HHS Public Access

Author manuscript

Med Image Comput Comput Assist Interv. Author manuscript; available in PMC 2018 September 01.

Published in final edited form as:

Med Image Comput Comput Assist Interv. 2017 September ; 10433: 602–610. doi: 10.1007/978-3-319-66182-7_69.

A Sparse Bayesian Learning Algorithm for White Matter Parameter Estimation from Compressed Multi-shell Diffusion MRI

Pramod Kumar Pisharady[1], **Stamatios N Sotiropoulos**[2,3], **Guillermo Sapiro**[4,5], and **Christophe Lenglet**[1]

[1]CMRR, Radiology, University of Minnesota, Minneapolis, Minnesota, USA

[2]Centre for Functional MRI of the Brain (FMRIB), University of Oxford, UK

[3]Sir Peter Mansfield Imaging Centre, School of Medicine, University of Nottingham, UK

[4]Electrical and Computer Engineering, Duke University, Durham, USA

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

Abstract

We propose a sparse Bayesian learning algorithm for improved estimation of white matter fiber parameters from compressed (under-sampled q-space) multi-shell diffusion MRI data. The multi-shell data is represented in a dictionary form using a non-monoexponential decay model of diffusion, based on continuous gamma distribution of diffusivities. The fiber volume fractions with predefined orientations, which are the unknown parameters, form the dictionary weights. These unknown parameters are estimated with a linear un-mixing framework, using a sparse Bayesian learning algorithm. A localized learning of hyperparameters at each voxel and for each possible fiber orientations improves the parameter estimation. Our experiments using synthetic data from the ISBI 2012 HARDI reconstruction challenge and in-vivo data from the Human Connectome Project demonstrate the improvements.

Keywords

Sparse Bayesian learning; linear un-mixing; multi-shell; diffusion MRI; sparse signal recovery

1 Introduction

Acquisition of images using diffusion MRI (dMRI) and representation of the MR signal using compartment models facilitate extraction of microstructural features of brain white matter [1]. In particular, estimation of orientations and volume fractions of the anisotropic compartments in these models helps infer the white matter fiber anatomy [2]. Accurate estimation of these parameters is challenged by the relatively limited spatial and angular resolutions of acquired dMRI data. Advances in magnetic field strength have significantly improved spatial resolution [3], though it may lead to increased scanning time. Recently developed multi-shell dMRI acquisition protocols, which improved the angular resolution, may further increase the length of the scanning time. Compressed sensing methods, which

require fewer measurements within a voxel, are effective ways to deal with the increased scan time.

Finding volume fractions and fiber directions with a large number of possible fiber orientations is an ill-posed problem. Considering the fact that the number of crossing fiber populations within a voxel is limited, we propose a sparse signal recovery algorithm, to address this issue. The algorithm which is based on sparse Bayesian learning (SBL) [4], is useful for improved inference from data with under-sampled q-space (i.e. lower number of diffusion encoding directions).

The seminal work by Tipping on SBL [4] using automatic relevance determination (ARD) [5] provides a framework for obtaining sparse solutions to regression and classification problems. The sparsity of parameters is enforced by selection of appropriate *prior* probability distributions for the parameters to be estimated. A mixture of zero-mean Gaussian distributions with individual hyperparameters for variance prior distributions promotes *relevance learning* [4]. The hyperparameters associated independently with every weight moderate the strength of the prior and govern the variances of the Gaussian scale mixture, by learning the hyperparameters from the data. We exploit this learning approach in SBL and estimate the fiber orientations and corresponding volume fractions from a dictionary representation of dMRI data. We use SBL for *selection* of the fiber orientations from a large number of possible fiber orientations and for *un-mixing* the corresponding volume fractions.

Our approach is different from existing methods [6–10] for reconstruction from under-sampled dMRI data in several aspects. The above works utilized basis-based transforms and exploited the sparsity in the basis dictionary representation. In this work we use a dictionary formulation of the dMRI data, but we consider the multiple anisotropic components (corresponding to fibers) and the single isotropic component in the diffusion model as the *end-members* in an *un-mixing* problem [11], and recover these end-members using an SBL based linear un-mixing approach. Previous study [12] has shown that l_1 norm minimization based approaches for promoting sparsity, which are widely used in spherical deconvolution based methods, have the drawback of inconsistency with the *sum-to-one* constraint (i.e., the physical constraint that the volume fractions of anisotropic and isotropic compartments within a voxel sum to unity). We demonstrate that sparse Bayesian learning within a linear un-mixing framework can address the sum-to-one and non-negativity (volume fractions ≥ 0) constraints, simultaneously promoting sparsity. The approach in SBL is typically much sparser as it is based on the notion of setting weights to zero (rather than constraining them to small values), and as it offers probabilistic predictions without the need to set additional regularization parameters.

