffusion along any diffusion direction, therefore enabling fiber crossing detection (3).

HARDI requires a relatively large number of diffusion encodings to accurately recover the ODF (4), limiting its applicability to clinical studies. The use of more complex excitation techniques combined with parallel imaging techniques (pMRI) (5–8) has reduced acquisition times of diffusion weighted MRI (DW-MRI), typically by a factor of three. Recently, pMRI and compressed sensing (CS)[1] methods have been combined to reduce the acquisition time of structural MRI (9–16), diffusion spectrum imaging (DSI) [14–19], HARDI (23–25), and dynamic MRI (26–34), leading to combined acceleration factors of 9 and higher.

Compressed sensing of HARDI data sets has been proposed in the past (24,35–37), where the diffusion directions are pseudo-randomly under-sampled and the skipped directions are recovered (interpolated) using sparse representation models and CS reconstruction methods. However, these under-sampled measurements are not linear combination of the diffusion signal (just a subsample of them), and hence, do not fully exploit the power of CS as an inverse model reconstruction technique. More recent work directly estimated the ODF from under-sampled single-shell (38) and multi-shell HARDI data (21–23,25,39) using CS methods. Under-sampled HARDI data can be seen as a CS measurement of the ensemble average propagator (EAP) (21–23,25,39) or the ODF (38), fully exploiting the power of CS. However, most of these previous works rely on the classical definition of the ODF (3). In this work, we develop a multi-task Bayesian compressive sensing framework for DW-MRI, following our previous work by Duarte-Carvajalino et al. (37), to directly estimate the constant solid angle ODF (CSA-ODF) (40) from multi-shell compressed sensed HARDI

[1]See for example rich references for CS in http://dsp.rice.edu/cs.

data sets. The CSA-ODF was recently introduced to update the commonly used ODF definition (3), providing a mathematically correct probabilistic distribution function (PDF) expression for the ODF. The classical definition of the ODF ignores the quadratic growth of the volume element in *q*-space with respect to its distance from the origin, leading to an inaccurate formulation that requires artificial post-processing such as normalization and sharpening (40–43). The CSA-ODF accounts for this quadratic growth, leading to a correct PDF formulation that does not require artificial normalization (integrating to one) or sharpening.

Compared to the work by Tristán-Vega et al. (38), we use multi-shell data, with signal-to-noise ratios (SNRs) commonly found in in-vivo HARDI data,[2] and a Bayesian CS approach (37,44). In addition, previous work on CS of the CSA-ODF provide only maximum *a posteriori* (MAP) (45) estimates of the ODF and/or diffusion-weighted volumes. We use here Bayesian inference to estimate the full posterior distributions of the CSA-ODF and diffusion-weighted volumes, providing point estimates, as well as uncertainty estimates of those quantities. The uncertainty estimates allow us to evaluate the confidence on the reconstructed data. Since the reconstructed point estimates may have large variance, the additional characterization of full posterior distribution, information not available with classical MAP CS methods, is as important as the reconstructed point estimates.

In addition to the proposed Bayesian compressed sensing of the CSA-ODF (bcsCSA-ODF), we provide comparison with the CSA-ODF (40) implemented in the FSL library[3] (46) to compute the ODFs from the original and under-sampled HARDI data sets. CSA-ODF relies on a bi-exponential model for the diffusion signal decay and analytic solution for multi-shell HARDI data (with *b*-values in arithmetic progression). We also compare with a generalized CSA-ODF (gCSA-ODF) bi-exponential algorithm (41) that can handle under-sampled HARDI, with arbitrary *q*-space sampling, and computes the CSA-ODFs and "missing" diffusion-weighted volumes by using optimally regularized SHs interpolation and the bi-exponential model. Both gCSA-ODF and CSA-ODF are established data-fitting techniques which do not explicitly exploit signal sparsity and have performances that can more readily be compared with the more sophisticated CS algorithm. The comparison with more conventional algorithms is used to illustrate the advantage of applying a sophisticated CS algorithm. The gCSA-ODF and bcsCSA-ODF models have been implemented in FSL and will be available on the NITRC web site.

The results demonstrate that, at realistic SNR levels, the proposed Bayesian CS framework performs better than using simpler angular interpolation methods such as the CSA-ODF and gCSA-ODFs for multi-shell images. We demonstrate the feasibility of acceleration factors equal to or greater than four.

---

[2]The authors of (38) use SNRs of 100 and 40 in their synthetic data.
[3]http://fsl.fmrib.ox.ac.uk/fsl/fsl-4.1.9/fdt/fdt_utils.html

METHODS

Constant Solid Angle ODF

For a given spatial location (voxel) of a HARDI dataset, the probability of diffusion along the direction $\hat{\mathbf{u}}$ can be modeled by the constant solid angle orientation distribution function (CSA-ODF) $\varphi(\hat{\mathbf{u}})$ (38,40)

$$\phi(\hat{\mathbf{u}}) = -\frac{1}{8\pi^2}\mathscr{R}_{\hat{\mathbf{u}}}\{\nabla^2 x(\mathbf{q})\}, \quad [1]$$

where, $\mathscr{R}_{\hat{\mathbf{u}}}\{f(\mathbf{q})\} = \iint f(\mathbf{q})\delta(\mathbf{q}^{\mathsf{T}}\hat{\mathbf{u}})d\mathbf{q}$ is the Funk-Radon transform, δ the Delta-Dirac function, $x(\mathbf{q}) = S(\mathbf{q})/S(0)$ the normalized signal attenuation at a given diffusion direction $\hat{\mathbf{q}}$ and magnitude $q(\mathbf{q} = q\hat{\mathbf{q}})$, $S(\mathbf{q})$ is the measured HARDI signal in q-space, and $S(0)$ the non diffusion-weighted ($\mathbf{q} = \mathbf{0}$) signal. Limiting the diffusion signal to concentric spheres and introducing a radial mono-exponential model, we obtain $x(\mathbf{q})$ for any $\mathbf{q}$ from a given measurement $x(q_0\hat{\mathbf{q}})$ at a specific q_0 value as $x(\mathbf{q}) = [x(q_0\hat{\mathbf{q}})]^{q^2/q_0^2}$ (40). With this assumption, it has been shown that [1] is equivalent to (40)

$$\phi(\hat{\mathbf{u}}) = \frac{1}{4\pi} + \frac{1}{16\pi^2}\mathscr{R}_{\hat{\mathbf{u}}}\left\{\nabla_b^2 ln\left(-ln\left(x(\mathbf{q})\right)\right)\right\}, \quad [2]$$

where $\nabla_b^2 = \frac{1}{\tan\theta}\frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial^2\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\varphi}$ is the Laplace-Beltrami operator in spherical coordinates (θ, φ). In practice, E in [2] is regularized in such a way that $ln(-ln(x))$ is smooth (40).

We should emphasize that there is, in practice, no need to carry the numerical derivation and integration indicated in [2] to compute the CSA-ODF. The computation of the CSA-ODF is greatly facilitated by the fact that the spherical harmonics (SH) basis are eigen-functions of the Funk Radon transform (47–49). Indeed, with $\tilde{E} = ln(-ln(x(\mathbf{q})))$, $\tilde{E}$ can be represented in the SH basis as

$$\tilde{x} = ln\left(-ln\left(x(\mathbf{q})\right)\right) \approx \sum_{r=1}^{R} c_r Y_r(\hat{\mathbf{u}}), \quad [3]$$

where R is the number of elements of the SH basis used in the approximation, $Y_r(\hat{\mathbf{u}})$ the real and symmetric modified SH basis (49), and c_r the SH coefficients computed using least-squares. Then, the CSA-ODF can be obtained *explicitly* from the SH representation of $\tilde{E}$ as (40,49)

$$\phi(\hat{\mathbf{u}}) = \sum_{r=1}^{R} c_r' Y_r(\hat{\mathbf{u}}), \quad c_r' = \begin{cases} \frac{1}{2\sqrt{\pi}}, & r=1 \\ -\frac{1}{8\pi}(-1)^{\frac{k_r}{2}}\frac{1\times3\times\cdots\times(k_r+1)}{2\times4\times\cdots\times(k_r-2)}c_r, & r>1 \end{cases}, \quad [4]$$

where $-k_r(k_r + 1)$ are the eigenvalues of the Laplace-Beltrami operator. This result indicates that there is an explicit linear relationship between the SH representation of the CSA-ODF

and the SH representation of the double logarithm of the diffusion signal. We build on this in the next sections.

Bayesian Compressed Sensing of the CSA-ODF (bcsCSA-ODF)

Consider a multi-shell HARDI experiment, with a total of M diffusion directions (encodings) spread over B shells and N voxels. Let $\mathbf{x}$ be the $M \times 1$ measured HARDI attenuations at a given voxel in vector form and $\mathbf{Y}_{M \times R}$ the SH basis, with R harmonic coefficients. Then [3] and [4] can be expressed in the SH basis using matrix-vector notation as[4]

$$(\phi)_{\text{SH}} \approx \begin{bmatrix} 0 & 0 \\ 0 & \Lambda \end{bmatrix}_{R \times R} \mathbf{Y}^+_{R \times M} \tilde{\mathbf{x}} + \begin{bmatrix} 1/\left(2\sqrt{\pi}\right) \\ 0 \end{bmatrix}_{R \times 1}, \tilde{\mathbf{x}} = \left[ln\left(-ln(x_j)\right)\right]_{1 \le j \le M}, \quad [5]$$

where $(\varphi)_{\text{SH}}$ corresponds to the SH representation of the CSA-ODF, and $\mathbf{Y}^+ = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T$ is the pseudo-inverse of the SH basis $\mathbf{Y}$. Hence, $\mathbf{Y}^+\tilde{\mathbf{x}}$ provides the (least-squares) estimated SH coefficients c_r indicated in [3], and Λ provides the scaling factors multiplying the SH coefficients c_r in [4], for $r > 1$. Notice that for $r = 1$, the scaling factor is zero, since the Laplace-Beltrami operator eliminates the first harmonic coefficient c_1 corresponding to the (constant) average component of the signal $\tilde{\mathbf{x}}$. Hence, the $1/\left(2\sqrt{\pi}\right)$ term in [5] is required to obtain $c_1' = \frac{1}{2\sqrt{\pi}}$ in [4].

Equation [5] is also equivalent to

$$(\phi)_{\text{SH}} \approx \begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \Lambda \end{bmatrix} \mathbf{Y}^+\tilde{\mathbf{x}} + \begin{bmatrix} 1/\left(2\sqrt{\pi}\right) - \gamma^{-1}c_1 \\ 0 \end{bmatrix}, \quad [6]$$

for any $\gamma \quad 0$. We introduce here the variable γ in order to make the matrix $\begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \Lambda \end{bmatrix}$ invertible.

