



## CDMRI 2012

# MICCAI 2012 workshop on Computational Diffusion MRI

Held in Conjunction with MICCAI 2012  
Nice, France, October 5th, 2012.

<http://cmic.cs.ucl.ac.uk/cdmri12/>

### Organizers:

Eleftheria Panagiotaki

University College London, UK

Lauren O'Donnell

Harvard Medical School, USA

Thomas Schultz

MPI Intelligent Systems, Germany

Gary Hui Zhang

University College London, UK

## Preface

The 2012 MICCAI Workshop on Computational Diffusion MRI (CDMRI) was held on October 5th, in Nice, France, under the auspices of the 15th International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2012. It is already the fourth event in a successful series, following the exciting and well-attended workshops in 2008, 2010, and 2011.

The 20 original research papers collected in this proceedings volume clearly demonstrate that the field remains as vibrant and diverse as ever. From fundamental theoretical work on mathematical diffusion modeling, to the development and evaluation of robust algorithms for tractography and connectivity mapping, diffusion MRI continues to provide intricate mathematical and computational challenges. Compressive sensing is being explored to reduce the measurement times of this demanding imaging modality; classic image processing and image analysis tasks such as registration and segmentation remain non-trivial when they have to be performed on this complex and noisy type of image data.

Publicly available research software is indispensable for scientific progress in a computationally demanding field such as diffusion imaging. Therefore, for the first time, this year's workshop also offered the opportunity to present live demos of research software, and the four software packages NeurInfarct, Camino, DTI-TK, and MITK Diffusion have been demonstrated to the workshop audience.

We are confident that computational research as it is exemplified by the contributions at the CDMRI workshop will continue to provide a unique insight into the microstructure of living tissue, enable in-vivo connectivity mapping of the brain, give fundamental new insights in neuroscience and neuroanatomy, and support a widespread transfer of diffusion MRI into the clinic.

We would like to express our gratitude to members of the Program Committee for ensuring the quality of the presented work, as well as chairing sessions at the meeting.

It is our distinct pleasure to welcome participants to CDMRI 2012, and to provide this record of the exciting work represented at the workshop.

September 2012

Eleftheria Panagiotaki  
Lauren O'Donnell  
Thomas Schultz  
Gary Hui Zhang

# Table of Contents

<b>Preface</b>	i
<b>Table of Contents</b>	iii
<b>Organization</b>	vi
<b>Oral Session I: Tractography and Applications</b>	1
Utilising Measures of Fiber Dispersion in White Matter Tractography . . . . .	2
<i>Matthew Rowe, Hui Zhang, and Daniel Alexander</i>	
Simultaneous ODF Estimation and Robust Probabilistic Tractography from HARDI . . . . .	13
<i>Hasan Ertan Çetingül, Laura Dumont, Mariappan S. Nadar, Paul M. Thompson, Guillermo Sapiro, and Christophe Lenglet</i>	
A Unified Tractography Framework for Comparing Diffusion Models on Clinical Scans . . . . .	25
<i>Christian Baumgartner, Oleg Michailovich, Jim Levitt, Ofer Pasternak, Sylvain Bouix, Carl-Fredrik Westin, and Yogesh Rathi</i>	
The Effect of Reorientation of the Fibre Orientation Distribution on Fibre Tracking . . . . .	33
<i>Daan Christiaens, Thijs Dhollander, Frederik Maes, Stefan Sunaert, and Paul Suetens</i>	
<b>Oral Session II: Registration</b>	45
Registration of Spherical Functions from High Angular Resolution Diffusion Imaging using the Heat Kernel Signature and Mobius Transformation . . . . .	46
<i>Liang Zhan, Yalin Wang and Paul M. Thompson</i>	
Detection of Unique Point Landmarks in HARDI Images of the Human Brain	58
<i>Henrik Skibbe and Marco Reisert</i>	
<b>Oral Session III: Diffusion Modeling and Experimental Design</b>	66
Generalized Constant Solid Angle ODF and Optimal Acquisition Protocol for Fiber Orientation Mapping . . . . .	67
<i>Amith Kamath, Iman Aganj, Junqian Xu, Essa Yacoub, Kamil Ugurbil, Guillermo Sapiro, and Christophe Lenglet</i>	

Decomposition of Higher-Order Homogeneous Tensors and Applications to HARDI.....	79
<i>Evgeniya Balmashnova, Andrea Fuster, and Luc M. J. Florack</i>	
Dynamic Diffusion Basis Functions for Axon Fiber Structure Estimation from DW-MRI .....	90
<i>Omar Ocegueda and Mariano Rivera</i>	
Complete Fourier Direct Magnetic Resonance Imaging (CFD-MRI) Equations for Diffusion MRI Signal.....	102
<i>Alpay Özcan</i>	
<b>Poster Session</b>	114
Multi-class DTI Segmentation: A Convex Approach.....	115
<i>Yuchen Xie, Ting Chen, Jeffrey Ho, and Baba C. Vemuri</i>	
A Computational Framework for Experimental Design in Diffusion MRI ....	124
<i>Emmanuel Caruyer and Rachid Deriche</i>	
A Framework for Multi-task Bayesian Compressive Sensing of Diffusion Weighted MRI.....	132
<i>Julio M. Duarte-Carvajalino, Christophe Lenglet, Kamil Ugurbil, Steen Moeller, Lawrence Carin, and Guillermo Sapiro</i>	
Introduction of a Quantitative Evaluation Method for White Matter Tractography using a HARDI-based Reference.....	145
<i>Peter F. Neher, Bram Stieltjes, Hans-Peter Meinzer, and Klaus H. Fritzsche</i>	
Constrained Diffusion Kurtosis Imaging Using Ternary Quartics and MLE ..	153
<i>Tristan Milne, Aurobrata Ghosh, and Rachid Deriche</i>	
Generalized Invariants of a 4th order tensor: Building blocks for new biomarkers in dMRI .....	165
<i>Aurobrata Ghosh, Théodore Papadopoulo, and Rachid Deriche</i>	
Anatomical Tissue Probability Priors for Tractography.....	174
<i>Gabriel Girard and Maxime Descoteaux</i>	
White matter structure assessment from reduced HARDI data using low-rank polynomial approximations .....	186
<i>Yaniv Gur, Fangxiang Jiao, Stella Xinghua Zhu, and Chris R. Johnson</i>	

Statistical Analysis of Maximum Density Path Deformation Fields in White Matter Tracts .....	198
--	-----

*Julio Villalon-Reina, Gautam Prasad, Shantanu H. Joshi, Maria Jalbrzikowski, Arthur Toga, Carrie Bearden, and Paul M. Thompson*

A minimum cost approach to connectivity from orientation distribution functions via efficient multi-directional graph propagation .....	210
---	-----

*Ipek Oguz, Alexis Boucharin, Wenyu Lu, Clement Vachet, Francois Budin, Yundi Shi, and Martin Styner*

### **Demo Session**

NeurInfarct: a pipeline for the segmentation of infarct core and prediction infarct growth using acute-phase clinical diffusion-weighted MRI

Camino: an open-source software toolkit for diffusion MRI processing

DTI-TK: a spatial normalization and atlas construction toolkit optimized for examining white matter morphometry using DTI data

MITK Diffusion: this application offers a selection of image analysis algorithms for the processing of diffusion-weighted MR images

# Organization

## Workshop Organizers

Eleftheria Panagiotaki	University College London, UK
Lauren O'Donnell	Harvard Medical School, USA
Thomas Schultz	MPI Intelligent Systems, Germany
Gary Hui Zhang	University College London, UK

## Program Committee

Nagesh Adluru	University of Wisconsin, USA
Andrew Alexander	University of Wisconsin, USA
Alfred Anwander	MPI Human Cognitive & Brain Sciences, Germany
Sylvain Bouix	Harvard Medical School, USA
Ai Wern Chung	St George's University of London, UK
Moo K Chung	University of Wisconsin, USA
Jon Clayden	University College London, UK
Philip Cook	University of Pennsylvania, USA
Olivier Coulon	CNRS, France
Kathleen Curran	University College Dublin, Ireland
Fani Deligianni	Imperial College London, UK
Rachid Deriche	INRIA, France
Maxime Descoteaux	Sherbrooke University, Canada
Ivanna Drobnjak	University College London, UK
Tim Dyrby	Copenhagen University Hospital, Denmark
Saad Jbabdi	University of Oxford, UK
Gordon Kindlmann	University of Chicago, USA
Jan Klein	Fraunhofer MEVIS, Germany
Uwe Klose	University Hospital Tübingen, Germany
Gemma Nedjati-Gilani	University College London, UK
Ofer Pasternak	Harvard Medical School, USA
Xavier Pennec	INRIA, France
Cyril Poupon	NeuroSpin, France
Yogesh Rathi	Harvard Medical School, USA
Marco Reisert	University Medical Center Freiburg, Germany
Bart ter Haar Romeny	Eindhoven University of Technology, Netherlands
Torben Schneider	University College London, UK
Kiran Seunarine	University College London, UK

Stamatis Sotiropoulos  
Jacques-Donald Tournier  
Nicolas Toussaint  
Pew-Thian Yap  
Paul Yushkevich

University of Oxford, UK  
Florey Neuroscience Institutes, Australia  
King's College London, UK  
University of North Carolina at Chapel Hill, USA  
University of Pennsylvania, USA

# Oral Session I: Tractography and Applications

# Utilising measures of fiber dispersion in white matter tractography

Matthew Rowe, Hui Zhang, Daniel Alexander

Centre for Medical Image Computing, Department of Computer Science, University  
College London, UK  
`matthew.rowe.09@ucl.ac.uk`

**Abstract.** This paper proposes a new tractography algorithm utilising measures of fiber dispersion derived from diffusion weighted magnetic resonance (DW-MR) imaging. Tractography estimates connectivity by integrating a pathway from a seed point following directional information derived from DW-MR images. Current tractography techniques follow a discrete set of directions given in each voxel of a DW-MR image and probabilistic techniques account for noise induced uncertainty on those discrete directions. Histological evidence suggests that fiber orientation dispersion exists in areas of white matter such as the centrum semiovale, representing a continuum of potential fiber orientations which cannot be accurately summarised by a limited set of discrete directions. Recent studies have shown that measures of fiber dispersion in brain white matter can be directly measured from DW-MR imaging data and explicitly represented in the orientation distribution function (ODF) of a voxel, but such measures have yet to be used in guiding tractography algorithms. We present a tracking algorithm which makes use of ODFs which account for underlying fiber dispersion to trace potential fiber pathways, we compare this method with traditional tracking methods on simulated data and *in vivo* human data, showing that measures of fiber dispersion can aid tractography in finding connectivity commonly missed by current tractography methods.

## 1 Introduction

Tractography is a powerful tool to probe the geometric structure of white matter non-invasively *in vivo* from diffusion weighted magnetic resonance images. Tractography algorithms estimate connectivity between different functional brain regions, giving us insight into brain function of great importance to neurological knowledge and understanding [1]. Tractography can also provide us with information on white matter structure which is difficult or impossible to define using anatomical images alone, aiding in surgical planning [2].

Tractography algorithms estimate connectivity by integrating pathways through a DW-MR image volume to estimate potential connections between different regions of the brain. Current deterministic and probabilistic tractography techniques follow a discrete set of directions given by diffusion tensor imaging [3]

or more sophisticated multifiber techniques such as Qball [4] and PASMRI [5], which account for multiple fiber populations per voxel. Probabilistic tractography accounts for noise-induced uncertainty in the dominant direction.

A significant flaw of current tractography approaches is the assumption of a discrete set of fiber directions per voxel. Post mortem dissection confirms that in addition to crossing fiber configurations, regions of fanning fibers exist in brain regions such as the centrum semiovale [6]. Current tractography techniques do not explicitly address the continuum of potential directions available in such regions due to underlying fiber dispersion.

Global tractography [7–9] provides an elegant solution to resolving contentious voxel fiber configurations such as crossing and fanning by solving for the entire projectome simultaneously. The drawback of a global approach is however the practicality of obtaining solutions. Many global tractography implementations require extensive computing resources or have long running times of several weeks on standard hardware. Combined with the fact that reaching the global minimum of such an astronomical optimisation is a practical impossibility, global tractography is not necessarily an ideal solution.

Recent studies [10–12] have demonstrated that intra-voxel fiber dispersion can be estimated using DW-MRI. These estimates of fiber dispersion within a voxel can provide tractography algorithms the appropriate basis to fully explore potential fiber pathways in regions of dispersing fiber structure such as the centrum semiovale, avoiding the potential for false negative connections due to the lack of coverage of the potential trajectories in fanning regions based on the assumption of discrete voxel fiber directions. Although Kaden *et al* [11] estimates fiber dispersion in each voxel this information is only used to sample the possible orientation at the initial tracking location, subsequent tracking is limited to following the principle direction in each incident voxel.

In this paper we present the first tracking algorithm utilising estimates of intra-voxel fiber dispersion to explore the connectome thoroughly. The algorithm makes use of distributions derived from fitting models intrinsically incorporating fiber dispersion to the diffusion weighted MR data using NODDI (neurite orientation dispersion and density imaging) [13]. Section 2.1 describes the method of obtaining the ODFs based on underlying fiber dispersion and in section 2.2 we outline the details of the tracking algorithm. In section 3.1 we present a comparison with traditional tracking methods on simulation data, showing the advantages of utilising dispersive ODFs. In section 3.4 we apply the algorithm to *in vivo* data and compare with results from traditional tracking.

## 2 Methods

### 2.1 DispersionODF

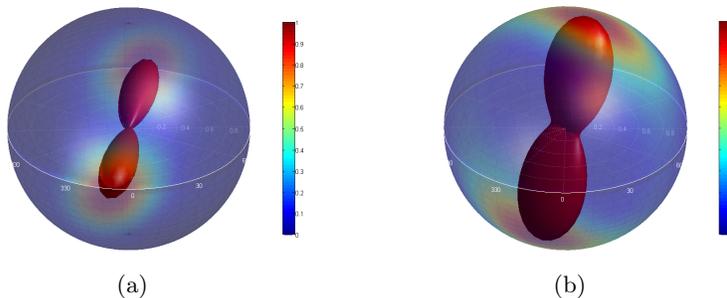
To obtain the dispersive ODFs from *in vivo* data we use NODDI [13], which produces estimates of the orientation distribution function (ODF). In [13] the ODFs are estimated by fitting a Watson distribution to the data. Here we extend the technique to account for cylindrically assymetrical dispersion by fitting

a Bingham distribution, the details of which will be reported elsewhere. The models are fitted to the data with the routine described in [14]. We find  $\kappa_1$  and  $\kappa_2$  for the Bingham model in every voxel where the Bingham distribution  $f : \mathbb{S}^2 \mapsto \mathbb{R}^+$  is described as:

$$f(\mathbf{n}) = F_1\left(\frac{1}{2}, \frac{3}{2}, \kappa_1, \kappa_2\right)^{-1} \exp[\kappa_1(\boldsymbol{\mu}_1 \cdot \mathbf{n})^2 + \kappa_2(\boldsymbol{\mu}_2 \cdot \mathbf{n})^2] \quad (1)$$

$\kappa_1$  and  $\kappa_2$  are parameters defining the degree of dispersion along the axes  $\boldsymbol{\mu}_1$  and  $\boldsymbol{\mu}_2$  respectively.  $F_1(1/2, 3/2, \kappa_1, \kappa_2)$  is the hypergeometric function. A chosen fiber signal model is used to estimate the MR signal from a fiber orientation distribution described by the above Bingham distribution and  $\kappa_1, \kappa_2, \boldsymbol{\mu}_1$  and  $\boldsymbol{\mu}_2$  are optimised to find the best description for the data. An isotropic compartment is also added to the signal model. This provides us with a distribution representing fiber dispersion derived directly from a model fitted to the data in each voxel, representing an accurate estimate of the potential orientation dispersion of the underlying tissue.

The Bingham distribution is an antipodially symmetric distribution, which, in contrast to the Watson distribution, is not cylindrically symmetric, giving a measure of the degree of orientation dispersion separately for two orthogonal axes which are perpendicular to the principle direction. The Bingham distribution therefore nicely summarises the degree of fiber dispersion in each voxel along two separate, orthogonal axes. Figure 1 shows Bingham distributions fitted in a voxel in the corpus callosum 1(a) and another voxel in the centrum semiovale 1(b) showing how the Bingham distribution captures the higher degree of dispersion in the centrum semiovale where histology confirms the presence of fanning fiber structures.



**Fig. 1.** Example Bingham distributions fitted in single voxels in the mid-sagittal corpus callosum 1(a) and deep within the centrum semiovale 1(b). The red interior structure shows the form of the distribution with probability density plotted with respect to polar radius. The probability density is also projected onto the translucent outer sphere in colour. The colour key depicts proportion of maximum probability density.

## 2.2 Tracking Algorithm

Our tracking algorithm adapts the strategy used by Friman in [15]. This approach allows us to employ a fully probabilistic framework exploring distributions based on fiber dispersion while applying suitable priors to limit curvature. Friman’s approach accommodates only the uncertainty in the principle diffusion direction induced by noise, image artifacts and partial volume effects; we instead incorporate underlying fiber dispersion directly in the ODF used to guide the tracking.

Specifically, to propagate the track through the image, starting from a seed, we choose a propagation direction  $v_i$  from a distribution formed from the product of the local ODF and a prior on the allowable deviation from the previous direction  $v_{i-1}$ :

$$P(\hat{v}_i|\hat{v}_{i-1}, \kappa_1, \kappa_2) = \frac{P(\hat{v}_i|\kappa_1, \kappa_2)P(\hat{v}_i|\hat{v}_{i-1})}{P(\hat{v}_i)}, \quad (2)$$

where  $P(\hat{v}_i|\kappa_1, \kappa_2)$  is the Bingham distribution described above.

For the angular Prior  $P(\hat{v}_i|\hat{v}_{i-1})$ , we use a distribution given by:

$$P(\hat{v}_i|\hat{v}_{i-1}) = \begin{cases} (\hat{v}_i^T \hat{v}_{i-1})^\gamma, & \text{if } \hat{v}_i^T \hat{v}_{i-1} \geq 0. \\ 0, & \text{if } \hat{v}_i^T \hat{v}_{i-1} < 0. \end{cases} \quad (3)$$

Sampling from this joint distribution allows exploration of the potential path directions in dispersive fiber regions while regularizing the curvature of the path.  $\gamma$  defines the strength of the curvature prior. Low values accommodate large degrees of deviation per track step, exploring more of the dispersion in each voxel, however, this also produces highly irregular tracts. Higher values promote smooth, slowly curving pathways which correspond to known tract geometries. Experiments on synthetic data suggest that a good choice for gamma is 24 (see section 3.2).

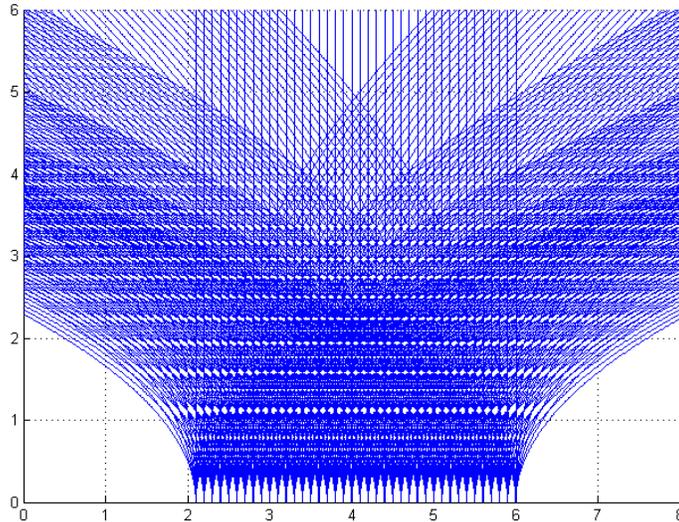
Due to the complexities and potential computational costs of various methods of sampling the continuous PDF given in equation 2 we choose, like Friman, to approximate the continuous PDF with a discrete PDF. By using a sufficiently large number of points spread evenly across the unit sphere and evaluating the PDF on the unit vector defined by each point and the origin, it is trivial to then draw a sample from this discrete PDF. As Friman, we use 2562 directions derived from the vertices on the unit sphere of a fourfold tesellation of an icosahedron.

## 3 Experiments and results

### 3.1 Synthetic experiments

For the experiments on synthetic data, we created a numerical phantom structure mimicking a region of dispersing fibers shown in Figure 2. The blue lines show a subset of the strands which form the structure to illustrate the geometry of the phantom. The region is 8 voxels wide by 6 voxels tall, with the voxel grid

shown behind the phantom. Each of the blue strands was subdivided into line segments of length much smaller than the voxel scale.



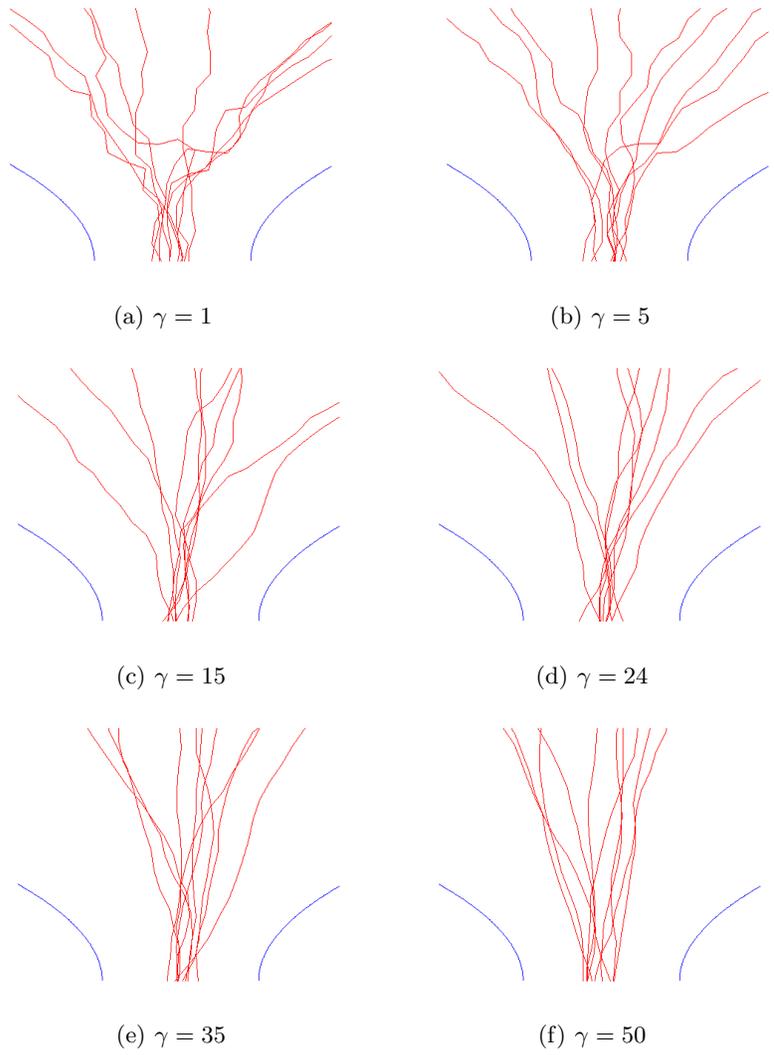
**Fig. 2.** Structure of synthetic dispersing phantom. The blue lines show a subset of the strands forming the structure, for illustration.

### 3.2 Determination of $\gamma$

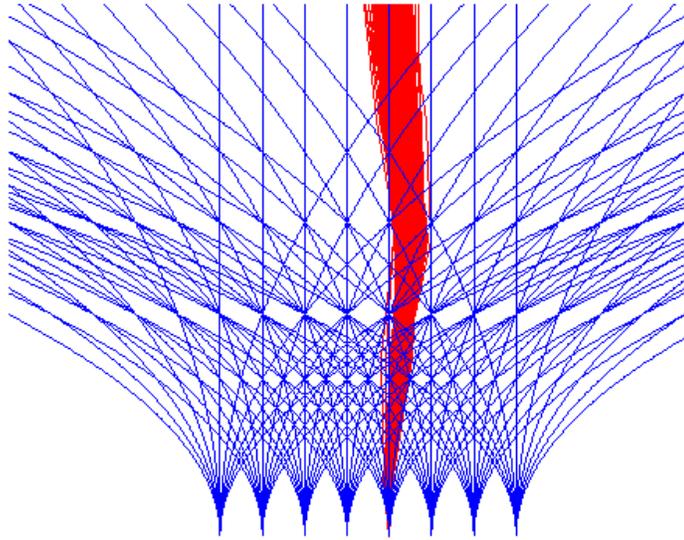
We use the phantom described in section 3.1 to obtain the most appropriate value of  $\gamma$ . ODFs were derived by fitting Bingham distributions directly to the phantom structure in each voxel. We then tracked from a region of the base of the phantom using the algorithm described in section 2.2 with a range of values of  $\gamma$  (Figure 2). Figure 3(a) shows that at low values of  $\gamma$  such as 1, irregular tracks result, however, at a significantly higher value of  $\gamma = 50$  (Figure 3(f)), such a strong prior on curvature can limit the potential trajectories of the tracks, hence limiting exploitation of the dispersive ODFs. Satisfactory results can be achieved for a range of intermediate values. For this demonstration of the algorithm we choose  $\gamma = 24$  (Figure 3(e)), however,  $\gamma$  may reasonably be tuned within a range of values for other applications if necessary.

### 3.3 Tracking on synthetic data

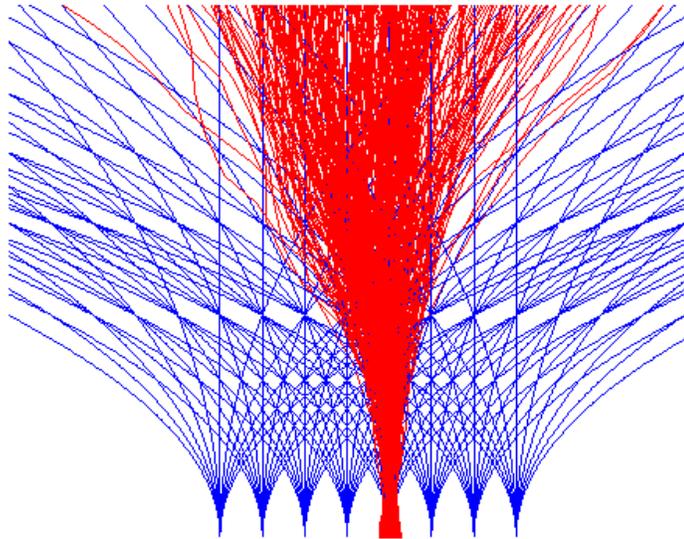
To evaluate the effectiveness of utilising measures of underlying fiber dispersion in tractography, we created a simulated DW-MR dataset based on the phan-



**Fig. 3.** Tracking through a synthetic region of dispersion utilising different values of the constant  $\gamma$  in equation 3. The blue lines represent the extremities of the phantom described in section 3.1.



(a)



(b)

**Fig. 4.** Tractography based on standard PICO tractography techniques (Figure 4(a)) and using the tracking algorithm described in section 2.2 (Figure 4(b)). The Blue lines represent a sparse selection of the underlying fibers of the phantom described in section 3.1. The red lines represent the tracking result.

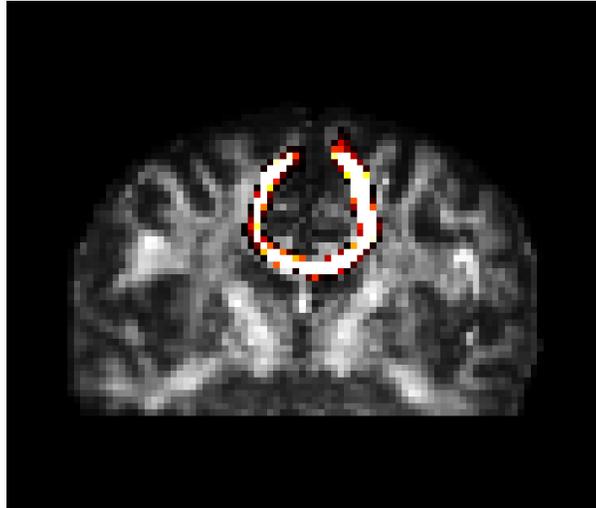
tom described above and tracked using traditional PICO tractography [16] for comparison. We then tracked using the algorithm described in section 2.2, using ODFs derived directly from the phantom structure. The DW-MR signal was simulated using a 30 direction gradient scheme [17] as the basis for a signal simulation using the diffusion tensor model. A diffusion tensor signal with  $d_{\perp} = 3.5101 \times 10^{-4} \text{mm}^2/\text{s}$  and  $d_{\parallel} = 2 \times 10^{-3} \text{mm}^2/\text{s}$  was simulated with a  $b$ -value of  $1000 \text{ s}/\text{mm}^2$  for each line segment of the phantom and signals for all line segments residing in each voxel were summed and normalised. We then fit diffusion tensors to the artificial data derived from the phantom using the open source diffusion MRI toolkit Camino [18]. Bingham distributions were fit to the phantom fiber structure and used to track from a seed point at the base using the algorithm described in section 2.2 (Figure 4(b)). This is compared against standard PICO tracking (Figure 4(a)) using Camino.

### 3.4 *in vivo* data

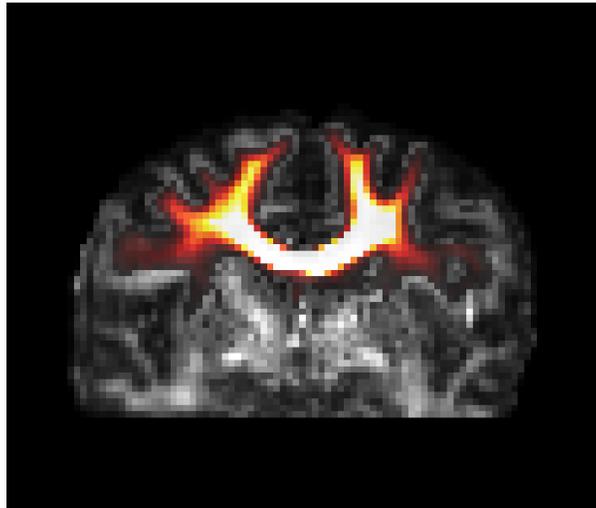
We also test our tracking algorithm on *in vivo* data of a subject. DW-MR images of a healthy male were acquired on a clinical 3T Philips system with isotropic voxels of 2mm, TE=78ms, TR=12.5, with one 30 direction shell and one 60 direction shell with  $b$ -values of  $1000 \text{ s}/\text{mm}^2$  and  $2000 \text{ s}/\text{mm}^2$  respectively. This dataset is the same as that used in [13]. The Camino toolkit was then used to fit the diffusion tensor to the data and perform standard PICO tractography from a single voxel seed in the mid-sagittal corpus callosum. Figure 3 demonstrates the performance of both standard PICO tractography (Figure 5(a)) and the algorithm presented in section 2.2 (Figure 5(b)) tracking from a seed voxel in the mid sagittal corpus callosum. For both tracking examples, 5000 streamlines in total are propagated from a single seed voxel. Tracts are terminated upon entry into a grey matter mask extracted from a T1 weighted image using Freesurfer [19] which is then coregistered to the diffusion weighted image.

### 3.5 Discussion

The experiments on simulated data detailed in section 3.1 show that tracking with traditional methods which ignore underlying fiber dispersion risks a large amount of false negative connections due to the limited exploration of the underlying fiber structure in regions exhibiting fiber dispersion. Figure 4(a) shows the algorithm presented in section 2.2 explores connectivity more thoroughly in such regions in this simple synthetic phantom. Figure 4 shows that including dispersion in tractography allows greater exploration of potential connectivity throughout the peripheral cortex. From a single voxel seed in the mid-sagittal corpus callosum standard PICO tractography streamlines are directed vertically and do not explore the continuum of potential routes through the centrum semiovale (Figure 5(a)), whereas tractography based on the algorithm described in section 2.2 explores connectivity spread throughout the peripheral cortex (Figure 5(b)), the tract density is evenly spread through the intermediate region of



(a)



(b)



**Fig. 5.** Maximum intensity projection map of tractography based on standard pico tractography techniques (Figure 5(a)) and using the tracking algorithm described in section 2.2 (Figure 5(b)) overlaid on FA map.

the centrum semiovale, where histological study verifies there exists fanning fiber structure.

A problem that the algorithm does not currently address is that of fanning polarity. Local estimates of dispersion are symmetric, and hence don't distinguish the direction in which the fibers disperse. There is currently no known method to determine fanning polarity on the voxel scale [20], however recent work by Savadjiev [21] has made progress towards resolving the polarity of a fanning configuration by leveraging local voxel information. In future work, we plan to investigate the inclusion of such methods to further refine tractography including dispersion.

## 4 Conclusion

In this article, we present the first tracking algorithm (to our knowledge) utilizing measures of intra-voxel fiber dispersion to explore the connectome. Including direct measures of intra-voxel fiber dispersion in tractography shows clear advantages in thoroughly exploring potential connections, exploring connectivity commonly missed by current tractography implementations.

## References

1. Kasturi, N., Lichtman, J.W.: The rise of the 'projectome'. *Nature Methods* **4**(4) (2007) 307–308
2. Winston, G.P., Yogarajah, M., Symms, M.R., McEvoy, A.W., Micallef, C., Duncan, J.S.: Diffusion tensor imaging tractography to visualize the relationship of the optic radiation to epileptogenic lesions prior to neurosurgery. *Epilepsia* **52** (2011) 1430–1438
3. Basser, P.J., Mattiello, J., LeBihan, D.: MR Diffusion tensor spectroscopy and imaging. *Biophys J.* **66**(1) (1994) 259–267
4. Tuch, D.S.: Q-Ball Imaging. *Magnetic Resonance in Medicine* **52** (2004) 1358–1372
5. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems* **19** (2003) 1031–1046
6. House, E.I., Panksy, B.: *A Functional Approach to Neuroanatomy*. McGraw Hill (1960)
7. Kreher, B.W., Mader, I., Kiselev, V.G.: Gibbs Tracking: A Novel Approach for the Reconstruction of Neuronal Pathways. *Magnetic Resonance in Medicine* **60** (2008) 953–963
8. Fillard, P., Poupon, C., Mangin, J.F.: A Novel Global Tractography Algorithm Based on an Adaptive Spin Glass Model. *MICCAI* **60** (2009) 927–934
9. Sherbondy, A.J., Dougherty, R.F., Ananthanarayanan, R., Modha, D.S., Wandell, B.A.: Think Global, Act Local; Projectome Estimation with Bluematter. *MICCAI* (2009) 861–868
10. Zhang, H., Hubbard, P.L., Parker, G.J.M., Alexander, D.C.: Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage* **56** (2011) 1301–1315
11. Kaden, E., Knosche, T.R., Anwender, A.: Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *NeuroImage* **37** (2007) 474–488

12. Sotiropoulos, S.N., Behrens, T.E.J., Jbabdi, S.: Ball and rackets: Inferring fiber fanning from diffusion-weighted MRI. *NeuroImage* **60** (2012) 1412–1425
13. Zhang, H., Schneider, T., Wheeler-Kingshott, C., Alexander, D.C.: NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* **61** (2012) 1000–1016
14. Alexander, D.C., Hubbard, P.I., Hall, M.G., Moore, E.A., Ptito, M., Parker, G.J.M., Dyrby, T.B.: Orientationally invariant indices of axon diameter and density from diffusion MRI. *Neuroimage* **52(4)** (2010) 1374–1389
15. Friman, O., Farneback, G., Westin, C.F.: A Bayesian Approach for Stochastic White Matter Tractography. *Transactions on Medical Imaging* **25(8)** (2006) 965–978
16. Parker, J.M., Haroon, H.A., Wheeler-Kingshott, C.A.M.: A Framework for a Streamline-Based Probabilistic Index of Connectivity (PICO) Using a Structural Interpretation of MRI Diffusion Measurements. *Journal of Magnetic Resonance Imaging* **18** (2003) 242–254
17. Jones, D.K., Horsfield, M.A., Simmons, A.: Optimal strategies for measuring diffusion in anisotropic systems by MRI. *Magnetic Resonance in Medicine* **42(3)** (1999) 515–525
18. Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., Alexander, D.C.: Camino: Open-Source Diffusion-MRI Reconstruction and Processing. *ISMRM* (2006) 2759
19. Freesurfer: <http://surfer.nmr.mgh.harvard.edu/>
20. Jbabdi, S., Johansen-Berg, H.: Tractography: Where do we go from here? *Brain Connectivity* **1(3)** (2011) In Press
21. Savadjiev, P., Campbell, J.S.W., Descoteaux, M., Deriche, R., Pike, G.B., Siddiqi, K.: Labeling of ambiguous subvoxel fibre bundle configurations in high angular resolution diffusion MRI. *NeuroImage* **41** (2008) 58–68

# Simultaneous ODF Estimation and Robust Probabilistic Tractography from HARDI

Hasan Ertan Çetingül<sup>1</sup>, Laura Dumont<sup>2</sup>, Mariappan S. Nadar<sup>1</sup>,  
Paul M. Thompson<sup>3</sup>, Guillermo Sapiro<sup>4</sup>, and Christophe Lenglet<sup>4,5</sup>

<sup>1</sup>Siemens Corporation, Corporate Research and Technology, Princeton, NJ, USA

<sup>2</sup>Engineering School of Chemistry, Physics, and Electronics (CPE Lyon), Lyon, France

<sup>3</sup>Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA, USA

<sup>4</sup>Dept. of Electrical & Computer Engineering, U of Minnesota, Minneapolis, MN, USA

<sup>5</sup>Center for Magnetic Resonance Research, U of Minnesota, Minneapolis, MN, USA

{hasan.cetingul,mariappan.nadar}@siemens.com; laura.dumont@cpe.fr;  
thompson@loni.ucla.edu; {guille,clenglet}@umn.edu

**Abstract.** We consider the problem of extracting white matter fiber tracts from high angular resolution diffusion imaging (HARDI) data while simultaneously estimating the local fiber orientation profile. Prior work showed that if parametric mixture models such as mixtures of tensors are used to characterize diffusion, nonlinear state-space filtering techniques (e.g., the unscented Kalman filter (UKF)) can provide a solution to this problem. In this work, we describe how to use the UKF on HARDI data modeled by orientation distribution functions (ODFs), a more generic and versatile diffusion model, by considering the spherical harmonic representation of the HARDI signal as the state. We apply the residual bootstrap method to develop a UKF-based probabilistic tractography method, which has the capability of rejecting outliers, i.e., erroneous trajectories, by measuring the abrupt changes in the state covariance. We demonstrate the robustness of the proposed method by means of experiments on synthetic, phantom, and real HARDI data.

## 1 Introduction

Quantitative characterization of the white matter (WM) circuitry of the human brain is an important problem in neuroradiology, with applications in detecting, diagnosing, and tracking the progression of neurological diseases. Diffusion magnetic resonance imaging (DMRI) can describe the WM circuitry by capturing variations in water diffusion patterns. More specifically, DMRI quantifies Brownian motion of water molecules in biological tissues and produces images that are weighted according to the amount (i.e., anisotropy) and direction of diffusion. The anisotropy arises from the presence of axonal membranes and myelin, thereby allowing the mapping of neural fiber tracts in WM. State-of-the-art DMRI techniques such as high angular resolution diffusion imaging (HARDI) enables the reconstruction of the orientation distribution function (ODF) [1–3], which offers improved accuracy in resolving intra-voxel complexities over the diffusion tensor (DT) model, currently the de facto standard for clinical applications.

Conventional WM tractography methods consider a volume of diffusion models (DTs, ODFs, etc.) estimated from the HARDI signal as input, and extract trajectories of the underlying fiber tracts using deterministic or probabilistic approaches (see [4–8] and references therein). Contrary to these techniques, which assume estimation of diffusion models and tractography to be two problems independent of each other, Malcolm et al. [9] considered tracking as a causal process (a common assumption in deterministic approaches) and proposed to simultaneously solve the aforementioned problems using an unscented Kalman filter (UKF). This method assumes a diffusion model with a fixed number of mixtures of diffusion tensors. In other words, the parameters of a multi-tensor model form the state. The UKF recursively updates the state, provides the covariance of the estimate, and determines the most appropriate direction to follow. In theory, this strategy is fast (since it avoids estimation at every voxel) and robust to noise (due to the nature of filtering). However, since the state vector is the concatenation of several elements (e.g., principal eigenvectors/eigenvalues and mixture weights) with different statistical properties, parameter selection and preservation of physically meaningful states during the evolution are vital for accuracy and robustness in tracking. This problem is addressed in [10] using an intrinsic UKF framework in which the states (based on the two-tensor model) are guaranteed to be in the space of DTs, i.e.,  $3 \times 3$  symmetric positive definite matrices. Nevertheless, to the authors’ knowledge, *mixture models* are an integral part of these methods and *reproducibility* of tracking results remains unanswered.

In this work, we present an UKF-based probabilistic tractography framework for HARDI data modeled by ODFs. Our approach considers the spherical harmonic representation of the HARDI signal as the state and enforces nonnegativity of the reconstructed ODFs during the state transition [11]. The mode detection scheme proposed in [12] is used to identify the directions at which the ODFs attain their peaks. We apply the *residual bootstrap* method [8] to develop the probabilistic framework, which computes a large number of trajectories from each seed point by propagating streamlines along the modes that are most consistent with the previous direction. Our method produces *likelihood maps* that can be further improved by rejecting outliers, i.e., erroneous trajectories, by measuring changes in the state covariance. We quantify and depict the uncertainty (equivalently the robustness or reproducibility) in tracking through crossing fibers in synthetic and phantom data, as well as WM fiber tracts in a real HARDI dataset.

## 2 Diffusion Representation using Spherical Harmonics

In a typical DMRI protocol, one acquires images of the MR signal attenuation  $\{S_n \doteq S(\theta_n, \phi_n)\}$  along  $N$  different gradient directions  $\{(\theta_n, \phi_n)\}_{n=1}^N$  on the 2-sphere  $\mathbb{S}^2$ , as well as baseline images  $S_0$  with no diffusion sensitization.<sup>1</sup> To represent diffusion at a specific spatial location, we consider the single shell q-ball imaging formulation within constant solid angle [3]. According to this formulation, the

<sup>1</sup>  $\theta_n \in [0, \pi]$  and  $\phi_n \in [0, 2\pi)$  are the polar and azimuthal angles, respectively, representing the direction  $(\theta_n, \phi_n)$  in spherical coordinates.

value of the ODF in the direction  $(\vartheta, \varphi)$  (equivalent to  $(\theta_n, \phi_n)$  in the transformed space) is given by  $p(\vartheta, \varphi) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \text{FRT}\{\nabla_b^2 \ln(-\ln(S_n/S_0))\}$ , where  $\nabla_b^2$  is the Laplace-Beltrami operator on  $\mathbb{S}^2$  and FRT is the Funk-Radon transform. Since  $S_n$  is in practice real and symmetric, the real spherical harmonics (SHs)  $Y_j(\theta_n, \phi_n)$  [13] are used to approximate the signal as  $\ln(-\ln(S_n/S_0)) \approx \sum_{j=1}^J c_j Y_j(\theta_n, \phi_n)$ , where  $c_j \in \mathbb{R}$  is the SH coefficient associated with the  $j$ -th basis function.<sup>2</sup> The SH expansion of the signals  $\{\ln(-\ln(S_n/S_0))\}_{n=1}^N$  can also be written in vector form as  $\mathbf{s} \approx B\mathbf{c}$ , where  $\mathbf{s} = [\ln(-\ln(S_1/S_0)), \dots, \ln(-\ln(S_N/S_0))]^\top \in \mathbb{R}^N$  and  $[Y_1(\theta_n, \phi_n), \dots, Y_J(\theta_n, \phi_n)]$  is the  $n$ -th row of  $B \in \mathbb{R}^{N \times J}$ . Once the least-squares solution of  $\mathbf{s} \approx B\mathbf{c}$  for  $\mathbf{c} \in \mathbb{R}^J$  is found, a discrete representation of the ODF can be reconstructed using the fact that SHs are eigenfunctions of  $\nabla_b^2$  and FRT [3, 14]. The reconstruction method in [14] also enforces the ODFs  $p$  to be nonnegative and it is used, in this work, to initialize the UKF.

### 3 The UKF for Simultaneous Estimation and Tracking

The main idea in the UKF-based tractography methods [9, 10, 15] is to estimate and recursively update the diffusion model given the measured signal at a spatial location and trace the fiber by propagating in the direction consistent with the previous directions of propagation. The four components of the UKF are the *state*  $\mathbf{x}$  (the diffusion model), *state transition function*  $f$  (to predict how the diffusion model changes), *observation function*  $h$  (to predict the observed signal for a particular state), and *measurement*  $\mathbf{y}$  (the acquired HARDI signal). In our framework, we consider the SH coefficient vector  $\mathbf{c} = [c_1, \dots, c_J]^\top$  as the state  $\mathbf{x}$  and identity dynamics for the state transition  $f$  assuming that the diffusion profile does not change drastically in the vicinity of the location of interest. The measurement  $\mathbf{y}$  is the signals  $\{S_n/S_0\}$  at that location, and the observation  $h$  is the reconstruction of the signals as  $\{\exp(-\exp(\sum_j c_j Y_j(\theta_n, \phi_n)))\}$ .

#### 3.1 Review of the UKF Formulation and the UKF-ODF Algorithm

The UKF uses a sampling technique known as the unscented transform (UT), which calculates the statistics of a random variable undergoing a nonlinear transformation. The UKF uses the state transition model to predict the next state (and observation) via the UT and the measurement to update the state estimate. One iteration of this recursive algorithm can be outlined as follows.

Let  $\mathbf{x}_k \in \mathbb{R}^J$  and  $P_k \in \mathbb{R}^{J \times J}$  be the estimated mean and covariance matrix of the current state at discrete time  $k$ . In the prediction stage, a set  $\mathcal{X}_k = \{\chi_i\}_{i=1}^{2J+1} \subset \mathbb{R}^J$  of  $2J+1$  sigma points with associated weights  $\{w_i\} \subset \mathbb{R}^+$  is generated by perturbing  $\mathbf{x}_k$  such that  $\chi_0 = \mathbf{x}_k$  with weight  $w_0 = \kappa/(J+\kappa)$ , and  $\chi_i = \mathbf{x}_k + [\sqrt{(J+\kappa)P_k}]_i$ ,  $\chi_{i+J} = \mathbf{x}_k - [\sqrt{(J+\kappa)P_k}]_i$  with weights  $w_i = w_{i+J} = 1/(2J+2\kappa)$  for  $1 \leq i \leq J$ .<sup>3</sup> Next, the predicted sample set of states

<sup>2</sup> We consider the SH basis of degree  $L=4$  and hence  $J=(L+1)(L+2)/2 = 15$ .

<sup>3</sup>  $[A]_i$  denotes the  $i$ -th column of the matrix  $A$  and  $\kappa$  is a scaling parameter set to 0.01.

$\mathcal{X}_{k+1|k}$  is obtained by propagating  $\mathcal{X}_k$  through the state transition function, i.e.,  $\mathcal{X}_{k+1|k} = \{f(\chi_i)\} = \{\hat{\chi}_i\}$ . These sample states are used to calculate the mean  $\hat{\mathbf{x}}_{k+1|k}$  of the predicted system state as well as its covariance  $P_{xx}$ , which is corrupted by the process noise with covariance  $Q_{\mathbf{c}} \in \mathbb{R}^{J \times J}$ . Then the predicted set of observations is obtained as  $\mathcal{Y}_{k+1|k} = \{h(\hat{\chi}_i)\} = \{\gamma_i\}$ . These observations are used to calculate the mean of the predicted measurement  $\hat{\mathbf{y}}_{k+1|k}$  as well as its covariance  $P_{yy}$ , which is corrupted by the measurement noise with covariance  $R_s \in \mathbb{R}^{N \times N}$ . The Kalman gain  $K = P_{xy}P_{yy}^{-1} \in \mathbb{R}^{J \times N}$  is finally used to correct our prediction and obtain the mean and covariance of the updated state as  $\mathbf{x}_{k+1} = \hat{\mathbf{x}}_{k+1|k} + K(\mathbf{y}_t - \hat{\mathbf{y}}_{k+1|k})$  and  $P_{k+1} = P_{xx} - KP_{yy}K^\top$ , respectively. The reader is referred to seminal papers such as [16] for further details on the UKF.

**The Algorithm** The UKF-ODF algorithm is initialized, at user-specified seed points, by estimating the SH representation of the HARDI signal and reconstructing the ODFs via the constrained optimization strategy in [14], which enforces nonnegativity of the ODFs. For each seed point on the fiber of interest, the SH coefficient vector  $\mathbf{c}$  is taken as the initial state  $\mathbf{x}_0$  and the number and locations of the mode(s) of the ODF are identified using the weighted spherical mean shift clustering [12]. These modes represent the directions to be followed at the current spatial location. At the  $k$ -th iteration, we predict the new state as  $\hat{\mathbf{x}}_{k+1|k} = f(\mathbf{x}_k) = \mathbf{x}_k$ , use the set of predicted observations along with the measured signal interpolated at the current spatial location, and find the new state  $\mathbf{x}_{k+1}$ .<sup>4</sup> This estimate is used to reconstruct the ODF whose modes (directions to be followed at  $k + 1$ ) are subsequently identified. The UKF-ODF algorithm produces one streamline trajectory by propagating forward in the direction consistent with the previous direction until a user-defined termination criterion (e.g., high curvature, low generalized fractional anisotropy (GFA) [1], etc.) is met.

### 3.2 Probabilistic Tractography using the Residual Bootstrap

We develop an UKF-based probabilistic tractography framework, which we call as **rbUKF-ODF**, by integrating the UKF-ODF algorithm into the *residual bootstrap* method [8]. Through residual bootstrapping, one can consider a trajectory given by UKF-ODF as an outcome of a probabilistic experiment and use it to construct the *likelihood maps*, which represent the probability that a track (or trajectory) emanating from a seed point passes through other spatial locations.

Consider the SH expansion of the HARDI signal in §2, which can be written in vector form as  $\mathbf{s} = B\mathbf{c} + \boldsymbol{\epsilon}$  with  $\boldsymbol{\epsilon}$  being the noise vector. The signal  $\hat{\mathbf{s}}$  predicted by the least-squares fit to the measured signal  $\mathbf{s}$  is given as

$$\hat{\mathbf{s}} = B\hat{\mathbf{c}} = B(B^\top B)^{-1}B^\top \mathbf{s} \doteq H\mathbf{s}. \quad (1)$$

Now consider the residual vector  $\hat{\boldsymbol{\epsilon}} = [\hat{\epsilon}_1, \dots, \hat{\epsilon}_N]^\top = \mathbf{s} - \hat{\mathbf{s}}$  and the raw residuals corrected for leverage as

$$\hat{\epsilon}_i^c = \hat{\epsilon}_i / \sqrt{1 - H_{ii}}, \quad (2)$$

<sup>4</sup>  $Q_{\mathbf{c}}$  and  $R_s$  are assumed diagonal with entries  $q_{\mathbf{c}} \in [0.05, 0.1]$  and  $r_s \approx 0.02$ , which are selected according to the spectra of the SH coefficients and the HARDI signal.

where  $H_{ii}$  is the  $i$ -th diagonal entry of  $H$  in (1). In the residual bootstrap method, the entries of the vector  $\hat{\epsilon}^c$  are randomly chosen with replacement to form a new bootstrapped residual  $\hat{\epsilon}^*$ . Then, a *synthetic bootstrap realization*  $\hat{\mathbf{s}}^*$  of  $\mathbf{s}$  can be generated as  $\hat{\mathbf{s}}^* = \hat{\mathbf{s}} + \hat{\epsilon}^*$ . By generating  $N_b$  bootstrap realizations from the measured signal  $\mathbf{s}$  and running the UKF-ODF algorithm on each realization from a seed point of interest, one can produce  $N_b$  trajectories emanating from that seed point. The resulting trajectories are used to obtain the likelihood map of the region of interest by assigning to each voxel the ratio of the number of trajectories passing through that voxel to  $N_b$ .

**Inferring Track Uncertainty from the UKF** It is worth noting that the covariance  $P_k$  provides a measure of confidence in the  $k$ -th estimation. In particular, the Frobenius norm of  $P_k$  can be considered as a scalar measure of uncertainty at  $k$  [17]. This information can be used to eliminate outliers, i.e., erroneous trajectories, and improve the quality of the likelihood maps. More specifically, one can eliminate a trajectory (or portion of a trajectory) before computing the likelihood map if there is a drastic increase in the Frobenius norm of the covariance  $\{P_k\}$ , i.e., if

$$\|P_{k+1}\|_F - \|P_k\|_F > \tau, \quad (3)$$

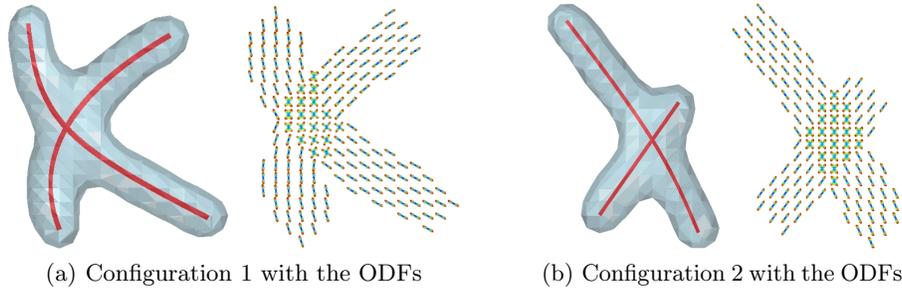
for some empirically-tuned threshold  $\tau > 0$ . [17] uses the moving average of the Frobenius norm of the covariance as a measure of track uncertainty to detect erroneous tracks, whereas, in this work, we propose to analyze the sequence  $(\|P_0\|_F, \|P_1\|_F, \|P_2\|_F, \dots)$  to eliminate the erroneous portions of the trajectories.

## 4 Method Validation

### 4.1 Experiments on Synthetic Data

Experiments on synthetic data evaluate the sensitivity of the **rbUKF-ODF** method to noise (without outlier rejection) and compare its performance to that of the filtered 2-tensor tractography method (with parameters given in [15]) extended into a probabilistic one (**rbUKF-2T**) using the residual bootstrapping method. We simulate diffusion weighted images of two randomly generated configurations, each of which contains two intersecting fiber bundles (see Figs. 1(a)-1(b)). The centerline of a fiber bundle is formed by fitting cubic splines through three randomly selected points in a  $30 \times 30$  lattice. We use the two-tensor model in [13], where the HARDI signals  $\{S_n\}_{n=1}^N$  at  $N = 81$  gradient directions, with  $S_0 = 1$  and  $b = 2,000$  s/mm<sup>2</sup>, are simulated to represent an isotropic background and 1- or 2-fiber ODFs according to the shape of a fiber centerline. Noisy signals are generated by adding Rician (complex Gaussian) noise with zero mean and standard deviation  $\sigma = S_0/\zeta$ , where  $\zeta$  is the signal-to-noise ratio (SNR).

Figs. 2(a) and 2(b) show the likelihood maps of both configurations, obtained using **rbUKF-2T** (top rows) and **rbUKF-ODF** (bottom rows) at different levels of SNR (from left to right: 30, 20, 10, 5 dB). Initiated at two different seed points (shown in yellow), the UKF-based tractography is performed on  $N_b = 1,000$  bootstrapped realizations of the signal. Superimposed on the GFA maps of the



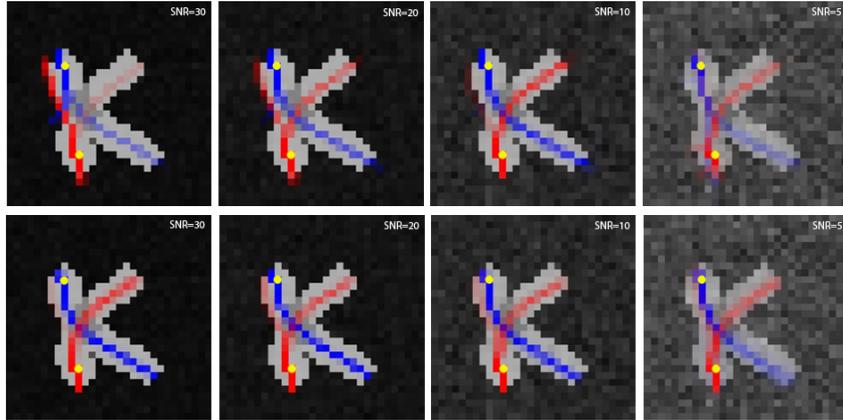
**Fig. 1.** Two synthetic configurations with fiber centerlines shown in red and the corresponding ODFs reconstructed for visualization purposes using the method in [14].

synthetic configurations at the corresponding SNR level, the probability of connectivity is depicted using the colors red and blue for the two fiber bundles in the configurations (i.e., dark red/blue~high probability, light red/blue~low probability). It is observed that **rbUKF-ODF** produces a higher likelihood of connectivity along the true fiber centerlines and less dispersion than **rbUKF-2T** does. More specifically, for configuration 1, **rbUKF-ODF** produces considerably less dispersion (i.e., less false negatives) when SNR = 30, 20 dB. Yet the performance of the methods is comparable when SNR = 5 dB, and **rbUKF-2T** slightly outperforms **rbUKF-ODF** when SNR = 10 dB. For configuration 2, **rbUKF-ODF** greatly outperforms **rbUKF-2T** at all levels of SNR, with trajectories almost flawlessly following the fiber centerlines when SNR = 30, 20, 10 dB. We observe that for a considerable number of realizations, the two-tensor model does not accurately resolve the crossing configuration (especially when SNR = 30, 20 dB), and hence **UKF-2T** produces erroneous trajectories. These results demonstrate that by removing the dependency of the UKF on a more restrictive diffusion model, it is possible to resolve higher degrees of complexity and obtain more accurate likelihood maps.

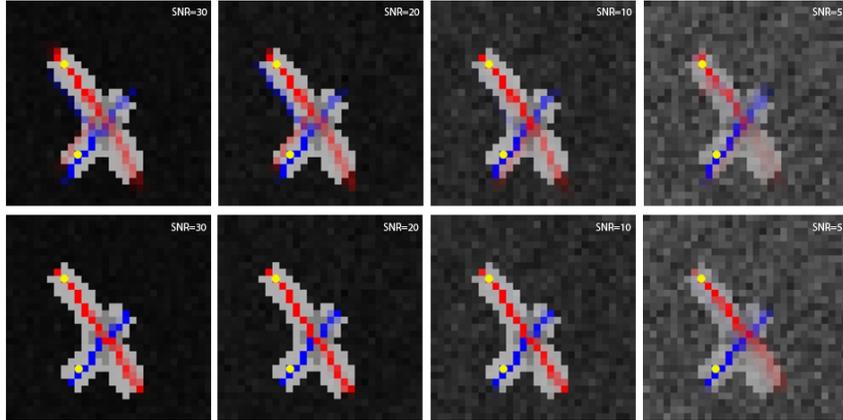
## 4.2 Experiments on Phantom Data

Next, we test the proposed method on the biological phantom in [18], constructed from excised rat spinal cords and designed to have crossing fiber bundles (see Fig. 3(a)). The diffusion weighted images of the phantom were acquired with a  $40 \times 9$  image matrix (40 slices with an isotropic spatial resolution of 2.5 mm) and a diffusion sensitization at  $b = 1,300$  s/mm<sup>2</sup> applied along a set of 90 gradient directions with 10 baseline images. For visualization purposes, we reconstruct the ODFs at  $y=4$  slice (see Fig. 3(b)) and show the underlying orientation profiles.

Fig. 3(c) shows the color-coded trajectories obtained using **UKF-ODF** (initiated at two different seeding regions marked in yellow) from the bootstrapped realizations of the HARDI signal. It is observed that **UKF-ODF** is successful in resolving the crossing configuration and delineating the centerlines of the fibers. To depict the track uncertainties, we plot the same trajectories in Fig. 3(d) using  $\{\|P_k\|_F\}$  as the colormap. One can clearly see in the zoomed regions that when a trajectory



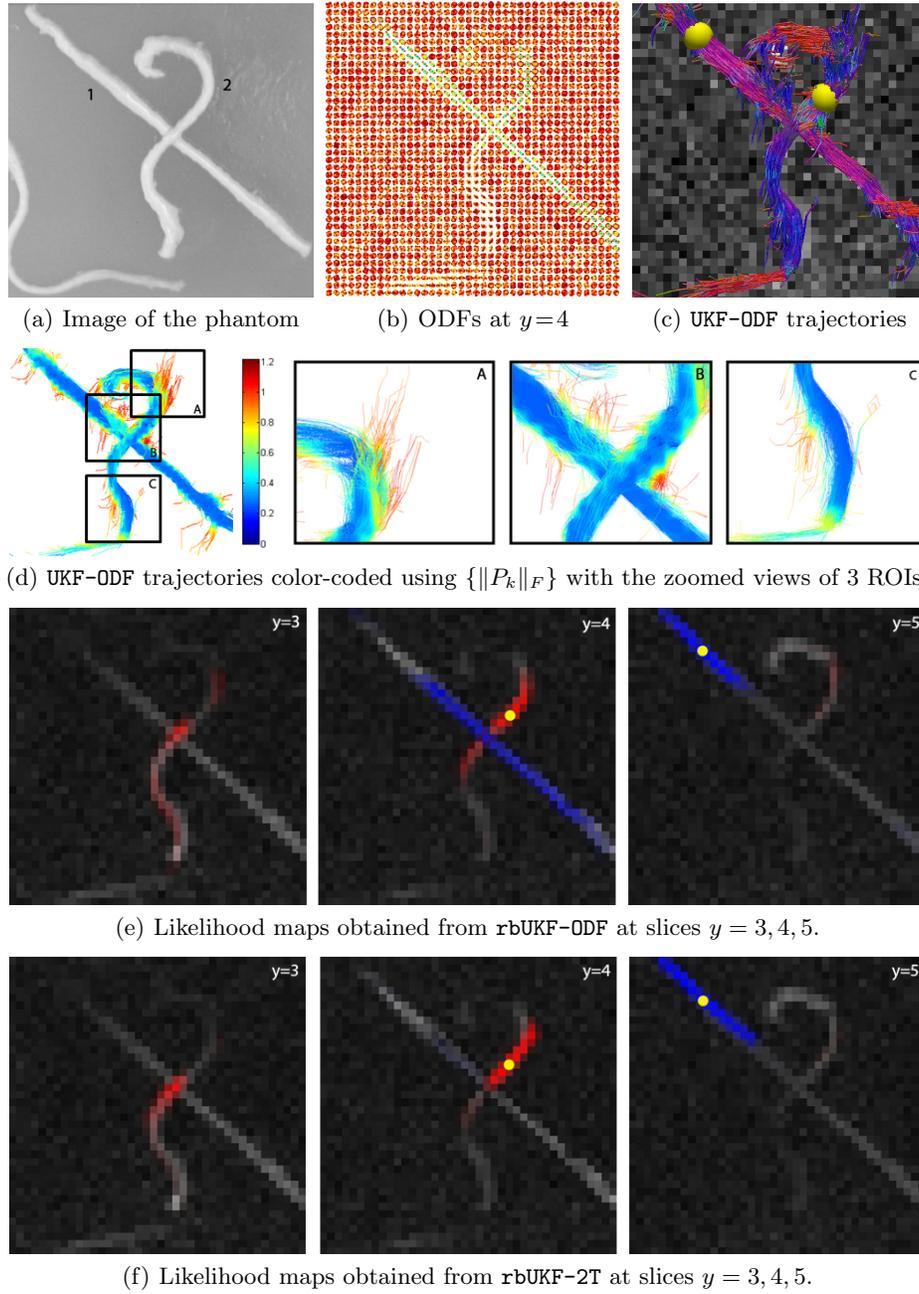
(a) Likelihood maps for synthetic configuration 1



(b) Likelihood maps for synthetic configuration 2

**Fig. 2.** Likelihood maps (superimposed on the GFA maps) of the configurations computed using `rbUKF-2T` (top two rows) and `rbUKF-ODF` (bottom two rows), without the covariance-based outlier rejection, at different levels of SNR. The probability of connectivity is depicted using the colors blue and red for the fibers (dark~high, light~low).

deviates from the fiber centerline, the norm of the covariance increases suggesting that confidence to that portion of the trajectory decreases. Superimposed on the GFA maps, the likelihood maps of fiber 1 (resp. fiber 2) are obtained using `rbUKF-ODF` (with  $N_b = 1,000$ ) and shown in blue (resp. red) in Fig. 3(e) following the same visualization rule in §4.1. We observe that `rbUKF-ODF` produces likelihood maps that are in accordance with the fiber centerlines. In particular, the small thickness of the fibers does not adversely affect the performance of UKF-ODF and the intra-voxel complexity in the crossing region is resolved by the ODF model. Thus, the resulting likelihood maps are more reliable than the maps produced by `rbUKF-2T`, which do not fully delineate fiber 1 (see Fig. 3(f)).



**Fig. 3.** (a-b) Image of the phantom from [18] and the reconstructed ODFs at  $y=4$  slice, (c) Trajectories obtained using UKF-ODF, (d) Trajectories color-coded using the values of  $\|P_k\|_F$ , (e-f) Likelihood maps (superimposed on the GFA maps) computed using **rbUKF-ODF** and **rbUKF-2T** at  $y = 3, 4, 5$ , and depicted using the colormap in Fig. 2.

### 4.3 Experiments on Real Data

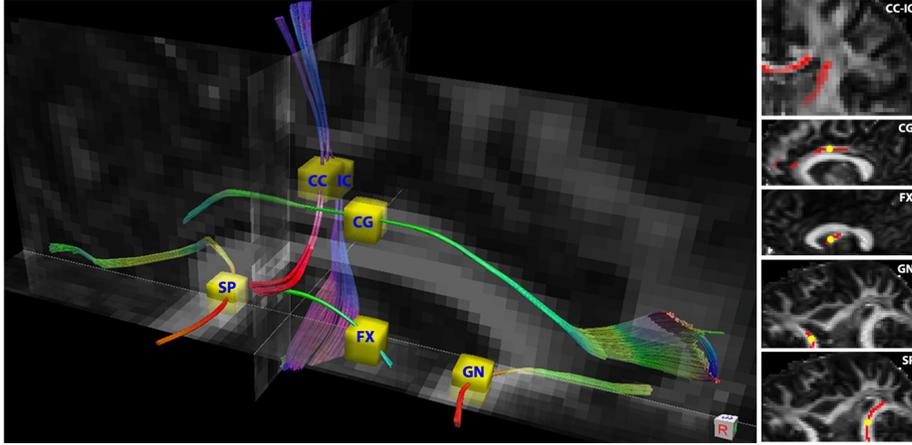
We finally test the **rbUKF-ODF** algorithm on a human brain dataset containing structural, functional, and diffusion MR images provided for the Pittsburgh Brain Connectivity Challenge (PBCC Spring 2009).<sup>5</sup> The diffusion weighted images were acquired with a  $128 \times 128$  image matrix, (68 slices with an isotropic spatial resolution of 2 mm), and a diffusion sensitization at  $b = 1,500 \text{ s/mm}^2$  applied along a set of 256 gradient directions with 29 baseline images. In the experiments, we consider a volume of size  $40 \times 71 \times 30$  containing the left hemisphere of the brain. We apply **rbUKF-ODF** and **rbUKF-2T** (with  $N_b = 1,000$ ) on several white matter fiber tracts such as the body, splenium (SP), and genu (GN) of the corpus callosum (CC), the internal capsule (IC), the cingulum (CG), and the fornix (FX). One seed point per tract is selected in the CG, FX, GN, and SP, whereas two seed points per tract are selected for the CC and IC to better illustrate how successfully the two diffusion models resolve intravoxel fiber complexities (i.e., partial volume effects) between these two fiber tracts.

Fig. 4 shows not only the trajectories computed from the bootstrapped realizations of the HARDI signal using **UKF-2T** and **UKF-ODF**, but also the resulting likelihood maps depicted on selected axial, coronal, and sagittal slices. First, we observe that both the two-tensor and ODF models successfully resolve intravoxel complexities in the region (shown in the coronal slices at the top right of Figs. 4(a)-4(b)) where the CC and IC meet. The trajectories generated by **UKF-2T** and **UKF-ODF** are dense and in accordance with the neuroanatomy. The dispersion in the bottom portion of the trajectories (generated by **UKF-ODF**) for the IC is due to the slightly low GFA threshold we selected for stopping. At this point, it is worth noting that dispersion due to the residual bootstrap method is not to be confused with anatomical dispersion. Bootstrap dispersion is due to noise (refer to §3.2) and should be viewed as a measure of *robustness* or *reproducibility* for the tractography algorithm under study [8]. From this perspective, in the analysis of the CG, the reproducibility of the results of **rbUKF-ODF** is higher than that of **rbUKF-2T** as **UKF-ODF** produces trajectories almost solely in the CG bundle. More specifically, the trajectories produced by **UKF-2T** for the CG show large bootstrap dispersion (i.e., bundles of trajectories that “bifurcate” around the region indicated by one of the white arrows in Fig. 4(b)) towards the corpus callosum. **UKF-ODF**, on the other hand, is able to recover the anterior and posterior loops of the CG, which are usually difficult to extract due to high curvature. The dispersion in the anterior portion of the trajectories (generated by **UKF-ODF**) is due to low GFA threshold for stopping, similar to the case for the internal capsule.

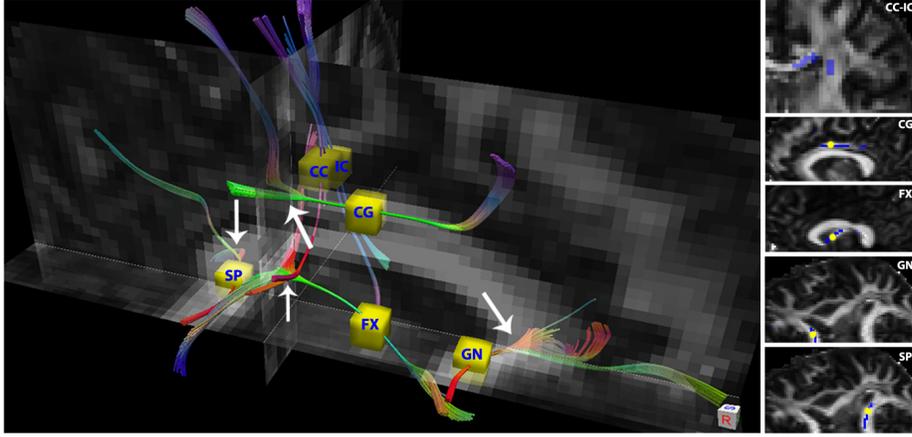
We reach a similar conclusion for the other WM neural fiber tracts. In particular, since the trajectories (generated by **UKF-ODF**) emanating from the middle portion of the FX show no severe bootstrap dispersion, **rbUKF-ODF** can be viewed as an algorithm that is more robust than **rbUKF-2T**, which produces likelihood maps with larger dispersion (due to the region indicated by the

<sup>5</sup> [pub.lrdc.pitt.edu/?q=2009a-home](http://pub.lrdc.pitt.edu/?q=2009a-home)

corresponding white arrow). For the fourth and fifth WM fiber tracts of interest (i.e., the GN and SP, respectively), **rbUKF-2T** presents large dispersion in the forceps minor (resp. forceps major) and some spurious connections, whereas **rbUKF-ODF** generates considerably less bootstrap dispersion and hence it is more robust in delineating these fiber tracts.



(a) Delineating WM neural fiber tracts using **rbUKF-ODF**.



(b) Delineating WM neural fiber tracts using **rbUKF-2T**.

**Fig. 4.** Trajectories emanating from the seed points shown in yellow (larger images on the left) and the corresponding likelihood maps superimposed on the axial, coronal, and sagittal GFA maps (five images in the right columns). The regions with “large” bootstrap dispersion (e.g., bundles of trajectories that “bifurcate”) are indicated by the white arrows. In the likelihood maps, the probability of connectivity is depicted using the colors blue and red for **rbUKF-2T** and **rbUKF-ODF**, respectively.

## 5 Discussions and Conclusions

In this work, we first explain how to use the UKF-based tractography techniques, which were shown to be successful for simultaneous estimation (of diffusion models) and tractography, for HARDI data characterized by ODFs. The resulting algorithm, which we refer to as **UKF-ODF**, uses the SH representation of the HARDI signal as the state, guarantees that the reconstructed ODFs are nonnegative, and produces a trajectory by considering the modes of the ODFs as candidate directions to follow [11]. We propose to apply the residual bootstrap method to develop a UKF-based probabilistic tractography algorithm called **rbUKF-ODF**, which has the capability of rejecting erroneous trajectories by measuring changes in the state covariance. Experiments on synthetic, phantom, and real data demonstrate the advantages of our method over its two-tensor based version (**rbUKF-2T**) in terms of accuracy in tract delineation and dispersion in likelihood maps under noisy conditions. In addition, even though **UKF-ODF** uses a higher-dimensional state vector and an iterative ODF mode detection scheme, we observe that **UKF-ODF** and **UKF-2T** have comparable computation times.

However, we should also mention that, in some cases, increased accuracy and reliability in tracking may come with an undesirable outcome as **UKF-ODF** is prone to producing slightly less smooth streamline trajectories. We believe that generation of sigma points as well as perturbation of the predicted states and measurements play a critical role in obtaining smooth trajectories. Accordingly, future work includes investigating other state space representations that will formalize the UKF into an intrinsic (manifold-constrained) formulation [19] on the Riemannian manifold of ODFs, as discussed in [10] for the space of diffusion tensors. In particular, we will investigate ideas from [20, 21] to perturb the states (i.e., spherical harmonic coefficients) in an intrinsic manner. We will also consider using a faster mode detection strategy that directly operates on the spherical harmonic coefficients of the ODFs to further reduce computation time.

**Acknowledgments:** This work was supported in part by NIH grants R01 EB008432, P41 RR008079 (NCCR), P41 EB015894 (NIBIB), P30 NS057091, P30 NS0576408, and the Human Connectome Project (U54 MH091657) from the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research. L. Dumont contributed to this work while she was at Siemens Corporation, Corporate Research and Technology, Princeton, NJ, USA.