ARD has been used for data-adaptive estimation of fiber parameters [2], avoiding data unsupported model complexities. The relevance learning in the proposed approach, which explicitly models sparsity, enhances the relevance determination by tuning the variance prior hyperparameters individually and independently for each possible fiber orientations. The proposed algorithm exploits the spatial redundancy in data representation better, and it improves the estimation of fiber orientations and volume fractions. We extend a dictionary

formulation of the dMRI data [13], using a multi-shell non-monoexponential model of diffusion [14], and propose an SBL based estimation algorithm for improved inference from single resolution multi-shell dMRI.

2 Methods

2.1 Dictionary Representation of Multi-shell Data

Multi-shell acquisitions using high b-values have the benefit of improved angular contrast, compared to single-shell acquisition schemes. The diffusion decay curve along any given gradient direction is shown to depart from *monoexponetial* decay to *non-monoexponential* decay, especially at b-values > 1500 s/mm^2. To address this issue, Jbabdi *et al.* proposed a non-monoexponential model [14] with a continuous gamma distribution of diffusivities, as an extension to the ball & stick model [1]. We use the non-monoexponential model for the dictionary representation of the dMRI data. The attenuation signal is given by [14],

$$A^k=\frac{S^k}{S^0}=\int_0^\infty A^k(d)p(d)\mathrm{d}d \quad (1)$$

where S^k is the signal after application of k^{th} diffusion-sensitizing gradient with direction g_k and b-value b_k, S^0 is the signal without diffusion gradient, $A^k(d)$ is the attenuation signal given by ball & stick model corresponding to a single diffusivity d, and $p(d) = \Gamma(\delta, \beta)$.

The integral in Equation (1) is evaluated analytically to represent the attenuation signal as [14],

$$A^k=\frac{S^k}{S^0}=\left(1-\sum_{n=1}^{N}f_n\right)\left(\frac{\beta}{\beta+b_k}\right)^\delta+\sum_{n=1}^{N}f_n\left(\frac{\beta}{\beta+b_k(g_k^T\upsilon_n)^2}\right)^\delta, \quad (2)$$

where f_n is the volume fraction of anisotropic compartment with orientation υ_n. The measured signal at a voxel is the sum of the attenuation signal and measurement noise,

$$y^k=S^k/S^0+\eta^k. \quad (3)$$

Based on (2) and (3), the measured signal along all K diffusion-sensitizing directions can be written in a *dictionary* form (4) as

$$y=\begin{pmatrix} \left(\frac{\beta}{\beta+b_1}\right)^\delta & \left(\frac{\beta}{\beta+b_1(g_1^T\upsilon_1)^2}\right)^\delta & \cdots & \left(\frac{\beta}{\beta+b_1(g_1^T\upsilon_N)^2}\right)^\delta \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{\beta}{\beta+b_K}\right)^\delta & \left(\frac{\beta}{\beta+b_K(g_K^T\upsilon_1)^2}\right)^\delta & \cdots & \left(\frac{\beta}{\beta+b_K(g_K^T\upsilon_N)^2}\right)^\delta \end{pmatrix}\cdot\begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_N \end{pmatrix}+\eta, \quad (4)$$

where $f_0 = \left(1 - \sum_{n=1}^{N} f_n\right)$, $f_n \geq 0$. Hence $\mathbf{y} = \mathbf{E}\mathbf{f} + \eta$, where $\mathbf{E}$ represents the local dictionary matrix for the diffusion data and $\mathbf{f}$ is the sparse vector representation of the data in the dictionary $\mathbf{E}$. The non-zero entries in $\mathbf{f}$ define the number and volume fraction of fibers in a voxel. The possible orientations of anisotropic components in the dictionary (second column onwards) are pre-specified and formed using a 5^{th} order icosahedral tessellation of the sphere with 10242 points. With this dictionary formulation the problem of finding the number of fibers, volume fractions and orientations reduces to accurately estimating the sparse vector $\mathbf{f}$. The estimation of $\mathbf{f}$ is detailed in the following sections.