Given that one key component in compressive sensing is the sparse representation of the signal in a sparsifying basis or dictionary,[5] and the SH basis does not constitute a good sparsifying basis for HARDI data (24,35,50), we represent the measured HARDI attenuations $\mathbf{x}$ using a more adequate sparsifying dictionary

$$\mathbf{x} = \mathbf{\Psi}_{M \times L} \vartheta + \eta, \quad [7]$$

where $\vartheta(L \times 1)$ is the sparse representation of the attenuation signal at a given voxel in a sparsifying dictionary $\mathbf{\Psi}$, and η is the representation noise at that voxel. As previously reported (24,35,37,50), we choose $\mathbf{\Psi}$ to be the Ridgelets basis (though we could learn this

---

[4]This is the actual implementation of [2] in FSL.
[5]See http://dsp.rice.edu/cs for some additional references on compressive sensing.

from data if desired), since it has been shown that it constitutes a good sparsifying dictionary for HARDI data (24,35,50). Replacing **x** given by [7] in [1]

$$\phi \approx -\frac{1}{8\pi^2}\mathscr{R}_{\hat{\mathbf{u}}}\{\nabla_b^2(\mathbf{\Psi}(\hat{\mathbf{u}})\boldsymbol{\vartheta})\}, \quad [8]$$

where we have used the fact that $\mathbf{\Psi}(\hat{\mathbf{u}})$ has no radial component, hence the Laplacian reduces to the Laplace-Beltrami operator. Since the Laplace-Beltrami and Funk-Radon transforms commute,

$$\phi \approx -\frac{1}{8\pi^2}\nabla_b^2\mathscr{R}_{\hat{\mathbf{u}}}\{\mathbf{\Psi}(\hat{\mathbf{u}})\boldsymbol{\vartheta}\}=-\frac{1}{8\pi^2}\nabla_b^2(\mathscr{H}(\hat{\mathbf{u}})\boldsymbol{\vartheta}), \quad [9]$$

where $\mathscr{R}_{\hat{\mathbf{u}}}\{\mathbf{\Psi}(\hat{\mathbf{u}})\boldsymbol{\vartheta}\} = \mathscr{H}(\hat{\mathbf{u}})\boldsymbol{\vartheta}$, and $\mathscr{H}(\hat{\mathbf{u}})$ corresponds to the Funk-Radon transform of the Ridgelets basis, as specified in (50). We can use finite differences to approximate [9] (see Appendix) and express the CSA-ODF in vector form as $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$, where **D** is a dictionary for the CSA-ODF, derived from the Ridgelets basis. The equations to be derived here apply to any CSA-ODF dictionary (including learned ones); we use the Ridgelets basis for convenience (see next section).

We have now two representations of the CSA-ODF. The first in terms of the SH basis ([6]), and another one in terms of a sparsifying dictionary ([9]) that can be represented in matrix-vector form as $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$. Since, it is the same CSA-ODF, both representations should be equivalent. However, since $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$ is in a different basis, we need to transform this representation in terms of the SH coefficients so that both representations are in the same basis. The SH representation of $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$ can be obtained with the pseudo-inverse as $(\boldsymbol{\varphi})_{SH} \approx \mathbf{Y}^+(\mathbf{D}\boldsymbol{\vartheta})$. Representing $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$ in terms of the SH basis does not affect the sparse representation $\boldsymbol{\vartheta}$, $(\boldsymbol{\varphi})_{SH} \approx (\mathbf{Y}^+\mathbf{D})\boldsymbol{\vartheta}$, where $\mathbf{Y}^+\mathbf{D}$ corresponds to the new sparsifying dictionary in the SH basis. Hence, [6] and [9] must correspond to the CSA-ODF in SH coordinates,

$$(\phi)_{\text{SH}} \approx \begin{bmatrix} \gamma^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda} \end{bmatrix}\mathbf{Y}^+\tilde{\mathbf{x}}+\begin{bmatrix} 1/(2\sqrt{\pi}) -\gamma^{-1}c_1 \\ \mathbf{0} \end{bmatrix} \approx \mathbf{Y}^+\mathbf{D}\boldsymbol{\vartheta}. \quad [10]$$

We can solve for $\tilde{\mathbf{x}}$ in [10] as (see Appendix)

$$\mathbf{x}^* \approx \mathbf{PDJ}, \quad \mathbf{x}^* \approx (\tilde{\mathbf{x}}-c_1\mathbf{Y}_1)+\gamma/(2\sqrt{\pi})\,\mathbf{Y}_1, \mathbf{P}=\mathbf{Y}\begin{bmatrix} \gamma & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{-1} \end{bmatrix}\mathbf{Y}^+, \quad [11]$$

where $\mathbf{Y}_1$ corresponds to the first column of the SH basis. $(\tilde{\mathbf{x}} - c_1\mathbf{Y}_1)$ in [11] corresponds to the variable component of $\tilde{\mathbf{x}}$, since the term $c_1\mathbf{Y}_1$ is the mean value of $\tilde{\mathbf{x}}$. $\begin{bmatrix} \gamma & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{-1} \end{bmatrix}$ corresponds to the inverse of $\begin{bmatrix} \gamma^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda} \end{bmatrix}$ in [6], explaining the auxiliary variable γ 0 in [5].

Eq. [11] implicitly uses a set of M diffusion directions $U_M = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_M\}$ that define the CSA-ODF signal $\mathbf{x}^*(U_M)$. However, this equation applies to any number of diffusion directions. Let $U_m = \{\hat{\mathbf{u}}_1, \cdots, \hat{\mathbf{u}}_m\} \subset U_m$ be the diffusion directions of an under-sampled HARDI experiment, with $m < M$ diffusion directions. Let $\mathbf{x}_s^* = \mathbf{x}^*(U_m)$, $\mathbf{P}_s = \mathbf{P}(U_m)$, $\mathbf{D}_s = \mathbf{D}(U_m)$ be respectively, the under-sampled HARDI diffusion signal, and $\mathbf{P}_s$, $\mathbf{D}_s$ the projection matrix and dictionary of the CSA-ODF $\varphi(U_m)$, then [11] for under-sampled data becomes

$$\mathbf{x}_s^* \approx \mathbf{P}_s \mathbf{D}_s \boldsymbol{\vartheta}, \ \mathbf{x}_s^* = \tilde{\mathbf{x}}(U_m) - c_1 \boldsymbol{Y}_1(U_m) + \gamma / \left(2\sqrt{\pi}\right) \boldsymbol{Y}_1(U_m). \quad [12]$$

A key requirement in CS is that the sampling should be incoherent with the signal of interest (here, the CSA-ODF). This requirement is satisfied with high probability when the sampling is random or pseudo-random (see, for instance (10,17–19,21,22,24,25,35,38,45,50)). As in our previous work (37), we use the optimal pseudo-random Dirac sampling of multi-shell HARDI data proposed by Caruyer et al. (4), which provides diffusion directions that uniformly cover the unit sphere when truncated before the acquisition is complete and satisfy the incoherence principle (35).

Notice that since the signal of interest is the CSA-ODF $\boldsymbol{\varphi} \approx \mathbf{D}\boldsymbol{\vartheta}$. Then with respect to the CSA-ODF, $\mathbf{P}_s$ corresponds to the CS matrix and $\mathbf{x}_s^*$ corresponds to the compressed samples of the CSA-ODF, i.e., $\mathbf{x}_s^*$ is an under-sampled linear combination of the full CSA-ODF ([12]), as required to exploit CS. Note that it is not just a sub-sampled signal as in more standard approaches (21–23,25,38,39). On the other hand, $\mathbf{x}^*$ is also a signal of interest and **PD** can be seen as its sparsifying dictionary. However, the samples $\mathbf{x}_s^*$ are not a linear combination of the full $\mathbf{x}^*$ signal, and hence the compressed sensed signal is the CSA-ODF, not the attenuations.

Another key requirement in CS is the sparse representation on a given dictionary, which can be exploited by using ℓ_0 or ℓ_1 minimization, sometimes in combination with additional regularizing constraints such as total variation (19,35). We use multi-task Bayesian compressed sensing (MT-BCS), (37), that enforces sparsity via sparsifying priors, provides spatial regularization via probabilistic clustering, and allows the estimation of the full posterior distribution of the signal of interest (while other approaches only give the maximum *a posteriori* or MAP).

Bayesian Compressed Sensing

We use MT-BCS (37,44,51) to obtain the full posterior distribution of the estimated CSA-ODF and diffusion signal using sparsity inducing priors and shared hyper-priors that model similar spatial sparsity patterns between diffusion signals from different spatial locations in the brain. Specifically,

$$p(\boldsymbol{\vartheta}_i | \boldsymbol{\alpha}_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}_i}), \ \boldsymbol{\Lambda}_{\boldsymbol{\alpha}_i} = diag\{\ \alpha_{i1}^{-1} \ \cdots \ \alpha_{iL}^{-1} \ \}, \ i = 1, \ldots, N, \quad [13]$$

is the a priori distribution of ϑ_i ([7], [10]–[12]), modeled as a multidimensional Gaussian with zero mean and diagonal covariance $\Lambda_{\boldsymbol{\alpha}_i}$. This prior promotes sparse representations, when the hyper-priors follow a Gamma distribution (44,51). Hyper-parameters $\{\boldsymbol{\alpha}_i\}_{i=1}^N$ in [13] model information sharing among different diffusion signal. Specifically, diffusion signal from different voxels are clustered using a nonparametric Dirichlet process (DP) mixture model. A Dirichlet process $G(\lambda, G_0)$ is a distribution of distributions with scaling variable λ and base distribution G_0. The DP prior models uncertainty in the distributions, hence, it is rich enough to approximate any possible distribution. This makes it more powerful than parametric probabilistic models (52). In particular, the DP is used here as a tool for non-parametric clustering, as explained next.