## References

1. Tuch, D.: Q-ball imaging. *Magnetic Resonance in Medicine* **52**(6) (2004) 1358–1372
2. Tristan-Vega, A., Westin, C.F., Aja-Fernandez, S.: Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. *NeuroImage* **47**(2) (2009) 638–650
3. 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. *Magnetic Resonance in Medicine* **64**(2) (2010) 554–566

4. Friman, O., Farnebäck, G., Westin, C.F.: A Bayesian approach for stochastic white matter tractography. *IEEE Transactions on Medical Imaging* **25**(8) (2006) 965–978
5. Behrens, T., Berg, H.J., Jbabdi, S., Rushworth, M., Woolrich, M.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* **34**(1) (2007) 144–155
6. Zhang, F., Hancock, E., Goodlett, C., Gerig, G.: Probabilistic white matter fiber tracking using particle filtering and von Mises-Fisher sampling. *Medical Image Analysis* **13** (2009) 5–18
7. Descoteaux, M., Deriche, R., Knösche, T., Anwander, A.: Deterministic and probabilistic tractography based on complex fiber orientation distributions. *IEEE Transactions on Medical Imaging* **28**(2) (Feb. 2009) 269–286
8. Jeurissen, B., Leemans, A., Jones, D., Tournier, J.D., Sijbers, J.: Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. *Human Brain Mapping* **32** (2011) 461–479
9. Malcolm, J., Shenton, M., Rathi, Y.: Filtered multitensor tractography. *IEEE Transactions on Medical Imaging* **29**(9) (2010) 1664–1675
10. Cheng, G., Salehian, H., Hwang, M., Howland, D., Forder, J., Vemuri, B.: A novel unscented Kalman filter for tractography from HARDI. In: *IEEE International Symposium on Biomedical Imaging*. (2012) 534–537
11. Çetingül, H., Nadar, M., Thompson, P., Sapiro, G., Lenglet, C.: Simultaneous ODF estimation and tractography in HARDI. In: *International Conference of the IEEE Engineering in Medicine and Biology Society*. (2012)
12. Çetingül, H., Vidal, R.: Intrinsic mean shift for clustering on Stiefel and Grassmann manifolds. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2009) 1896–1902
13. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine* **58**(3) (2007) 497–510
14. Goh, A., Lenglet, C., Thompson, P., Vidal, R.: Estimating orientation distribution functions with probability density constraints and spatial regularity. In: *Medical Image Computing and Computer Assisted Intervention*. (2009) 877–885
15. Lienhard, S., Malcolm, J., Westin, C.F., Rathi, Y.: A full bi-tensor neural tractography algorithm using the unscented Kalman filter. *EURASIP Journal of Advances in Signal Processing* **77** (2011)
16. Wan, E., van der Merwe, R.: The unscented Kalman filter for nonlinear estimation. In: *IEEE Symposium on Adaptive Systems for Signal Processing, Communications and Control*. (2000) 153–158
17. Rathi, Y., Malcolm, J., Bouix, S., Westin, C.F., Shenton, M.: False positive detection using filtered tractography. In: *International Symposium on Magnetic Resonance in Medicine*. (2010) 4019
18. Campbell, J.S., Siddiqi, K., Rymar, V.V., Sadikot, A.F., Pike, G.B.: Flow-based fiber tracking with diffusion tensor and q-ball data: Validation and comparison to principal diffusion direction techniques. *NeuroImage* **27**(4) (2005) 725–736
19. Sipos, B.: Application of the manifold-constrained unscented Kalman filter. In: *IEEE/ION Position, Location and Navigation Symposium*. (2008) 30–43
20. Goh, A., Lenglet, C., Thompson, P., Vidal, R.: A nonparametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. *NeuroImage* **56**(1) (2011) 1181–1201
21. Çetingül, H., Afsari, B., Wright, M., Thompson, P., Vidal, R.: Group action induced averaging for HARDI processing. In: *IEEE International Symposium on Biomedical Imaging*. (2012) 1389–1392

# A Unified Tractography Framework for Comparing Diffusion Models on Clinical Scans

Christian Baumgartner<sup>1</sup>, O. Michailovich<sup>2</sup>, J. Levitt<sup>1</sup>, O. Pasternak<sup>1</sup>, S. Bouix<sup>1</sup>, C-F Westin<sup>1</sup> and Yogesh Rathi<sup>1</sup>

<sup>1</sup> Brigham and Women’s Hospital, Harvard Medical School, Boston,

<sup>2</sup> Department of Electrical Engg., University of Waterloo, Canada

**Abstract.** In this paper, we compare several parametric and non-parametric models of diffusion in a unified framework that allows simultaneous model estimation and tractography. The framework uses the Unscented Kalman Filter (UKF) to compare several variants of spherical harmonics (SH), i.e., SH with sharpening, spherical deconvolution, SH with solid angle and also several parametric models like single and two-tensor models with and without an additional “free water” component. We estimate all these models and perform tractography using the same optimizer, namely, the UKF. Comparison is done by tracing two fiber bundles whose connectivity is known from the literature on human anatomy. We trace these fiber bundles on 10 healthy subjects and compare how well each of these models perform on clinical in-vivo scans. For quantitative comparison, we propose two new measures; traceability and coverage, both of which capture how well each method performs in terms of tracing known fiber bundles. Our results show that, the two-tensor with “free-water” model performs very well for both these measures.

## 1 Introduction

Diffusion magnetic resonance imaging (dMRI) allows neuroscientists to investigate how neurons originating from one region connect to other regions, or how well-defined these connections may be. For such studies, the quality of the results relies heavily on the chosen fiber representation and the method of reconstructing pathways.

To begin studying the microstructure of fibers, we need a model to interpret the diffusion weighted signal. Such models fall broadly into two categories: parametric and nonparametric. One of the simplest parametric models is the diffusion tensor which describes a Gaussian estimate of the diffusion orientation and strength at each voxel. While robust, this model can be inadequate in cases of mixed fiber presence or more complex orientations, and so to handle more complex diffusion patterns, various alternatives have been introduced: weighted mixtures [1,2,3,4,5], higher order tensors [6], and directional functions [7]. In contrast, nonparametric techniques estimate a fiber orientation distribution function (fODF) describing an arbitrary configuration of fibers. For this estimation, several techniques have been proposed, among them Q-ball [2], spherical harmonics [8,9], spherical deconvolution [10,11,12,7], reconstruction with constant solid angle [13], and diffusion orientation transform [14].

All of those methods rely on different techniques for reconstructing, and regularizing the pathways. Deterministic tractography using the single tensor model simply follows

the principal diffusion direction, whereas multi-fiber models use various techniques for determining the number of fibers present. While parametric methods directly describe the principal diffusion directions, interpreting the fODFs from model independent representations typically involves a separate algorithm to determine the number and orientation of diffusion patterns present [15,10,9,16,17].

### 1.1 Our Contributions

In this study we propose a single unified optimization framework for estimating the parameters of any diffusion model. We place some of the most commonly used models into an Unscented Kalman Filter (UKF) framework as described in [18]. While the method in [18] used only the one and two-tensor models, we propose to extend it to the case of nonparametric models, such as spherical harmonics. The UKF framework has the advantage of estimating the model parameters and performing tractography simultaneously, resulting in an inherent regularization of the tracts at the same time. We compare several different diffusion models on clinical in-vivo scans on two very well studied fiber bundles: The tract that connects the anterior limb of the internal capsule to certain areas of the frontal lobe, and the fibers connecting the thalamus to the cortical regions. Furthermore, we propose two new measures for in-vivo comparison of fiber tracts resulting from different tractography techniques.

## 2 Methods

In Section 2.1 we describe the different models compared in this paper. Section 2.2 outlines how the models can be put into state space form and how an unscented Kalman filter can be used for estimation of the model parameters while performing tractography.

### 2.1 Fiber Representations

We chose a total of seven diffusion representation models to investigate in detail. Four parametric mixture models, and three nonparametric techniques.

Each representation strives to consolidate a diffusion signal  $\mathbf{s} = [s_1, \dots, s_n]^T \in \mathbb{R}^n$ , measured along the corresponding gradients,  $\mathbf{q}_1, \dots, \mathbf{q}_n \in \mathbb{S}^2$  (on the unit sphere), with the underlying model. The parametric models studied in this work, are all based on the general formulation of a mixture of Gaussian tensors,

$$s(\mathbf{q}_i) = s_0 \sum_j w_j e^{-b \mathbf{q}_i^T D_j \mathbf{q}_i}, \quad (1)$$

where  $s_0$  is a baseline signal intensity,  $b$  is an acquisition-specific constant,  $w_j$  are convex weights, and  $D_j$  is a tensor matrix describing a diffusion pattern. From this general mixture model, we derive four specific representations of the diffusion pattern.

In the simplest case, we assume only one tensor component of ellipsoidal shape (1T), i.e., one principal diffusion direction  $\mathbf{m}$  with eigenvalue  $\lambda_1$  and the remaining orthonormal directions have equal eigenvalues  $\lambda_2 = \lambda_3$  (as in [4,19,7]).

It has been suggested that tractography may be improved in some cases by modeling the diffusion properties of free water separately from the diffusion in brain tissue [5]. Under the assumption that the free water is unrestricted by microstructural barriers, it can be modeled by a diagonal isotropic tensor  $D_{iso}$  with just one eigenvalue  $\lambda_{iso} = 3 \cdot 10^{-3} mm^2/s$ , the apparent diffusion coefficient of free water at body temperature. Consequently, the 1T model can be extended into a mixture of an ellipsoidal component

modeling the diffusion along axons, and an isotropic component modeling the diffusion of free water:

$$s(\mathbf{q}_i) = s_0 \left( w_1 e^{-b\mathbf{q}_i^T D \mathbf{q}_i} + w_2 e^{-b\mathbf{q}_i^T D_{iso} \mathbf{q}_i} \right).$$

This will be called the one tensor model with free water estimation (1T-FW).

In the presence of fiber crossings the single tensor models may perform poorly, as they are unable to pick up the second fiber direction [20,18]. One can extend the 1T, and 1T-FW models by adding an additional ellipsoidal diffusion tensor to each, giving the two-tensor (2T):

$$s(\mathbf{q}_i) = \frac{s_0}{2} e^{-b\mathbf{q}_i^T D_1 \mathbf{q}_i} + \frac{s_0}{2} e^{-b\mathbf{q}_i^T D_2 \mathbf{q}_i},$$

and the two-tensor model with free water estimation (2T-FW)

$$s(\mathbf{q}_i) = s_0 \left( \frac{w_1}{2} e^{-b\mathbf{q}_i^T D_1 \mathbf{q}_i} + \frac{w_1}{2} e^{-b\mathbf{q}_i^T D_2 \mathbf{q}_i} + w_2 e^{-b\mathbf{q}_i^T D_{iso} \mathbf{q}_i} \right).$$

Note, that for the last two cases the ellipsoidal tensors are equally weighted. This is justified by the fact, that the eigenvalues of the ellipsoidal tensors adjust to fit the signal in much the same way a fully weighted model would adjust.

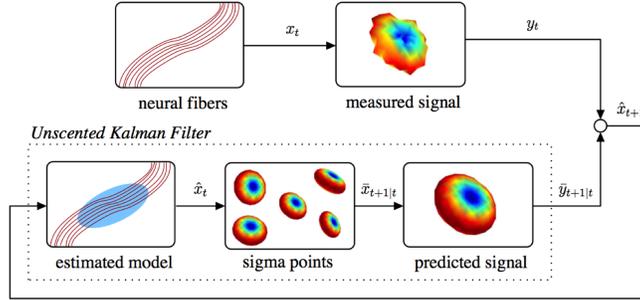
Instead of making prior assumptions regarding the shape of the signal, it is possible to estimate the diffusion orientation density function (dODF) or the underlying fiber orientation density function (fODF), directly from the signal. This has the advantage that any number of fibers can be estimated without prior knowledge. However, the principal diffusion direction(s) are not inherently available, but must be extracted from the fODF. Such an approach was taken in [9], where the dODF is first estimated followed by a “sharpening transform” to obtain the fODF. We will call this method the “sharpened spherical harmonics” or simply spherical harmonics (SH).

Spherical deconvolution (SD) as described in [12,21] (i.e., filtered spherical deconvolution) makes the assumption that the diffusion characteristics of all fiber populations found in the brain are identical in everything but orientation. Based on this, the signal can be expressed as the convolution on the unit sphere between the fODF and a single fiber response function that can be estimated from the data itself. The fODF can then be found through a spherical deconvolution. This calculation is reduced to a simple matrix inversion when the signal is represented in a spherical harmonics basis,

$$s(\mathbf{q}_i) = \sum_{j=0}^R c_j Y_j(\mathbf{q}_i), \quad (2)$$

where  $Y_j(\cdot)$  is the modified symmetric, real and orthonormal spherical harmonics basis, and  $c_j$  are its coefficients [22,12].

Another method to deblur the dODF was proposed by [13]. The infinitesimal element of the dODF was rewritten to cover a cone of constant solid angle, as opposed to a cylindrical shape as was done in [9,22,20]. This takes into account the quadratic growth of the volume element with respect to its distance from the origin. By following through the calculations the relation of the dODF to the signal is derived as,  $dODF(\mathbf{q}_i) = \frac{1}{4\pi} + \frac{1}{16\pi^2} FRT \{ \nabla_b^2 \ln(-\ln s(\mathbf{q}_i)) \}$ , where  $\nabla_b^2$  is the Laplace-Beltrami operator and  $FRT$  is the Funk-Radon Transform. This method directly estimates the fODF from the signal. We will call this method “solid angle” spherical harmonics (SA).



**Fig. 1:** System overview illustrating relation between the neural fibers, the scanner signals, and the unscented Kalman filter as it is used to estimate the local model. At each step, the filter uses its current model state to reconstruct a synthetic signal and then compares that against the actual signal from the scanner in order to update its internal model state.

## 2.2 Estimating the Fiber Model

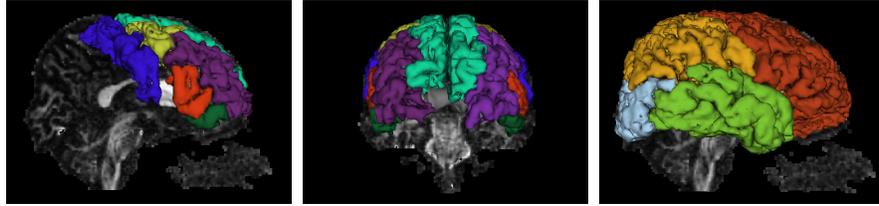
In order to estimate the models described in the last section we chose the Unscented Kalman Filtering approach described in our earlier work [18]. As opposed to the classical linear Kalman filter, the UKF can perform nonlinear estimation as is needed for the highly nonlinear reconstruction of the signals (for the parametric models).

Given the measured scanner signal at a particular voxel, we want to estimate the underlying model parameters that explain this signal. At each step, we examine the measured signal at that position, estimate the underlying model parameters, and propagate forward in the most consistent direction, i.e. the component direction most aligned with the incoming vector. Recursive estimation in this manner greatly improves the accuracy of resolving individual orientations and yields inherently smooth tracts despite the presence of noise and uncertainty. Fig. 1 illustrates this filtering process. In principal, any model can be estimated in this manner.

In order to put the studied models into state-space form, we need to define four filter components for each: 1) The system state  $\mathbf{x}$ : the model parameters, 2) The state transition  $f[\cdot]$ : how the model changes as we trace the fiber, 3) The observation  $h[\cdot]$ : how the signal appears given a particular state, 4) The measurement  $\mathbf{y}$ : the actual signal obtained from the scanner.

For all models we assume the state transition  $f[\cdot]$  to be identity, which reflects the assumption, that the local fiber configuration does not undergo drastic change from one position to the next. Furthermore, the observation  $h[\cdot]$  is simply the reconstruction of,  $\mathbf{y} = \mathbf{s} = [s_1, \dots, s_n]^T$ , where  $\mathbf{s}$  is given by the respective models for the non-parametric models, and the spherical harmonics reconstruction given in Eq. 2 for the nonparametric models.

For the two diffusion tensor based models with free water estimation (1T, 2T) the state consists of the principal diffusion direction  $\mathbf{m}$  and the major and minor eigenvalues  $\lambda_1, \lambda_2$  of each tensor. For the free water cases (1T-FW, 2T-FW), additionally the state is augmented by the weight  $w_1$ . Note, that it is not necessary to include  $w_2$  since the weights are convex, and consequently  $w_2 = 1 - w_1$ . The state for the spherical harmonics based methods is the vector of coefficients  $c_j$  in the SH basis (see Eq. 2).



**Fig. 2: Left and Middle (a):** R-L and A-P view of the desired and undesired connectivity regions for fibers passing through the anterior limb of the internal capsule (white). Desired: superior-frontal gyrus (turquoise), rostral-middle-frontal gyrus (purple), pars-triangularis (red), and pars-orbitalis (green). Undesired: precentral gyrus (blue), and caudal-middle-frontal gyrus (yellow). **Right (b):** The 4 lobes: frontal (red), temporal (green), parietal (yellow), and occipital lobe (blue).

### 3 Experiments

The seven diffusion models from Section 2.1 were compared on brain scans of 10 healthy human subjects. The images were acquired using a 3-Tesla magnet and 51 distinct diffusion gradient directions at  $b = 900s/mm^2$ .

We assessed the quality of the tractography for each method based on two very well studied fiber tracts. First, we investigated the fibers that travel through the anterior limb of the internal capsule to certain areas of the frontal lobe. This connectivity is known from postmortem and histology studies [23]. Based on information provided to us by an expert neuroanatomist the fiber bundle should pass through the anterior limb of the internal capsule and connect to the superior-frontal gyrus, rostral-middle-frontal gyrus, pars triangularis, and the pars orbitalis. Furthermore, the **fibers do not connect** to the adjoining areas of precentral gyrus, and the caudal-middle-frontal gyrus (see Fig. 2a).

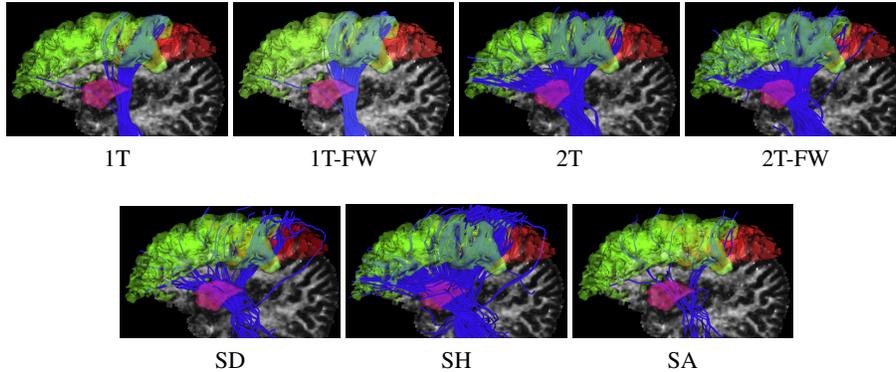
Secondly, we investigated the thalamo-cortical connections as described in [24]. The thalamus is connected to all the four lobes of the brain and a good tractography method should be able to trace these connections (see Fig. 2b for the 4 lobes). The segmentation of the internal capsule was obtained as defined in [25] whereas the cortical segmentations were obtained with Freesurfer.

The quality of the tractography of the first fiber bundle was assessed using two novel measures (Eq. 3). The traceability  $\mathcal{T}$ , measures the proportion of the fibers in a bundle that go to the desired region. A fiber was counted as  $N_{pos}$  if it passed through, or ended in one of the desired regions. If it went through one of the undesired regions it was counted as  $N_{neg}$ . Note, that the measure becomes negative for a case where the number of fibers connecting to an undesired region is larger than the number of correct fibers.

The second measure, coverage  $\mathcal{C}$ , is given by the ratio of voxels in the target region passed by a fiber  $V_{fiber}$ , to the total number of voxels in the target region  $V_{target}$ . While the traceability will assign a high score to fibers that go straight to one region without branching, the coverage will assign a high score only to fibers that cover a large area of the target region, which is expected from the anatomy.

$$\mathcal{T} = \frac{N_{pos} - N_{neg}}{N_{total}}, \quad \mathcal{C} = \frac{V_{fiber}}{V_{target}} \quad (3)$$

To generate the fiber tracts we began by seeding each voxel once, and traced the fibers for each of the diffusion models until the anisotropy threshold was reached. We



**Fig. 3:** Fiber bundles passing through the anterior limb of the internal capsule (purple) to the target region (green), and the non-target region (red) for each of the 7 examined methods.

implemented each method by closely following the corresponding papers (albeit in the UKF framework). For SH, and SD, we chose the order of the spherical harmonics  $L = 6$ , for SA we set  $L = 4$ . The peak extraction of the respective ODFs was performed following the method described in [26].

From the resulting tractography of each hemisphere, we extracted the fibers that pass through the anterior limb of the internal capsule, or the thalamus. These filtered fiber sets served as starting point to calculate the scores described above.

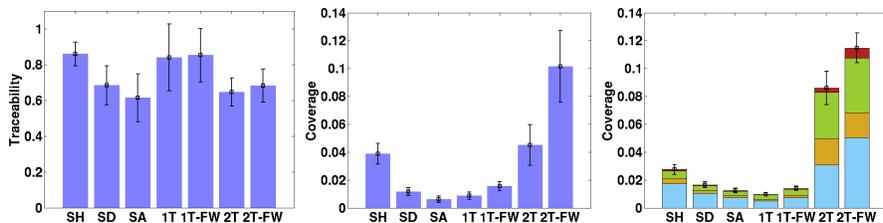
The resulting fibers for the tract through the internal capsule are shown in Fig. 3. The tracability and coverage scores for this tract are shown in Fig. 4. The percentage per lobe, and the coverage of the fibers passing through the thalamus are shown in Fig. 4. The coloring of the bars corresponds to the lobes in Fig. 2b.

**Anterior Internal capsule results:** From the results, it is clear that, while the 1T and 1T-FW methods get the best score for traceability, they perform very poorly in coverage  $\mathcal{C}$  (as can be seen in Fig. 2a). Thus, these methods have little false positives but they do not connect to all parts of the target regions (as is known from human anatomy). The 2T, 2T-FW and SH methods perform well in both measures  $\mathcal{T}$ , and  $\mathcal{C}$ . However, the method 2T-FW seems to perform the best, providing a nice balance between traceability and coverage.

**Thalamo-cortical results:** For this fiber bundle, the results demonstrate that the parametric models (2T, 2T-FW) perform significantly better in terms of connecting all the different lobes of the brain compared to the nonparametric methods. Thus, while all methods find the thalamo-cortical connections to each of the lobes, the coverage of the 2T and 2T-FW is significantly better. This is in tune with known anatomical results and using the probabilistic tractography method of [24].

## 4 Discussion and Conclusions

In this work, we proposed the UKF framework to estimate several parametric and non-parametric diffusion models and subsequently compared their ability to trace two known fiber bundles in the human brain. The comparison in this work truly demonstrates the differences in each model since the optimization framework is the same for



**Fig. 4: Left and Middle:** Resulting scores for traceability and coverage for the fiber bundle passing through the anterior limb of the internal capsule. **Right:** Coverage of the fibers passing through the thalamus. The colors correspond to the lobes outlined in Fig. 2b. All scores are averaged over 10 subjects.

all the models. We should also note that, there has been a lot of work in the literature on comparing individual models, either using Monte-Carlo simulations at a single voxel level or on a diffusion phantom (Fiber cup) or on in-vivo data as in [27]. However, in this work we propose to perform comparison using the same optimization framework for all diffusion models and compare the ability of each of these methods to trace the fiber bundle of interest (connectivity known from the anatomy) on a clinical scan.

While, the one tensor methods are significantly better in terms of traceability, however, they perform poorly in terms of coverage of the target region. On the other hand the sharpened spherical harmonics, and the two 2-Tensor methods (2T, 2T-FW) perform very well in terms of coverage for both the fiber bundles. For the fibers passing through the thalamus, the 2T-FW method outperform all other methods, in term of coverage. Furthermore, the 2-Tensor methods find some fibers in the frontal, and parietal lobe, where the other methods find almost none. We should note that, the proposed work is by no means an exhaustive comparison of the different models, however, it does provide a way to compare the various diffusion models proposed in the literature in a single optimization framework. Further, the results shown in this work are for a particularly low b-value ( $b = 900 \text{ s/mm}^2$ ). It is expected that the methods like SA and SD will perform significantly better at higher b-values.

## References

- Alexander, A., Hasan, K., Tsuruda, J., Parker, D.: Analysis of partial volume effects in diffusion-tensor MRI. *Magnetic Resonance in Medicine* **45** (2001) 770–780
- Tuch, D., Reese, T., Wiegell, M., Makris, N., Belliveau, J., Wedeen, V.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* **48** (2002) 577–582
- Kreher, B., Schneider, J., Mader, I., Martin, E., Hennig, J., Il’yasov, K.: Multitensor approach for analysis and tracking of complex fiber configurations. *Magnetic Resonance in Medicine* **54** (2005) 1216–1225
- Parker, G., Alexander, D.C.: Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Phil. Trans. R. Soc. B* **360** (2005) 893–902
- Pasternak, O., Sochen, N., Gur, Y., Intrator, N., Assaf, Y.: Free water elimination and mapping from diffusion mri. *Magnetic Resonance in Medicine* **62**(3) (2009) 717–730
- Basser, P.J., Pajevic, S.: Spectral decomposition of a  $4^{th}$ -order covariance tensor: Applications to diffusion tensor MRI. *Signal Processing* **87** (2007) 220–236

7. Kaden, E., Knösche, T., Anwander, A.: Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *NeuroImage* **37** (2007) 474–488
8. Anderson, A.: Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine* **54**(5) (2005) 1194–1206
9. Descoteaux, M., Deriche, R., Anwander, A.: Deterministic and probabilistic Q-ball tractography: from diffusion to sharp fiber distributions. Technical Report 6273, INRIA (2007)
10. Jian, B., Vemuri, B.: A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. *Trans. on Med. Imag.* **26**(11) (2007) 1464–1471
11. Jansons, K., Alexander, D.C.: Persistent angular structure: New insights from diffusion MRI data. *Inverse Problems* **19** (2003) 1031–1046
12. Tournier, J.D., Calamante, F., Gadian, D., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* **23** (2004) 1176–1185
13. 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. *Magnetic Resonance in Medicine* **64**(2) (2010) 554–566
14. Özarslan, E., Shepherd, T., Vemuri, B., Blackband, S., Mareci, T.: Resolution of complex tissue microarchitecture using the diffusion orientation transform. *NeuroImage* **31**(3) (2006) 1086–1103
15. Zhan, W., Yang, Y.: How accurately can the diffusion profiles indicate multiple fiber orientations? A study on general fiber crossings in diffusion MRI. *J. of Magnetic Resonance* **183** (2006) 193–202
16. Bloy, L., Verma, R.: On computing the underlying fiber directions from the diffusion orientation distribution function. In: *Medical Image Computing and Computer Assisted Intervention (MICCAI)*. (2008) 1–8
17. Schultz, T., Seidel, H.: Estimating crossing fibers: A tensor decomposition approach. *Trans. on Visualization and Computer Graphics* **14**(6) (2008) 1635–1642
18. Malcolm, J., Shenton, M., Rathi, Y.: Filtered multitensor tractography. *Medical Imaging, IEEE Transactions on* **29**(9) (2010) 1664–1675
19. Friman, O., Farneäck, G., Westin, C.F.: A Bayesian approach for stochastic white matter tractography. *Trans. on Med. Imag.* **25**(8) (2006) 965–978
20. Tuch, D.: Q-ball imaging. *Magnetic Resonance in Medicine* **52** (2004) 1358–1372
21. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35** (2007) 1459–1472
22. Descoteaux, M., Deriche, R., Knoesche, T., Anwander, A.: Deterministic and probabilistic tractography based on complex fiber orientation distributions. *Trans. on Med. Imag.* **28**(2) (2009) 269–286
23. Carpenter, M.: *Core text of neuroanatomy*, ed 4, baltimore, 1991
24. Behrens, T., Johansen-Berg, H., Woolrich, M., Smith, S., Wheeler-Kingshott, C., Boulby, P., Barker, G., Sillery, E., Sheehan, K., Ciccarelli, O., et al.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature neuroscience* **6**(7) (2003) 750–757
25. Mori, S., Wakana, S., Van Zijl, P., Nagae-Poetscher, L.: *MRI atlas of human white matter*. *Am Soc Neuroradiology* (2005)
26. Descoteaux, M.: *High Angular Resolution Diffusion MRI: from Local Estimation to Segmentation and Tractography*. PhD thesis, University of Nice-Sophia Antipolis (2008)
27. Yo, T., Anwander, A., Descoteaux, M., Fillard, P., Poupon, C., Knösche, T.: Quantifying brain connectivity: A comparative tractography study. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009* (2009) 886–893

# The Effect of Reorientation of the Fibre Orientation Distribution on Fibre Tracking

Daan Christiaens<sup>1,2</sup>, Thijs Dhollander<sup>1,2</sup>, Frederik Maes<sup>1,2</sup>, Stefan Sunaert<sup>1,3</sup>,  
and Paul Suetens<sup>1,2</sup>

<sup>1</sup> Medical Imaging Research Center (MIRC), KU Leuven, Leuven, Belgium

<sup>2</sup> Center for Processing Speech and Images (PSI), Department of Electrical Engineering (ESAT), Faculty of Engineering, KU Leuven, Leuven, Belgium

<sup>3</sup> Department of Radiology, University Hospitals of the KU Leuven, Leuven, Belgium

**Abstract.** Diffusion weighted imaging (DWI) allows to delineate neural fibres, based on local, directional information of the diffusion of water. Due to its directional nature, the local information needs to be reoriented upon image transformation, in order to preserve correspondence to the anatomy. In this work, we show that reorientation of the fODF with preservation of volume fractions (PVF) affects both deterministic and probabilistic fibre tracking. We identify the main causes for this, and validate them on synthetic and real brain DWI data. The problem is not with the PVF reorientation itself, but rather with the fODF reconstruction, its use in fibre tracking, and the influence of the seeds.

## 1 Introduction

Diffusion weighted imaging (DWI) is a magnetic resonance imaging (MRI) modality that measures the diffusion of water *in vivo*, along the direction of a diffusion gradient. High angular resolution diffusion imaging (HARDI) applies many, evenly distributed gradient directions to obtain a spherical distribution of the diffusion process in every voxel. Under the assumption that the diffusion is hindered by the tissue structure, and axon myelination in particular, DWI can estimate the axon directions in every voxel in the image. The fibre orientation distribution function (fODF) [1–3] represents the probability of fibres in a certain direction, based on HARDI data and given a convolution kernel  $K(\theta)$ . The fODF is defined by

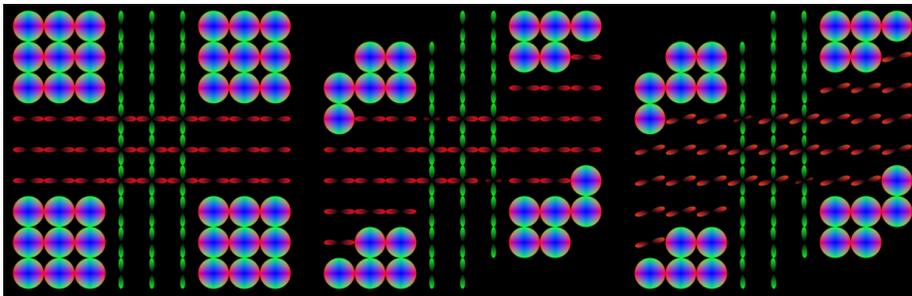
$$S(\theta, \phi) = K(\theta) * \text{fODF}(\theta, \phi) , \quad (1)$$

where  $S(\theta, \phi)$  is the measured HARDI signal and  $*$  denotes spherical convolution. The fODF reconstruction from HARDI data is then referred to as *spherical deconvolution*. Descoteaux et al. [3] have established linear relations between the HARDI signal and the fODF, using a basis of real, symmetric, orthonormal spherical harmonics (SH).

Fibre tracking aims at reconstructing the white matter (WM) connections, by “walking” along the estimated fibre direction. Deterministic fibre tracking methods follow the peaks, i.e., the local maxima, of the fODF. Probabilistic

methods draw random samples from a probability distribution like the fODF, and hence allow to assess the *sensitivity* of the tracking to local errors [4].

Registration, i.e., spatial alignment of images, is a common requirement in medical image analysis, e.g., for comparing images of different patients. All registration methods aim to optimize a *similarity measure* in the parameter space of a spatial transformation. After the transformation, image resampling on a regular grid is often required. The directional nature of DWI data imposes an additional requirement on the transformation, known as *reorientation* [5–7]. As shown in Fig. 1, the local fibre directions in each image need to be corrected for the transformation, as to preserve the coherence in the underlying tissue structure. In this work, we investigate how reorientation can affect fibre tracking.



**Fig. 1.** The need for reorientation: A vertical shearing is applied to the fODF image on the left. Without reorientation (*middle*) the correspondence to the anatomical bundles is lost. Reorientation (*right*) corrects the fODFs for the transformation.

## 2 fODF Reorientation and Fibre Tracking

### 2.1 Preservation of volume fractions

The fODF represents the fraction of fibres in each direction. If the image transformation locally compresses the fibre distribution on the sphere, the amplitude of the fODF should increase, similar to a change-of-variables of a PDF. This principle is referred to as *preservation of volume fractions* (PVF), and is formally stated as

$$\text{fODF}(\theta, \phi) d\Omega = \text{fODF}'(\theta', \phi') d\Omega' , \quad (2)$$

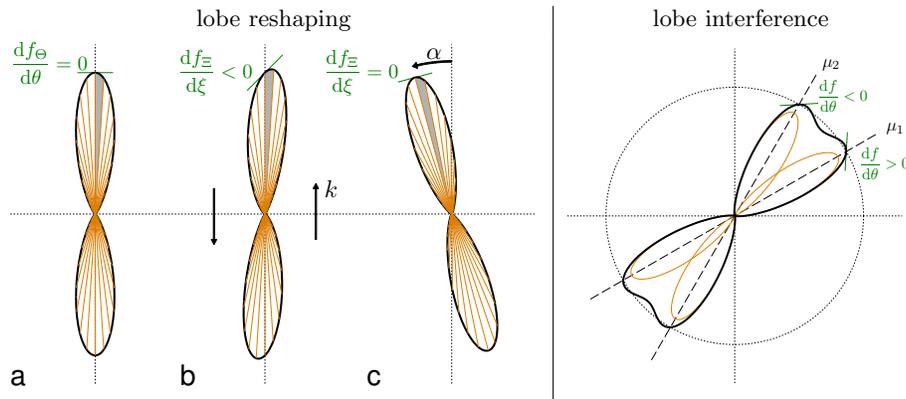
where  $d\Omega = \sin \theta d\theta d\phi$  is a small surface patch on the unit sphere. The primed symbols represent the coordinates and the fODF after transformation. From this equation, Hong et al. [5] have derived that

$$\text{fODF}'(\theta', \phi') = \text{fODF}(\theta, \phi) \frac{\sin \theta}{\sin \theta'} \frac{1}{|\det(\mathbf{J}_{\Omega})|} , \quad (3)$$

where  $\mathbf{J}_\Omega$  is the Jacobian of the angular transformation from  $(\theta, \phi)$  to  $(\theta', \phi')$ , which can be derived from the deformation field at every voxel. Raffelt et al. [6] have presented a computationally efficient method that models the fODF as a weighted sum of SH  $\delta$ -functions. Dhollander et al. [7] have translated this method to the signal space and added an isotropic volume fraction. The use of one of these PVF methods generally gives rise to a result such as the one in Fig. 1, where the angles between the peaks of the fODF can change upon reorientation.

## 2.2 Deterministic fibre tracking

In order to study the effect of PVF reorientation on deterministic fibre tracking, we must look at its influence on the local maxima of the fODF. We identify 2 effects that occur in theory, i.e., *lobe reshaping* and *lobe interference*, and one additional effect that occurs due to aliasing in the SH basis.



**Fig. 2.** The effect of lobe reshaping and lobe interference on the local maxima.

**Lobe reshaping** PVF reorientation will generally change the shape of the fODF, and consequently alter the position of the local maxima. To avoid interference from other lobes, we restrict our study to single-fibre voxels. In addition, we aim to avoid unnecessary mathematical complexity by limiting the study to 2-D functions in polar coordinates, which can best be thought of as a cross-section of the fODF in the plane of the transformation. We define  $f_\Theta(\theta)$  as the (2-D) fODF of the original fibre direction  $\Theta$ ,  $f_\Xi(\xi)$  as the fODF of the transformed fibre direction  $\Xi$ , and  $g: \Theta \rightarrow \Xi: \theta \mapsto g(\theta)$  as a monotonously increasing function that defines the local reorientation of the fODF. Equation (3) then becomes

$$f_\Xi(\xi) = \frac{dg^{-1}(\xi)}{d\xi} f_\Theta(g^{-1}(\xi)) . \quad (4)$$

Let  $\theta^*$  be the maximum of  $f_\Theta(\theta)$ , i.e.,  $\frac{df_\Theta(\theta^*)}{d\theta} = 0$  and  $\frac{d^2f_\Theta(\theta^*)}{d\theta^2} < 0$ . The derivative of (4) at the transformed maximum  $\xi^* = g(\theta^*)$  is then given by

$$\frac{df_\Xi(\xi^*)}{d\xi} = f_\Theta(\theta^*) \frac{d^2g^{-1}(\xi^*)}{d\xi^2}, \quad (5)$$

In general, this expression is not equal to 0, which proves that the peak of the transformed fODF does not correspond to the transformed peak of the original fODF. This is illustrated in Fig. 2b. In the special case where  $g$  is a pure rotation, the second derivative of  $g$  is 0 and  $\xi^*$  is a local maximum of  $f_\Xi$ . Hence, for a rigid rotation, the maxima of the transformed fODF correspond to the transformed maxima of the original fODF, as shown in Fig. 2.

**Lobe interference** As a second effect of reorientation, the lobes of the fODF in a voxel with crossing fibres can move towards or away from each other. Each lobe inevitably has a certain width, for several reasons. (i) The bandwidth of the fODF reconstruction is constrained by the spatial distribution of the HARDI samples in  $q$ -space. More samples increase the angular resolution and hence allow for narrower lobes. (ii) The noise regularization of the various fODF reconstruction techniques, whether by low-pass filtering, constrained SD or the finite order of the SH-basis, puts an additional constraint on the bandwidth. (iii) The true underlying fibre structure in one voxel is expected to have a certain spread as well, especially due to the partial volume effect. As a result of their width, the lobes can interfere with each other, and thereby influence the location of the maxima. This effect is illustrated in the right panel of Fig. 2. Note that lobe interference occurs as a general side effect of fODF reconstruction, even without reorientation. However, in its ability to change the local angular fibre density, fODF reorientation can increase or decrease the effect of interference.

**Aliasing in the SH basis** The SH basis can be thought of as the analogy of a Fourier series on a sphere. By this analogy, the SH  $\delta$ -function does not have an exact representation in a basis of finite order, but will rather have side lobes due to aliasing. In Raffelt’s method, a mixture of  $\delta$ -functions is fitted to the fODF, in such a way that the net effect of the side lobes is zero. However, when the  $\delta$ -functions are reoriented, the side lobes start to interfere in a complex pattern, and the net effect will no longer be zero. This can influence the peaks of the fODF and consequently fibre tracking.

### 2.3 Probabilistic fibre tracking

Probabilistic fibre tracks are generated step-by-step, by drawing random samples from the fODF. As opposed to the deterministic tracking, it is difficult to compare tracks based on correspondences, as the property of having one track per seed is lost. Instead, we can look at the statistical distribution of all generated fibre tracks before and after transformation. Ideally, the distribution of the tracks in

the transformed image should match the transformed distribution of the tracks in the original (untransformed) image.

At the local scale, the fODF is defined as the probability distribution of the direction of all tracks within a single voxel. Reorientation with preservation of volume fractions ensures, by definition, that the transformed fODF is equivalent to the distribution of the direction of the transformed tracks (within the limits of partial volume effects). This suggests that the statistical distribution of the tracks in the transformed image is equivalent to the distribution of transformed tracks of the original image.

However, in the above reasoning, it is implicitly assumed that the distribution of tracks, generated by probabilistic fibre tracking, matches the distribution of the true fibres that generate the fODF. While, in theory, this is indeed the goal of fibre tracking, in practice, it is generally not the case for several reasons. First of all, the reconstructed fODF does not entirely correspond to the true directional fibre density, for reasons associated with partial voluming and aliasing. Secondly, and more importantly, the *topological* density of the generated tractography will generally differ from the true fibre density. While the number of fibres that cross a volume element in white matter are a physiological property of the tissue, the number of generated tracks strongly depends on the distribution of the seeds. Moreover, if seeds are uniformly distributed in the original and the transformed image, the *seeding density* of the transformed original tracks differs as well.

### 3 Experimental Set-up

#### 3.1 DWI data

**Synthetic image** We have created a synthetic image of a perpendicular crossing using a multi-tensor model. Horizontal and vertical diffusion tensors with eigenvalues  $\lambda_1 = 0.0018 \text{ mm}^2/\text{s}$  and  $\lambda_2 = \lambda_3 = 0.0006 \text{ mm}^2/\text{s}$  are created and converted to signals in  $q$ -space using 75 equally distributed gradient directions and  $b = 3000 \text{ s}/\text{mm}^2$ . At the crossing, we calculate the average of both reconstructed signals. The background is assumed to be isotropic, and is generated from a spherical diffusion tensor with eigenvalues  $\lambda_1 = \lambda_2 = \lambda_3 = 0.002 \text{ mm}^2/\text{s}$ . The voxel size is isotropic and equal to 1 mm.

**Real brain image** DWI data was obtained from one healthy volunteer, using a Siemens 3T scanner at an isotropic voxel size of 2.5 mm. The  $S_0$  image was acquired as the average of 10 regular T2-weighted images ( $b = 0 \text{ s}/\text{mm}^2$ ). In addition, 75 diffusion-weighted images  $S(b, \mathbf{g})$  were recorded at  $b = 2800 \text{ s}/\text{mm}^2$ .

#### 3.2 fODF estimation

The fibre orientation in each voxel, represented by the fODF, is estimated with constrained spherical deconvolution (CSD) [2] in an SH-basis of order 6. We use

the software package MRtrix<sup>4</sup> [4]. The diffusion kernel is estimated directly from the DWI data, by aligning and averaging the signal profiles within a mask of single fibre voxels, retrieved as the voxels with  $FA > 0.7$  [1, 2, 4].

### 3.3 Tractography algorithm

Fibre tracking is performed using MRtrix, with a standard Euler stepping method. In deterministic tracking, the direction of the next step is determined by of the fODF peak closest to the current direction. In probabilistic tracking, the next step is a random sample, drawn from the fODF using rejection sampling [4]. The step size is set to 0.2 mm. The fODF in each step is retrieved from trilinear interpolation on the SH coefficients. Tracks are terminated if the amplitude of the fODF peak falls below a threshold, set to 0.1, or if the local radius of curvature is smaller than 1 mm. Seeds are either distributed uniformly in a spherical ROI, or within a (full brain) mask, specified by the user.

### 3.4 Image transformation and reorientation

The synthetic image is submitted to a shearing along the  $y$ -axis, i.e.,  $\mathbf{F} = [1 \ 0; k \ 1]$ , where the parameter  $k = \tan \alpha$  is determined by the shearing angle  $\alpha$ . Different values for  $\alpha$  are evaluated, ranging from  $10^\circ$  to  $60^\circ$ . The transformation and resampling are performed in MATLAB<sup>5</sup> on the raw DWI data. Reorientation is done on the signal values as well, using the method of Dhollander et al. [7] with 1000 evenly distributed *alpha* functions ( $\delta$ -functions translated to signal space).

The real brain image is submitted to a non-linear transformation. The deformation field is obtained from registration to another subject, recorded under similar conditions, using the registration algorithm described in [8].

### 3.5 Distance measure

We will use the *current distance*  $D_{cd}$  [9, 10] to compare the full set of tracks in the transformed image to the transformed tracks of the original image, for different harmonic orders. The current distance between 2 sets of tracks  $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$  and  $\mathcal{B} = \{B_1, B_2, \dots, B_m\}$ , where each track  $A_k$  and  $B_l$  is a sequence of points  $\mathbf{a}_i$  and  $\mathbf{b}_j$ , is defined as

$$D_{cd}(\mathcal{A}, \mathcal{B}) = \kappa(\mathcal{A}, \mathcal{A}) + \kappa(\mathcal{B}, \mathcal{B}) - 2 \kappa(\mathcal{A}, \mathcal{B}), \quad (6)$$

where  $\kappa(\mathcal{A}, \mathcal{B})$  represents the similarity between  $\mathcal{A}$  and  $\mathcal{B}$  [10]. This similarity function is defined as

$$\kappa(\mathcal{A}, \mathcal{B}) = \sum_{A_k \in \mathcal{A}} \sum_{B_l \in \mathcal{B}} \sum_{\mathbf{a}_i \in A_k} \sum_{\mathbf{b}_j \in B_l} G_\sigma(\|\mathbf{b}_j - \mathbf{a}_i\|) (\delta_1^1 \mathbf{a}_i \cdot \delta_1^1 \mathbf{b}_j), \quad (7)$$

<sup>4</sup> Freely available at <http://www.nitrc.org/projects/mrtrix/> (GNU GPL)

<sup>5</sup> The MathWorks Inc., Natick, MA (<http://www.mathworks.com/products/matlab/>)

where  $G_\sigma(x)$  is a Gaussian kernel with standard deviation  $\sigma$  (equal to the voxel size in our experiments), and the operator  $\delta_1^1$  denotes the first order central difference, i.e., the tangent vector of the track in that point. As such, the current distance takes both the distance between points and the difference in direction into account. The Gaussian kernel ensures that the relative weight of points on distant tracks decreases.

## 4 Results

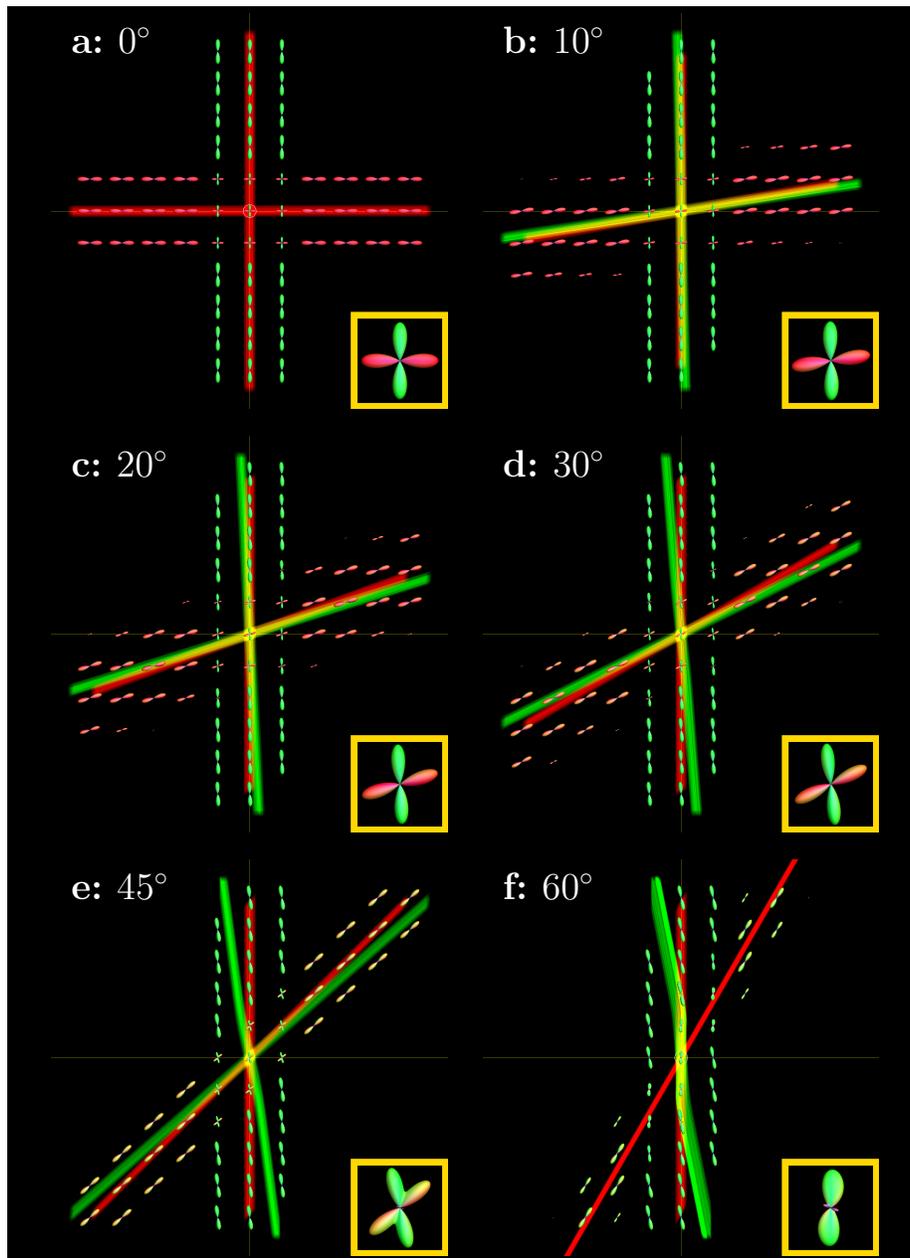
### 4.1 Synthetic data

The fODF reconstruction of the synthetic image before and after shearing is shown in Figs. 3 and 4. Deterministic fibre tracking is initiated from a spherical seeding region of radius 0.2mm at the centre of the crossing. The resulting tracks are shown in red in Fig. 3a. These tracks are submitted to the image transformation, and then compared to the outcome of tracking on the transformed (and reoriented) fODFs, shown in green in Fig. 3b–f. For shearing angles under  $30^\circ$  the green tracks are straight lines. There is a clear deflection with respect to the transformed original tracks in red. The angular difference between the red and green tracks at the crossing is not stronger than in the single fibre voxels in the distal areas. This suggests that lobe interference has little effect if the shearing angle is small with respect to the angle of the crossing. For larger shearing angles, as shown in Figs. 3e–f, the effect of lobe interference gains importance, up to the point where both lobes merge and only one fibre direction is detected.

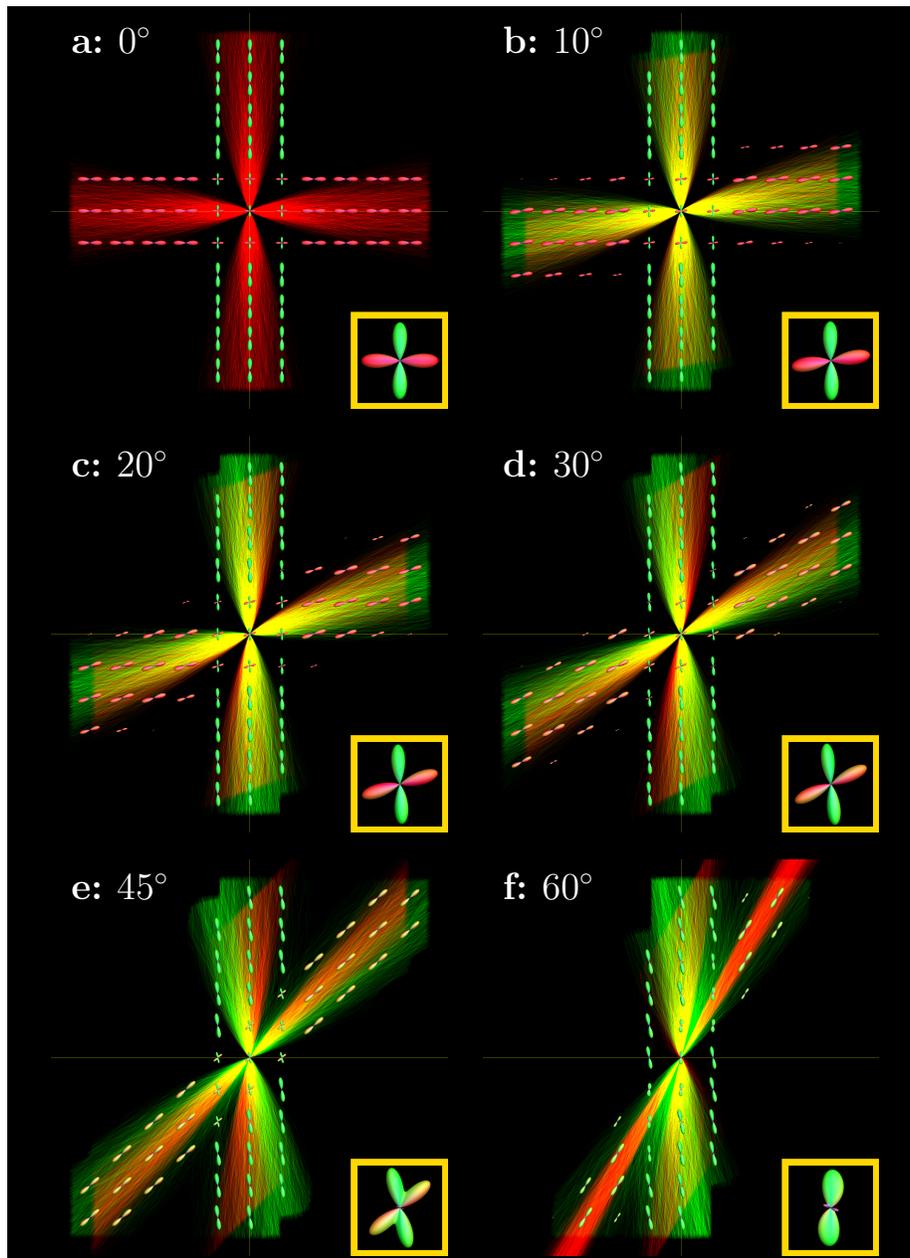
The result of probabilistic fibre tracking in the same set-up, and initiated from a single seed point at the centre of the crossing, is shown in Fig. 4. The overall distribution of the green tracks deflects from the distribution of the red tracks in much the same way as for deterministic tracking, because the fODFs have been reoriented identically. For large shearing angles, as shown in Fig. 4f, probabilistic tracking does succeed to recover some of the tracks in the horizontal (now displaced) fibre bundle, unlike deterministic tracking.

### 4.2 Full brain tracking

In the second experiment, we submit the real brain image to the non-linear transformation and reorient the signal values using PVF in the signal space. A full-brain tractography is then generated using 50 000 equally distributed seed points within a brain mask, both in the original image and in the transformed image. As in the previous experiment, we transform the original tracks to the new image and compare the result (red) to the tracks of the transformed image (green). The current distance  $D_{cd}$  between both track sets is reported in Table 1, both for deterministic and probabilistic fibre tracking and for varying harmonic order  $L$ . We observe a decrease of the current distance for increasing harmonic order. Furthermore, the current distance at a particular harmonic order is smaller for probabilistic tracking than for deterministic tracking.



**Fig. 3.** Results of deterministic tracking on synthetic data. **(a)** Original image of crossing fibres. The initial tracking is shown in red. **(b)–(f)** Image after vertical shearing. The original tracks are transformed using the same shearing and shown in red. The outcome of fibre tracking in the transformed image is shown in green. Yellow then indicates overlap between both track distributions.



**Fig. 4.** Results of probabilistic tracking on synthetic data. **(a)** Original image of crossing fibres. The initial tracking is shown in red. **(b)–(f)** Image after vertical shearing. The original tracks are transformed using the same shearing and shown in red. The outcome of fibre tracking in the transformed image is shown in green. Yellow then indicates overlap between both track distributions.

**Table 1.** Current distance  $D_{cd}$ .

( $\times 10^8$ )	deterministic	probabilistic
$L = 6$	2.811	1.502
$L = 8$	1.666	1.165
$L = 10$	1.064	0.911

## 5 Discussion and Conclusion

The fODF is, by definition, the probability distribution of the fibre directions in a voxel, and hence contains the maximal available information for fibre tracking. However, fibre reconstruction is still limited by the partial volume effect of the discrete measurements. It is for example not possible to discriminate between crossing and kissing fibres on a voxel scale. Deterministic fODF-based fibre tracking methods determine the direction of tracking as the peak of the fODF, closest to the current direction of the track. It has been pointed out in literature that the peaks do not necessarily correspond to the exact fibre directions due to noise or interference from other track directions [1–3]. Probabilistic methods draw random samples from the fODF to guide each step of the tracking.

We have shown that reorientation with preservation of volume fractions (PVF) affects the local maxima of the fODF, and hence deterministic fibre tracking. The peaks of a reoriented fODF do not always correspond to the reoriented peaks of the original fODF. We have identified 2 effects, directly associated with the PVF, that can cause this error, i.e., lobe reshaping and lobe interference, and an additional effect that hampers the PVF and influences the peaks as well, i.e., aliasing in the SH basis. Lobe reshaping occurs when a non-rigid transformation induces an unequal amount of compression in the local fibre density on the sphere. By preservation of volume fractions, local compression of the fODF will rightly lead to an increase of the amplitude of the fODF. Consequently, unequal compression will cause an unequal rescaling of the fODF, which can result in a shift of the local maximum. Lobe interference is a general effect in fODF reconstruction, that occurs mainly at sharp fibre crossings. It is relevant in the context of reorientation as well, as a non-rigid transformation can change the angle between the fibre directions. We have shown that, if the transformation pushes two fibre directions closer to each other, lobe interference will cause an additional shift of the peaks of the fODF towards each other, up to the point where both peaks merge and only one fibre direction can be retrieved. Due to the combination of lobe reshaping and lobe interference, fibre tracking on the non-rigidly transformed fODFs will produce a different result than transforming the result of fibre tracking on the original fODFs. The problem is not with the reorientation or the idea of PVF, but rather with the principle of deterministic fibre tracking itself. The local maximum of the fODF is in fact the mode of a PDF, which is known to be unstable upon a change of variables.

In our experiment on synthetic data, we have noticed that the green tracks indeed differ from the red tracks. However, the deflection is not according to the

predicted lobe reshaping. This suggests that the PVF reorientation is hampered by the limited order of the SH basis, that causes side lobes in the  $\delta$ -functions in Raffelt’s method due to aliasing. Upon reorientation, these side lobes will interfere in a complex pattern, that can influence the shape of the main lobes and hence the local maxima and the direction of tracking. In a recent paper, Raffelt et al. [11] have suggested the use of *apodised*  $\delta$ -functions to avoid the influence of the side lobes.

During the course of our work, we have noticed that the reorientation methods of Raffelt et al. [6] and Dhollander et al. [7] suffer both from aliasing artefacts. Moreover, Raffelt et al. [11] report identical effects when using Hong’s method [5]. This is surprising, as Hong’s method is a direct implementation of PVF reorientation, and in no way uses  $\delta$ -functions that can introduce side lobes. We therefore suspect that the aliasing artefacts might have a more profound cause, i.e., the sparsity of the reoriented samples or  $\delta$ -function directions. The unequal distribution of the samples might indeed introduce Gibbs ringing artefacts when fitted in a SH-basis. The apodised  $\delta$ -functions might then perform better because they are wider, rather than because of the reduced side lobes. This remark is made as a point of discussion, and should be the topic of further research.

Probabilistic fibre tracking, initiated from a single seed, is in principle not influenced by lobe reshaping and interference, but does suffer from aliasing artefacts in the SH basis. As a result, the distribution of tracks of the transformed image deflects from the transformed distribution of tracks of the original image. Additionally, if a large number of randomly distributed seeds is used, e.g., for full-brain tractography, the density of the generated tracks might be unfair with respect to the anatomy (e.g., longer tracks will have more seeds and hence a larger track density). Moreover, the seeding distribution will generally differ between the original and transformed images.

In real data, we have reported a decrease of the current distance between the full brain tractography after transformation and the transformed tractography of the untransformed image for increasing harmonic order. For higher orders of the SH basis, the lobes are less wide. As such, the effects of lobe reshaping and interference will be smaller, and the influence on the tracking will be reduced. For a given order, the current distance is smaller for probabilistic tracking than for deterministic tracking, which might suggest that probabilistic fibre tracking is less affected by the reorientation.

In the case of a global, rigid transformation, the reorientation is a pure rotation that will not affect the shape of the fODF in any way. The distribution of the seeds in the original and transformed images will be equal as well. As such, we expect that rigid registration has no effect on fibre tracking, apart from interpolation effects. In the case of non-rigid registration, it is better to do the fibre tracking on the untransformed image, i.e., *before* the registration, and submit the outcome to the transformation. In the case that fibre tracking *needs* to be done after registration, e.g., for comparison to an atlas, one has to be aware of the deflection of the tracks, caused by the registration.

In conclusion, based on theoretical considerations and experiments on synthetic and real HARDI data, we have shown that spatial transformations can influence the outcome of fibre tracking, both for deterministic and probabilistic methods.

## References

1. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* **23**(3) (2004) 1176–1185
2. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**(4) (May 2007) 1459–1472
3. Descoteaux, M., Deriche, R., Knösche, T.R., Anwander, A.: Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE transactions on medical imaging* **28**(2) (February 2009) 269–286
4. Tournier, J.D., Calamante, F., Connelly, A.: MRtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology* **22**(1) (2012) 53–66
5. Hong, X., Arlinghaus, L.R., Anderson, A.W.: Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data. *Magnetic Resonance in Medicine* **61**(6) (2009) 1520–1527
6. Raffelt, D., Tournier, J.D., Fripp, J., Lepore, N., Crozier, S., Connelly, A., Salvado, O.: Non-Rigid Registration of Diffusion Weighted Images using Fibre Orientation Distributions. In: *Proceedings 17th Scientific Meeting, International Society for Magnetic Resonance in Medicine*. Volume 17. (2009) 1387
7. Dhollander, T., Van Hecke, W., Maes, F., Sunaert, S., Suetens, P.: Spatial transformations of high angular resolution diffusion imaging data in Q-space. In: *Proc. MICCAI 2010 workshop on computational diffusion MRI*. (2010) 73–83
8. Dhollander, T., Veraart, J., Van Hecke, W., Maes, F., Sunaert, S., Sijbers, J., Suetens, P.: Feasibility and Advantages of Diffusion Weighted Imaging Atlas Construction in Q-Space. In Fichtinger, G., Martel, A., Peters, T., eds.: *Proc. MICCAI 14th international conference on medical image computing and computer assisted intervention*. Volume 6892 of *Lecture Notes in Computer Science.*, Springer (2011) 166–173
9. Vaillant, M., Glaunès, J.: Surface matching via currents. In Christensen, GE and Sonka, M., ed.: *Proc. Information Processing in Medical Imaging*. Volume 3565 of *Lecture Notes in Computer Science*. (2005) 381–392
10. Jiao, F., Phillips, J.M., Stinstra, J., Krger, J., Varma, R., Hsu, E., Korenberg, J., Johnson, C.R.: Metrics for Uncertainty Analysis and Visualization of Diffusion Tensor Images. In Liao, H., Edwards, P.J.T.T., Pan, X., Fan, Y., Yang, G.Z., eds.: *Medical Imaging and Augmented Reality*. Volume 6326 of *Lecture Notes in Computer Science.*, Springer (2010) 179–190
11. Raffelt, D., Tournier, J.D., Crozier, S., Connelly, A., Salvado, O.: Reorientation of fiber orientation distributions using apodized point spread functions. *Magnetic Resonance in Medicine* **67**(3) (March 2012) 844–55

## Oral Session II: Registration

# Registration of Spherical Functions from High Angular Resolution Diffusion Imaging using the Heat Kernel Signature and Möbius Transformation

L. Zhan<sup>1</sup>, Y. Wang<sup>2</sup>, P.M. Thompson<sup>1</sup>

1. Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, CA, USA
2. Computer Science and Engineering, School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ, USA

**Abstract.** High angular resolution diffusion imaging (HARDI) is a powerful variant of diffusion MRI, which images the 3D profile of diffusion at each imaged location in the brain. At each voxel, this leads to an orientation density function (ODF) expressing the probability density of water diffusion in each direction on the unit sphere. As diffusion is greatest along the brain’s axons, these functions are used to map fiber trajectories (tractography) and fiber integrity. To average and compare this data across subjects, we developed a new method based on the heat kernel signature and ‘Möbius voting’ to identify and align peaks of diffusivity on the sphere. We compare our method to standard coordinate-based averaging, and it helps to reinforce consistent features in both synthetic data and real HARDI data. This scheme to compare and integrate HARDI data may be helpful in population-based analyses of brain integrity and connectivity.

**Keywords:** HARDI, ODF, Heat Kernel Signature, Möbius Transformation

## 1 Introduction

High angular resolution diffusion imaging (HARDI) offers several advantages for studying the local integrity, geometry and connectivity of white matter fibers in the living brain. When fibers cross or mix in the same voxel, measures based on the widely-used diffusion tensor can be biased (e.g., fractional anisotropy estimates tend to be too low), in the ~40% of white matter voxels where more than one dominant fiber direction is detectable [1, 2]. From HARDI, orientation distribution functions (ODF) may be defined as the radial projection of the spherical diffusion function. Despite the rich information in HARDI datasets – yielding an entire spherical function at each point in a 3D image – the statistical study of diffusion images have mainly focused on scalars derived from the tensor model. The vast majority of studies still examine fractional anisotropy (FA) as the main target of study. [3]

Precise voxel-level comparison of HARDI datasets cannot be performed across subjects without removing some of the confounding structural variability. The diffusion functions can be considered as defined on the anatomy of each individual, and if nonlinear registration of the anatomy is used for normalization, there will be better agreement among the diffusion functions defined on the anatomy. Better structural alignment generally improves the power to detect changes related to disease, development, and aging, and improves the accuracy of segmentation and post-processing of HARDI data [4].

Thus, there has been a recent flurry of HARDI registration algorithms [5-10]. Some work has attempted to register fields of spherical functions, treating the spherical function as a probability density, and defining distances between them based on information theory, such as the Fisher-Rao metric, or symmetrized Kullback-Leibler distance [11]. Such metrics may be integrated over the whole brain, and their variational derivative may be computed with respect to tunable parameters of 3D vector fields used to register the images [12]. Chiang et al. [12] found that 3D fluid registration of diffusion images gave more accurate alignment when the tensors or ODFs were aligned using information theory, but the approach was limited because there was no attempt to define corresponding features (such as peaks of diffusivity) on the ODFs across subjects. Because of this, corresponding features were not reinforced in the ODFs across subjects, when data were averaged and compared.

In the standard, single-tensor, DTI model, the fitted diffusion tensor has only one dominant direction (its principal eigenvector), so registration is straightforward as it only involves aligning one direction to the other (although clearly there may be ambiguous or difficult cases if the principal eigenvector is not unique or if the 2 or 3 highest eigenvalues are close in magnitude). By contrast, the ODF model can often have multiple dominant directions. This leads to difficulties in peak matching, unless an explicit effort is made to match them. In this study, we propose a new ODF registration method, designed to reinforce common features in populations of diffusion images. As the ODF is defined on the unit sphere, we first use the heat kernel signature (HKS) to detect the peaks in the spherically-parameterized functional domain. We then use stereographic projection to map the 3D spherical surface to the 2D complex plane, and match the HKS peaks in complex plane using Möbius transformation. We then use the inverse stereographic projection to pull-back the mapping to the 3D space. Using this method, ODF multiple peaks may be matched very quickly, giving robust results in line with intuition.

## 2 Method description

### 2.1 ODF Computation using the Tensor Distribution Function

We adopted the Tensor Distribution Function (TDF) theory [13, 14] to calculate the ODF. The space of symmetric positive definite 3x3 matrices was denoted by  $\bar{D}$ . The probabilistic ensemble of tensors, represented by a tensor distribution function (TDF)  $P$ , is

defined on the tensor space  $\bar{D}$  that best explains the observed diffusion-weighted signals. We used the least-squares approach with the gradient descent defined in [13] to solve for an optimal TDF  $P^*$ . From the TDF  $P^*$ , the ODF may be computed analytically [15]. All ODFs were rendered using 642 point samples, determined using an icosahedral approximation of the unit sphere. The TDF-based ODF calculation can be expressed in **Equation 1** (Please refer to [13] for explanation in details). In addition, 1280 faces were determined using Delaunay triangulation. Every local maximum of the 3D surface may be one of the possible dominant directions. To find the local maxima, we adapted Heat Kernel Signature theory.

$$\begin{cases} S_{cal}(q) = \int_{D \in \bar{D}} P(D) \cdot e^{(-bq^T D q)} dD \\ P^* = \operatorname{argmin}_p \sum_i (S_{obs}(q_i) - S_{cal}(q_i))^2 \\ ODF(\tilde{x}) = \frac{1}{4\pi} \sum_{D \in \bar{D}} \frac{P^*(D)}{\det(D)^{\frac{1}{2}} (\tilde{x}^T \cdot D^{-1} \cdot \tilde{x})^{\frac{3}{2}}} \end{cases} \quad (1)$$

## 2.2 Peak detection using the Heat Kernel Signature

The Heat Kernel Signature (HKS) was first proposed by [16]. Let  $M$  be a compact Riemannian manifold, possibly with a boundary. The heat diffusion process over  $M$  is governed by the heat equation:

$$\Delta_M \mu(x, t) = -\frac{\partial \mu(x, t)}{\partial t} \quad (2)$$

Here  $\Delta_M$  is the Laplace-Beltrami operator of  $M$ . If  $M$  has boundaries, the Dirichlet boundary condition  $\mu(x, t)=0$  for all  $x \in \partial M$  and all  $t$ . Given an initial heat distribution  $f: M \rightarrow \mathbb{R}$ , let  $H_t(f)$  denote the heat distribution at time  $t$ , namely  $H_t(f)$  satisfies the heat equation for all  $t$ , and  $\lim_{t \rightarrow 0} H_t(f) = f$ .  $H_t$  is called the heat operator. It is well known [17] that for any  $M$ , there exists a function  $k_t(x, y): \mathbb{R}^+ \times M \times M \rightarrow \mathbb{R}$  such that

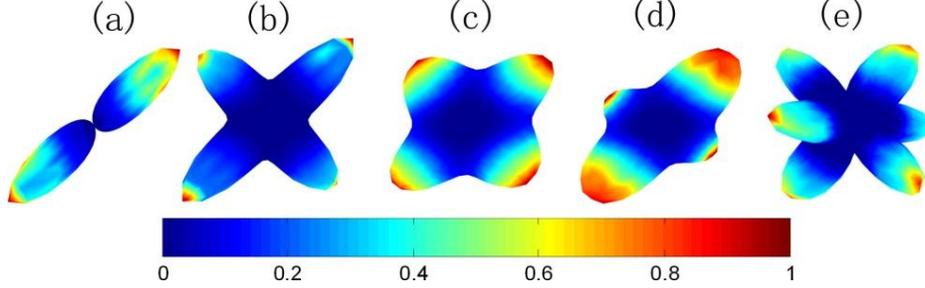
$$H_t f(x) = \int_M k_t(x, y) f(y) dy \quad (3)$$

Where  $dy$  is the volume form at  $y \in M$ . The minimum function  $k_t(x, y)$  that satisfies Eq. (3) is called the heat kernel. Then, given a point  $x$  on the manifold  $M$ , the Heat Kernel Signature,  $HKS(x)$  is defined as a function over the temporal domain:

$$HKS(x): \mathbb{R}^+ \rightarrow \mathbb{R}, HKS(x, t) = k_t(x, x) \quad (4)$$

The HKS has a very nice property: the surface HKS function of a point  $x$  is directly related to the Gaussian curvature on a surface at  $x$  [16], which is an intrinsic geometric property of the shape. This method offers the advantages of inelastic deformation invariance and is somewhat robust to topological noise. So it offers a reliable way for us to compute surface Gaussian curvature related statistics, e.g., for peak detection (**Figure 1**). HKS was computed using 1-100 (in unit increments) as the artificial time and then normalized by the maximum value. We chose the HKS values at  $t=1$  as the HKS values for each ODF. Here, the local maxima of the ODFs were selected, based on the amplitude of the HKS value at each point in the spherical surface. For each ODF, we pick  $n$  points - let

us assume  $p_i$  from ODF1 and  $q_i$  from ODF2 and  $i=1, 2, \dots, n$ . **Fig. 1** shows a few computed HKS examples on various ODFs. Once we define local maxima ( $p$  and  $q$ ) for the ODFs, the next step is to find a diffeomorphism to match the ODFs. To match multiple peaks at the same time, we use Möbius transformation theory, as every Möbius transformation is a bijective conformal map of the Riemann sphere to itself.



**Fig. 1.** Heat Kernel Signature for several typical kinds of orientation density functions (ODFs). (a) HKS for one fiber; (b) HKS for two sharply concentrated fibers crossing at 90 degrees; (c) HKS for two less concentrated fibers crossing at 90 degrees; (d) HKS for two unequally weighted fibers crossing at 90 degrees; (e) HKS for three fibers, each with an equal weighting. *Red colors* indicate high HKS values and *blue colors* indicate low HKS values. A high HKS value indicates a high Gaussian curvature point, i.e., a potential ODF local maximum.

### 2.3 Peak matching using Möbius Transformation and Möbius Voting

As any ODF is defined on a spherical domain, we can use stereographic projection to project the 3D spherical surface into the 2D complex plane. In the Cartesian coordinates  $(x, y, z)$  on the sphere and  $(X, Y)$  on the plane, the projection and its inverse are given by the formulae:

$$\begin{aligned} (X, Y) &= \left( \frac{x}{1-z}, \frac{y}{1-z} \right) \\ (x, y, z) &= \left( \frac{2X}{1+X^2+Y^2}, \frac{2Y}{1+X^2+Y^2}, \frac{-1+X^2+Y^2}{1+X^2+Y^2} \right) \end{aligned} \quad (5)$$

In geometry, a Möbius transformation of the plane is a rational function of the form:

$$f(z) = \frac{az+b}{cz+d} \quad (6)$$

- involving one complex variable  $z$ ; here the coefficients  $a, b, c, d$  are complex numbers satisfying  $ad - bc \neq 0$ . In order to compare two ODFs, it is desirable to adjust the conformal mapping to match the geometric features on the ODF as well as possible. So we use the least-squares approach to compute an optimal Möbius transformation [18].

$$\{a, b, c, d\} = \arg \min \sum_{i=1}^n g(z_i) \left| \frac{az_i + b}{cz_i + d} - \tau_i \right|^2 \quad (7)$$

where  $z_i$  is the stereographic projection of  $p_i$ ,  $\tau_i$  is the projection of  $q_i$ .  $g$  is the conformal factor from the plane to the sphere, which may be simplified as:

$$g(z) = \frac{4}{1+z\bar{z}} \quad (8)$$

with  $\bar{z}$  is the complex conjugate of complex number  $z$ . To ensure the peaks are matched in the right order, we use the ‘‘Möbius voting’’ concept [19]. Prior work [19] reported an algorithm to automatically discover surface point correspondences by optimizing Möbius transformation parameters. Isometries are a subset of the Möbius group. As the Möbius group is low-dimensional (6 DOF for the topological sphere), one can compute a closed-form Möbius transformation after selecting three matching points. By permuting matching points, one can obtain ‘‘votes’’ for predicted correspondences by measuring the mismatch energy values. We borrow this idea here to find the best diffeomorphism to match different ODFs. The Möbius group is much bigger than congruence, e.g., voxel rotation, which has been typically used for HARDI registration before [5]. Möbius transformation still only has limited numbers of parameters, making subsequent optimization work simpler. In this study, we permute all possible combinations of matched features to minimize the  $L^2$  norm between the transformed versions of ODF1 and ODF2 (if we assume we are mapping ODF1 to ODF2).

## 2.4 Overall framework

The current framework is summarized as follows:

- (a) Use a preliminary whole-brain registration method (affine or non-linear) to register the entire diffusion image from one subject to a target subject. This ensures all relevant brain tissues are roughly matched;
- (b) Calculate the ODF at each voxel for both the source image (to be registered) and the target image;
- (c) Compute the HKS for each ODF in both source and target and determine the peak number and locations for each ODF;
- (d) For corresponding ODFs between source and target, use Möbius transformation (and voting) to achieve finer-scale alignment of peaks.