2.2 Hierarchical Bayesian Framework

The learning algorithm in SBL is based on a hierarchical Bayesian Framework (Fig. 1). A mixture of zero-mean Gaussian distributions with individual hyperparameters controlling the variances is used as the prior on the parameter to be estimated (the volume fractions with predefined orientations here). Gamma distributions are used as *hyperpriors*, which form the priors over the hyperparameters. The mixture of Gaussians with hyperparameters associated independently with every weight was shown equivalent to using a product of Student-t priors, once the hyperparameters are integrated out [4]. Mathematically the prior over volume fractions is given by,

$$p(f/\alpha) = \prod_{n=1}^{N} \mathcal{N}(f_n|0, \alpha_n^{-1}), \quad (5)$$

where the hyper-parameter α_n controls the variance of individual Gaussians. The update procedure for α_i (detailed in Subsection 2.3) is such that many of the α are pushed to higher values, adapting to the data. The variance $1/\alpha$ of the corresponding Gaussians are pushed towards zero which forces the corresponding weights to be zero (or negligibly small), leading to a sparse solution.

2.3 Sparse Bayesian Learning Based Linear Un-mixing Inference

Assuming Gaussian noise, the likelihood of the data can be expressed as

$$p(y|f, \alpha, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{K}{2}} e^{-\frac{\|y - \mathrm{E}f\|_2^2}{2\sigma^2}}, \quad (6)$$

where σ^2 is the variance of the error in representation of y using dictionary $\mathbf{E}$ and volume fractions $\mathbf{f}$. Let $f^+ = [f_1, \ldots, f_{n_0}]^T$ be the volume fractions with n_0 non-zero anisotropic components, then f^+ belongs to a simplex S,

$$S = \left\{ f^+ | f_n \geq 0, \forall n = 1, \ldots, n_0, \sum_{n=1}^{n_0} f_n \leq 1 \right\}. \quad (7)$$

We follow the sparse inference procedure detailed in [4] (page 215, equations (7) to (13)) with a modification to resolve the linear un-mixing constraints [11]. We introduce non-negativity and sum-to-one constraints to the volume fractions posterior computation, to propose sparse linear un-mixing inference:

$$p(f^+|y,\alpha,\sigma^2)\sim e^{-(f^+-\mu_f)^T\Lambda_f^{-1}(f^+-\mu_f)}1_S(f^+), \quad (8)$$

where

$$\Lambda_f=[\sigma^{-2}(E_{n0}^+-e_0u^T)^T(E_{n0}^+-e_0u^T)+A]^{-1}, \quad (9)$$

and

$$\mu_f=\sigma^{-2}\Lambda_f(E_{n0}^+-e_0u^T)^T(y-e_0), \quad (10)$$

with u a $1 \times n_0$ vector, $[1, \ldots, 1]^T$, and $A = diag(\alpha_0, \alpha_1, \ldots, \alpha_N)$. E_{n0}^+ contains the columns of E that correspond to n_0 non-zero coefficients in f^+ (the *effective dictionary*) and e_0 is the first column in the dictionary, which corresponds to the isotropic compartment. $1_S(f^+)$ in (8) is 1 if $f^+ \in S$ and 0 otherwise.

Each hyper-parameter α_n in A are updated iteratively [4] as per $\alpha_n^{new}=\gamma_n/\mu_n^2$ where $\gamma_n = 1 - \alpha_n * \Lambda_{nn}$ and Λ_{nn} is the n^{th} diagonal element of the posterior volume fractions covariance (9). The noise variance σ^2 is updated as

$$(\sigma^2)^{new}=\frac{\|(y-e_0)-(E_{n0}^+-e_0u^T)\mu_f\|^2}{K-\sum_n\gamma_n}. \quad (11)$$

3 Experiments and Results

3.1 Synthetic Data from HARDI Reconstruction Challenge

We performed experiments using the test dataset from the *HARDI reconstruction challenge* organized as part of the ISBI 2012 conference [15]. The synthetic data is generated using the test data phantom (16×16×5 voxels) and the data simulation algorithm the challenge organizers released. Rician noise is added to the data with SNR 10. We used the gradient tables from the Human Connectome Project (HCP) [16] to simulate the data. The full dataset had 270 diffusion measurements (and 18 b_0 measurements) with three b-values (1000, 2000, and 3000 s/mm^2). We under-sampled the data by a factor of up to 6 (45 measurements) in our experiments.