Hyper-parameters $\boldsymbol{\alpha}_i$ are modeled as independent identically distributed random variables drawn from a distribution generated by G. Let $\{\boldsymbol{\alpha}_k^*\}_{k=1}^K$, $K \leq N$, be the distinct values taken by $\{\boldsymbol{\alpha}_i\}_{i=1}^N$, then (44)

$$p\left(\boldsymbol{\alpha}_i | \boldsymbol{\alpha}^{-i}\right) = \frac{\lambda}{\lambda+N-1}G_0 + \frac{1}{\lambda+N-1}\sum_{k=1,k\neq i}^K n_k^{-i}\delta_{\boldsymbol{\alpha}_k^*}, \; i=1,\ldots,N, \quad [14]$$

where $\boldsymbol{\alpha}^{-i} = \{\boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_{i-1} \cdots \boldsymbol{\alpha}_{i+1} \cdots \boldsymbol{\alpha}_N\}$, n_k^{-i} is the number of times $\boldsymbol{\alpha}_k^*$ is in $\boldsymbol{\alpha}^{-i}$, and $\delta_{\boldsymbol{\alpha}_k^*}$ the distribution concentrated at $\boldsymbol{\alpha}_k^*$. Eq. [14] explains the sharing and non-parametric properties of DPs: new hyper-parameters $\boldsymbol{\alpha}_{i+1}$ can either use the value of one of the existing hyper-parameters $\boldsymbol{\alpha}^{-i}$ or a new one (generated from G_0). The random variable λ controls the balance between the creation of new hyper-parameters and sharing with the existing hyper-parameters. Let $\mathbf{z}_i$ be a variable that indicates how much of each hyper-parameter $\boldsymbol{\alpha}_j^*$ is used in the mixture model $\boldsymbol{\alpha}_i = \sum_{k=1}^K z_{ik}\boldsymbol{\alpha}_k^*$. Hence $\mathbf{z}_i$ provides a probabilistic clustering of the data, based on how much of each hyper-parameter $\{\boldsymbol{\alpha}_k^*\}_{k=1}^J$ is used at each voxel. Since G can be modeled as $G = \sum_{k=1}^\infty w_k\delta_{\boldsymbol{\alpha}_k^*}$, using a stick-breaking process (53), the complete DP-MT-CS probabilistic model is given by

$$
\begin{aligned}
\mathbf{x}_{is}^* | \boldsymbol{\vartheta}_i, \alpha_0 &\sim \mathcal{N}(\mathbf{\Phi}\boldsymbol{\vartheta}_i, \alpha_0\mathbf{I}), \mathbf{\Phi} \overset{\text{def}}{=} \mathbf{P}_s\mathbf{D}_s \\
\boldsymbol{\vartheta}_i | \mathbf{z}_i \{\boldsymbol{\alpha}_k^*\}_{k=1}^K &\sim \mathcal{N}\left(\mathbf{0}, \Lambda_{\boldsymbol{\alpha}_{\mathbf{z}_i}^*}\right), \\
\mathbf{z}_i | \{w_k\}_{k=1}^k &\sim Multinomial(w_1, \cdots, w_K), \\
w_k &= \pi_k \prod_{l=1}^{k-1}(1-\pi_l), \\
\pi_k | \lambda &\sim Beta(1, \lambda), \\
\lambda | e, f &\sim Gamma(e, f), \\
\boldsymbol{\alpha}_k^* | c, d &\sim \prod_{k=1}^K Gamma(c, d), \\
\alpha_0 &\sim Gamma(a, b),
\end{aligned}
\quad [15]
$$

where $1 \leq i \leq N, 1 \leq k \leq K$, and a, b, c, d, e, f are non-informative hyper-parameters determining the distribution of the random variables in the model.

Bayesian inference using the DP-MT-CS model is provided by Bayes rule

$$p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon}) = \frac{p(\mathbf{x}_s^*|\mathbf{H})p(\mathbf{H}|\boldsymbol{\Upsilon})}{\int p(\mathbf{x}_s^*|\mathbf{H})p(\mathbf{H}|\boldsymbol{\Upsilon})d\mathbf{H}}, \quad [16]$$

where $\mathbf{H} = \left\{ \{\boldsymbol{\vartheta}_i\}_{i=1}^N, \left\{\boldsymbol{\alpha}_j^*\right\}_{k=1}^K, \{\mathbf{z}_i\}_{i=1}^N, \{\pi_j\}_{j=1}^J, \alpha_0, \lambda \right\}$ are the hidden model variables, $\mathbf{x}_s^* = [\mathbf{x}_{si}^*]_{i=1}^N$ the $M \times N$ matrix of CS HARDI measurements, and $\mathbf{Y} = \{a, b, c, d, e, f\}$ the hyper-parameters determining the distribution of the hidden variables. Eq. [16] cannot be solved analytically, due to the mathematical intractability of the integration term in the denominator. An iterative variational method can be used to seek a distribution $q(\mathbf{H})$ that approximates $p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon})$. Let us consider the logarithm of the marginal likelihood (44),

$$log(p(\mathbf{x}_s^*|\boldsymbol{\Upsilon})) = \mathscr{L}(q(\mathbf{H})) + \mathscr{D}_{KL}(q(\mathbf{H}), p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon})), \mathscr{L}(q(\mathbf{H})) = \int q(\mathbf{H})log\left(\frac{p(\mathbf{x}_s^*|\boldsymbol{\Upsilon})p(\mathbf{H}|\boldsymbol{\Upsilon})}{q(\mathbf{H})}\right), \quad [17]$$

where $\mathscr{D}_{KL}(q(\mathbf{H}), p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon}))$ is the Kullback-Leibler (KL) divergence between the approximate $q(\mathbf{H})$ and the true posterior $p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon})$. The approximation of the true posterior $p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon})$ by $q(\mathbf{H})$ can be achieved by minimizing $\mathscr{D}_{KL}(q(\mathbf{H}), p(\mathbf{H}|\mathbf{x}_s^*, \boldsymbol{\Upsilon}))$. Since the KL divergence is non-negative and $p(\mathbf{x}_s^*|\boldsymbol{\Upsilon})$ is fixed given $\mathbf{Y}$, then minimizing the KL divergence is the same as maximizing $\mathcal{L}(q(\mathbf{H}))$ in [17] (44).

Maximizing $\mathcal{L}(q(\mathbf{H}))$ is tractable using functional derivatives if we define

$$q(\mathbf{H}) = q(\alpha_0)q(\lambda)\Pi_{k=1}^K q(\pi_j)\Pi_{i=1}^N q(\mathbf{z}_i)\Pi_{i=1}^N q(\boldsymbol{\vartheta}_i)\Pi_{k=1}^K q(\boldsymbol{\alpha}_j^*). \quad [18]$$

In particular we assume $q(\boldsymbol{\vartheta}_i) \sim \mathscr{N}(\boldsymbol{\mu}_i, \Sigma_i)$, where $\boldsymbol{\mu}_i$ and Σ_i are estimated by iteratively maximizing $\mathcal{L}(q(\mathbf{H}))$. Hence, by Eq. [8] and Eq. [11]

$$\begin{cases} \hat{\mathbf{x}}_i^* \sim \mathscr{N}(\mathbf{PD}\boldsymbol{\mu}_i, \mathbf{PD}\sum_i(\mathbf{PD})^{\mathrm{T}}) \\ \hat{\phi}_i \sim \mathscr{N}(\mathbf{D}\boldsymbol{\mu}_i, \mathbf{D}\sum_i\mathbf{D}^{\mathrm{T}}) \end{cases}, i = 1, \dots, N, \quad [19]$$

Eq. [19] provides the full posterior of $\mathbf{x}_i^*$ and CSA-ODF at each voxel. From here, it is simple to obtain (see Eq. [12])

$$\tilde{\mathbf{x}}_i \sim \mathscr{N}(\tilde{\varepsilon}_i, \mathbf{PD}\sum_i(\mathbf{PD})^{\mathrm{T}}), \tilde{\varepsilon}_i \approx \mathbf{PD}\boldsymbol{\mu}_i + (c_1 - \gamma/(2\sqrt{\pi}))\mathbf{Y}_1. \quad [20]$$

It is easy to see that $\boldsymbol{\varepsilon}_i = [exp(-exp(\tilde{\varepsilon_{ij}}))]_{1 \ j \ M}$ (see Eq. [5]) provides a point estimate of the diffusion-weighted signal, $\mathbf{x}_i$. We did not use Eq. [7] here to obtain $\mathbf{x}_i$, because there were several approximations made in Eqs. [10]–[12] and the mean value of $\tilde{\mathbf{x}}$ was removed, so it is more accurate to estimate $\mathbf{x}_i$ as indicated above. However, it is difficult to obtain an estimate of the uncertainty of $\mathbf{x}_i$, given the highly non-linear relationship between $\tilde{\mathbf{x}_i}$ and $\mathbf{x}_i$.

Here, we can use Eq. [7] to provide a reasonable estimate of the uncertainty (covariance matrix) in $\mathbf{x}_i$ as

$$\mathbf{x}_i \sim \mathcal{N}(\varepsilon_i, \mathbf{\Psi} \sum_i \mathbf{\Psi}^{\mathrm{T}}). \quad [21]$$

An advantage of using a known dictionary $\mathbf{\Psi}$ for the attenuation signal $\mathbf{x}_i$ is that it is related to the CSA-ODF, Eqs. [7] and [9]. Compared to using directly a dictionary for the CSA-ODF, it allows us to obtain uncertainty estimates for both the CSA-ODF and attenuation signal.

HARDI Data Sets

We tested the proposed approach on two kinds of HARDI data sets:

- Synthetic multi-shell HARDI data with non-staggered (same diffusion directions on all shells) and staggered (complementary diffusion directions on all shells) protocols.

- *In-vivo* non-staggered multi-shell HARDI datasets.

The synthetic HARDI data was generated using the analytic Camino diffusion MRI toolkit[6] that employs realistic diffusion models (54). Rician noise was added in Camino to produce SNRs of 25, and 15, where $SNR = E/\sigma$, E being the magnitude of the noise-free attenuation signal and σ the standard deviation of the noise (55). The diffusion model consists of two fibers crossing at angles 45°, 60°, and 90°, where one fiber was fixed at (45°, 45°) in spherical coordinates, while the other was rotated 45°, 60°, and 90° with respect to the first. Fiber crossings with angles around 30° could not be detected using the CSA-ODF. The reason for this is that the ODF requires the product of the mean apparent diffusion coefficient (ADC) and b-values to be close to one (40). This condition is satisfied in practice, for the range of b-values commonly used and the mean ADC in the brain parenchymia (0.7 μm^2/ms). However, Camino produces lower ADC values (~0.07 μm^2/ms) for all models and parameters tested (see also Figure 1b in (56)). Fiber crossings reported in the literature using the CSA-ODF and angles around 30 degrees (40,41) were obtained using ADC values close to those found in the brain, a much simpler model consisting of two anisotropic diffusion tensors, no noise, and a large SH order (8). We limit ourselves here to those angles that could be discriminated using Camino's state-of-the-art analytic models, Rician noise, and a maximum SH order of six.

The volume fractions for each fiber were 0.3 for the intra-axonal, 0.05 for the extra-axonal components, and volume fraction of 0.3 for the isotropic compartment. Following (54), the intra-axonal compartment is modeled as a cylinder of radius 4 μm with a Gaussian phase distribution (GPD) and diffusivity 1.7 μm^2/ms. The extra-axonal compartment is modeled as a zeppelin with diffusivity 1.7 μm^2/ms along, and 0.2 μm^2/ms across the fiber orientation. Finally, the isotropic component is modeled as a sphere of radius 4 μm and diffusivity 0.8 μm^2/ms. Simulated standard Stejskal-Tanner pulse-gradient spin-echo pulse sequence was

[6]http://cmic.cs.ucl.ac.uk/camino/

50 *mT/m*, TE = 100m*s*, δ = 20m*s*. Multi-shell non-staggered synthetic HARDI signal was generated with three shells and a total of *M*=399 diffusion directions (133 gradients per shell), and three b-values, *b*=2000, 4000, and 6000 *s/mm*². Multi-shell staggered synthetic HARDI signal was generated with three shells and a total of *M*=400 diffusion directions (b=2000, 4000, and 6000 *s/mm*²), where 134 were allocated in the first shell, 133 in the second and 133 in the last one. The sampling protocol follows (4), which ensures that CS "partial" samples (taking the first *m* gradients) cover the whole shell almost uniformly. CS single and multi-shell HARDI datasets were generated with accelerations one (no CS), two, four, and eight, corresponding to *m* = *M*, *M*/2, *M*/4, and *M*/8.