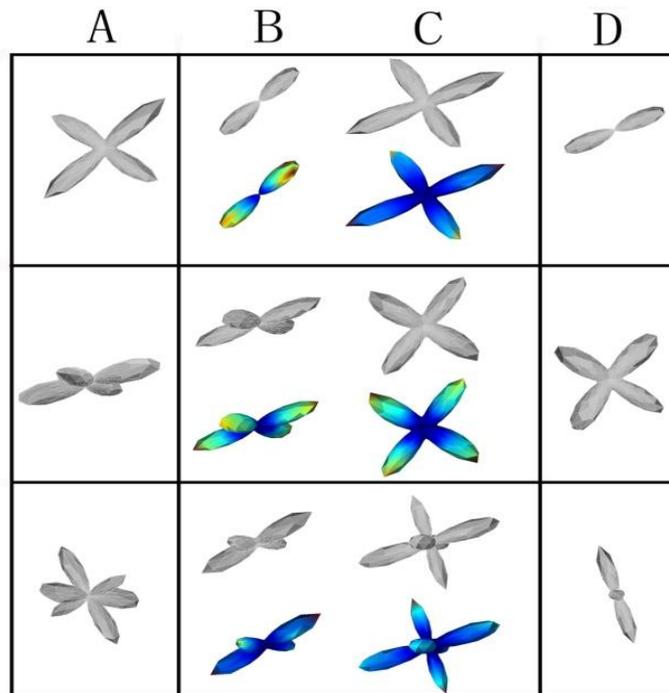
## 3 Experimental Results

### 3.1 ODF Rotation in a Synthetic Example

We studied different registration situations to match ODFs with different numbers of peaks ranging from 1 to 3 peaks. In a prior work [1], it was found that the number of detectable peaks in a voxel generally lies between 1 and 3, with more than one peak in more

than 40% of voxels. A voxel with more than 3 fibers crossing is less plausible for neurobiological reasons, unless the voxels are so large that many tracts are present and in such case a very high angular sampling would be needed to resolve more peaks. To illustrate the simplest case, multiple fibers were created, crossing at 90 degrees with equal volume fractions. Here we chose  $\lambda_1/\lambda_2=5:1$  as the ratio of eigenvalues for each cylinder-shape tensor (FA= 0.77, typical for white matter in the brain) to generate simulated ODFs using discrete mixtures of Gaussian distributions (Eq. 9). The basic idea for ODF registration is to match the dominant directions. Our registrations included matching pairs of ODFs with different numbers of peaks, with results shown in Fig. 2.

$$S(q) = S_0 \cdot \sum_{i=1}^n w_i \cdot e^{-b \cdot q^T \cdot D_i \cdot q} \quad \text{where} \quad \sum_{i=1}^n w_i = 1 \quad (9)$$



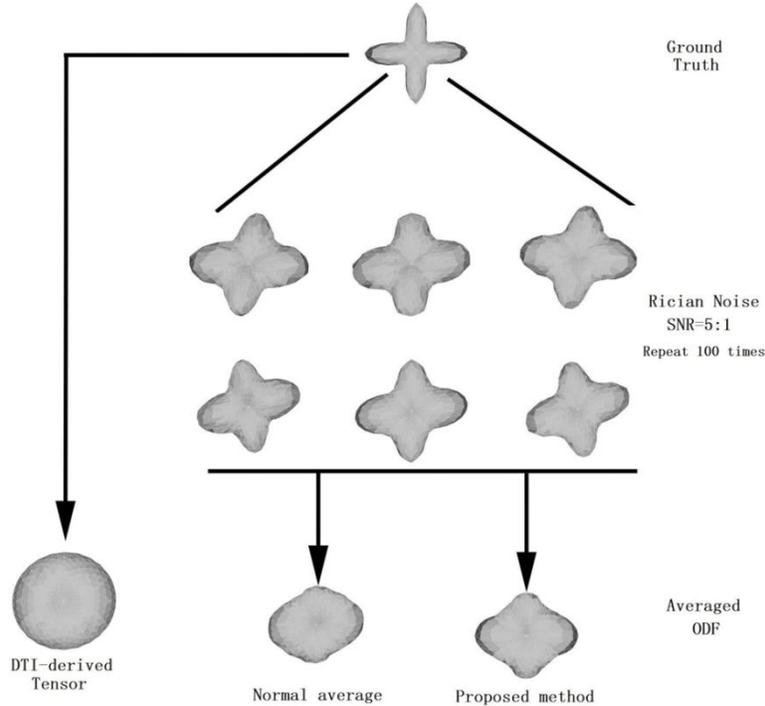
**Fig. 2. Illustration of ODF rotation in the synthetic dataset.** In each row, two ODFs and their corresponding HKS are shown in Columns B and C; column A shows the ODF computed from registering C to B, and column D shows the ODF computed from registering B to C. The first row shows the registration between 1 and 2, the second row for 2 vs. 2 and the last row is for 2 vs. 3.

The symmetric Kullback–Leibler distance (sKL) between the average registered ODFs and the template ODF was used to quantitatively evaluate the proposed method, according to Eq. 10:

$$sKL = \frac{1}{2} \cdot \sum_{i=1}^{642} \left[ p_i * \log \left( \frac{p_i}{q_i} \right) + q_i * \log \left( \frac{q_i}{p_i} \right) \right] \quad (10)$$

In this equation,  $p_i$  and  $q_i$  ( $i=1, 2 \dots 642$ ) are the sampled points on the two ODF surfaces, as described in Section 2.1.

In the simulation experiment (**Fig. 3**), we generated two fibers crossing at 90 degrees, and diffusion-weighted measurements were sampled from 94 directions (same as the protocol used for the human brain experiments in Section 3.2). Rician noise (SNR=5) in was added to each of the diffusion-weighted measurements and the ODF was computed using TDF theory (Eq. 1). We repeated this process 100 times, and we then calculated the averaged ODF with and without the adjustment of our proposed method. The sKL between the original ODF and averaged ODF - without adjustment - is 2.22 times higher than the sKL between original ODF and averaged ODF with the adjustment of our proposed method. This suggests how noise may affect the results, and how our proposed method should improve the statistical results. With the added noise, the averaged ODF – once it has been adjusted with our proposed method – offers a significant advantage in recognizing the dominant directions, compared to using the averaged ODF without adjustment. In the next section, we evaluate our methods on human brain data.

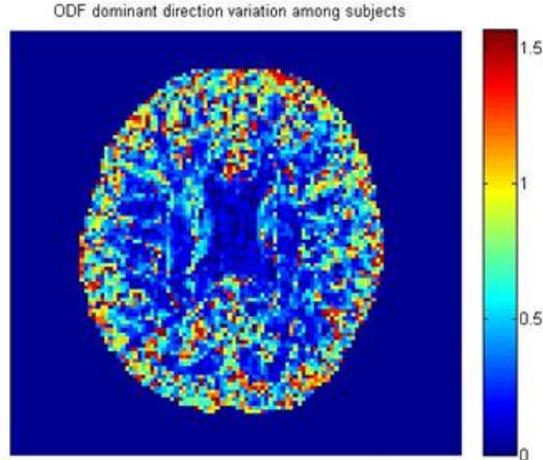


**Fig. 3. Validation of our proposed method on a synthetic dataset.** Here we define the ground truth as two fibers, crossing at 90 degrees. For this simulation, diffusion-weighted measurements were generated in each of 94 directions – the same gradient vector table as is used for our human datasets – and Rician noise (SNR=5) was added. Then the ODF was reconstructed using TDF theory. This process was repeated 100 times and all ODFs were averaged with and without adjustment by the proposed method. Without the alignment of peaks, the average ODF is very close to a sphere, and it is hard to see any evidence of peaks in the resulting ODF at all. The ODF that is created after peak matching is still not sharp – as it results from averaging many very noisy datasets, but it has more readily identifiable peaks. There is an evident advantage in retaining the dominant directions in the averaged ODF.

### 3.2 Real Brain ODF Transformation

Diffusion-weighted scans were acquired from 10 healthy adult participants (mean age:  $23.8 \pm 2.4$  years, 5 male) on a 4T MRI scanner. 105 gradient images were collected including 11 baseline ( $b_0$ ) images with no diffusion sensitization and 94 diffusion-weighted images ( $b$ -value:  $1159 \text{ s/mm}^2$ ) (see [20] for details). Firstly, a non-linear registration method [21] was used to match the overall brain shape. After this pre-registration, TDF-derived ODFs for each subject were computed.

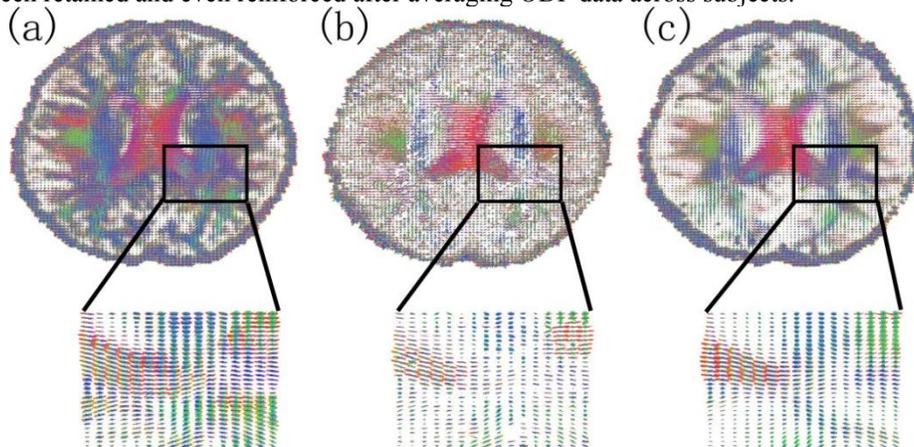
Before describing the adjustment of ODFs using our proposed methods, we first evaluate the necessity of performing ODF adjustment when computing ODF-based statistics. Here, we investigated the variation in the ODFs’ dominant directions, among ten registered subjects [21]. The dominant direction of the ODF was defined as the direction with the maximum ODF surface value. **Fig. 4** shows the maximum dominant direction discrepancy (in units of radians) across ten subjects’ corresponding voxels, for one brain slice. To further quantify the result, we calculated the variability in the dominant direction for the ROI-based mean ODF. In the corpus callosum ROI, the mean variation was 0.1 radian ( $4.6^\circ$ ) while it was as high as 0.6 radian ( $33.8^\circ$ ) in the fiber crossing region of the superior longitudinal fasciculus for ten registered subjects. This enormous degree of variance suggests that the ODF peak adjustment would be beneficial for ODF-based statistics, or else comparable regions would not be averaged together.



**Fig. 4. ODF dominant direction variation across subjects.** This result represents the ODF dominant direction discrepancy among ten registered subjects, in units of radians. The higher the value is, the greater is the disagreement in the ODF dominant directions among registered subjects. If fiber direction vary this much even in registered data, it makes sense to investigate further alignment of ODF peaks, before comparing ODFs across subjects.

Then we chose one subject as the target, we used our method to more accurately register all other subjects’ ODF to the target. (Note that more complex methods could be used to select a target, e.g. the subject with least summed distance to the others, in some appropriate metric). **Fig. 5a** is the template subject’s ODF; **Fig. 5b** shows the directly averaged ODFs across 10 subjects, using standard pointwise averaging of the ODF data across subjects, in spherical coordinates. **Fig. 5c** shows averaged registered ODFs across 10 subjects, after using the methods proposed in this paper. Compared to **Fig. 5a**, **Fig. 5b**’s main fiber tracts are blurred away due to the variance among the subjects, and due to the lack of

a method to align and constructively reinforce them. In **Fig. 5c**, the main fiber tracts have been retained and even reinforced after averaging ODF data across subjects.



**Fig. 5.** ODF plots from individuals, and computation of population averages. **(a)** A typical subject's ODF field - this subject's scan was used as the target for registration, and other subjects' data was aligned to it; **(b)** directly averaged ODF across 10 subjects, without any ODF alignment - note the blurring of the features; **(c)** averaged registered ODF across 10 subjects, using the methods proposed in this paper. The colors represent the dominant fiber direction, mapped as a 3D vector to the RGB space: *red* for left-right, *blue* for superior-inferior, and *green* for anterior-posterior.

## 4 Conclusion and Future Work

Here we proposed a general method - based on the HKS and Möbius transformation - to register fields of spherical functions, such those that arise in analyses of diffusion imaging data from a population. The novelty of our work is twofold. First, we use HKS, which is proportional to Gaussian curvature, to objectively locate peaks; second, we optimize the matching energy between ODF by Möbius voting, to find the best transformation parameters.

Our paper has some limitations. It may not always be meaningful to align ODFs with different numbers of peaks; in this case, going further and explicitly extracting global paths for the fibers may help to resolve ambiguities. This may also help to compute a spatial coherent matching field across voxels. In future, we plan further numerical analysis of the accuracy of the ODF re-orientation at an individual level and of the accuracy when averaging entire ODF fields. Comparative studies with other ODF registration methods (e.g., with fidelity metrics based on information theory) would also be of interest.

Perhaps surprisingly, diffusion imaging has exquisite angular resolution but this is generally thrown away and reduced to a single scalar measure per voxel, prior to cross subject

comparisons and population analyses. Here we show that by matching peaks on ODFs, we can average data in way that constructively reinforces the available features. This effort should be useful in ongoing projects to define characteristic patterns of fiber integrity and geometry in disease, as well as changes with disease progression and over the human lifespan. In addition, Möbius transformations can do more than just rotate the full ODF, and could be part of future Möbius-driven global nonlinear registration, using a cost function derived from the local Möbius transformation to drive a global 3D image flow and then apply the image flow back to the ODF. Here we show HKS and Möbius transformation can reliably used for local maxima detection and matching in ODF registration.

## References

1. Zhan L, Leow AD, Zhu S, Baryshev M, Toga AW, McMahon KL, de Zubicaray GI, Wright MJ, Thompson PM. : A novel measure of fractional anisotropy based on the tensor distribution function. *Med Image Comput Comput Assist Interv.* 12:845-852 (2009)
2. Leow, A.D., Zhan, L., Zhu, S.W., Hageman, N.S., Chiang, M.C., Barysheva, M., Toga, A.W., McMahon, K., de Zubicaray, G.I., Wright, M.J., Thompson, P.M.: White matter integrity measured by fractional anisotropy correlates poorly with actual individual fiber anisotropy. In *Proceedings of the 6<sup>th</sup> IEEE international Symposium on Biomedical Imaging: From Nano to Macro, ISBI2009*, 622-625 (2009).
3. Lee, A.D., Lepore, N., Brun, C., Chou, Y.Y., Barysheva, M., Chiang, M.C., Madsen, S.K., de Zubicaray, G.I., McMahon, K.L., Wright, M.J., Toga, A.W., Thompson, P.M.: Tensor-based analysis of genetic influences on brain integrity using DTI in 100 twins. *Med Image Comput Comput Assist Interv*, 12:967-974 (2009).
4. Jin, Y., Shi, Y.G., Jahanshad, N., Aganj, I., Sapiro, G., Toga, A.W., Thompson, P.M.: 3D elastic registration improves HARDI-derived fiber alignment and automated tract clustering. In *Proceedings of the 8<sup>th</sup> IEEE international Symposium on Biomedical Imaging: From Nano to Macro, ISBI2011*, 822-826 (2011).
5. Barmptoutis, A., Hwang, M.S., Howland, D., Forder, J.R., Vemuri, B.C.: Regularized positive-definite fourth order tensor field estimation from DW-MRI. *Neuroimage.* 45: 153-162 (2009).
6. Geng, X., Ross, T.J., Zhan, W., Gu, H., Chao, Y.P., Lin, C.P., Christensen, G.E., Schuff, N., Yang, Y.: Diffusion MRI registration using orientation distribution functions. *Inf Process Med Imaging*, 21:626–637 (2009).
7. Cheng, G., Vemuri, B.C., Carney, P.R., Mareci, T.H.: Non-rigid Registration of High Angular Resolution Diffusion Images Represented by Gaussian Mixture Fields. *Med Image Comput Comput Assist Interv*, 5761: 190-197 (2009).
8. Hong, X., Arlinghaus, L.R., Anderson, A.W.: Spatial normalization of the fiber orientation distribution based on high angular resolution diffusion imaging data. *Magn Reson Med.* 61:1520–1527 (2009).
9. Bloy, L., Verma, R.: Demons registration of high angular resolution diffusion images. *ISBI* 1013–1016 (2010).

10. Yap, P.T., Chen, Y., An, H., Yang, Y., Gilmore, J.H., Lin, W., Shen, D.: SPHERE: SPHERical Harmonic Elastic REgistration of HARDI data. *Neuroimage*. 55:545–556 (2011).
11. Goh, A., Lenglet, C., Thompson, P.M., Vidal, R.: A non-parametric Riemannian framework for processing high angular resolution diffusion images and its applications to ODF-based morphometry. *Neuroimage*. 56:1181–1201 (2011).
12. Chiang, M.C., Leow, A.D., Klunder, A.D., Dutton, R.A., Barysheva, M., Rose, S.E., McMahon, K.L., de Zubicaray, G.I., Toga, A.W., Thompson, P.M.: Fluid registration of diffusion tensor images using information theory. *IEEE Trans Med Imaging*, 27:442–456 (2008).
13. Leow, A.D., Zhu, S., Zhan, L., McMahon, K., de Zubicaray, G.I., Meredith, M., Wright, M.J., Toga, A.W., Thompson, P.M.: The tensor distribution function. *Magn Reson Med*. 61:205–214. (2009).
14. Jian, B., Vemuri, B.C.: A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted MRI. *IEEE Trans Med Imaging* 26:1464–1471. (2007).
15. 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*. 64:554–566. (2010).
16. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum* 28(5):1383–1392. (2009).
17. Hsu, E.P. *Stochastic Analysis on Manifolds*. American Mathematical Society (2002).
18. Gu, X.F., Wang, Y.L., Chan, T.F., Thompson, P.M., Yau S.T.: Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Trans. Med. Imag.*, 23(8):949–958 (2004).
19. Lipman, Y., Funkhouser, T.: Möbius voting for surface correspondence. *SIGGRAPH*, 72:1–12. (2009).
20. Zhan, L., Leow, A.D., Jahanshad, N., Chiang, M.C., Barysheva, M., Lee, A.D., Toga, A.W., McMahon, K.L., de Zubicaray, G.I., Wright, M.J., Thompson, P.M. : How does angular resolution affect diffusion imaging measures? *NeuroImage*. 49 (2):1357-1371. (2010).
21. Jahanshad, N., Lee, A.D., Barysheva, M., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Wright, M.J., Toga, A.W., Thompson, P.M.: Genetics Influences on Brain Asymmetry: A DTI Study of 374 Twins and Siblings. *Neuroimage* 15; 52(2):455-69 (2010).

# Detection of Unique Point Landmarks in HARDI Images of the Human Brain

Henrik Skibbe and Marco Reisert

Department of Radiology, University Medical Center Freiburg, Germany  
{henrik.skibbe, marco.reisert}@uniklinik-freiburg.de

**Abstract.** Modern clinical image acquisition techniques like diffusion magnetic resonance imaging (dMRI) or functional MRI (fMRI) allow us to study tissue and organisms in their natural volumetric configuration. Group studies often require the co-registration of images or partial image structures of different individuals. In such applications the detection of characteristic landmarks is often an indispensable prerequisite. However, the reliable detection of unique anatomical landmarks in medical images is a challenging problem which is often solved in a semi-automated or even manually manner. Landmarks are often sought to be located and analyzed at every position, and in any orientation. In particular the latter makes landmark detection challenging in 3D. In this paper we propose a new landmark detection system that reliably detects arbitrarily placed landmarks in High Angular Resolution Diffusion Images (HARDI) of the human brain.

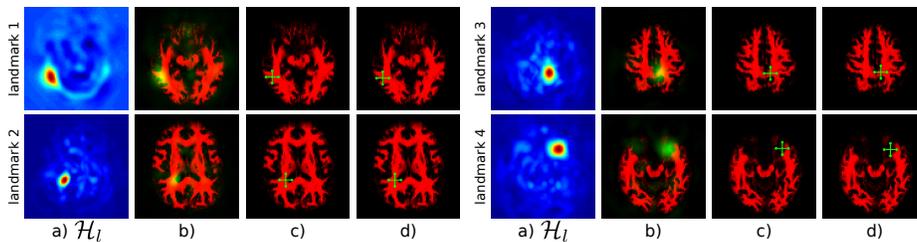
## 1 Introduction

The study of Magnetic Resonance (MR) imaging modalities is of great interest in fundamental neuroscience and medicine. MR imaging offers a wide range of different contrasts, providing scientists insights into anatomical and functional properties of the human brain. In this paper we focus on High Angular Resolution Diffusion Images (HARDI [8]) of the human brain. The HARDI-technique combines different measurement parameters to infer underlying tissue properties and allows for studying the neuronal fiber architecture in the human brain without harming the patient.

In this paper we introduce a new approach for detecting unique point landmarks in HARDI images of the human brain. The reliable detection of landmarks [1] plays an indispensable role in registering brain structures from different images and thus is an important prerequisite for many registration and segmentation algorithms [3]. Our approach is based on the computation of a dense feature map of the HARDI signal. Similar to [9], where features are used to find correspondences in scalar valued MR contrasts, we propose features offering a unique signature of a voxel’s surrounding in tensor-valued HARDI signals. Thanks to these features a large number of corresponding points can be reliably found in images of different individuals using a linear classifier. The parameters for the linear classifier are learned from a training set of landmarked images.

The challenges in our scenario are manifold: the HARDI technique leads to images with poor quality in terms of resolution and signal to noise ratio. Furthermore, the tensor valued HARDI signal can be considered as function  $\mathbf{f} : \mathbb{R}^3 \times S_2 \rightarrow \mathbb{R}$ , where  $S_2$  denotes the unit sphere in  $\mathbb{R}^3$ . This means at each voxel position  $\mathbf{x} \in \mathbb{R}^3$  we have an angular dependent measurement  $\mathbf{f}(\mathbf{x}, \mathbf{n})$  represented as a function on the unit sphere. This function represents the diffusion weighted MR signal with respect to different diffusion directions  $\mathbf{n}$ . Due to this fact we cannot use intensity based standard techniques to describe a landmark; suitable representations of the signal are required in order to uniquely detect landmarks within the images.

The contributions of the present paper are the following: (1) we derive new rotation invariant features for HARDI images of the human brain. These features are yielding a unique description of each voxel in the images which allows for reliably detecting landmarks within images of different individuals. Positioning landmarks is not restricted to specific areas or points; landmarks can be positioned anywhere in the brain. (2) We propose a linear coarse to fine classification scheme to detect a large number (more than several thousands) of different, unique landmarks in reasonable time. (3) We demonstrate the effectiveness of our method in an experiment based on a dataset of brain images of 21 healthy volunteers. Furthermore, preliminary results on datasets of patients with pathologies are promising. (4) The source code will be made publicly available upon acceptance of this paper.



**Fig. 1.** a) Differently weighted linear combinations of the feature images lead to different detection results (Maximum intensity projection). b) The outcome of the linear classifier at the desired landmark position (green) together with the white matter mask (red). c) Position of the global maximum of  $\mathcal{H}_l$ . d) The landmark reference position.

## 2 Landmark Detection in HARDI Images

Our landmark detection method is based on local, rotation invariant feature images. The idea is that corresponding to each voxel position  $\mathbf{x} \in \mathbb{R}^3$  in the HARDI signal there exists a feature vector  $\mathbf{F}(\mathbf{x}) := (F_1(\mathbf{x}), F_2(\mathbf{x}), \dots, F_N(\mathbf{x}))^T \in \mathbb{R}^N$  that uniquely describes the appearance of a voxel's surrounding. This includes

features describing the surrounding of a voxel in a coarser scale, thus the rough position of the voxel in the brain can be determined. On the other hand, the feature vector comprises features describing finer details of the close neighborhood of a voxel. Our experiments will show that the resulting features are highly discriminative. Such feature images of a HARDI signal of a human brain are shown in figure 2. Details are given later in this section.

Our assumption is that the features guarantee a unique representation of each landmark. Hence we can use the maximum response of a linear classifier to determine the landmark positions. This is probably the fastest way to detect the landmarks. In a supervised training step we learn weights for the linear classifier for all landmarks  $l = \{1, \dots, M\}$ . We denote by  $\mathcal{H}_l \in \mathbb{R}^3 \rightarrow \mathbb{R}$  the evidence image for the position of landmark  $l$ :

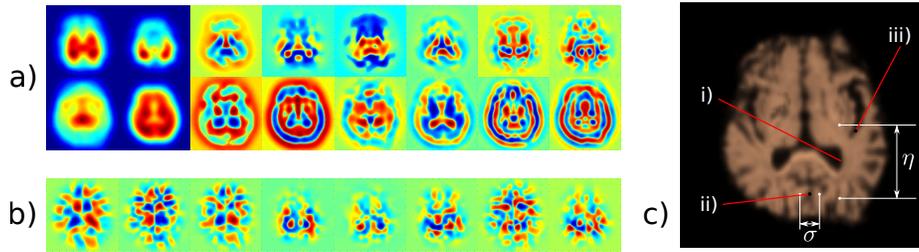
$$\mathcal{H}_l(\mathbf{x}) := \sum_i \alpha_i(l) F_i(\mathbf{x}) = \boldsymbol{\alpha}(l)^T \mathbf{F}(\mathbf{x}) \quad . \quad (1)$$

See figure 1 a) for some examples of  $\mathcal{H}_l$  for different landmarks  $l$ . With  $\boldsymbol{\alpha}(l) \in \mathbb{R}^N$  we denote the weights corresponding to landmark  $l$ . The global maximum of  $\mathcal{H}_l$  is considered as the prediction for a landmark position. It is worth noting that due to the fact that eq. (1) is linear we can use a simple least square fit  $\operatorname{argmin}_{\boldsymbol{\alpha}(l)} \|\sum_{\mathbf{x}} \mathcal{H}_l^{\text{train}}(\mathbf{x}) - \boldsymbol{\alpha}(l)^T \mathbf{F}^{\text{train}}(\mathbf{x})\|^2$  to compute the weights in a training step.  $\mathbf{F}^{\text{train}}$  are features of training images and  $\mathcal{H}_l^{\text{train}}$  are binary valued label images where the desired landmark position has been set to 1. Note that due to the sparseness of  $\mathcal{H}_l^{\text{train}}$  the system of equations can be solved for all landmarks simultaneously in a memory efficient way within seconds. Once the weights are determined we can find the landmarks in any new image by computing the features and attaching the weights according to eq. (1).

In our experiment we aim at detecting more than  $10^4$  landmarks. Successively computing the evidence images  $\mathcal{H}_l$  followed by the determination of the global maximum for each single landmark is far too computational expensive. Since the landmarks are unique we can determine the predicted landmark position in two steps: 1) We down-sample the feature images by a factor of 4 (we just consider every fourth voxel). We then compute the evidence  $\mathcal{H}_l(\mathbf{x})$  for each landmark voxel by voxel “on the fly” and store the position of the highest result for each landmark. 2) We then search for the precise location of the maxima in the surrounding of each previously stored position. Given the feature images we can find about 22000 landmarks in a  $100 \times 100 \times 69$  brain image in about 5 minutes<sup>1</sup>.

**Computing the feature images  $\mathbf{F}$ :** We utilize spherical tensor algebra (STA) [4, 5], which allows to compute dense, multi-scale feature images from HARDI data in an efficient and rotation invariant way. The use of STA is quite reasonable, because it is common to represent HARDI images by Spherical Harmonics (SH). The orientation distribution in each voxel is encoded in the SH basis providing a memory efficient and smooth way to handle the data.

<sup>1</sup> using 4 cores of an Intel Xeon CPU X7560 with 2.27GHz



**Fig. 2.** a) Rotation invariant features according to eq. (4) of a human brain (center slice). They are invariant with respect to reflection symmetry, too. b) We additionally use features that can clearly distinguish between the left and right hemisphere (the sign differs, eq. (5)) c) Features for a landmark i) are computed in two steps: We first densely compute local nonlinear image features ii). Based on these features in a larger neighborhood iii) around i) we form the final rotation invariant features.

We first decompose the HARDI signal  $\mathbf{f} : \mathbb{R}^3 \times S_2 \rightarrow \mathbb{R}$  into its basic angular frequency components by orthogonal projection onto the SH basis functions  $\mathbf{Y}^\ell : S_2 \rightarrow \mathbb{C}^{2\ell+1}$  voxel by voxel [4]. With each index  $\ell$  a certain angular frequency is represented. Since diffusion is symmetric, the HARDI signals are symmetric, too. Hence we only need to consider SH associated with an even index. Consequently we can represent  $\mathbf{f}$  in terms of  $\mathbf{Y}^\ell$  by  $\mathbf{f}(\mathbf{x}, \mathbf{n}) = \sum_{\ell \text{ even}}^{\infty} \mathbf{a}^\ell(\mathbf{x})^T \mathbf{Y}^\ell(\mathbf{n})$ . We denote by  $\mathbf{a}^\ell(\mathbf{x}) \in \mathbb{C}^{2\ell+1}$  the vector valued expansion coefficients representing the HARDI signal of  $\mathbf{f}$  at image position  $\mathbf{x} \in \mathbb{R}^3$  in the SH-domain.

In our framework we only use the coefficient images  $\mathbf{a}^0 : \mathbb{R}^3 \rightarrow \mathbb{C}$  and  $\mathbf{a}^2 : \mathbb{R}^3 \rightarrow \mathbb{C}^5$ .  $\mathbf{a}^0$  represents the mean of the signal.  $\mathbf{a}^2$  gives us information about the diffusion directions and heavily contributes in the white matter regions to the HARDI signal. Considering higher frequency components, i.e.  $\mathbf{a}^4$  in our experiments did not lead to better results.

The raw coefficients  $\mathbf{a}^0$  and  $\mathbf{a}^2$  are only describing the very local properties of the tissue and thus are far not sufficient to yield enough information to uniquely represent a landmark in the HARDI signal. Due to this reason we designed new nonlinear features for representing the neighborhood around a respective voxel. The resulting features are rotation invariant thus no pre-alignment of the images is required. The features are computed in two steps: First, non-linear local image features are densely computed in a close neighborhood of a voxel (Fig. 2 ii). A second step combines the non-linear features in a larger surrounding of a voxel and forms its unique feature signature (Fig. 2 iii).

STA provides two basic operations to deal with the coefficient images  $\mathbf{a}^0$  and  $\mathbf{a}^2$ . These operations do not alter the rotation behavior of the SH representation, that is, they allow to compute rotation invariant features in a systematic way. The first class of operations are finite difference operators that connect SH representations of different degrees by differentiations, so-called *spherical tensor derivatives*  $\nabla^n$  [5]. We distinguish between spherical up-derivatives, where  $n > 0$  and spherical down-derivatives where  $n < 0$ . The first operator increases

the tensor rank by  $n$ , the latter one decreases the tensor rank by  $n$ . The second class of operations are products that connect two different SH representations to form a new field with a different degree, called *spherical products*  $\circ_\ell : \mathbb{C}^{2\ell_1+1} \times \mathbb{C}^{2\ell_2+1} \rightarrow \mathbb{C}^{2\ell+1}$  [5]. They couple *spherical tensors* associated with different *orders*  $\ell_1, \ell_2$  to form new tensors of higher or lower order  $\ell$ .

We obtain the local non-linear image descriptors (Fig. 2 ii) ) in the following way: We first expand the local neighborhood of voxels in  $\mathbf{a}^0$  and  $\mathbf{a}^2$  in terms of spherical Gaussian derivatives by initially convolving the coefficient images with a Gaussian  $G_\sigma$  followed by successively computing the tensor derivatives [7]:

$$\mathbf{b}_0^\ell := \nabla^\ell (\mathcal{G}_\sigma * \mathbf{a}^0), \quad 0 \leq \ell < L, \quad \mathbf{b}_0^\ell(\mathbf{x}) \in \mathbb{C}^{2\ell+1} \quad (2)$$

$$\mathbf{b}_2^\ell := \nabla^{\ell-2} (\mathcal{G}_\sigma * \mathbf{a}^2), \quad 0 \leq \ell < L, \quad \mathbf{b}_2^\ell(\mathbf{x}) \in \mathbb{C}^{2(2+\ell)+1} \quad (3)$$

We denote by  $\mathbf{b}_0^\ell(\mathbf{x}), \mathbf{b}_2^\ell(\mathbf{x})$  the expansion coefficients of the local neighborhood of  $\mathbf{a}^0(\mathbf{x})$  and  $\mathbf{a}^2(\mathbf{x})$ , respectively. The upper bound  $L \in \mathbb{N}$  restricts the number of expansion coefficients. The size of the local neighborhood representations is defines by  $\sigma$ , the size of the Gaussian. Then we use the spherical tensor products  $\circ_\ell$  to form new, nonlinear representations  $(\mathbf{b}_a^{\ell_1}(\mathbf{x}) \circ_\ell \mathbf{b}_a^{\ell_2}(\mathbf{x})) \in \mathbb{C}^{2\ell+1}$  voxel by voxel.

Finally, we follow ideas proposed by the Harmonic Filter framework (HF) [5] to form the final rotation invariant large neighborhood descriptors (Fig. 2 iii) ). The HF is some kind of voting based approach for generic object detection where local image descriptors are voting for the presence of objects. Voting offers several advantages: The detection of objects is very robust with respect to occlusions, intra-class variations and deformations! In our framework we adopt the idea of voting and consider it as a collection of local descriptors in a voxel's surrounding. Mathematically this step coincides with the voting of the HF thus we gain rotation invariant features in the following way:

$$F_i(\mathbf{x}) := \mathcal{G}_\eta * (\nabla^{(-\ell)}(\mathbf{b}_a^{\ell_1} \circ_\ell \mathbf{b}_a^{\ell_2})) \quad \ell \leq L \quad (4)$$

A larger choice of  $\eta$  leads to image descriptors representing the rough position of the voxel in the brain. As small  $\eta$  leads to features representing local details of the HARDI signal. Figure 2 a) shows some examples of such features based on the HARDI signal of a human brain. Note that  $L \in \mathbb{N}$  restricts the number of possible products i.e. the number of feature images.

Considering our aims there exists one significant drawback of these features: they are invariant against reflection about an axis. Hence they can't distinguish the left and the right hemisphere. Figure 2 a) illustrates this problem. It is known that the spherical triple-correlation [2] yields complete rotation invariant features. Hence they must solve this issue. Based on this idea we designed new  $3^{rd}$  order rotation invariant differential features fitting into our framework that are variant with respect to reflections about an axis:

$$F_j(\mathbf{x}) := \mathcal{G}_\eta * (\nabla^{(-\ell_4)}((\mathbf{b}_a^{\ell_1} \circ_\ell \mathbf{b}_a^{\ell_2})) \circ_{\ell_4} \mathbf{b}_a^{\ell_3}), \quad \begin{array}{l} \ell_1 + \ell_2 + \ell_3 + \ell_4 \text{ is odd} \\ \text{and } \ell_4, \ell \leq L \end{array} \quad (5)$$

The proof is given in the appendix. Figure 2 b) shows some examples.

**Experiments** For our experiments 21 in vivo diffusion acquisitions of human brains were acquired on a Siemens 3T TIM Trio using an SE EPI sequence with a TE of 95 ms and a TR of 8.5 s and an effective b-value of 1000. One voxel corresponds to  $2mm^3$ . We used 67 directions which we entirely fit to spherical harmonics. For evaluating the performance of our detection system we conducted the following experiment: The 21 HARDI images of healthy volunteers have been co-registered using the SPM<sup>2</sup> toolkit. Based on this co-registration we found the position of 20685 landmarks (per image) in the original image domain. The landmarks were densely distributed in the brain white and gray matter (one landmark at every second voxel position with respect to X,Y, and Z direction). We consider these co-registered landmarks as a reference which we use for training the classifier and for evaluating our approach. It is worth noting that it is difficult to co-register the noisy, tensor-valued HARDI images thus that the true “ground-truth” can hardly be provided. But if the positions of the landmarks detected by our algorithm are similar to the co-registered positions, than there is high evidence that most of the landmarks have been detected correctly.

The HARDI images have been transformed to the SH-domain and features have been computed for each image as described above. We first computed the expansion coefficients based on eq. (2) and eq. (3) using a Gaussian with  $\sigma = 6mm$ . We experienced that the signal corresponding to the brain white matter leads to the most reliable features thus we additionally expanded  $\mathbf{a}^2$  with respect to a larger neighborhood  $\sigma = 5mm$  (eq. (3)). We then computed features based on eq. (4) and eq. (5) using  $L = 5$  and three different scales, namely with respect to  $\eta = 4, 8$  and  $12mm$ . We found these parameters via a leave-one-out parameter grid-search on the training set. Using these parameters we gained 528 discriminative feature images per HARDI signal.

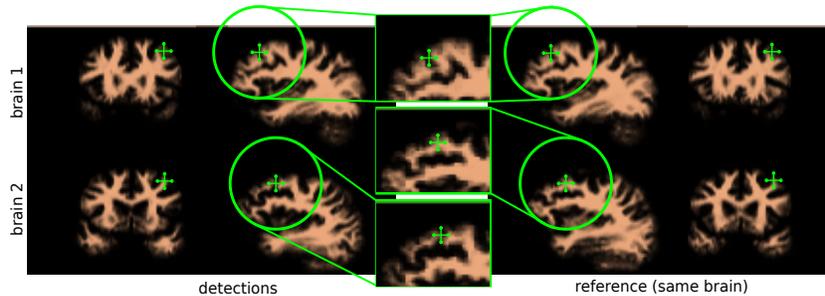
We used one third of the images (7 images) for training the linear classifier (eq. 1). The co-registered landmark positions were used for determining the filter parameters  $\alpha(l)$  for all landmarks. The remaining 14 images were used for evaluation.

**Results & Discussion** The displacements of the filter responses with respect to the co-registered reference positions can be found in table 1. The results clearly show that most of the detected landmark positions are very close to the ground

<sup>2</sup> SPM (Statistical Parametric Mapping version 5), <http://www.fil.ion.ucl.ac.uk/spm/>

**Table 1.** Correctly detected landmarks. Correctly means the detected landmark position is similar to its reference position. The column in red corresponds to the worst results. The table shows results for an increasing tolerated displacement.

tolerance	correctly detected landmarks for the 14 test datasets (totally 20685 unique landmarks)														
1 voxel	39.0%	26.6%	39.2%	36.5%	51.9%	40.4%	37.2%	51.0%	38.7%	49.4%	46.1%	38.7%	51.8%	44.7%	
2 voxel	81.5%	65.7%	79.8%	77.0%	90.1%	81.1%	79.3%	89.3%	81.0%	88.6%	86.6%	79.1%	88.5%	85.8%	
3 voxel	96.9%	90.7%	96.3%	95.4%	98.9%	96.7%	96.1%	98.5%	95.8%	97.5%	98.7%	95.7%	98.2%	97.8%	
4 voxel	99.4%	96.1%	99.4%	98.7%	99.8%	99.4%	99.0%	99.6%	98.1%	99.1%	99.8%	98.3%	99.4%	99.6%	
5 voxel	99.9%	97.9%	99.9%	99.6%	99.9%	99.9%	99.7%	99.9%	99.0%	99.4%	99.9%	99.3%	99.7%	100%	



**Fig. 3.** The white matter mask of the HARDI signal [6] together with the landmark positions. We counted a detection as successful if the predicted landmark position was close to its reference position. Since the filter encodes the appearance of local structures in several granularities it can use coarser representations to find the global position in the brain while finer representations are used to adapt the position according to the local neighborhood configuration. This often shows more plausible results than the reference positions as exemplarily shown in the second row. the last row).

truth. Only a very small number (in the worst case  $< 2.1\%$ ) of the landmark positions differ by more than 5 voxel, which corresponds to a displacement of  $1cm$ .

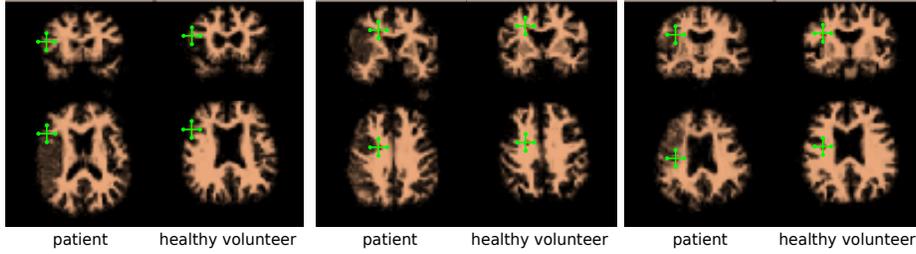
For a qualitative analysis of the results we computed a white matter mask directly from the HARDI signal [6]. This ensures consistency with the data. Figure 3 gives qualitative results. We observed that the detected landmarks were often more consistent with the local structure of the HARDI images than the co-registered reference location (Fig. 3). The detection of all 20685 landmarks takes about 16 minutes<sup>1</sup> per image.

We further conducted experiments based on five images of patients showing pathologies. Since no ground truth was available we visually compared the detection results of 50 landmarks with their position in images of healthy volunteers. Our approach was able to successfully detect all landmarks in the healthy areas of the brain. It is worth noting that we also detected the landmarks located very close to pathological areas (Fig. 4).

### 3 Conclusion

In this paper, we have presented a new framework that allows the detection of a large number of unique point landmarks within the tensor valued HARDI images of the human brain. Our experiment has shown that based on new image features in combination with a fast linear classifier the landmarks can be reliably detected in reasonable time. We make the source code publicly available after acceptance of this paper.

**Appendix** Eq. (5) is variant with respect to reflection about an axis: We consider w.o.l.g the reflection about the origin. It holds that  $\mathbf{Y}^\ell(-\mathbf{n}) = (-1)^\ell \mathbf{Y}^\ell(\mathbf{n})$ .



**Fig. 4.** Detected landmarks in patients with pathologies and healthy volunteers. Although the images strongly differ, the landmarks have been detected correctly in the healthy areas of the patients images.

Let  $f(\mathbf{r}, \mathbf{n}) : \mathbb{R}^3 \times S_2 \rightarrow \mathbb{R}$ . When  $f(\mathbf{r}, \mathbf{n}) = \sum_{\ell} (\mathbf{a}^{\ell}(\mathbf{r}))^T \mathbf{Y}^{\ell}(\mathbf{n})$  we have  $f'(\mathbf{r}, \mathbf{n}) = f(-\mathbf{r}, -\mathbf{n}) = \sum_{\ell} (\mathbf{b}^{\ell}(-\mathbf{r}))^T \mathbf{Y}^{\ell}(\mathbf{n})$ , with  $\mathbf{b}^{\ell}(-\mathbf{r}) = (-1)^{\ell} \mathbf{a}^{\ell}(\mathbf{r})$ . If  $\ell_1 + \ell_2 + \ell_3$  is odd, then  $((\mathbf{b}^{\ell_1}(-\mathbf{r}) \circ_{\ell_1} \mathbf{b}^{\ell_2}(-\mathbf{r})) \circ_{\ell_4} \mathbf{b}^{\ell_3}(-\mathbf{r})) = (-1)^{\ell_1 + \ell_2 + \ell_3} ((\mathbf{a}^{\ell_1}(\mathbf{r}) \circ_{\ell_1} \mathbf{a}^{\ell_2}(\mathbf{r})) \circ_{\ell_4} \mathbf{a}^{\ell_3}(\mathbf{r}))$ . With  $(\nabla^{(-\ell)} \mathbf{b}^{\ell})(-\mathbf{r}) = (-1)^{\ell} (\nabla^{(-\ell)} \mathbf{a}^{\ell})(\mathbf{r})$  we can conclude that if  $\ell_1 + \ell_2 + \ell_3 + \ell_4$  is odd, then  $(\nabla^{(-\ell_4)} ((\mathbf{b}^{\ell_1} \circ_{\ell_1} \mathbf{b}^{\ell_2}) \circ_{\ell_4} \mathbf{b}^{\ell_3}))(-\mathbf{r}) = -(\nabla^{(-\ell_4)} ((\mathbf{a}^{\ell_1} \circ_{\ell_1} \mathbf{a}^{\ell_2}) \circ_{\ell_4} \mathbf{a}^{\ell_3}))(\mathbf{r})$ .

## References

1. Jimin Liu, Wenpeng Gao, Su Huang, and W. L. Nowinski. A Model-Based, Semi-Global Segmentation Approach for Automatic 3-D Point Landmark Localization in Neuroimages. *IEEE Trans. Med. Imaging*, 27(8):1034–1044, 2008.
2. R.Kondor. *Group theoretical methods in machine learning*. PhD thesis, Columbia University, 2008.
3. K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn. Landmark-based elastic registration using approximating thin-plate splines. *IEEE Trans. Med. Imaging*, 20(6):526–534, jun 2001.
4. M. Rose. *Elementary Theory of Angular Momentum*. Dover Publications, 1995.
5. M Schlachter, M Reisert, C Herz, F Schluermann, S Lassmann, M Werner, H Burkhardt, and O Ronneberger. Harmonic Filters for 3D Multi-Channel Data: Rotation Invariant Detection of Mitoses in Colorectal Cancer. *IEEE Trans. Med. Imaging*, 29(8):1485–1495, 2010.
6. H. Skibbe, M. Reisert, and H. Burkhardt. Gaussian neighborhood descriptors for brain segmentation. In *Proc. of the MVA*, Nara, Japan, 2011.
7. H. Skibbe, M. Reisert, T. Schmidt, T. Brox, O. Ronneberger, and H. Burkhardt. Fast Rotation Invariant 3D Feature Computation utilizing Efficient Local Neighborhood Operators. *IEEE Trans. on PAMI*, 2012.
8. DS. Tuch, RM. Weisskoff, JW. Belliveau, and VJ. Wedeen. High angular resolution diffusion imaging of the human brain. In *Proc. of the ISMRM*, Philadelphia, USA, 1999.
9. Z. Xue, D. Shen, and C. Davatzikos. Determining correspondence in 3-D MR brain images using attribute vectors as morphological signatures of voxels. *IEEE Trans. Med. Imaging*, 23(10):1276–1291, October 2004.

## Oral Session III: Diffusion Modeling and Experimental Design

# Generalized Constant Solid Angle ODF and Optimal Acquisition Protocol for Fiber Orientation Mapping

Amith Kamath<sup>1</sup>, Iman Aganj<sup>3,4</sup>, Junqian Xu<sup>2</sup>, Essa Yacoub<sup>2</sup>, Kamil Ugurbil<sup>2</sup>,  
Guillermo Sapiro<sup>5</sup>, and Christophe Lenglet<sup>2</sup>

<sup>1</sup> Electrical and Computer Engineering, University of Minnesota.

<sup>2</sup> CMRR, Department of Radiology, University of Minnesota.

<sup>3</sup> Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School.

<sup>4</sup> Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

<sup>5</sup> Electrical and Computer Engineering, Duke University.

{kama0149@umn.edu}

**Abstract.** Recent advances in diffusion MRI have allowed for improved understanding of the white matter connectivity. Models like the Diffusion Tensor, diffusion Orientation Distribution Function (ODF) with a mono-exponential signal decay have shown good fiber reconstruction accuracies. More complex radial signal decay models, like the bi-exponential model, have been shown to better approximate the in-vivo diffusion signal. In this paper, we generalize the Constant Solid Angle ODF (CSA-ODF) algorithm to handle any  $q$ -space sampling and exploit the bi-exponential model. Simulation results to optimize the reconstruction and acquisition parameters are described. Finally, the algorithm is validated on human brain data. Our generalized CSA-ODF model performs optimally with 200  $q$ -space data points distributed over three shells acquired at  $b = 1000, 2000s/mm^2$  and in the range  $[3000, 6000]s/mm^2$  for the third shell. Crossings up to about 30 degrees can be recovered, and fiber orientations can be detected with a precision of about 1 degree.

## 1 Introduction

In the past decade, there has been significant interest in diffusion Weighted Magnetic Resonance (MR) Imaging for applications related to brain connectivity mapping [1,2]. This has been facilitated by faster MR acquisitions and improved signal modeling. The fiber orientation can be approximated using models like Diffusion Tensor Imaging (DTI) [3], Ball and Stick model [4], Diffusion Orientation Transform (DOT) [5], diffusion ODF [6] based models, Constrained Spherical Deconvolution (CSD) [7] and compartment-based models as in Panagiotaki et al. [8]. Aganj et al [9] introduced a correction of the original ODF model [6], called the CSA-ODF model, which implicitly sharpens the peaks, and is mathematically accurate. In this case, the fiber orientations are estimated using maxima extraction [10, 11] methods.

The utility of acquiring diffusion-weighted data over multiple  $q$ -space shells has been discussed previously by many groups [9, 12–16]. Diffusion Spectrum Imaging (DSI) [17] achieves a similar purpose by sampling  $q$ -space on a uniform 3D Cartesian grid. Current  $q$ -ball methods typically rely on sampling diffusion signals on a single shell in  $q$ -space, based on which a radial decay model can be estimated. This places a restriction on the types of signal decay models that can be used. Acquiring multiple signal values along a single gradient direction allows estimating more complex decay models like the bi-exponential model, which is known to be more accurate [18], but open new computational challenges.

**Contributions:** We explore an approach where the acquisition is done over a finite set of  $b$ -values in *staggered* gradient directions, so as to allow a better angular resolution, along with providing radial information of the diffusion signal decay. Given a finite number of measurements on a particular shell, we investigate how to accurately interpolate such spherical data to be able to retrieve samples in any orientation. Computationally efficient methods in estimating the ODF have been explored in detail in [19]. An intelligent choice of Spherical Harmonic (SH) bases allows for a linear and regularized algorithm. We further refine previous investigations on the optimal choice of regularization for SH fitting.

The CSA-ODF model [9] has been shown to perform better than “artificially” sharpened models, and this work focuses on extending this method to support generic sampling of  $q$ -space. We study the optimal set of  $b$ -values for the multi-shell acquisition for this new generalized CSA-ODF algorithm. We also investigate the minimum number of gradient directions to use in order to minimize distortions on ODF reconstruction and uncertainty on fiber estimation. We finally investigate the minimum achievable separating angle between fibers and the uncertainty on the localization of fiber orientation with the proposed acquisition and reconstruction protocol.

The rest of the paper is organized as follows. A short overview of the CSA-ODF model is presented in Section 2. We then describe the extension of this algorithm to generic  $q$ -space sampling. Section 3 includes extensive simulation studies and results. This is followed by results to validate the algorithm on human brain data.

## 2 Methods

### 2.1 $q$ -ball Estimation

The  $q$ -ball method [6] estimates fiber orientations from high angular resolution diffusion imaging (HARDI) data.  $q$ -ball imaging is a natural choice for fiber orientation estimates, as the data is acquired on a spherical shell with different orientations rather than on a grid in  $q$ -space. Using more complex radial decay models can provide a better understanding of the diffusion process and hence the axonal configuration of the white matter.

A modified SH basis was used in [19–21] for representing spherical distributions. This leads to an analytic computation of the Funk Radon Transform (FRT) where the ODF is obtained as a linear combination of SH bases. The CSA-ODF model [9] corrects a mathematical inaccuracy, due to which the ODF

was originally measured over a cylindrical profile rather than within the intended solid angle  $\Omega$ . The SH coefficients usually denoted by  $Y_l^m$  ( $l$  is the order,  $m$  is the phase factor), are the basis functions for complex functions on the sphere. It is assumed that ODFs are antipodally symmetric, and consequently, even values of  $l$  are used to decompose the data. Increasing the order of approximation reduces the width of the ODF lobes and hence increases the angular resolution as well as, to some extent, sensitivity to noise. With this kind of basis functions, the normalized signal  $S(\theta_i, \phi_i)$ , for  $K$  points ' $i = 1, 2, \dots, K$ ' on the  $q$ -shell is decomposed as

$$S(\theta_i, \phi_i) = \sum_{j=1}^N c_j Y_j(\theta_i, \phi_i) \quad (1)$$

where  $j$  indexes the even ordered SH bases, for ' $l = 0, 2, 4, \dots, l_{max}$ ' and ' $m = -l, \dots, 0, \dots, l$ '.  $c_j$  are the  $N$  SH coefficients. These coefficients completely determine the shape of the ODF.

For discrete samples on a single shell, this can be written as an overdetermined linear system, where the number of gradient directions along which the signal is acquired is higher than  $N$ . The SH coefficients are usually estimated using a least-squares minimization procedure, but this is known to be sensitive to noise, especially preponderant in high  $b$ -value datasets with lower SNR. Several regularization schemes have therefore been suggested, especially Tikhonov [20] and Laplace Beltrami regularizations [19], which smooth the reconstruction, at the expense of losing sharp features of the fiber distribution. This trade-off is essential in the interpolation of the signal values for multi-shell reconstruction, and a detailed analysis is presented in Section 3.2.

## 2.2 Multishell structure

Diffusion MR signal acquisitions are characterized by two important parameters: the  $b$ -value(s) (typically specified in  $s/mm^2$ ) and the gradient table. For typical diffusion levels in the human brain ( $0.0021mm^2/s$ ) [21], it has been observed that a single  $b$ -value in the order of  $2200 - 2800s/mm^2$  is optimal to reconstruct crossing fibers [21]. The gradient table selects points on the shell and encodes the spatial coordinates of points where the signal is acquired.

This type of sampling technique can be extended to multiple concentric shells of different  $b$ -values [22], so that the diffusion signal is measured at different levels of decay. Keeping the orientation fixed, as the  $b$ -value is varied, signal samples can be measured, which allows a straightforward estimation of higher order diffusion models along each orientation independently. We refer to this acquisition scheme as the *aligned* multi-shell scheme. It is also possible to distribute points uniformly (and incrementally) over several shells in order to maximize angular coverage, as demonstrated in [23]. In this setting, a wider range of orientations is covered. We refer to this acquisition scheme as the *staggered* multi-shell scheme [23].

## 2.3 Generalized CSA-ODF Algorithm

We generalize the previous work [9] on CSA-ODF by i) optimally fitting data on each shell using SH series and adaptive regularization, and ii) handling any

combination of  $b$ -values to fit a bi-exponential model (when at least three shells are present). We limit the discussion in this section to three shells, but this approach can be used over any number of shells, as illustrated in the experimental section. Once the signal  $S$ , sampled in  $K$  gradient directions has been fitted to a SH series of a given order (4, 6, or 8 in general), yielding  $N$  coefficients  $c_i$ , it is straightforward to estimate the signal value at a new orientation  $\hat{z}_j$  using Equation (1). We investigate in Section 3.2 optimal regularization choices at various SNR levels, SH orders and for various fiber configuration using Generalized Cross-Validation (GCV) [24].

With signal values available in any direction, it becomes possible to estimate the three parameters of a bi-exponential model. Aganj et al. [9] derive an analytic formula to estimate these parameters, under the constraint that the shells must have  $b$ -values in arithmetic progression. We generalize this to any set of  $b$ -values, using a trust region based optimization scheme (as implemented in MATLAB). The signal decay along a given direction is thus fitted to a bi-exponential model as follows: Along  $\hat{z}_j$ , the signal is estimated as

$$X_{\hat{z}_j}(b) = \alpha_{\hat{z}_j} e^{-\delta_{\hat{z}_j}^{(1)} b} + (1 - \alpha_{\hat{z}_j}) e^{-\delta_{\hat{z}_j}^{(2)} b} \quad (2)$$

where  $b$  is the  $b$ -value, and  $\alpha_{\hat{z}_j}$  is the volume fraction of the first decay component.  $\delta_{\hat{z}_j}^{(1)}$  and  $\delta_{\hat{z}_j}^{(2)}$  are the decay rates for the corresponding compartment along  $\hat{z}_j$  usually referred to as “slow” and “fast” compartment [18]. The ODF can now be calculated as

$$ODF(\hat{z}_j) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \text{FRT} \{ \nabla_b^2 (\alpha_{\hat{z}_j} \ln(\delta_{\hat{z}_j}^{(1)}) + (1 - \alpha_{\hat{z}_j}) \ln(\delta_{\hat{z}_j}^{(2)})) \} \quad (3)$$

The SH coefficients  $c_i$  of the double-logarithm of the signal are computed such that

$$\alpha_{\hat{z}_j} \ln(\delta_{\hat{z}_j}^{(1)}) + (1 - \alpha_{\hat{z}_j}) \ln(\delta_{\hat{z}_j}^{(2)}) = \sum_{i=1}^N c_i Y_i(\theta_j, \phi_j) \quad (4)$$

The SH coefficients  $c_{\text{ODF}}$  of the ODF are then obtained through a linear transformation of the  $c_i$  as described in [9]. We describe the method in detail in Algorithm 1.  $K$  signal acquisitions ( $S$ ) are supposed available along  $K$  gradient directions encoded by  $bvecs$ . The  $i^{\text{th}}$  element of shellID encodes the  $b$ -value at which the  $i^{\text{th}}$  acquisition in  $S$  was made.  $T$  is the matrix containing values taken by each considered spherical harmonic along each direction in  $bvecs$ .  $S_{\text{interp}}$  is the interpolated signal on each  $q$ -shell.  $\alpha$  and  $\delta^{(1,2)}$  are the parameters of the bi-exponential model.  $X$  is the estimated bi-exponential decay for each direction, and  $c_{\text{ODF}}$  is the SH coefficients of the ODF thus estimated. In order to illustrate the performance of the bi-exponential model fit, we provide in Fig. 1 the signal decay profile as a function of  $b$ -values along various directions for a single- and two-fiber configurations. The cylinder-dot-zeppelin compartment model [8] was used to generate data at SNR 15, along 300 directions. Along the fiber direction of the single fiber model, when averaged over the closest 30 datapoints, we observe a fast diffusivity of  $0.0017 \text{mm}^2/\text{s}$  with a volume fraction of 0.525 and a slow diffusivity of  $0.0005 \text{mm}^2/\text{s}$  with a volume fraction of 0.474. Across the

---

**Algorithm 1** : Generalized CSA-ODF
 

---

**Input:**  $S_{(K,1)}$ ,  $bvecs_{(K,3)}$ ,  $b\text{-value}_{(3,1)}$  (for 3 shells),  $ShellID_{(K,1)}$ .

**Output:**  $c_{ODF}$

**for**  $i = 1$  to 3 **do**

$S_i \leftarrow S$  where  $ShellID == b\text{-value}(i)$

$c_i = (T_i^T T_i + \lambda_i L)^{-1} T_i^T S_i$

$S_{interp_i} \leftarrow T_J * c_i$  {where  $J$  indexes an arbitrary gradient table.}

**end for**

**for**  $j = 1$  to number of directions in  $J$  **do**

$X_j(\alpha_j, \delta_j^{(1)}, \delta_j^{(2)}) \leftarrow \text{bi-expfit}(S_{interp})$  {bi-expfit as defined in Equation (2)}

**end for**

$c_{ODF} \leftarrow ODF((T_J^T T_J)^{-1} T_J^T X_j)$  {the ODF function as defined in Equation (3)}.

---

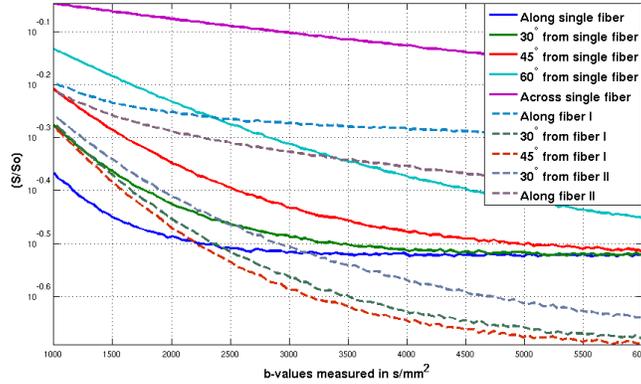


Fig. 1: Bi-exponential decay for the single- and two-fiber configurations.

fiber direction, when averaged over the closest 30 datapoints, we observe a fast diffusivity of  $0.0029mm^2/s$  with a volume fraction of 0.296 and a slow diffusivity of  $0.0002mm^2/s$  with a volume fraction of 0.68. These values are of the same order as the parameters we use to simulate the data. Moreover, we verify [8] that diffusion along the cylinder is bi-exponential, while largely monoexponential across the cylinder. Similar results were found in the two-fiber case along both cylinders (bi-exponential) and in the orthogonal direction.

### 3 Protocol Optimization

#### 3.1 Simulation setup

We rely on realistic diffusion models available in Camino [25]. We use three compartments for single-fiber configurations and five compartments for two-fiber configurations, with volume fractions 0.6 for the intra-axonal, 0.1 for the extra-axonal, and 0.3 for the isotropic compartment, as described in [8]. This is scaled appropriately for the two-fiber case. The intra-axonal compartment is modeled as a cylinder of radius  $0.004mm$  with a gaussian phase distribution (GPD) and diffusivity  $0.0017mm^2/s$ . The extra-axonal compartment is modeled

as a zeppelin with diffusivity  $0.0017\text{mm}^2/\text{s}$  along the fiber, and  $0.0002\text{mm}^2/\text{s}$  across it. Finally, the isotropic component is modeled as a dot. Eleven  $b$ -values:  $[1000, 1500, 2000, \dots, 6000]\text{s}/\text{mm}^2$  are calculated using the following pulse sequence parameters  $G = 0.05\text{ T/m}$ ,  $\text{TE} = 0.1\text{s}$ ,  $\delta = 0.02\text{s}$  and varying  $\Delta$ . In all experiments, we choose a maximum number of 300 gradient directions divided across three shells. For ground truth signal and associated ODF, we choose 300 directions on eleven shells (from the 11  $b$ -values), at SNR 40. As a measure of similarity of ODFs, the  $L^2$  norm of the difference in the SH coefficients is used.

### 3.2 Optimal regularization analysis:

The first step of the generalized CSA-ODF algorithm creates a continuous representation of the diffusion signal on each shell. To that end, we fit SH series of order 4, 6 or 8. The regularized least squares solution for the inverse problem of estimating the SH coefficients is given by:

$$c = (T^T T + \lambda I)^{-1} T^T S \quad (5)$$

where  $c$  is the vector of SH coefficients,  $S$  is the normalized signal,  $I$  is the identity matrix, and  $\lambda$  is the regularization parameter.

Methods to estimate the optimal value for  $\lambda$  have been discussed in [24, 26, 27]. The need for regularization arises from the small singular values of  $T$ . Small variations in  $S$  leads to high variation in the SH coefficients  $c$ . Regularized inversion is parameterized by  $\lambda$  which ideally is between the highest and lowest singular value of  $T$ . Laplace-Beltrami regularization has been extensively used

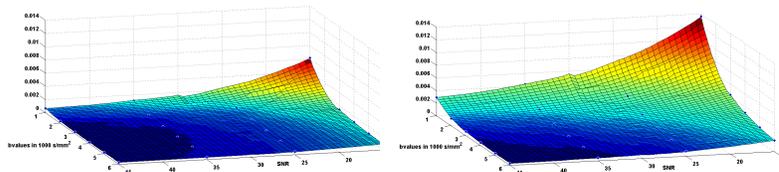


Fig. 2: Surface plot of  $\lambda$  for order 4 (left), and order 6 (right) for the two-fiber configuration as a function of SNR and  $b$ -value.

in the context of ODF reconstruction mostly because of the simple form of the Laplace-Beltrami operator for functions described by SH series [19]. It is controlled by replacing the identity matrix in Equation (5) by  $L$  (the square of the Laplace Beltrami matrix) whose diagonal elements are  $(l + 1)^2 l^2$ , where  $l$  is the order of the SH at the corresponding column.

$\lambda$  needs to be adjusted as a function of the acquisition and reconstruction parameters. We carried out an exhaustive investigation of the optimal choice of regularization for eleven  $b$ -values in the range  $[1000, 1500, \dots, 6000]\text{s}/\text{mm}^2$ , five SNR=5, 15, 25, 35, and 45, three SH reconstruction orders of 4, 6 and 8, and finally, three fiber configurations with 1, 2 or 3 orientations. The optimal  $\lambda$  value is calculated using the GCV method [24], rather than the L-curve method, since the L-curve method assumes the discrete Piccard conditions [26], which is not

Table 1: Optimal  $\lambda$  values for three-fiber, order 6 reconstruction ( $b$ -value in  $s/mm^2$ ).

$b$ -value	SNR: 5	SNR: 15	SNR: 25	SNR: 35	SNR: 45
1000	0.083302	0.014496	0.011214	0.006020	0.004279
1500	0.038631	0.008837	0.004006	0.001763	0.002086
2000	0.047775	0.006534	0.002520	0.001391	0.001055
2500	0.017271	0.004824	0.001556	0.000823	0.000573
3000	0.021266	0.002089	0.001388	0.000670	0.000429
3500	0.026	0.002072	0.001057	0.000470	0.000318
4000	0.027045	0.001593	0.0006	0.00037	0.000319
4500	0.01063	0.002516	0.000865	0.000676	0.000243
5000	0.024093	0.001576	0.000795	0.000448	0.000453
5500	0.022867	0.001241	0.000788	0.000383	0.000391
6000	0.020809	0.001715	0.001317	0.000419	0.000353

true for Equation (5). Fig. 2 shows surface plots of  $\lambda$ , for varying  $b$ -values and SNRs. Two SH orders are represented in the case of a two-fiber configuration. Each data point is obtained as the average optimal  $\lambda$  for 100 repetitions with randomly selected angles between fibers. Table 1 provides numerical values in the three-fiber case. The trends show that lesser regularization is necessary for higher SNRs, and that the optimal  $\lambda$  can deviate quite significantly from the commonly used 0.006 value [19], especially at SNR=25 and  $b$ -values greater than  $1500s/mm^2$ . Interestingly, at all orders, regularization slightly decreases with increased  $b$ -value for a given SNR. This behavior is consistent across the three investigated configurations. Eight other tables similar to Table 1 are available, therefore making it possible to adjust the choice of  $\lambda$  to specific experimental conditions, as done in the rest of the paper. We note that SH parameterization of the ODF (Equation (3)) is done without regularization.

### 3.3 Optimal $b$ -value analysis:

We study the effects of varying the sets of  $b$ -values to accurately estimate the ODF. We restrict ourselves to fourteen representative sets of three  $b$ -values. We will demonstrate in the following that adding a fourth shell only improve performance marginally. We restrict the range up to a  $b$ -value of  $6000s/mm^2$  as most of the bi-exponential nature of the diffusion signal is concentrated in the range of  $1500 - 4000s/mm^2$ . We fix the total number of gradient directions to 300 (distributed as 21, 86, 193 over the shells) [23]. All SNRs, single- and two-fiber configurations and SH orders 4 and 6 are tested. The corresponding optimal  $\lambda$  is always chosen from Section 3.2. In each case, as for all experiments from now on, 20 repetitions are performed and the mean result reported. Results of this simulation are shown in Fig. 3. The set  $[1000, 2000, 6000]s/mm^2$  shows lower error at SNR=15, 25 or 40 for an order 4 reconstruction and one-fiber configuration. Similar trends are observed at SNR 5 although, as expected, with much higher reconstruction error levels. Results are consistent with Fig. 1, where the signal decay exhibits the highest deviation from the mono-exponential behavior in the vicinity of  $b$ -value  $2000s/mm^2$ .  $b$ -value sets that skip this critical region show errors that are considerably larger as they do not capture the bi-exponential nature of the diffusion signal.  $b$ -value sets  $[1000, 2000, 3000/4000/5000]s/mm^2$

exhibit similar performance but we focus on the set  $[1000, 2000, 6000]s/mm^2$  to demonstrate the performance of the bi-exponential fitting algorithm in the most non-uniform sampling case.  $b$ -value  $[1000, 2000, 6000] s/mm^2$  show supe-

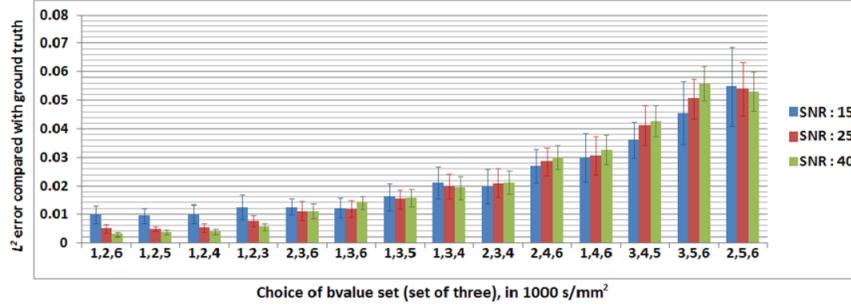


Fig. 3: Reconstruction error for order 6, single-fiber configuration.

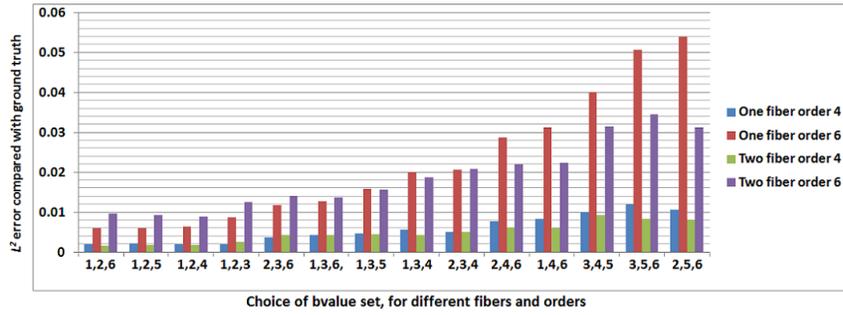


Fig. 4: Average reconstruction error for four combinations of SH order and fibers.

Table 2: Reconstruction error for varying order, SNR and fiber configurations for two representative  $b$ -value sets.

		Order: 4		Order: 6	
		one-fiber	two-fiber	one-fiber	two-fiber
SNR: 5	$b=1,2,6$	0.03383	0.02842	0.04709	0.03997
	$b=1,4,6$	0.03808	0.02938	0.05965	0.04820
SNR: 25	$b=1,2,6$	0.00138	0.00111	0.00498	0.00868
	$b=1,4,6$	0.00791	0.00577	0.03063	0.02152

rior reconstruction in all our tests as seen in Fig. 4. Clearly,  $b$ -value sets including  $1000$  and  $2000s/mm^2$  and with  $3000, 4000, 5000$  or  $6000s/mm^2$  as third shell outperform other combinations. We will consider  $[1000, 3000, 5000]$  and  $[1000, 4000, 6000]s/mm^2$  respectively as “average” and “poor” candidates for subsequent comparisons. For completeness, Table 2 demonstrates, for “good” and “poor” choices of  $b$ -value, the relative insensitivity of the reconstruction to fiber configurations (a very desirable property for brain data) as well as the difference in error (one order of magnitude) between SNR 5 and 25.

### 3.4 Optimal number of gradient directions:

Reconstruction error is studied for a reduced number of total points, with all the other parameters remaining unchanged. This provides a quantitative measure of the loss in reconstruction accuracy, in comparison with the ground truth. Fig. 5 (for a two-fiber configuration at SNR 25) shows how the error decreases as the number of sample points increases. It is seen that the curves are nearly flat after around 200 directions. Moreover,  $b$ -value sets previously identified as “good”, “average” and “poor” preserve their relative performance with increasing number of directions. Additionally, we compare the *aligned* sampling scheme

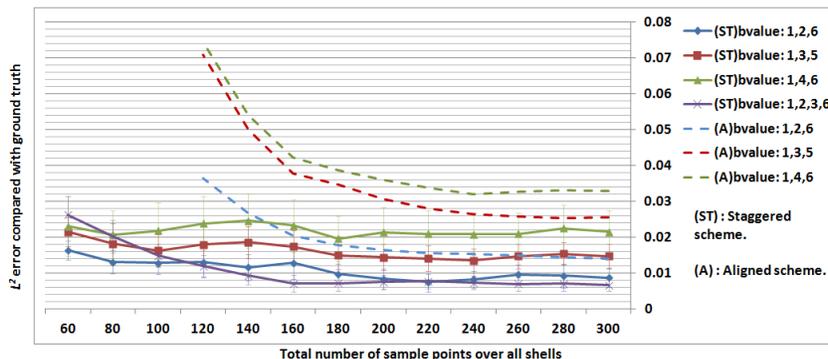


Fig. 5: Reconstruction error as a function of the number of directions acquired for different cases of  $b$ -value for two-fiber configuration at SNR 25, order 6 reconstruction.

with the *staggered* scheme for the same total number of data points over all shells, e.g: 300 *staggered* points divided as 21,86 and 193 are compared with 300 *aligned* points divided as 100, 100 and 100. Fig. 5 shows the reconstruction errors in both cases. The interpolation step in the *staggered* case does not introduce significant distortions. Moreover, this approach clearly provides a better angular resolution and improves ODF reconstruction. Fig. 5 also shows that a four-shell [1000, 2000, 3000, 6000] $s/mm^2$  acquisition does not provide an appreciable increase in accuracy. It therefore illustrates that three points are likely sufficient to estimate the three parameters  $\alpha$ ,  $\delta^{(1)}$  and  $\delta^{(2)}$ .

### 3.5 Angular resolution and error

We investigate the minimum achievable separating angle between fiber orientations using the *staggered* acquisition scheme with 200 directions,  $b$ -values of [1000, 2000, 6000] $s/mm^2$ , SH order 8 and SNR 40. It has been demonstrated that angles around 30 degrees can be recovered when using order 8 SH series [20]. Fig. 6 shows the reconstruction in steps of 10 degrees from orthogonal to a single-fiber configuration, with an order 8 reconstruction. The \* indicates the minimum angle between which we resolve crossings, which precisely happens between 35 and 30 degrees. Table 3 summarizes the angular error achieved with the same sampling scheme but varying SH orders and SNR for an orthogonal two-fiber configuration. For each case, 100 repetitions were performed and the mean  $\pm$  standard deviation of the angular error between the estimated ODF

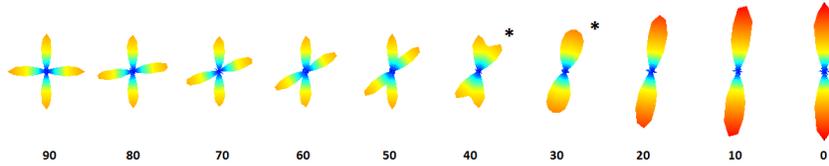


Fig. 6: Angular resolution in the two-fiber case, varying from orthogonal-fiber to a single-fiber configuration.

Table 3: Angular error performance (degrees) for an orthogonal-fiber configuration.

mean $\pm$ std	SNR: 5	SNR: 15	SNR: 25	SNR: 40
Order: 4	$5.3759 \pm 2.0902$	$1.5826 \pm 0.6079$	$1.0886 \pm 0.4610$	$0.7299 \pm 0.2736$
Order: 6	$5.4046 \pm 2.1040$	$1.5992 \pm 0.6564$	$1.1093 \pm 0.4642$	$0.7463 \pm 0.2725$
Order: 8	$5.4309 \pm 2.1006$	$1.6184 \pm 0.6215$	$1.0920 \pm 0.4685$	$0.7356 \pm 0.2769$

maxima [10] and the true orientation are reported. Even at low SNR, the maximum error is around 5 degrees. For realistic SNR, an average error of 1 degree is achieved which outperforms many existing ODF reconstruction technique [7,28].