Daducci *et al.* [15] reported the results of the challenge and compared 20 algorithms used for recovering the intra-voxel fiber structures. We used the reported results in [15] as well as the

results of *BedpostX* algorithm [2] for comparison of our results. The algorithm performance is evaluated using two criteria, the correct assessment of the number of fiber populations: *success rate* (12) and the error in orientation estimation: *angular precision* (13) [15]

$$\text{Success Rate (SR)} = \left(1 - \frac{|M_{\text{true}} - M_{\text{estimated}}|}{M_{\text{true}}}\right) \times 100, \quad (12)$$

where M_{true} and $M_{\text{estimated}}$ are, respectively, the true and estimated number of fiber compartments in a voxel.

$$\text{Angular Precision (AP)} = \frac{180}{\pi} \arccos\left(|\mathbf{d}_{\text{true}} \cdot \mathbf{d}_{\text{estimated}}|\right), \quad (13)$$

where $\mathbf{d}_{\text{true}}$ and $\mathbf{d}_{\text{estimated}}$ are, respectively, the true and estimated fiber orientations in a voxel.

Fig. 2 compares the mean SR and AP (across 1280 voxels) of our algorithm with that of *BedpostX* (multi-shell) [2] under different under-sampling factors, as well as with the top five algorithms reported in [15]. It also shows the standard deviation in AP as error bar. On comparison with *BedpostX*, the proposed SBL based approach provides higher SR and improved AP (lower error), with lower uncertainty. The algorithm provides reasonably stable performance with increase in under-sampling factor. On comparison with the top five algorithms [17, 7, 18–20] reported in [15] (which are not matched for the number of diffusion measurements), the proposed method provides the best performance in terms of SR as well as AP when the full dataset (270 samples) is used. The proposed method also provides the best success rate (69.70%) even with an under-sampling factor of 6 (45 samples), but at this under-sampling factor the AP of the algorithm decreases below two of the compared methods (NN-L2 [7] and L2-L2 [19]) which used similar number of samples (48 and 37 respectively).

3.2 In-vivo Data from the Human Connectome Project

We performed in-vivo experiments using the exemplar dataset (subject ID: 100307) from the HCP [16]. The image size is 145×174×145, with 1.25 mm^3 isotropic resolution. The full dataset has 270 diffusion measurements. The reported results (Fig. 3, upper panels and lower left panel) are with an under-sampling factor of two (135 samples). Our results are compared with multi-shell *BedpostX* [2] and the multi-shell multi-tissue constrained spherical deconvolution (msmt-CSD) [21] (the implementation available in MRtrix3 is used, with the default parameters). We used the Connectome Workbench [22] from the HCP for visualizing the results of our algorithm and *BedpostX*. MRtrix (*mrview*) is used for visualizing the orientation distribution functions (ODFs) from CSD.

On comparison the proposed method provides better detection of crossing fibers (highlighted with red and blue arrows in Fig. 3). The lower estimation uncertainty of the proposed method on comparison to *BedpostX* can be observed in the areas with *cleaner*

orientations (highlighted with green arrows). The improved detection of second and third fibers is also demonstrated through a graph (Fig. 3, lower-right panel) showing detected second and third fiber crossings at two representative regions of interests (ROIs); the left superior longitudinal fasciculus (SLF) and left posterior corona radiata (PCR), with respect to *acceleration* (under-sampling from 270 to 67 samples). On comparison to *BedpostX*, the proposed method exhibits greater robustness in the number of detected crossings with increase in acceleration (for example, only 1.2% decrease in the number of detected second fibers in the left PCR compared to a 17.3% decrease with *BedpostX*, at an acceleration factor of four). We noted similar comparisons for the corresponding ROIs on the right side too.