In-vivo multi-shell HARDI data sets were obtained on the WU-UMN Human Connectome Project (57) 3T Connectome Skyra system (Siemens, Erlangen, Germany), with various spatial and angular resolutions. The scanner was equipped with SC72 gradients operating at 100 *mT/m* maximum gradient amplitude with a maximum slew rate of 91 *T/m/s* for diffusion encoding. Dataset 1, with spatial resolution 1.5×1.5×1.5 mm³, has three shells with 133 diffusion directions per shell at *b*=1000, 2000, and 3000 *s/mm*² and 10 additional b0s per shell. Dataset 2, with spatial resolution 1.25×1.25×1.25 mm³, has three shells with 128 diffusion directions per shell at *b*=1500, 2500, and 3500 *s/mm*² and 26 b0s. The third data set, with spatial resolution 1.25×1.25×1.25 mm³, has six shells with 128 diffusion directions per shell at *b*=1500, 2500, 3500, 5000, 7000, and 10,000 *s/mm*² and 28 b0s. Each dataset was corrected for geometric and eddy current distortions, using information from acquisitions in opposite phase-encoding directions, as well as head motion (58). The gradient table for these images follow the protocol proposed in (4). Simulated CS HARDI datasets were constructed with accelerations one, two, four, and eight. An acceleration factor of one means that all the data is used (no CS). Acceleration factors of two, four and eight means that only half, one quarter, and one eight of the diffusion-weighted volumes are used as input to the proposed algorithm.

The research protocol used in this investigation was approved by the Institutional Review Board of the University of Minnesota. All subjects provided informed written consent prior to participating in the research.

Finally, parameters *a* and *b* are defined as $a = b = Variance\{\tilde{E}\}$, while *c*, *d*, *e*, *f* in [15] are set initially to very small values and are estimated from the data. Experimentally, we found that $\mathbf{z}_i = \begin{bmatrix} \frac{1}{K} & \cdots & \frac{1}{K} \end{bmatrix}^{\mathrm{T}}$ provides a good initialization, with $K = O(N)$. Also, the best γ in [6], [10]–[12], was found by trial and error as $\gamma = 10^{-2}$.

Ground Truth Definition and Evaluation

The ground truth for the fiber orientations in the synthetic data sets is given by the knowledge of the true fiber orientations. We also generated synthetic images without noise and a full set of diffusion directions (acceleration one) that provide ground truth for the corresponding diffusion-weighted volumes. The ground truth CSA-ODFs are computed from this "clean" data using each method and all the diffusion directions (acceleration one). Finally, the "ground truth" for the fiber orientations, CSA-ODFs, and diffusion-weighted volumes of the *in-vivo* HARDI datasets is obtained from each method studied in this work

with the full dataset (acceleration one). Notice that there is obviously not a single ground-truth for the *in-vivo* HARDI datasets. Hence using the estimated CSA-ODFs using only one of these methods as ground truth is favorably biased towards that method. The best we can do to mitigate this caveat is to compare each method at accelerations two and higher with the same method using the full dataset (acceleration one). This allows us to evaluate their relative performance as the acceleration factor increases, with respect to the best-case scenario, i.e., without CS and a relatively high number of diffusion directions.

We evaluate results using the metrics (see Eq. [19])

$$FOE(\phi, \hat{\phi}) = \frac{1}{N} \sum_{i=1}^{N} \left| \cos^{-1} \left(\frac{180}{\pi} \hat{\mathbf{u}}(\phi_i)^{\mathrm{T}} \hat{\mathbf{v}}(\hat{\phi}_i) \right) \right|$$

$$KL(\phi, \hat{\phi}) = \frac{1}{2N} \sum_{i=1}^{N} \left(\sum_{j=1}^{J} ln \left(\frac{\phi_{ij}}{\hat{\phi}_{ij}} \right) \phi_{ij} \delta\Omega_j + \sum_{j=1}^{J} ln \left(\frac{\hat{\phi}_{ij}}{\phi_{ij}} \right) \hat{\phi}_{ij} \delta\Omega_j \right), \; \delta\Omega_j = \Delta\theta \Delta\varphi sin\theta_j \quad [22]$$

$$D_{SH}(\phi, \hat{\phi}) = \frac{1}{N} \sum_{i=1}^{N} \left\| (\phi_i)_{\mathrm{SH}} - (\phi_i)_{\mathrm{SH}} \right\|_2^2,$$

where FOE is the Fiber Orientation Error (degrees) computed from the known (synthetic images) or estimated (*in-vivo* data sets) fiber orientations $\hat{\mathbf{u}}(\varphi_i)$ and the fiber orientations $\hat{\mathbf{v}}(\hat{\varphi_i})$ from the known (φ_i) and estimated CSA-ODF ($\hat{\varphi}$). Fiber orientations are approximated by the CSA-ODF peaks using functionality provided in CSA-ODF (FSL) (42)). $KL(\varphi, \hat{\varphi})$ is the Kullback-Leibler divergence in spherical coordinates (59), between φ and $\hat{\varphi}$, from the CS HARDI data sets. $\delta\Omega$ stands for the discrete approximation of the differential of the solid angle Ω, and $J \gg M$ stands for the number of tessellation points in the unit sphere. $D_{SH}(\varphi, \hat{\varphi})$ stands for the mean Euclidean distance between the SH coefficients of φ_i and $\hat{\varphi}$. These metrics are used routinely to compare ODFs (3,59–61). The order of the SH basis was six for the synthetic data and four for the in-vivo data (as recommended in (40,62,63)).

RESULTS

Synthetic Data

Multi-shell non-staggered—Figure 1[7] compares the FOE of the estimated CSA-ODF from under-sampled HARDI synthetic multi-shell non-staggered HARDI data using bcsCSA-ODF, gCSA-ODF, and CSA-ODF at SNRs 15, 25 and fiber crossing angles: 45°, 60°, and 90°. CSA-ODF reconstructions cannot be obtained at acceleration eight, since at that acceleration the number of diffusion directions per shell is lower than the number of SH coefficients (28 in this case). bcsCSA-ODF significantly outperforms CSA-ODF and gCSA-ODF in terms of the FOE, for all acceleration factors up to eight, except for the case of SNR 15 and fiber crossing angle of 45°, where the CSA-ODF is better up to an acceleration factor of six. The CSA-ODF provides a reduction in the FOE of up to 10 degrees. The worst performance of the CSA-ODF corresponds to fibers crossing at 45°, SNR 15 and acceleration eight. This is due to the need of a better initialization of the variational Bayesian inference algorithm as the acceleration, SNR and order of the SH basis increase.

---

[7]Error bars on all graphs indicate the standard deviation of the mean errors.

Figure 2 indicates that the SH distance and KL divergence are at least 70% lower for the proposed bcsCSA-ODF than CSA-ODF (even at acceleration one). The SH distance is 18–88% lower for the bcsCSA-ODF than the gCSA-ODF, while the KL divergence is up to 90% lower for the bcsCSA-ODF than the gCSA-ODF. As the acceleration increases, the difference between the bcsCSA-ODF and gCSA-ODF reduces. As the fiber crossing angle increases, the difference between the bcsCSA-ODF and gCSA-ODF reduces, since fiber crossings close to 90° have CSA-ODFs with more separated modes than smaller angles.

Multi-shell staggered—Figure 3 compares the FOE of the estimated CSA-ODF from under-sampled HARDI synthetic multi-shell staggered HARDI data using bcsCSA-ODF, gCSA-ODF, and CSA-ODF at SNRs 15, 25 and fiber crossing angles 45°, 60°, and 90°. bcsCSA-ODF significantly outperforms CSA-ODF and gCSA-ODF in terms of the FOE, for all acceleration factors up to eight, except for the case of a fiber crossing angle of 45°, where the CSA-ODF is better up to an acceleration factor of five. The CSA-ODF provides a reduction in the FOE of up to 16 degrees. The worst performance of the CSA-ODF corresponds to fibers crossing at 45° and accelerations above five. Comparing with Figure 2, it can be seen that the differences with the gCSA-ODF are larger for staggered data than with non-staggered data, since the gCSA-ODF performs worse with staggered diffusion directions. The reason for this might be that the gCSA-ODF requires a preliminary SHs interpolation to obtain the same diffusion directions on all shells so that the bi-exponential model can be fitted for all diffusion directions. This interpolation is not necessary with non-staggered diffusion directions.

Figure 4 indicates that the SH distance and KL divergence are at least 73% lower for the proposed bcs-CSA-ODF than CSA-ODF (even at acceleration one). The SH distance is up to 92% lower for the bcs-CSA-ODF than the gCSA-ODF, while the KL divergence 45–93% lower for the bcsCSA-ODF than the gCSA-ODF. As the acceleration and fiber crossing angle increases, the difference between the bcsCSA-ODF and gCSA-ODF reduces too as with the non-staggered case.

In-vivo Data

Figure 5 compares the FOE of the estimated CSA-ODF from under-sampled HARDI multi-shell data using the proposed bcsCSA-ODF, gCSA-ODF, and CSA-ODF. Notice that CSA-ODF was used only on one dataset, since the other datasets were not compatible with the CSA-ODF requirements.[8] The results indicate that bcsCSA-ODF does perform better than gCSA-ODF (and CSA-ODF) up to acceleration four. The largest benefit of bcsCSA-ODF occurs at acceleration two.

Figure 6 shows the KL divergence and SH distance for the three models considered here. Here, the bcsCSA-ODF has a 55–75% lower error than CSA-ODF and (17–50%) lower than gCSA-ODF, especially showing significant improvements at acceleration two. Notice that the KL divergence and SH distance metrics show bigger differences relative to acceleration one (Figure 6) than the differences observed on the fiber orientations (Figure 5). The reason

---

[8]CSA-ODF requires b-values that follow an arithmetic progression, starting with zero.

for this is that KL divergence and SH distance are metrics that apply directly to the estimated "full" ODFs, which is the focus of the bcsCSA-ODF model, while fiber orientation depends on peak detection.

Figure 7 compares the CSA-ODFs estimated using gCSA-ODF and the proposed bcsCSA-ODF at acceleration one and acceleration four. Differences in the CSA-ODF between acceleration one and four are visually lower for the proposed bcsCSA-ODF.

Figure 8 shows the KL divergence at acceleration four for these three methods, where we can see how the KL varies spatially (left) and its distribution (right). The lowest KL divergence is obtained by using bcs-CSA-ODF, followed by gCSA-ODF and then CSA-ODF.

Figure 9 shows the histogram of the CSA-ODF and diffusion-weighted volumes uncertainties (see Eq. [22]) obtained by normalizing and averaging the covariance matrices and then using these covariances to generate zero-mean random variables, which distribution can be approximated via histograms. As can be seen from the histograms, the uncertainty (dispersion) of the CSA-ODF remains almost the same until acceleration four, growing to about 20 times for acceleration 8. On the other hand, the uncertainty for the estimated diffusion-weighted volumes seems to be fairly insensitive to acceleration.