#### 4 Application to human brain data

Finally, our generalized CSA-ODF algorithm was applied to human brain data from a healthy volunteer [29], acquired on a 3T Siemens scanner with the following parameters: Voxel size  $1.5mm$  isotropic, TR/TE=3200/77ms, 133 aligned DWI at  $b = [1000, 2000, 3000]s/mm^2$  and 10  $b_0$ . ODF reconstructions are presented in Fig. 7. Fiber crossings in the centrum semiovale and various other complex white matter areas are successfully recovered. Moreover, we down-sampled the data by half (from 133 to 67 directions) and verified (Fig. 7) that the ODF reconstruction still performed very well. This is expected from our simulations, although a staggered scheme was not available for this dataset (in which case we would expect even better results).

#### 5 Conclusion

We have generalized the CSA-ODF model to account for a flexible acquisition scheme, and provided an analysis of the optimal parameters for the reconstruction of the ODF. It is observed that specific sets of  $b$ -values (e.g. [1000, 2000, 6000]  $s/mm^2$ ) perform clearly better, which can be explained by looking deeper into the bi-exponential nature of the diffusion signal decay. For this particular reconstruction algorithm, three-shell acquisitions perform nearly as well as four-shell acquisitions, and around 200 gradient directions are sufficient for good angular resolution as well as accuracy.

#### Acknowledgements:

This work was partly supported by NIH grants R01 EB008432, P41 RR008079 (NCRR), P41 EB015894 (NIBIB), P30 NS057091, P30 NS5076408, and the Human Connectome Project (U54 MH091657) from the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research.

#### References

1. K. K. Seunarine and D. C. Alexander. Chapter 4 - multiple fibers: Beyond the diffusion tensor. In *Diffusion MRI*, pages 55 – 72. Academic Press, 2009.

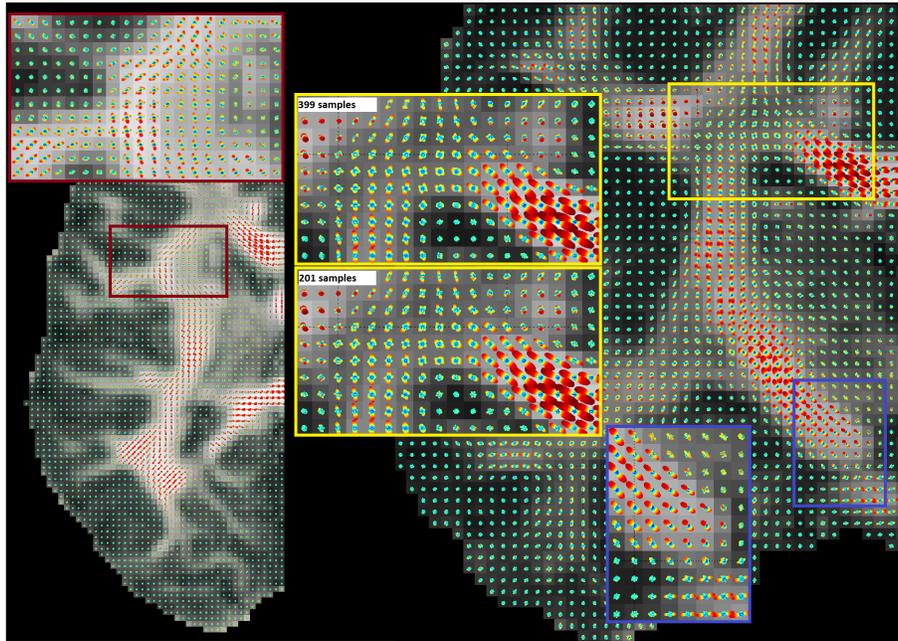


Fig. 7: The generalized CSA-ODF algorithm applied to brain data, superimposed on a Generalized Fractional Anisotropy (GFA) map with an *aligned* sampling scheme.

2. D.C. Alexander. An introduction to computational diffusion MRI: the diffusion tensor and beyond. *Visualization and processing of tensor fields*, pages 83–106, 2006.
3. P. J. Basser, J. Mattiello, and D. L. Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.
4. T. E. J. Behrens, M. W. Woolrich, M. Jenkinson, H. Johansen-Berg, R. G. Nunes, S. Clare, P. M. Matthews, J. M. Brady, and S. M. Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.*, 50(5):1077–1088, 2003.
5. E. Ozarslan, T. M. Shepherd, B. C. Vemuri, S. J. Blackband, and T. H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31(3):1086 – 1103, 2006.
6. D. S. Tuch. Q-ball imaging. *Magn. Reson. Med.*, 52(6):1358–1372, 2004.
7. J.-D. Tournier, C.-H. Yeh, F. Calamante, K.-H. Cho, A. Connelly, and C.-P. Lin. Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *NeuroImage*, 42(2):617 – 625, 2008.
8. E. Panagiotaki, T. Schneider, B. Siow, M. G. Hall, M. F. Lythgoe, and D. C. Alexander. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. *NeuroImage*, 59(3):2241 – 2254, 2012.
9. I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel. Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle. *Magn. Reson. Med.*, 64(2):554–566, 2010.
10. I. Aganj, C. Lenglet, and G. Sapiro. ODF maxima extraction in spherical harmonic representation via analytical search space reduction. *MICCAI*, pages 84–91, 2010.

11. A. Ghosh and R. Deriche. Extracting geometrical features & peak fractional anisotropy from the ODF for white matter characterization. In *ISBI*, pages 266–271, 2011.
12. S. Jbabdi, S. N. Sotiropoulos, A. M. Savio, M. Grana, and T. E. J. Behrens. Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magnetic Resonance in Medicine*, 2012.
13. M. Descoteaux, R. Deriche, D. L. Bihan, J.-F. Mangin, and C. Poupon. Multiple q-shell diffusion propagator imaging. *Medical Image Analysis*, 15(4):603 – 621, 2011. Special section on IPMI 2009.
14. M. H. Khachaturian, J. J. Wisco, and D. S. Tuch. Boosting the sampling efficiency of q-ball imaging using multiple wavevector fusion. *Magn. Reson. Med.*, 57(2):289–296, 2007.
15. H.-E. Assemlal, D. Tschumperlé, and L. Brun. Evaluation of q-space sampling strategies for the diffusion magnetic resonance imaging. *MICCAI '09*, pages 406–414, 2009.
16. A. Tristán-Vega, S. Aja-Fernández, and C.-F. Westin. On the blurring of the funk—radon transform in q—ball imaging. *MICCAI '09*, pages 415–422, 2009.
17. V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn. Reson. Med.*, 54(6):1377–1386, 2005.
18. Y. Assaf and Y. Cohen. Chapter 7 - inferring microstructural information of white matter from diffusion MRI. In *Diffusion MRI*, pages 127 – 146. 2009.
19. M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magn. Reson. Med.*, 58(3):497–510, 2007.
20. C. P. Hess, P. Mukherjee, E. T. Han, D. Xu, and D. B. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magn. Reson. Med.*, 56(1):104–117, 2006.
21. D.C. Alexander and G.J. Barker. Optimal imaging parameters for fiber-orientation estimation in diffusion MRI. *Neuroimage*, 27(2):357–367, 2005.
22. L. Zhan, A. D. Leow, I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, N. Harel, A. W. Toga, and P. M. Thompson. Differential information content in staggered multiple shell HARDI measured by the tensor distribution function. In *ISBI'11*, pages 305–309, 2011.
23. E. Caruyer, J. Cheng, C. Lenglet, G. Sapiro, T. Jiang, and R. Deriche. Optimal Design of Multiple Q-shells experiments for Diffusion MRI. In *MICCAI Workshop on CDMRI'11*, September 2011.
24. G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):pp. 215–223, 1979.
25. P.A. Cook, Y. Bai, S. Nedjati-Gilani, K. K. Seunarine, M. G. Hall, G. J. Parker, and D.C. Alexander. Camino: Open-source diffusion-MRI reconstruction and processing. *14th scientific meeting of the ISMRM*, 2759, 2006.
26. P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4):pp. 561–580, 1992.
27. E. Caruyer and R. Deriche. Optimal Regularization for MR Diffusion Signal Reconstruction. In *ISBI*, May 2012.
28. K.-H. Cho, C.-H. Yeh, J.-D. Tournier, Y.-P. Chao, J.-H. Chen, and C.-P. Lin. Evaluation of the accuracy and angular resolution of q-ball imaging. *NeuroImage*, 42(1):262 – 271, 2008.
29. J. L. R. Andersson, J. Xu, E. Yacoub, E. Auerbach, S. Moeller, and K. Ugurbil. A comprehensive gaussian process framework for correcting distortions and movements in diffusion images. *ISMRM Annual Meeting*, May 2012.

# Decomposition of Higher-Order Homogeneous Tensors and Applications to HARDI

E. Balmashnova, A. Fuster and L.M.J. Florack

Eindhoven University of Technology, The Netherlands  
E.Balmachnova@tue.nl

**Abstract.** High Angular Resolution Diffusion Imaging (HARDI) holds the promise to provide insight in connectivity of the human brain in vivo. Based on this technique a number of different approaches has been proposed to estimate the fiber orientation distribution, which is crucial for fiber tracking. A spherical harmonic representation is convenient for regularization and the construction of orientation distribution functions (ODFs), whereas maxima detection and fiber tracking techniques are most naturally formulated using a tensor representation. We give an analytical formulation to bridge the gap between the two representations, which admits regularization and ODF construction directly in the tensor basis.

## 1 Introduction

Diffusion MRI provides information about the fiber structure of brain white matter in a noninvasive way. It is based on the measurement of the Brownian motion of water molecules in tissue, which can be related to the tissue's microstructure. In the last decade several techniques have been proposed, from Diffusion Tensor Imaging (DTI) [1,2] to the more general High Angular Resolution Diffusion Imaging (HARDI) [3]. In order to perform classification or fiber tracking one needs accurate estimation of the orientation distribution function (ODF) from the original HARDI signal.

Most HARDI approaches are based on two alternative function representations on the sphere. The first one is the well-known spherical harmonic decomposition [4,5] and the second one is the homogeneous tensor decomposition proposed by Özarslan and Mareci [6]. These representations have the same number of basis functions and span the same vector space provided that the rank of the tensor representation equals the spherical harmonics highest order. Each has its pros and cons. The spherical harmonic representation permits efficient ways of computing the ODF from the data, as demonstrated in the context of the diffusion orientation transform (DOT) [7], and Q-ball technique and its generalizations [8,9,10]. It also permits regularization in a relatively straightforward way [11,12,13]. On the other hand, the tensor representation attracts nowadays more and more attention due to applications such as ODF maxima extraction [14,15], the computation of rotationally invariant scalar measures [16,17,18] and for parametrization of the space of positive-definite higher-order tensors [19,20]. The tensorial representation offers the additional advantage that it has the natural format for a Finsler geometric

approach towards HARDI tractography and connectivity analysis [21,22,23]. On the other hand, the initial step of obtaining the diffusion ODF requires a SH representation, after which a linear transformation to a tensor description is performed [24]. Florack and Balmashnova [12] offer a third representation, based on an inhomogeneous tensor decomposition. Their representation reconciles the tensor representation with the regularization rationale obviating the need for an intermediate spherical harmonic decomposition. A first attempt to construct the diffusion ODF directly in tensor basis was recently presented [25]. Their recipe requires lengthy manual derivation for each tensor order, which becomes very cumbersome for high orders. In this work, however, we present an analytic approach.

In this paper we describe how to obtain the diffusion ODF (Q-ball and constant solid angle) and regularization analytically for a tensor representation of arbitrary order directly from data evidence, by fitting a homogeneous tensor to the data and then analytically decomposing the resulting homogeneous tensor in order to benefit from the inhomogeneous representation. The final algorithm for regularization or ODF computation is based on simple matrix multiplication of the tensor coefficients.

## 2 Higher-Order Tensor Representations

An interesting alternative to spherical harmonics is provided by a higher-order tensor representation (higher than order two). Originally Özarslan and Mareci [6] proposed a homogeneous high-rank tensor representation<sup>1</sup>

$$S_n^{\text{hom}}(y) = D^{i_1 \dots i_n} y_{i_1} \dots y_{i_n}, \quad (1)$$

where  $y = (y_1, y_2, y_3)$  is a unit vector and  $D$  is a higher-order diffusion tensor of rank  $n$ . This is a generalization of the diffusion tensor model in which a 2-rank tensor is used. These high-rank Cartesian diffusion tensors can be computed avoiding the computationally costly spherical harmonics. Regularization plays a very important role for robustness purposes. In case of a homogeneous high-rank tensor representation, the only explicit type of regularization that is easily enforced is by constraining the rank (lower rank implies higher degree of regularization) as explained in [26]. In fact, DTI is reasonably robust because of the 2nd order rank constraint explicitly built in. The alternative inhomogeneous higher-order tensor decomposition resolves this robustness problem by exploiting redundancy of the polynomial basis on the sphere in such a way that terms of the same order in the resulting decomposition become eigenfunctions under regularization [12]. The nonregularized data decomposition has the form

$$S_n^{\text{inhom}}(y) = \sum_{j=0}^n D^{i_1 \dots i_j} y_{i_1} \dots y_{i_j} \quad (2)$$

in the redundant basis of polynomials  $\{y_{i_1} \dots y_{i_j} \mid j = 0, 2, \dots, n\}$  on the unit sphere<sup>2</sup>. The idea is to encode in the higher-order part the residual information which cannot

<sup>1</sup> Here and henceforth summation convention for repeated indices applies.

<sup>2</sup> It is assumed that  $S_n(y) = S_n(-y)$ , whence only even order monomials are used.

be revealed by lower-order terms. The algorithm for computing the coefficients has a hierarchical structure [12]:

1.  $D^0$  is found by minimization of the energy ( $\Omega$  denotes the unit sphere, and  $d\Omega$  the appropriate measure)

$$E_0(D^0) = \int_{\Omega} (S(y) - D^0)^2 d\Omega, \quad (3)$$

2. If all terms of order up to  $j - 1$  are known, the  $j$ -th order coefficients are obtained from minimization of the energy

$$E_j(D^{i_1 \dots i_j}) = \int_{\Omega} \left( (S(y) - \sum_{k=0}^{j-1} D^{i_1 \dots i_k} y_{i_1} \dots y_{i_k}) - D^{i_1 \dots i_j} y_{i_1} \dots y_{i_j} \right)^2 d\Omega \quad (4)$$

The solution requires only one data term of the resulting linear system to be computed numerically, analytic expressions for all other integrals are given in [12]. Although both tensor decompositions, homogeneous and inhomogeneous, are equivalent on the sphere in the sense that  $S_n^{\text{inhom}}(y) = S_n^{\text{hom}}(y)$ , the inhomogeneous one has important advantages. It allows to regularize and construct the ODF directly, without going to a spherical harmonics basis. This is possible due to the following properties:

1. The polynomials  $D^{i_1 \dots i_k} y_{i_1} \dots y_{i_k}$  for fixed  $k$  belong to the span  $\{Y_{km}(y), m = -k, \dots, k\}$  of the spherical harmonics of the same order  $k$ .
2. The polynomials  $D^{i_1 \dots i_k} y_{i_1} \dots y_{i_k}$  are eigenfunctions of the Laplace-Beltrami operator  $\Delta_{\text{LB}}$ , thus in practice they are easily regularized by Tikhonov regularization [11,12].
3.  $S_n^{\text{inhom}}(y, t)$  satisfies the heat equation on the unit sphere with the initial condition  $S_n^{\text{inhom}}(y, 0) = S_n^{\text{inhom}}(y)$ .

However, the inhomogeneous representation also poses a few drawbacks:

1. Due to the fact that the algorithm requires several steps with least squared fittings, it is less robust compared to one for the homogeneous representation.
2. It requires to store more coefficients for each point. Instead of the  $(n+1)(n+2)/2$  coefficients required in the homogeneous case,  $(n+2)(n+4)(2n+3)/24$  are needed, unless one is able to make all mutual dependencies among the coefficients explicit beforehand (which is not an easy task).

### 3 Decomposition of Homogeneous Tensors

Starting point is a homogeneous tensor representation, Eq. (1), fitted to HARDI data (see, for example, [27]). We propose to decompose such a homogeneous tensor into the inhomogeneous representation given by Eq. (2). Instead of tensor index notation

we adopt the multi-index notation used in [19]. The signal  $S$  can then be written as a homogenous polynomial of even order  $n$ :

$$S_n^{\text{hom}}(y) = \sum_{|\alpha|=n} D_\alpha y^\alpha = \sum_{\alpha_1+\alpha_2+\alpha_3=n} D_{\alpha_1\alpha_2\alpha_3} y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} \quad (5)$$

For implementation purposes we store the coefficients in Eq. (5) as a  $\frac{1}{2}(n+1)(n+2)$ -dimensional vector  $\mathbf{D}$  and the monomials as a vector  $\mathbf{v}$ , so that

$$S_n^{\text{hom}}(y) \stackrel{\text{def}}{=} \mathbf{D}^T \mathbf{v}. \quad (6)$$

For example, in the case  $n = 2$ ,  $\mathbf{v}^T = (y_1^2, y_1 y_2, y_2^2, y_2 y_3, y_3^2, y_1 y_3)$ . Given such a homogeneous polynomial matched to the data  $S$  we now wish to map the known coefficients  $D_\alpha$  to coefficients  $D_n^{2\nu}$  of the aforementioned equivalent inhomogeneous representation<sup>3</sup>, Eq. (2), in such a way that  $D_n^{2\nu}(y)$  belongs to the span of spherical harmonics of order  $2\nu$ :

$$S_n(y) = S_n^{\text{inhom}}(y) \stackrel{\text{def}}{=} \sum_{\nu=0}^{n/2} D_n^{2\nu}(y). \quad (7)$$

The vectors of coefficients are related by

$$\mathbf{D} = \sum_{\nu=0}^{n/2} \mathbf{D}^{2\nu}. \quad (8)$$

Any polynomial can be decomposed into a sum of harmonic polynomials  $h_l$ , as follows:

$$S_n(y) = \sum_{\nu=0}^{n/2} r^{n-2\nu} h_{2\nu}(y) \quad (9)$$

where  $h_l$  is a homogeneous harmonic polynomial of order  $l$  (i.e.  $\Delta h_l(y) = 0$ , where  $\Delta$  is the Laplacian operator in 3D) and  $r^2 = y_1^2 + y_2^2 + y_3^2$ . The property of spherical harmonics being homogeneous harmonic polynomials on the unit sphere is of crucial importance, since it allows us to transfer well-developed algorithms for HARDI relative to a spherical harmonic basis to the inhomogeneous polynomial basis. This allows us to conclude that the polynomial  $r^{n-k} h_k$  restricted to the unit sphere for fixed  $k$  belongs to the span of the spherical harmonics of the same order,  $\{Y_{km}(y), m = -k, \dots, k\}$ . The polynomial  $r^{2\nu} h_{n-2\nu}$  is explicitly given by [28]:

$$r^{n-2\nu} h_{2\nu} = \frac{(4\nu+1)!!}{(n-2\nu)!!(n+2\nu+1)!!} \sum_{\mu=0}^{\nu} \frac{(-1)^\mu (4\nu-2\mu-1)!!}{(2\mu)!!(4\nu-1)!!} r^{2(\mu+\frac{n}{2}-\nu)} \Delta^{\mu+\frac{n}{2}-\nu} S_n(y) \quad (10)$$

We derive the exact formula by substituting Eq. (5) into Eq. (10) and using the binomial expansions for both  $r^{n-2(\nu-\mu)}$  and  $\Delta^{\frac{n}{2}-(\nu-\mu)}$ , with changes in the order of summation<sup>4</sup>

$$D_n^{2\nu}(y) \stackrel{\text{def}}{=} r^{n-2\nu} h_{2\nu} = \sum_{|\alpha|=n} D_\alpha^{2\nu} y^\alpha \quad (11)$$

<sup>3</sup> The reverse mapping from inhomogeneous to homogeneous representation is trivial by inserting powers of  $r^2 = y_1^2 + y_2^2 + y_3^2$  into the monomials in Eq. (5).

<sup>4</sup> Caveat:  $\alpha$  is multi-index,  $n$  is integer, so  $D_n^{2\nu}(y)$  and  $D_\alpha^{2\nu}$  should not be confused.

where for all  $|\alpha| = n$  (cf Eq. (5))

$$D_\alpha^{2\nu} = \sum_{|\beta|=n} D_\beta c_\alpha^{2\nu,\beta}, \quad (12)$$

$$c_\alpha^{2\nu,\beta} = \frac{(4\nu+1)!!}{(n-2\nu)!!(n+2\nu+1)!!} \sum_{\mu=0}^{\nu} (-1)^\mu \sum_{|\gamma|=\mu+\frac{n}{2}-\nu}^* \frac{(4\nu-2\mu-1)!!((\mu+\frac{n}{2}-\nu)!)^2 \beta!}{(2\mu)!!(4\nu-1)!!\gamma!(\alpha-\beta+2\gamma)/2!(\beta-2\gamma)!} \quad (13)$$

where \* means that summation applies only for the terms for which the following conditions are satisfied:

$$(\beta - \alpha \leq 2\gamma \leq \beta) \quad \text{and} \quad (\alpha_i - \beta_i \text{ even}), \quad (14)$$

Note that in multi-index notation  $\alpha! = \alpha_1! \alpha_2! \alpha_3!$ , and  $\alpha < \beta$  implies  $\alpha_i < \beta_i$  for all indices  $i$ . The values  $c_\alpha^{2\nu,\beta}$  do not depend on the data and have to be computed only once. From Eq. (13) we can construct the vector of coefficients  $\mathbf{D}^{2\nu}$ , Eq. (12), as follows:

$$\mathbf{D}^{2\nu} = C^{2\nu} \mathbf{D}, \quad \nu = 0, \dots, \frac{n}{2}. \quad (15)$$

The  $\frac{1}{2}(n+1)(n+2) \times \frac{1}{2}(n+1)(n+2)$  matrices  $C^{2\nu}$  sum up to the identity:

$$\sum_{\nu=0}^{n/2} C^{2\nu} = I. \quad (16)$$

This decomposition gives the coefficient vector  $\mathbf{D}^{2\nu}$  corresponding to a polynomial from the span of the spherical harmonics of order  $2\nu$ ,  $\{Y_{2\nu,m}(y), m = -2\nu, \dots, 2\nu\}$ . Since most approaches use properties of spherical harmonics, this decomposition allows one to use the same techniques in the tensor representation. We illustrate this point in the next section with several examples.

## 4 Applications

**Example 1: Regularization.** Laplace-Beltrami regularization is widely used in HARDI since it is a natural smoothing via diffusion on the sphere [11,12]. Let  $S_n^{\text{SH}}(y)$  be a spherical harmonic decomposition of the signal  $S$

$$S_n^{\text{SH}}(y) = \sum_{l=0}^n \sum_{m=-l}^l a_{lm} Y_{lm}(y), \quad (17)$$

and  $\Delta_{\text{LB}}$  the Laplace-Beltrami operator on the unit sphere, then the regularized signal can be readily obtained [12]

$$S_n^{\text{SH}}(y, t) \equiv e^{t\Delta_{\text{LB}}} S_n^{\text{SH}}(y) = \sum_{l=0}^n \sum_{m=-l}^l a_{lm}(t) Y_{lm}(y), \quad (18)$$

where  $a_{lm}(t) = e^{-l(l+1)t} a_{lm}$ . Note that the coefficient transformation depends only on  $l$  and not on  $m$ . This allows us to apply this technique to the tensor representation. The polynomial with the coefficients  $\mathbf{D}^{2\nu} = C^{2\nu} \mathbf{D}$  belongs to span of spherical harmonics of order  $2\nu$ . Therefore, the regularized polynomial will have the following coefficients

$$\mathbf{D}^{2\nu}(t) = e^{-2\nu(2\nu+1)t} C^{2\nu} \mathbf{D}. \quad (19)$$

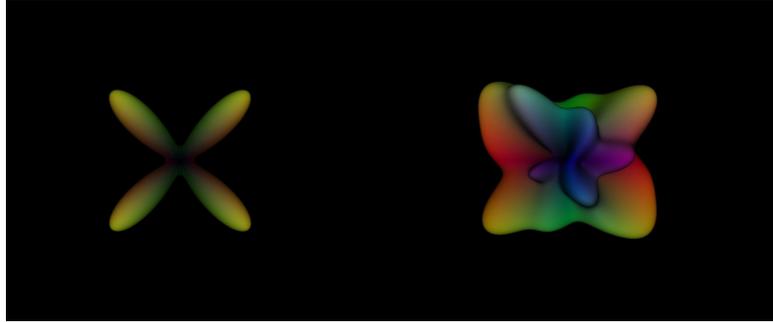
The resulting regularization matrix for the whole polynomial is thus

$$C_{\text{LB}}(t) = \sum_{\nu=0}^{n/2} e^{-2\nu(2\nu+1)t} C^{2\nu} \quad (20)$$

and the regularized polynomial can be written as

$$D(y_1, y_2, y_3, t) = \mathbf{D}^T C_{\text{LB}}(t) \mathbf{v}, \quad (21)$$

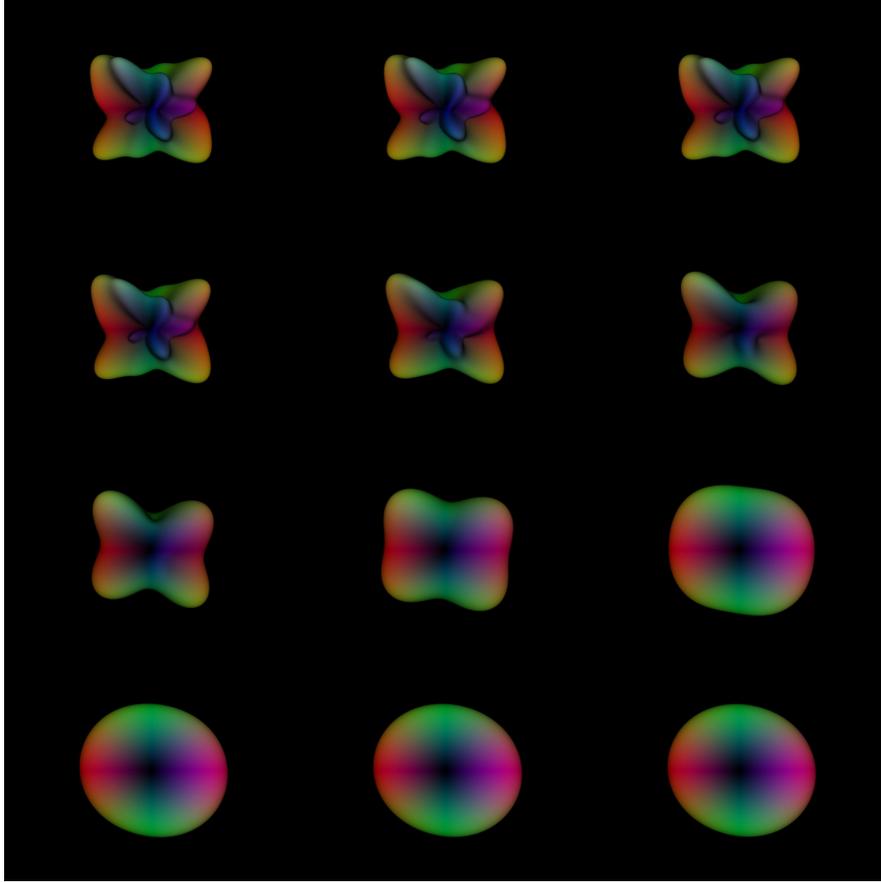
The parameter  $t$  can be interpreted as inverse angular resolution. Figures 1 and 2 illustrate the importance of the scale parameter. In the case of noisy data, low  $t$ -values prevent correct detection of fiber orientations. In the interval  $t \in [0.05, 0.15]$  the two most pronounced modes agree with the crossing fiber orientations with an angular error of less than  $9^\circ$ , and a minimal error of  $6^\circ$ .



**Fig. 1.** Left: Synthetic noise-free profile induced by two crossing fibers at an angle of 90 degrees. Right: Same profile with added Rician noise.

**Example 2: Q-Ball Imaging.** A model-free diffusion ODF reconstruction scheme has been introduced by Tuch [10]. The diffusion probability function  $P(\mathbf{r})$  is related to the measured MR diffusion signal  $E(\mathbf{q})$ , where  $E(\mathbf{q}) = S(\mathbf{q})/S_0$  ( $S_0$  is the non-diffusion-weighted image), by a Fourier integral

$$P = \mathcal{F}[E] \quad (22)$$



**Fig. 2.** Regularized profiles produced from the right image in Fig. 1 up to  $n = 8$ , with scale  $t \in [0.007, 1.0]$  exponentially sampled.

where  $\mathbf{r}$  is the displacement vector and  $\mathbf{q}$  is the diffusion wave-vector, which is proportional to the applied magnetic field gradient in the MRI scanner. The three-dimensional probability density function  $P(\mathbf{r})$  contains information about the tissue microstructure. A problem with using Eq. (22) directly via DFT is that acquisition data  $E(\mathbf{q})$  for all  $\mathbf{q}$  is needed. This is not feasible since it leads to long acquisition times and the need for high values of  $|\mathbf{q}|$ , causing serious noise problems. By considering the diffusion orientation function (ODF) for a direction  $\mathbf{u}$  defined by the radial projection of the diffusion function

$$\Psi(u) = \int_0^\infty P(ru)dr, \quad (23)$$

where  $u$  is a unit vector, the orientation structure of  $P(\mathbf{r})$  can be described.

There are several similar analytical Q-ball algorithms [11,13,29]. In this example we use the approach of Descoteaux et al., based on the 3D Funk-Hecke theorem, since it

can be adopted for higher-order tensor representations. The Funk-Radon transform  $\mathcal{G}$  of a spherical harmonic decomposition, Eq. (17), can now be written as

$$\Psi(u) = \mathcal{G}[S_n](u) = 2\pi \sum_{l=0}^n P_l(0) a_{lm} Y_{lm}(u). \quad (24)$$

where  $P_{2\nu}(0) = (-1)^\nu \frac{(2\nu-1)!!}{(2\nu)!!}$  and the Funk-Radon transform is defined as

$$\mathcal{G}[f](u) \stackrel{\text{def}}{=} \int \int_{|w|=1} f(w) \delta(u^T w) dw. \quad (25)$$

Here again the coefficient transformation depends only on the spherical harmonic order, and as in the regularization example, we can apply the result to our inhomogeneous tensor representation:

$$\mathbf{D}_{\text{ODF}}^{2\nu} = 2\pi P_{2\nu}(0) C^{2\nu} \mathbf{D}. \quad (26)$$

Therefore, we can construct the transformation matrix for the whole polynomial as

$$C_{\text{ODF}} = 2\pi \sum_{\nu=0}^{n/2} P_{2\nu}(0) C^{2\nu} \quad (27)$$

and finally we have

$$\Psi(u) = \mathbf{D}^T C_{\text{ODF}} \mathbf{V}, \quad (28)$$

The regularization and ODF reconstruction steps, given by Eqs. (18) and (28), can be combined in one for implementation purposes. Therefore, the whole algorithm for obtaining the diffusion ODF in tensor representation has two steps involving linear operations only:

1. Fit an  $n$ th order homogeneous polynomial to the signal by solving a linear system of equations.
2. Compute the regularized ODF coefficients  $\mathbf{D}_{\text{ODF}}(t) = C_{\text{ODF}} C_{\text{LB}}(t) \mathbf{D}$ .

**Example 3: Constant Solid Angle Orientation Distribution Function.** Several flaws of the Q-ball approach were pointed out in [30]. One of them is that the ODF definition does not deal with volume elements in a proper way, the adequate definition being

$$\Psi(u) = \int_0^\infty P(ru) r^2 dr. \quad (29)$$

This definition has been used to compute the ODF in [8]. However, this approach is not model-free as it requires the Stejskal-Tanner assumption about monoexponential decay of the signal [31]. The modified formula for the ODF, say instead of Eq. (24), is

$$\Psi(u) = \frac{1}{4\pi} + \frac{1}{16\pi^2} \mathcal{G}[\Delta_{\text{LB}} \ln(-\ln(E(y)))](u), \quad (30)$$

In this case the same approach as in the original Q-ball gives

$$C_{\text{ODF}} = -\frac{1}{8\pi} \sum_{\nu=0}^{n/2} P_{2\nu}(0) 2\nu(2\nu+1) C^{2\nu} \quad (31)$$

and the final expression becomes

$$\Psi(u_1, u_2, u_3) = \frac{1}{4\pi} + \mathbf{D}^T C_{\text{ODF}} \mathbf{V} \quad (32)$$

where  $\mathbf{D}$  is a vector of tensor coefficients of  $\ln(-\ln(E(y)))$ .

## 5 Conclusions

We have shown that a particular inhomogeneous polynomial representation of HARDI data on the unit sphere has certain theoretical and practical merits. We have derived the formulas for decomposition of a homogeneous polynomial of arbitrary order on the sphere in terms of such an inhomogeneous polynomial. The resulting decomposition requires a simple matrix multiplication of the vector of the coefficients obtained from the measurement data. We also show how this decomposition may be used for regularization and closed-form Q-ball representation. The resulting diffusion ODF is equivalent to the one obtained from the well-established analytical Q-ball algorithm. The advantage is that this approach does not require a detour via spherical harmonic decomposition, but instead can be obtained directly using a suitable higher order tensor formalism. This is of interest in many applications where the tensor formalism is the preferred choice.

The diffusion ODF can be used for fiber tracking techniques. In the most straightforward approach the tracking is performed by following the principle directions. For this reason maxima detection of the ODF is one of the major focuses, and a tensor representation has proven to be helpful [14,15]. We also would like to point out that in maxima detection techniques, regularization plays a crucial role. For this reason we have combined the Q-ball technique with regularization. In the case where the regularization parameter is sufficiently high, the blurred function has only one maximum. Decreasing regularization parameter leads to a splitting into two and subsequently more local maxima. Therefore, the hierarchical structure can be exploited in a coarse-to-fine maxima detection framework. Finally, the proposed method might also have applications in the context of Diffusional Kurtosis Imaging and in statistical models of second-order diffusion tensors, for decomposition of fourth-order kurtosis and covariance tensors.

## References

1. Basser, P.J., Mattiello, J., Le Bihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance* **103** (1994) 247–254

2. Basser, P.J., Mattiello, J., Le Bihan, D.: MR diffusion tensor spectroscopy and imaging. *Biophysics Journal* **66**(1) (1994) 259–267
3. Tuch, D., Reese, T., Wiegell, M., Makris, N., Belliveau, J., van Wedeen, J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine* **48**(6) (2004) 1358–1372
4. Alexander, D.C., Barker, G.J., Arridge, S.R.: Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine* **48**(2) (2002) 331–340
5. Frank, L.R.: Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine* **47**(6) (2002) 1083–1099
6. Özarslan, E., Mareci, T.H.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution imaging. *Magnetic Resonance in Medicine* **50** (2003) 955–965
7. Özarslan, E., Shepherd, T.M., Vemuri, B.C., Blackband, S.J., Mareci, T.H.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage* **31** (2006) 1086–1103
8. Aganj, I., Lenglet, C., Sapiro, G.: Odf reconstruction in q-ball imaging with solid angle consideration. In: *ISBI'09: Proceedings of the Sixth IEEE international conference on Symposium on Biomedical Imaging*, Piscataway, NJ, USA, IEEE Press (2009) 1398–1401
9. Tristan-Vega, A., Westin, C.F., Aja-Fernandez, S.: A new methodology for the estimation of fiber populations in the white matter of the brain with the funk-radon transform. *NeuroImage* **49** (2010) 1301–1315
10. Tuch, D.S.: Q-ball imaging. *Magnetic Resonance in Medicine* **52** (2004) 1358–1372
11. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine* **58**(3) (2007) 497–510
12. Florack, L., Balmashnova, E.: Decomposition of high angular resolution diffusion images into a sum of self-similar polynomials on the sphere. In Bayakovsky, Y., Moiseev, E., eds.: *Proceedings of the Eighteenth International Conference on Computer Graphics and Vision, GraphiCon'2008, Moscow, Russia, June 23–27, 2008, Moscow State University* (2008) 26–31
13. Hess, C., Mukherjee, P., Han, E., Xu, D., Vigneron, D.: Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine* **56**(1) (2006) 104–117
14. Bloy, L., Verma, R.: On computing the underlying fiber directions from the diffusion orientation distribution function. In Metaxas, D.N., Axel, L., Fichtinger, G., Székely, G., eds.: *MICCAI (1)*. Volume 5241 of *Lecture Notes in Computer Science.*, Springer (2008) 1–8
15. Ghosh, A., E.Tsagaridas, Descoteaux, M., Comon, P., B.Mourrain, Deriche, R.: A polynomial based approach to extract the maxima of an antipodally symmetric spherical function and its application to extract fiber directions from the orientation distribution function in diffusion mri. In: *MICCAI:08 Workshop on Computational Diffusion MRI, New York, USA (10/09/2008 2008)*
16. Özarslan, E., Vemuri, B., Mareci, T.: Generalized scalar measures for diffusion mri using trace, variance and entropy. In: *Magn. Reson. Med.* 53:866876. *Biophysical Journal* 94(7) 28092818 Özarslan et al. (2005)
17. Ghosh, A., Deriche, R.: Extracting geometrical features & peak fractional anisotropy from the ODF for white matter characterization. In: *ISBI. (2011)* 266–271
18. Fuster, A., van de Sande, J., Astola, L., Poupon, C., Velterop, J., ter Haar Romeny, B.M.: Fourth-order tensor invariants in high angular resolution diffusion imaging. In Zhang, G.H., Adluru, N., eds.: *Proc. MICCAI Workshop on Computational Diffusion MRI. (2011)* 54–63
19. A. Barmpoutis, M. S. Hwang, D.H.J.R.F., Vemuri, B.C.: Regularized positive-definite fourth-order tensor field estimation from dw-mri. *NeuroImage* **45**(1 sup.1) (March 2009) 153–162

20. Qi, L., Yu, Wu, E.X.: Higher order positive semidefinite diffusion tensor imaging. *SIAM Journal on Imaging Sciences* **3**(3) (2010) 416–433
21. Astola, L., Florack, L.: Finsler geometry on higher order tensor fields and applications to high angular resolution diffusion imaging. In Tai, X.C., Mørken, K., Lysaker, M., Lie, K.A., eds.: *SSVM*. Volume 5567 of *Lecture Notes in Computer Science.*, Springer (2009) 224–234
22. Florack, L., Balmashnova, E., Astola, L., Brunenberg, E.: A new tensorial framework for single-shell high angular resolution diffusion imaging. *Journal of Mathematical Imaging and Vision* **3**(38) (August 13 2010) 171–181 Published online: DOI 10.1007/s10851-010-0217-3.
23. Melonakos, J., Pichon, E., Angenent, S., Tannenbaum, A.: Finsler active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30**(3) (2008) 412–423
24. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine* **56**(2) (2006) 395–410
25. Astola, L., Jalba, A., Balmashnova, E., Florack, L.: Finsler streamline tracking with single tensor orientation distribution function for high angular resolution diffusion imaging. *Journal of Mathematical Imaging and Vision* **41** (2011) 170–181
26. Florack, L., Balmashnova, E.: Two canonical representations for regularized high angular resolution diffusion imaging. In Alexander, D., Gee, J., Whitaker, R., eds.: *MICCAI Workshop on Computational Diffusion MRI*, New York, USA, September 10, 2008. (2008)
27. Descoteaux, M.: High angular resolution diffusion MRI: From local estimation to segmentation and tractography. In: PhD thesis. (2008)
28. Avery, J., Antonsen, F.: Evaluation of angular integrals by harmonic projection. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling* **85**(1-3) (March 1993) 33–42
29. Anderson, A.W.: Measurement of fiber orientation distribution using angular resolution diffusion imaging. *Magnetic Resonance in Medicine* **54** (2005) 1194–1206
30. Barnett, A.: Theory of q-ball imaging redux: Implications for fiber tracking. *Magnetic Resonance in Medicine* **62**(4) (October 2009) 910–923
31. Stejskal, E.O., Tanner, J.E.: Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *Journal of Computational Physics* **42** (1965) 288–292

# Dynamic Diffusion Basis Functions for Axon Fiber Structure Estimation from DW-MRI

Omar Ocegueda      Mariano Rivera  
{omar@cimat.mx, mrivera@cimat.mx}

Center for Research in Mathematics

**Abstract.** We present an algorithm for dictionary learning of diffusion basis functions for recovering the intra-voxel axon fiber tract geometry from diffusion-weighted magnetic resonance images of the brain. The input diffusion weighted signal can be represented as a sparse linear combination of the limiting dictionary. Synthetic experiments using the 2012 HARDI Reconstruction Challenge dataset demonstrate that, by using the limiting dictionary, the representation of the input signal is sparser and the reconstruction error is consistently lower than using state-of-the-art *basis-pursuit* algorithms, which use a fixed dictionary of diffusion functions.

## 1 Introduction

Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) is a method for measuring the diffusion of water molecules in biological tissue. When applied to the brain, DW-MRI can be used to reconstruct the geometric structure of the axon fiber tracts, which in turn can be exploited to study *in vivo* the neuronal connectivity of the brain. In 1994, Basser *et al.* [2] proposed the Diffusion Tensor (DT) as a plausible model for the DW-MRI signal:

$$S_p(q_k) = S_0(p) \exp(-\tau q_k^T D_p q_k) \quad (1)$$

where  $S_0(p)$  is a standard T2 image at voxel  $p$ ,  $D_p$  is a symmetric positive definite tensor and  $\tau$  is the effective diffusion time. The coefficient  $-\log \frac{S_p(q_k)}{S_0(p)}$  is known as the Apparent Diffusion Coefficient (ADC) diffusion vector  $q_k$ , for consistency with this naming convention, we will refer to the function  $-\log \frac{S_p(\cdot)}{S_0(p)}$  as the *Apparent Diffusion Function* (ADF) at voxel  $p$ . In practice, the output of the scanner is the signal  $S_p$ , sampled at a small set of diffusion vectors  $Q = \{q_k\}_{k=1}^n \subset \mathcal{S}^3$  (known as the *acquisition scheme*) which is the same for all the voxels in the volume. Therefore, the output of the scanner at voxel  $p$  may be regarded as a vector in  $\mathfrak{R}^n$ ,  $v_{p,k} = S_p(q_k)$ .

Under the DT model, the shape of the ADF is determined by the tensor  $D_p$ . The Principal Diffusion Direction (PDD) of a tensor  $D_p$  is its eigenvector with the largest eigenvalue. According to the DT model, the diffusion of water

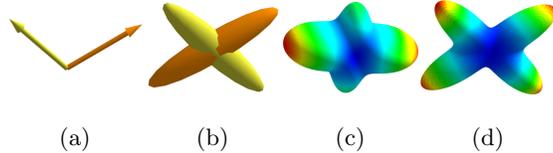


Fig. 1: Graphical representations of a 2-compartment multi-tensor. a) Principal Diffusion Directions (PDD's), b) Iso-surfaces of the diffusion tensors, c) Apparent Diffusion Function (ADF), d) Orientation Diffusion Function (ODF), *i.e.* the probability of diffusion in a given direction [14].

molecules is maximum along the PDD. The most important limitation of the Diffusion Tensor model is that it does not correctly model the signal observed at voxels where two or more fibers cross. To overcome this limitation, Tuch [14] developed the Multi-Tensor (MT) model:

$$S_p(q_k) = S_0(p) \sum_{j=1}^m \beta_j \exp(-\tau q_k^T D_j q_k), \quad (2)$$

where the nonnegative coefficients  $\{\beta_j\}_{j=1}^m$  indicate the contribution of the  $j$ -th tensor (also known as the  $j$ -th *compartment*) to the ADF. In other words, in a voxel containing an axon fiber crossing, the contribution of each bundle to the observed signal is modeled using one individual diffusion tensor. Figure 1 depicts four graphical representations of a MT with two compartments. Over the years, DW-MRI data acquisition schemes have evolved to become increasingly more complex and accurate. The simplest techniques, like Diffusion Tensor Imaging (DTI) [2] [3] require a low acquisition time but provides low angular detail, resulting in low-quality reconstructions, while most complex schemes like Diffusion Spectrum Imaging (DSI) [5] [15] provide the highest imaging quality but their acquisition time is too large, which prevents its use in a clinical environment. In recent years, many techniques have been proposed aiming either at reducing the acquisition time or increasing the reconstruction quality.

The rest of the paper is organized as follows. In Section 2, we describe the Diffusion Basis Function (DBF) reconstruction algorithm proposed by A. Ramírez *et al.* [12], which our method is based on, and explain its limitations. In Section 3, we present our proposed Dynamic DBF reconstruction algorithm. Experimental results are reported in Section 4. We discuss the properties and limitations of our proposal in Section 5 and conclude in Section 6 with a summary of our work. Our contributions are: (i) the development of an iterative Dynamic DBF construction algorithm with the property that the input DW-MRI signal can be represented as a sparse positive linear combination of the elements of the limiting DBF; (ii) the development of a variation of K-means for orientation vectors; (iii) the development of an axon fiber estimation algorithm from DW-MRI that presents a consistently lower reconstruction error than the state-of-the-art basis-pursuit re-

construction algorithms according to our experiments performed on the publicly available 2012 HARDI Reconstruction Challenge dataset<sup>1</sup>.

## 2 Previous work

One of the most important drawbacks of the MT model (Eq. 2) is the computational complexity of the optimization required to estimate its parameters, which makes most optimization techniques prone to get stuck at local minima. The most successful approach to fit the MT model is the basis-pursuit algorithm proposed by A. Ramírez *et al.* [12]. In their work, a basis of diffusion functions (DBF)  $\Phi = [\phi_1, \phi_2, \dots, \phi_m]$  was defined, where each  $\phi_i$  is of the form

$$\phi_i(q_k) = S_0 \exp(-\tau q_k^T D_i q_k). \quad (3)$$

To define the shape of the diffusion functions, the vector of eigenvalues of  $D_i$  (which is referred to as the *diffusivity profile*) is assumed to be of the form  $\Lambda = (\lambda_R, \lambda_R, \lambda_L)$ , where  $\lambda_R$  is the *radial diffusivity* and  $\lambda_L > \lambda_R$  is the *longitudinal diffusivity*. Therefore, by using the spectral decomposition of  $D_i$  (and using  $\|q_k\| = 1$ ), equation 3 can be written in terms of the PDD,  $p_i$ , of  $D_i$  and the diffusivity profile as:

$$\phi_i(q_k) = S_0 \exp(-\tau(\lambda_L - \lambda_R)(q_k^T p_i)^2) \exp(-\tau\lambda_R). \quad (4)$$

The assumption that the two smallest eigenvalues are equal to each other and significantly smaller than the longitudinal diffusivity is based on the observed properties of very coherent white matter bundles in the brain where, typically,  $\lambda_L \approx 5 \times \lambda_R$  [13] [12] [6]. The diffusivity profile is estimated by selecting a region of interest (ROI) in the brain that does not contain fiber crossings (*e.g.* the corpus callosum) [12] and fitting one single DT to each voxel, the diffusivity profile is then computed as the average diffusivity profile along the ROI. The PDD's  $P = \{p_1, p_2, \dots, p_m\} \subset \mathcal{S}^3$  of the tensors are chosen to uniformly cover a unit hemisphere in  $\mathbb{R}^3$ . Since the acquisition scheme is the same for all the elements of  $\Phi$ , the DBF is fully characterized by  $P$  and  $\Lambda$ .

Once the DBF has been built, the input DW-MRI signal is modeled as a positive linear combination of the elements in the dictionary:

$$S_p \approx \Phi\beta. \quad (5)$$

The estimation problem is reduced to a nonnegative least squares problem [7]:

$$\min_{\beta \in \mathbb{R}^m} \|\Phi\beta - S\|_2^2, \text{ s.t. } \beta \geq 0. \quad (6)$$

Ideally, only the diffusion functions whose PDD is aligned to the *true* axon fiber directions should be assigned a nonzero coefficient, thus, it is reasonable to expect a sparse optimal coefficient vector  $\hat{\beta}$ . However, in practice this is not the

<sup>1</sup> <http://hardi.epfl.ch/index.html>

case. Figure 2b depicts a typical situation in which the optimal  $\hat{\beta}$  is not sparse. As a consequence of the lack of sparsity, a post-processing of the optimal coefficient vector  $\hat{\beta}$  is necessary in order to correctly estimate the number of fibers in the voxel. For each PDD  $p_j \in P$ , its  $K$ -neighborhood  $N_K(p_j) \subset P$  is defined as the  $K$  elements in  $P$  with smallest angle with  $p_j$ . The  $K$ -neighborhood of the PDD's with nonzero  $\hat{\beta}$  coefficient define a graph  $G(V, E)$ , where the set of vertices is  $V = \{p_i \in P | \hat{\beta}_i > 0\}$ . There is an edge  $e_{i,j} \in E$  between  $p_i$  and  $p_j$  if and only if  $p_i \in N_K(p_j)$  or  $p_j \in N_K(p_i)$ . The *clusters*, are the connected components of  $G$ . The *centroid* (a unit orientation vector) of each cluster is computed as a convex linear combination of its elements. Finally, the diffusion functions, whose PDD is defined by the centroids (with the same diffusivity profile  $\Lambda$ ) are reported as the components of the multi-tensor [12], in particular, the estimated number of compartments is the number of connected components of  $G$ .

One of the main drawbacks of the clustering process based on connected components is that the estimated number of compartments depends on the neighborhood size  $K$ , for example, in the extreme case of setting  $K = m - 1$ , the number of connected components of  $G$  will always be 1, while setting  $K = 0$ , the number of connected components will be precisely the number of vertices  $|V|$ . Thus, by setting  $K$  to a small value, the number of compartments will be overestimated, while setting it to a large value, the number of compartments will be underestimated. Another drawback is that two PDD's  $p_a, p_b$  whose angle is large (they are far apart from each other) may be wrongly grouped together if the  $\hat{\beta}$  vector contains a chain of positive coefficients assigned to PDD's  $p_1, p_2, \dots, p_t$  such that  $p_a \in N_K(p_1)$ ,  $p_1 \in N_K(p_2)$ , ...,  $p_{t-1} \in N_K(p_t)$ ,  $p_t \in N_K(p_b)$ , thus forming a connected component.

The above limitations are a consequence of using a *static* DBF to represent any arbitrary DW-MRI signal, which in general do not admit a sparse representation under the static base. Many attempts have been done to promote the  $\hat{\beta}$  coefficient vector to be sparse [1][8][9][10]. The most direct approaches add a penalty to the  $\beta$  coefficient vector in equation (6), yielding the *Positive-LASSO* model:

$$\min_{\beta \in \mathbb{R}^m} \|\Phi\beta - s\|_2^2 + \lambda \|\beta\|_1, \text{ s.t. } \beta \geq 0, \quad (7)$$

where the hyper parameter  $\lambda$  controls the sparsity of the solution. The main drawback of directly enforcing the solution to be sparse is that important information about the orientation of the *true* diffusion function is dropped, thus increasing the angular error of the reconstruction. The fundamental problem behind these sparsity-promoting approaches is that an arbitrary multi-tensor signal does not admit, in general, a sparse representation as a positive linear combination of a relatively small static DBF. The multi-tensor signal will admit such a sparse representation only if the *true* diffusion tensors are contained in the DBF. However, attempting to build a unique DBF under which any arbi-

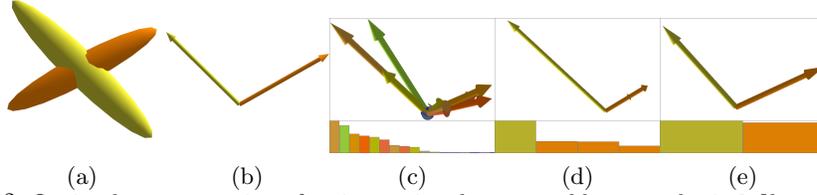


Fig. 2: Optimal representation of an input signal generated by a synthetic 2-fiber crossing. The PDD of each diffusion function with a nonzero coefficient is depicted as an arrow with length proportional to its corresponding  $\hat{\beta}$  coefficient. The sorted coefficients of  $\hat{\beta}$ , are depicted as a histogram. (a) Synthetic multi-tensor. (b) PDD's of the synthetic multi-tensor. (c) Optimal representation using a static DBF, 17 diffusion functions were assigned a positive coefficient. (d) Optimal representation using the proposed dynamic DBF (without promoting sparsity). Notice that the signal admits a sparser representation under the dynamic DBF (three out of four nonzero elements are highly correlated). (e) Optimal representation using the proposed dynamic DBF after promoting sparsity.

bitrary input signal can be sparsely represented would require a large number of elements.

### 3 Dynamic DBF construction

Our Dynamic DBF construction algorithm is based on the following observation. Let  $\Phi$  be a DBF with PDD's  $P = \{p_1, p_2, \dots, p_m\}$  uniformly distributed along the unit hemisphere in  $\mathbb{R}^3$  and diffusivity profile  $\Lambda = (\lambda_R, \lambda_R, \lambda_L)$ . We randomly generated a set of vectors  $U \subset \mathcal{S}^3$  uniformly distributed along the unit sphere (independent from  $P$ ) and for each  $u \in U$  we created a synthetic diffusion function  $S^{(u)}$  from the tensor model (Eq. 1) using  $u$  as the PDD of the tensor and  $\Lambda$  as the diffusivity profile. We then solved Eq. 6:

$$\hat{\beta}^{(u)} = \arg \min_{\beta \in \mathbb{R}^m} \|\Phi\beta - S^{(u)}\|_2^2, \text{ s.t. } \beta \geq 0. \quad (8)$$

Figure 3 depicts the scatter plot of pairs  $(\angle(u, p_i), \hat{\beta}_i^{(u)})$ ,  $\forall u \in U, \hat{\beta}_i^{(u)} > 0$ . The behavior depicted in figure 3 is not surprising, it shows that, *with high probability*, the solution  $\hat{\beta}^{(u)}$  of the Non-negative least squares problem (Eq. 7), assigns a higher coefficient  $\hat{\beta}_i^{(u)}$  to diffusion functions in  $\Phi$  whose PDD  $p_i$  makes a smaller angle with the PDD  $u$  of the *true* tensor that generated the signal  $S^{(u)}$ . More specifically:

**Property 1:** Let  $\hat{\beta}^{(u)}$  be the solution of Eq. 8 for a given DBF  $\Phi$  with PDD's  $P = \{p_1, p_2, \dots, p_m\}$  uniformly distributed along the unit sphere  $\mathcal{S}^3 \subset \mathbb{R}^3$ . Let  $S^{(u)}$  be the input signal generated from a diffusion tensor with PDD  $u$ . If  $\hat{\beta}_j^{(u)} > \hat{\beta}_i^{(u)} > 0$ , then *with high probability*,  $\angle(u, p_i) > \angle(u, p_j)$ .

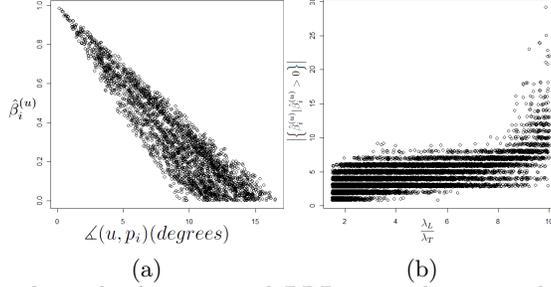


Fig. 3: (a) For each randomly generated PDD  $u$  on the unit sphere, the synthetic diffusion function  $S^{(u)}$  was generated with PDD  $u$  and the same diffusivity profile as the DBF  $\Phi$ . After solving eq. 8, the positive coefficients  $\hat{\beta}_i^{(u)} > 0$  and their corresponding PDD's  $p_i$  define a point on the scatterplot  $(\angle(u, p_i), \hat{\beta}_i^{(u)})$ , where  $\angle(u, p_i)$  is the angle between  $p_i$  and the *true* PDD  $u$ . Note that a large  $\hat{\beta}_i^{(u)}$  coefficient is strong evidence that its corresponding PDD  $p_i$  is oriented towards the *true* PDD of the tensor (small angle). (b) Number of nonzero coefficients (sparsity) of the optimal vector  $\hat{\beta}^{(u)}$  as a function of the ratio  $\frac{\lambda_L}{\lambda_R}$  used to create the DBF (small jitter was added for visualization purposes).

Although a rigorous proof of Property 1 is difficult to derive, it can be verified experimentally. For example, in the experiment previously described, the empirical probability described above satisfies:

$$\widehat{P} \left( \angle(u, p_i) > \angle(u, p_j) \mid \hat{\beta}_j^{(u)} > \hat{\beta}_i^{(u)} > 0 \right) > \frac{9}{10} \quad (9)$$

It is not difficult to show that whenever Property 1 holds, there exists an arc  $A = \widehat{p_i, p_j}$  on the unit sphere (whose endpoints are PDD's with positive  $\hat{\beta}$  coefficients  $\hat{\beta}_i < \hat{\beta}_j$ ) whose bisector  $\mu_{i,j} \in A$  satisfies  $\angle(u, \mu_{i,j}) < \angle(u, p_i)$ . Therefore, the coefficient  $\hat{\beta}_{i,j}^{(u)}$  assigned to  $\mu_{i,j}$  in the optimal solution of Eq.8 using a DBF  $\Phi'$  with PDD's  $P' = P \cup \{\mu_{i,j}\}$  will satisfy  $\hat{\beta}_i^{(u)} < \hat{\beta}_{i,j}^{(u)}$ . In other words, the DBF was *improved* by adding  $\mu_{i,j}$ , in the sense that  $\mu_{i,j}$  is *closer* to  $u$  than  $p_i$ .

A second property that can be observed through experimentation is that there is a correlation between the sparsity (*i.e.*, the number of nonzero coefficients) of the optimal vector  $\hat{\beta}^{(u)}$  and the isotropy of the diffusivity profile. Figure 3b depicts the sparsity of the optimal vector  $\hat{\beta}^{(u)}$  as a function of the ratio of the longitudinal diffusivity,  $\lambda_L$ , to the radial diffusivity,  $\lambda_R$ , of the profile used to create the DBF. Note that more isotropic DBF's result in sparser representations. Inspired on the above observations, we enhance the DBF by adding the bisector of every pair of PDD's which were assigned a positive coefficient:

$$\mu_{i,j} \leftarrow \frac{p_i + p_j}{\|p_i + p_j\|}, \quad \forall i, j : \hat{\beta}_i > 0, \hat{\beta}_j > 0. \quad (10)$$

When no further improvement is achieved by refining the DBF, we increase the isotropy of the diffusivity profile and continue the DBF refinement. The Dynamic DBF construction is summarized in Algorithm 1. Figure 4 depicts five iterations of the Dynamic DBF algorithm. Note that the optimal representation of the input signal using the limiting DBF does not contain the correct number of compartments (two, in this case), the reason is that some elements of the limiting DBF are highly correlated and the nonnegative least squares algorithm [7] tends to assign positive coefficients to all elements that are correctly aligned to the *true* tensors of the input signal. However, using the limiting DBF, sparsity can be safely promoted by solving for example the *Positive-LASSO* model (Eq. 7) or the model recently proposed by R. Aranda *et al.*[1]. An advantage of the later approach over more conventional algorithms for promoting sparsity, like the Positive-LARS algorithm [4], for example, is that it does not rely on the *one at a time* condition, which is not satisfied by the limiting DBF, since its components are highly correlated.

---

### Algorithm 1 Dynamic DBF Construction

---

**Require:** Target DW-MRI signal  $S_r$  at target voxel  $r \in \Omega$   
**Require:** Initial diff. profile  $\lambda^{(0)} = (\lambda_T^{(0)}, \lambda_T^{(0)}, \lambda_L)$ , as proposed in [12]  
**Require:** Initial PDD's of the DBF  $P^{(0)} = \{p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)}\}$ , as proposed in [12]  
**Require:** Isotropy increase rate  $\gamma > 1$ , see line 17 (default:  $\gamma = 1.01$ )  
**Require:** Minimum anisotropy required  $K > 0$ :  $\lambda_L$  will be larger than  $K\lambda_R$  (default:  $K = 3$ )  
**Require:** Tolerance  $\tau > 0$  (default:  $\tau = 10^{-9}$ )  
**Require:** Maximum nonzero coefficients allowed  $L$  (default:  $L = 6$ )

- 1: Compute initial DBF  $\Phi^{(0)}$  using  $P^{(0)}$  and  $\lambda^{(0)}$
- 2:  $\hat{\beta}^{(0)} \leftarrow 0$
- 3:  $\epsilon^{(0)} \leftarrow \|S_r\|$
- 4:  $t \leftarrow 0$
- 5: **repeat**
- 6:      $t \leftarrow t + 1$
- 7:     Solve  $\hat{\beta}^{(t)} = \arg \min_{\beta \in \mathbb{R}^m} \|\Phi^{(t-1)}\beta - S_r\|^2$  s.t.  $\beta \geq 0$
- 8:      $\epsilon^{(t)} \leftarrow \|\Phi^{(t-1)}\hat{\beta}^{(t)} - S_r\|$
- 9:     **for all**  $p_i^{(t-1)} \in P^{(t-1)} : \hat{\beta}_i > 0$  **do**
- 10:         **for all**  $p_j^{(t-1)} \in P^{(t-1)} : i < j, \hat{\beta}_j > 0$  **do**
- 11:              $\mu_{i,j} \leftarrow \frac{p_i^{(t-1)} + p_j^{(t-1)}}{\|p_i^{(t-1)} + p_j^{(t-1)}\|}$
- 12:         **end for**
- 13:     **end for**
- 14:      $\Psi^{(t-1)} \leftarrow \{\mu_{i,j} | i < j, \hat{\beta}_i > 0, \hat{\beta}_j > 0\}$
- 15:      $P^{(t)} \leftarrow \Psi^{(t-1)} \cup \{p_i^{(t-1)} \in P^{(t-1)} | \hat{\beta}_i > 0\}$
- 16:     **if**  $[\epsilon^{(t-1)} - \epsilon^{(t)} < \tau]$  and  $[\gamma\lambda_T^{(t-1)} < \frac{1}{K}\lambda_L]$  **then**
- 17:          $\lambda^{(t)} = (\gamma\lambda_T^{(t-1)}, \gamma\lambda_T^{(t-1)}, \lambda_L)$
- 18:     **else**
- 19:          $\lambda^{(t)} = \lambda^{(t-1)}$
- 20:     **end if**
- 21:     Compute new DBF  $\Phi^{(t)}$  using  $P^{(t)}$  and  $\lambda^{(t)}$
- 22: **until**  $[\epsilon^{(t-1)} - \epsilon^{(t)} < \tau]$  and  $[[\gamma\lambda_T^{(t-1)} \geq \frac{1}{K}\lambda_L]$  or  $[\|\hat{\beta}^{(t)}\|_0 \leq L]$
- 23: **return**  $P^{(t)}$

---

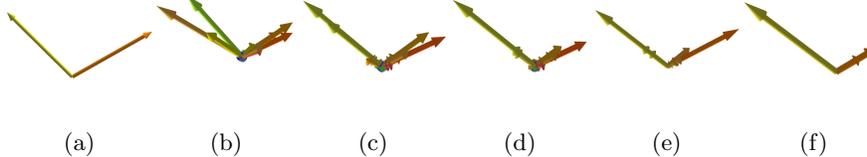


Fig. 4: Dynamic DBF evolution for the signal corresponding to the 2-compartment multi-tensor depicted in Fig. 2a. (a) Ground Truth. (b) Optimal representation using the initial static DBF (1 iteration). (c-f) Optimal representation using the dynamic DBF at iterations 18, 37, 56 and 75, respectively.

### 3.1 Clustering algorithm for the Dynamic DBF

In the absence of noise, the optimal representation of the input signal as a sparse linear combination of the elements of the limiting Dynamic DBF effectively identifies the number of compartments of the input signal (see Table 2). However, when noise is added, the number of nonzero coefficients over-estimates the number of compartments. Thus, clustering is still needed. Notice that the clustering based on connected-components proposed by A. Ramírez *et al.* [12] cannot be used with the Dynamic DBF, since its elements are no longer uniformly distributed along the unit hemisphere.

Given a set of vectors  $C = \{x_1, x_2, \dots, x_m\}$ , we define its *orientation centroid*  $\mu$

Comp.	$\ \hat{\beta}\ _0$	No noise				SNR=40				SNR=10			
		1	2	3	4...48	1	2	3	4...15	1	2	3	4...12
1	0.0%	0.0%	51.3%	48.7%	0.0%	0.3%	33.0%	66.7%	0.0%	3.5%	20.5%	76.0%	
2	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	100%	0.0%	0.0%	1.0%	99.0%	
3	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	100%	0.0%	0.0%	0.7%	99.3%	

Table 1: Number of nonzero  $\beta$  coefficients in the optimal representation (Eq. 6) of the 1,280 input signals of the HARDI Challenge spatially-coherent training set, using the Static DBF proposed by A. Ramírez *et al.*[12]. The number of nonzero coefficients does not reflect the number of compartments, even in the absence of noise.

Comp.	$\ \hat{\beta}\ _0$	No noise				SNR=40				SNR=10			
		1	2	3	> 3	1	2	3	4...5	1	2	3	4...5
1	100.0%	0.0%	0.0%	0.0%	65.1%	23.4%	10.3%	1.3%	27.2%	39.1%	27.2%	6.4%	
2	0.0%	100.0%	0.0%	0.0%	0.0%	72.1%	26.7%	1.2%	0.0%	28.1%	55.8%	16.1%	
3	0.0%	50.0%	50.0%	0.0%	0.0%	47.9%	50.7%	1.4%	0.0%	27.9%	57.9%	14.3%	

Table 2: Number of nonzero  $\beta$  coefficients in the optimal representation (Eq. 6) of the 1,280 input signals of the HARDI Challenge spatially-coherent training set, using the limiting Dynamic DBF. In the absence of noise, the number of nonzero coefficients coincides with the *true* number of compartments when the signal comes from 1 and 2 compartments (rows 1 and 2, respectively).

as

$$\mu_C = \arg \max_{u \in \mathbb{R}^n: \|u\|_2=1} \sum_{x_i \in C} (x_i^T u)^2 = \arg \max_{u \in \mathbb{R}^n: \|u\|_2=1} u^T X X^T u, \quad (11)$$

where  $X = [x_1, x_2, \dots, x_m]$ . Therefore  $\mu_C$  is the eigenvector of  $X X^T$  corresponding to its maximum eigenvalue. On the other hand, given a set of centroids  $M = \{\mu_{C_1}, \mu_{C_2}, \dots, \mu_{C_k}\}$ , the closest element  $\hat{\mu}(x)$  in  $M$  to an arbitrary unit vector  $x$  is

$$\hat{\mu}(x) = \arg \max_{\mu \in M} (x^T \mu)^2. \quad (12)$$

Equations 11 and 12 define a variation of the K-means algorithm that is well suited for orientation vectors.

## 4 Experiments

Our experiments were conducted on the publicly available datasets used in the 2012 HARDI Reconstruction Challenge, developed in the context of the 2012 IEEE International Symposium on Biomedical Imaging. The first test set consists of 9,100 synthetic independent signals (they are independent in the sense that they have no spacial structure). The second test set consists of 1,280 voxels arranged in five  $16 \times 16$  slices from a 3D synthetic phantom, of smooth fibers. We used an acquisition scheme consisting of 48 diffusion orientations and a  $b = 1500$  value.

One of the limitations of the static DBF model is that the number of compartments is consistently over-estimated [11]. One way to overcome this limitation is to increase the parameter  $K$  (the neighborhood size) of the clustering algorithm. In order to illustrate the effect of increasing the neighborhood size, we compared two versions of the static DBF. The first version (referred to as Static DBF-A) uses the default configuration,  $K = 6$ . Notice that this algorithm was ranked  $3^{rd}$  among the participants of the HARDI reconstruction Challenge. The second version (referred to as Static DBF-B) uses  $K = 16$ , this value was the configuration that yielded the best results on the spatially-coherent training set. Our proposal is referred to as Dynamic DBF. The evaluation takes into account two criteria. First, the number of wrongly estimated compartments is reported in Figs. 5 and 6 for the independent-voxel and spatially-coherent sets, respectively. The number of underestimated compartments (denoted  $N-$ ), the number of over-estimated compartments (denoted  $N+$ ), and the total wrongly estimated compartments (denoted  $N_T$ ) are reported separately. The results for the spatially-coherent set are summarized in Fig. 6. Notice that the Static DBF-A algorithm effectively reduces  $N+$  but at the cost of an increase of  $N-$ , which results in an overall decrease in performance, as can be seen from the increase in  $N_T$ . The Dynamic DBF algorithm consistently reduces  $N+$  without significantly increasing  $N-$ , as can be seen from the consistent reduction of  $N_T$ . The second evaluation criterion is the angular precision. The angular error of an estimated voxel  $v_E$ , with respect to the ground truth voxel  $v_G$  is computed as follows. Assume  $v_E$  was estimated

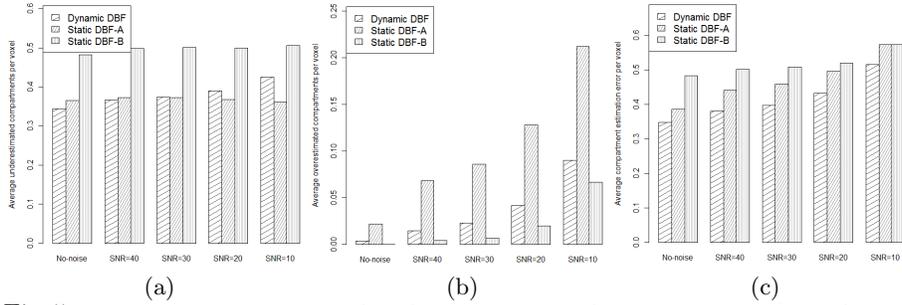


Fig. 5: Average compartment estimation error on the independent-voxels testing set. (a) Underestimated number of compartments. (b) Overestimated number of compartments. (c) Total wrongly estimated compartments.

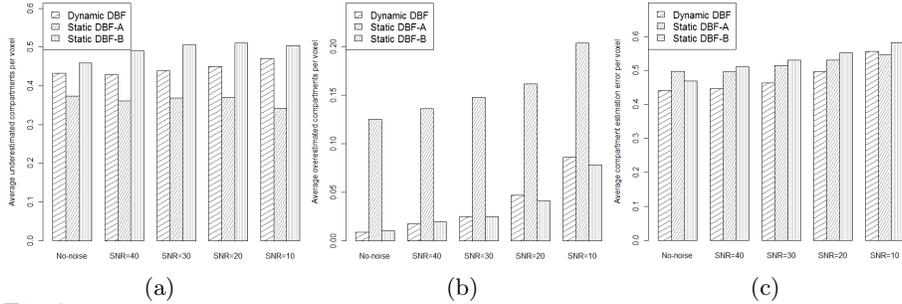


Fig. 6: Average compartment estimation error on the spatially-coherent testing set. (a) Underestimated number of compartments. (b) Overestimated number of compartments. (c) Total wrongly estimated compartments.

to have  $a$  compartments, and the *true* number of compartments (from  $v_G$ ) is  $b$ , then only  $\min\{a, b\}$  compartments from  $v_E$  can be assigned to the compartments from  $v_G$ . We solve this assignment problem using the Hungarian Algorithm. The assignment is done in such a way that the sum of the angles of the PDD's of the paired compartments is minimized (in other words, we compute the best possible assignment in terms of sum of angles). The angular error is computed as the average angle of the paired PDD's. The angular error for different levels of noise is reported in Figs. 7a and 7b for the independent-voxel and spatially-coherent sets, respectively. Note that the Dynamic DBF consistently reduces the angular error.

## 5 Discussion

The pipeline of the proposed algorithm for fiber tract geometry reconstruction is very similar to the standard pipeline that uses a static DBF, it consists of an estimation step (computation of the optimal  $\hat{\beta}$  coefficient), and a clustering

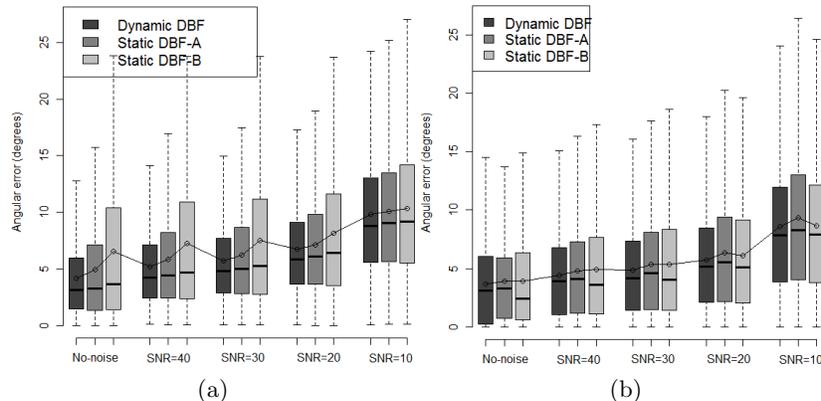


Fig. 7: Angular error for different levels of noise. The line joining the boxplots depicts the average angular error. (a) Independent-voxels testing set. (b) Spatially-coherent testing set.

step. Since using the limiting DBF does not eliminate the need of clustering, it may seem pointless to use the Dynamic DBF instead of the original static DBF approach proposed by A. Ramírez *et al.*[12], however, as shown in the experiments, using the Dynamic DBF consistently reduces the overestimation of the number of compartments (without significantly increasing its underestimation) while consistently reducing the angular error of the reconstruction as well. Although the computational time required to fit the Dynamic DBF is higher than the original DBF formulation that uses a static DBF, it is reasonably low: the original DBF (static) has a processing speed of 435.4 voxels per second, while using the dynamic DBF the processing speed reduces to 75.7 voxels per second running on a Intel i5 processor at 1.7 GHz (using only one core). Notice that this approach is easily parallelizable, since the processing is voxel-wise, for example using a modest computer with 4-cores at 1.7 GHz, a realistic brain volume of around  $10^6$  white-matter voxels could be processed in less than an hour. Convergence may still be improved by taking advantage of the spatial regularity of the anatomical structure of the brain: the limiting Dynamic DBF of a voxel is expected to be similar to the limiting Dynamic DBF of at least one of its neighbors.

## 6 Conclusion

We presented an algorithm for dynamically constructing a dictionary of diffusion functions for recovering the intra-voxel axon fiber tract geometry from diffusion-weighted magnetic resonance images of the brain. The limiting DBF has the property that the input DW-MRI signal can be represented as a sparse positive linear combination of the elements of the DBF, which in general cannot be done using a static DBF. Experimental results show that using the limiting Dynamic

DBF consistently reduces the overestimation of the number of compartments while consistently reducing the angular error of the reconstruction.

**Acknowledgments.** The authors would like to thank Dr. Alonso Ramírez-Manzanares and the anonymous reviewers for their useful comments that helped us to improve the quality of this paper. This research was supported in part by CONACyT grants 131369 and 169178.