4 Discussion and Conclusion

We proposed an SBL based sparse signal recovery algorithm for estimation of white matter fiber parameters from multi-shell single resolution dMRI data. The elements of our overcomplete dictionary for each voxel are obtained from an icosahedral tessellation of the sphere, with 10242 possible fiber directions. The estimated fiber orientations are approximated to the nearest pre-specified orientation during the learning process. The worst-case error due to this approximation is 1.18 degrees. The number of possible fiber directions can be increased for slightly better orientation accuracy, at the expense of computational time. The major benefit from the SBL based algorithm is the more accurate estimation of dictionary weights: the improved detection of single and crossing fiber populations, as reflected by the reported higher success rates. The algorithm has also shown very good performance in orientation estimation and reliability with under-sampling. The results from the HCP data we presented are well-representative of the results in other subjects we analyzed. We did not note any significant bias with Rician noise, though our model assumed Gaussian noise.

The current implementation of our algorithm takes about 5.7 seconds to process one voxel with a CPU speed of 2.6 GHz, which is approximately 2 times slower compared to *BedpostX*. To speed-up the processing, we parallelized the algorithm using OpenMP. It takes an average time of 219 milliseconds / voxel on a server with 26 processors. The computational performance of the algorithm will be improved using GPU/CUDA in our future work.

Acknowledgments

This work was partly supported by NIH grants P41 EB015894, P30 NS076408, and the Human Connectome Project (U54 MH091657).

References

1. Behrens TE, Woolrich MW, et al. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. MRM. 2003; 50:1077–1088. [PubMed: 14587019]
2. Behrens TE, Berg HJ, Jbabdi S, et al. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? Neuroimage. 2007; 34:144–155. [PubMed: 17070705]
3. Vu AT, Auerbach E, Lenglet C, et al. High resolution whole brain diffusion imaging at 7t for the human connectome project. Neuroimage. 2015; 122:318–331. [PubMed: 26260428]

4. Tipping ME. Sparse bayesian learning and the relevance vector machine. Journal of Machine Learning Research. 2001; 1:211–244.
5. MacKay DJC. Bayesian methods for backpropagation networks. Springer, Models of Neural Networks III. 1994:211–254. chapter 6.
6. Duarte-Carvajalino JM, Lenglet C, et al. Estimation of the CSA-ODF using bayesian compressed sensing of multi-shell hardi. MRM. 2014; 72:1471–1485. [PubMed: 24338816]
7. Manzanares AR, et al. Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. IEEE TMI. 2007; 26:1091–1102.
8. Rathi Y, Michailovich O, Setsompop K, Bouix S, Shenton ME, Westin CF. Sparse multi-shell diffusion imaging. MICCAI. 2011; 2011:58–65.
9. Tristan-Vega A AWC. Probabilistic ODF estimation from reduced HARDI data with sparse regularization. MICCAI. 2011; 2011:182–190.
10. Aranda R, Manzanares AR, Rivera M. Sparse and adaptive diffusion dictionary for recovering intra-voxel white matter structure. MedIA. 2015; 26:243–255.
11. Dobigeon N, et al. Semi-supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery. IEEE TSP. 2008; 56:2684–2695.
12. Daducci A, et al. Sparse regularization for fiber odf reconstruction: From the suboptimality of l2 and l1 priors to l0. Medical Image Analysis. 2014; 18:820–833. [PubMed: 24593935]
13. Pisharady PK, et al. Sparse Bayesian inference of white matter fiber orientations from compressed multi-resolution diffusion MRI. MICCAI. 2015; 2015:117–124.
14. Jbabdi S, et al. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. MRM. 2012; 68:1846–1855. [PubMed: 22334356]
15. Daducci A, et al. Quantitative comparison of reconstruction methods for intravoxel fiber recovery from diffusion mri. IEEE TMI. 2014; 33:384–399.
16. Essen DCV, Smith SM, Barch DM, et al. The wu-minn human connectome project: An overview. Neuroimage. 2013; 80:62–79. [PubMed: 23684880]
17. Rodriguez EJC, Medina YI, Alemn-Gmez Y, Melie-Garca L. Deconvolution in diffusion spectrum imaging. Neuroimage. 2010; 50:136–149. [PubMed: 19962440]
18. Ozarslan E, Shepherd TM, et al. Resolution of complex tissue microarchitecture using the diffusion orientation transform. Neuroimage. 2006; 31:1086–1103. [PubMed: 16546404]
19. Rodriguez EJC, Lin CP, Medina YI, Yeh CH, Cho KH, Melie-Garcia L. Diffusion orientation transform revisited. Neuroimage. 2009; 49:1326–1339. [PubMed: 19815083]
20. Tournier JD, Calamante F, Connelly A. Robust determination of the fibre orientation distribution in diffusion mri. Neuroimage. 2007; 35:1459–1472. [PubMed: 17379540]
21. Jeurissen B, et al. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion mri data. Neuroimage. 2014; 103:411–426. [PubMed: 25109526]
22. Marcus DS, Harwell J, et al. Informatics and data mining: Tools and strategies for the human connectome project. Frontiers in Neuroinformatics. 2011; 5:1–12. [PubMed: 21472085]