DISCUSSION

For under-sampled multi-shell HARDI data, the proposed bcsCSA-ODF model performs very well compared to the original CSA-ODF and the generalized CSA-ODF models when the number of diffusion-weighted volumes per shell is at least one quarter of the original image, corresponding to 32–33 volumes per shell, for the in-vivo data, i.e. 96–99 for the three shells. Further reduction in the acquisition time can be achieved using parallel imaging techniques (pMRI) (5–8). Even larger acceleration factors could also be achieved using trained dictionaries rather than pre-specified basis as indicated in (19).

Errors for the proposed bcsCSA-ODF increase with acceleration and noise, with the exception of the third fiber orientation error in Figure 6. There are several reasons for this. ODF peak detection is a hard problem that depends on the selected thresholds and is affected by spurious lobes in ODFs arising from noise. It has also been recognized that there are limits in CS in terms of acceleration and noise level (64,65).

Figures 1–4 show that even at acceleration one (no compression) it is advantageous to use CS since the estimated CSA-ODFs have lower errors than using CSA-ODF or gCSA-ODF. This is due in part to the denoising capabilities of sparse representation and also to the fact that the full dataset can be seen as the CS version of a larger dataset (say acceleration ½ for instance) so it may contain information not seen by the other algorithms.

The number of steps required to maximize $\mathcal{L}(q(\mathbf{H}))$ in [17] depends on the tolerance used to determine convergence and the initialization. The time complexity of the variational Bayesian inference is determined by the cost of estimating the precision matrix Σ_i in [19], [20] on every voxel, which requires the inversion of an $M \times M$ matrix (44). Hence, the time

complexity is $O(\kappa_1 NM^3)$, where N is the number of voxels, M the number of samples, and κ_1 the number of iterations (20 in all our simulations). However, in contrast to (44,51) where the CS matrix is random, here Φ (see Eq. [15]) is fixed and Σ_i needs to be computed only once at each iteration, hence, the precision matrix can be computed in $O(\kappa_1 M^3)$ time. On every voxel there are also matrix by vector operations, hence the time complexity of the proposed bcsCSA-ODF is $O(\kappa_1 M^3) + O(\kappa_1 NL^2)$, where L correspond to the number of ridgelets basis. This reduces to $O(\kappa_1 NL^2)$, since in general $N \gg L > M$. The time complexity of the CSA-ODF is determined by the estimation of the SH coefficients, which requires only matrix by vector computations (see [5]), i.e. $O(NMR)$, where R corresponds to the number of SH coefficients. The time complexity of gCSA-ODF is determined by the cost of fitting the bi-exponential model on each voxel and for each diffusion direction. Using an iterative nonlinear least squares fitting, the computational complexity is $O(\kappa_2 NMB^3)$, with $\kappa_2 \sim 10^4$ and B corresponds to the number of shells in the diffusion image.

The proposed bcsCSA-ODF is computationally more demanding than the CSA-ODF and the gCSA-ODF. The higher computational complexity of the proposed CSA-ODF is in agreement with the computational complexity of typical CS algorithms, CS signal reconstruction is more complex than traditional signal reconstruction. Computationally more efficient CS methods have being proposed recently and could also be employed here within the proposed CSA-ODF framework.

Comparisons with other CS algorithms that focus on the CSA-ODF should also be conducted in the future. The bcsCSA-ODF could be further improved by using dictionaries trained for CSA-ODFs and HARDI data, and also by changing the implicit diffusion signal decay (mono-exponential) model by a more accurate representation of the white matter microstructure in Eq. [2] such as the bi-exponential model or other more complex microstructure models.

CONCLUSION

In summary, this work presents, for the first time, a multi-task Bayesian compressive sensing approach to simultaneously estimate the full posterior of the CSA-ODF and diffusion-weighted volumes from multi-shell HARDI acquisitions. It demonstrates improvement of the quality of acquired datasets via CS de-noising, and accurate estimation of the CSA-ODF, as well as enables a reduction in the acquisition time by a factor of two to four, especially when "staggered" q-space sampling schemes are used.

Acknowledgments

References

1. Basser PJ, Jones DK. Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. NMR biomed. 2002; 15:456–467.10.1002/nbm.783 [PubMed: 12489095]

2. Behrens TEJ, Berg HJ, Jbabdi S, Rushworth MFS, Woolrich MW. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage. 2007; 34:144–55.10.1016/j.neuroimage.2006.09.018 [PubMed: 17070705]

3. Tuch DS. Q-ball imaging. Magn Reson Med. 2004; 52:1358–1372.10.1002/mrm.20279 [PubMed: 15562495]

4. Caruyer, E.; Cheng, J.; Lenglet, C.; Sapiro, G.; Jiang, T.; Deriche, R. MICCAI. Vol. 11. Toronto, Canada: 2011. Optimal design of multiple q-shells experiments for diffusion MRI.

5. Moeller S, Yacoub E, Olman Ca, Auerbach E, Strupp J, Harel N, U urbil K. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magn Reson Med. 2010; 63:1144–1153.10.1002/mrm. 22361 [PubMed: 20432285]

6. Feinberg D, Moeller S, Smith SM, Auerbach E, Ramanna S, Gunther M, Glasser MF, Miller KL, Ugurbil K, Yacoub E. Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. PloS One. 2010; 5:e15710.10.1371/journal.pone.0015710 [PubMed: 21187930]

7. 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. Magn Reson Med. 2011; 67:1210–1224.10.1002/mrm.23097 [PubMed: 21858868]

8. Setsompop K, Cohen-Adad J, Gagoski Ba, Raij T, Yendiki A, Keil B, Wedeen VJ, Wald LL. Improving diffusion MRI using simultaneous multi-slice echo planar imaging. NeuroImage. 2012; 63:569–80.10.1016/j.neuroimage.2012.06.033 [PubMed: 22732564]

9. Block KT, Uecker M, Frahm J. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. Magn Reson Med. 2007; 57:1086–1098.10.1002/mrm.21236 [PubMed: 17534903]

10. Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med. 2007; 58:1182–1195.10.1002/mrm.21391 [PubMed: 17969013]

11. Maleh, R. Efficient sparse approximation methods for medical imaging. University of Michigan; 2009.

12. Chartrand, R. Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data. Int Simp Biomed Imaging; Boston, MA: 2009. p. 262-265.

13. Khare K, Hardy CJ, King KF, Turski Pa, Marinelli L. Accelerated MR imaging using compressive sensing with no free parameters. Magn Reson Med. 2012; 68:1450–1457.10.1002/mrm.24143 [PubMed: 22266597]

14. Seeger M, Nickisch H, Pohmann R, Schölkopf B. Optimization of k-space trajectories for compressed sensing by Bayesian experimental design. Magn Reson Med. 2010; 63:116–126.10.1002/mrm.22180 [PubMed: 19859957]

15. Weller, DS.; Polimeni, JR.; Grady, L.; Wald, LL.; Adalsteinsson, E.; Goyal, VK. IEEE Int Symp Biomedical Imaging. Chicago: IEEE Int. Symp. Biomed. Imag; 2011. Evaluating sparsity penalty functions for combined compressed sensing and parallel MRI; p. 1589-1592.

16. Weller, DS.; Polimeni, JR.; Grady, L.; Wald, LL.; Adalsteinsson, E.; Goyal, VK. Accelerated parallel magnetic resonance imaging reconstruction using joint estimation with a sparse signal Model. IEEE SSP; Ann Arbor, MI, US: 2012. p. 221-224.

17. Menzel MI, Tan ET, Khare K, Sperl JI, King KF, Tao X, Hardy CJ, Marinelli L. Accelerated diffusion spectrum imaging in the human brain using compressed sensing. Magn Reson Med. 2011; 66:1226–1233.10.1002/mrm.23064 [PubMed: 22012686]

18. Bilgic, B.; Setsompop, K.; Cohen-Adad, J.; Yendiki, A.; Wald, LL.; Adalsteinsson, E. MICCAI. Nice, France: 2012. Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries; p. 1-9.

19. Bilgic B, Setsompop K, Cohen-Adad J, Yendiki A, Wald LL, Adalsteinsson E. Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries. Magn Reson Med. 2012 epub:1–8. 10.1002/mrm.24505

20. Gramfort, A.; Poupon, C.; Descoteaux, M. CDMRI. Nice, France: MICCAI; 2012. Sparse DSI : learning DSI structure for denoising and fast imaging; p. 288-296.

21. Merlet, S.; Caruyer, E.; Deriche, R. MICCAI. Nice, France: 2012. Parametric Dictionary Learning for Modeling EAP and ODF in Diffusion MRI; p. 10-17.

22. Merlet, S.; Deriche, R. MICCAI. Beijin, China: 2010. Compressed sensing for accelerated EAP recovery in diffusion MRI.

23. Merlet, S.; Cheng, J.; Ghosh, A.; Deriche, R. Spherical polar Fourier EAP and ODF Reconstruction via compressed sensing in diffusion MRI. IEEE Int. Symp. Biomedical Imaging; Chicago. 2011. p. 365-371.

24. Rathi Y, Michailovich O, Setsompop K, Bouix S, Shenton ME, Westin CF. Sparse multi-shell diffusion imaging. MICCAI. 2011; 14:58–65. [PubMed: 21995013]

25. Cheng, J.; Jiang, T.; Deriche, R. MICCAI. Vol. 3. Nice, France: 2012. Nonnegative definite EAP and ODF estimation via a unified multi-shell HARDI reconstruction; p. 313-321.

26. Jung H, Sung K, Nayak KS, Kim EY, Ye JC. k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. Magn Reson Med. 2009; 61:103–116.10.1002/mrm. 21757 [PubMed: 19097216]

27. Jung H, Ye JC. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques. Int J Imag Syst Tech. 2010; 20:81–98.10.1002/ima.20231

28. Hsiao A, Lustig M, Alley MT, Murphy M, Chan FP, Herfkens RJ, Vasanawala SS. Rapid pediatric cardiac assessment of flow and ventricular volume with compressed sensing parallel imaging volumetric cine phase-contrast MRI. Am J Roentgenol. 2012; 198:W250–W259.10.2214/AJR. 11.6969 [PubMed: 22358022]

29. Knobloch V, Boesiger P, Kozerke S. Sparsity transform k-t principal component analysis for accelerating cine three-dimensional flow measurements. Magn Reson Med. 2012:1–11.10.1002/ mrm.24431

30. Niebergall A, Zhang S, Kunay E, Keydana G, Job M, Uecker M, Frahm J. Real-time MRI of speaking at a resolution of 33 ms: Undersampled radial FLASH with nonlinear inverse reconstruction. Magn Reson Med. 2012 epub. 10.1002/mrm.24276

31. Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. Magn Reson Med. 2010; 64:767–776.10.1002/mrm.22463 [PubMed: 20535813]

32. Trzasko JD, Haider CR, Borisch Ea, Campeau NG, Glockner JF, Riederer SJ, Manduca A. Sparse-CAPR: highly accelerated 4D CE-MRA with parallel imaging and nonconvex compressive sensing. Magn Reson Med. 2011; 66:1019–1032.10.1002/mrm.22892 [PubMed: 21608028]

33. Prieto C, Usman M, Wild JM, Kozerke S, Batchelor PG, Schaeffter T. Group sparse reconstruction using intensity-based clustering. Magn Reson Med. 2012 epub:1–11. 10.1002/mrm.24333

34. Li W, Griswold M, Yu X. Fast cardiac T(1) mapping in mice using a model-based compressed sensing method. Magn Reson Med. 2012; 68:1127–1134.10.1002/mrm.23323 [PubMed: 22161952]

35. Michailovich O, Rathi Y, Dolui S. Spatially regularized compressed sensing for high angular resolution diffusion imaging. IEEE Trans Image Process. 2011; 30:1100–1115.10.1109/TMI. 2011.2142189

36. Dolui, S.; Kuurstra, A.; Michailovich, OV. Proc SPIE. Vol. 8314. San Diego, CA: SPIE; 2012. Rician compressed sensing for fast and stable signal reconstruction in diffusion MRI.