## References

1. R. Aranda, M. Rivera, and A. Ramírez-Manzanares. Improved diffusion basis functions fitting and metric distance for brain axon fiber estimation. In *Proc. The 5th Pacific Rim conference on Advances in Image and Video Technology - Volume Part II*, PSIVT'11, pages 36–47, Berlin, Heidelberg, 2012. Springer-Verlag.
2. P. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259 – 267, 1994.
3. P. J. Basser and C. Pierpaoli. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance*, 213(2):560 – 570, 2011.
4. B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Annals of Statistics*, 32:407–499, 2004.
5. P. Hagmann. *From diffusion MRI to brain connectomics*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2005.
6. R. Henson. Introduction to functional magnetic resonance imaging: Principles and techniques. *Brain*, 126(9):2110, Sept. 2003.
7. C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*, chapter 23, page 161. 1974.
8. M. Mani, M. Jacob, A. Guidon, C. Liu, A. Song, V. Magnotta, and J. Zhong. Acceleration of high angular and spatial resolution diffusion imaging using compressive sensing. In *Proc. IEEE International Symposium on Biomedical Imaging*, pages 1–4, 2012.
9. M. Mani, J. Mathews, and J. Zhong. Compressed sensing reconstruction of multi-tensor models. In *Proc. 2012 IEEE International Symposium on Biomedical Imaging Workshop on HARDI Reconstruction*, page 11, 2012.
10. M. Paquette and M. Descoteaux. Sparse L1-L1 Multi-Tensor Imaging at the Price of DTI. In *Proc. 2012 IEEE International Symposium on Biomedical Imaging Workshop on HARDI Reconstruction*, page 9, 2012.
11. A. Ramirez-Manzanares, P. A. Cook, M. Hall, M. Ashtari, and J. C. Gee. Resolving Axon Fiber Crossings at Clinical b-values: An Evaluation Study. *Journal of Medical Physics*, 38(9):5239–53, 2011.
12. A. Ramirez-Manzanares, M. Rivera, B. C. Vemuri, P. Carney, and T. Mareci. Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. *IEEE Transactions on Medical Imaging*, 26(8):1091–1102, 2007.
13. D. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, and J. Belliveau. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582, 2002.
14. D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, Dec. 2004.
15. V. J. Wedeen, P. Hagmann, W. Y. Tseng, T. G. Reese, and W. R. M. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54:1377–1386, 2005.

# Complete Fourier Direct Magnetic Resonance Imaging (CFD–MRI) Equations for Diffusion MRI Signal

Alpay Özcan

Health Research, Arlington Innovation Center,  
Virginia Polytechnic Institute and State University, 900 N. Glebe Road, Arlington  
VA, 22203, USA.,  
alpay@vt.edu  
<http://www.aic.ncr.vt.edu>

**Abstract.** By carefully re-examining the first principles of diffusion weighted magnetic resonance imaging (DW–MRI) signal formation and deriving the mathematical model from scratch, the foundation for an accurate and unifying Fourier based theory is constructed. The derivations are specifically obtained for DW–MRI signal by including imaging gradients and without taking the magnitude of the signal using particle methods instead of the conventional partial differential equations approach. The signal is shown to be the Fourier transform of the joint distribution of number of the magnetic moments (at a given location at the initial time) and their displacement integrals which augments the  $k$ -space by three more dimensions. The joint distribution function is recovered by applying the Fourier transform to the complete high-dimensional data set. A physically meaningful real valued distribution function is obtained after phase corrections are applied for the re-establishment of Hermitian symmetry in the signal. Consequently, the method is fully unconstrained and directly presents the distribution of displacement integrals without any assumptions such as symmetry or Markovian property. The joint distribution function is visualized with isosurfaces overlaid on the distribution map of the number of magnetic moments with low mobility. The presentation provides a highly detailed level of microstructural information. The results are demonstrated on the experimental data obtained from an ex-vivo baboon brain.

**Keywords:** magnetic resonance imaging, diffusion weighted imaging, Fourier transform

## 1 Introduction

Diffusion weighted (DW) MRI's ability to estimate molecular displacements provides information about the microstructure within which the motion takes place. The technique has successfully been used in a variety of clinical and research

areas such as cancer diagnosis [1], white matter integrity assessment [2], in monitoring of neurological diseases such as multiple sclerosis [3] and disorders [4] as well as neonatal development [5].

However, the existence of several methods for modeling the DW–MRI signal is proof that the field has not yet reached a satisfactory consensus. A thorough mathematical analysis and a description of the subject and the background of existing methods is presented in [6]. Herein, to provide a unified approach, the MRI signal in the presence of diffusion sensitizing gradients is derived by carefully re-examining the first principles of DW–MRI signal formation and deriving the mathematical model from scratch using particle methods in the spirit of the work of McCall et al. [7] rather than starting from the conventional partial differential equations (PDE) of Stejskal and Tanner [8]. The outcome is the Complete Fourier Direct (CFD) MRI method [9] that establishes a high dimensional Fourier relationship between the signal and the joint distribution of position and displacement integrals of the particles.

## 2 Signal Formation

The derivation of the CFD–MR signal model is accomplished by simply expressing the magnetization of each magnetic moment. For this purpose, the effect of spin–spin relaxation can be neglected. The evolution of the transverse magnetization of the  $i^{th}$  magnetic moment is described in a standard manner by a rotating magnetization vector:

$$m_i(t) = \exp(-j\gamma\Omega_i) m_i(t_0). \quad (1)$$

Here,  $\gamma$  is the gyromagnetic ratio, the transverse magnetization vector,  $m_i$ , is written in complex number form with  $m_i(t_0)$  denoting the initial magnetization tipped to the transverse plane,

$$\Omega_i = \int_{t_0}^t G(x_i, \tau) \cdot x_i(\tau) d\tau \quad (2)$$

describes the phase as a function of the magnetic field gradients  $G(x, t) \in \mathbb{R}^3$ , and the position of the magnetic moment is  $x_i \in \mathbb{R}^3$ .

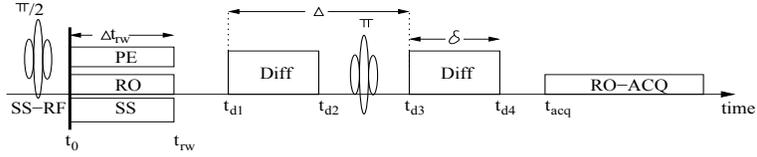
The signal originates from all of the magnetic moments

$$M(t) = \sum_i m_i(t) = \sum_i \exp(-j\gamma\Omega_i) m_i(t_0). \quad (3)$$

In order to precisely obtain the CFD–MR signal, the time evolution of the phase,  $\Omega_i$ , in Eq. 2 must be explicitly written on the right hand side of Eq. 3 by expressing  $x_i$ . The time dependent position of each magnetic moment is given as:

$$x_i(t) = x_i(t_0) + w_i(t) \quad (4)$$

where  $w_i(t) \in \mathbb{R}^3$  represents the displacement of the magnetic moment from its initial position with  $w_i(t_0) = 0$ . The function  $w_i(t)$  could express any kind of displacement such as Brownian motion, molecular movement in biological tissue with different medium and obstacles, coherent motion or any combination thereof. The function does not have to be differentiable but it must be continuous to make physical sense since a magnetic moment cannot disappear at a given point and reappear at another. No other requirements and/or assumptions about motion properties in the time domain, deterministic and/or stochastic, such as being Markovian, is required in the derivations of the CFD model. Based on



**Fig. 1.** The pulsed–gradient spin–echo (PGSE) pulse sequence and the definition of the variables used in the calculations. RO is for read out, PE for phase encode, SS for is the slice select gradient, Diff marks the motion sensitizing pulses and ACQ stands for digital acquisition period. In practice, the MR pulse sequences implement the rewind (rw) gradients such that the amplifiers are turned on and off at the same times. The magnetic field gradient pulses during the radio frequency (RF) pulses are not shown because their effect is overwhelmed by the RF pulse (see also the text).

this model, the phase of the  $i^{\text{th}}$  magnetic moment during the digital acquisition period of the pulsed–gradient spin–echo (PGSE) sequence of Fig. 1 (for  $l_{\text{mr}} = 2$  dimensional imaging slice) is obtained after tedious but routine derivations as follows. In the beginning of each acquisition, a  $\pi/2$  radio frequency (RF) pulse tips the longitudinal magnetization to the transversal plane. The initial time,  $t_0$ , is chosen accordingly to be the end time of the RF pulse when the magnetization is fully tipped to the transversal plane. Using the definitions of the variables in Fig. 1, the evolution of the phase is described as:

1) First, imaging gradients for read out rewinding,  $G_{\text{ro}} \in \mathbb{R}^3$ , phase encoding,  $G_{\text{pe}} \in \mathbb{R}^3$  and slice select  $G_{\text{ss}} \in \mathbb{R}^3$  on the time interval  $[t_0, t_{\text{rw}}]$  are turned on. After these gradients are applied the phase and the magnetization become

$$\begin{aligned}
\Omega_{\text{rw}} &= \int_{t_0}^{t_{\text{rw}}} (G_{\text{ro}}(\tau) + G_{\text{pe}}(\tau) + G_{\text{ss}}(\tau)) \cdot x_i(\tau, t_0) d\tau \\
&= (G_{\text{ro}} + G_{\text{pe}} + G_{\text{ss}}) \cdot \left( \Delta t_{\text{rw}} x_i(t_0) + \int_{t_0}^{t_{\text{rw}}} w_i(\tau) d\tau \right) \\
&= \Delta t_{\text{rw}} (G_{\text{ro}} + G_{\text{pe}} + G_{\text{ss}}) \cdot x_i(t_0)
\end{aligned}$$

$$+(G_{\text{ro}} + G_{\text{pe}} + G_{\text{ss}}) \cdot \int_{t_0}^{t_{\text{rw}}} w_i(\tau) d\tau. \quad (5)$$

Note that the calculations are carried out assuming ideal gradient amplifiers resulting in rectangular shaped gradient pulses in the formulation. Routine derivations for trapezoidal shapes are omitted for clarity.

2) The  $\pi$  RF pulse between the diffusion gradient pulses,  $G_{\text{D}} \in \mathbb{R}^3$ , and the accompanying slice select gradient provide theoretical sign reversal of the phase. The slice select gradient during this period does not encode magnetic moment motion into the signal (thus it is not shown in Fig. 1). This is a direct consequence of the RF pulse imposes a strong magnetic field that changes all of the entries of the main magnetic field  $B_0$ . In comparison, the slice select gradient's modification of the third component of the  $B_0$  vector is negligible. Therefore the strong (RF) magnetic field rotates the magnetization vector  $m_i$  of each magnetic moment without being affected by diffusion.

Since  $x_i(t_{dk}) = x_i(t_0) + w_i(t_{dk})$ ,  $k = 1, \dots, 4$ ; at  $t = t_{d4}$

$$\begin{aligned} \Omega_{\text{D}} &= \int_{t_{d3}}^{t_{d4}} G_{\text{D}} \cdot x_i(\tau, t_0) d\tau - \int_{t_{d1}}^{t_{d2}} G_{\text{D}} \cdot x_i(\tau, t_0) d\tau \quad (6) \\ &= G_{\text{D}} \cdot \left( (t_{d4} - t_{d3}) x_i(t_0) + \int_{t_{d3}}^{t_{d4}} w_i(\tau) d\tau \right) \\ &\quad - G_{\text{D}} \cdot \left( (t_{d2} - t_{d1}) x_i(t_0) + \int_{t_{d1}}^{t_{d2}} w_i(\tau) d\tau \right) \\ &= G_{\text{D}} \cdot \left( \int_{t_{d3}}^{t_{d4}} w_i(\tau) d\tau - \int_{t_{d1}}^{t_{d2}} w_i(\tau) d\tau \right) \\ &\quad + ((t_{d4} - t_{d3}) - (t_{d2} - t_{d1})) G_{\text{D}} \cdot x_i(t_0). \quad (7) \end{aligned}$$

The sign change on  $\Omega_{\text{rw}}$  is the effect of the  $\pi$  pulse, which is incorporated in the expression for  $\Omega_{\text{D}}$ . If the diffusion gradient times are equal, i.e.,  $t_{d4} - t_{d3} = t_{d2} - t_{d1} = \delta$ , then the last term in Eq. 7 is equal to zero, erasing the influence of initial position from motion encoding part of the signal. As a result, the formulation with displacement integrals,  $\int w_i(\tau) d\tau$ , is a more meaningful choice than the center of mass (COM) of random walk,  $\int x(\tau) d\tau$ , introduced in [10].

3) The last part of the sequence is where the data are collected:

$$\Omega_{\text{ro}}(t) = \int_{t_{\text{acq}}}^t G_{\text{ro}} \cdot (x_i(t_0) + w_i(\tau)) d\tau = (t - t_{\text{acq}}) G_{\text{ro}} \cdot x_i(t_0) + \int_{t_{\text{acq}}}^t G_{\text{ro}} \cdot w_i(\tau) d\tau, \quad (8)$$

$$\Omega_i(t, G_{\text{pe}}, G_{\text{ss}}, G_{\text{D}}) =$$

$$(t - t_{\text{acq}} - \Delta t_{\text{rw}}) G_{\text{ro}} \cdot x_i(t_0) - \Delta t_{\text{rw}} G_{\text{pe}} \cdot x_i(t_0) - \Delta t_{\text{rw}} G_{\text{ss}} \cdot x_i(t_0) + G_{\text{D}} \cdot W^{\text{d}} + G_{\text{ro}} \cdot W_i^{\text{acq}}(t) - (G_{\text{ro}} + G_{\text{pe}} + G_{\text{ss}}) \cdot W_i^{\text{rw}} \quad (9)$$

Equation 9 incorporates three integrals of the displacement  $w_i(t) : (W_i^{\text{d}}, W_i^{\text{acq}}, W_i^{\text{rw}})$  corresponding to the displacement integrals for diffusion (d), analog to digital conversion acquisition (acq) and initial rewind (rw) gradient time periods respectively

$$W_i^{\text{d}} = \int_{t_{\text{d3}}}^{t_{\text{d4}}} w_i(\tau) d\tau - \int_{t_{\text{d1}}}^{t_{\text{d2}}} w_i(\tau) d\tau, W_i^{\text{acq}}(t) = \int_{t_{\text{acq}}}^t w_i(\tau) d\tau, W_i^{\text{rw}} = \int_{t_0}^{t_{\text{rw}}} w_i(\tau) d\tau. \quad (10)$$

Moreover, in Eq. 9, the first two terms that multiply the initial position  $x_i(t_0)$  are the foundation of MR imaging and can be written in the  $k$ -space formalism as

$$k_{\text{mr}}(t, G_{\text{pe}}) = (t - t_{\text{acq}} - \Delta t_{\text{rw}}) G_{\text{ro}} - \Delta t_{\text{rw}} G_{\text{pe}} \quad (11)$$

where  $G_{\text{ro}}$  and  $\Delta t_{\text{rw}}$  are fixed, and  $t$  and  $G_{\text{pe}}$  define the sampling points in  $k$ -space. Also, the term  $\varphi_{\text{slice}} = -\Delta t_{\text{rw}} G_{\text{ss}} \cdot x_i(t_0)$  is constant for all  $i$  because the slice select axis component of  $x_i(t_0)$  is the slice position.

Without loss of generality, the coordinate frame used for the equations is the imaging coordinate frame defined by the read-out, phase encode and slice select gradients,  $G_{\text{ro}} = [g_{\text{ro1}}, 0, 0]$ ,  $G_{\text{pe}} = [0, g_{\text{pe2}}, 0]$ ,  $G_{\text{ss}} = [0, 0, g_{\text{ss3}}]$ . Using this notation, Eq. 11 implies that  $t = k_{\text{mr1}}/g_{\text{ro1}} + t_{\text{acq}} + \Delta t_{\text{rw}}$  and  $g_{\text{pe2}} = -k_{\text{mr2}}/\Delta t_{\text{rw}}$  where  $k_{\text{mr}}$  has its components equal to

$$[k_{\text{mr1}}, k_{\text{mr2}}, k_{\text{mr3}}] = [(t - t_{\text{acq}} - \Delta t_{\text{rw}}) g_{\text{ro1}}, -\Delta t_{\text{rw}} g_{\text{pe2}}, 0].$$

With this,  $W_i^{\text{acq}}(t)$  becomes a function of  $k_{\text{mr1}}$  and it is denoted by  $W_i^{\text{acq}}(k_{\text{mr1}})$ . The coefficients of  $W^{\text{rw}}$  in Eq. 9 are written as a vector:

$$k_{\text{rw}} = [-g_{\text{ro1}}, k_{\text{mr2}}/\Delta t_{\text{rw}}, -g_{\text{ss3}}], \quad (12)$$

which is an *affine* function of  $k_{\text{mr}}$ .

By defining  $k_{\text{D}} = G_{\text{D}}$ , Eq. 9 is rewritten concisely:

$$\Omega_i = k_{\text{mr}} \cdot x_i(t_0) + k_{\text{D}} \cdot W_i^{\text{d}} + k_{\text{rw}} \cdot W_i^{\text{rw}} + W_{i,1}^{\text{acq}}(k_{\text{mr1}}) g_{\text{ro1}} + \varphi_{\text{slice}}, \quad (13)$$

reflecting the partition of the effect of initial position and displacement integrals on the phase of each magnetic moment. Since  $\varphi_{\text{slice}}$  is constant for all  $i$ , it is taken out of Eq. 13 with the appropriate rotation of the magnetization coordinate frame on a slice by slice basis.

In Eq. 13, the pairs  $(k_{\text{mr}}, x_i(t_0))$ ,  $(k_{\text{D}}, W_i^{\text{d}})$  and  $(k_{\text{rw}}, W_i^{\text{rw}})$  are clearly Fourier dual variables while  $W_i^{\text{acq}}$  introduces dependence on magnetic moment motions. The derivations of the total signal expression in the next section clarifies this Fourier relationship.

### 3 Complete Fourier Direct MRI

The total CFD–MRI signal comes from all of the magnetic moments and it is the sum of individual magnetizations with the real valued initial magnetization  $m_i(t_0) = m_0$ . Equation 3 can be re–written using the relationship in Eq. 13 using the Fourier variables of Section 2:

$$S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}}) = m_0 \sum_i \exp(-j\gamma \Omega_i(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}})). \quad (14)$$

The dimension of the signal domain is  $(l_{\text{mr}} + 3 + 3)$  but, whereas  $(k_{\text{mr}}, k_{\text{D}}) \in \mathbb{R}^{l_{\text{mr}}+3}$  are independent variables,  $k_{\text{rw}} \in \mathbb{R}^3$  is affinely dependent on  $k_{\text{mr}}$  by Eq. 12. Accordingly, the CFD  $k$ –space variable is defined by joining the independent variables:  $k_{\text{cfd}} = (k_{\text{mr}}, k_{\text{D}}) \in \mathbb{R}^{l_{\text{mr}}+3}$ .

Since the properties of interest, e.g., tissue structure, are location dependent, instead of adding up the magnetization magnetic moment by magnetic moment, the sum in Eq. 14 can be carried out in *the joint position–displacement integral space*. For this purpose, the total CFD distribution function,  $P_{\text{cfd}}^{\text{total}}(x, W)$ , is defined here as the number of magnetic moments with the initial position  $x \in \mathbb{R}^3$  at time  $t_0$ , possessing the displacement integral values of  $W = (W^{\text{d}}, W^{\text{rw}}, W_1^{\text{acq}}) \in \mathbb{R}^7$ . By definition,  $P_{\text{cfd}}^{\text{total}}$  is real valued and positive semi–definite.

All of the individual magnetic moment signals coming from the infinitesimal volume element  $dx dW$  around the (position, displacement integral) point  $(x, W)$  must be integrated to form the DW–MRI signal. The sum in Eq. 14 becomes (absorbing  $m_0$  into  $P_{\text{cfd}}^{\text{total}}$  for ease of notation):

$$S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}}) = \int P_{\text{cfd}}^{\text{total}}(x, W) \times \exp(-j\gamma(k_{\text{mr}} \cdot x + k_{\text{D}} \cdot W^{\text{d}} + k_{\text{rw}} \cdot W^{\text{rw}} + W_1^{\text{acq}}(k_{\text{mr}1})g_{\text{ro}1})) dx dW \quad (15)$$

defining the Fourier transform of the total CFD distribution function  $P_{\text{cfd}}^{\text{total}}$  evaluated at the frequencies specified by  $(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}})$ :

$$S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}}) = \mathcal{F}\{P_{\text{cfd}}^{\text{total}}\}(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}}) \quad (16)$$

with nonlinearities added by  $W_1^{\text{acq}}$ . Accordingly, the only physical constraint that guides the signal reconstruction is the following.

Since  $P_{\text{cfd}}^{\text{total}}$  is real valued,  $S_{\text{cfd}}$  is Hermitian symmetric.

Experimentally,  $P_{\text{cfd}}^{\text{total}}$  cannot be obtained by the inversion of Eq. 16 *via* discrete Fourier transform (DFT) because  $k_{\text{rw}}$  cannot be independently sampled at necessary points (e.g., on a Cartesian grid) due to the affine dependence of  $k_{\text{rw}}$  on  $k_{\text{mr}}$  in Eq. 12. However,  $P_{\text{cfd}}^{\text{total}}$  is not the main point of interest for discovering the microstructural properties of the environment. Among the elements of  $W$ , the focus is on the most descriptive MRI observable,  $W^{\text{d}}$ . The objective is to marginalize  $W^{\text{rw}}$  from Eq. 15 to obtain the distribution

$$P_{\text{cfd}}(x, W^{\text{d}}) = \int P_{\text{cfd}}^{\text{total}}(x, W^{\text{d}}, W^{\text{rw}}) dW^{\text{rw}} = \mathcal{F}_{(k_{\text{mr}}, k_{\text{D}})}^{-1}\{S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, 0)\}. \quad (17)$$

without being able to sample at  $k_{\text{rw}} = 0$ .

With this issue at hand and the disruption of the Hermitian property of  $S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, k_{\text{rw}})$  (and therefore  $S_{\text{cfd}}(k_{\text{mr}}, k_{\text{D}}, 0)$ ) under experimental conditions, the estimation of  $P_{\text{cfd}}$  is achieved by applying systematic phase corrections during the computation of the transform in Eq. 17. By refraining from using the signal's magnitude both the energy and information content are simultaneously preserved also permitting asymmetry in  $P_{\text{cfd}}$ . The starting point the regular spin-echo image:  $I_{k_{\text{mr}}}^{\text{complex}}((x_{\text{ro}}, x_{\text{pe}}), 0, k_{\text{rw}}) = \mathcal{F}_{k_{\text{mr}}}^{-1}\{S_{\text{cfd}}(k_{\text{mr}}, 0, k_{\text{rw}})\}$ . The theoretically real valued image  $I_{k_{\text{mr}}}$  becomes complex valued,  $I_{k_{\text{mr}}}^{\text{complex}} = I_{k_{\text{mr}}} \exp(j \angle I_{k_{\text{mr}}}^{\text{complex}})$  due to the systematic echo time shifting from the middle of the read-out period that creates a linear phase shift in the physical read-out axis and other effects, such as motion, along the phase-encode axis. The phase shifts are approximated by polynomials in each imaging axis

$$r_{\text{ro}}(x_{\text{ro}}) \approx \angle I_{k_{\text{mr}}}^{\text{complex}}((x_{\text{ro}}, 0), 0), \quad r_{\text{pe}}(x_{\text{pe}}) \approx \angle I_{k_{\text{mr}}}^{\text{complex}}((0, x_{\text{pe}}), 0). \quad (18)$$

The corrections are then applied systematically at each value of  $k_{\text{D}}$  according to:

$$I_{k_{\text{mr}}}(x, k_{\text{D}}) = I_{k_{\text{mr}}}^{\text{complex}}(x, k_{\text{D}}) \exp(-j(r_{\text{ro}}(x_{\text{ro}}) + r_{\text{pe}}(x_{\text{pe}}))). \quad (19)$$

Afterwards, the Fourier transform in  $k_{\text{D}}$ -space,

$$P_{\text{cfd}}(x, W^{\text{d}}) = \mathcal{F}_{k_{\text{D}}}^{-1}\{I_{k_{\text{mr}}}(x, k_{\text{D}})\}, \quad (20)$$

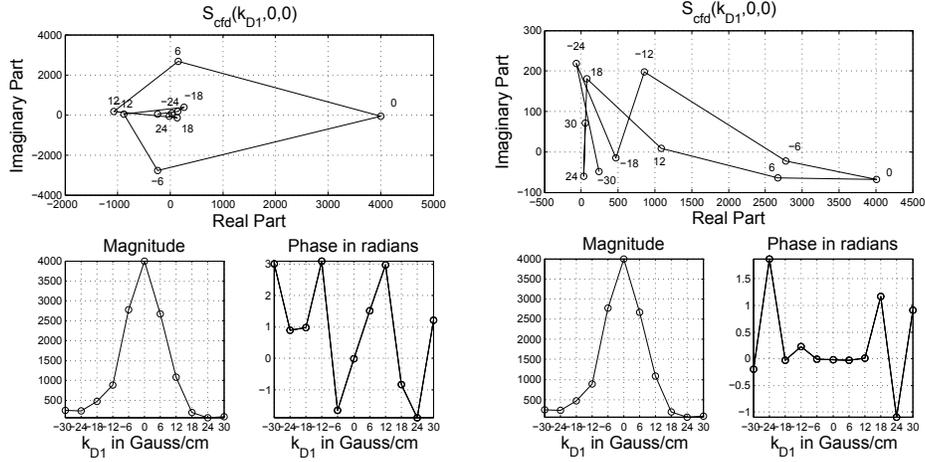
is evaluated sequentially in each  $k_{\text{D}}$ -dimension. At each step before executing the transform, the Hermitian property in  $k_{\text{D}}$ -space is re-established,  $I_{k_{\text{mr}}}(x, -k_{\text{D}}) = I_{k_{\text{mr}}}^*(x, k_{\text{D}})$  as demonstrated in Fig 2.

## 4 Experimental Setup and Analysis Methods

A fixed baboon brain immersed in 4% paraformaldehyde was used for the experiments. The primate was prematurely delivered on the 125<sup>th</sup> day and sacrificed on the 59<sup>th</sup> day after delivery. All animal husbandry, handling, and procedures were performed at the Southwest Foundation for Biomedical Research, San Antonio, Texas. Animal handling and ethics were approved to conform to American Association for Accreditation of Laboratory Animal Care (AAALAC) guidelines. Further details of the preparation are explained in [11].

The experiments were carried out on a 4.7 Tesla MR scanner (Varian NMR Systems, Palo Alto, CA, USA) with a 15 cm inner diameter gradient system, 45 Gauss/cm maximum gradient strength and 0.2 ms rise time using a cylindrical quadrature birdcage coil (Varian NMR Systems, Palo Alto, CA, USA) with 63 mm inner diameter.

CFD-MRI data were obtained using the standard pulsed-gradient spin-echo multi-slice sequence. The  $k_{\text{mr}}$ -space was sampled to result in images of  $128 \times 128$  pixels with a field of view  $64 \times 64 \text{ mm}^2$  and 0.5 mm slice thickness. The  $k_{\text{D}}$ -space was sampled in a uniformly spaced Cartesian grid in a cube



**Fig. 2.** On the top row the Nyquist plots of uncorrected (left) and corrected (right)  $S_{\text{cfd}}$  obtained from the experimental data described in Section 4. The plots show data acquired at each diffusion gradient value  $k_{D1}$  on the complex plane. On the bottom row are the magnitude and phase plots of the data. Uncorrected data (bottom row left, second column) exhibit a linear phase shift around 0 frequency, indicative of coherent motion. After the phase corrections obtained using the polynomial  $0.266 k_{D1}$  estimated from the points  $k_{D1} = -6, 0, 6, 12$  Gauss/cm, the magnitude is unchanged but signal's imaginary part is smaller for the corrected values visible in the vertical axis spans of Nyquist plots and the phase plots.

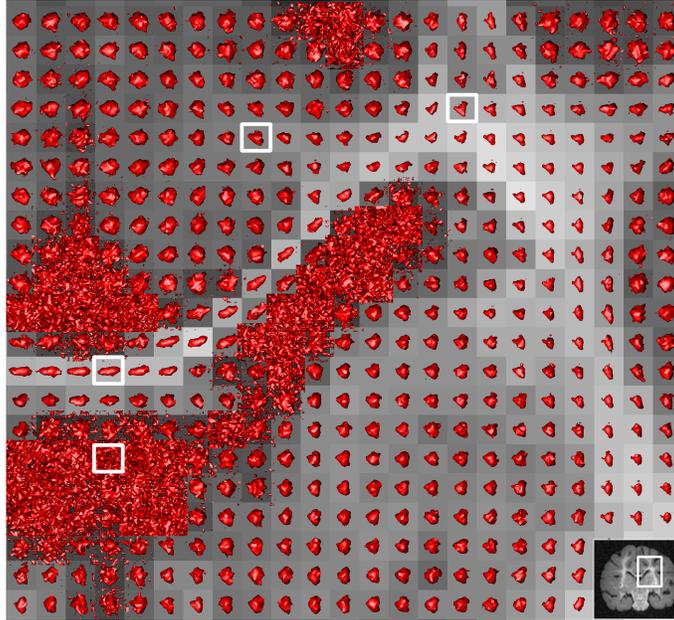
$[-30 \text{ Gauss/cm}, 30 \text{ Gauss/cm}]^3$  with  $11 \times 11 \times 11$  voxels, i.e., 6 Gauss/cm sampling intervals at each dimension. In regular MRI terms, 1331 images per slice were obtained. The repetition time  $T_R = 1$  s, echo time  $T_E = 56.5$  ms, diffusion pulse time offset  $\Delta = 30$  ms and diffusion pulse duration  $\delta = 15$  ms were used.

After the acquisition, the data were transferred to a two quad core 2.3GHz Intel Xeon® cpu and 8GB memory Dell Precision Workstation 490 running Windows XP® 64-bit operating system. The DWI data were placed in a 5-dimensional array in the computer memory and the discrete Fourier transform was computed along with the phase corrections as recounted in Section 3. Afterwards, isosurfaces were computed as explained in Section 5. In-house Matlab® (Mathworks, Natick, MA USA) programs were used for all of the computations and to display the graphics and maps.

## 5 Visualization of the CFD Distribution

In the case of a two dimensional MR slice ( $l_{\text{mr}} = 2$ ), the domain of  $P_{\text{cfd}}$  is the five dimensional real space  $P_{\text{cfd}} : \mathbb{R}^{l_{\text{mr}}} \times \mathbb{R}^3 \rightarrow \mathbb{R}$ , which makes the visualization challenging. The task is realized by using two components: the presentation of the number of magnetic moments at the initial position and the three

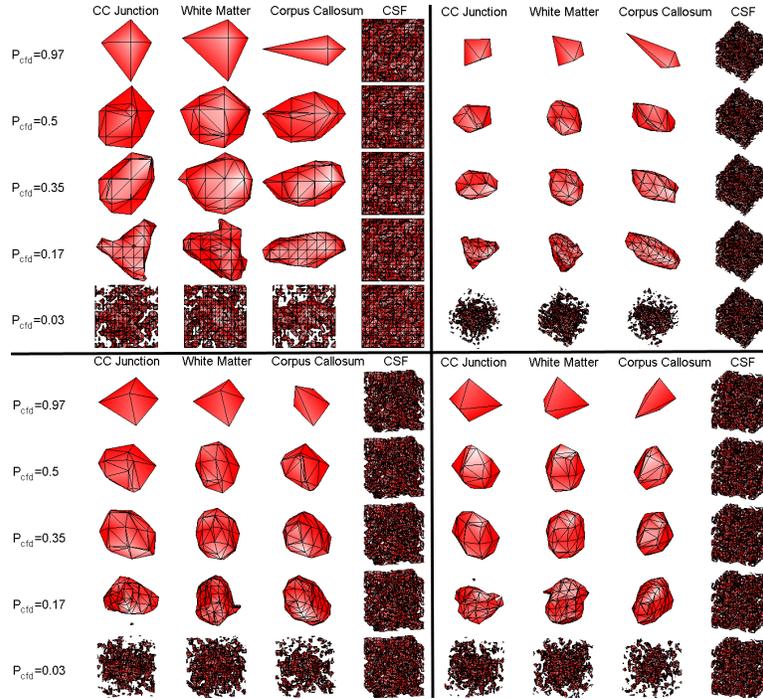
dimensional distribution of the displacement integral values. The former is accomplished by using  $P_{\text{cfd}}(x, 0)$  as the background image, e.g., in Fig. 3, because in a restricted environment without coherent motion, the function  $P_{\text{cfd}}(x, W^{\text{d}})$  assumes its maximum at  $W^{\text{d}} = 0$  since the majority of the magnetic moments do not travel too far from their initial position. If there is coherent motion, linear portion of phase corrections rectify it anyway.



**Fig. 3.** The isosurfaces ( $\bar{P}_{\text{cfd}} = 0.17$ ) picked from the area around the corpus callosum (CC) and external capsule (EC) junction. The area of interest is indicated on the right bottom corner. Starting from left bottom going clockwise, the sample pixels of Fig. 4 are from cerebrospinal fluid (CSF), CC, White Matter (WM) and CC and EC junction respectively. The isosurfaces for different  $c$ -values are shown in Fig. 4.

In order to present the displacement integral component with the sole purpose of obtaining a single scale across the image,  $P_{\text{cfd}}$  is normalized at each location  $x$ :  $\bar{P}_{\text{cfd}}(x, W^{\text{d}}) = P_{\text{cfd}}(x, W^{\text{d}})/P_{\text{cfd}}(x, 0)$ . The second part of visualization is achieved by overlaying on  $P_{\text{cfd}}(x, 0)$  the isosurfaces of  $\bar{P}_{\text{cfd}}$  with a level value  $c$  ( $0 < c \leq 1$ ), i.e.  $\{W^{\text{d}} \in \mathbb{R}^3 : \bar{P}_{\text{cfd}}(x, W^{\text{d}}) = c\}$ . For the sake of an objective assessment, a common  $c$ -value is adopted throughout the image after the normalization creating a different isosurface at each location. For biological tissue, relevant information is provided by the values of  $P_{\text{cfd}}$  away from  $c = 0$ , i.e., the outskirts of  $P_{\text{cfd}}$  corresponding to the small number of magnetic moments that travel further away, portraying the microstructure. The effect of level value choice is summarized in Fig. 4 displaying isosurfaces from four different pixels within the baboon brain marked in Fig. 3:

- 1) Too high values do not provide enough structural information (see the first rows in Fig. 4).
- 2) The appropriately informative value depends on the properties of the motion (thus of the microstructure) at a given location (compare the columns of Fig. 4).
- 3) Too low values force the isosurfaces to become extremely noisy (see last rows of Fig. 4).



**Fig. 4.** Using multiple  $c$ -values, the isosurfaces from the area around the CC and EC junction marked in Fig. 3 are shown with different view angles in  $\mathbf{R}^3$  at each subfigure in the quadrants. Less and less information is available with an increasing  $c$ -value. In contrast, if the value is chosen to be too low, all of the isosurfaces are immersed in noise. Regardless of the  $c$ -value, the CSF isosurface is noisy because the diffusion pulse times, which are appropriate for diffusion in tissue, are too long for the displacement in the fluid.

In Figs. 3 and 4, highly structured areas, e.g., the corpus callosum (CC), demonstrate ellipsoidal shapes. In contrast, at the junctions, the isosurfaces indicate crossing fibers (see the ‘T’ shaped isosurface in Fig. 4) at the CC–external capsule (EC) junction. The organization in white matter presents shapes in between. These properties are highly informative for tractographic studies.

Overall, the isosurfaces are not constrained to given forms like Gaussians, spherical harmonics or to any expansions. In fact, they are typically not even symmetric. They are general and direct.

## 6 Conclusion and Future Studies

CFD–MRI brings clarifications to the nature of the DW–MR signal as well as conceptual simplifications over existing methods [12]. The narrow pulse approximation (negligible duration of diffusion sensitizing magnetic field gradient pulse) is not necessary because CFD–MRI uses the displacement integrals, which incorporate directly pulse durations, rather than the displacements themselves. In other words,  $P_{\text{cfd}}$  does not involve the propagator which is the basis of spectral methods [13]. Furthermore, this also removes the necessity to model the motion as a stochastic process relieving in conjunction the constraints and derivations, such as being Markovian, in the existing methods.

The high dimensional Fourier relationship with the support of Hermitian property preservation (to avoid the usage of signal magnitude) results in completely free distribution functions that display asymmetry as demonstrated in Figs. 3 and 4 in an evidence based manner. Accordingly, re-establishment of Hermitian symmetry requires future development of more sophisticated systematic phase corrections starting from the ones introduced herein. From a practical point of view, this mathematically challenging task also requires a deeper investigation and better understanding of the origin of DW–MRI artifacts, specifically the ones that disturb the Hermitian property.

The clinical aim of future studies is the establishment of connections between biological properties and the CFD distribution function,  $P_{\text{cfd}}$ . The challenge is, therefore, the classification of different microstructures, healthy and pathological, despite lack of analytical descriptions of  $P_{\text{cfd}}$  for even simple geometries. A biological phantom, with known geometry, constructed with two nerve fiber bundles and agar gel has been recently investigated as an initial step towards that goal [14].

The analysis and visualization of  $P_{\text{cfd}}$  need to be designed to define and enhance CFD–MRI surrogate markers for diagnosis and prognosis of disease, in the mentality of the accomplishments with DTI markers in [3]. Naturally, these markers will provide much deeper information than the existing DW–MRI methods. However, the effect of integration on the displacement has to be carefully studied, for example, in order to make axon caliber assessments for white matter integrity estimation. Likewise, the observation that in DW–MRI the displacement is from the initial magnetic moment position indicates that fiber tractography methods cannot exactly determine the location of tracts within a pixel. It will be necessary to introduce new tools, e.g., from differential geometry, for better characterization, categorization and visualization of three dimensional motion portion of  $P_{\text{cfd}}$ .

### Acknowledgement

This study was supported, in part, by the Washington University Small Animal Imaging Resource (U24-CA83060), the NIH/NINDS grant Biomarkers and Pathogenesis of MS (P01-NS059560), and the US Army grant, NeuroPerformance Imaging (W23RYX1089N603). Special thanks to J. Quirk, T. Conturo,

S-K. Song, K. Uğurbil, G. Sapiro, K. Wong, S. K. Mun and Z-H. Cho for valuable discussions.

## References

1. Song, S.K., Qu, Z., Garabedian, E.M., Gordon, J.I., Milbrandt, J., Ackerman, J.J.H.: Improved magnetic resonance imaging detection of prostate cancer in a transgenic mouse model. *Cancer Research* **62**(5) (March 2002) 1555–1558
2. Budde, M.D., Kim, J.H., Liang, H.F., Schmidt, R.E., Russell, J.H., Cross, A.H., Song, S.K.: Toward accurate diagnosis of white matter pathology using diffusion tensor imaging. *Magnetic Resonance in Medicine* **57**(4) (April 2007) 688–695
3. Song, S.K., Sun, S.W., Ramsbottom, M.J., Chang, C., Russell, J., Cross, A.H.: Demyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *NeuroImage* **17**(3) (November 2002) 1429–1436
4. Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C., Thompson, A.: Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. *The Lancet Neurology* **7**(8) (2008) 715–727
5. McKinstry, R.C., Mathur, A., Miller, J.H., Özcan, A., Snyder, A.Z., Schefft, G.L., Almlı, C.R., Shiran, S.I., Conturo, T.E., Neil, J.J.: Radial organization of developing preterm human cerebral cortex revealed by non-invasive water diffusion anisotropy MRI. *Cerebral Cortex* **12** (December 2002) 1237–1243
6. Özcan, A., Wong, K.H., Larson-Prior, L., Cho, Z.H., Mun, S.K.: Background and mathematical analysis of diffusion mri methods. *International Journal of Imaging Systems and Technology* **22**(1) (2012) 44–52
7. McCall, D.W., Douglass, D.C., Anderson, E.W.: Self-diffusion studies by means of nuclear magnetic resonance spin-echo techniques. *Berichte der Bunsengesellschaft für Physikalische Chemie* **67**(3) (1963) 336–340
8. Stejskal, E.O., Tanner, J.: Spin diffusion measurements: Spin echoes in the presence of a time-dependent field. *The Journal of Chemical Physics* **42**(1) (January 1965) 288–292
9. Özcan, A.: A new model for diffusion weighted MRI: Complete fourier direct MRI. In: Proceedings of the 32st Annual International Conference of the IEEE EMB Society, Buenos Aires, Argentina (September 2010) 2710–2713 PAP - EMBC10.1722.
10. Mitra, P.P., Halperin, B.I.: Effects of finite gradient-pulse widths in pulsed-field-gradient diffusion measurements. *Journal of Magnetic Resonance. Series A* **113**(1) (March 1995) 94–101
11. Kroenke, C.D., Bretthorst, G.L., Inder, T.E., Neil, J.J.: Diffusion MR imaging characteristics of the developing primate brain. *NeuroImage* **25**(4) (May 2005) 1205–1213
12. Özcan, A.: Comparison of the complete fourier direct MRI with existing diffusion weighted MRI methods. In: Proceedings of the 2011 IEEE International Symposium on Biomedical Imaging, Chicago, Illinois, USA (April 2011) 931–934
13. Callaghan, P.T.: Principles of Nuclear Magnetic Resonance Microscopy. Oxford University Press (1991)
14. Özcan, A., Quirk, J., Wang, Y., Wang, Q., Sun, P., Spees, W., Song, S.K.: The validation of complete fourier direct mr method for diffusion mri via biological and numerical phantoms. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. (30 2011-sept. 3 2011) 3756–3759

## Poster Session

# Multi-class DTI Segmentation: A Convex Approach

Yuchen Xie<sup>1</sup>      Ting Chen<sup>2</sup>      Jeffrey Ho<sup>1</sup>      Baba C. Vemuri<sup>1\*</sup>

<sup>1</sup>Department of CISE, University of Florida, Gainesville, FL

<sup>2</sup>IBM Almaden Research Center, San Jose, CA

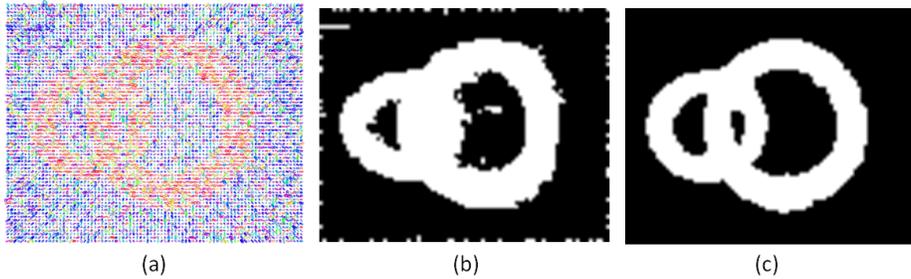
**Abstract.** In this paper, we propose a novel variational framework for multi-class DTI segmentation based on global convex optimization. The existing variational approaches to the DTI segmentation problem have mainly used gradient-descent type optimization techniques which are slow in convergence and sensitive to the initialization. This paper on the other hand provides a new perspective on the often difficult optimization problem in DTI segmentation by providing a reasonably tight convex approximation (relaxation) of the original problem, and the relaxed convex problem can then be efficiently solved using various methods such as primal-dual type algorithms. To the best of our knowledge, such a DTI segmentation technique has never been reported in literature. We also show that a variety of tensor metrics (similarity measures) can be easily incorporated in the proposed framework. Experimental results on both synthetic and real diffusion tensor images clearly demonstrate the advantages of our method in terms of segmentation accuracy and robustness. In particular, when compared with existing state-of-the-art methods, our results demonstrate convincingly the importance as well as the benefit of using more refined and elaborated optimization method in diffusion tensor MR image segmentation.

## 1 Introduction

Diffusion tensor magnetic resonance imaging (DT-MRI) introduced by Basser et al. [1] provides a powerful modality for measuring the diffusion of water molecules in tissues by modeling the local diffusivity in each voxel as a  $3 \times 3$  symmetric positive-definite (SPD) matrix. Because diffusion tensor images (DTI) are able to reveal in-vivo the anatomical tissue structures such as white matter fibers in the brain and muscle fibers in the heart, numerous methods and algorithms have been proposed in the literature for DTI images. In particular, DTI segmentation often plays an indispensable and crucial role in diffusion tensor image analysis, hence has attracted immense attentions in recent years. Current DTI segmentation algorithms can be broadly classified into three categories: level-set-based methods ([14, 15, 9, 13]), combinatorial optimization based methods such as graph cuts ([16]), and clustering-based methods ([18, 8]). Wang et al. [14] presented the first variational method to segment tensor fields using region-based active contour model with Euclidean distance. In [15], the square root of the symmetrized Kullback-Leibler (KL) divergence which measures the difference between tensors was incorporated into the active contour framework. Although the symmetrized

---

\* This research was in part supported by the NIH grant NS066340 to BCV.



**Fig. 1.** Two-class DTI segmentation. (a) Synthetic DTI image with noise. (b) The binary segmentation result using the method proposed in [15]. (c) The DTI segmentation result using the propose method. The improvement in accuracy is clearly visible.

KL divergence used in [15] is not a true distance, using it as a similarity measure often produces reasonable segmentation results. Lenglet et al. proposed a unified surface evolution framework to segment diffusion tensor images using the geodesic distance on the manifold of symmetric positive-definite matrices in [9] and obtained more accurate segmentations of the corpus callosum in diffusion tensor brain images. In [13], a robust measure, the total KL divergence, was introduced and applied in DTI segmentation problem, which leads to good segmentation results even in the presence of large amounts of noise and outliers. Since these variational models are solved using gradient-descent type methods with non-convex cost functions, the usual thorny issue of local minima is difficult to avoid or resolve satisfactorily. In order for those level-set-based methods to work well, good initialization and robust numerical algorithms are needed to prevent the techniques from being stuck in local minima, which undoubtedly diminishes the applicability of these DTI segmentation algorithms. On the other hand, a graph cuts based method was introduced for global DTI segmentation ([16]). While the combinatorial approach is able to produce globally optimal solutions, the method can only deal with binary segmentation and require interactive user inputs. Furthermore, grid bias can be significant when the images are under-sampled. Meanwhile, clustering-based methods [18, 8] are able to automatically solve the multi-class segmentation problem; however, without the explicit optimization of an energy function, they often cannot produce the same level of accuracy in segmentation as the energy minimization methods.

In recent years, convex relaxation technique has been utilized to solve several well-known variational models in image processing. The seminal work by Chan et. al. [5] showed that by using the characteristic functions to represent regions, the non-convex two-phase Chan-Vese problem [6] can be formulated as a convex problem in a space that is one dimension higher. The optimal solution of the original problem can be determined from the optimal solution of the new optimization problem, which can be efficiently computed due to the convexification. For the multi-class segmentation problem, the globally optimal solution cannot be obtained in polynomial time. However, Pock et al. [11, 3] proposed a method to relax the non-convex problem to its local convex envelop, and they show experimentally that the solution of the relaxed problem is usually very

close to the true global minimizer of the original problem. These recently-published optimization methods using convex relaxation offer several significant advantages over most of the aforementioned earlier methods in that they are free of grid bias and can be efficiently solved and easily accelerated on parallel architectures [4].

In this paper, we propose a variational framework for multi-class DTI segmentation. Different from [14, 15, 9, 13], our method is able to simultaneously compute robust multi-label partitions for a diffusion tensor image instead of the binary partition by using the advanced convex relaxation techniques. The minimizer of the relaxed problem can be efficiently computed using well-established convex optimization methods such as the first-order primal-dual algorithms [4], and the solution of the original problem can be obtained by a simple thresholding procedure. In particular, no contour initializations are required for computing image partitions in our method, and it avoids the thorny issue of local minima that has plagued most existing DTI segmentation methods using variational approaches.

Both synthetic and real tensor images are used to validate the proposed method. Thanks to the convex relaxation, our method is robust to noisy tensor fields, and Figure 1 convincingly displays one example. For the synthetic tensor image corrupted with significant noise, the region-based active contour method proposed in [15] stops at an unsatisfactory solution that is clearly a local minimum, and in comparison, the proposed method correctly segments out the double-ring shape. While both methods use the symmetrized KL divergence as their similarity measure between tensors and their respective energy functions are also similar (at least, in principle), the difference is in the optimization that in our case, is able to utilize the convex relaxation that is known to be reasonably tight [4] and produces solutions that are qualitatively superior as demonstrated in the Figure 1.

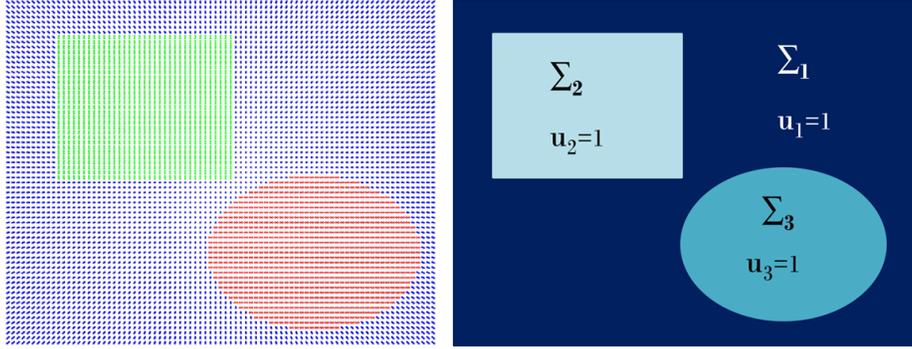
The rest of this paper is organized as follows. In section 2, we present the proposed framework for multi-phase DTI segmentation. In section 3, we describe a primal-dual algorithm to solve the partition problem. Experimental results for synthetic tensor images and real DTI data are presented in section 4, and we conclude the paper with a short summary in section 5.

## 2 Multi-class DTI Segmentation Method

Let  $\mathbf{T}$  denote a diffusion tensor image defined on the domain  $\Omega \subset \mathbb{R}^m$  ( $m = 2$  for 2D image and  $m = 3$  for 3D image), and we consider  $\mathbf{T}$  as a function  $\mathbf{T} : \Omega \rightarrow \mathbf{P}(m)$ , where  $\mathbf{P}(m)$  is the space of symmetric positive-definite (SPD) matrices. The multi-class DTI segmentation can be formulated as a partition problem using the piece-wise constant variational model

$$\min_{\{\Sigma_i\}, \{\mathbf{T}_i\}} \sum_{i=1}^k \int_{\Sigma_i} d^2(\mathbf{T}(x), \mathbf{T}_i) dx + \frac{\lambda}{2} \sum_{i=1}^k \mathbf{Per}(\Sigma_i; \Omega), \quad (1)$$

where  $\{\Sigma_i\}_{i=1}^k$  is a partition of  $\Omega$  with  $k$  regions (labels) such that  $\bigcup_{i=1}^k \Sigma_i = \Omega$  and  $\Sigma_i \cap \Sigma_j = \emptyset, \forall i \neq j$ . The first term is the data term with  $d$  denoting the similarity measure for the tensors  $\mathbf{T}(x)$  and the mean tensor  $\mathbf{T}_i$  in the region  $\Sigma_i$ .  $\mathbf{Per}(\Sigma_i; \Omega)$



**Fig. 2. Left:** Synthetic three-class tensor image. **Right:** The corresponding label map with the region partition.

denotes the perimeter of the set  $\Sigma_i$ .  $\lambda$  is a nonnegative parameter. Since the common surfaces of the partition are counted twice in the sum of perimeters, the second term represents the total interface energy (area)  $\bigcup_{i < j} \partial \Sigma_i \cap \partial \Sigma_j$ . A three-class example is shown in Figure 2.

Several tensor distance measure can be incorporated in the DTI segmentation framework discussed above, and this paper considers the following three choices. The simplest is the Euclidean distance between SPD matrices based on the Frobenius norm. For DTI applications, the geodesic distance is the more appropriate choice of metric [7, 10, 9]. The geodesic distance on  $\mathbf{P}(m)$  is given by the following formula

$$d_G(\mathbf{T}_1, \mathbf{T}_2) = \sqrt{\text{tr}(\text{Log}^2(\mathbf{T}_1^{-1/2} \mathbf{T}_2 \mathbf{T}_1^{-1/2}))}, \quad (2)$$

where  $\text{tr}$  denotes the trace of a matrix and  $\text{Log}$  denotes the matrix logarithm. Because the diffusion of the water molecules is modeled by a Gaussian distribution in DT-MRI, it is natural to consider information theoretical “distances” between Gaussian distributions to measure the similarity between diffusion tensors. [15] introduced the symmetrized KL divergence as a new affine-invariant discriminant of diffusion tensors

$$d_{sKL}^2(\mathbf{T}_1, \mathbf{T}_2) = sKL(\mathbf{T}_1, \mathbf{T}_2) = \frac{1}{4}(\text{tr}(\mathbf{T}_1^{-1} \mathbf{T}_2 + \mathbf{T}_2^{-1} \mathbf{T}_1) - 2n) \quad (3)$$

where  $\mathbf{T}_1$  and  $\mathbf{T}_2$  are  $n \times n$  SPD matrices. [13] proposed a novel measure, the total KL divergence,

$$d_{tKL}^2(\mathbf{T}_1, \mathbf{T}_2) = tKL(\mathbf{T}_1, \mathbf{T}_2) = \frac{\log(\det(\mathbf{T}_1^{-1} \mathbf{T}_2)) + \text{tr}(\mathbf{T}_2^{-1} \mathbf{T}_1) - n}{2\sqrt{c + \frac{(\log(\det \mathbf{T}_2))^2}{4} - \frac{n(1+\log 2\pi)}{2} \log(\det \mathbf{T}_2)}} \quad (4)$$

where  $c = 3n/4 + n^2 \log(2\pi)/2 + (n \log(2\pi))^2/4$ , and in the cited paper, experimental results have shown that the total KL divergence is intrinsically robust to noise and outliers.

The computation of the mean tensor  $\mathbf{T}_i^*$  over the region  $\Sigma_i$  is essential in the piece-wise constant region-based framework. It follows the definition of the Karcher mean

$$\mathbf{T}_i^* = \operatorname{argmin}_{\mathbf{T}_i \in \mathbf{P}(m)} \int_{\Sigma_i} d^2(\mathbf{T}(x), \mathbf{T}_i) dx. \quad (5)$$

It is known that the tensorial mean for both symmetrized KL and total KL divergences can be computed in closed-form [15, 13]. However, there is no closed-form formula for computing mean on  $\mathbf{P}(m)$  using the Riemannian metric. However because  $\mathbf{P}(m)$  has non-positive sectional curvature, the intrinsic mean is unique and can be efficiently computed by a gradient-descent type algorithm as described in [7].

Given the appearance model  $\mathbf{T}_i^*$ , we can use the characteristic functions  $u_i = \chi_{\Sigma_i}$  ( $u_i(x) \in \{0, 1\}$ ,  $i = 1, \dots, k$ ) to rewrite the energy functional for DTI segmentation:

$$E(\mathbf{u}) = \sum_{i=1}^k \int_{\Omega} u_i(x) d^2(\mathbf{T}(x), \mathbf{T}_i^*) dx + \frac{\lambda}{2} \sum_{i=1}^k \int_{\Omega} |Du_i|, \quad (6)$$

where  $D$  denotes the distributional derivative,  $\mathbf{u} = (u_1, \dots, u_k)$ , with the point-wise constraint  $\sum_{i=1}^k u_i(x) = 1$ . Although the cost functional  $E$  is convex with respect to  $\mathbf{u}$ , the set of the binary functions  $u_i$  is non-convex. Thus a convex relaxation is needed for the optimal image partition. One straightforward way is to relax the range of functions  $u_i$  from  $\{0, 1\}$  to  $[0, 1]$  as in [17]. Using the paired calibration method proposed in [11], a reasonably tight convex relaxation of the cost function  $E(\mathbf{u})$  can be constructed [4] for the image partition problem. The final relaxed convex optimization problem is given as

$$\min_{\mathbf{u} \in \mathbf{S}} \sum_{i=1}^k \int_{\Omega} u_i(x) d^2(\mathbf{T}(x), \mathbf{T}_i^*) dx + \lambda \int_{\Omega} \sup_{\mathbf{v} \in \mathbf{K}} (\mathbf{v} \cdot D\mathbf{u}) \quad (7)$$

where the convex sets  $\mathbf{S}$  and  $\mathbf{K}$  are defined as

$$\mathbf{S} = \left\{ \mathbf{u} = (u_1, \dots, u_k) \mid u_i(x) \in [0, 1], \sum_{i=1}^k u_i(x) = 1 \right\} \quad (8)$$

$$\mathbf{K} = \{ \mathbf{v} = (v_1, \dots, v_k) : \Omega \rightarrow \mathbb{R}^{m \times k} \mid \|v_i - v_j\|_2 \leq 1, 1 \leq i < j \leq k \}. \quad (9)$$

We note that the second term in Equation (7) is a convex relaxation of the total variation term in the original cost function in Equation (6). Given the minimizer  $\mathbf{u}$  of the relaxed problem, we can use the simple max-rounding scheme

$$u_i^*(x) = \begin{cases} 1 & \text{if } u_i(x) = \max(u_1(x), \dots, u_k(x)), \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

to obtain  $\mathbf{u}^*$  for the original non-convex problem.

### 3 Numerical Algorithm

In this section, we briefly describe the numerical algorithm for solving the partition problem (7). For more complete details including convergence proofs, we refer the reader to [4]. First, we discretize the image domain  $\Omega$  to  $\Omega^h$ , where  $h$  denotes the discrete spatial width. Let  $a^h$  indicate the size of base element in the discrete setting,  $a^h = h^2$  for 2D images and  $a^h = h^3$  for 3D images. We use  $\mathbf{u}^h$  and  $\mathbf{v}^h$  to represent the discretized primal and dual variables, respectively. Let  $\phi_i(\mathbf{T}(x)) = d^2(\mathbf{T}(x), \mathbf{T}_i^*)$  and  $\phi^h = (\phi_1^h, \dots, \phi_k^h)$ . Thus the problem (7) can be rewritten in discrete setting as

$$\min_{\mathbf{u}^h \in \mathbf{S}} \max_{\mathbf{v}^h \in \mathbf{K}} a^h \sum_{i=1}^k \sum_{x^h \in \Omega^h} u_i^h \phi_i^h + \lambda a^h \sum_{i=1}^k \sum_{x^h \in \Omega^h} v_i^h \cdot \nabla^h u_i^h \quad (11)$$

where  $\nabla^h$  is the discrete gradient operator. Following the work by Chambolle and Pock [4], we use a first-order primal-dual algorithm to solve the minimax problem in (11). Let  $\tau_p, \tau_d > 0$  represent the primal and dual time steps. We initialize  $(\bar{\mathbf{u}}^h)^0 = (\mathbf{u}^h)^0 = 0$  and  $(\mathbf{v}^h)^0 = 0$ . Then we use the following update scheme for  $n > 0$

$$\begin{cases} (\mathbf{v}^h)^{n+1} = \Pi_{\mathbf{K}}((\mathbf{v}^h)^n + \tau_d(\nabla^h(\bar{\mathbf{u}}^h)^n)), \\ (\mathbf{u}^h)^{n+1} = \Pi_{\mathbf{S}}((\mathbf{u}^h)^n + \tau_p(\operatorname{div}^h(\mathbf{v}^h)^{n+1} - \phi^h)), \\ (\bar{\mathbf{u}}^h)^{n+1} = 2(\mathbf{u}^h)^{n+1} - (\mathbf{u}^h)^n, \end{cases} \quad (12)$$

where  $\Pi_{\mathbf{K}}$  and  $\Pi_{\mathbf{S}}$  are projection operators. The projection onto the simplex  $\mathbf{S}$  can be performed by using thresholding operators in linear time. The projection  $\Pi_{\mathbf{K}}$  is much harder to compute. We first describe  $\mathbf{K}$  as a finite intersection of simple convex sets and then use the iterative algorithm proposed in [2] to compute  $\Pi_{\mathbf{K}}$ .

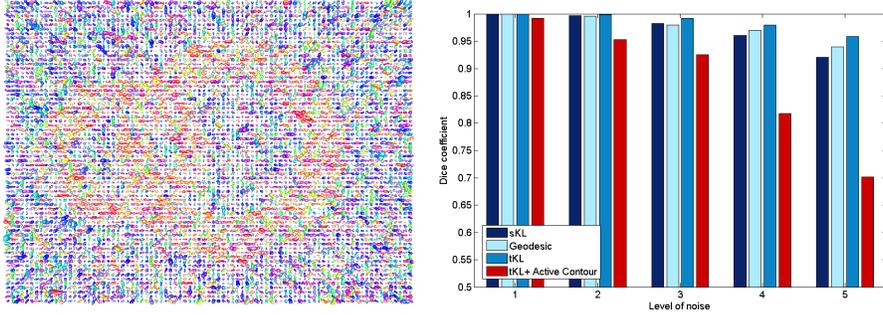
For the DTI segmentation problem, we apply a modified k-means clustering to obtain the intrinsic means of the SPD matrices as the initialization of  $\mathbf{T}_i$ . Then we perform the optimal tensor field partition and update the appearance model  $\mathbf{T}_i^*$  by solving the problem (5) in each iteration.

### 4 Experiments

In this section, we evaluate the propose DTI segmentation framework using both synthetic and real datasets. We show the comparisons between the state-of-the-art variational DTI segmentation method [13] and our segmentation framework using three different tensor distance measures: geodesic distance, symmetrized KL divergence and total KL divergence.

#### 4.1 Synthetic data

We first evaluate the proposed DTI segmentation methods with three different metrics on the synthetic tensor fields. Because previous variational DTI segmentation methods [14, 15, 9, 13] can only deal with binary cases and the robust total KL divergence has

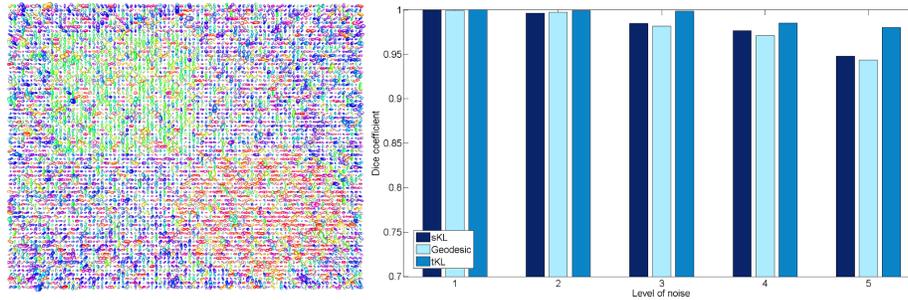


**Fig. 3.** Experimental results on synthetic two-class tensor fields. **Left:** Synthetic tensor image with Rician noise. **Right:** Dice score comparison for segmentation results using total KL active contour [13] and our methods with symmetrized KL, geodesic distance and total KL.

been proved to be the best measure for tensor comparison among those three metrics, we only list the method proposed in [13] as a representative of the variational DTI segmentation models in the binary DTI segmentation comparison. The tensor images were generated using the simulation model in [12] with 81 gradient directions and diffusion weighting  $b = 1500s/mm^2$ . For instance, a synthetic  $69 \times 86$  three-label tensor image is shown in Figure 2. We add different levels of Rician noises (from 10% to 50%) on the clean tensor field and test the robustness of different segmentation methods. The synthetic two-label and three-label tensor images with noise level 5 are shown in Figure 3 and Figure 4, respectively. Regardless the distance measure used, our method consistently provides more accurate segmentation results when compared with the method in [13], especially for tensor images corrupted with significant noise. For the multi-class DTI segmentation problem, the comparison of the DTI segmentation dice coefficients is among our methods with different tensor distance measures as displayed in Figure 4. We can readily observe that all of our methods perform well on noisy tensor fields and the method using robust total KL divergence outperforms others as the noise level increases.

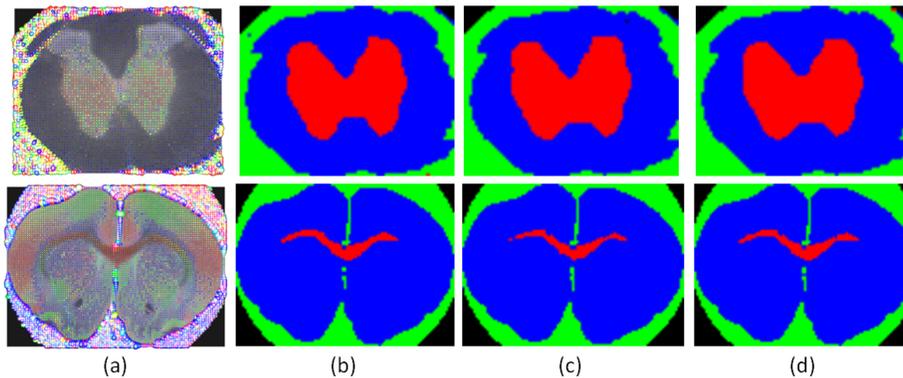
## 4.2 Real diffusion tensor images

In this experiment, we test the proposed DTI segmentation framework on two real diffusion tensor images of the spinal cord and corpus callosum of a rat, as shown in Figure 5. The data were acquired using a PGSE with  $TR=1.5s$ ,  $TE=28.3ms$ ,  $bandwidth=35KHz$ . Twenty-one diffusion weighted images with a  $b$ -value of  $1250s/mm^2$  were collected. The sizes of the regions of interest for rat spinal cord and corpus callosum are  $71 \times 61$  and  $101 \times 74$ , respectively. Because the diffusion tensors at background and liquid also need to be distinguished, we use four colored labels in the experiments: black (background), green (liquid), red (spinal cord or corpus callosum) and blue (other rat neural tissues). Figure 5 shows the DTI segmentation results using the proposed framework



**Fig. 4.** Experimental results on synthetic three-class tensor fields. **Left:** Synthetic tensor image with Rician noise. **Right:** Dice score comparison for segmentation results using symmetrized KL, geodesic distance and total KL.

with different tensor distance measures. The results confirm that the DTI segmentation methods based on convex relaxation do provide robust and accurate segmentation results on real diffusion tensor images.



**Fig. 5.** Experimental results on real diffusion tensor images. **Top row:** Spinal cord of a rat. **Bottom row:** Corpus callosum of a rat. **(a)** Diffusion tensor images. **From (b) to (d):** our segmentation results using symmetrized KL, geodesic distance and total KL respectively.

## 5 Conclusions

In this paper, we have presented a DTI segmentation framework that utilizes recent advances in continuous energy optimization. Unlike the traditional level-set-based approach, we are the first to use characteristic functions to represent image regions with

different labels and introduce the convex relaxation in the DTI segmentation problem. Given the appearance model, the non-convex image partition problem is relaxed to a convex problem which can be efficiently solved using a first-order primal-dual algorithm. Compared with traditional region-based active contour methods which are sensitive to the contour initialization, our DTI segmentation method is robust and completely automatic without requiring user inputs. Experimental results on both synthetic and real diffusion tensor images validate the proposed method. Furthermore, they also demonstrate the importance and benefit of using more refined and elaborated optimization method in DT-MRI image segmentation, producing consistently superior results.

## References

1. Basser, P., Mattiello, J., Lebihan, D.: Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance-Series B* 103(3), 247–254 (1994)
2. Boyle, J., Dykstra, R.: A method for finding projections onto the intersection of convex sets in hilbert spaces. *Lecture Notes in Statistics* 37(28-47), 4 (1986)
3. Chambolle, A., Cremers, D., Pock, T., et al.: A convex approach to minimal partitions. Technical Report, Ecole Polytechnique, France (2011)
4. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision* 40, 120–145 (2011)
5. Chan, T., Esedoglu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J. on Applied Mathematics* 66(5), 1632–1648 (2006)
6. Chan, T., Vese, L.: Active contours without edges. *IEEE TIP* 10(2), 266–277 (2001)
7. Fletcher, P., Joshi, S.: Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing* 87(2), 250–262 (2007)
8. Goh, A., Vidal, R.: Segmenting fiber bundles in diffusion tensor images. In: *ECCV* (2008)
9. Lenglet, C., Rousson, M., Deriche, R.: Dti segmentation by statistical surface evolution. *IEEE Trans. on Medical Imaging* 25(6), 685–700 (2006)
10. Pennec, X., Fillard, P., Ayache, N.: A riemannian framework for tensor computing. *International Journal of Computer Vision* 66(1), 41–66 (2006)
11. Pock, T., Chambolle, A., Cremers, D., Bischof, H.: A convex relaxation approach for computing minimal partitions. In: *CVPR* (2009)
12. Söderman, O., Jönsson, B.: Restricted diffusion in cylindrical geometry. *Journal of magnetic resonance. Series A* 117(1), 94–97 (1995)
13. Vemuri, B., Liu, M., Amari, S., Nielsen, F.: Total bregman divergence and its applications to dti analysis. *IEEE Trans. on Medical Imaging* 30(2), 475–483 (2011)
14. Wang, Z., Vemuri, B.: Tensor field segmentation using region based active contour model. In: *ECCV* (2004)
15. Wang, Z., Vemuri, B.: DTI segmentation using an information theoretic tensor dissimilarity measure. *IEEE Trans. on Medical Imaging* 24(10), 1267–1277 (2005)
16. Weldeslassie, Y., Hamarneh, G.: DT-MRI segmentation using graph cuts. In: *SPIE Medical Imaging*. vol. 6512 (2007)
17. Zach, C., Gallup, D., Frahm, J., Niethammer, M.: Fast global labeling for real-time stereo using multiple plane sweeps. In: *Vision, modeling, and visualization*. p. 243 (2008)
18. Ziyang, U., Tuch, D., Westin, C.: Segmentation of thalamic nuclei from dti using spectral clustering. *MICCAI* (2006)

# A Computational Framework for Experimental Design in Diffusion MRI

Emmanuel Caruyer, Rachid Deriche

Athena Project-Team, Inria Sophia Antipolis - Méditerranée, France.

**Abstract.** In this work, we develop a computational framework for optimal design of experiment in parametric signal reconstruction. We apply this to the optimal design of one dimensional  $q$ -space,  $q$ -ball imaging and multiple  $q$ -shell experimental design. We present how to construct sampling scheme leading to minimal condition number, and compare to state-of-the-art sampling methods. We show in particular a better noise performance of these scheme through Monte-Carlo simulations for the reconstruction of synthetic signal. This demonstrates the impact of this computational framework on acquisition in diffusion MRI.

## 1 Introduction

Diffusion MRI investigates the properties of tissue microstructure from the analysis of water molecules displacement. The diffusion characteristics, such as the ensemble average propagator or the orientation distribution function, are related to the diffusion signal attenuation through continuous transforms. Since then, the first step in the processing pipeline is usually a parametric estimation of the diffusion signal, from a series of discrete measurements. The number of samples in diffusion MRI is limited to keep the acquisition time compatible with in-vivo measurements. Therefore, the choice of sampling points in the  $q$ -space is critical for a proper reconstruction and quantitative analysis of diffusion characteristics.

The question of sampling efficiency has been widely studied for parametric estimation in diffusion MRI. Several approaches were proposed to uniformly arrange points on the sphere in  $q$ -ball imaging (QBI), using an analogy between the sampling directions and a pair of antipodal electric charges [1, 2], or geometric constructions [3]. For the reconstruction of diffusion tensor MRI, the noise performance has been studied through the minimization of the condition number [4, 5]. In  $q$ -space MRI, several studies on multiple shell sampling [6–9] focused on the efficiency of various sampling strategies, but they do not provide a method to systematically improve the noise performance.

In this work, we give a general method for optimal design of experiment in parametric signal reconstruction. We apply this to the optimal design on one dimensional  $q$ -space experiment,  $q$ -ball imaging and multiple  $q$ -shell experimental design. In the last section, we compare the proposed method to state-of-the-art sampling strategies.

## 2 Theory

### 2.1 Parametric estimation of the diffusion signal

The diffusion signal is approximated in a finite, orthonormal basis of functions

$$\forall \mathbf{q} \in \Omega, \quad E(\mathbf{q}) = \sum_{i=1}^R c_i f_i(\mathbf{q}), \quad (1)$$

where  $\Omega \subset \mathbb{R}^3$ . Depending on the application, we have  $\Omega = \mathbb{R}$  (1D diffusion signal),  $\Omega = \mathcal{S}^2$  (QBI) or  $\Omega = \mathbb{R}^3$  ( $q$ -space imaging).

Provided  $K$  measurements  $y_k = E(\mathbf{q}_k)$  of the signal at wavevectors  $\mathbf{q}_k$ , the coefficients  $\hat{c}_i$  are estimated by least squares. Put in matrix form, we write

$$\hat{\mathbf{c}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y}. \quad (2)$$

$\mathbf{H}$  is the design matrix, and has entries  $H_{ki} = f_i(\mathbf{q}_k)$ .

### 2.2 Information matrix and optimal design

We present a general method to choose the sampling points  $\mathbf{q}_k$  to optimize the noise performance. A useful index for noise performance and stability of the reconstruction is the condition number  $\kappa(\mathbf{M}) = \lambda_{\max}(\mathbf{M})/\lambda_{\min}(\mathbf{M})$  of the information matrix  $\mathbf{M} = 1/K \mathbf{H}^T \mathbf{H}$ , where  $K$  is the number of measurements. The condition number is an upper bound to the error propagation from the measurements to the coefficients estimates. The optimal value of  $\kappa(\mathbf{M})$  is 1, in which case the information matrix is proportional to the identity  $\mathbf{I}_R$ .

The coefficients of the information matrix  $M_{ij}$  can be interpreted as the approximation of the continuous dot product  $\langle f_i, f_j \rangle$

$$M_{ij} = \frac{1}{K} \sum_{k=1}^K f_i(\mathbf{q}_k) f_j(\mathbf{q}_k) \approx \int_{\Omega} f_i(\mathbf{q}) f_j(\mathbf{q}) d\Omega(\mathbf{q}) = \delta_{ij} \quad (3)$$

The basis is orthonormal, hence if this approximation is exact,  $\mathbf{M} = \mathbf{I}_R$ , and the associated condition number equals 1. This naturally introduces the notion of quadrature formula, and its generalization to higher dimension, called the cubature formula.

**Definition 1.** A cubature formula for the integral  $\mathcal{I} = \int_{\Omega} g(\mathbf{q}) d\Omega(\mathbf{q})$  is a collection of nodes  $\mathbf{q}_s$  and weights  $\omega_s$  such that

$$\mathcal{I} = \sum_{s=1}^S \omega_s g(\mathbf{q}_s) \quad (4)$$

If such a cubature formula exists for the integral in Eq. 3, then we place the sampling points at nodes  $\mathbf{q}_s$ , and the number of repetitions  $K_s$  at node  $\mathbf{q}_s$  is proportional to the weight  $\omega_s$ .

### 3 Methods

In this section, we derive cubature formulae for the simple harmonic oscillator basis [10] (1D diffusion signal), spherical harmonic basis [11] (QBI) and spherical polar Fourier basis [12] ( $q$ -space imaging).