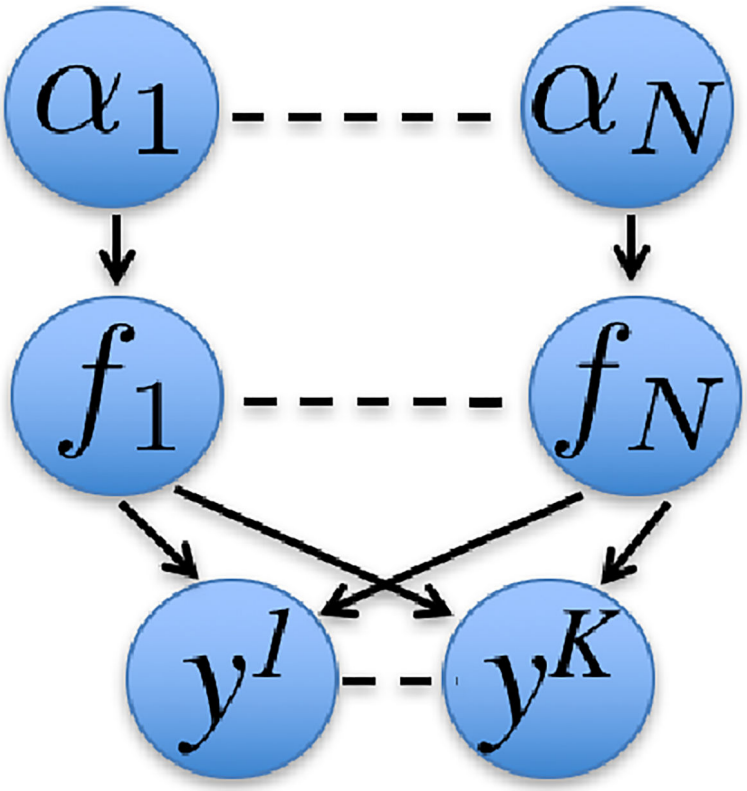
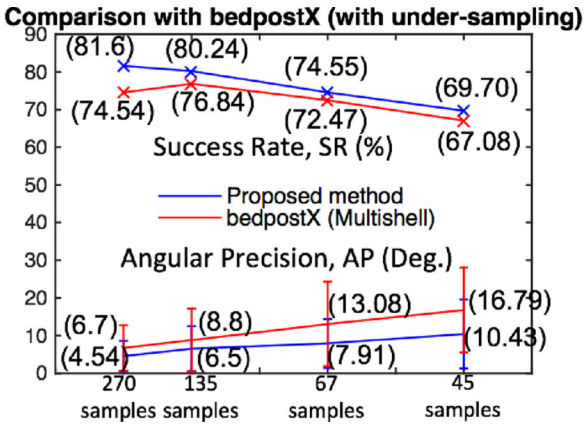


Fig. 1.

The hierarchical Bayesian framework used in the proposed algorithm. y^k is the measured signal along diffusion gradient direction k. f_n is the n-*th* component of the anisotropic volume fractions vector and α_n is the hyperparameter in the prior distribution of f_n.



Comparison with 5 top methods reported in Daducci *et al.*

Method	SR (%)	AP (Deg.)	# samples
Proposed	81.6	4.54	270
DSI_{LR} [17]	64.0	7.5	257
NN-L2 [7]	60.3	6.5	48
DOT [18]	57.06	11.22	60
L2-L2 [19]	55.29	7.84	37
CSD [20]	49.12	11.08	60

Fig. 2.

Comparison of SR and AP (mean across 1280 voxels). Left panel shows the variation in SR and AP with increase in *acceleration* (under-sampling factor). The error bars shown represent the standard deviation in AP. Right panel shows comparisons with top five algorithms reported in [15].