37. Duarte-Carvajalino, JM.; Lenglet, C.; Ugurbil, K.; Carin, L.; Sapiro, G. MICCAI. Nice, France: 2012. A framework for multi-task Bayesian compressive sensing of DW-MRI.

38. Tristán-Vega A, Westin C-F. Probabilistic ODF estimation from reduced HARDI data with sparse regularization. MICCAI. 2011; 14:182–90. [PubMed: 21995028]

39. Ye W, Vemuri BC, Entezari A. An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators. IEEE Int Symp Biomedical Imaging IEEE. 2012:940–943.

40. Aganj I, Lenglet C, Sapiro G, Yacoub E, Ugurbil K, Harel N. Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. Magn Reson Med. 2010; 64:554–566.10.1002/mrm.22365.Reconstruction [PubMed: 20535807]

41. Kamath, A.; Aganj, I.; Xu, J.; Yacoub, E.; Ugurbil, K. MICCAI. Nice, France: 2012. Generalized Constant Solid Angle ODF and Optimal Acquisition Protocol for Fiber Orientation Mapping; p. 67-78.

42. Aganj, I.; Lenglet, C.; Sapiro, G. MICCAI. Vol. 13. Beijin, China: Springer; 2010. ODF maxima extraction in spherical harmonic representation via analytical search space reduction; p. 84-91.

43. Tristán-Vega A, Westin C-F, Aja-Fernández S. A new methodology for the estimation of fiber populations in the white matter of the brain with the Funk-Radon transform. NeuroImage [Internet]. 2010; 49:1301–15.10.1016/j.neuroimage.2009.09.070

44. Qi, Y.; Liu, D.; Dunson, D.; Carin, L. IEEE ICML. Helsinki, Finland: 2008. Bayesian multi-Task compressive sensing with Dirichlet process priors.

45. Bilgic B, Goyal VK, Adalsteinsson E. Multi-contrast reconstruction with Bayesian compressed sensing. Magn Reson Med. 2011; 66:1601–1615.10.1002/mrm.22956 [PubMed: 21671267]

46. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. Fsl. NeuroImage. 2012; 62:782–90.10.1016/j.neuroimage.2011.09.015 [PubMed: 21979382]

47. Anderson AW. Measurement of Fiber Orientation Distributions Using High Angular Resolution Diffusion Imaging. 2005; 54:1194–1206.10.1002/mrm.20667

48. Hess CP, Mukherjee P, Han ET, Xu D, Vigneron DB. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. Magn Reson Med [Internet]. 2006; 56:104–17.10.1002/mrm.20931

49. Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Regularized, fast, and robust analytical Q-ball imaging. Magn Reson Med. 2007; 58:497–510.10.1002/mrm.21277 [PubMed: 17763358]

50. Michailovich O, Rathi Y. On approximation of orientation distributions by means of spherical ridgelets. IEEE T Image Process. 2010; 19:461–477.10.1109/TIP.2009.2035886

51. Ji S, Dunson D, Carin L. Multitask compressive sensing. IEEE T Signal Proces. 2009; 57:92–106.

52. Ferguson TS. A Bayesian analysis of some nonparametric problems. The Annals of Statistics. 1973; 1:209–230.

53. Sethuraman J. A constructive definition of the Dirichlet priors. Statistica Sinica. 1994; 4:639–650.

54. Panagiotaki E, Schneider T, Siow B, Hall MG, Lythgoe MF, Alexander DC. Compartment models of the diffusion MR signal in brain white matter: a taxonomy and comparison. NeuroImage. 2012; 59:2241–2254.10.1016/j.neuroimage.2011.09.081 [PubMed: 22001791]

55. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med. 2008; 34:910–914. [PubMed: 8598820]

56. Ferizi, U.; Panagiotaki, E.; Schneider, T.; Wheeler-kingshott, C.; Alexander, D. White Matter Models of In Vivo Diffusion MRI Human Brain Data : A Statistical Ranking. 16th Conference on Medical Image Understanding and Analysis; Washington DC. 2012. p. 1-6.

57. Van Essen DC, Ugurbil K, Auerbach E, et al. The Human Connectome Project: A data acquisition perspective. NeuroImage. 2012; 62:2222–2231.10.1016/j.neuroimage.2012.02.018 [PubMed: 22366334]

58. Jesper, LRA.; Junquian, X.; Essa, Y.; Edward, A.; Steen, M.; Ugurbil, K. ISMRM. Melbourne, Australia: 2012. A comprehensive Gaussian Process framework for correcting distortions and movements in diffusion images.

59. Chiang, M.; Klunder, AD.; Mcmahon, K.; Zubicaray, GI.; De Wright, J.; Toga, AW.; Thompson, PM. Inf Process Med Imaging. Kerkrade, Netherlands: 2007. Information-theoretic analysis of brain white matter fiber orientation distribution functions; p. 172-182.

60. Zhan, L.; Leow, AD.; Barysheva, M., et al. MICCAI. London: 2009. Investigating the uncertainty in multi-fiber estimation in High Angular Resolution Diffusion Imaging.

61. Geng X, Ross TJ, Gu H, Shin W, Zhan W, Chao Y-P, Lin C-P, Schuff N, Yang Y. Diffeomorphic image registration of diffusion MRI using spherical harmonics. IEEE transactions on medical imaging. 2011; 30:747–58.10.1109/TMI.2010.2095027 [PubMed: 21134814]

62. Côté M-A, Boré A, Girard G, Houde J-C, Descoteaux M. Tractometer: online evaluation system for tractography. MICCAI [Internet]. 2012; 15:699–706.

63. Zhan L, Franc D, Patel V, Jahanshad N, Jin Y, Mueller BA. How do spatial and angular resolution affect brain connectivity maps from diffusion MRI? IEEE Int Symp Biomedical Imaging. 2013:1–14.10.1109/ISBI.2012.6235469

64. Davenport, Ma; Laska, JN.; Treichler, JR.; Baraniuk, RG. The Pros and Cons of Compressive Sensing for Wideband Signal Acquisition: Noise Folding versus Dynamic Range. IEEE Transactions on Signal Processing. 2012; 60:4628–4642.10.1109/TSP.2012.2201149

65. Donoho DL, Maleki A, Montanari A. The Noise-Sensitivity Phase Transition in Compressed Sensing. IEEE Transactions on Information Theory. 2011; 57:6920–6941.10.1109/TIT. 2011.2165823

Appendix

Discrete Approximation in Eq. [9]

Using finite differences to approximate $\nabla_b^2(\mathscr{H}(\hat{\mathbf{u}})\vartheta)$ in Eq. [9]

$$\nabla_b^2(\mathscr{H}(\hat{\mathbf{u}}(\varphi,\theta))\vartheta) \approx \mathbf{D}\vartheta$$

where

$$\mathbf{D} = \frac{1}{\tan\theta}\left(\frac{\mathscr{H}(\varphi,\theta+\Delta)-\mathscr{H}(\varphi,\theta-\Delta)}{2\Delta}\right)$$
$$+\frac{(\mathscr{H}(\varphi,\theta+\Delta)-2\mathscr{H}(\varphi,\theta)+\mathscr{H}(\varphi,\theta-\Delta)}{\Delta^2}$$
$$+\frac{1}{\sin^2\theta}\frac{(\mathscr{H}(\varphi+\Delta,\theta)-2\mathscr{H}(\varphi,\theta)+\mathscr{H}(\varphi-\Delta,\theta))}{\Delta^2},$$

is the dictionary for the ODF. For sufficiently small and θ $\pi, 0$ ϕ 2π. $\mathscr{H}$ is explicitly defined in (50).

Derivation of Eq. [11]

Let us rewrite Eq. [10] here for convenience

$$\begin{bmatrix} \gamma^{-1} & 0 \\ 0 & \Lambda \end{bmatrix}\boldsymbol{Y}^+\tilde{\mathbf{x}}+\begin{bmatrix} 1/(2\sqrt{\pi})-\gamma^{-1}c_1 \\ 0 \end{bmatrix} \approx \boldsymbol{Y}^+\mathbf{D}\vartheta,$$

multiplying both sides by $\begin{bmatrix} \gamma & 0 \\ 0 & \Lambda^{-1} \end{bmatrix}$

$$\boldsymbol{Y}^+\tilde{\mathbf{x}}+\begin{bmatrix} \gamma & 0 \\ 0 & \Lambda^{-1} \end{bmatrix}\begin{bmatrix} 1/(2\sqrt{\pi})-\gamma^{-1}c_1 \\ 0 \end{bmatrix} \approx \begin{bmatrix} \gamma & 0 \\ 0 & \Lambda^{-1} \end{bmatrix}\boldsymbol{Y}^+\mathbf{D}\vartheta,$$
$$\boldsymbol{Y}^+\tilde{\mathbf{x}}+\begin{bmatrix} \gamma/(2\sqrt{\pi})-c_1 \\ 0 \end{bmatrix} \approx \begin{bmatrix} \gamma & 0 \\ 0 & \Lambda^{-1} \end{bmatrix}\boldsymbol{Y}^+\mathbf{D}\vartheta.$$