#### 3.1 Optimal design in one dimensional $q$ -space MRI

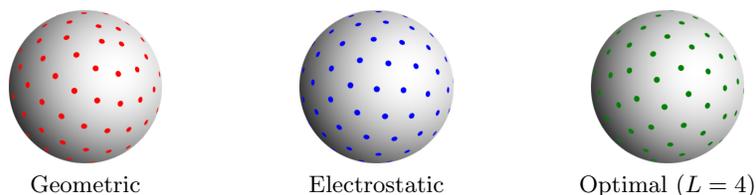
The simple harmonic oscillator basis for the reconstruction of real diffusion signal in one dimension [10] is given by  $\Phi_i(q, u) = \kappa_i(u) \exp(-2\pi^2 q^2 u^2) L_i^{-1/2}(4\pi^2 q^2 u^2)$ , where  $u$  is a characteristic length,  $L_i^{-1/2}$  the generalized Laguerre polynomial of degree  $i$  and  $\kappa_i(u)$  a normalization constant.

Put back into the general framework presented in Section 2.2, we have  $\Omega = \mathbb{R}$ , and the basis functions are  $f_i = \Phi_i$ . For the evaluation of the dot product in Eq. 3, we use the substitution  $x = 4\pi^2 q^2 u^2$ , so that

$$\langle \Phi_i, \Phi_j \rangle = 2\pi u \kappa_i(u) \kappa_j(u) \int_0^\infty L_i^{-1/2}(x) L_j^{-1/2}(x) x^{-1/2} e^{-x} dx. \quad (5)$$

When the basis is truncated to order  $N$ , the evaluation of Eq. 5 reduces to the problem of Gauss-Laguerre quadrature [13]. The optimal samples are  $q_s = \sqrt{x_s}/2\pi u$ , with  $K_s$  repetitions, where  $K_s$  is proportional to  $x_s e^{x_s} / [L_N^{-1/2}(x_s)]^2$  and the nodes  $x_s, s = 1 \dots N + 1$  are the roots of  $L_{N+1}^{-1/2}$ .

#### 3.2 Optimal design in $q$ -ball imaging



**Fig. 1.** Arrangements of  $K = 50$  points on the unit sphere.

The real, symmetric spherical harmonic basis  $\{Y_{lm}\}$  truncated to order  $L$  has dimension  $R = (L+1) \cdot (L+2)/2$ . Put back into the general framework presented in Section 2.2, we have  $\Omega = \mathcal{S}^2$ , and the basis functions are  $f_i = Y_i$ , where  $i(l, m) = 1, 2, \dots, R$  for  $(l, m) = (0, 0), (2, -2), \dots, (L, L)$ .

This basis is equivalent to the basis of harmonic polynomial of degree  $L$  on  $\mathcal{S}^2$ , for which cubature formulae exist and are called spherical design.

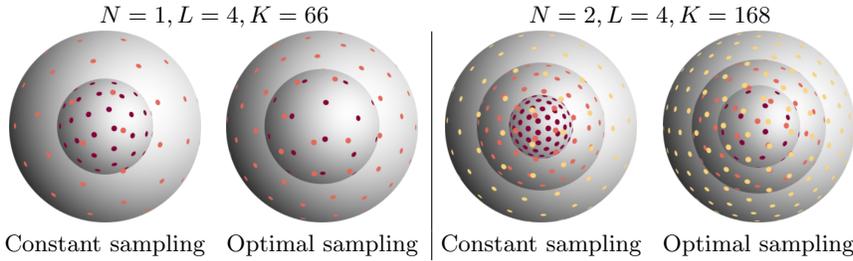
**Definition 2.** A spherical  $t$ -design [14] is a sequence of  $K$  points  $\mathcal{X} = (\mathbf{u}_k)$ ,  $k = 1 \dots K$  on the unit sphere, such that the integral of any polynomial  $p(x, y, z)$

of degree at most  $t$  over the sphere is equal to the average value of the polynomial on  $\mathcal{X}$ :

$$\frac{1}{K} \sum_{k=1}^K p(u_{kx}, u_{ky}, u_{kz}) = \int_{\mathcal{S}^2} p(\omega) d^2\omega. \quad (6)$$

If the sampling directions  $\mathbf{u}_k$  form a spherical  $2L$ -design, then the approximation in Eq. 3 is exact and  $\kappa(\mathbf{M}) = 1$ . For the construction of a spherical  $2L$ -design with antipodal symmetry, we rely here on the equivalence criterion in [15]. An example of optimal direction set for order  $L = 4$  is presented on Fig. 1, and compared to a geometric [3] and an electrostatic [1, 2] arrangements of points.

### 3.3 Optimal design in $q$ -space imaging



**Fig. 2.** 2-shell (left) and 3-shell (right) acquisition protocols: regular multiple  $q$ -shell with constant number of points per shell, and optimal arrangement with minimal condition number for reconstruction in SPF basis.

The truncated spherical polar Fourier (SPF) basis [12] is able to represent the diffusion signal in the whole  $q$ -space. To represent a continuous signal, which verifies  $E(\mathbf{0}) = 1$ , we have recently proposed a slightly modified version of the SPF basis [16]. We reconstruct the signal as  $E(q \cdot \mathbf{u}) = \exp(-q^2/2\zeta) + \sum_{nlm} a_{nlm} C_{nlm}(q \cdot \mathbf{u})$ . The basis functions are  $C_{nlm}(q \cdot \mathbf{u}) = F_n(q) Y_{lm}(\mathbf{u})$ , with

$$F_n(q) = \chi_n \frac{q^2}{\zeta} \exp\left(-\frac{q^2}{2\zeta}\right) L_n^{5/2}\left(\frac{q^2}{\zeta}\right), \quad (7)$$

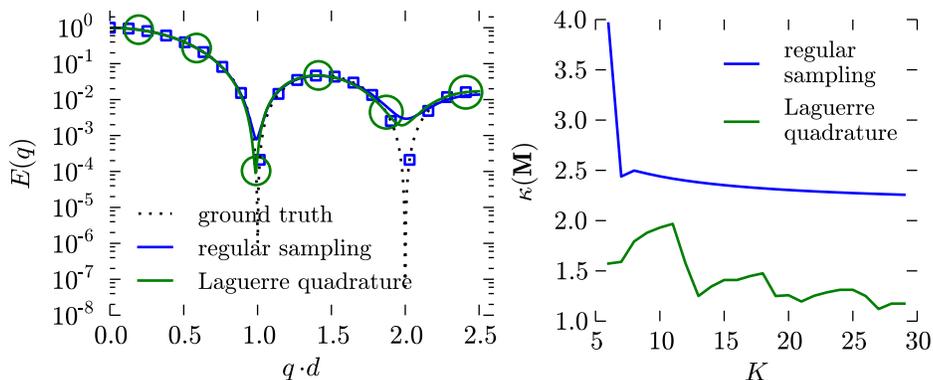
$\zeta$  is a scale factor,  $\chi_n$  a normalization constant, and  $Y_{lm}$  is the real spherical harmonic function. When the radial and angular truncation orders are  $N$  and  $L$ , respectively, this basis has dimension  $R = N \cdot (L + 1) \cdot (L + 2)/2$ . Put back into the general framework presented in Section 2.2, we have  $\Omega = \mathbb{R}^3$ , and the basis functions are  $f_i = C_i$ , where  $i(n, l, m) = 1, 2, \dots, R$  for  $(n, l, m) = (0, 0, 0), (0, 2, -2), \dots, (N, L, L)$ .

For the construction of an optimal design for this basis, we build on the findings of the previous two sections. We show that the radial part of the integral in Eq. 3 reduces to a Gauss-Laguerre quadrature problem, while the angular part reduces to a spherical design problem.

Therefore we propose a design on  $N + 1$  spheres in the  $q$ -space. The shell  $s$  has radius  $q_s = \sqrt{\zeta x_s}$ , where  $x_s$  is the  $s^{\text{th}}$  root of  $L_{N+1}^{5/2}$ . The number of points  $K_s$  on shell  $s$  should be proportional to  $\omega_s = \exp(-x_s)/[x_s(L_N^{5/2}(x_s))^2]$ . Finally, the points on each sphere should form a spherical  $2L$ -design. Example of points sets generated with this method are depicted on Fig. 2. They are compared to multiple shell sampling where the shell radii are evenly spaced, and the number of points equal on each shell, as suggested in [7].

## 4 Experiments and Results

### 4.1 Quadrature formula in one dimensional $q$ -space MRI



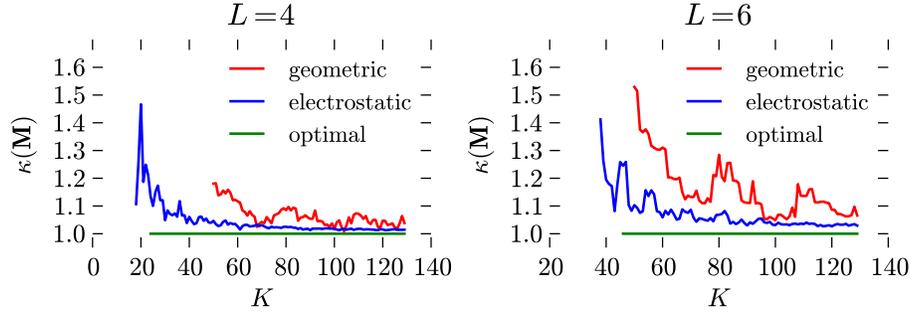
**Fig. 3.** Evaluation of Gauss-Laguerre quadrature for one dimensional  $q$ -space. (Left) An example of signal and its reconstruction. The blue squares and the green circles represent the regular and quadrature samples respectively. The radii of the circles are proportional to the number of repeated acquisitions. (Right) Condition number of the information matrix.

In this section, we show the feasibility of one dimensional  $q$ -space signal reconstruction from a set of measurements on a limited support size. We compare the Gauss-Laguerre quadrature to a regular sampling on the range  $[0, q_{\max}]$ .

We plot on Fig. 3 an example of reconstruction of a diffusion signal corresponding to the restricted diffusion between two parallel planes, separated by distance  $d$  [10]. The truncation order in the basis was set to  $N = 5$ , and the corresponding Gauss-Laguerre quadrature works on 6 nodes. The result for a total of 20 acquisitions is visually identical to the reconstruction from a regular sampling. Besides, the associated information matrix is better conditioned for the quadrature sampling. The reason why the condition number is not exactly 1 in this case is that the quadrature weight  $\omega_s$  is approximated by the number of repetitions at node  $q_s$ , which is an integer.

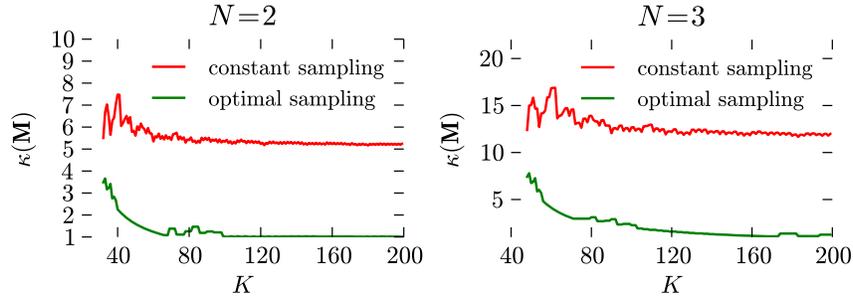
## 4.2 Evaluation of conventional schemes in $q$ -ball imaging

We evaluate and report on Fig. 4 the noise performance of point sets generated with electrostatic analogy [1,2] and by geometrical construction [3], for the reconstruction of SH coefficients of the diffusion signal. We compare these sampling methods to the proposed, optimal point set based on spherical design.



**Fig. 4.**  $q$ -ball imaging: condition number of the information matrix corresponding to the truncated SH up to order  $L = 4$  (left) and  $L = 6$  (right), of the electrostatic, geometric and optimal point sets. The geometric configurations are only provided for  $K \geq 50$ , as this method is reported to be dedicated to large  $K$  by the author in [3]. Our proposed, optimal design is based on spherical design, and therefore exists for  $K \geq 24$  at order  $L = 4$ , and for  $K \geq 46$  at order  $L = 6$  [15]. By construction, the condition number associated is exactly 1.

## 4.3 Multiple $q$ -shell and three dimensional signal reconstruction

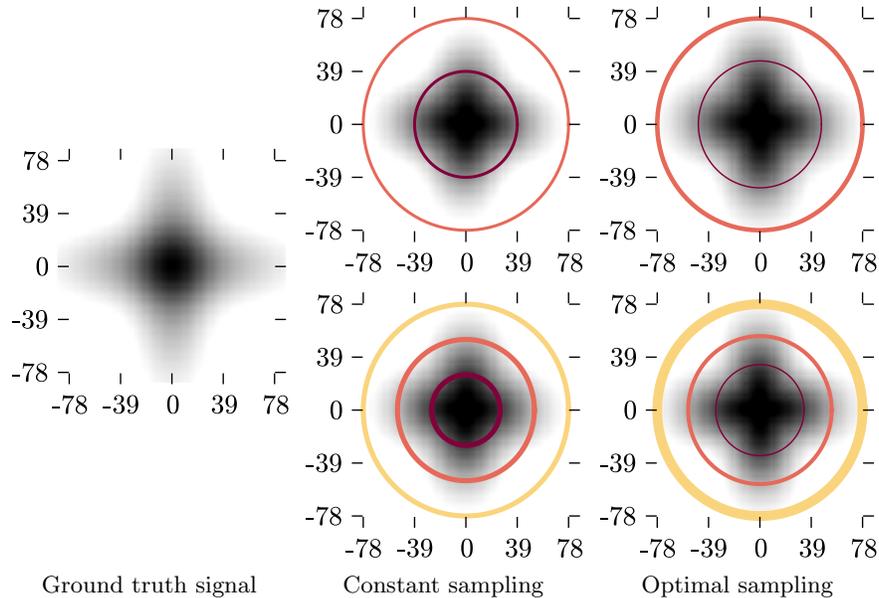


**Fig. 5.** Condition number of multiple shell sampling, corresponding to the SPF basis.

Using the results on Gauss-Laguerre quadrature and spherical designs, we generated optimal sampling schemes on multiple shells for the reconstruction in

the SPF basis. We compare this to the sampling strategy with shell radii evenly spaced and constant number of points per shell proposed in [7]. The condition number for the reconstruction in SPF basis is reported on Fig. 5.

We also simulate both methods, for the sampling and reconstruction of a synthetic diffusion signal corresponding to a mixture of Gaussian. Visual reconstruction is reported on Fig. 6, and quantitative comparison on Fig. 7.

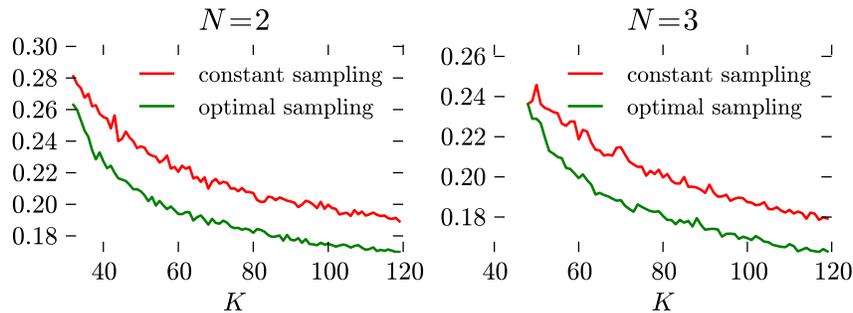


**Fig. 6.** 2 shell (top) and 3 shell (bottom) sampling in action, for the reconstruction of a synthetic diffusion signal corresponding to a mixture of Gaussian, simulating fiber crossing. The circles represent the sampling shells, and the line widths are proportional to the number of points per shell.

## 5 Conclusion

In this work, we develop a computational framework for optimal design of experiment in diffusion MRI. For the reconstruction of 1D, spherical and 3D signal, we propose sampling scheme with minimal condition number. Monte-Carlo simulations confirm this result, as the signal to noise ratio of the parameters estimated from optimal sampling scheme is improved with respect to conventional sampling scheme, for the same number of acquisitions.

As a conclusion to this study, we claim that a sampling method is optimal for a reconstruction in a given basis and a given order. In addition to the technical and physical limitations of the imaging system, the choice of the type of reconstruction must be taken into account when designing the acquisition protocol.



**Fig. 7.** Mean squared error for a synthetic diffusion signal corresponding to a mixture of Gaussian, with Rician noise (SNR=25).

## References

1. Jones, D., Horsfield, M., Simmons, A.: Optimal strategies for measuring diffusion in anisotropic systems by MRI. *Magn Reson Med* **42**(39) (August 1999) 515 – 525
2. Jansons, K.M., Alexander, D.C.: Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems* **19** (2003) 1031–1046
3. Koay, C.G.: A simple scheme for generating nearly uniform distribution of antipodally symmetric points on  $S^2$ . *J. Comput. Science* **2**(4) (2011) 377–381
4. Papadakis, N.G., Xing, D., Huang, C.L., Hall, L.D., Carpenter, T.A.: A comparative study of acquisition schemes for DTI. *J Magn Reson* **137**(1) (1999) 67–82
5. Skare, S., Hedehus, M., Moseley, M.E., Li, T.Q.: Condition number as a measure of noise performance of diffusion tensor data acquisition schemes with MRI. *J Magn Reson* **147**(2) (December 2000) 340–352
6. Caruyer, E., Cheng, J., Lenglet, C., Sapiro, G., Jiang, T., Deriche, R.: Optimal design of multiple q-shells experiments for diffusion MRI. In: MICCAI Workshop on Computational Diffusion MRI - CDMRI'11, Toronto, Canada (September 2011)
7. Assemlal, H.E., Tschumperlé, D., Brun, L.: Evaluation of q-space sampling strategies for the diffusion magnetic resonance imaging. In: MICCAI, London (2009)
8. Portnoy, S., Ye, W., Entezari, A., Blackband, S.J., Vemuri, B.C.: A novel interlaced sampling scheme for multi-shell q-space MR microscopy. In: ISMRM. (2011)
9. Daducci, A., McEwen, J., Van De Ville, D., Thiran, J.P., Wiaux, Y.: Harmonic analysis of spherical sampling in diffusion MRI. In: ISMRM. (2011) 3929
10. Ozarslan, E., Koay, C., Basser, P.: Simple harmonic oscillator based estimation and reconstruction for 1D q-space MR. In: ISMRM. Volume 16. (2008) 35
11. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical q-ball imaging. *Magn Reson Med* **58**(3) (2007) 497–510
12. Assemlal, H., Tschumperlé, D., Brun, L.: Efficient and robust computation of pdf features from diffusion MR signal. *Medical Image Analysis* **13**(5) (2009) 715–729
13. Abramowitz, M., Stegun, I.A.: *Handbook of mathematical functions*. Dover (1970)
14. Delsarte, P., Goethals, J., Seidel, J.: Spherical codes and designs. *Geometriae Dedicata* **6** (1977) 363–388 10.1007/BF03187604.
15. Goethals, J.M., Seidel, J.J.: *Cubature formulae, polytopes and spherical designs*. Springer-Verlag, New York (1981) 203–218
16. Caruyer, E., Deriche, R.: Diffusion mri signal reconstruction with continuity constraint and optimal regularization. *Medical Image Analysis* (2012, In press)

# A Framework for Multi-task Bayesian Compressive Sensing of DW-MRI

Julio M. Duarte-Carvajalino,<sup>1</sup> Christophe Lenglet,<sup>1</sup> Kamil Ugurbil,<sup>1</sup> Steen Moeller,<sup>1</sup> Lawrence Carin,<sup>2</sup> Guillermo Sapiro<sup>2</sup>

<sup>1</sup> Center for Magnetic Resonance Research, Minneapolis, USA.

<sup>2</sup> Electric and Computer Engineering Department, Duke University, Durham, USA.

**Abstract.** We present a framework to significantly reduce the acquisition time of diffusion-weighted magnetic resonance (DW-MR) data. The proposed approach is based on multi-task Bayesian compressive sensing (MT-BCS). The MT-BCS framework models the statistical relationships within groups (clusters) of diffusion signals, thereby better exploiting the data redundancy. In addition to providing accurate reconstruction of DW-MR data from its compressed sensing samples, this framework also generates probabilistic uncertainty estimates. Experiments validating the proposed framework are performed using synthetic and in-vivo DW-MR data sets. The results indicate a significant reduction in the number of required diffusion volumes, hence reducing acquisition time, by a factor of at least four, sacrificing little or nothing in terms of reconstruction accuracy. This framework can be applied to any diffusion protocol, high angular resolution diffusion imaging (HARDI) in particular, and it can be naturally combined with parallel MR imaging for a combined acceleration factor of at least 12. Finally, the proposed framework can also be exploited to enhance previously acquired data.

**Keywords:** Bayesian compressive sensing, diffusion-weighted MR, Dirichlet process, magnetic resonance imaging, multi-task.

## 1 Introduction

Diffusion weighted magnetic resonance imaging (DW-MRI) is an MRI modality that characterizes the anisotropic diffusion of water molecules in biological tissues, such as brain white matter, along several directions. This information can be used to estimate the local orientation of white matter fiber bundles, since water diffusion is favored along the axonal fibers, providing critical information for neuroscience and clinical studies [1]. The diffusion tensor (DT) formalism [2] was introduced to describe the three-dimensional nature of anisotropic diffusion in biological tissues. The corresponding diffusion tensor images (DTI) have been used to investigate the anatomical structure of white matter, via tractography and connectivity mappings, enabling the study of a wide range of neurological disorders, brain tumors, brain development, and aging. However, the diffusion tensor model breaks down when two or more fiber bundles cross; between one-third and two-thirds of imaging voxels in the human brain's white matter might

contain multiple fiber crossings [3]. DW-MRI techniques [4] such as diffusion spectrum imaging (DSI) [5, 6, 8] and q-ball Imaging (QBI) [7], have been proposed to overcome the limitations of DTI and recover fiber crossing information. DSI samples the diffusion domain (q-space) on a Cartesian grid of size  $O(n^3)$  at each image voxel, leading to relatively long acquisition times, which severely limits its clinical application (acceleration techniques have been proposed, see below). Q-ball imaging, on the other hand, favors angular resolution to map fiber orientations and samples the q-space on a single or multiple concentric spherical shells, requiring only  $O(n^2)$  diffusion volumes [7]. Nevertheless, q-ball imaging often uses a relatively large number of diffusion volumes to accurately recover the water diffusion orientation distribution function (ODF) [9, 10], leading also to relatively long acquisition times. Novel techniques of image acceleration such as parallel imaging can be used to perform multi-slice acquisition providing accelerations by acquiring multiple slices simultaneously [11–14] that can be combined with CS to accelerate image acquisition.

Recently, several research efforts have been made to reduce the acquisition time of MRI using compressive sensing (CS) theory [16, 17, 19–22], in combination with MRI parallel imaging [23–26], for multiple contrast MRIs [27], and dynamic MRIs [28–30]. However, little has been done to reduce the acquisition time of DW-MRI. Classic CS theory has been recently proposed to reduce the number of q-samples in DW-MRI for both DSI [15] and QBI [31–33] sensing protocols. These studies show that compressed sensed DSI and QBI remain diagnostically useful with up to four acceleration factors in the diffusion domain.

In this work we present a multi-task Bayesian compressive sensing (MT-BCS) framework developed in [34, 35]. A key difference with the relevant previous CS work [15, 31, 32] is that we use Bayesian inference to estimate the DW-MRI from CS samples in q-space using probabilistic priors such as sparsity and statistical relationships within groups (clusters) of diffusion samples. Contrary to traditional CS, where the reconstruction is performed independently on each voxel, the BCS reconstruction is performed on groups of voxels, sharing common statistical properties, hence solving multiple CS tasks simultaneously (multi-tasking). Previous work accounts for the spatial redundancy of diffusion data by imposing a total variation constrain; in this work, diffusion signals are explicitly clustered based on their sparsity patterns, accounting thus for the spatial redundancy. In addition to estimating the DW-MR signals from their CS measurements, the proposed MT-BCS framework automatically provides the uncertainty of those estimates. This framework can be applied to any DW-MR sensing protocol such as DSI and QBI. Experimental results indicate that a factor of at least four in acceleration is obtained.

## 2 Methods

### 2.1 Compressive Sensing of Diffusion Weighted MRI

Let  $\mathbf{W} = \{\mathbf{W}_0, \mathbf{W}_1, \dots, \mathbf{W}_K\} \in \mathbb{R}^{N_x \times N_y \times N_z \times K}$  be the DW-MR data, where  $\mathbf{W}_0$  represents the b0 volume (no gradient direction),  $\{\mathbf{W}_k\}_{k=1}^K$  the  $K$  diffusion

volumes (gradient directions in QBI, Cartesian samples in DSI), and  $N_x, N_y, N_z$  the number of voxels along each spatial dimension (sagittal, coronal, and axial in radiological coordinates, for instance). The normalized diffusion volumes are  $\tilde{\mathbf{W}}_k = \mathbf{W}_k \oslash \mathbf{W}_0$ , where  $\oslash$  is the Hadamard division. Let  $\tilde{\mathbf{h}}_k = \text{Vect}\{\tilde{\mathbf{W}}_k\}$  be the vectorization of the diffusion volumes of size  $N \times 1$ , where  $N = N_x \times N_y \times N_z$ . Hence, the DW-MR data can be conveniently represented by a matrix  $\mathbf{S}_{K \times N} = [\mathbf{s}_1 \dots \mathbf{s}_N]$ , where each column  $\{\mathbf{s}_i\}_{i=1}^N$  ( $K \times 1$ ) represents the diffusion signal (in q-space) at each voxel. Let  $\Psi_{K \times M}$ ,  $K \leq M$ , be a sparsifying dictionary such that the diffusion signals can efficiently be represented as

$$\mathbf{S}_{K \times N} = \Psi_{K \times M} \boldsymbol{\theta}_{M \times N} + \boldsymbol{\eta}_{K \times N}, \quad \boldsymbol{\theta} = [\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_N], \quad \|\boldsymbol{\theta}_i\|_0 \ll M, \quad (1)$$

where  $\boldsymbol{\theta}$  is the sparse representation of the diffusion signals in the dictionary  $\Psi$  and  $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \alpha_0^{-1} \mathbf{I})$  is an additive Gaussian representation ‘‘noise.’’ It should be pointed here that the measurement noise in diffusion weighted MRIs is considered to be Riccian distributed [36]. Nevertheless, assuming additive Gaussian noise makes the Bayesian inference mathematically tractable and leads to very good estimates, as shown here as well as in previous studies [33] (note also that this is the noise of the representation, and not the physical noise). Finally, let  $\Phi_{L \times K}$ ,  $L < K$ , be a CS matrix such that

$$\mathbf{Y}_{L \times N} = \Phi_{L \times K} \Psi_{K \times M} \boldsymbol{\theta}_{M \times N} + \Phi_{L \times K} \boldsymbol{\eta}_{K \times N}, \quad (2)$$

where  $\mathbf{Y}_{L \times N}$  represents the CS DW-MR dataset. Under-sampling the q-space corresponds to selecting  $L < K$  rows of the dictionary  $\Psi_{K \times M}$ , such that  $\tilde{\Psi}_{L \times M} = \Phi_{L \times K} \Psi_{K \times M}$  is an equivalent dictionary requiring less diffusion samples

$$\mathbf{Y}_{L \times N} = \tilde{\Psi}_{L \times M} \boldsymbol{\theta}_{M \times N} + \tilde{\boldsymbol{\eta}}_{L \times N}. \quad (3)$$

Classical CS theory ensures that  $\mathbf{S}_{K \times N}$  can be reconstructed from the CS signals  $\mathbf{Y}_{L \times N}$  with high accuracy using constrained  $l_0$  or  $l_1$  minimization, as long as the signals can be represented sparsely with an appropriate dictionary  $\Psi$  and the CS sampling is incoherent with the basis. The work of [15, 33] shows that wavelets provide a good sparsifying dictionary for DSI data sets, while [31, 32] show that spherical ridgelets [37] constitutes an appropriate sparsifying dictionary for QBI data. It has also been recognized that pure random CS of the k-space in MRI is impractical, adds incoherent artifacts, and does not fully cover the signal space [16–18, 22, 28]. Hence, deterministic, pseudo-random and optimized CS strategies have been proposed that satisfy the incoherence principle. Similarly, in DW-MRI, pseudo-random sampling of the q-space has been proposed, using random-Gaussian sampling centered in the Cartesian q-space grid for DSI [15], and pseudo-random Dirac sampling fully covering the unit semi-sphere for QBI [37] that satisfy the incoherence principle, as shown in [31]. In this work, we use the optimal pseudo-random Dirac sampling of multi-shell QBI proposed in [9], which provides gradient directions that remain approximately optimal (cover the unit semi-sphere) when truncated before the acquisition is complete. The mutual coherence between CS Dirac sampling of the q-space and

spherical ridgelets is as low as using pure random sampling, with the advantage of covering the whole hemisphere [31]. Using standard CS optimization methods, a point estimate  $\hat{\boldsymbol{\theta}}_{M \times N}$  of the sparse representation of the data in the dictionary  $\boldsymbol{\Psi}$  can be obtained from the CS signals  $\mathbf{Y}_{L \times N}$ , and from there an estimate  $\hat{\mathbf{S}}_{K \times N} = \boldsymbol{\Psi}_{K \times M} \hat{\boldsymbol{\theta}}_{M \times N}$  of the DW-MR data with sufficiently large number of diffusion volumes ( $K$ ) to robustly estimate the ODF.

## 2.2 Bayesian Inference

Rather than just obtaining a deterministic point estimate of  $\boldsymbol{\theta}$ , we propose here to use Bayesian CS to perform inference on the distribution of  $\boldsymbol{\theta}$  using a sparsity prior and a shared hyper-prior that models similar sparsity patterns between diffusion signals. Specifically,

$$p(\boldsymbol{\theta}_i | \boldsymbol{\alpha}_i) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}_i}), \quad \boldsymbol{\Lambda}_{\boldsymbol{\alpha}_i} = \text{diag}\{\alpha_{i1}^{-1}, \dots, \alpha_{iM}^{-1}\}, \quad i = 1, \dots, N, \quad (4)$$

is the a priori distribution of  $\boldsymbol{\theta}_i$ , modeled as a multidimensional Gaussian with zero mean and diagonal covariance  $\boldsymbol{\Lambda}_{\boldsymbol{\alpha}}$ . As shown in [34, 35], this prior promotes sparse representations. Hyper-parameters  $\{\boldsymbol{\alpha}_i\}_{i=1}^N$  in (4) model information sharing among diffusion signals (tasks). Specifically, diffusion signals are clustered using a nonparametric Dirichlet process (DP) mixture model.<sup>3</sup> A Dirichlet process  $G(\lambda, G_0)$  is a distribution of distributions with scaling variable  $\lambda$  and base distribution  $G_0$ . Hyper-parameters  $\boldsymbol{\alpha}_i$  are modeled as independent identically distributed (*iid*) random variables drawn from a distribution generated by  $G(\lambda, G_0)$ . Let  $\{\boldsymbol{\alpha}_j^*\}_{j=1}^J$  be the distinct values taken by  $\{\boldsymbol{\alpha}_i\}_{i=1}^N$ , then [35]

$$p(\boldsymbol{\alpha}_i | \boldsymbol{\alpha}^{-i}) = \frac{\lambda}{\lambda + N - 1} G_0 + \frac{1}{\lambda + N - 1} \sum_{j=1, j \neq i}^J n_j^{-i} \delta_{\boldsymbol{\alpha}_j^*}, \quad (5)$$

where  $\boldsymbol{\alpha}^{-i} = \{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{i-1}, \boldsymbol{\alpha}_i, \dots, \boldsymbol{\alpha}_N\}$ ,  $n_j^{-i}$  the number of times  $\boldsymbol{\alpha}_j^*$  is in  $\boldsymbol{\alpha}^{-i}$ , and  $\delta_{\boldsymbol{\alpha}_j^*}$  the distribution concentrated at  $\boldsymbol{\alpha}_j^*$ . From (5) it can be seen that  $\boldsymbol{\alpha}_i$ , given all other observations  $\boldsymbol{\alpha}^{-i}$ , has a probability  $\frac{\lambda}{\lambda + N - 1}$  of being drawn from distribution  $G_0$ , or it can be drawn from the existing  $\{\boldsymbol{\alpha}_j^*\}_{j=1}^J$ , with probabilities proportional to the corresponding cluster sizes  $n_j^{-i}$ . This explains the sharing and non-parametric properties of DPs: A new sample can either join one of the previous hyper-parameters or it can create a new one (from  $G_0$ ). The random variable  $\lambda$  controls the balance between the creation of new hyper-parameters and sharing with the existing hyper-parameters. The likelihood of CS measurements  $\{\mathbf{Y}_i\}_{i=1}^N$  given  $\boldsymbol{\theta}_i$  and  $\alpha_0$  is given by  $p(\mathbf{Y}_i | \boldsymbol{\theta}_i, \alpha_0) \sim \mathcal{N}(\tilde{\boldsymbol{\Psi}}\boldsymbol{\theta}_i, \alpha_0^{-1}\mathbf{I})$ .

$G$  can be modeled as the mixture  $G = \sum_{j=1}^{\infty} w_j \delta_{\boldsymbol{\alpha}_j^*}$ , using a stick-breaking process [38]. Hence, the DP multi-task CS (DP-MT-CS) model is defined by

$$\mathbf{Y}_i | \boldsymbol{\theta}_i, \alpha_0 \sim \mathcal{N}(\tilde{\boldsymbol{\Psi}}\boldsymbol{\theta}_i, \alpha_0^{-1}\mathbf{I}), \quad \boldsymbol{\theta}_i | \mathbf{z}_i, \{\boldsymbol{\alpha}_j^*\}_{j=1}^J \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda}_{\boldsymbol{\alpha}_i^*}),$$

<sup>3</sup> The DP is non-parametric since the number of clusters is not determined a priori.

$$\mathbf{z}_i | \{w_j\}_{j=1}^J \sim \text{Multinomial}(\{w_j\}_{j=1}^J), \quad w_j = \pi_j \prod_{l=1}^{j-1} (1 - \pi_l), \quad \pi_j \sim \text{Beta}(1, \lambda),$$

$$\lambda | e, f \sim \text{Gamma}(e, f), \quad \boldsymbol{\alpha}_j^* | c, d \sim \prod_{m=1}^M \text{Gamma}(c, d), \quad \alpha_0 \sim \text{Gamma}(a, b), \quad (6)$$

where  $1 \leq i \leq N, 1 \leq J < \infty, a, b, c, d, e, f$  are non-informative hyper-parameters determining the distribution of the random variables in the model, and  $\mathbf{z}_i$  corresponds to a draw from a multinomial distribution of weights  $\{w_j\}_{j=1}^J$  such that  $\boldsymbol{\alpha}_i = \sum_{j=1}^J z_{ij} \boldsymbol{\alpha}_j^*$ . Notice that  $\mathbf{z}_i$  is a variable that indicates how much of each hyper-parameter  $\boldsymbol{\alpha}_j^*$  is used in the mixture to model  $\boldsymbol{\alpha}_i$ . Hence it provides a probabilistic (soft) clustering of the data. Bayesian inference using the DP-MT-CS model is provided by Baye's rule

$$p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon}) = \frac{p(\mathbf{Y} | \mathbf{H}) p(\mathbf{H} | \boldsymbol{\Upsilon})}{\int p(\mathbf{Y} | \mathbf{H}) p(\mathbf{H} | \boldsymbol{\Upsilon}) d\mathbf{H}}, \quad (7)$$

where  $\mathbf{H} = \{\{\boldsymbol{\theta}_i\}_{i=1}^N, \{\boldsymbol{\alpha}^*\}_{j=1}^J, \{\mathbf{z}_i\}_{i=1}^N, \{\pi_j\}_{j=1}^J, \alpha_0, \lambda\}$  are the hidden model variables,  $\mathbf{Y}$  the CS measurements, and  $\boldsymbol{\Upsilon} = \{a, b, c, d, e, f\}$  the hyper-parameters determining the distribution of the hidden variables. Equation (7) cannot be solved analytically, due to the mathematical intractability of the integration term in the denominator. A variational method can be used to seek a distribution  $q(\mathbf{H})$  that approximates  $p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon})$

$$p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon}) \approx \text{arg}_{q(\mathbf{H})} \min \{ \mathcal{D}_{KL}(q(\mathbf{H}), p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon})) \} \quad (8)$$

where  $\mathcal{D}_{KL}(q(\mathbf{H}), p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon}))$  is the Kullback-Leibler divergence between the approximate  $q(\mathbf{H})$  and the true posterior  $p(\mathbf{H} | \mathbf{Y}, \boldsymbol{\Upsilon})$ . The solution to (8) is tractable using functional derivatives if we define (see more details in [35])

$$q(\mathbf{H}) = q(\alpha_0) q(\lambda) \prod_{j=1}^J q(\pi_j) \prod_{i=1}^N q(\mathbf{z}_i) \prod_{i=1}^N q(\boldsymbol{\theta}_i) \prod_{i=j}^J q(\boldsymbol{\alpha}_j^*). \quad (9)$$

In particular  $q(\boldsymbol{\theta}_i) \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i^{-1})$ , where  $\boldsymbol{\mu}_i$  and  $\boldsymbol{\Sigma}_i$  are estimated by iteratively solving (8). Hence,

$$\hat{\mathbf{S}}_i \sim \mathcal{N}(\boldsymbol{\Psi} \boldsymbol{\mu}_i, \boldsymbol{\Psi} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Psi}^T), \quad (10)$$

providing thus a point estimate  $\hat{\mathbf{S}}_i = \boldsymbol{\Psi} \boldsymbol{\mu}_i$  and its uncertainty  $\boldsymbol{\Psi} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\Psi}^T$ .

Parameters  $a, b, c, d, e, f$  are set to very small values to have a non-informative prior on  $\alpha_0$  and  $\lambda$ , being thus estimated from the data. Experimentally, we found that  $\mathbf{z}_i = [\frac{1}{J} \dots \frac{1}{J}]$  provides a good initialization, with  $J = \log_2(N)$ .

## 3 Experimental Results

### 3.1 Dataset

We first generated a synthetic DW-MR image with 128 gradient directions and three fibers that cross in the center (Figure 1) using the fanDTasia toolbox.<sup>4</sup>

<sup>4</sup> <http://www.cise.ufl.edu/~abarm pou/lab/fanDTasia/>.

In-vivo QBI data was obtained on a 7T Siemens Avanto magnet using a single refocused 2D single shot spin echo EPI sequence, 1.5 mm isotropic voxels, 128 directions at  $b=1500 \text{ s/mm}^2$  and 15 b0s. This dataset consists of four DW-MR series corresponding to four different subjects [39]. Each DW-MR dataset was corrected for eddy current distortions and head motion using a 12 degrees-of-freedom linear registration with the initial b0 image. The data was also corrected for geometric distortions by unwrapping the images using a field map (see [39] for details). We refer to these images as S1 to S4. The gradient table for these images (synthetic and real data) follow the protocol proposed in [9], which provides gradient directions that remain approximately optimal when truncated before the acquisition is complete (as it happens here with CS). An eroded FA mask was employed ( $FA \geq 0.2$ ) for the real data set in order to limit the number of voxels to white matter regions with one or more fiber tracks.

Compressed sensed DW-MR datasets were constructed by taking the first 64, 32, or 16 gradient directions of the full (128 gradient directions) data set. Hence, the full data set is considered as “ground truth” to evaluate the quality of the reconstructed data. The reconstruction results were evaluated as

$$\begin{aligned}
 MSE &= \frac{1}{N} \sum_{i=1}^N \|\mathbf{s}_i - \Psi \hat{\theta}_i\|_2^2, & D_{SH}(\varphi, \hat{\varphi}) &= \frac{1}{N} \sum_{i=1}^N \sum_{l=1}^{n_{SH}} \|\zeta_{il} - \hat{\zeta}_{il}\|_2^2, \\
 FOE &= \frac{1}{N} \sum_{i=1}^N \arccos |\mathbf{u}_0^T \hat{\mathbf{u}}| \frac{180}{\pi}, & &
 \end{aligned} \tag{11}$$

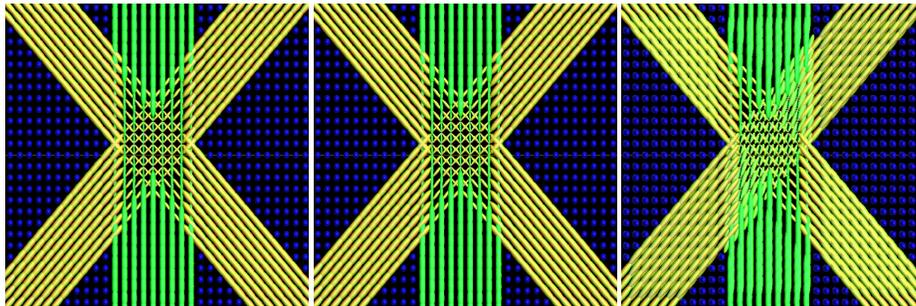
where  $MSE$  is the mean square error between the intensity values of the original and reconstructed DW-MRIs;  $D_{SH}$  the mean square error between the coefficients of the spherical harmonics (SH) representation of the fiber orientation distribution (FOD) in the original ( $\zeta$ ) and reconstructed diffusion dataset ( $\hat{\zeta}$ ), truncated to  $n_{SH}$  harmonic coefficients; and  $FOE$  is the fiber orientation error between the peaks of the FOD in the original ( $\mathbf{u}_0$ ) and reconstructed ( $\hat{\mathbf{u}}$ ) diffusion dataset. Notice that  $N$  here represents the number of voxels that lie within the FA mask. The FOD is estimated following [40] with harmonics of degree 12 for the synthetic image and 8 for the in-vivo data.<sup>5</sup> We use here the FOD rather than the ODF since it provides sharper estimates of fiber orientations down to relatively smaller crossing angles than with the ODF [41].

Best results using the proposed DP-MT-BCS framework with in-vivo data were obtained using a sliding window along the axial axis that processes three neighboring axial slices and returns estimates for the center slice. Figure 1a shows the FODs of the original synthetic image with 128 gradient directions. There are regions containing single fibers and regions containing two and three fiber crossings. Figures 1b and 1c show the FODs estimated from the reconstructed synthetic image using 32 and 16 diffusion directions, respectively. The FODs of the reconstructed DW-MRI using 32 gradient directions are close to the original DW-MRI, while strong differences can be observed on the estimated FODs using the reconstructed DW-MRI with just 16 gradient directions.

<sup>5</sup> Code can be found at <http://www.brain.org.au/software/mrtrix/index.html>.

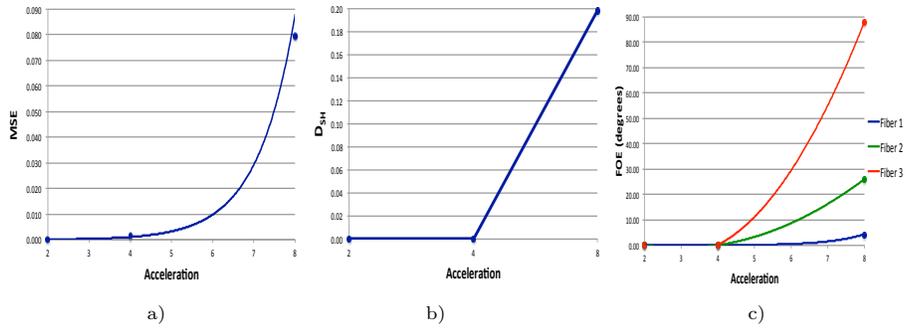
### 3.2 Results

Figure 2 shows the  $MSE$ ,  $D_{SH}$ , and  $FOE$  as a function of the CS acceleration. As can be seen, the error is negligible for accelerations of 2 and 4, while an acceleration of 8 produces large errors that can seriously affect fiber orientation mapping, and tractography. Figure 3a shows the clustering automatically obtained from the DP-MT-BCS, where each color indicates different regions (clusters) that coincide with the different fiber orientations and fiber crossings in the DW-MRI. Figure 3b shows the estimated uncertainty for each acceleration and diffusion volume, obtained 1,000 Gaussian random samples with zero mean and covariance given by (10), then computing the standard deviation for each diffusion volume. Figure 3b shows the estimated a priori uncertainty for each acceleration factor (first 64, 32, and 16 volumes for acceleration factors of 2, 4, and 8 respectively), where it can be seen that the uncertainty increases as the acceleration factor increases, for the estimated volumes. Notice that the uncertainty is about the same for the CS volumes used and for all the acceleration factors. This makes sense, since the minimization considers only the known CS diffusion volumes and hence the uncertainty there is lower than for the estimated volumes.

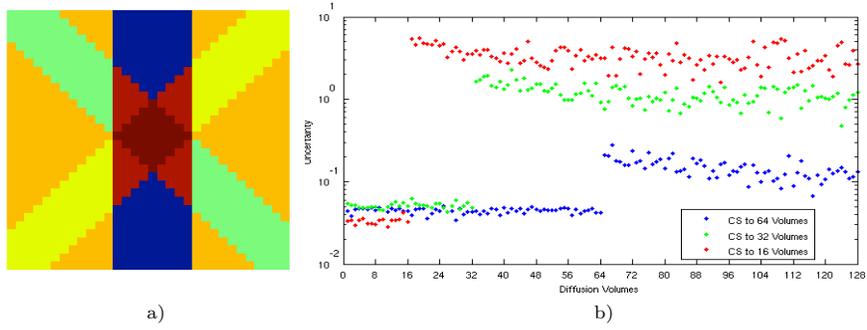


**Fig. 1.** a) Image with 128 gradient directions, b) Image reconstructed from 32 gradient directions, c) Image reconstructed from 16 gradient directions.

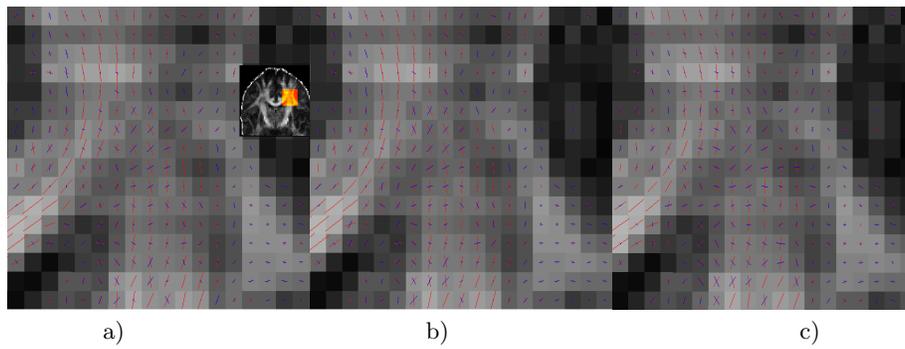
We must point out here that the estimated uncertainty assumes no informative distributions. Hence, large uncertainties do not necessarily mean large estimation errors, as can be appreciated in figures 1 and 2, where the error is very low for accelerations 2 and 4. Despite this, the uncertainty estimates allow us to compare the effects of the different CS acceleration factors. From this point of view, an acceleration of two provides the lowest predicted uncertainty on all the estimated diffusion volumes using CS. Figure 4a shows the first three peaks (fiber bundle directions) of the estimated FODs in the original DW-MRI of S2, while figures 4b and 4c shows the first two peaks of the estimated FODs in the reconstructed DW-MRI using accelerations of 4 and 8, respectively. Figures 5 and 6 show the  $MSE$ ,  $D_{SH}$ , and  $FOE$  for every one of the subjects and ac-



**Fig. 2.** Reconstruction error of the synthetic DW-MRIs: a)  $MSE$ ; b)  $D_{SH}$ ; c) FOE.



**Fig. 3.** a) Clustering; b) Uncertainty in the estimation (standard deviation).



**Fig. 4.** a) First two fiber directions in the DW-MR image of S2 using a) All volumes; b) 32 diffusion volumes; c) 16 diffusion volumes.

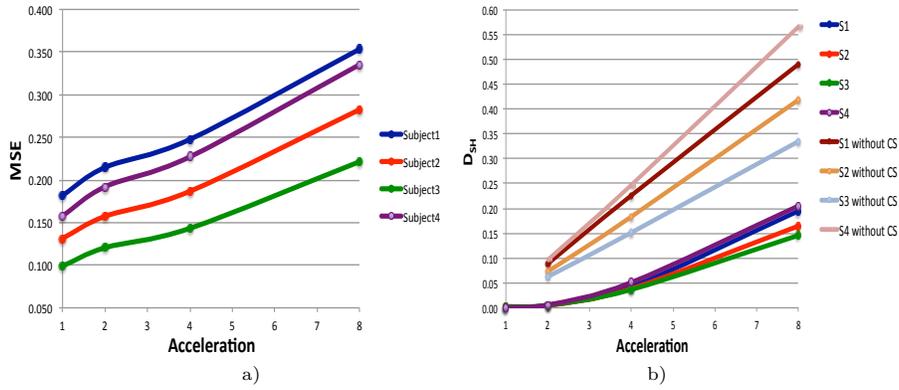


Fig. 5. Reconstruction error for the in-vivo DW-MRIs: a)  $MSE$ ; b)  $D_{SH}$ .

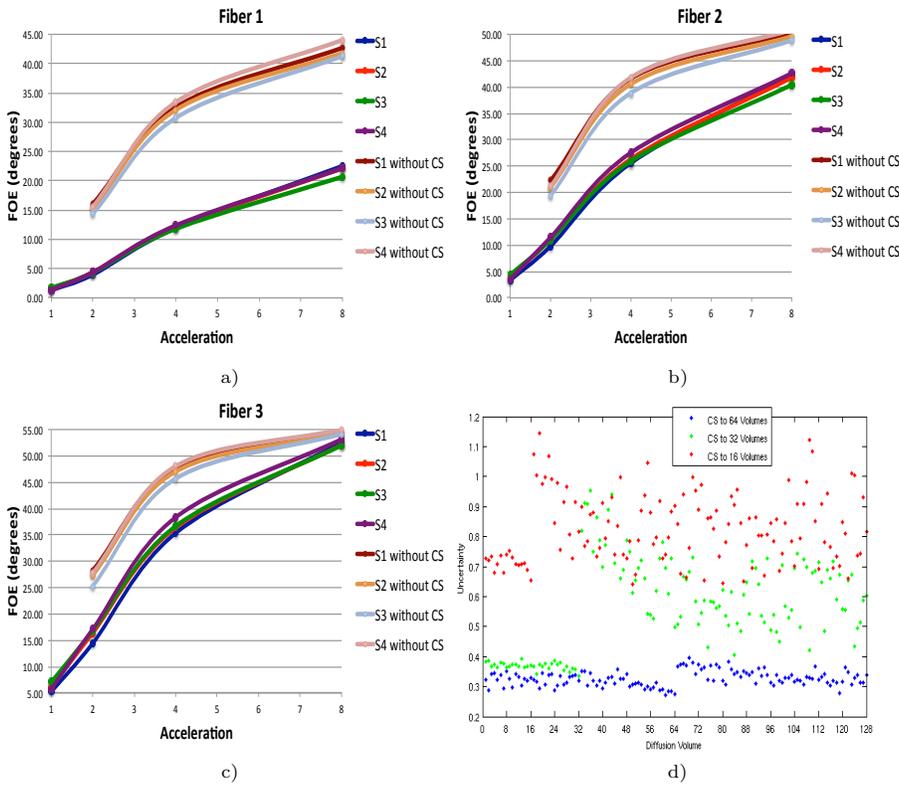
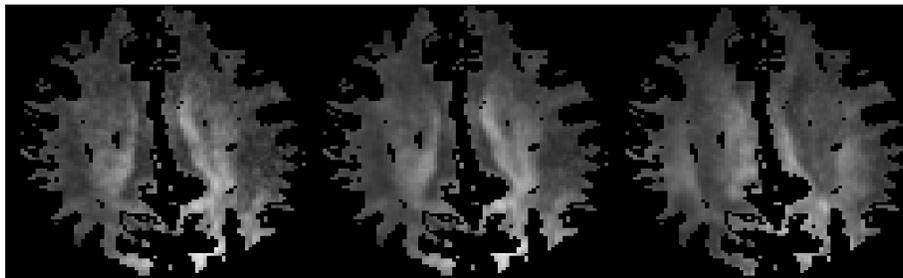


Fig. 6. Reconstruction error for the in-vivo DW-MRIs: a) FOE first peak; b) FOE second peak; c) FOE third peak; d) Uncertainty of the reconstruction.

celerations considered. Note that we have included an acceleration of one (i.e., no acceleration), the purpose of which is to measure the difference between the original and Bayesian “reconstructed” DW-MRIs without CS. This difference can be attributable only to approximation errors in the sparse representation of the DW-MRI and/or denoising, providing thus a baseline. In order to evaluate the importance of CS, we have also included the reconstruction errors, when using the first 64, 32, and 16 volumes, without using BCS to reconstruct the 128 gradient directions of the original DW-MRI. As indicated in figures 5 and 6, BCS does significantly reduces the reconstruction error compared to simple sub-sampling the DW-MRI. An acceleration of 8 using BCS can still produce good results if we are only interested in the strongest peak of the FODs. An acceleration of 4 using BCS is sufficient for resolving up to two fiber crossings. To recover three fiber crossings, an acceleration factor of at most two is required.



**Fig. 7.** a) Diffusion volume (72) of the original DW-MRI, masked with the FA mask; b) Diffusion volume (72) of the reconstructed DW-MRI with acceleration one; c) Closest diffusion volume (162) of the estimated DW-MRI with 256 gradient directions.

Figures 5b and 6 indicate that the DP-MT-BCS framework can be easily applied to various datasets without parameter tuning, since the  $D_{SH}$  and  $FOE$  are basically the same for all subjects and accelerations. The differences in the  $MSE$  (Figure 5a) might be due to different noise levels between acquisitions, differences in the denoising effect of the sparse representation, or errors in the representation itself. Figure 6d shows the uncertainty in the estimation for each acceleration (computed as indicated before) and diffusion volume. As before, the a priori uncertainty indicates that two is the best acceleration, which also agrees with the values presented in figures 5 and 6, where an acceleration of two allows the estimation of up to three fibers. Certainly, higher accelerations are possible, sacrificing some accuracy in the reconstruction. Figure 7a shows a slice of diffusion volume 72 of the original DW-MRI, while Figure 7b shows the corresponding (Bayesian) reconstructed volume 72, with no acceleration. As expected, the reconstructed DW-MRI looks cleaner and smoother. Figure 7c shows volume 162 of a DW-MRI constructed considering 256 gradient directions. The first 128 gradient directions are the same as the original DW-MRI, while the remaining 128 are extra gradient directions that complement the first 128 (on the same half sphere), obtained following [9]. Volume 162 in this new (256

gradient directions) DW-MRI is the closest gradient direction to volume 72 in the original DW-MRI. The new DW-MRI is obtained by assuming that the original DW-MRI was a CS version of a fictitious DW-MRI with 256 gradient directions, thus achieving higher angular resolution in q-space. As can be noted in Figure 7c, the new diffusion signal seems to provide information that is complementary to figures 7a and 7b.

## 4 Conclusion

We presented a robust and mathematically well-founded Bayesian compressive sensing framework to reduce the acquisition time of DW-MRIs. We can achieve acceleration factors of at least four, with minimum loss of information, reducing significantly the acquisition time of DW-MR data. Higher acceleration factors are certainly possible, given the quasi-linear relationship between the reconstruction error and acquisition acceleration found in the in-vivo data set. Additional work is required to model the Riccian noise in DW-MR data sets, which might further improve reconstruction accuracy. We plan also to consider multiple-shells in HARDI datasets.

## Acknowledgements

This work was partly supported by NIH grants R01 EB008432, R01 EB008645, S10 RR026783, P41 RR008079 (NCRR), P41 EB015894 (NIBIB), P30 NS057091, P30 NS0576408, the Human Connectome Project (U54 MH091657) from the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research, and the W.M. KECK Foundation. Additional support was received from DARPA and NSSEFF. The authors would like to thank Dr. Noam Harel for providing the human diffusion MRI data used in this paper.

## References

1. Basser, P. J., Jones, D. K.: Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR Biomed.* 15, 456–467 (2002)
2. Basser, P. J., Mattiello, J., Bihan, D. L.: Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson. B* 103, 247–254 (1994)
3. Behrens, T. E. J., Johansen-Berg, H., Jbabdi, S., Rushworth, M. F. S., Woolrich, M. W.: Probabilistic diffusion tractography with multiple fibre orientations. What can we gain? *Neuroimage* 34(1), 144–155 (2007)
4. Tuch, D.: Diffusion MRI of complex tissue structure. Ph.D. Thesis, Harvard University and Massachusetts Institute of Technology. (2002)
5. Callaghan, P.T., Eccles, C.D., Xia, Y.: Rapid Communication: NMR microscopy of dynamic displacements: k-space and q-space imaging. *J. Phys. E Scientific Instruments* 21, 820–822 (1988)
6. Wiegell, M., Larsson, H., Wedeen, V.J.: Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology* 217, 897–903 (2000)

7. Tuch, D.: Q-ball imaging. *Magn. Reson. Med.* 52(6), 1358–1372 (2004)
8. Wedeen, V. J., Wang, R. P., Schmahmann, J. D., Benner, T., Tseng, W. Y. I., Dai, G., Pandya, D. N., Hangmann, P., D’Arceuil, H., de Crespigny, A. J.: Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *Neuroimage* 41(4), 1267–1277 (2008)
9. Caruyer, E., Cheng, J., Lenglet, C., Sapiro, G., Jiang, T., Deriche, R.: Optimal design of multiple q-shells experiments for diffusion MRI. In: MICCAI Workshop on Computational Diffusion MRI (CDMRI’11), Toronto, Canada (2011)
10. Koay, C. G., Özarslan, E., Johnson, K. M., Meyerand, M. E.: Sparse and optimal acquisition design for diffusion MRI and beyond. *Medical Physics* 39(5), 2499–2511 (2012)
11. Setsompop, K., Gagoski, B. A., Polimeni, J. R., Witzel, T., Wedeen, V. J., Wald, L. L.: Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Mag. Reson. Med.* 67, 1210–1224 (2012)
12. Moeller, S., Auerbach, E., van de Moortele, P.-F., Adriany, G., Ugurbil, K.: fMRI with 16 fold reduction using multiband multislice sampling. In: *Proc. Int. Soc. Mag. Reson. Med.* 16, 2366 (2008)
13. Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., Ugurbil, 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.* 63, 1144–1153 (2010)
14. Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., Glasser, M. F., Miller, K. L., Ugurbil, K., Yacoub, E.: Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS ONE* 5, e15710 (2010)
15. Menzel, M. I., Tan, E. T., Khare, K., Sperl, J. I., King, K. F., Tao, X., Hardy, C. J., Marinelli, L.: Accelerated diffusion spectrum imaging in the human brain using compressed sensing. *Magn. Reson. Med.* 66(5), 1226–1233 (2011)
16. Block, K. T., Uecker, M., Frahm, J.: Undersampled Radial MRI with Multiple Coils. Iterative Image Reconstruction Using a Total Variation Constraint. *Magn. Reson. Med.* 57:1086–1098 (2007)
17. Lustig, M., Donoho, D., Pauly, J. M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* 58, 1882–1195 (2007)
18. Usman, M., Batchelor, P. G.: Optimized Sampling Patterns for Practical Compressed MRI, in SAMPTA, 2009.
19. Maleh, R.: Efficient sparse approximation methods for medical imaging. PhD Thesis, University of Michigan (2009)
20. Chartrand, R.: Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data. In: *IEEE Int. Symposium on Biomedical Imaging: From Nano to Macro (ISBI ’09)*, Boston, MA, US (2009)
21. Khare, K., Hardy, C. J., King, K. F., Turski, P. A., Marinelli, L.: Accelerated MR Imaging using compressive sensing with no free parameters. *Magn. Reson. Med.* 67(5), 1226–1233 (2011)
22. Seeger, M., Nickisch, H., Pohmann, R., Schölkopf, B.: Optimization of k-space trajectories for compressed sensing by bayesian experimental design. *Magn. Reson. Med.* 63(1), 116–126 (2010)
23. Murphy, M., Alley, M., Demmel, J., Keutzer, K., Vasanawala, S., Lustig, M.: Fast l1-SPIRiT compressed sensing parallel imaging MRI: Scalable parallel implementation and clinically feasible runtime. *IEEE Trans. Med. Imaging* (2012)

24. Vasanawala, S. S., Alley, M. T., Hargreaves, B. A., Barth, R. A., Pauly, J. M., Lustig, M.: Improved pediatric MR imaging with compressed sensing. *Radiology*, 256(2), 607–616 (2010)
25. Chang, C.-H., Ji, J.: Compressed sensing MRI with multichannel data using multicore processors. *Magn. Reson. Med.* 64(4), 1135–1139 (2010)
26. Smith, D. S., Gore, J. C., Yankeelov, T. E., Welch, E. B.: Real-time compressive sensing MRI reconstruction using GPU computing and split Bregman methods. *Int. J. Biomed. Imaging*, vol. 2012, Article ID 864827, 6 p. (2012)
27. Bilgic, B., Goyal, V. K., Adalsteinsson, E. Multi-contrast reconstruction with bayesian compressed Sensing. *Magn. Reson. Med.* 66(6):1601–1615 (2011)
28. Jung, H., Sung, K., Nayak, K. S., Kim, E. Y., Ye, J. C.: k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI. *Magn. Reson. Med.* 61(1), 103–116 (2009)
29. Jung, H., Ye, J. C.: Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video. *Int. J. Imag. Syst. Tech.* 20(2), 81-98 (2010)
30. Hsiao, A., Lustig, M., Alley, M. T., Murphy, M., Chan, F. P., Herfkens, R. J., Vasanawala, S. S.: Rapid Pediatric Cardiac Assessment of Flow and Ventricular Volume With Compressed Sensing Parallel Imaging Volumetric Cine Phase-Contrast MRI. *AJR Am. J. Roentgenol.* 198(3), W250–259 (2012)
31. Michailovich, O. V., Rathi, Y., Dolui, S.: Spatially regularized compressed sensing for high angular resolution diffusion imaging. *IEEE Trans. Med. Imaging* 30(5), 1100–1115 (2011)
32. Rathi, Y., Michailovich, O. V., Setsompop, K., Bouix, S., Shenton, M. E., Westin, C.-F.: Sparse multi-shell diffusion imaging. In: *MICCAI Proc. 14th Int. Conf. on medical image computing and computer-assisted intervention - Part II (CDMRI'11)*, Toronto, Canada (2011)
33. Dolui, S., Kuurstra, A., Michailovich, O. V.: Rician compressed sensing for fast and stable signal reconstruction in diffusion MRI. In: *Proc. SPIE 8314 San Diego, CA, US* (2012)
34. Ji, S., Dunson, D., Carin, L.: Multi-task compressive sensing. *IEEE Trans. Signal Processing* 57(1), 92–106 (2009)
35. Qi, Y., Liu, D., Dunson, D., Carin, L.: Bayesian multi-task compressive sensing with Dirichlet priors. In: *Proc. IEEE Int. Conf. Machine Learning (ICML)*, Helsinki, Finland (2008)
36. Gudbjartsson, H., Patz, S.: The Rician distribution of noisy MRI data. *Mag. Reson. Med.* 34(6), 910–914 (1995)
37. Michailovich, O. V., Rathi, Y.: On approximation of orientation distributions by means of spherical ridgelets. *IEEE Trans. Image Process.* 19(2), 461–477 (2010)
38. Sethuraman, J.: A constructive definition of the Dirichlet prior. *Statistica Sinica*, 2, 639–650 (1994)
39. Lenglet, C., Abosch, A., Yacoub, E., De Martino, F., Sapiro, G., Harel, N.: Comprehensive in vivo Mapping of the Human Basal Ganglia and Thalamic Connectome in Individuals Using 7T MRI. *PLoS ONE* 7(1), e29153 (2012)
40. Tournier, J. D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35 (4), 1459–1472 (2007)
41. Tournier, J. D., Yeh, C.-H., Calamante, F., Chod, K.-H., Connelly, A., Lin, C.-P.: Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *Neuroimage* 42(2), 617–625 (2008)

# Introduction of a Quantitative Evaluation Method for White Matter Tractography using a HARDI-based Reference

Peter F. Neher<sup>1</sup>, Bram Stieltjes<sup>2</sup>, Hans-Peter Meinzer<sup>1</sup>, and  
Klaus H. Fritzsche<sup>1,2,\*</sup>

<sup>1</sup> German Cancer Research Center (DKFZ), Medical and Biological Informatics,  
Heidelberg, Germany

<sup>2</sup> German Cancer Research Center (DKFZ), Quantitative Image-based Disease  
Characterization, Heidelberg, Germany

**Abstract.** The diffusion tensor imaging (DTI) Tractography Challenge MICCAI 2011 featured a review board of highly recognized medical experts and prominent research groups working in the field of tractography. The workshop revealed the limitations of current tractography algorithms in crossing fiber regions when clinical image data quality is provided and spotlighted a critical need for objective and quantitative evaluation methods to stimulate a clinically relevant advancement in the field. Here we address this issue by proposing a reference-based validation procedure that allows for objective and quantitative assessment of tractography results obtained from DTI data of clinical quality. The method employs high angular resolution diffusion images (HARDI) as reference datasets and is evaluated by applying it to two well-known tractography algorithms. Their different capabilities in the reconstruction of crossing fiber regions were successfully revealed by the validation procedure. We conclude the paper with an outlook on open access of data and public validation opportunities.

**Keywords:** Tractography Evaluation, Fiber Tracking, Diffusion-weighted Imaging, Diffusion Tensor Imaging, HARDI, Reference-based Validation Procedure

## 1 Introduction

Diffusion based tractography noninvasively provides insight into the course of white matter pathways in the living human brain. However, good tract representation in research settings with high quality data is strongly hampered by a reduced image quality in a clinical setting. This issue was confirmed during the DTI Tractography Challenge MICCAI 2011 where a neurosurgically highly relevant task had to be solved on clinical data [1]. According to the workshops medical review board, none of the participants could satisfactorily solve the task

---

\* Corresponding author. E-mail address: k.fritzsche@dkfz-heidelberg.de.

of reconstructing the corticospinal tract, one of the most prominent tracts in the brain. More than one third of all image voxels contain crossing fibers [2] and thus challenge tensor-based tractography with regions of uncertainty that have to be resolved by some kind of intelligent mechanism or apriori knowledge. We believe, also motivated by the success story of the organ segmentation challenges at MICCAI, that the key aspect in making a significant step forward in the field of diffusion tractography research will be the development of quantitative tools for automatic evaluation of these mechanisms.

Most current validation techniques are either based on synthetic [3] or physical [4] phantoms that can only represent limited aspects of *in-vivo* data. Existing approaches for quantitative tractography evaluation on real patient data have so far been limited to the evaluation of reproducibility, e.g. by bootstrap analysis [5], or the comparison of fiber envelopes by the STAPLE algorithm [6]. Evaluation by medical experts suffers from a high inter and intra rater variability. Due to the complexity and diversity of fiber bundles in the brain, manually defined gold standards are difficult to establish and not commonly used in tractography. An optimal evaluation method would be automated and generically applicable to all fiber structures in the brain, yielding quantitative measures of reproducibility, validity and integrity of tracking algorithms especially in complex fiber configurations.

While many of the advanced diffusion acquisition and modeling techniques like high angular resolution diffusion imaging (HARDI) are not applicable in clinical settings due to extensive acquisition and processing times, we believe that they can provide important means for the validation of tractography algorithms. We propose a reference-based validation procedure that uses HARDI data as basis of comparison for tractography obtained from DTI. The method proposed here follows the validation framework for medical image processing introduced by Jannin *et al.* [7, 8] and is particularly useful for the evaluation of tractography results in crossing regions.

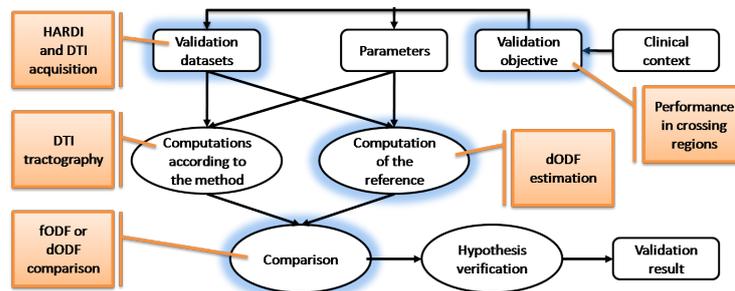
## 2 Methods

The proposed reference-based validation framework is depicted in Fig. 1 and described in more detail in section 2.1. In our experiments the procedure was applied to two well known tractography algorithms (section 2.2).

### 2.1 Validation framework

**Validation objective:** The objective is the validation of tractography algorithms with a focus on their ability to correctly reconstruct crossing regions in the brain from DTI datasets that do not explicitly model multiple directions per voxel.

**Validation datasets:** The validation dataset consists of HARDI and corresponding DTI datasets. The DTI data can either be acquired in the same session as the HARDI dataset or it can be subsampled from the HARDI dataset

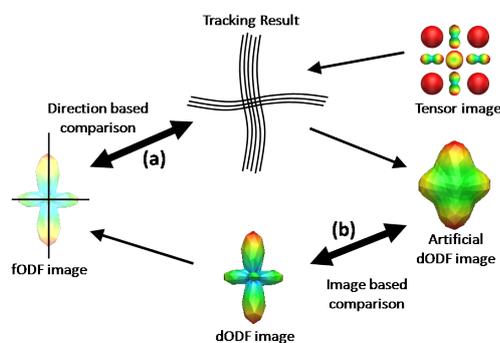


**Fig. 1.** Schematic depiction of the proposed reference-based validation procedures for quantitative tractography evaluation. The main contributions of this work are highlighted in blue. Adapted from [8].

using the desired b-value, spatial resolution, number of gradient directions and gradient distribution scheme.

**Computation of the reference:** In order to obtain a reference, diffusion orientation distribution functions (dODF) are reconstructed from the HARDI dataset [9].

**Comparison:** We propose two options to compare the DTI tractography to the reference: an analysis of the voxel-wise fiber directions (fODF) as illustrated in Fig. 2a or a comparison based on the dODF images as shown in Fig. 2b.



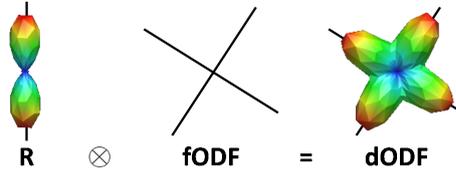
**Fig. 2.** Illustration of the comparison of the tractography result with the reference image based on directions (a) and dODF images (b).

*Direction-based comparison:* For a direction-based comparison, fODFs have to be calculated from the reference image as well as from the tractography result.

The relation between dODF, R and fODF is illustrated in Fig. 3. The fODFs are calculated from the reference image using spherical deconvolution as described by Tournier *et al.* [10]. The deconvolution kernel R is estimated by aligning and averaging the most anisotropic dODFs in the image. The fODFs of the fiber tractography result at a spatial position  $x$  are calculated by averaging the fiber fragments in the neighborhood of  $x$  that is defined by an isotropic Gaussian with variance  $\sigma$ :

$$\text{fODF}_{\text{trac}}(x, n) = \frac{1}{\sqrt{2\pi}} \sum_{f \in \mathbb{F}} \int e^{-\frac{1}{2} \frac{|x-x'|^2}{\sigma^2}} \delta(t(x') - n) dx' , \quad (1)$$

where  $\delta$  is the Dirac delta function,  $\mathbb{F}$  is the set of all fibers and  $t(x)$  is the fiber tangent at position  $x$  on fiber  $f$ . An angular threshold is used to combine clusters of fODF maxima that indicate a common underlying fiber direction.



**Fig. 3.** Illustration of the relation between the kernel R, fODF and dODF. Adapted from [11].

Several measures can be evaluated for comparisons based on the fODF directions:

1. Sensitivity and specificity of the detection of crossing fiber voxels.
2. Root-mean-square (RMS) error of the number of detected directions per voxel.
3. RMS angular error of detected direction angles.
4. Visual inspection.

*Image-based comparison:* For an image-based comparison, an artificial dODF image has to be calculated from the tracking result. This can be achieved by a convolution of the fODF (c.f. equation 1) with the convolution kernel R as illustrated in Fig. 3. The resulting dODF images can either be compared directly or via derived quantities like the generalized fractional anisotropy:

1. RMS error of derived quantities or of the dODFs.
2. Jensen-Shannon divergence [12] of the dODFs.
3. Distance measures based on euclidean metrics [13] of the dODFs.
4. Visual inspection.

## 2.2 Experiments

The images used in our experiments were obtained using a single-shot EPI sequence on a 3T MR scanner, 64 different gradient directions with three repetitions, 2.5 mm isotropic resolution and two shells at b-values of 1000 mm/s<sup>2</sup> and 3500 mm/s<sup>2</sup>. All images were corrected for head motion and eddycurrent effects using an affine registration to the first baseline image employing FSL-FLIRT ([www.fmrib.ox.ac.uk/fsl/flirt/](http://www.fmrib.ox.ac.uk/fsl/flirt/)). The gradient directions were corrected according to these transformations. The 192 images at b=3500 mm/s<sup>2</sup> were used as HARDI reference image. A DTI image of clinical quality was extracted from the b=1000 mm/s<sup>2</sup> images by selecting 30 unique gradient directions.

The validation procedure was tested on a local streamlining approach based on the well known FACT algorithm [14] and a global probabilistic approach called Gibbs tracking [15]. The local algorithm was automatically seeded in every voxel and the FA stopping criterion was set to 0.2. The Gibbs tracking was run with 10<sup>8</sup> iterations. These two tracking algorithms were chosen to illustrate the capability of our method to detect the well known limitations of FACT to resolve crossing fiber configurations as well as the expected better performance of a global approach in this situation that was already shown during MICCAI’s Fiber Cup 2009 [16].

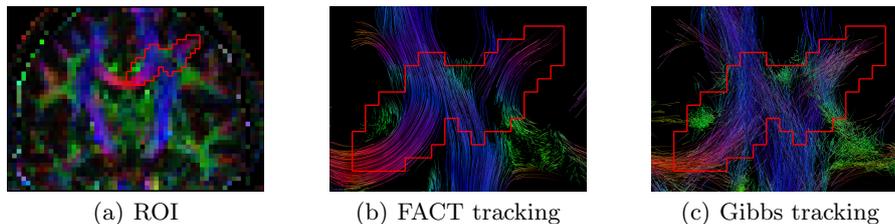
We illustrate the direction-based part of our validation approach (c.f. Fig. 2a) by calculating the sensitivity and specificity of the algorithms’ capabilities to detect voxels containing crossing fibers and the image-based part (c.f. Fig. 2b) by generating an artificial dODF image for each tracking result and comparing it visually to the reference. In all experiments we focused on one representative crossing of the corpus callosum and the corona radiata. This area was manually labelled by a radiologist.

## 3 Results

Fig. 4 shows both fiber tractography results in the region of interest. Both algorithms show a high specificity in detecting crossing voxels (Tab. 1). The low sensitivity of the FACT algorithm shows that almost no fiber crossings could be resolved. The global Gibbs tracking was capable of resolving about 26% of all crossing voxels. The most intuitive way to evaluate the tracking result in the image domain is by visual inspection. The different dODF images are depicted in Fig. 5 and are discussed in the next section.