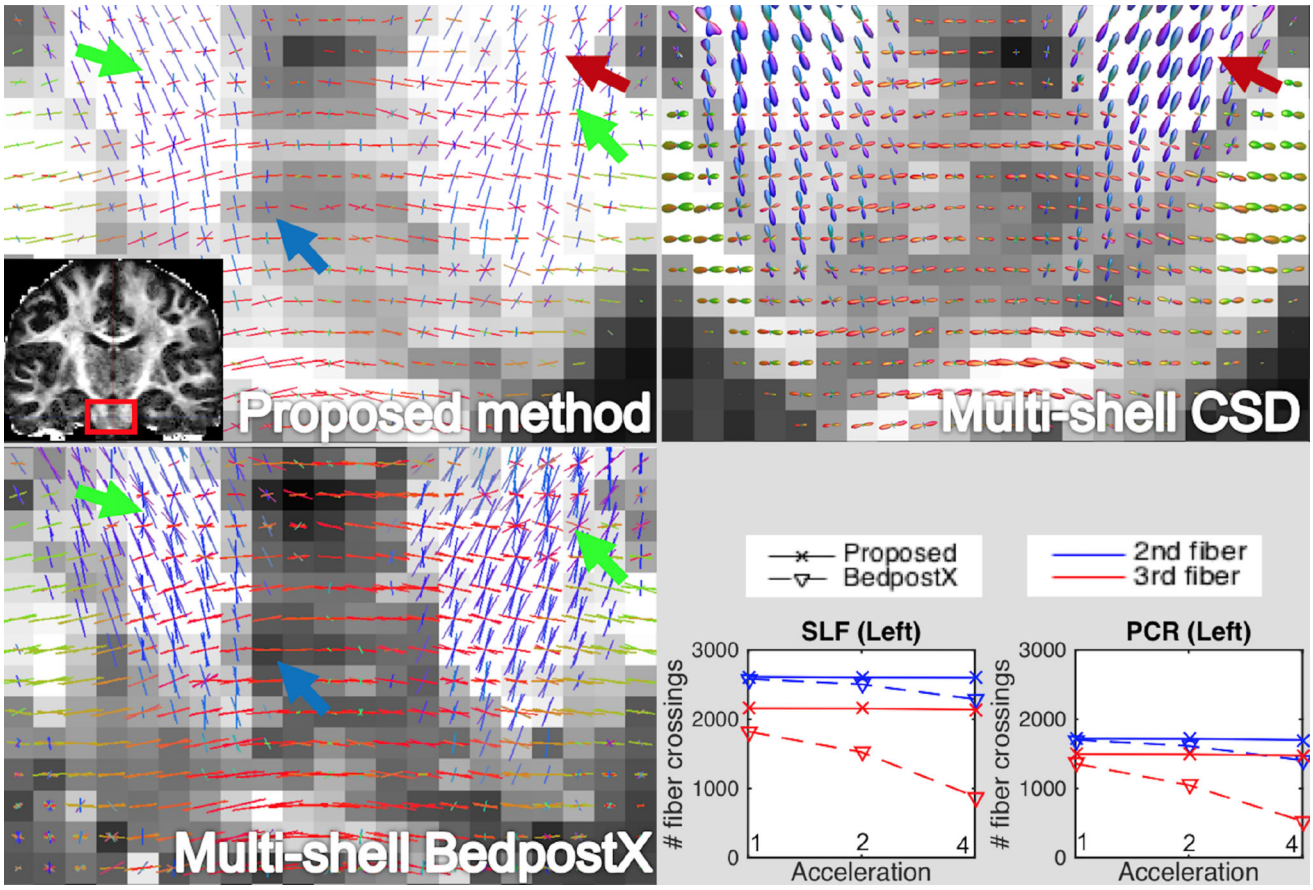


Fig. 3.

Comparison between proposed method, CSD, and *BedpostX*. Upper panels and lower-left panel show color coded orientation estimates (ODFs in the case of CSD) at the pons region highlighted in the inset view. The background is the sum of anisotropic volume fractions for the proposed method and *BedpostX*, and FA for the CSD. The areas highlighted with arrows depict the improvements; the better detection of fiber crossings (red and blue arrows) and the lower estimation uncertainty (green arrows). Lower right panel shows the detected number of second (blue) and third (red) fiber crossings at two representative ROIs, and its variation with acceleration (under-sampling factor).