Now $\tilde{\mathbf{x}} \approx \mathbf{Y}\mathbf{Y}^+\tilde{\mathbf{x}}$, hence multiplying by $\mathbf{Y}$

$$\tilde{\mathbf{x}}+\left[\begin{array}{ccc} \boldsymbol{Y}_1 & \cdots & \boldsymbol{Y}_R \end{array}\right]\left[\begin{array}{c} \gamma/(2\sqrt{\pi})-c_1 \\ \mathbf{0} \end{array}\right] \approx \boldsymbol{Y}\left[\begin{array}{cc} \gamma & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{-1} \end{array}\right]\boldsymbol{Y}^{+}\mathbf{D}\vartheta,$$

$$\tilde{\mathbf{x}}+(\gamma/(2\sqrt{\pi})-c_1)\boldsymbol{Y}_1 \approx \boldsymbol{Y}\left[\begin{array}{cc} \gamma & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{-1} \end{array}\right]\boldsymbol{Y}^{+}\mathbf{D}\vartheta.$$

and we arrive at Eq. [11] by defining $\mathbf{x}^{*}=(\tilde{\mathbf{x}}-c_1\boldsymbol{Y}_1)+\gamma/(2\sqrt{\pi})\boldsymbol{Y}_1$ and

$$\mathbf{P}=\boldsymbol{Y}\left[\begin{array}{cc} \gamma & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}^{-1} \end{array}\right]\boldsymbol{Y}^{+}.$$
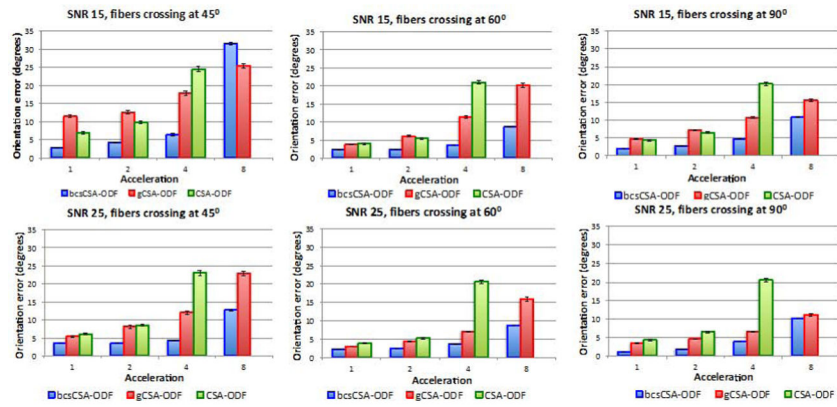
Figure 1.
Fiber orientation errors (FOEs) estimated from under-sampled synthetic multi-shell non-staggered HARDI data. First row corresponds to the FOEs of the estimated ODFs for a signal to noise ratio (SNR) of 15, while the second row shows the FOEs for a SNR of 25. All the figures indicate the FOEs using the proposed bcsCSA-ODF, gCSA-ODF, and CSA-ODF for three fiber crossing angles: 45°, 60°, and 90°. Acceleration one corresponds here to the full 399 gradients.
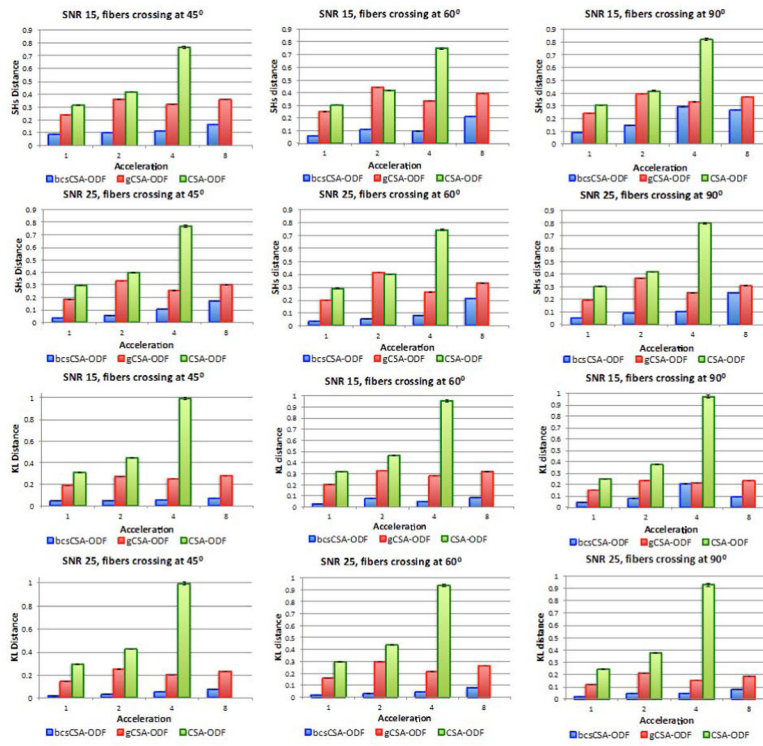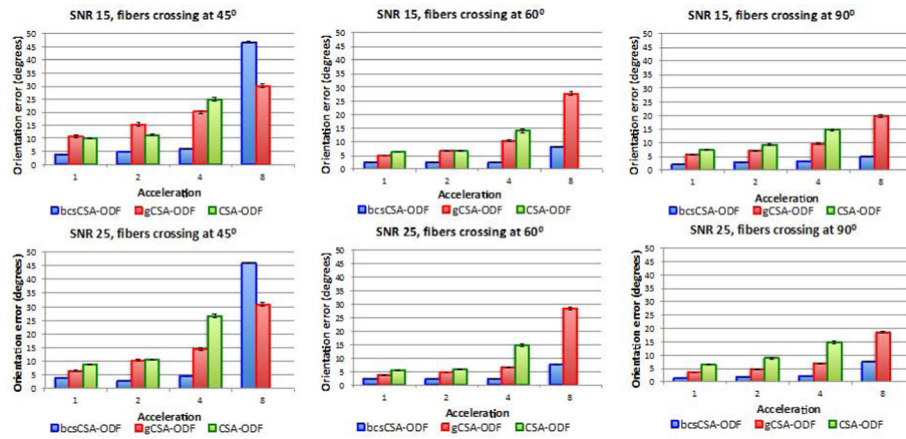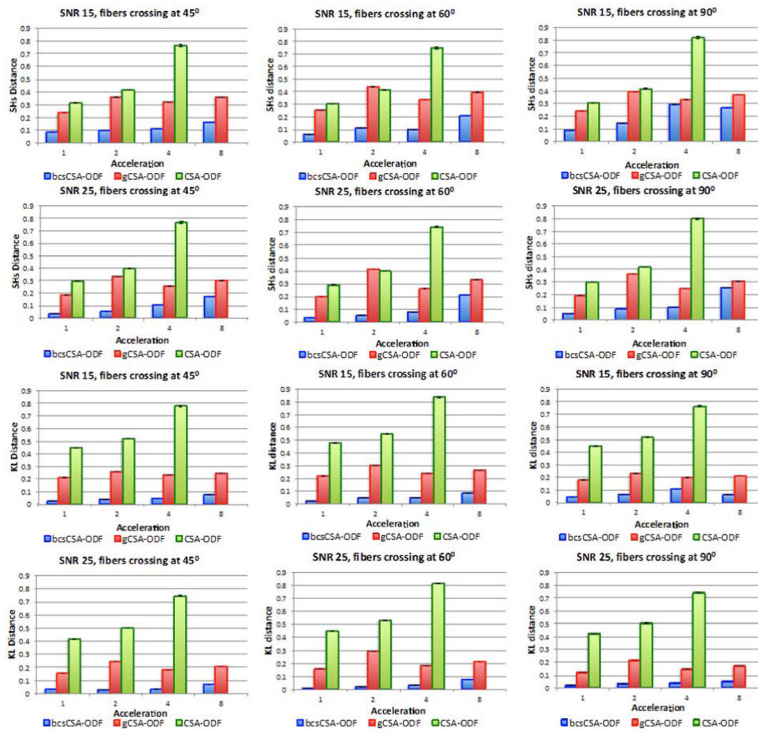
Figure 2.
Euclidean SH distance (first and second rows) and the Kullback-Leibler divergence (third and fourth rows) for the estimated ODFs from under-sampled synthetic multi-shell non-staggered HARDI data using bcsCSA-ODF, gCSA-ODF, and CSA-ODF (no CS algorithm) at two different signal to noise ratios, 15 and 25; and three fiber crossing angles, 45°, 60°, and 90°. Acceleration one corresponds here to 399 gradients.
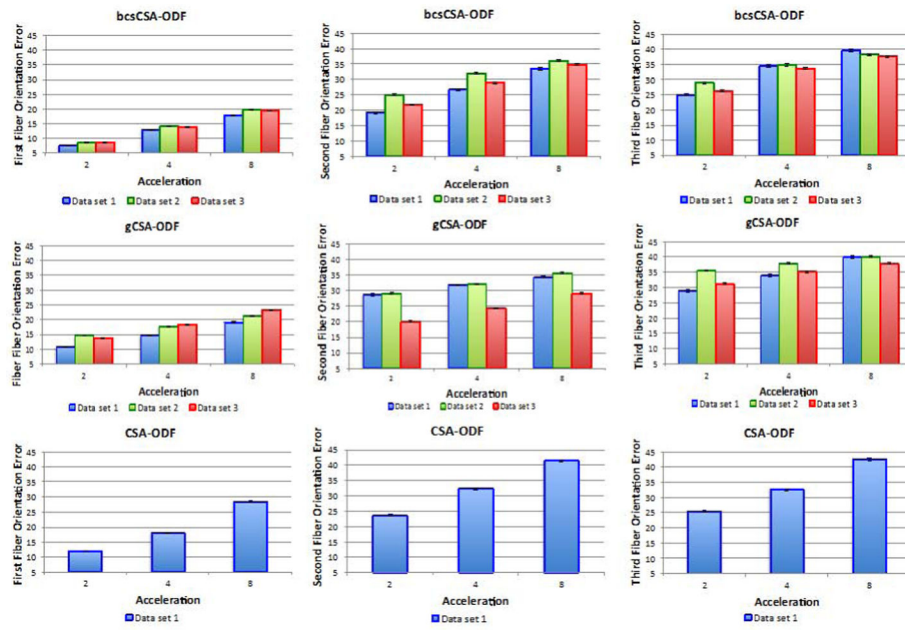
Figure 3.
Fiber orientation errors (FOEs) estimated from under-sampled synthetic multi-shell <u>staggered</u> HARDI data. First row corresponds to the FOEs of the estimated ODFs for a signal to noise ratio (SNR) of 15, while the second row shows the FOEs for a SNR of 25. All the figures indicate the FOEs using the proposed bcsCSA-ODF, gCSA-ODF, and CSA-ODF for three fiber crossing angles, 45°, 60°, and 90°. Acceleration one corresponds here to the full 399 gradients.

Figure 4.
Euclidean SH distance (first and second rows) and the Kullback-Leibler divergence (third and fourth rows) for the estimated ODFs from under-sampled synthetic multi-shell staggered HARDI data using bcsCSA-ODF, gCSA-ODF, and CSA-ODF (no CS algorithm) at two different signal to noise ratios, 15 and 25; and three fiber crossing angles, 45°, 60°, and 90°. Acceleration one corresponds here to 399 gradients.