Algorithm	Sensitivity	Specificity
Gibbs	26%	86%
FACT	3%	97%

**Table 1.** Sensitivity and specificity of the two tracking algorithms to detect voxels containing crossing fibers.



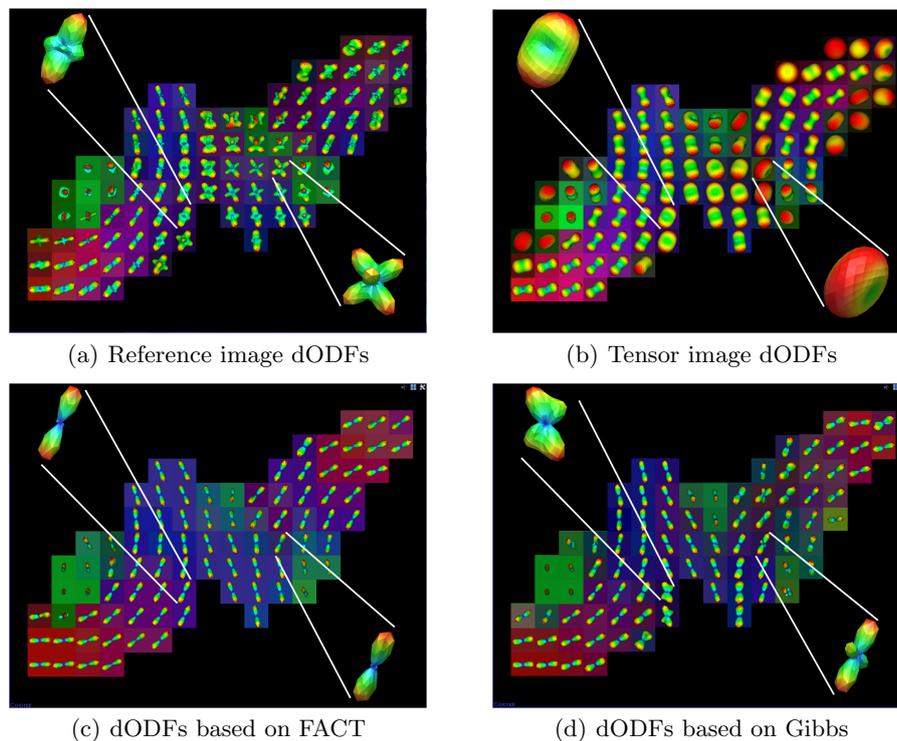
**Fig. 4.** Reconstructed tensor image overlaid with the region of interest (red) containing the crossing of corpus callosum and corona radiata (a) and the tractography results of FACT (b) and Gibbs (c) tracking in the selected ROI.

## 4 Discussion

This paper proposes a reference-based approach to the automatic evaluation of fiber tracking algorithms using high quality HARDI data as basis of comparison for results generated from DTI data of clinical quality. The proposed method can be used to evaluate several different measures of tracking accuracy and integrity. It is especially valuable for the examination of the algorithms' performance in crossing regions and other situations that cannot be resolved by clinical DTI data. To our knowledge this is the first approach to quantitatively evaluate the performance of fiber tracking algorithms on clinical DTI data with a HARDI reference yielding objective and easy to interpret results like sensitivity and specificity.

The method was demonstrated by comparing a local streamlining approach based on the well known FACT algorithm to a global tractography method called Gibbs tracking. Their different capabilities in the reconstruction of crossing fiber regions were successfully revealed by the validation procedure. The crossing sensitivity of the streamline approach shows that almost no crossing fiber was detected. The global approach shows a somewhat higher detection rate. These findings are fortified by the visual analysis of the fiber tracts and visual comparison of dODF images and are consistent with the experiences made during the DTI Tractography Challenge MICCAI 2011.

While the presented evaluation procedure and corresponding results are very promising, there is still room for improvement considering the current analysis pipeline: The reconstruction of dODFs and fODFs from the HARDI acquisition would potentially be more accurate when employing a multi-shell q-ball reconstruction scheme instead of selecting only the higher q-shell as reference image. The analytical calculation of dODF maxima from the spherical harmonic coefficients should also be considered as an alternative to the deconvolution of the signal. In future experiments we will start incorporating more regions of the brain and consider the complete set of proposed indices of comparison in our experiments. Finally, one important limitation should be noted: the proposed validation scheme is, of course, limited by the quality and potential ambiguities



**Fig. 5.** Reference (a) and DTI (b) images in comparison with the artificial dODF images generated from the two tractography results (c and d).

of the reference images (e.g. if crossing and kissing fiber configurations should be distinguished).

Besides the further exploration of the methods proposed here we are planning to achieve a general improvement of the evaluation and comparability of tracking methods developed by groups around the world. Following the model of the Fiber Cup where phantom datasets and access to evaluation methods was made publicly available, we plan to establish a similar framework for the evaluation of *in-vivo* datasets. The first step will be the acquisition of multiple high quality datasets from different subjects and the construction of a large image database containing "quality pyramids" of test images with multiple levels of spatial resolution, number of gradient directions and number of q-shells. It is planned to provide open access to this database, allowing scientists in the whole community to test their tracking algorithms on common datasets and to submit their results for evaluation.

## References

1. Pujol, S., Kikinis, R., Golby, A., Gerig, G., Styner, M., Wells, W., Westin, C.F., Gouttard, S. In: DTI Tractography for Neurosurgical Planning: A Grand Challenge, MICCAI (September 2011)
2. Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* **34**(1) (Jan 2007) 144–155
3. Delputte, S., Fieremans, E., Dedeene, Y., D’Asseler, Y., Achten, R., Lemahieu, I., de Walle, R.V.: Quantitative validation of white matter fiber tractography by use of an anatomically realistic synthetic diffusion tensor phantom. In: Proc. Intl. Soc. Mag. Reson. Med. 14. (2006)
4. Moussavi-Biugui, A., Stieltjes, B., Fritzsche, K., Semmler, W., Laun, F.B.: Novel spherical phantoms for q-ball imaging under in vivo conditions. *Magnetic Resonance in Medicine* **65** (2011) 190–194
5. Jones, D.K., Pierpaoli, C.: Confidence mapping in diffusion tensor magnetic resonance imaging tractography using a bootstrap approach. *Magnetic Resonance in Medicine* **53** (2005) 1143–1149
6. Pujol, S., Westin, C.F., Whitaker, R., Gerig, G., Fletcher, T., Magnotta, V., Bouix, S., Kikinis, R., III, W.M.W., Gollub, R.: Preliminary results on the use of staple for evaluating dt-mri tractography in the absence of ground truth. In: Proc. Intl. Soc. Mag. Reson. Med. 17. (2009)
7. Jannin, P., Fitzpatrick, J., Hawkes, D., Pennec, X., Shahidi, R., Vannier, M.: Validation of medical image processing in image-guided therapy. *IEEE Trans Med Imaging* **21**(12) (2002) 1445–1409
8. Jannin, P., Grova, C., Maurer, C.: Model for defining and reporting reference-based validation protocols in medical image processing. *Int J CARS* **1** (2006) 63–73 10.1007/s11548-006-0044-6.
9. Aganj, I., Lenglet, C., Sapiro, G.: Odf reconstruction in q-ball imaging with solid angle consideration. In: Sixth IEEE International Symposium on Biomedical Imaging. (2009)
10. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage* **23** (2004) 1176–1185
11. Lenglet, C., Campbell, J., Descoteaux, M., Haro, G., Savadjiev, P., Wassermann, D., Anwander, A., Deriche, R., Pike, G., Sapiro, G., Siddiqi, K., Thompson, P.: Mathematical methods for diffusion mri processing. *NeuroImage* **45** (2009) 111–122
12. Cohen-Adad, J., Descoteaux, M., Wald, L.L.: Quality assessment of high angular resolution diffusion imaging data using bootstrap on q-ball reconstruction. *Journal of Magnetic Resonance Imaging* **33** (2011) 1194–1208
13. Pasternak, O., Sochen, N., Basser, P.J.: The effect of metric selection on the analysis of diffusion tensor mri data. *NeuroImage* **49** (2010) 2190–2204
14. Fillard, P., Gerig, G.: Analysis tool for diffusion tensor mri. In: In Proc. of Medical Image Computing and Computer-Assisted Intervention, Springer (2003) 967–968
15. Reisert, M., Mader, I., Anastasopoulos, C., Weigel, M., Schnell, S., Kiselev, V.: Global fiber reconstruction becomes practical. *Neuroimage* **54** (2011) 955–962
16. Fillard, P., Descoteaux, M., Goh, A., Gouttard, S., Jeurissen, B., Malcolm, J., Ramirez-Manzanares, A., Reisert, M., Sakaie, K., Tensaouti, F., Yo, T., Mangin, J.F., Poupon, C.: Quantitative evaluation of 10 tractography algorithms on a realistic diffusion mr phantom. *Neuroimage* **56**(1) (May 2011) 220–234

# Constrained Diffusion Kurtosis Imaging Using Ternary Quartics & MLE

Tristan Milne, Aurobrata Ghosh, and Rachid Deriche

Project Team Athena, INRIA Sophia Antipolis - Méditerranée, France

**Abstract.** We present a ternary quartic based approach with an improved gradient based optimization scheme for diffusion kurtosis imaging to estimate constrained and physically realistic diffusion and kurtosis tensors. We account for the signal noise by considering a maximum likelihood estimation based on the Rician noise model. Diffusion kurtosis imaging (DKI) is a recent important improvement over the diffusion tensor imaging (DTI) model that quantifies the degree of non-Gaussian diffusion in a tissue using a 3D 4<sup>th</sup> order tensor. However, DKI estimation needs to consider three constraints to be physically relevant. By adopting the implicit ternary quartic parameterization which allows to elegantly impose a positivity constraint on the kurtosis tensor and by employing gradient based optimization schemes, we show dramatically improved performance in terms of estimation time and quality. We derive the mathematical framework and show results on extensive synthetic data experiments.

**Keywords:** Diffusion Kurtosis Imaging, Ternary Quartics, Maximum Likelihood Estimation, Rician noise, Constrained Optimization

## 1 Introduction

Diffusion Kurtosis Imaging (DKI) [1, 2] is a recent extension to the diffusion tensor imaging (DTI) [3] signal model, under the q-space formalism [4], that attempts to characterize the underlying tissue by quantifying the amount of non-Gaussian diffusion. DTI estimates a 3D  $2^{nd}$  order diffusion tensor  $\mathbf{D}$  which quantifies diffusion anisotropy. It is a widely used model in diffusion MRI (dMRI) since  $\mathbf{D}$  and its scalar indices (“biomarkers”) have proven useful for inferring the white-matter’s microstructure and for discerning changes in the microstructure’s integrity. However, DTI makes the strong simplifying assumption that the diffusion is Gaussian, whereas, it has been argued that diffusion of water molecules in structurally rich biological tissues, like the brain, would deviate from Gaussianity due to the complexity of the underlying microstructure.

The q-space formalism, therefore, considers a generic unknown distribution –  $P$ , the ensemble average propagator (EAP), usually non-Gaussian – to describe the diffusion phenomenon in tissues with complex microstructures [4]. The EAP is related to the diffusion signal from the Stejskal-Tanner pulsed gradient spin echo (PGSE) experiment by the Fourier transform  $E(\mathbf{q}, \Delta) = S(\mathbf{q}, \Delta)/S_0 =$

$\int_{\mathbb{R}^3} P(\mathbf{r}, \Delta) e^{i2\pi\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}$ , where  $\Delta$  is the diffusion time (time between the gradient pulses) and  $\mathbf{q} = \gamma\delta\tilde{\mathbf{g}}/(2\pi)$ , with  $\gamma$  the gyromagnetic ratio,  $\delta$  the pulse application time and  $\tilde{\mathbf{g}}$  the gradient vector such that  $\delta\tilde{\mathbf{g}}$  is finite ( $\delta \ll \Delta$ ),  $\|\tilde{\mathbf{g}}\| = g$  and  $\mathbf{g} = \tilde{\mathbf{g}}/g$ . Since  $P$  is real and symmetric,  $E$  is also its characteristic function and the Taylor expansion of  $\ln(E)$ , with respect to  $\mathbf{q}$ , reveals the second cumulant of  $P$  to be proportional to  $\mathbf{D}$ . The succeeding (even ordered) cumulants characterize  $P$  and in particular measure  $P$ 's deviation from Gaussianity.

Based on this idea, DKI, considers the fourth order expansion of the diffusion signal, to estimate a 3D 4<sup>th</sup> order tensor  $\mathcal{W}$ , the kurtosis tensor, in addition to  $\mathbf{D}$ . The kurtosis, or the *excess kurtosis*, measures the ‘‘peakedness’’ of the EAP to quantify its deviation from Gaussianity. It has been proposed that in cerebral tissues the complex microstructure comprised of cell membranes, boundaries, various intracellular and extracellular compartments with and without exchange can substantially affect the diffusion process resulting in non-Gaussian diffusion, which can be quantified by measuring the kurtosis [1, 2]. DKI based scalar indices have, therefore, gained popularity for characterizing cerebral tissue and have shown promising sensitivity at discerning microstructural changes due to development and disease (references in [5, 6]).

Diffusion being a physical phenomenon, the tensors  $\mathbf{D}$  and  $\mathcal{W}$  have to fulfill conditions to be physically relevant (see Sec. 2). (1)  $\mathbf{D}$  should be positive definite. (2a) The excess kurtosis along any direction,  $K_{app}$ , has a lower bound – theoretically  $-2$  – but *platykurtosis* ( $K_{app} < 0$ ) is unlikely in biological tissue, where due to the complex microstructure *leptokurtosis* ( $K_{app} > 0$ ) is more likely [1, 5, 6]. (2b) However, although unlikely the lower bound in biological tissue is in theory  $-3/7$  [1]. (3)  $K_{app}$  also has an upper bound which ensures the convergence of the Taylor expansion of the DKI model. However, these constraints are often violated by an unconstrained estimator because the signal is corrupted by noise. Therefore, to accurately estimate physically meaningful quantities it is important to estimate the tensors with the constraints. It is also important to consider the noise model, which is known to be Rician, since it is crucial at low signal-to-noise-ratios (SNRs) [6].

In [7] the authors proposed a rather complex iterative minimization scheme to estimate  $\mathbf{D}$  and  $\mathcal{W}$  with the constraints (1) & (2b). However, instead of constraint (3) they considered  $E(\mathbf{q}) \geq 0$  on each q-shell and ignored the Rician noise-model. In [8] the authors only considered constraints (1) & (2a). These were imposed by a Cholesky parameterization of  $\mathbf{D}$  and  $\mathbf{W}$ , the  $6 \times 6$  matrix representation of  $\mathcal{W}$ . However, this makes  $\mathbf{W}$  positive definite over  $\mathbb{R}^6$ , which is a much stronger constraint than (2a) [9]. Further, they too ignored the Rician noise model. A convex quadratic programming approach was presented in [5], where constraints (1), (2a) & (3) were all applied, but only along the discrete signal acquisition directions. Although this ensured a unique minimizer, the constraints were applied only along finite directions and the approach also didn't account for the Rician noise. [6] presented an approach that considered all constraints (1), (2a) & (3) and incorporated the Rician noise model by a maximum likelihood estimation (MLE). (1) was applied by the Cholesky parameterization of  $\mathbf{D}$  and

(2a) & (3) were applied along discrete directions as discontinuous penalties. A Nelder-Mead scheme was used to solve the MLE. However, the heuristic Nelder-Mead scheme doesn't use the function gradient, therefore, although appropriate for discontinuous functions, it suffers from slow convergence rates.

In this paper, we propose a twofold improvement over the work in [6]. Our first contribution is in the ternary quartic (TQ) parameterization of  $\mathcal{W}$ , which is the correct approach for parameterizing a non-negative 3D 4<sup>th</sup> order tensor [10, 9, 11]. The TQ approach implicitly imposes (2a), which both speeds up computations and ensures non-negativity for *all* directions. Our second contribution is in a gradient based (MLE) optimization scheme, which dramatically improves both computation time and quality of the two tensors over the Nelder-Mead approach. This is possible since we don't impose discontinuous penalties for the constraints. Overall we consider both the Rician noise model, by adopting the MLE approach, and all three constraints (1), (2a) & (3). We validate the robustness and speed of our DKI estimation on extensive synthetic data experiments.

## 2 Background

First we briefly resume the DKI model and the constraints. Since  $E$  is the characteristic function of the EAP,  $\ln(E)$  is its cumulant generating function. Therefore in the Taylor expansion of  $\ln(E)$

$$\text{1D : } \ln(E(q)) = \sum_{n=1}^{\infty} \kappa_n \frac{(i\gamma\delta g)^n}{n!}, \quad (1)$$

$$\text{3D : } \ln(E(\mathbf{q})) = \sum_{n=1}^{\infty} K_{j_1 j_2 \dots j_n}^{(n)} g_{j_1} g_{j_2} \dots g_{j_n} \frac{(i\gamma\delta g)^n}{n!}, \quad (2)$$

$\kappa_n$  represent  $n^{\text{th}}$  order cumulants and  $K_{j_1 j_2 \dots j_n}^{(n)}$  represent the coefficients of the symmetric  $n^{\text{th}}$  order cumulant tensors  $\mathcal{K}^{(n)}$ . The repeated indices in the 3D formula indicate summation over the indices and  $g_{j_k}$  are components of  $\mathbf{g}$ . Further, since  $P$  is symmetric, all odd order cumulant (tensors) vanish. The expansion truncated at the fourth order is

$$\text{1D : } \ln(E(q)) = -\frac{(\gamma\delta g)^2}{2!} \kappa_2 + \frac{(\gamma\delta g)^4}{4!} \kappa_4, \quad (3)$$

$$\text{3D : } \ln(E(\mathbf{q})) = -(\gamma\delta g)^2 K_{j_1 j_2}^{(2)} g_{j_1} g_{j_2} / 2! + (\gamma\delta g)^4 K_{j_1 j_2 j_3 j_4}^{(4)} g_{j_1} g_{j_2} g_{j_3} g_{j_4} / 4!, \quad (4)$$

where  $\kappa_2 = 2\Delta D$ , with  $D$  the diffusion coefficient and  $\mathcal{K}^{(2)} = 2\Delta \mathbf{D}$ . Excess kurtosis, which measures the departure from Gaussianity, is defined as  $K = \kappa_4 / \kappa_2^2$  in 1D and defining  $b \approx \gamma^2 \delta^2 g^2 \Delta$ , the 1D DKI signal can be written as  $E(q) = e^{-bD + (b^2 D^2 K)/6}$ . Similarly in 3D, the excess kurtosis,  $K_{app}$ , along any direction  $\mathbf{g}$  can be defined as  $K_{app}(\mathbf{g}) = (K_{j_1 j_2 j_3 j_4}^{(4)} g_{j_1} g_{j_2} g_{j_3} g_{j_4}) / (K_{j_1 j_2}^{(2)} g_{j_1} g_{j_2})^2$  and the 3D DKI signal turns out to be [1]

$$E(b, \mathbf{g}) = \exp \left( -b \cdot D_{j_1 j_2}^{(2)} g_{j_1} g_{j_2} + b^2 \cdot (\text{tr}(\mathbf{D})/3)^2 \cdot W_{j_1 j_2 j_3 j_4}^{(4)} g_{j_1} g_{j_2} g_{j_3} g_{j_4} / 6 \right), \quad (5)$$

where the kurtosis tensor  $\mathcal{W}$  is defined from  $K_{app}(\mathbf{g}) = (K_{j_1 j_2 j_3 j_4}^{(4)} g_{j_1} g_{j_2} g_{j_3} g_{j_4}) / (K_{j_1 j_2}^{(2)} g_{j_1} g_{j_2})^2 = ((tr(\mathbf{D})/3)^2 W_{j_1 j_2 j_3 j_4} g_{j_1} g_{j_2} g_{j_3} g_{j_4}) / (D_{j_1 j_2} g_{j_1} g_{j_2})^2$ . Estimating  $\mathbf{D}$  and  $\mathcal{W}$  instead of  $\mathcal{K}^{(2)}$  and  $\mathcal{K}^{(4)}$  allows to work only with the  $b$ -value instead of the acquisition parameters  $\Delta$ ,  $\delta$  &  $\tilde{\mathbf{g}}$ . We also define  $\hat{\mathcal{K}} = (tr(\mathbf{D})/3)^2 \mathcal{W}$ . Replacing  $\mathcal{W}$  by  $\hat{\mathcal{K}}$  in Eq. (5) makes it log-linear in the coefficients of  $\mathbf{D}$  and  $\hat{\mathcal{K}}$  [5]

$$E(b, \mathbf{g}) = \exp\left(-b \cdot D_{j_1 j_2}^{(2)} g_{j_1} g_{j_2} + b^2 \cdot \hat{\mathcal{K}}_{j_1 j_2 j_3 j_4} g_{j_1} g_{j_2} g_{j_3} g_{j_4} / 6\right). \quad (6)$$

Both  $\mathbf{D}$  and  $\mathcal{W}$  (or  $\hat{\mathcal{K}}$ ) are symmetric tensors, therefore,  $\mathbf{D}$  has 6 unique coefficients and  $\mathcal{W}$  (or  $\hat{\mathcal{K}}$ ) has 15 unique coefficients.

We consider the following constraints to ensure that the estimated tensors  $\mathbf{D}$  and  $\mathcal{W}$  (or  $\hat{\mathcal{K}}$ ) are physically relevant.

1.  $\mathbf{D}$  is positive definite since it encodes the apparent diffusion coefficient ( $D_{app}(\mathbf{g}) = D_{j_1 j_2} g_{j_1} g_{j_2} = \mathbf{g}^T \mathbf{D} \mathbf{g}$ ) and negative diffusion is non-physical.
2.  $K_{app}(\mathbf{g}) \geq 0 \forall \mathbf{g} \in S^2$  – we consider the EAP to be either Gaussian ( $K_{app} \equiv 0$ ) or leptokurtic ( $K_{app} > 0$ ), which is the expected behaviour of diffusion in structurally rich cerebral tissues.
3.  $K_{app}(\mathbf{g}_i) \leq 3/(b \cdot D_{app}(\mathbf{g}_i)) \forall \mathbf{g}_i \in \mathbf{N}$ , where  $\mathbf{N}$  represents the set of signal acquisition directions. This constraint was derived to find an upper bound for the  $b$ -value that ensured the convergence of the Taylor expansion [2], but it can also serve as an upper bound on  $K_{app}$  when  $b < 3000\text{s}/\text{mm}^2$ .

Note that (#2) is for *all* directions, since we want to recover non-negative kurtosis along *all* directions. This is much stronger than the discrete constraints in [5, 6]. In comparison (#3) is only taken along the acquisition directions to ensure that the estimation results in convergent DKI signals along the data directions.

### 3 Materials and Methods

Here we first derive the formulations for imposing the constraints. Then we present a constrained weighted least squares (CWLS) approach for estimating the DKI tensors without the noise model. Finally we present a constrained MLE (CMLE) approach with the Rician noise model to account for acquisition noise. Although we only derive the gradients for the CWLS, which enables us to use a gradient based optimization scheme, computing the CMLE gradients is straightforward. We consider Eq. (6) as the DKI signal since it can be log-linearized. It is also simple to compute  $K_{app}$  and  $\mathcal{W}$  from  $\hat{\mathcal{K}}$  whenever necessary (Sec. 2).

The constraints of Sec. 2 are imposed as follows.

1. We consider the Cholesky parameterization [6],  $\mathbf{D} = \mathbf{L}\mathbf{L}^T$ , where  $\mathbf{L}$  is lower triangular. Therefore,  $D_{app}(\mathbf{g}) = \mathbf{g}^T \mathbf{L}\mathbf{L}^T \mathbf{g} = tr(\mathbf{G}\mathbf{L}\mathbf{L}^T)$ , where  $\mathbf{G} = \mathbf{g}\mathbf{g}^T$ . From [12],  $\nabla_{\mathbf{L}} D_{app}(\mathbf{g})$ , the gradient is equal to the six elements of the lower triangular matrix  $\frac{\partial}{\partial \mathbf{L}} D_{app}(\mathbf{g}) = 2\mathbf{G}\mathbf{L}$ .

2. In terms of  $\hat{\mathcal{K}}$ ,  $K_{app}(\mathbf{g}) = (\hat{K}_{j_1 j_2 j_3 j_4} g_{j_1} g_{j_2} g_{j_3} g_{j_4}) / (D_{app}(\mathbf{g}))^2$ , therefore,  $K_{app}(\mathbf{g}) \geq 0$  if the homogeneous TQ form of  $\hat{\mathcal{K}}$  is non-negative,  $\hat{\mathcal{K}}(\mathbf{g}) = \hat{K}_{j_1 j_2 j_3 j_4} g_{j_1} g_{j_2} g_{j_3} g_{j_4} \geq 0$ . This allows us to use Hilbert's theorem on non-negative TQs as shown in [10, 9, 11]. We use [11] to parameterize  $\hat{\mathcal{K}}$  as  $\hat{\mathcal{K}}(\mathbf{g}) = \mathbf{X}^T \mathbf{U} \mathbf{X}$ , where  $\mathbf{X} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \mathbf{x}_3^T]^T$ , with  $\mathbf{x}_i$  representing the six coefficients of the  $i^{th}$  quadratic form, and  $\mathbf{U}$  is the  $18 \times 18$  block diagonal matrix ( $\mathbf{W}$  in [11]), with each  $6 \times 6$  diagonal block entry as  $\mathbf{V} = \mathbf{v} \mathbf{v}^T$  where  $\mathbf{v}^T = [g_1^2, g_2^2, g_3^2, g_1 g_2, g_1 g_3, g_2 g_3]$ . Therefore, estimating the 18 unknowns of  $\mathbf{X}$  to compute the 15 unknowns of  $\hat{\mathcal{K}}$ , via the Gram matrix, implicitly imposes the non-negativity constraint on  $\hat{\mathcal{K}}$  and  $K_{app}$ . Finally the gradient  $\nabla_{\mathbf{X}} \hat{\mathcal{K}}(\mathbf{g}) = 2\mathbf{U} \mathbf{X}$  ( $\mathbf{U}$  symmetric) [11].
3. This can be rewritten in terms of  $\mathbf{D}$  and  $\hat{\mathcal{K}}$  with the above parameterizations as  $C_3(\mathbf{g}) = \hat{\mathcal{K}}(\mathbf{g}) - 3D_{app}(\mathbf{g})/b = \mathbf{X}^T \mathbf{U} \mathbf{X} - 3tr(\mathbf{G} \mathbf{L} \mathbf{L}^T)/b < 0$ . We compute its gradient to be  $\nabla_{\mathbf{L}} C_3(\mathbf{g}) = -(3/b) \nabla_{\mathbf{L}} D_{app}(\mathbf{g})$  and  $\nabla_{\mathbf{X}} C_3(\mathbf{g}) = \nabla_{\mathbf{X}} \hat{\mathcal{K}}(\mathbf{g})$ .

The Cholesky and TQ parameterizations of  $\mathbf{D}$  and  $\hat{\mathcal{K}}$  impose the constraints (#1) & (#2) implicitly. Therefore, estimating the 6 unknown coefficients of  $\mathbf{L}$  and the 18 unknown coefficients of  $\mathbf{X}$ , to reconstruct  $\mathbf{D}$  and  $\hat{\mathcal{K}}$ , ensures (#1) & (#2) without the need for applying them explicitly. Incorporating these the DKI signal can be written as

$$E(b, \mathbf{g}) = \exp\left(-b \cdot tr(\mathbf{G} \mathbf{L} \mathbf{L}^T) + b^2 \cdot \mathbf{X}^T \mathbf{U} \mathbf{X} / 6\right). \quad (7)$$

Constraint (#3) is the only one that is applied explicitly – along the signal acquisition directions  $\mathbf{N}$ .

The objective function of the CWLS estimator can be formulated from the log-linearized DKI signal in Eq. (7) using the constraint-parameterizations as

$$\mathcal{E}_{CWLS}(\mathbf{L}, \mathbf{X}) = \frac{1}{2} \sum_{j=1}^N \omega_j \left( \ln \left( \frac{S_j}{S_0} \right) + b_j tr(\mathbf{G}_j \mathbf{L} \mathbf{L}^T) - \frac{b_j^2}{6} \mathbf{X}^T \mathbf{U}_j \mathbf{X} \right)^2 \quad (8)$$

$$\text{subject to } C_3(\mathbf{L}, \mathbf{X}, \mathbf{g}_j) = \mathbf{X}^T \mathbf{U}_j \mathbf{X} - 3tr(\mathbf{G}_j \mathbf{L} \mathbf{L}^T) / b_j < 0, \quad j = 1..N,$$

where  $\omega_j = E_j^2 / \sigma_j^2$  with  $E_j = (S_j / S_0)$  and  $\sigma_j^2$  the variance of  $E_j$ . Eq. (8) forms a constrained nonlinear programming problem in the 24 variables of  $\Theta = [\text{vec}(\mathbf{L})^T, \mathbf{X}^T]^T$  where clearly both the objective function  $\mathcal{E}_{CWLS}(\mathbf{L}, \mathbf{X})$  and the inequality constraint  $C_3(\mathbf{L}, \mathbf{X}, \mathbf{g}_j)$  are twice differentiable w.r.t.  $\mathbf{L}$  and  $\mathbf{X}$ .

There exist a number of sophisticated iterative algorithms for efficiently solving such kinds of well behaved constrained problems. These include barrier methods, interior point methods, active-set methods, sequential quadratic programming (SQP), etc. A number of these methods are available under the MATLAB Optimization Toolbox. In particular, we solve our constrained nonlinear programming problems, including Eq. (8), using a MATLAB implementation of SQP. All such solvers, SQP including, require the gradients of both the objective and the constraint functions. The solvers are even more efficient when the gradients can be computed analytically. The gradient of the constraint was already computed to be  $\nabla_{\Theta} C_3(\Theta, \mathbf{g}_j) = [\nabla_{\mathbf{L}} C_3(\mathbf{g}_j)^T, \nabla_{\mathbf{X}} C_3(\mathbf{g}_j)^T]^T$ . The gradient

of  $\mathcal{E}_{CWLS}(\Theta)$  is  $\nabla_{\Theta}\mathcal{E}_{CWLS}(\Theta) = [\nabla_{\mathbf{L}}\mathcal{E}_{CWLS}(\Theta)^T, \nabla_{\mathbf{X}}\mathcal{E}_{CWLS}(\Theta)^T]^T$  where

$$\nabla_{\mathbf{L}}\mathcal{E}_{CWLS}(\Theta) = \sum_{j=1}^N \omega_j b_j \left( \ln \left( \frac{S_j}{S_0} \right) + b_j \text{tr}(\mathbf{G}_j \mathbf{L} \mathbf{L}^T) - \frac{b_j^2}{6} \mathbf{X}^T \mathbf{U}_j \mathbf{X} \right) \nabla_{\mathbf{L}} D_{app}(\mathbf{g}_j) \quad (9)$$

$$\nabla_{\mathbf{X}}\mathcal{E}_{CWLS}(\Theta) = \sum_{j=1}^N \omega_j \frac{-b_j^2}{6} \left( \ln \left( \frac{S_j}{S_0} \right) + b_j \text{tr}(\mathbf{G}_j \mathbf{L} \mathbf{L}^T) - \frac{b_j^2}{6} \mathbf{X}^T \mathbf{U}_j \mathbf{X} \right) \nabla_{\mathbf{X}} \hat{\mathcal{K}}(\mathbf{g}_j) \quad (10)$$

Although, easy to implement and computationally light, the CWLS estimator doesn't take into consideration the correct noise model of the diffusion signal. The magnitude diffusion signal is known to be independently Rice distributed and the maximum likelihood estimator (MLE), which, based on the Rician distribution attempts to maximize the joint likelihood function of all observations, is known to be asymptotically unbiased [13, 6]. Taking this into consideration has been shown to be especially crucial at low SNRs. The Rician distribution of an observation of the diffusion signal was shown to be [13, 6]

$$p(E_j | E(b_j, \mathbf{g}_j; \Theta), \sigma) = \frac{E_j}{\sigma^2} \exp \left( \frac{-E_j^2 - E(b_j, \mathbf{g}_j; \Theta)^2}{2\sigma^2} \right) I_0 \left( \frac{E_j E(b_j, \mathbf{g}_j; \Theta)}{\sigma^2} \right), \quad (11)$$

where  $I_0$  is the zeroth order modified Bessel function of the first kind and  $\sigma^2$ , the variance of  $E_j$  or the noise ratio, can either be approximated from the image background or also be considered as an unknown that can be estimated along with  $\Theta$ . We consider the form in Eq. (7) for  $E(b_j, \mathbf{g}_j; \Theta)$  to incorporate the constraints (#1) and (#2) implicitly.

The constrained MLE (CMLE) objective function that maximizes the joint log-likelihood of all observations based on the above Rician distribution can be derived, for the unknowns  $\Theta$ , to be

$$\mathcal{E}_{CMLE}(\Theta) = \sum_{j=1}^N \ln p(E_j | E(b_j, \mathbf{g}_j; \Theta), \sigma) \quad (12)$$

$$\text{subject to } C_3(\mathbf{L}, \mathbf{X}, \mathbf{g}_j) = \mathbf{X}^T \mathbf{U}_j \mathbf{X} - 3\text{tr}(\mathbf{G}_j \mathbf{L} \mathbf{L}^T)/b_j < 0, \quad j = 1..N.$$

This again forms a well behaved, twice differentiable, constrained nonlinear programming problem that can be efficiently solved by a SQP solver. In our experiments, again we use the MATLAB implementation. The gradient of  $\mathcal{E}_{CMLE}(\Theta)$  can be easily derived from the CWLS example above, i.e. Eqs. (9) & (10).

Thus, we have derived two formulations, the CWLS and the CMLE, to accurately estimate the constrained DKI tensors. We have also derived the gradients of the constrained problems to take advantage of efficient optimization solvers. Finally, in the CMLE, we have also incorporated the Rician noise model to make our estimator robust to signal noise.

## 4 Experiments and Results

In this section we validate the accuracy, robustness and speed of our proposed method by conducting extensive experiments on synthetic data. The synthetic

data was generated using the multi-tensor model with 1, 2 and 3 crossing fiber configurations randomly weighted and oriented in space. The diffusion of a single fiber compartment was modelled by a diagonal tensor  $\tilde{\mathbf{D}} = \text{diag}(\lambda_1, \lambda_2, \lambda_2) \times 10^{-3} \text{mm}^2/\text{s}$ , where the  $\lambda_i$ 's were taken in the ranges typically found in the human brain  $\lambda_1 \in [1, 2]$  and  $\lambda_2 \in [0.1, 0.6]$  [14]. This resulted in fractional anisotropy (FA) values in the range of  $\sim [0.52, 0.93]$ .

To validate the robustness to noise, we corrupted the signal with Rician distributed noise with SNR values in the range of [10, 30]. Although on real data, the noise ratio  $\sigma$ , needs to be estimated from the acquired signal, since we conducted experiments on synthetic data this was known to us  $\sigma = 1/\text{SNR}$ . To validate the accuracy of our constrained estimators, we followed [6] and considered the ground-truths to be the DKI tensors estimated from noise-free signals using the CWLS formulation. Note however, that for the multi-tensor model  $D_{app}(\mathbf{g})$  and  $K_{app}(\mathbf{g})$  can be computed analytically [2].

We followed the experimental design proposed in [6] and based our signal generation on their optimal choice of  $b$ -values for DKI (see reference [21] in [6] for an extensive study). The synthetic data was generated along 140 gradient directions with 25, 40 and 75 gradients isotropically distributed over the shells with  $b = 700, 1000, 2800 \text{s}/\text{mm}^2$ . The gradient directions were optimally distributed on every shell using the algorithm in [15], which is available online.

We conducted experiments A) to quantify the effects of considering the TQ parameterization and gradient based optimization schemes – with analytically computed gradients – in terms of improved speed and quality, B) to evaluate the accuracy and robustness of the CWLS and CMLE formulations we have proposed and C) to determine the degree of constraint violations for unconstrained estimators to motivate the importance of constrained DKI estimation.

First, to evaluate the contributions of the TQ parameterization and the analytical gradient based optimization, we conducted experiments on 10 randomly generated voxels with an SNR level of 15. We compared four estimation schemes based on estimation time and quality (i) UNM: unconstrained Nelder-Mead, (ii) LSTR: unconstrained gradient based estimator solved by a large scale trust region based scheme (MATLAB: `fminunc`), (iii) D-CSQP: a gradient based scheme which applies the constraints like in [6] – Cholesky for (#1) and discrete constraints along acquisition directions for (#2) & (#3), solved by the SQP scheme and (iv) TQ-SQP: where we followed Sec. 3: constraints (#1) & (#2) were applied using Cholesky and TQ parameterizations and (#3) was applied along acquisition directions – solved by the SQP scheme using analytical gradients. Note that our MATLAB Nelder-Mead based estimators (unconstrained and constrained) were extremely slow at converging. Therefore, we considered UNM only on 1 voxel and used a gradient based solver (D-CSQP) to emulate the approach in [6] (which already improves its performance).

To arrive at a fair comparison, we standardized all numerical parameters. We considered similar stopping tolerances for the optimization solvers and also initialized all optimizations with similar initial points for each voxel. The speed and quality of the four estimators were evaluated using five different metrics

	UNM*	LSTR <sup>+</sup>	D-CSQP	<b>TQ-SQP</b>
$t$	24m 30s	2m 38s	14m 42s	9.43s
$f_{min}$	0.06	0.07	0.06	0.06
$\ \nabla f\ $	-	4.12	0.3	0.07
MSE-MD	4e-9	1e-8	5e-9	4e-9
MSE-MK	0.4	1.5e+3	0.3	4e-11

**Table 1.** WLS: Estimation time is dramatically decreased by using the TQ parameterization and an analytical gradient based optimization. UNM\*: The results of UNM are biased since only one voxel has been considered, it is kept for comparing the estimation time. LSTR<sup>+</sup>: For WLS we could use the simple linear pseudoinverse solution but we use a gradient based solver for comparing the estimation time.

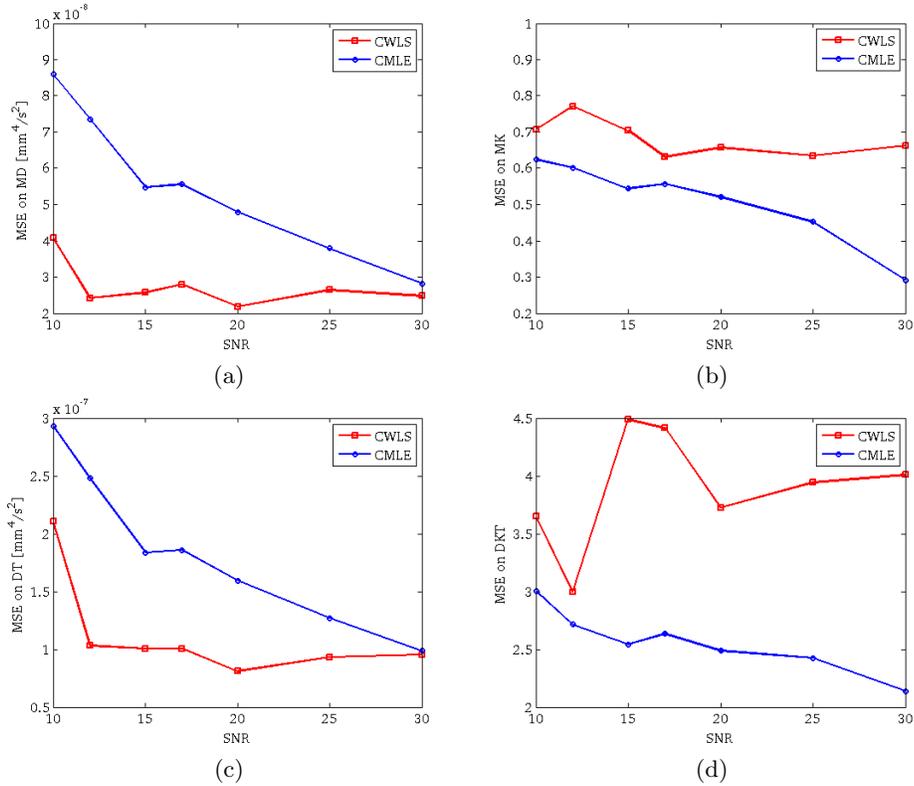
	UNM*	LSTR	D-CSQP	<b>TQ-SQP</b>
$t$	20m	2m 40s	26m 12s	48s
$f_{min}$	180	173	191	188
$\ \nabla f\ $	-	2e+3	189	330
MSE-MD	2e-8	2e-9	2e-8	5e-8
MSE-MK	0.17	443	1.53	0.24

**Table 2.** MLE: Estimation time is again dramatically decreased by using the TQ parameterization and an analytical gradient based optimization. UNM\*: as above.

(a)  $t$ : the estimation time, (b)  $f_{min}$ : the goodness of fit, (c)  $\|\nabla f\|$ : the first order optimality criterion, (d) MSE-MD: mean squared error (MSE) on mean diffusivity (MD) and (e) MSE-MK: MSE on mean kurtosis (MK). We carried out the computations on a 64-bit quad-core computer with 8 Gb of RAM and each core running at 3.20 GHz. The results are presented in Table 1 for the WLS estimator and in Table 2 for the ML estimator.

As expected, incorporating gradient based optimization schemes with analytically computed gradients dramatically reduces the estimation time for both the WLS and ML estimators. Clearly applying explicit constraints is very expensive. This is expected since every explicit constraint requires additional computations, implying that every discrete explicit constraint comes at a high price. Therefore, the TQ parameterization, which converts (#2) to an implicit constraint is highly valuable. Also, from the quality of the convergence, it is clear that implicit constraints perform better. Explicit constraints tend to distort the search space, which affects the quality of their solutions. Furthermore, when one doesn't take advantage of the gradients of the constraints and the objective function, clearly the optimization suffers a heavy setback as seen for the Nelder-Mead scheme.

Overall, therefore, TQ-SQP, the TQ parameterization with the gradient based SQP optimizer using analytically computed gradients stands out with dramatically superior performance in terms of both time and quality. TQ-SQP, both in the CWLS and the CMLE formulations has the fastest estimation time by far – orders of magnitude faster than all the other estimation schemes. Furthermore, its estimation quality is also generally better than the other schemes. At its worse it is comparable to any of the other methods. Indeed, TQ-SQP seems

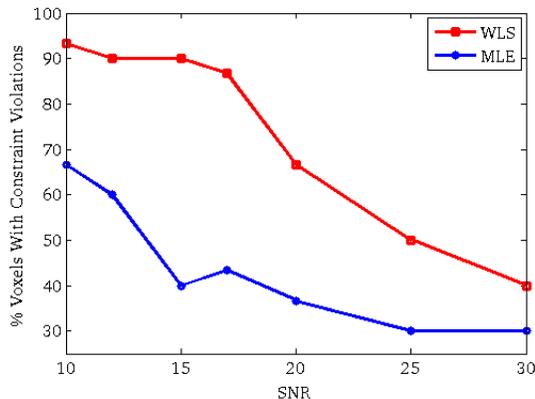


**Fig. 1.** Accuracy and robustness of our proposed CWLS and CMLE formulations on synthetically generated data with varying SNR. Results are from 100 voxels for each SNR and they compare the CWLS to the CMLE. (a) MSE-MD, (b) MSE-MK, (c) MSE-DT and (d) MSE-DKT.

to be the method of choice. Last, Tables 1 & 2 highlight the importance of the three constraints from the two MSE errors of the unconstrained schemes.

Next we test the accuracy and robustness to noise of the CWLS and CMLE formulations we have proposed in Sec. 3. We conduct experiments on 100 randomly generated voxels with SNR varying in  $[10,30]$ . First we compute the ground-truths as described above, then for each SNR we estimate the two DKI tensors in every voxel using the CWLS and CMLE formulations with TQ parameterization and analytical gradients. We then evaluate the accuracy and robustness of the two estimators by comparing the estimated tensors to the ground-truth tensors using the four quantities proposed in [6]: (a) MSE-MD, (b) MSE-MK, (c) MSE-DT: MSE on the coefficients of  $\mathbf{D}$  and (d) MSE-DKT: MSE on the coefficients of the kurtosis tensor  $\mathcal{W}$ . The results are presented in Fig. 1.

The graphs in Fig. 1 confirm the trends of the results in [6] qualitatively. Quantitatively the results are slightly different since we use different fiber con-



**Fig. 2.** Violations of (#3) when explicit constraints are not considered and the importance of constrained DKI estimation. (#1) and (#2) were applied implicitly.

figurations to generate the synthetic data. We have considered single, double and triple fiber configurations instead of only the two compartment model used in [6]. Further, we have considered different configurations in each voxel and our single fiber model spans a range. Qualitatively, nonetheless, we perceive that the CMLE is more robust than the CWLS and has more consistent error plots.

In case of the diffusion tensor, we find that both the CWLS and the CMLE perform well and the errors they commit are in the orders of  $10^{-7}$ . The advantage of the CMLE, however, becomes evident in the case of the kurtosis tensor and mean kurtosis. The CWLS estimator is clearly highly sensitive to the SNR and the errors it commits fluctuate with the SNR. As expected, the Rician noise-model based CMLE, on the other hand, turns out to be robust across all SNR values and, therefore, has the upper hand. This is in agreement with [6] and highlights the importance of the Rician noise-model based ML estimator.

To determine the importance of the constraints and motivate their necessity, we finally conduct an experiment where we do not impose any explicit constraints, i.e. (#3). We only consider the constraints (#1) and (#2) that can be applied implicitly by the Cholesky and TQ parameterizations. We estimate the “semi-unconstrained” DKI tensors and check for the percentage of voxels where (#3) is violated. The results are presented in Fig. 2. The trend in the graph clearly indicates the need for even constraint (#3) when (#1) and (#2) are implicitly accounted for. Further, the WLS violates the constraint more often than the MLE. Both approaches violate (#3) more with decreasing SNR. Fig. 2, in conjunction with the MSE errors of the unconstrained estimation schemes presented in the Tables 1 & 2 emphasize the importance of the three constraints while estimating the DKI tensors.

## 5 Conclusion

DKI is a recent extension to the DTI signal model that characterizes the tissue microstructure by measuring the non-Gaussian diffusion. It is proving to be valuable for discerning changes in the cerebral tissue and is, therefore, an interesting dMRI signal model. DKI estimates a 4<sup>th</sup> order kurtosis tensor in addition to the 2<sup>nd</sup> order diffusion tensor. However, to be physically relevant the DKI tensors have to satisfy three constraints. We have presented a method to incorporate two of the constraints implicitly and the third explicitly. Our ternary quartic based parameterization of the kurtosis tensor allows to elegantly impose the non-negativity constraint. Further, we have proposed a gradient based optimization scheme that dramatically improves both the estimation time and quality of the tensors. Finally, we have also incorporated the Rician noise model to render our constrained estimator robust to signal acquisition noise.

**Acknowledgements:** This work was partially supported by the ANR project NucleiPark and the France-Parkinson Association.

## References

1. Jensen, J.H., Helpert, J.A., Ramani, A., Lu, H., Kaczynski, K.: Diffusional Kurtosis Imaging: The Quantification of Non-Gaussian Water Diffusion by Means of Magnetic Resonance Imaging. *Magnetic Resonance in Medicine* **53**(6) (2005) 1432–1440
2. Jensen, J.H., Helpert, J.A.: MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR in biomedicine* **23**(7) (2010) 698–710
3. Basser, P.J., Mattiello, J., Le Bihan, D.: Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance* **B**(103) (1994) 247–254
4. Callaghan, P.T.: Principles of nuclear magnetic resonance microscopy. Oxford University Press, Oxford (1991)
5. Tabesh, A., Jensen, J.H., Ardekani, B.A., Helpert, J.A.: Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. *Magnetic Resonance in Medicine* **65**(3) (2011) 823–836
6. Veraart, J., Van Hecke, W., Sijbers, J.: Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a Rician noise model. *Magnetic Resonance in Medicine* **66**(3) (2011) 678–686
7. Barmpoutis, A., Zhuo, J.: Diffusion kurtosis imaging: Robust estimation from DW-MRI using homogeneous polynomials. In: Proceedings of the 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2011, March 30 - April 2, 2011, Chicago, Illinois, USA, IEEE (April 2011) 262–265
8. Ardekani, B.A., Tabesh, A., Jensen, J.H., Helpert, J.A., Bachman, A., Kushner, H.: An Improved Method for Diffusional Kurtosis Estimation. In: Proceedings ISMRM. (May 2010)
9. Barmpoutis, A., Hwang, M.S., Howland, D., Forder, J.R., Vemuri, B.C.: Regularized Positive-Definite Fourth-Order Tensor Field Estimation from DW-MRI. *NeuroImage* **45**(1 sup.1) (March 2009) 153–162

10. Barmpoutis, A., Jian, B., Vemuri, B.C.: Symmetric Positive 4th Order Tensors & their Estimation from Diffusion Weighted MRI. In: Information Processing in Medical Imaging (IPMI 2007). (2007) 308–319
11. Ghosh, A., Moakher, M., Deriche, R.: Ternary quartic approach for positive 4th order diffusion tensors revisited. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. (June 2009) 618–621
12. Petersen, K.B., Pedersen, M.S.: The Matrix Cookbook (October 2008) Version 20081110.
13. Sijbers, J., den Dekker, A., Scheunders, P., Van Dyck, D.: Maximum-likelihood estimation of Rician distribution parameters. Medical Imaging, IEEE Transactions on **17**(3) (1998) 357–361
14. Canales-Rodríguez, E.J., Melie-García, L., Iturria-Medina, Y.: Mathematical description of q-space in spherical coordinates: Exact q-ball imaging. Magnetic Resonance in Medicine **61**(6) (2009) 1350–1367
15. Caruyer, E., Cheng, J., Lenglet, C., Sapiro, G., Jiang, T., Deriche, R.: Optimal Design of Multiple Q-shells experiments for Diffusion MRI. In: MICCAI Workshop on Computational Diffusion MRI - CDMRI'11, Toronto, Canada (September 2011)

# Generalized Invariants of a $4^{th}$ order tensor: Building blocks for new biomarkers in dMRI

Aurobrata Ghosh, Théo Papadopoulo, and Rachid Deriche

Project Team Athena, INRIA Sophia Antipolis - Méditerranée, France

**Abstract.** This paper presents a general and complete (up to degree 4) set of invariants of 3D  $4^{th}$  order tensors with respect to  $\mathcal{SO}_3$ . The invariants to  $\mathcal{SO}_3$  for the  $2^{nd}$  order diffusion tensor are well known and play a crucial role in deriving important biomarkers for DTI, e.g. MD, FA, RA, etc. But DTI is limited in regions with fiber heterogeneity and DTI biomarkers severely lack specificity.  $4^{th}$  order tensors are both a natural extension to DTI and also form an alternate basis to spherical harmonics for spherical functions. This paper presents a systematic method for computing the  $\mathcal{SO}_3$  invariants of 3D  $4^{th}$  order tensors, derives relationships between the new (generalized) invariants and existing invariants and shows results on synthetic and real data. It also presents, hitherto unknown, new invariants for  $4^{th}$  order tensors. Analogously to DTI, these new invariants can perhaps form building blocks for new biomarkers.

**Keywords:** Invariants,  $4^{th}$  order tensors, homogeneous forms/polynomials,  $\mathcal{SO}_3$ , basic invariants, principal invariants, biomarkers.

## 1 Introduction

The invariants of a  $2^{nd}$  order diffusion tensor, to  $\mathcal{SO}_3$  the rotation group, play a fundamental role in diffusion tensor imaging (DTI). They form the building blocks for a wide range of popularly used biomarkers such as mean diffusivity (MD), fractional anisotropy (FA), relative anisotropy (RA), etc. [1]. The importance of the diffusion tensor's invariants is highlighted by the importance of biomarkers in diffusion MRI (dMRI). Biomarkers are crucial in analyzing changes in white matter related to development, degeneration and disease.

However, DTI is limited in regions with heterogeneous fiber configurations. Furthermore, DTI based biomarkers severely lack specificity – not only due to the coarse resolution of dMRI with respect to the actual axons, but also due to the over-simplification of the DTI model. Therefore, the MD or the FA are affected similarly by a variety of disparate types of physical changes or acquisition settings (partial voluming). Hence, there is a strong need for conceiving of new scalar measures for characterizing the integrity of white matter – especially from higher order models which describe the microstructure with greater accuracy.

Higher order tensors and in particular  $4^{th}$  order tensors are a natural generalization of the  $2^{nd}$  order diffusion tensor. Initially proposed in generalized DTI (GDTI) [2] for describing complex apparent diffusion coefficient (ADC) profiles,

they also appear as a bijective alternate basis to the spherical harmonic basis for describing generic functions on the sphere [2].  $4^{th}$  order tensors, therefore, can be used for describing a number of well-known spherical functions from higher order models in dMRI. Furthermore,  $4^{th}$  order tensors have been extensively studied and important mathematical frameworks have been developed for estimating ADCs with the positivity constraint [3–6]

In this paper, we present the general and a complete (up to degree 4) set of invariants of 3D  $4^{th}$  order tensors with respect to  $\mathcal{SO}_3$ . There exists no simple approach for either computing or for interpreting these invariants as for  $2^{nd}$  order tensors. However, analogously to DTI, we believe that these invariants would form the building blocks for new biomarkers. We call these the “general” invariants to distinguish them from the integrity basis of  $4^{th}$  order tensors that only compute certain types of invariants – the basic ( $S4_i$ ) and the principal ( $J4_i$ ) [7, 8]. To our knowledge, this approach has never been presented before.

There exist two known methods for computing the invariants of 3D  $4^{th}$  order tensors to  $\mathcal{SO}_3$ . In [9] the authors integrate a function of the tensor over the sphere to compute a scalar measure (generalized trace) invariant under 3D rotations. This provides a way for computing the invariants. However, there is no systematic way for choosing which functions to integrate, therefore, there is no systematic way of generating a linearly independent set of invariants using this method. In [7, 8] the authors compute the integrity basis – the 6 basic and the algebraically dependent 6 principal invariants from the spectral decomposition of the  $6 \times 6$  matrix representation of the  $4^{th}$  order tensor. However, these are designed to be invariant to  $\mathcal{SO}_6$ . Therefore, this approach is limited since symmetric 3D  $4^{th}$  order tensors have 15 independent coefficients and any 3D rotation can be described by 3 parameters, implying that 3D  $4^{th}$  order tensors should have  $12 = 15 - 3$  algebraically independent invariants to  $\mathcal{SO}_3$ .

In this paper, we work out a systematic approach for computing the polynomial invariants which are complete up to a given degree. Our approach is based on the polynomial interpretation of tensors. It is well known that tensors can be rewritten as homogeneous polynomials or “forms” by a rearrangement of the indices. For example a 3D  $4^{th}$  order tensor is simply a ternary quartic form – a homogeneous degree-4 polynomial in 3 variables. After presenting the theory and deriving their formulae, we conduct experiments on synthetic datasets in an attempt to understand these invariants and finally present results on an in vivo human cerebral dataset.

## 2 Materials and Methods

This section describes a systematic method to generate the polynomial invariants of a form under the group  $\mathcal{SO}_3$  of the rotations of  $\mathcal{R}^3$ . The method will be illustrated with the simple example of 2-forms (quadratic forms). Then, the results of applying the method to 4-forms (quartic forms) will be described.

## 2.1 Invariants of a form under group transformations

Let  $\mathcal{G}$  denote a transformation group operating on  $\mathcal{R}^3$ . We further assume that  $\mathcal{G}$  is a Lie group, i.e. a group that depends continuously on set of  $n$  parameters. Let  $\mathcal{H}$  be a family of objects of  $\mathcal{R}^3$  depending on  $p$  parameters. Since  $\mathcal{G}$  operates on  $\mathcal{R}^3$ , it induces transforms over the family  $\mathcal{H}$ . Invariants of  $\mathcal{H}$  under the group  $\mathcal{G}$  are expressions involving the  $p$  parameters of  $\mathcal{H}$  that remain unchanged under the transformations induced by the group  $\mathcal{G}$ . Polynomial invariants correspond to those expressions which are polynomials in the  $p$  parameters of  $\mathcal{H}$ . Invariant theory states that all polynomial invariants can be expressed as polynomial expressions of a finite set of  $p - n$  basic polynomial invariants.

As an example,  $4^{th}$  order forms over  $\mathcal{R}^3$ , or ternary quartics, depend on  $p = 15$  parameters. Thus, there are  $12 = 15 - 3$  basic polynomial invariants of  $4^{th}$  order forms under the rotation group of  $\mathcal{R}^3$  ( $n=3$ ).

## 2.2 Generators of $\mathcal{SO}_3$

The method proposed in the next section is quite general but relies on rational parameterization of the group  $\mathcal{G}$ . Such a parameterization is thus described hereafter for  $\mathcal{SO}_3$ , which is the group of interest in the context of this work. Euler angles is a way to parameterize rotations. Any rotation  $\mathbf{R}$  of  $\mathcal{SO}_3$  can be written as the composition of three elementary rotations with angles  $\alpha, \beta, \gamma$ :

$$\mathbf{R} = \mathbf{R}_Z(\alpha)\mathbf{R}_Y(\beta)\mathbf{R}_Z(\gamma), \quad (1)$$

where  $\mathbf{R}_Y(\theta)$  and  $\mathbf{R}_Z(\theta)$  denote respectively rotations around the  $Y$  and  $Z$  axis with the angle  $\theta$ . Introducing  $t = \tan \frac{\theta}{2}$ ,  $\mathbf{R}_Y(\theta)$  and  $\mathbf{R}_Z(\theta)$  can be written as:

$$\mathbf{R}_Y(\theta) = \begin{bmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{bmatrix}, \quad \mathbf{R}_Z(\theta) = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{with } c = \frac{1-t^2}{1+t^2} \text{ and } s = \frac{2t}{1+t^2}.$$

Because of Eq. (1), a polynomial expression separately invariant under both the two sub-families  $\mathbf{R}_Y(\theta)$  and  $\mathbf{R}_Z(\theta)$  (for any values of  $\theta$ ) is also invariant under the whole group  $\mathcal{SO}_3$ . The next section describes how to find such polynomial expressions. It will be applied sequentially with the subgroups  $\mathbf{R}_Y(\theta)$  and  $\mathbf{R}_Z(\theta)$  to generate polynomial expressions invariant under  $\mathcal{SO}_3$ .

## 2.3 Systematic generation of polynomial invariants

Let us denote by  $\mathbf{S}(t)$  the group under which invariants are sought.  $t$  is the parameter which allows the rational representation of the Lie group. Let  $\mathbf{m} = [x, y, z]$  be a vector of the 3D space. Let  $f_q(\mathbf{m}) = \sum_{i+j+k=q} \mu_{ijk} c_{ijk} x^i y^j z^k$  denote a homogeneous form of degree  $q$  with coefficients  $c_{ijk}$  (there are  $p$  such coefficients).  $\mu_{ijk}$  are constants eventually used to represent weights of the monomial  $x^i y^j z^k$ .

For example, a 2-form is defined by  $f_2(\mathbf{m}) = c_{200}x^2 + 2c_{110}xy + 2c_{101}xz + c_{020}y^2 + 2c_{011}yz + c_{002}z^2$  ( $p_2 = 6$ ).

For every  $t$ ,  $\mathbf{S}(t)$  defines a change of coordinates  $\mathbf{m}' = \mathbf{S}(t)\mathbf{m}$ . Applying this change to  $f_q$  yields a new form of degree  $q$   $f'_q(\mathbf{m}') = \sum_{i+j+k=q} \mu_{ijk} c'_{ijk} x^i y^j z^k$ .  $c'_{ijk}$  are rational expressions in  $t$  and linear expressions in the coefficients  $c_{ijk}$ .

Let  $I_d(\{c_{ijk}\})$  be a polynomial expression of degree  $d$  with coefficients  $m_r, r = 1..R$ . Without loss of generality,  $I_d$  can be restricted to be a homogeneous polynomial.  $I_d$  is invariant iff  $I_d(\{c_{ijk}\}) - I_d(\{c'_{ijk}\}) = 0$ . The numerator of this expression is a polynomial in  $t$  and  $c_{ijk}$  which must be identically 0. Consequently, each of its coefficients is null yielding a set of linear constraints on the coefficients  $m_r$ . This system potentially has multiple solutions providing a generative family of the polynomial invariants of degree  $d$ . The full algorithm is described in 1.

---

**Algorithm 1** Generation of a basis of the set of homogeneous polynomial invariants of degree  $d$  of a form  $f$  under  $\mathcal{SO}_3$ .

---

**Require:** A polynomial expression  $f(\mathbf{m})$  with coefficients  $c_{ijk}$ .

Start with a generic homogeneous polynomial  $I_d(\{c_{ijk}\})$  of degree  $d$  in the variables  $c_{ijk}$  with coefficients  $m_{ijk}$ .

Initialize the parameter list  $P = \{m_{ijk}\}$ .

**for all**  $\mathbf{S}(t) \in \{\mathbf{R}_Z(t), \mathbf{R}_Y(t)\}$  **do**

    Compute coefficients  $c'_{ijk}$  of  $f(\mathbf{S}(t)\mathbf{m})$  induced by the change of coordinates  $\mathbf{S}(t)$ .

    Compute the numerator of  $D(t, P) = I_d(\{c_{ijk}\}) - I_d(\{c'_{ijk}\})$ .

    Extract all the coefficients of the polynomial  $D(t, P) \rightarrow L$ .

    ▷  $L$  is a linear system in the coefficients of  $P$ .

    Solve  $L$  and substitute the solution in  $I_d(\{c_{ijk}\})$ .

    ▷  $L$  is not a square system, its solution contains some of the unknowns  $m_{ijk}$ .

    Update  $P$  as the list of parameters in the solution of  $L$ .

**end for**

Extract the coefficients of  $I_d$  with respect to  $P \rightarrow Invs$ .

▷  $Invs$  is a basis of the set of the sought polynomial invariants.

**return**  $Invs$

---

Considering the 2-form  $f_2$ , an invariant of degree 1 can be written as  $I_d(\{c_{ijk}\}) = \sum_{l \in Ind} m_l c_l$ , with  $Ind = \{200, 020, 002, 110, 101, 011\}$  and  $P = \{m_l, l \in Ind\}$ . Using the family of rotations  $\mathbf{R}_Z(t)$ , the numerator of  $I_d(\{c_{ijk}\}) - I_d(\{c'_{ijk}\}) = 0$  (divided by  $2t$ ) can be written as:

$$2(m_{200} - m_{020}) ((c_{020} - c_{200})t - c_{110}(1 + t^2)) + m_{110}(1 - t^2)(c_{020} - c_{200}) - (1 + t^2)(c_{011}(m_{011}t - m_{101}) + c_{101}(m_{101}t + m_{011})) - 4m_{110}c_{110}t = 0,$$

which yields  $m_{200} = m_{020}, m_{110} = m_{011} = m_{101} = 0$ . This means that  $I_d(\{c_{ijk}\}) = m_{200}(c_{200} + c_{020}) + m_{002}c_{002}$  with a new set of parameters  $P = \{m_{200}, m_{002}\}$ . Applying the same procedure to this new polynomial with  $\mathbf{R}_Y(\theta)$  yields the familiar trace invariant  $c_{200} + c_{020} + c_{002}$ . In this case, the last linear system has only one solution (up to an arbitrary scale factor)  $m_{200} = m_{002}$ .

## 2.4 Linear independence and simplification of the invariants

Using algorithm 1, we get  $I_d$  a basis of all homogeneous polynomial invariants of degree  $d$ . Yet nothing ensures that those invariants are independent of those of lower degrees. This section explains how to construct a linearly independent set of invariants up to degree  $d$ . Algorithm 2 is recursive: assuming that a linearly independent set  $J_{d-1}$  of invariants up to degree  $d-1$  is available, it computes  $J_d$  given  $I_d$ . In this paper,  $J_1$  is always non empty set and contains one single polynomial (so it is independent). Algorithm 2 makes use of a function  $matrix(K_d)$ , where  $K_d$  is a set of homogeneous polynomials of degree  $d$ . This function returns a matrix whose entries are the coefficients of the polynomials. Each row of the matrix corresponds to a polynomial of  $K_d$ . Each column corresponds to a monomial of degree  $d$  (ordered arbitrarily). If the rank of this matrix is the same as the number of polynomials in  $K_d$ , then the system  $K_d$  is linearly independent.

---

**Algorithm 2** Generation of a basis of the set of polynomial invariants up to degree  $d$ .

---

```

Initialize  $J_d = J_{d-1}$ .
Compute a basis  $K_d$  of the homogeneous polynomial of degree  $d$  generated by  $J_{d-1}$ .
▷  $I_d$  is assumed to be sorted in ascending order in terms of simplicity, where
▷ simplicity is for example the number of terms of the polynomial.
for all  $I_d^i \in I_d$  do
  if  $rank(matrix(K_d \cup \{I_d^i\})) = rank(matrix(K_d)) + 1$  then
    ▷  $I_d^i$  is linearly independent from the polynomials in  $K_d$ , add it.
     $K_d = K_d \cup \{I_d^i\}$ .
     $J_d = J_d \cup \{I_d^i\}$ .
  end if
end for
return  $J_d$ 

```

---

Note that  $M_d = matrix(J_d)$  can also be used to minimize the number of terms in the polynomials in  $J_d$ . Indeed, each column of  $M_d$  corresponds to a linear system on the polynomials in  $J_d$  which cancels the monomial corresponding to that column. Thus all columns provide a set of linear equations. This set can be explored combinatorially to find the linear combinations of the polynomials in  $J_d$  which cancel the most monomials.

## 2.5 Application to $2^{nd}$ order forms

Given  $f_2$ , we introduce the two vectors  $\mathbf{V}_1 = [c_{110}, c_{101}, c_{011}]$  and  $\mathbf{V}_2 = [c_{200}, c_{020}, c_{002}]$ . In this case, it is known that polynomial invariants can be generated using  $6 - 3 = 3$  polynomials  $J_{2_1}, J_{2_2}, J_{2_3}$  which correspond to the coefficients of the characteristic polynomial  $\mathbf{F}$  of the  $2^{nd}$  order tensor associated to  $f_2$  (in particular  $J_{2_1} = trace(\mathbf{F})$  and  $J_{2_3} = det(\mathbf{F})$ ). Using algorithm 1 for degrees 1, 2 and 3

gives:

$$\begin{aligned}
I_1 &= \{I_1^1\} && \text{with } I_1^1 = J2_1 = c_{002} + c_{020} + c_{200} \\
I_2 &= \{I_2^1, I_2^2\} && \text{with } I_2^1 = J2_2 \text{ and } I_2^2 = 2\|\mathbf{V}_1\|^2 + \|\mathbf{V}_2\|^2 \\
I_3 &= \{I_3^1, I_3^2, I_3^3\} && \text{with } I_3^1 = J2_1^3, I_3^2 = J2_1(I_2^1 - I_2^2) \text{ and } I_3^3 = J2_3.
\end{aligned}$$

Algorithm 2 gives successively  $J_1 = \{J2_1\}$ ,  $J_2 = \{J2_1, J2_2\}$  and  $J_3 = \{J2_1, J2_2, J2_3\}$ .  $J_3$  is exactly the set obtained using  $\mathbf{F}$ . Algorithm 2 also reveals (as expected) that invariants of higher degrees (4 to 10) can all be expressed using those of  $J_3$ .

## 2.6 Application to 4<sup>th</sup> order forms

Using the proposed method on 4-forms (15 parameters) yields 12 linearly independent invariants up to degree 4. The coefficients of  $f_4$  can be grouped in 5 vectors  $\mathbf{V}_1 = [c_{112}, c_{121}, c_{211}]$ ,  $\mathbf{V}_2 = [c_{220}, c_{202}, c_{022}]$ ,  $\mathbf{V}_3 = [c_{130}, c_{301}, c_{013}]$ ,  $\mathbf{V}_4 = [c_{310}, c_{103}, c_{031}]$  and  $\mathbf{V}_5 = [c_{400}, c_{040}, c_{004}]$ . As in section 2.3, each coefficient  $c_{ijk}$  is attached to the monomial  $x^i y^j z^k$  and the multiplicity  $\mu_{ijk}$  are respectively 12, 6, 4, 4 and 1 for the coefficients appearing in respectively  $\mathbf{V}_1$ ,  $\mathbf{V}_2$ ,  $\mathbf{V}_3$ ,  $\mathbf{V}_4$  and  $\mathbf{V}_5$ . Algorithm 2 finds 12 linearly independent invariants up to degree 4: 1 invariant of degree 1, 2 of degree 2, 4 of degree 3 and 5 of degree 4. These ‘‘general’’ invariants will be named  $G4_i$ ,  $i = 1..12$ . The first invariants are:

$$\begin{aligned}
G4_1 &= 2S_2 + S_5, && G4_2 = 12C_{11} + 6C_{22} + 4(C_{33} + C_{44}) + C_{55}, \\
G4_3 &= 4(C_{13} + C_{14} + C_{34} - C_{11}) - 3C_{22} - 2(C_{25} + D_2) - D_5,
\end{aligned}$$

where  $S_i = \sum_{j=1}^3 \mathbf{V}_{ij}$ ,  $C_{ij} = \mathbf{V}_i \cdot \mathbf{V}_j$ ,  $D_i = \sum_{j \neq k} \mathbf{V}_{ij} \mathbf{V}_{ik}$ . The other invariants (up to degree 4) are too complicated to be written here and will be provided on demand as maple code. It has not yet been possible to apply the method of section 2.3 to invariants of degree 5 and 6. While in this paper, we only prove the linear independence between the  $G4_i$ , it is easy to check that the first 3 invariants (up to degree 2) are actually algebraically (and not only linearly) independent. One of the invariants of degree 2  $G4_3$  is, therefore, new (not a basic/principal invariant). Note that  $G4_3$  could still be a consequence of e.g. the principal invariants  $J4_i$ ,  $i = 1..6$  as a power of  $G4_3$  might be written as a combination of  $\{J4_i\}$ . In any case,  $G4_3$  is simpler.

## 2.7 Relation with previously known invariants

The simplification method described in section 2.4 can also be used to express one invariant family as functions of other invariant families. Since the invariants obtained in the previous section are complete up to degree 4, all previously known (basic & principal) invariants (up to degree 4) can be expressed in terms

of the invariants  $G4_i, i = 1..12$ .

$$J4_1 = -G4_1, \quad J4_2 = \frac{1}{2}(G4_1^2 - G4_2), \quad J4_3 = -\frac{1}{6}G4_1^3 + \frac{1}{2}G4_1G4_2 - \frac{1}{3}G4_5$$

$$J4_4 = \frac{G4_1^4}{24} - \frac{5}{12}G4_1^2G4_2 + \frac{1}{3}G4_1G4_5 + \frac{1}{9}G4_1G4_7 + \frac{G4_2^2}{72} - \frac{1}{3}G4_2G4_3 - \frac{G4_9}{36} - \frac{G4_{11}}{504}$$

Thus, the principal invariants up to degree 4 can all be expressed in terms of the general invariants. This cannot be done with  $J4_5$  and  $J4_6$ . Therefore, in the end, we have 16 invariants which are linearly independent. Non-linear dependency among those is much more complicated and under investigation. Using the same method, trace or basic invariants can be expressed as functions of the principal invariants, i.e. basic invariants up to degree 4 are also consequences of the general invariants:

$$S4_1 = -J4_1, \quad S4_2 = J4_1^2 - 2J4_2, \quad S4_3 = -J4_1^3 + 3J4_1J4_2 - 3J4_3$$

$$S4_4 = J4_1^4 - 4J4_1^2J4_2 + 4J4_1J4_3 + 2J4_2^2 - 4J4_4$$

$$S4_5 = -J4_1^5 + 5J4_1^3J4_2 - 5J4_1^2J4_3 - 5J4_1J4_2^2 + 5J4_1J4_4 + 5J4_2J4_3 - 5J4_5$$

$$S4_6 = J4_1^6 - 6J4_1^4J4_2 + 6J4_1^3J4_3 + 9J4_1^2J4_2^2 - 6J4_1^2J4_4 - 12J4_1J4_2J4_3 +$$

$$6J4_1J4_5 - 2J4_2^3 + 6J4_2J4_4 + 3J4_3^2 - 6J4_6$$

### 3 Experiments and Results

Although any spherical function from higher order models either in the tensor basis or the spherical harmonic basis can be used as a source 4<sup>th</sup> order tensor, we adopt the simplest GDTI model with least square approximation [2]. However, computing these invariants isn't limited to GDTI.

On synthetic data we test the numerical stability of the 12 invariants to arbitrary 3D rotations to verify their validity. Synthetic diffusion signals in a voxel are generated using  $S(\mathbf{g}_i) = \sum_{k=1}^{N=1,2} (1/N) \exp(-b\mathbf{g}_i^T \mathbf{D}_k \mathbf{g}_i)$  with  $\mathbf{D}_k = \mathbf{R}_k^T \text{diag}(\lambda_1, \lambda_2, \lambda_2) \mathbf{R}_k$  where  $\mathbf{R}_k$  are 3D rotations and  $\{\lambda_1 = 2, \lambda_2 = 0.1\} \times 10^{-3} \text{mm}^2/\text{s}$ . Along the x-axis of the dataset we vary the crossing angle from 0° – 90° in 12 steps. Along the y-axis we vary the anisotropy of each fiber by varying  $\lambda_2$  from 0.1 –  $\lambda_1$  in 11 steps. And along the z-axis we vary the volume by increasing each  $\lambda_i$  tenfold over 10 steps.

We generate a second dataset with identical fiber configurations as the above, but with each voxel oriented randomly (Fig. 1a,b). We first estimate the 4<sup>th</sup> order tensors then compute the 12 invariants for both datasets  $\{G4_i\}$  &  $\{G4'_i\}$  and finally compute the ratio  $|G4_i|/|G4_i - G4'_i|$  for each invariant. For reference we also conduct the same experiment with FA from 2<sup>nd</sup> order tensors and with  $\{S4_i\}$  &  $\{J4_i\}$ . The results are presented in Fig. 1c where taller bars indicate bigger ratios and greater numerical stability to rotation. It is interesting to note that 4<sup>th</sup> order tensor invariants show far greater stability to rotation than FA,

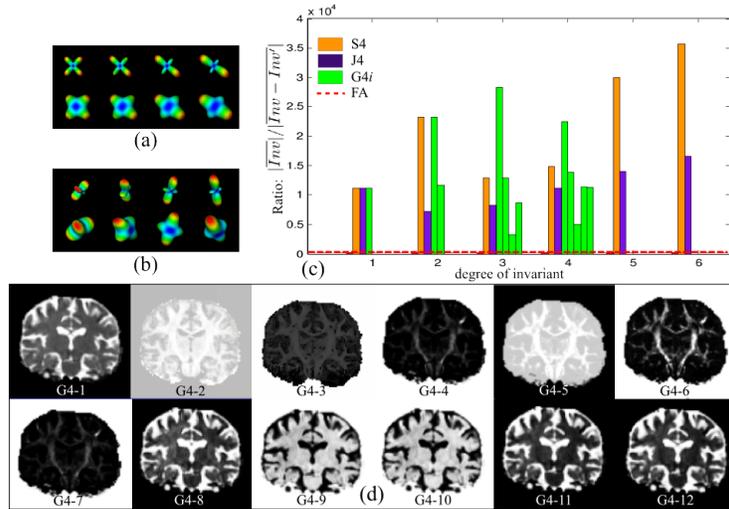
which seems to imply that the estimation of  $4^{th}$  order tensors is more stable to rotated diffusion data than the estimation of  $2^{nd}$  order diffusion tensors.