Figure 5.
Fiber orientation errors estimated from under-sampled in-vivo multi-shell non-staggered HARDI data. First row shows the FOEs of the ODFs estimated using bcsCSA-ODF. Second row shows the FOEs computed using gCSA-ODF. Third row shows the FOEs using CSA-ODF on the under-sampled data, directly (no CS algorithm).
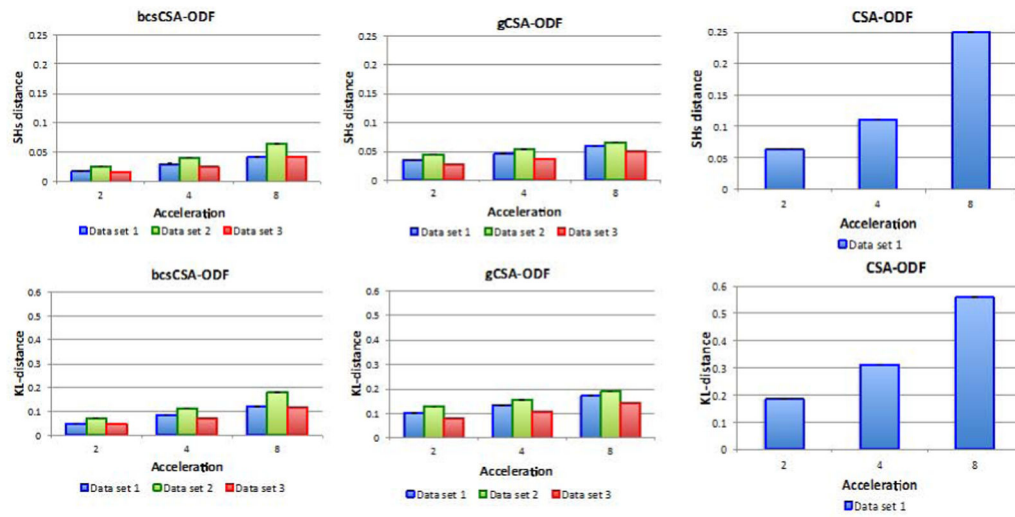
Figure 6.
Euclidean SH distance (first row) and Kullback-Leibler divergence (second row) for the estimated ODFs from under-sampled in-vivo multi-shell non-staggered HARDI data using bcsCSA-ODF, gCSA-ODF, and CSA-ODF (no CS algorithm).
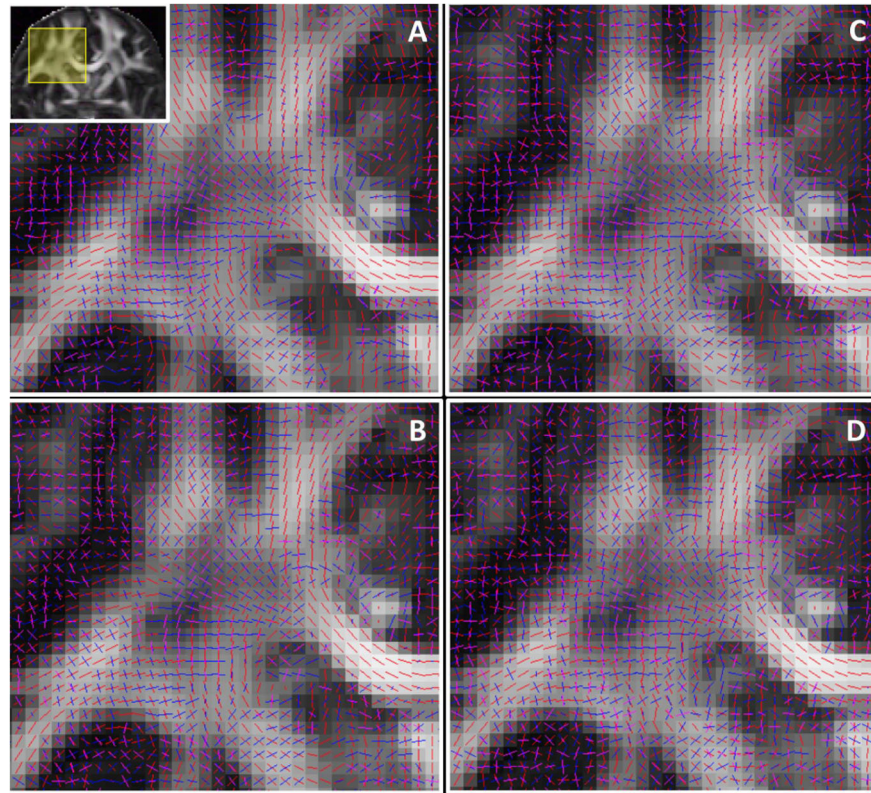
Figure 7.
Visualization of the estimated three principal fiber orientations (ODF peaks) on a coronal view in the region of the *centrum semiovale*, and obtained using A) bcs-CSA-ODF at acceleration one, B) bcs-CSA-ODF at acceleration four, C) gCSA-ODF at acceleration one, D) gCSA-ODF at acceleration four.
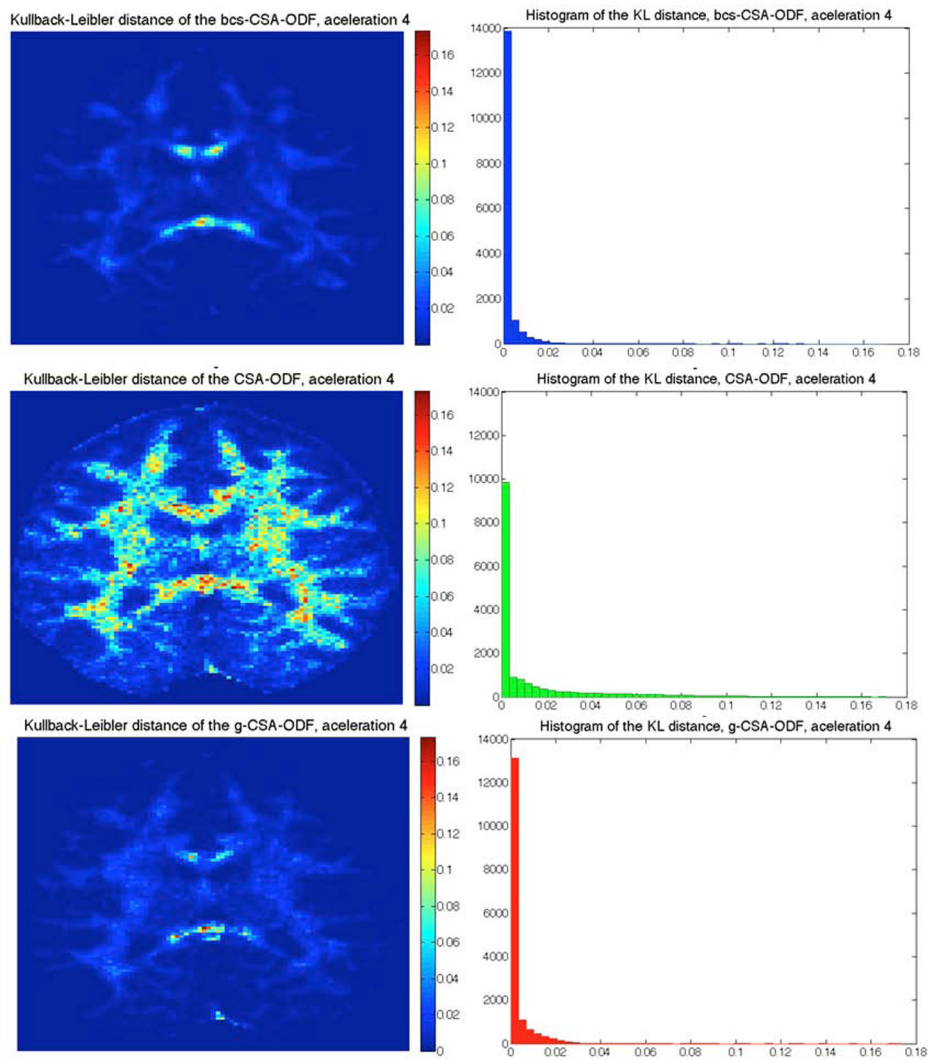
Figure 8.
Kullback-Leibler divergence of the reconstructed ODF from under-sampled in-vivo HARDI data (acceleration four) using bcsCSA-ODF (top), gCSA-ODF (middle), and CSA-ODF (bottom). Left images correspond to a color-coded representation of the KL divergence that goes from cyan (low) to magenta (high). Right images correspond to the histogram of the KL divergence. Results shown are for dataset 1 (see Figure 6).
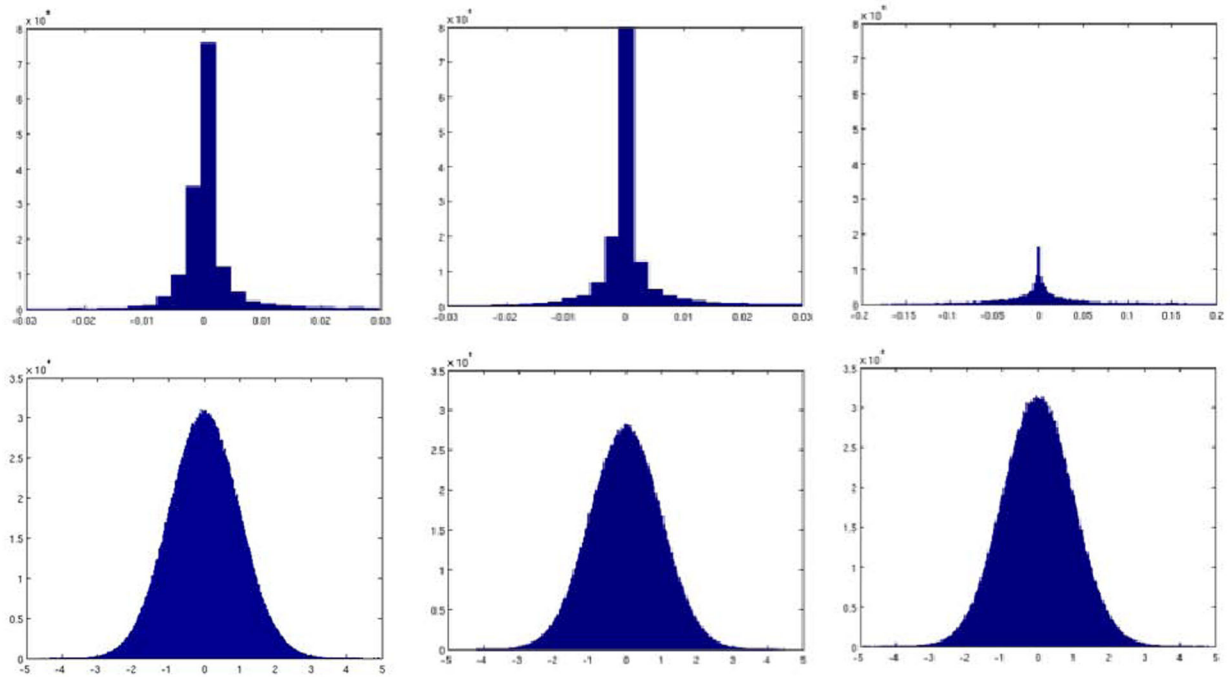
Figure 9.
First row indicates the normalized ODF uncertainty for acceleration one (left), four (center), and eight (right). Second row indicates the normalized volume estimation uncertainty for acceleration one (left), four (center), and eight (right). Uncertainties were compute on dataset 1.