Finally we also conduct experiments on an in vivo human cerebral dataset that was acquired on a 3T Siemens scanner, with 60 gradient directions and a  $b$ -value of  $1000s/mm^2$  [10]. The resulting 12 invariants are presented in Fig. 1d.

## 4 Conclusion

We described a method to systematically generate a linearly independent set of homogeneous polynomial invariants up to degree  $d$  under  $\mathcal{SO}_3$ . This method was used to produce the invariants of  $2^{nd}$  and  $4^{th}$  order forms or tensors (for the  $4^{th}$  order only up to degree 4). The tools also allowed us to express the relations between the new invariants and those already known (basic & principal). Furthermore some of the new invariants were found to be algebraically independent of the known invariants, and many were of lower degree. Experimentally we confirmed the numerical stability of the new invariants to  $\mathcal{SO}_3$ . Future work will explore the algebraic (i.e. non linear) dependency among the new invariants and will attempt to produce a complete set of 12 invariants with simple expressions.

**Acknowledgements:** This work was partially supported by the ANR project NucleiPark and the France-Parkinson Association.



**Fig. 1.** (a, b) Synthetic data ADC with increasing crossing angles, isotropy and (b) random orientations. (c) Rotation invariance: taller bars (ratios) indicate greater numerical stability,  $G4_i$  sorted by degree (sec. 2.6). (d) 12 invariants from human dataset.

## References

1. Basser, P.J.: Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine* **8** (1995) 333–344
2. Özarslan, E., Mareci, T.H.: Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution imaging. *Magnetic Resonance in Medicine* **50** (2003) 955–965
3. Barmpoutis, A., Jian, B., Vemuri, B.C.: Symmetric Positive 4th Order Tensors & their Estimation from Diffusion Weighted MRI. In: *Information Processing in Medical Imaging (IPMI 2007)*. (2007) 308–319
4. Ghosh, A., Descoteaux, M., Deriche, R.: Riemannian framework for estimating symmetric positive definite 4th order diffusion tensors. In Metaxas, D.N., Axel, L., Fichtinger, G., Széke, G., eds.: *Medical Image Computing and Computer-Assisted Intervention-MICCAI*. Volume 5242 of *Lecture Notes in Computer Science.*, New York, NY, USA, Springer (September 2008) 858–865
5. Barmpoutis, A., Hwang, M.S., Howland, D., Forder, J.R., Vemuri, B.C.: Regularized Positive-Definite Fourth-Order Tensor Field Estimation from DW-MRI. *NeuroImage* **45**(1 sup.1) (March 2009) 153–162
6. Ghosh, A., Moakher, M., Deriche, R.: Ternary quartic approach for positive 4th order diffusion tensors revisited. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. (June 2009) 618–621
7. Basser, P.J., Pajevic, S.: Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI. *Signal Processing* **87** (2007) 220–236
8. Fuster, A., van de Sande, J., Astola, L., Poupon, C., Velterop, J., Romeny, B.M.t.H.: Fourth-order Tensor Invariants in High Angular Resolution Diffusion Imaging. In: *CDMRI Workshop – MICCAI*. Volume 6891 of *LNCS.*, Springer (2011)
9. Özarslan, E., Vemuri, B.C., Mareci, T.H.: Generalized scalar measures for diffusion mri using trace, variance and entropy. *Magnetic Resonance in Medicine* **53**(4) (2005) 866–876
10. Anwender, A., Tittgemeyer, M., von Cramon, D.Y., Friederici, A.D., Knosche, T.R.: Connectivity-Based Parcellation of Broca’s Area. *Cerebral Cortex* **17**(4) (2007) 816–825

# Anatomical Tissue Probability Priors for Tractography

Gabriel Girard and Maxime Descoteaux

Sherbrooke Connectivity Imaging Laboratory (SCIL), Computer Science Department,  
Université de Sherbrooke, Sherbrooke, Qc, Canada

**Abstract.** When visualizing streamlines produced by tractography algorithms, it is striking to observe the number of them that stop prematurely or that stop in anatomically impossible regions. We propose using the T1-weighted image tissue segmentation probability maps as *a priori* information for tractography to address this issue. A relaxation of the stopping criterion using tissue probabilities is first presented. It is then incorporated in a flexible sequential Monte Carlo modular add-on applicable to most streamline tractography algorithm. Results on high angular resolution diffusion imaging (HARDI) tractography show higher length on curved bundles or tight tracking paths, which agree better with known anatomy, for both deterministic and probabilistic algorithms.

## 1 Introduction

Tractography is the algorithmic procedure that estimates white matter (WM) fiber bundles following the local model estimated from diffusion-weighted (DW) magnetic resonance imaging (MRI). Streamline tractography outputs a sequence of tri-dimensional (3D) spatial points. Each sequence is called a streamline or tract, and represents an estimate of the link between two anatomically connected brain regions. A tractogram is the set of streamlines calculated by a tractography algorithm. Tractography algorithms are based on a set of parameters that determine which direction to follow and when to stop the tracking process.

From an initial position, the tractography algorithm follows diffusion orientations in the forward and backward directions, using a pre-determined step size ( $\Delta t$ ), until a stopping criterion is reached. Typically, stopping criteria are: i) when the tracking takes a step outside the tracking mask, ii) when the radius of curvature between two consecutive steps is smaller than a minimum radius of curvature  $R$ , or iii) when the streamline length is greater than or smaller than pre-determined lengths ( $\delta_{\min}$ ,  $\delta_{\max}$ ). The algorithm iterates over a set of initial positions (a subset of the mask, or the whole mask) or randomly generates initial positions within the tracking mask until a fixed number of streamlines are computed [1]. Stopping parameters will determine when the streamline is included or excluded of the tractogram [1]. In general, tractography is done inside a mask defined by a WM segmentation of the T1-weighted image or fractional anisotropy (FA) thresholded mask.

But how does this tracking mask affect the tractography result? Surprisingly, few works have studied the effect of the tracking mask on tractography. Guevara *et al.* [2] show that a WM mask computed from a high resolution T1-weighted image produces richer and more accurate streamlines than a thresholded FA map. The discrete binary mask is an aggressive stopping criterion and can have a dramatic impact on tractography. When streamlines are visualized, a large amount of incomplete or prematurely stopping streamlines can be observed. Additionally, these streamlines stop in anatomically impossible regions [3]. For example, it is impossible that WM fibers end in the ventricles bordering large fiber bundles such as the corpus callosum (CC).

Recently, Smith *et al.* [4] proposed a method called Anatomically Constrained Tractography (ACT), which takes advantage of the tissue segmentation maps based on the structural T1-weighted image. They proposed relaxing the stopping criterion by using WM, Gray Matter (GM) and Cerebrospinal Fluid (CSF) probability maps. Thresholds over interpolated probability maps determine if a tract stops and if it is included or excluded from the tractogram. They also proposed a back-tracking approach to find a plausible tract when the tracking algorithm fails to find one.

Our work goes in the same direction as the work proposed in [4], using the T1-weighted image tissue probability segmentations as *a priori* information for tractography. Our paper provides a novel strategy using the WM, GM and CSF probability maps and changes the way tractography stopping events are triggered by taking advantage of the complete probability map. This strategy is incorporated in a general modular Sequential Monte Carlo Tractography (SMCT) framework that can be combined to most of the local streamline tractography algorithms (both deterministic and probabilistic). We show that this relaxation of the stopping criterion enhances the density of complex (e.g. high curvature or tight WM paths) streamline bundles, eliminates streamlines that are incomplete due to premature stop, increases the average length of tract bundles and can have a positive effect on brain connectivity. All this is done with a relatively low increase in computation time.

## 2 Streamline Tractography

This study focuses on using continuous probability maps to relax the discrete binary mask for tractography stopping criterion. Since these strategies are novel in the relaxation of the stopping criterion for tractography and do not represent new tractography algorithms as such, we compared and applied these relaxations to the state-of-the-art fibre orientation distribution function (fODF) deterministic and probabilistic algorithms. In-house implementations of these tractography algorithms is used, similar to those in MRtrix [1].

**Stopping Criterion** In this study, we put the emphasis on the stopping criterion and stress that using only a binary tracking mask leads to streamlines that stop in anatomically impossible regions. Namely, tracts connecting no cortical or sub-cortical GM regions.

To overcome this effect, exclusion and inclusion binary masks combined to a WM binary tracking mask can be used [1]. Tracts exiting the tracking mask and entering the CSF mask are excluded and those entering the GM masks are included. Since every voxel is either classified as WM, GM or CSF, all included tracts exiting the WM mask are those entering the GM mask. There is one exception: if the tracking process leads out of the DWI field of view (e.g. brainstem), the tracking is stopped and the tract is included since no more data is available.

Once the tractography is done, tracts with length within the interval  $[\delta_{min} = 10mm, \delta_{max} = 300mm]$  are included in the tractogram and excluded otherwise. It is also common to see a constraint on the minimum radius  $R$  of curvature [1], limiting the high angle variations of tracts and addressing the hypothesis of smoothness of WM fibers. Thus, this can lead to situations where no tracking direction is available in the aperture cone, which causes the tracking to stop. Since this always happens within the tracking mask, these tracts are excluded.

**Parameters** The deterministic and probabilistic tractography algorithms presented use a step size  $\Delta t = 0.2mm$ , which is 1/10 of the voxel size in our DWI dataset in this work. Moreover, the minimum radius of curvature  $R = 1mm$  for probabilistic tractography and  $R = 0.26mm$  for deterministic tractography, as suggested in [1]. The WM binary mask is used as the seeding region.

### 3 Methods

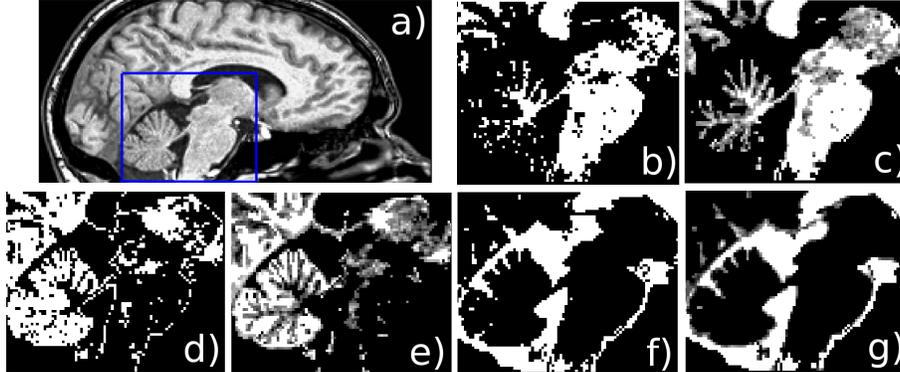
In this study, we formulate the hypothesis that a discrete binary mask (WM, GM, CSF) is too strong of a stopping criterion. It should be relaxed using continuous probability maps. This section presents 2 strategies to exploit these probability maps as relaxed stopping criterion.

#### 3.1 Continuous Maps Criterion

The T1-based FAST segmentation [5] assigns a probability of belonging to a class to each voxel. The probability at each voxel can represent an estimation of volume fraction. The resulting binary masks are the set of voxels having the highest probability for each class. Figure 1 shows the segmentation results using 3 classes (WM, GM, CSF): (b,d,f) are binary masks, (c,e,g) are probability maps. One can see that there are two major differences: i) Voxels near the boundary between distinct tissues are gray [0:black, 1:white], ii) The sub-cortical GM is mainly gray on both WM and GM maps. The discretization of these GM regions creates holes in the WM mask, which makes some WM paths tighter, making streamlines stop easily in these regions (see Figure 1 (b,d)). This problem is especially important when tracking corticospinal fibers or fibers involved in the motor system [6].

Instead of setting the probability to 1 or 0 for all classes, we propose exploiting the probabilities using a method called Continuous Maps Criterion (CMC). Let us assume the tracking process leads to the position  $pos$ . Trilinear interpolation can be done over the 3 probability maps in order to get the probability of each

**Fig. 1.** T1-weighted image (a) binary and probabilistic tissues segmentation using FAST with 3 classes: (b,c) WM, (d,e) GM and (f,g) CSF.



tissue at position  $pos$ :  $P_{pos}^{WM}$ ,  $P_{pos}^{GM}$  and  $P_{pos}^{CSF}$ . The hypothesis is that the amount of tracts passing through a voxel should be proportional to  $P_{pos}^{WM}$ . Similarly, the amount of included tracts should be proportional to  $P_{pos}^{GM}$ , and the amount of excluded tracts proportional to  $P_{pos}^{CSF}$ . This hypothesis directly uses the tissue probability in each voxel. Thus, if the segmentation algorithm cannot surely determine the class of a voxel, the CMC method will not apply thresholds to set deterministic tracking behaviors. Rather, the probability will proportionally change the tracking behaviors. The probability of continuing the tracking process is given by the equation:

$$P_{pos}^{continu} = \left( \frac{\alpha \cdot P_{pos}^{WM}}{\alpha \cdot P_{pos}^{WM} + P_{pos}^{GM} + P_{pos}^{CSF}} \right)^{\frac{\Delta t}{T1_{res}}}, \quad (1)$$

where  $\alpha$  is a weighting factor on the probability of stopping the tracking process. For example, using  $\alpha > 1$  increases the probability of tracking in low WM partial volume fraction regions. This can help to counterbalance errors in the registration or segmentation processes, such as in [6]. In this study, we compared results for  $\alpha = 1$  and  $\alpha = 4$ .

The probability maps are calculated from the T1-weighted image at resolution  $T1_{res} = 1mm$  isotropic. Using a step size  $\Delta t < T1_{res}$  will exponentially increase the probability of stopping the tractography and exponentially decrease the probability using  $\Delta t > T1_{res}$ . For example, let positions  $pos_a$  and  $pos_b$  be at a distance of  $2mm$  from each other, and  $P^{WM}$  uniformly equal to 0.8. If the tractography starts at  $pos_a$  and ends at  $pos_b$ , the probability of stopping with  $\Delta t = 1mm$  is  $1 - 0.8^2 = 0.36$ , using  $\Delta t = 0.5mm$  this probability becomes  $1 - 0.8^4 \approx 0.59$ , and using  $\Delta t = 2mm$  the probability is  $1 - 0.8 = 0.2$ . The hypothesis of the amount of tracts passing through a voxel is thus not respected with  $\Delta t \neq T1_{res}$ . The step size has a strong influence on the tract propagation using CMC. In order to respect the previous hypothesis, the probability of con-

---

**Algorithm 1** Continuous Maps Criterion Algorithm

---

**Require:**  $pos \leftarrow$  *New Tracking Position*  
**Require:**  $P_{pos}^{continu} \leftarrow$  (Eq.1)  
**Require:**  $P_{pos}^{include} \leftarrow$  (Eq.2)  
 $x_1, x_2 \sim U(0, 1)$  (*uniform distribution*  $[0, 1]$ )  
**if**  $x_1 < P_{pos}^{continu}$  **then**  
    **return** *continue tracking*  
**else**  
    **if**  $x_2 < P_{pos}^{include}$  **then**  
        **return** *stop tracking and include*  
    **else**  
        **return** *stop tracking and exclude*  
    **end if**  
**end if**

---

tinuing the tracking process (Eq. 1) is adjusted to give the same tracking effect with a smaller (or bigger)  $\Delta t$ .

If the tracking process stops, the probability of including the tract in the tractogram is given by the equation:

$$P_{pos}^{include} = \frac{P_{pos}^{GM}}{P_{pos}^{GM} + P_{pos}^{CSF}}, \quad (2)$$

which represents the possibility of an anatomical connection stopping at  $pos$ . Algorithm 1 outlines the CMC method, determining whether the tracking continues or stops. In case of stopping, the algorithm determines if the tract is included or excluded of the tractogram.

### 3.2 Sequential Monte Carlo Framework

The streamline tractography can be modeled as a state system evolving over time using noisy measurements (DWI, T1, etc.), where states are both the tracking position and the tracking status (for example in the WM or stopped in the GM), and are connected over time by a Markov chain. We propose to use a Sequential Monte Carlo (SMC) framework based on the previous state model to estimate the uncertainty in a tractography stopping event. The SMC or particle filter model has been widely used for robot localization [7] using sensor measurements to estimate the robot position. Recently, it has also been used for WM tractography [8, 9]. This section presents the SMC method followed by its implementation in the context of tractography and probability maps.

**SMC Method** SMC methods aim to estimate a sequence of target state variables  $X_{0:t} = \{X_k, k = 0, \dots, t\}$  from a sequence of observation variables  $Y_{0:t} = \{Y_k, k = 0, \dots, t\}$ . The goal is to sequentially estimate the posterior distribution  $p(X_k|Y_{0:k})$ .  $X_{0:t}$  is a first order Markov process such that  $X_k|X_{k-1} \sim p(X_k|X_{k-1})$  with a known initial distribution  $p(X_0)$  and  $Y_{0:k}$  are conditionally independent if  $X_{0:k}$  are known. The key idea of SMC is to represent the posterior

distribution  $p(X_{0:t}|Y_{0:t})$  by a set of random samples with associated weights and compute estimates based on the samples and weights [7, 10].  $\{x_{0:k}^i, w_t^i\}_{i=1}^N$  denotes the set of  $N$  discrete random samples that characterize the posterior distribution, where  $\{x_{0:k}^{(i)}, i = 1, \dots, N\}$  is the set of random samples,  $\{w_k^{(i)}, i = 1, \dots, N\}$  their associated weights with  $\sum_{i=1}^N w_k^{(i)} = 1$ . The weight of a sample  $x_k^{(i)}$  at time  $k$  corresponds to its weight at time  $k - 1$  times the likelihood of the observation  $y_k^{(i)}$  (see Eq. 5). Such a discrete model suffers of degeneracy since the variance of the weights increases over time, leading to a situation where all samples except one have a weight close to zero. To overcome this problem, the number of samples can be increased or a resampling method can be applied when a significant degeneracy is observed. The latter has been chosen in this study. The degeneracy problem can be observed when the number of effective samples  $N_{eff}$ , as described in Eq. 3, falls below some threshold  $N_T$  [10].

$$N_{eff} = 1 / \sum_{i=1}^N (w_k^{(i)})^2. \quad (3)$$

The resampling idea is to eliminate samples with small weights and concentrate on samples that have large weights. The resampling generates  $N$  new samples with equal weights from the current discrete estimation of  $p(X_{0:k}|Y_{0:k})$ . This method is also known as Sequential Importance Sampling particle filter [7, 10].

**SMC Tractography** This section presents a Sequential Monte Carlo Tractography (SMCT) framework to estimate a likely tract using probability maps whenever the tractography reaches a stopping criterion. The key idea is to back track  $K_b$  steps and compute a better likely tract after  $K = K_b + K_f$  steps, where  $K_b$  and  $K_f$  are respectively the number of backward and forward steps. If there are less tracking steps done than  $K_b$ ,  $K_b$  is set to the number of tracking steps done so far. The goal is to estimate a likely tract initialized with  $K_b$  steps before the stopping criterion is reached, and then go  $K_f$  forward steps to ensure the local stopping event is solved. That is, the tract stops correctly in the GM or the tract continues its propagation in the WM. If the tract stops in the GM, the tracking is done. If the tract is in the WM, the tractography continues normally until another stopping criterion is reached.

In the context of tractography  $x_k = [pos_k, status_k]$ , where  $pos_k$  is the tracking position at time  $k$  and  $status_k \in \{active, inactive\}$  represents the tracking process propagating in the WM (*active*) or stopped in the GM (*inactive*). If  $status_k = active$  then  $pos_{k-1}$  propagates to  $pos_k$  following a probabilistic method, otherwise the tracking reaches the GM and is stopped ( $pos_k = pos_{k-1}$  in this case). The  $status_k$  becomes *inactive* following equation:

$$P_{pos}^{inactive} = \left( \frac{P_{pos}^{GM}}{P_{pos}^{GM} + \alpha \cdot P_{pos}^{WM}} \right)^{\frac{\Delta t}{T_{1res}}}, \quad (4)$$

otherwise, it remains constant over time ( $status_k = status_{k-1}$ ). The probability of CSF is the observation  $Y_k$  in our model, thus, the weight  $w_k$  of a tract at position  $p_k$  is computed following equation:

$$w_k^{(i)} = w_{k-1}^{(i)} \cdot p(y_k^{(i)} | x_k^{(i)}) = w_{k-1}^{(i)} \cdot (1 - P_{pos_k}^{CSF})^{\frac{\Delta t}{T1res}}. \quad (5)$$

Equations 4 and 5 are related to  $\Delta t$  to ensure the step size does not change the weighting and stopping event of the algorithm, as it is in section 3.1. The SMCT uses a fixed number  $N$  of samples to estimate  $p(X_k)$ . The weight  $w_k^{(i)}$  of each sample  $x_k^{(i)}$  at time  $k$  is normalized over all samples to ensure that  $\sum_{i=1}^N w_k^{(i)} = 1$ . The resampling is done when  $N_{eff} < N/10$ , see Eq. 3.

SMCT will estimate the valid tract distribution around the stopping event and iteratively learn subsequent valid tract distribution from the previous one. The resulting tract is drawn from the final valid tract distribution. The key elements of the algorithm are:

1. Whenever a stopping criterion excluding a tract is reached, set  $p(X_{k=0})$  to  $N$  samples with position  $pos_0^{(i)}$  set to the  $K_b$  previous tracking position, its  $status_0^{(i)} = active$  and  $w_0^{(i)} = 1/N$ . Thus, all samples have the same initial position and weight.
2. Propagate the  $N$  samples from  $p(X_{k=0})$ , the initial tracts distribution, using a probabilistic method.
3. Update  $w_k^{(i)} \forall x_k^{(i)}$  samples using Eq. 5 and normalize  $w_k^{(i)}$  in order to have that  $\sum_{i=1}^N w_k^{(i)} = 1$ .
4. Resample the distribution if  $N_{eff} < N/10$ .
5. Propagate the *status* of each drawn sample from the previous iteration. The status becomes *inactive* following Eq. 4.
6. Update the position  $pos_k^{(i)} \forall x_k^{(i)}$  samples using a probabilistic method if the  $status_k^{(i)} = active$ , otherwise  $pos_k^{(i)} = pos_{k-1}^{(i)}$ .
7. Repeat steps 3 to 6  $K$  times,  $k = k + 1$ .
8. Draw a single sample from  $p(X_K)$ . If  $status = active$ , the tracking continues using the principal tractography algorithm, otherwise the tracking is done and the tract is included in the tractogram.

This algorithm generates multiple probabilistic tracts, encouraging tracts to follow the WM or end in the GM, and discouraging tracts to propagate in CSF. The output of the SMCT is either an inactive tract ending in the GM or an active tract in the WM. Otherwise, if at any iteration  $k$  the weights  $w_k^{(i)} = 0 \forall x_k^{(i)}$  samples, the tract is excluded from the tractogram because no valid tract is found (e.g.  $P_{pos_k}^{CSF} = 1 \forall x_k^{(i)}$ ).

In our implementation, we used both the fODF deterministic and probabilistic tractography algorithms as principal algorithm and the probabilistic fODF algorithm within the SMCT, see section 2 for details on these algorithms. The principal tractography algorithm is done until the tracking reaches a stopping criterion excluding the tract, as determined by the CMC method of section 3.1 or if there is no valid tracking direction. Whenever it happens, the SMCT is used. When no valid tracking direction is found for a sample within SMCT, the sample  $x_k^{(i)}$  is still in the WM, but cannot propagate further, thus his weight  $w_k^{(i)} = 0$ . Other parameters are  $N = 100$  samples,  $K_b = 10$  and  $K_f = 5$ , which corresponds to the tracking distance of 2 voxels in T1-weighted space, respectively 1

voxel using a  $\Delta t = 0.2mm$ . Estimating likely tracts after a tracking distance of three  $T1_{res}$  voxels is sufficient to find a plausible path for many tracts previously excluded due to stopping criterion. It also keeps computation time low.

## 4 Dataset

Diffusion-weighted images were acquired on a single volunteer along 64 uniformly distributed directions using a b-value of  $b = 1000s/mm^2$  and a single  $b = 0s/mm^2$  image using the single-shot echo-planar imaging (EPI) sequence on a 1.5 Tesla SIEMENS Magnetom (128x128 matrix, 2mm isotropic resolution, TR/TE 11000/98 ms and GRAPPA factor 2). An anatomical T1-weighted 1mm isotropic MPRAGE (TR/TE 6.57/2.52 ms) image was also acquired.

Diffusion data was first corrected for eddy currents and head motion using FSL ([www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl)), then upsampled to 1mm isotropic resolution using a trilinear interpolation [1, 11]. Diffusion Tensor (DT) estimation and corresponding Fractional Anisotropy (FA) map generation were done using MRtrix [1]. From this, the single fiber response function was estimated from all FA values above a threshold of 0.7, within the WM binary mask. This single fiber response was used as input to spherical deconvolution [12, 13] to compute the fiber orientation distribution function (fODF), with spherical harmonic order 8, at every voxel of the brain. In this work, we used the efficient implementation publicly available in MRtrix.

The T1-weighted image was registered to a 1mm isotropic DWI using *FLIRT*. The Brain Extraction tool (BET) and *FAST* [5] of FSL were also used to extract both binary and probabilistic maps of the WM, GM and CSF.

## 5 Results and discussion

We compared the results using binary mask, CMC and the recently proposed method ACT [4]. ACT is adapted to only use 3 maps (WM, GM, CSF) and not use the subcortical GM criterion to compare with our proposed methods. The results of 12 tractograms each containing 50,000 tracts are discussed: both deterministic (Det.) and probabilistic (Prob.) tractograms using a binary mask (bin.), ACT, CMC and SMCT. CMC\* and SMCT\* denote the use of  $\alpha = 4$ .

Table 1 shows the proportion of seeds resulting in included tracts in the tractograms as well as the proportion of tracts excluded due to the stopping criterion and length criterion. The use of CMC with the fODF deterministic tractography decreases the proportion of tracts excluded due to the stopping criterion by 10% with respect to its binary version and decreases this proportion by 36% using SMCT (10% and 39% respectively using fODF Prob. CMC and SMCT). Table 1 shows high percentage of excluded tract for tractography using ACT. We hypothesize that this is mainly due to not using subcortical GM maps, which, after visualization, can cause tract to propagate through subcortical GM and end in the CSF. Nonetheless, we compared our results to ACT.

Methods	Seeding			Tracts							
	% incl.		% excl.	Whole		PYT		Cg		SLF	
	Incl.	stop.	$\delta$	#	$\delta_\mu$ (mm)	#	$\delta_\mu$ (mm)	#	$\delta_\mu$ (mm)	#	$\delta_\mu$ (mm)
Det. bin.	20	57	24	50,000	38	662	80	417	76	1,025	106
Det. ACT	17	69	14	50,000	43	1,197	<b>99</b>	597	81	968	<b>113</b>
Det. CMC	23	47	30	50,000	38	714	89	373	76	1,030	105
Det. CMC*	25	50	25	50,000	39	979	<b>94</b>	428	82	894	<b>112</b>
Det. SMCT	<b>45</b>	21	34	50,000	<b>47</b>	958	<b>93</b>	357	<b>89</b>	1,527	<b>114</b>
Det. SMCT*	<b>50</b>	23	28	50,000	<b>49</b>	1,426	<b>98</b>	430	<b>92</b>	1,439	<b>119</b>
Prob. bin.	25	59	16	50,000	39	647	80	450	75	859	105
Prob. ACT	16	70	14	50,000	44	1,213	<b>96</b>	600	<b>83</b>	859	111
Prob. CMC	22	49	29	50,000	38	724	88	378	71	811	105
Prob. CMC*	23	51	21	50,000	39	1,013	<b>93</b>	429	77	780	107
Prob. SMCT	<b>45</b>	20	35	50,000	<b>47</b>	856	<b>92</b>	267	<b>83</b>	1,406	<b>115</b>
Prob. SMCT*	<b>51</b>	22	27	50,000	<b>50</b>	1,416	<b>97</b>	365	<b>86</b>	1,352	<b>118</b>

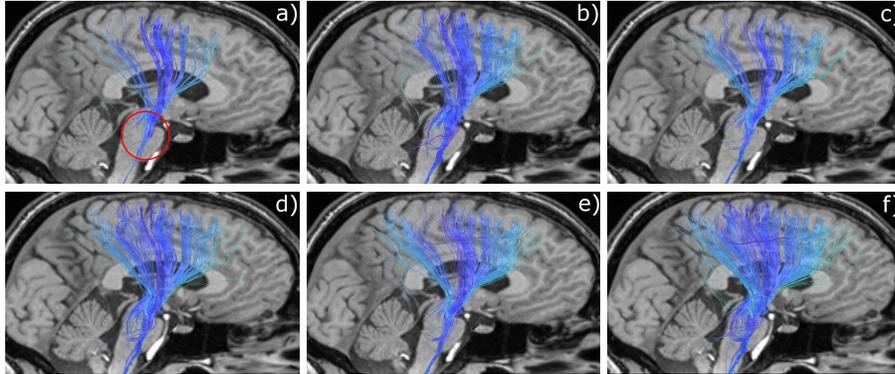
**Table 1.** The proportion of seeds resulting in included (% incl.) or excluded (% excl.) tracts, and the number of tracts in the whole tractogram, the pyramidal tracts (PYT), the cingulum (Cg) and the superior longitudinal fasciculus (SLF) with their average tract length ( $\delta_\mu$ ) for the 12 tractograms. The excluded tracts are split in 2 categories: i) excluded by the stopping criterion (stop.), ii) excluded by the  $\delta_{min}$  or  $\delta_{max}$  length criterion ( $\delta$ ). Bundles were obtained using manual ROI selection from whole tractograms.

Seeds resulting in tracts near complex tracking configurations (high curvature or tight WM paths) have more chance of being excluded due to stopping criterion and thus, an increase of the included tracts could provide a better representation of the actual WM fiber distribution. Note that this does not address the seeding issue of long bundles that traverse more voxels and result in an over estimation of longer bundles compared to shorter bundles with similar fiber density.

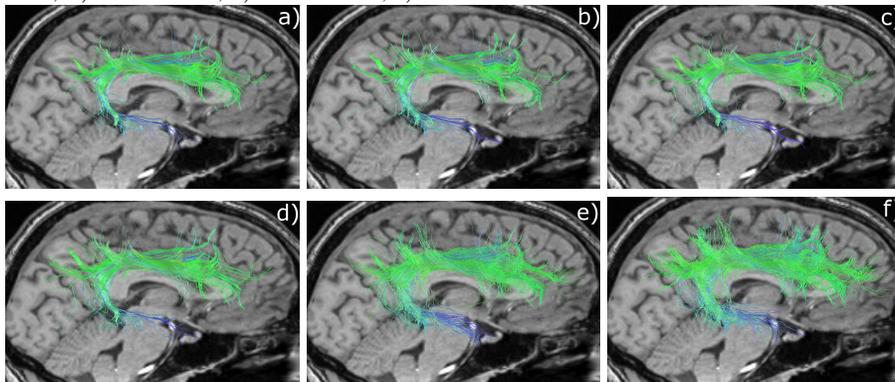
Table 1 also provides the tract density and average tract length of the whole tractogram, the pyramidal tracts (PYT), the cingulum (Cg) and the superior longitudinal fasciculus (SLF). The CMC has little effect on the density and length of tracts. ACT and CMC\* provide results with longer length than bin. and CMC. We hypothesize that this is due to the wider tracking area defined by the stopping criterion of ACT and the relaxed stopping probability of CMC\*. SMCT, and especially SMCT\*, also increase the average tract length of the 3 selected tract bundles and increase the tract density of the SLF.

One can see from Figure 2 (a) that most of the tracts stop early when entering in the brainstem. This is mainly due to the binarization of the probability maps (see Figure 1), which cuts several pathways through the brainstem. ACT, CMC and SMCT using the full probability maps can be observed in Figure 2 (b,c,d,e,f), where tracts more efficiently traverse the brainstem and some exit the image field of view, as expected. The use of SMCT provided a more uniform and dense PYT, providing longer tracts and producing qualitatively the best PYT reconstruction. We observed the same effect on probabilistic results (not shown due to space restriction).

**Fig. 2.** The 6 deterministic tractograms of the PYT. a) det. bin, b) det. ACT, c) det. CMC, d) det. CMC\*, e) det. SMCT, f) det. SMCT\*.



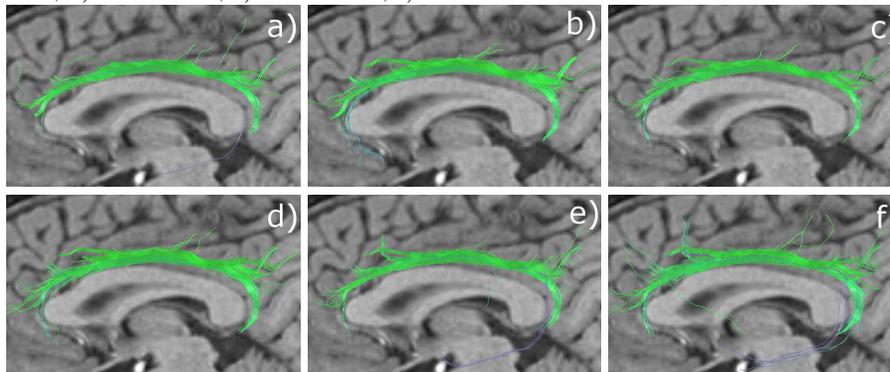
**Fig. 3.** The 6 deterministic tractograms of the SLF. a) det. bin, b) det. ACT, c) det. CMC, d) det. CMC\*, e) det. SMCT, f) det. SMCT\*.



SMCT increases tract densities on complex tract configurations as shown on the SLF in Figure 3 and expressed quantitatively in Table 1. SMCT also produces longer tracts than other methods on the Cg bundle (Figure 4), finding the full curving extent of the Cg. By helping the tracking of complex bundles, SMCT not only contributes to a more uniform tract density but also to find new tracts or longer paths that qualitatively agree better with known anatomy.

Finally, the back-tracking approach proposed by Smith *et al.* [4] was tested and showed a general increase in the number of seeds resulting in included tracts in the tractograms using binary mask, ACT or CMC for both probabilistic and deterministic tractography. This incremental back-tracking idea offers even more relaxation of the stopping criterion. We believe that combined to SMCT, it could potentially increase the general quality of the tractography. This issue will be further investigated and quantified in the future.

**Fig. 4.** The 6 deterministic tractograms of the Cg. a) det. bin, b) det. ACT, c) det. CMC, d) det. CMC\*, e) det. SMCT, f) det. SMCT\*.



**Computational Cost** Using our dataset, the SMCT was called on average 1.1 times per seed, resulting in an average computation time increase by a factor of 4 to obtain the same total number of tracts in the tractogram.

## 6 Conclusion

The results show that the use of the full probabilistic tissue segmentation maps increases the quality of tractograms compared to tractography done with binarized mask versions. Even more importantly, tissue probability maps allow the integration of a general framework solving local stopping issues by providing a likely plausible path when the tractography algorithm fails to find one. The SMCT takes advantages of global tractography algorithms [8, 9] and applies these advantages in a local manner. The results show that this approach increases the proportion of seeds resulting in actual tracts, which could be beneficial for connectivity studies. The SMCT combined with a fODF deterministic algorithm provides a hybrid tractography taking advantage of the regularity of the deterministic algorithm and richness of the probabilistic scheme. In our opinion, this gives a promising extension for deterministic tractography.

We think the CMC approach offers a more accurate way of dealing with the tractography mask than the use of a binary mask. This approach is of course dependent on the segmentation algorithm, as any method based on WM segmentation. FAST [5] provides good results, but other tools should be investigated, such as SPM ([www.fil.ion.ucl.ac.uk/spm/](http://www.fil.ion.ucl.ac.uk/spm/)). In particular, it would be of interest to study how CMC and ACT methods are sensitive to noise and perturbations, and how these methods can deal with less reliable tissue probability maps. Furthermore, the use of subcortical GM could be incorporated to CMC in a similar way as it proposed by Smith *et al.* [4]. Although CMC uses a probabilistic strategy and ACT uses a dynamic thresholding strategy, they both showed similar results in this study.

Since SMCT is a flexible modular add-on to most of the local streamline tractography algorithms, its implementation based on the DT is of interest, as DT tractography is still most commonly used in neuroscience applications. It could also be quite interesting to see how the CMC could be integrated to the particle filtering algorithm of Zhang *et al.* [8] and Pontabry *et al.* [9]. Conversely, it could be interesting to integrate the ingredients of their work on particle filtering for WM tractography to the SMCT framework.

One of the next challenges in tractography is to properly use *a priori* information from anatomy to improve the results and perform better brain connectivity. We think SMCT and CMC are a step forward to address this challenge.

## References

1. Tournier, J.D., Calamante, F., Connelly, A.: MRtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology* **22**(1) (2012) 53–66
2. Guevara, P., Duclap, D., Marrakchi-Kacem, L., Rivière, D., Cointepas, Y., Poupon, C., Mangin, J.F.: Accurate tractography propagation mask using t1-weighted data rather than fa. In: *Proc. ISMRM*. (2011) 2018
3. Côté, M.A., Boré, A., Girard, G., Houde, J.C., Descoteaux, M.: Tractometer: Online evaluation system for tractography. In: *Proc. MICCAI*. Volume 15. (2012)
4. Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A.: Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage* (2012)
5. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging* **20**(1) (2001) 45–57
6. Girard, G., Chamberland, M., Houde, J.c., Fortin, D., Descoteaux, M.: Neurosurgical tracking at the Sherbrooke Connectivity Imaging Lab ( SCIL ). In: *DTI Tractography Challenge - MICCAI Workshop*. (2012)
7. Doucet, A., de Freitas, N., Gordon, N.: *Sequential Monte Carlo Methods in Practice*. Springer (2001)
8. Zhang, F., Hancock, E.R., Goodlett, C., Gerig, G.: Probabilistic white matter fiber tracking using particle filtering and von Mises-Fisher sampling. *Medical image analysis* **13**(1) (2009) 5–18
9. Pontabry, J., Rousseau, F.: Probabilistic tractography using Q-ball modeling and particle filtering. In: *Proc. MICCAI*. Volume 14. (2011) 209–216
10. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A Tutorial on Particle Filters for Online Nonlinear / Non-Gaussian Bayesian Tracking. *IEEE Transaction on Signal Processing* **50**(2) (2002) 174–188
11. Dyrby, T., Lundell, H., Liptrot, M., Burke, W., Ptito, M., Siebner, H.: Interpolation of DWI prior to DTI reconstruction , and its validation. In: *Proc. ISMRM*. (2011)
12. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**(4) (2007) 1459–1472
13. Descoteaux, M., Deriche, R., Knösche, T.R., Anwander, A.: Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions in Medical Imaging* **28**(2) (2009) 269–286

# White matter structure assessment from reduced HARDI data using low-rank polynomial approximations

Yaniv Gur<sup>\*</sup>, Fangxiang Jiao<sup>\*</sup>, Stella Xinghua Zhu<sup>†</sup> and Chris R. Johnson<sup>\*</sup>

<sup>\*</sup> SCI Institute, University of Utah, 72 S. Central Campus Dr., SLC, UT 84112, USA

<sup>†</sup> Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong

**Abstract.** Assessing white matter fiber orientations directly from DWI measurements in single-shell HARDI has many advantages. One of these advantages is the ability to model multiple fibers using fewer parameters than are required to describe an ODF and, thus, reduce the number of DW samples needed for the reconstruction. However, fitting a model directly to the data using Gaussian mixture, for instance, is known as an initialization-dependent unstable process. This paper presents a novel direct fitting technique for single-shell HARDI that enjoys the advantages of direct fitting without sacrificing the accuracy and stability even when the number of gradient directions is relatively low. This technique is based on a spherical deconvolution technique and decomposition of a homogeneous polynomial into a sum of powers of linear forms, known as a *symmetric tensor decomposition*. The fiber-ODF (fODF), which is described by a homogeneous polynomial, is approximated here by a discrete sum of even-order linear-forms that are directly related to rank-1 tensors and represent single-fibers. This polynomial approximation is convolved to a single-fiber response function, and the result is optimized against the DWI measurements to assess the fiber orientations and the volume fractions directly. This formulation is accompanied by a robust iterative alternating numerical scheme which is based on the Levenberg-Marquardt technique. Using simulated data and in vivo, human brain data we show that the proposed algorithm is stable, accurate and can model complex fiber structures using only 12 gradient directions.

## 1 Introduction

In contrast to diffusion tensor imaging (DTI), High Angular Resolution Diffusion Imaging (HARDI) is an imaging technique that is capable of describing complex white matter structures such as crossing fibers. Given HARDI data, various reconstruction techniques are used to infer the fiber structures [1-6]. These techniques are primarily based on the reconstruction of an orientation distribution function (ODF) that describes the dominant diffusion directions. To recover the white matter fiber pathways, the dominant diffusion directions are extracted from the ODF. Since white matter connectivity maps are obtained

from tracking these directions, an accurate reconstruction of this information is crucial. This motivated the development of various analytical and numerical techniques to achieve this task. These techniques are mainly based on polynomial root-finding and high order ODF tessellation [7–9], or low-rank tensor approximations [10, 11]. However, the accuracy of these algorithms is limited by the ODF quality of reconstruction and its reconstruction order (i.e., the spherical harmonics truncation order). Also, since these algorithms introduce significant complexity, the complete process of ODF reconstruction, followed by orientations estimation, is inefficient. Multi-compartment models [3, 12–14], however, avoid the ODF estimation step by estimating the fiber parameters directly from the DWI measurements. This allows modeling multiple fiber orientations using a few number of parameters and, thus, reduce the number of DWI measurements needed for the reconstruction.

However, these models have two main disadvantages: First, to obtain the best results the correct number of fiber compartments has to be pre-selected. As was pointed out in [13], in single-fiber voxels the model will lose accuracy if fitting two-compartment models to the data. Second, the resulting non-linear optimization problem is unstable and initialization dependent since the objective function possesses local minima.

In this paper we present an alternative direct estimation technique that enjoys the advantages of direct fitting without sacrificing the stability, accuracy and robustness to noise of the algorithm. In addition, it allows accurate estimation of the orientations even in a case of over-fitting. This technique is based on spherical deconvolution, which is a powerful technique for modelling complex fiber structures by means of a fiber-ODF (fODF) [2]. It is known that the maximal accuracy of spherical deconvolution is achieved when the fODF is decomposed into rank-1 tensors that represent single fiber orientations [10, 11, 15]. Thus, the fODF estimation step is followed by a tensor decomposition. In this work we show that using rank-1 tensor fiber representations as linear-forms, these two distinct steps can be combined into one optimization problem that allows robust estimation of the fiber parameters directly from the DWI measurements.

The proposed approach is motivated by the *symmetric tensor decomposition* [16], that is, any homogeneous polynomial of order  $d$  may be decomposed into a sum of  $r$  distinct linear-forms of the same order. Since any spherical function with antipodal symmetry may be represented as an even-order homogeneous polynomial (or a symmetric higher-order tensor) [17], we can decompose an ODF or a fODF in a similar manner. Thus, we consider here a lower-rank polynomial approximation of a fODF in terms of even order linear-forms which are directly related to rank-1 tensors. In this approximation, each linear-form represents a single fiber and its coefficients correspond directly to the fiber orientation and the volume fraction (the mixing parameter). Similar to existing multi-compartment models, the fODF expansion in linear-forms is naturally positive-definite, and hence, no additional constraint that guarantees this property is required. The expansion’s coefficients are estimated via a spherical deconvolution operation such that each term is convolved to a single-fiber response and the result is opti-

mized against the HARDI measurements by means of the  $l_2$  norm. The resulting non-linear optimization problem is solved here using a novel iterative alternating scheme based upon the Levenberg-Marquardt technique and is shown to produce stable and accurate results.

In this paper we test the algorithm on simulated data as well as in vivo, human brain data. In both cases the set of gradient directions was sub-sampled from 96 (or 64 for the human brain data) down to 12 so we could explore the limitations of the algorithm and the decline in performance. We show that, in both cases (simulated and real data), sub-sampling from 64 to 32 directions does not change significantly the results. A performance decline is clearly observed when the set contains only 12 gradient directions, yet, the algorithm produces useable results and can reliably separate fibers crossing at 75 degrees and above.

This paper is organized as follows: In Sec. 2 we briefly review the spherical deconvolution approach and develop the new method in this context. In Sec. 3 we discuss the numerical optimization technique that we developed to solve the minimization problem. Finally, Sec. 4 is devoted to accuracy and stability studies and experiments on in-vivo, human brain data.

## 2 Spherical deconvolution via symmetric tensor decomposition

Spherical deconvolution is a common technique to recover major diffusion directions from DWI data [2]. It is based on a convolution between a spherically symmetric function, known as fODF, and an axially symmetric kernel that represents a single fiber response. Given a vector of  $n$  DWI measurements in the gradient directions, the fODF, denoted by  $F$ , is reconstructed by solving the following deconvolution problem:

$$\min_F \frac{1}{2} \sum_{i=1}^n \left\| S(\mathbf{g}_i, b) - S_0 \int_{S^2} F(\mathbf{v}) K(\mathbf{g}_i, \mathbf{v}) d\mathbf{v} \right\|^2. \quad (1)$$

This problem is solved for a fixed kernel,  $K$ , where its width is adjusted to the particular dataset. The resulting fODF represents a sum of spherical delta functions aligned with the fiber orientations and weighted by the volume fractions. This basic problem is solved by means of least-squares where the fODF is reconstructed by a pseudo-inverse operation. However, additional constraints such as fODF positivity leads to non-linear optimization problem [2].

In [16] it was shown that any homogeneous polynomial of order  $d$  may be decomposed into a sum of linear-forms of the same order such that:

$$F(x_1, x_2, \dots, x_l) = \sum_{i=1}^r \lambda_i f_i^d \quad (2)$$

where  $f_i = (\sum_{j=1}^l \alpha_j x_j)$ ,  $r$  is the polynomial *rank* and  $l$  is the polynomial *dimension*. This decomposition is known as *symmetric tensor decomposition* since

homogeneous polynomials are directly related to symmetric tensors. An algorithm to decompose a general homogenous polynomial was proposed in [16].

It is known that any spherical function with antipodal symmetry may be represented as an even-order homogeneous polynomial, where its order is equivalent to the truncation order of the corresponding spherical harmonics expansion [17]. Since a fODF may be represented as a homogenous polynomial, one may use [16] to compute its full-rank decomposition. However, a full-rank fODF encodes information on white matter fibers, as well as noise. Thus, it was proposed in [10, 11] to recover the fiber orientations via a lower-rank tensor approximation. This approximation was applied to the fODF and required its estimation first.

To combine the fODF reconstruction and the orientations estimation into one optimization problem, we first approximate the fODF using an equivalent lower-rank approximation by means of polynomial approximation (symmetric tensor decomposition) such that:

$$F(\mathbf{v}) \sim \sum_{i=1}^{\tilde{r}} \gamma_i f_i^d = \sum_{i=1}^{\tilde{r}} (\boldsymbol{\alpha}_i \cdot \mathbf{v})^d, \quad \tilde{r} < r, \quad (3)$$

where  $\boldsymbol{\alpha}_i \in \mathbb{R}^3$ ,  $\mathbf{v} \in S^2$  and each fiber aligned in direction  $\boldsymbol{\alpha}_i$  is identified with a linear form  $(\boldsymbol{\alpha}_i \cdot \mathbf{v})^d$ . The number of fibers to be estimated is determined by the approximation rank  $\tilde{r}$  and the expansion coefficients are defined as  $\gamma_i = \|\boldsymbol{\alpha}_i\|^d$ .

Next, we substitute (3) into (1). This leads to the following non-linear optimization problem:

$$\min_{\boldsymbol{\alpha}_j} \frac{1}{2} \sum_{i=1}^n \left\| S(\mathbf{g}_i, b) - S_0 \int_{S^2} \sum_{j=1}^{\tilde{r}} (\boldsymbol{\alpha}_j \cdot \mathbf{v})^d K(\mathbf{g}_i, \mathbf{v}) d\mathbf{v} \right\|^2. \quad (4)$$

This problem is solved for the coefficients of the linear-forms, three coefficients for each fiber, which are directly estimated from the DWI measurements. The fiber orientations and the volume fractions are derived as follows: Since each linear-form gets its maximum at the direction specified by  $\boldsymbol{\alpha}_j$ , given the optimal solution,  $\tilde{\boldsymbol{\alpha}}_j$ , the corresponding fiber orientation is simply  $\mathbf{u}_j = \frac{\tilde{\boldsymbol{\alpha}}_j}{\|\tilde{\boldsymbol{\alpha}}_j\|}$ . As we do not impose the constraint  $\sum_{j=1}^{\tilde{r}} \|\boldsymbol{\alpha}_j\|^d = 1$ , the corresponding volume fraction is given by  $w_j = \frac{\|\tilde{\boldsymbol{\alpha}}_j\|^d}{\sum_{j=1}^{\tilde{r}} \|\tilde{\boldsymbol{\alpha}}_j\|^d}$ .

### 3 Numerical optimization

To solve this non-linear optimization problem (4) we adopt the Levenberg-Marquardt (LM) technique. When  $\tilde{r} = 1$ , the three coefficients can be estimated accurately using a straightforward implementation of the LM. However, when  $r > 1$ , more coefficients are involved and estimating them at once provides poor results. To deal with the multi-fiber estimation case, we suggest an iterative alternating LM scheme. In this scheme, a complete update step is composed of  $\tilde{r}$

LM sub-steps. In each sub-step only the coefficient associated with a single fiber are updated while the other coefficients are kept fixed as described in Algorithm 1. In each iteration, one has to convolve the fiber estimate to the kernel. This operation is performed using a discrete spherical integration scheme [18]. In terms of convergence, we have found that the algorithm is very robust and converges for any initial guess.

---

**Algorithm 1** Alternating LM for  $\tilde{r} = 2$

---

- 1: Let  $I$  be the objective function defined in 4, and let  $J_k = \frac{\partial I}{\partial \alpha_k}$ ,  $k = 1, 2$ .
  - 2: Set  $t = 0$ .
  - 3: Initialize  $\alpha_k^t$ ,  $k = 1, 2$ .
  - 4: Compute  $\alpha_1^{t+1}$  using an LM update with respect to  $J_1(\alpha_1^t, \alpha_2^t)$  and a damping parameter  $\epsilon_1$ .
  - 5: Compute  $\alpha_2^{t+1}$  using an LM update with respect to  $J_2(\alpha_1^{t+1}, \alpha_2^t)$  and a damping parameter  $\epsilon_2$ .
  - 6: **if** converged **then**
  - 7: return  $\alpha_{1,2}^t$
  - 8: **else**
  - 9:  $t \leftarrow t + 1$
  - 10: goto 4
  - 11: **end if**
- 

## 4 Simulations

### 4.1 Synthetic data

To test the accuracy and stability of the algorithm we simulated two crossing fibers at 4 separation angles: 45, 60, 75 and 90, equal volume fractions and two b-values:  $b = 1500s/mm^2$  and  $b = 3000s/mm^2$ . The signal was simulated using the multi-tensor model:

$$S(\mathbf{g}_i, b) = S_0 \sum_{j=1}^2 w_j \exp(-b \mathbf{g}_i^T D_j \mathbf{g}_i) \quad (5)$$

where for each compartment we assume a prolate tensor with FA=0.8. The simulated signal was corrupted by Rician noise distribution as follows:

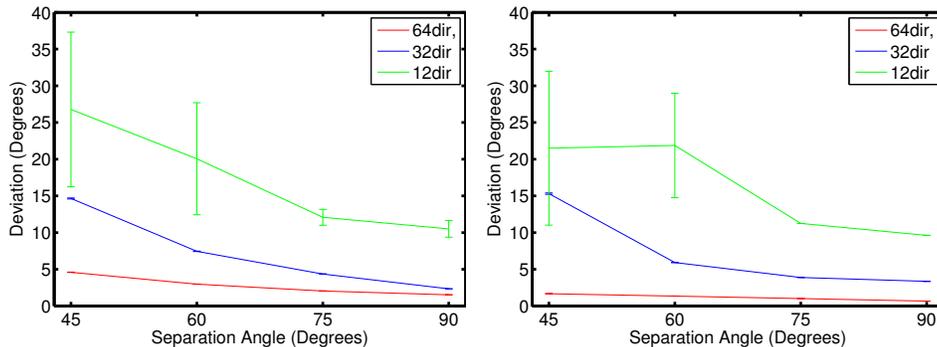
$$S_{\text{noisy}}(\mathbf{g}_i, b) = \sqrt{(S(\mathbf{g}_i, b) + n_1)^2 + n_2^2}. \quad (6)$$

where  $n_1, n_2 \sim \mathcal{N}(0, \sigma^2)$  and  $\sigma = \frac{S_0}{\text{SNR}}$ .

For all of the experiments presented below the SNR was set to 20 and for each separation angle the performance of the algorithm was evaluated on 200 noise realizations. The polynomial order is set to  $d = 8$  as this value gives an

optimal trade-off between the ability to resolve low separation angles and noise sensitivity at this SNR [2]. The single-fiber response kernel is described here by the Watson function  $K(\mathbf{g}_i, \mathbf{v}, \delta) = \exp(-\delta(\mathbf{g}_i^T \cdot \mathbf{v})^2)$  where  $\delta$  is a function of the  $b$ -value and the principal diffusivity,  $\mathbf{g}_i$  is the gradient direction and  $\mathbf{v}$  is the integration parameter. The Watson distribution is preferred here to the Bingham distribution used in [19]. As was pointed out in [19] the Bingham distribution accounts better for fiber-spread but decreases the angular separation power of the algorithm which is not a desired result here.

All the angular deviations reported here are calculated by summing up the deviations of the fibers from their closest ground-truth compartments, and dividing the result by the number of fibers (two fibers in our experiments). In the first experiment we tested the stability of the algorithm using 64, 32 and 12 gradient directions. For each separation angle we simulated a noisy signal at different separation angle, initialized the algorithm randomly 200 times and measured the mean angular deviation from the true orientations as well as the standard deviation. The results in Fig. 1 show that the algorithm provides stable performance in all of the simulated cases using 64 and 32 directions. The stability of the algorithm declines when the number of gradient directions is reduced to 12. In that case, only large separation angles can be detected reliably, especially when  $b=3000$ .



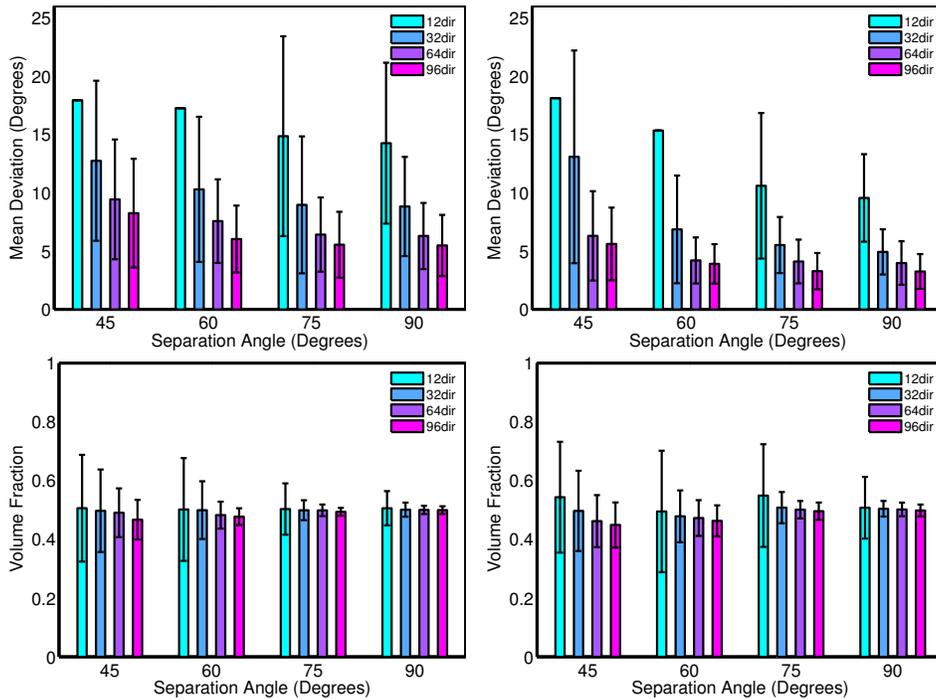
**Fig. 1.** Stability analysis for SNR=20. The experiments were carried on  $b=1500$  (left) and  $b=3000$  (right).

Next, we tested the accuracy of the algorithm. For each separation angle 200 noisy data instantiations were simulated and the angular deviation was measured separately for each instantiation. To test the performance decline under gradient directions sub-sampling, different sets of gradient directions were used: 96, 64, 32 and 12. The mean and the standard deviation of the collected results are depicted in Fig. 2. These results verify the stability test observations: The algorithm performs very well using 32 gradient directions and provide plausible results with only 12 gradient directions. Note that 12 directions are below the

minimal number of measurements required for a 4th-order ODF estimation (15 coefficients) without using sparse representations or super-resolution techniques.

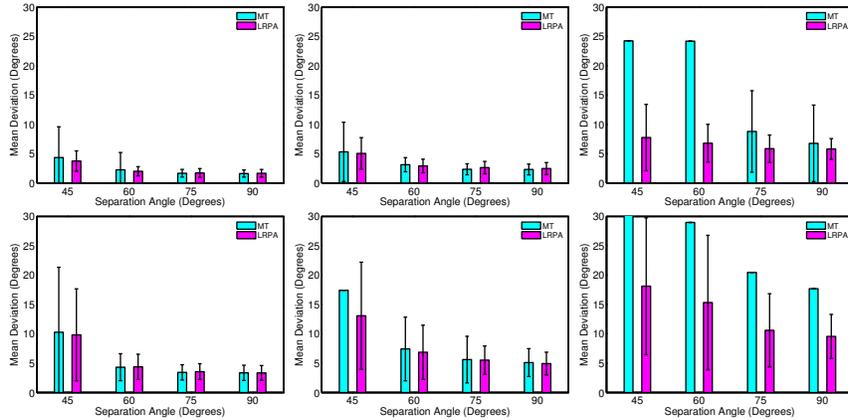
Finally, we compared our results to direct multi-tensor fitting. Since the signal was simulated using the multi-tensor model, theoretically, fitting this model to the data would give the best results. However, the comparison results depicted in Fig. 3 show that our algorithm is more stable, provide more accurate results and has a better separation resolution. To compare robustness to noise we added a dataset with SNR=40.

The multi-tensor fitting results were generated using Camino [20] with cylindrically symmetric tensors constraint. To get the best results, the diffusivities were set to the same values used to simulate the signal. The superiority of our algorithm in terms of accuracy and robustness to noise with only 12 samples is clearly shown in Fig. 3.



**Fig. 2.** Simulated data with SNR=20. Results of two b-values are presented: 1500 s/mm<sup>2</sup> (left) and 3000 s/mm<sup>2</sup> (right). The angular resolution is presented at the top and the volume fractions at the bottom. When the standard deviation exceeds the axis limit, we present the mean only.

In addition to the results reported here, our algorithm was compared with 12 different HARDI reconstruction techniques in ISBI'12 Workshop on HARDI



**Fig. 3.** Comparison of our Low Rank Polynomial Approximation (LRPA) and the Multi-Tensor (MT) model. The b-value is 3000 and the SNRs are 40 (top) and 20 (bottom). From left to right: 64, 32 and 12 gradient directions.

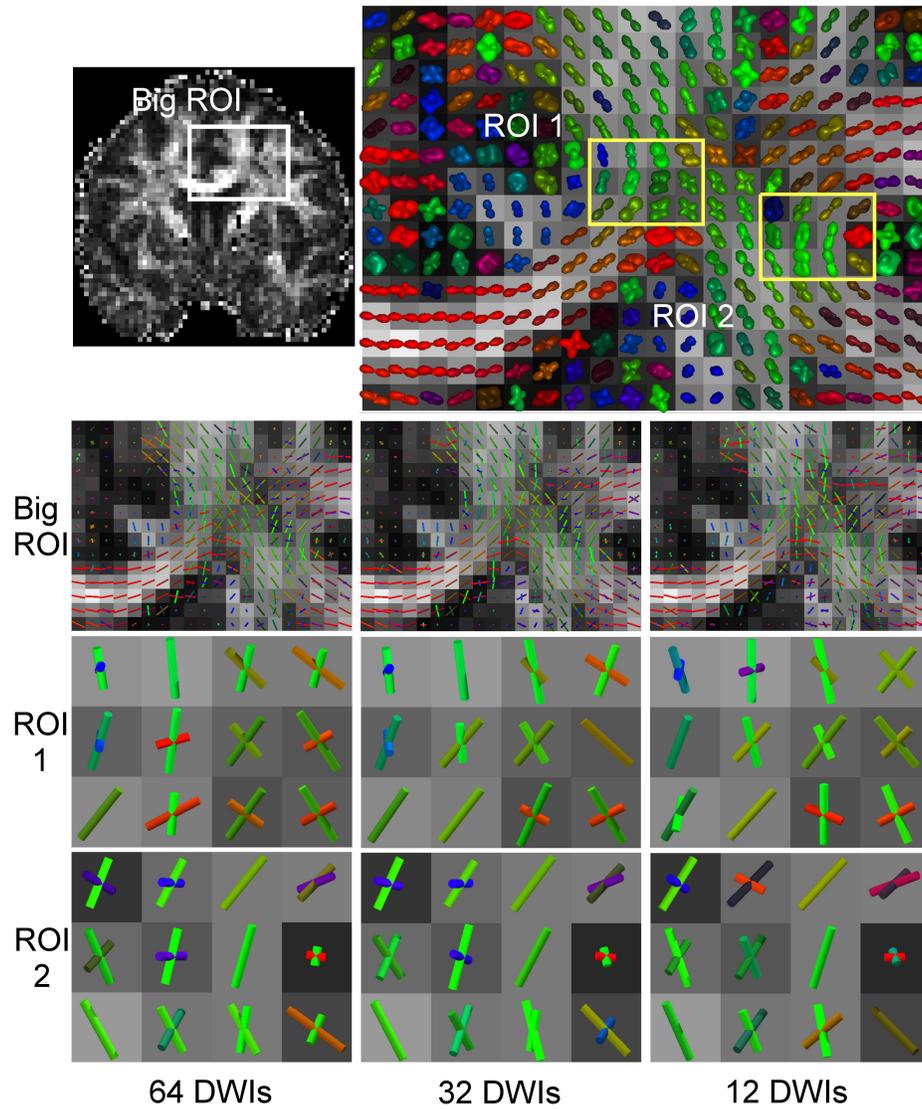
reconstruction [21]. The different algorithms were evaluated on a simulated 3D phantom as well as single voxels, in three SNR levels: 30, 20 and 10. Each algorithm was ranked based on the accuracy of fiber directions reconstruction, estimation of the number of fibers in each voxel and the ODF quality compared to the ground-truth. The number of gradient directions used for the reconstruction was also taken into account for the final ranking. Our algorithm (presented under the team name “The HOT gang”) was ranked first or second in all of the three final ranking options.

The measured running time for a Matlab implementation of our algorithm is on average 40ms per voxel. This was measured on a standard laptop with 2.4Ghz Intel Core i5 CPU and 4 GM RAM. The tested implementation is non-optimal and acceleration methods can be used to achieve faster convergence times.

## 4.2 Human brain data

The human brain data was acquired on a 3T Siemens Tim Trio scanner using a single-shot spin-echo EPI sequence and a b-value of 2000 s/mm<sup>2</sup>. One B0 image and 64 diffusion weighted images with a matrix size of 106 × 106 × 76 and a voxel volume of 2 mm<sup>3</sup> were acquired. The measured baseline SNR for these data was approximately 20. A white matter mask was registered to the data and the kernel width,  $\delta$ , was estimated by computing the mean principal diffusivity of all white matter voxels with FA > 0.7.

Selecting the number of fiber compartments per voxel is a significant challenge, especially in direct model fitting techniques. As was pointed out in [13], fitting a mixture model with more than one compartment to a single fiber voxel will yield inaccurate results. Therefore, the number of fibers has to be pre-selected by using statistical inference methods such as an  $F$ -test. As we have found out



**Fig. 4.** Coronal slice showing reconstruction results of a crossing fibers region in human brain. From left to right: 64, 32 and 12 gradient directions. The FA values are shown at the background. The ODFs were reconstructed using [4] and are only provided as a reference for the fibers' structure in this ROI.

by simulated data experiments, our technique can accurately resolve the fiber orientations even in a case of over-fitting. That is, fitting a convolution of a sum of two linear-forms to a single fiber voxel, will yield a dominant fiber with a high volume fraction and a fiber with a low volume fraction. The dominant fiber accurately matches the orientation of the single fiber. This suggests that we can estimate the orientations everywhere using a fixed number of linear-forms and, then, eliminate “weak” fibers by direct thresholding. This is a clear advantage over multi-compartment fitting techniques. Also, it was shown in [14] that weight-based thresholding provides more accurate estimates of the number of fibers compared to statistical inference methods.

For this brain data we set  $\tilde{r} = 3$  and learn the threshold from the high FA voxels that were used for the kernel parameter estimation. As most of these voxels lie in single tract regions, such as the corpus callosum, they presumably consist of single fibers. Thus, by applying a rank-two polynomial approximation to these voxels, the term with the lowest weight is likely to describe noise rather than a fiber. Indeed, the results show a high ratio between the weights of the first and the second term in these voxels. The threshold was then set as the average of the lowest weights where a value of 0.21 was computed. Thus, fibers with a volume fraction less than about 25% of the dominant volume fraction are considered as noise and eliminated.

To test the algorithm we have chosen the brain region where the corpus callosum (CC), the corona radiata (CR) and the superior longitudinal fasciculus (SLF) form a crossing pattern (Fig. 4). The results show that along the single tracts mostly one fiber model was selected whereas in the region where the different tracts cross, mostly two-fiber patterns were selected.

The fiber orientations were reconstructed using different sets of gradient directions which contain 64, 32 and 12 directions. The original set of gradient directions was sub-sampled using the algorithm described in [22]. The results of the fiber orientations and the number of fiber selected in each voxel are very similar whether 64 or only 32 directions were used. Crossing fibers are still presented using only 12 gradient directions although a performance decline is observed. This is due to loss of separation resolution and estimation accuracy in compliance with the simulated data observations. Yet, considering the noise level and the number of gradient directions used, these results are plausible. Note that we have not used any spatial regularization term in the current algorithm and we believe that these results can be improved by adding such regularization.

## 5 Conclusions

We presented a robust technique for the estimation of white matter fiber orientations and volume fractions directly from single-shell HARDI measurements. Similar to multi-compartment models, this technique avoids the complexity of extracting the orientations from the ODF and can model complex fiber structures using a few number of parameters. This technique relies on a low-rank homogeneous polynomial approximation by means of powers of linear-forms represent-

ing single fibers. An  $l_2$  optimization problem based on a spherical deconvolution technique is used to estimate the fiber orientations and the volume fractions. Our technique is accompanied by a robust iterative alternating Levenberg-Marquardt scheme. Using simulated data we showed that our algorithm provide accurate and stable results for low number of gradient directions and is favorable to direct multi-tensor fitting. We applied this algorithm to in vivo, human brain data and showed that the reconstructed orientations follow the major tracts and describe fiber intersection regions well. Furthermore, by sub-sampling the number of gradient directions we showed that plausible results can be obtained using only 12 gradient directions. The potential of this approach for reduced HARDI acquisition time is clear. In the future, we plan to deploy our new method within a tractography algorithm. In addition, accuracy and stability evaluations with respect to other reconstruction techniques, including sparsity-based methods, will be provided.

**Acknowledgments.** This work was made possible in part by the NIH/NIGMS Center for Integrative Biomedical Computing, 2P41 RR0112553-12.

## References

1. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. *Magn. Res. Med.* **58**(3) (2007) 497–510
2. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**(4) (2007) 1459–1472
3. Tuch, D.S., Reese, T.G., Wiegell, M.R., Makris, N., Belliveau, J.W., Wedeen, V.J.: High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn. Res. Med.* **48**(4) (2002) 577–582
4. Weldeselassie, Y., Barmpoutis, A., Atkins, M.S.: Symmetric positive-definite cartesian tensor orientation distribution functions (CT-ODF). In Jiang, T., Navab, N., Pluim, J., Viergever, M., eds.: MICCAI'10. Volume 6361 of LNCS., Springer, Heidelberg (2010) 582–589
5. Tuch, D.S.: Q-ball imaging. *Magn. Res. Med.* **52**(6) (2004) 1358–1372
6. Özarslan, E., Shepherd, T.M., Vemuri, B.C., Blackband, S.J., Mareci, T.H.: Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage* **31**(3) (2006) 1086–1103
7. Aganj, I., Lenglet, C., Sapiro, G.: ODF maxima extraction in spherical harmonic representation via analytical search space reduction. In: MICCAI'10. Volume 6361 of LNCS. (2010) 84–91
8. Bloy, L., Verma, R.: On computing the underlying fiber directions from the diffusion orientation distribution function. In: MICCAI'08. Volume 5241 of LNCS. (2008) 1–8
9. Ghosh, A., Wassermann, D., Deriche, R.: A polynomial approach for maxima extraction and its application to tractography in hardi. In: IPMI'11. Volume 6801 of LNCS. (2011) 723–734
10. Jiao, F., Gur, Y., Johnson, C.R., Joshi, S.: Detection of crossing white matter fibers with high-order tensors and rank-k decompositions. In: IPMI'11. Volume 6801 of LNCS. (2011) 538–549

11. Schultz, T., Seidel, H.P.: Estimating crossing fibers: A tensor decomposition approach. *IEEE TVCG* **14**(6) (2008) 1635–1642
12. Behrens, T., Berg, H.J., Jbabdi, S., Rushworth, M., Woolrich, M.
13. Alexander, D.: Multiple-fibre reconstruction algorithms for diffusion MRI. *Annals of the New York Academy of Sciences* **1046** (2005) 113–133
14. Schultz, T., Westin, C.F., Kindlmann, G.: Multi-diffusion-tensor fitting via spherical deconvolution: A unifying framework. In Jiang, T., Navab, N., Pluim, J., Viergever, M., eds.: *MICCAI'10*. Volume 6361 of LNCS., Springer, Heidelberg (2010) 673–680
15. Rathi, Y., Malcolm, J.G., Michailovich, O., Westin, C.F., Shenton, M.E., Bouix, S.: Tensor kernels for simultaneous fiber model estimation and tractography. *Magnetic Resonance in Medicine* **64**(1) (2010) 138–148
16. Brachat, J., Comon, P., Mourrain, B., Tsigaridas, E.P.: Symmetric tensor decomposition. *Linear Algebra and Applications* **433**(11-12) (2010) 1851–1872
17. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine* **56**(2) (2006) 395–410
18. Atkinson, K.: Numerical integration on the sphere. *J. Austral. Math. Soc.* **23** (1982) 332–347
19. Kaden, E., Knsche, T.R., Anwender, A.: Parametric spherical deconvolution: Inferring anatomical connectivity using diffusion MR imaging. *NeuroImage* **37** (2007) 474–488
20. Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., Alexander, D.C.: UCL Camino Diffusion MRI Toolkit. <http://cmic.cs.ucl.ac.uk/camino/> (2006)
21. Gur, Y., Jiao, F., Zhu, S.X., Johnson, C.R.: Fiber orientations assessment via symmetric tensor decomposition. In: *ISBI Workshop on HARDI reconstruction*, Barcelona, Spain (May 2012) <http://hardi.epfl.ch/>
22. Zhan, L., Leow, A.D., Jahanshad, N., Chiang, M.C., Barysheva, M., Lee, A.D., Toga, A.W., McMahon, K.L., de Zubicaray, G.I., Wright, M.J., Thompson, P.M.: How does angular resolution affect diffusion imaging measures? *NeuroImage* **49**(2) (2010) 1357 – 1371

# Statistical Analysis of Maximum Density Path Deformation Fields in White Matter Tracts

Julio Villalon-Reina<sup>1</sup>, Gautam Prasad<sup>1</sup>, Shantanu H. Joshi<sup>2</sup>, Maria Jalbrzikowski<sup>3</sup>, Arthur Toga<sup>2</sup>, Carrie Bearden<sup>3,4</sup>, Paul M. Thompson<sup>1</sup>

<sup>1</sup>Imaging Genetics Center, Laboratory of Neuro Imaging,  
UCLA School of Medicine, Los Angeles, CA, USA  
{thompson, julio.villalon}@loni.ucla.edu

<sup>2</sup>Laboratory of Neuro Imaging, UCLA School of Medicine,  
Los Angeles, CA, USA

<sup>3</sup>Department of Psychology, UCLA, Los Angeles, CA, USA

<sup>4</sup>Department of Psychiatry and Biobehavioral Sciences,  
Semel Institute for Neuroscience and Human Behavior, UCLA,  
Los Angeles, CA, USA

**Abstract.** Several methods have been developed to cluster the axonal fibers of the white matter into meaningful bundles, after performing whole-brain tractography. Here we use a previously described method to cluster fibers, represent them compactly as paths ('maximum density paths'), and compute homologies across a population using geodesic curve registration. We now extend this method to include an analysis of the deformation fields from the curve registration, allowing both fiber geometry and integrity to be analyzed. To evaluate differences between normal and diseased populations, we computed a population mean maximum density path and extracted the corresponding deformation fields matching each subject's tracts to this mean. By analyzing a cohort of children with 22q11.2 deletion syndrome (22q11.2DS) and typically developing children, we found significant relationships of tract geometry to performance on visual and verbal cognitive tests.

**Keywords.** Diffusion tensor imaging, tractography, white matter tracts, clustering, geodesic distance, shape analysis

## 1 Introduction

Profound changes in white matter morphology and integrity occur during development and aging as well as in neurogenetic and degenerative brain diseases [1,2]. DTI tractography algorithms have greatly advanced our ability to study white matter, by reconstructing the

trajectories of neural pathways following the dominant directions of water diffusion along them. The myriad of fibers that these algorithms are able to reconstruct has generated increased interest in tools for population and group comparisons of specific white matter tracts. Even so, to compare anatomy across subjects, it is vital to be able to bundle fibers together in a consistent way, and define metrics sensitive to local differences in fiber integrity and geometry.

White matter tracts interconnect a vast array of grey matter structures in the brain, transmitting information along neural pathways or tracts. In addition, fibers within each of these tracts may connect different neighboring grey matter areas. Thus, it is important to use a suitable representation of those specific fiber bundles to allow population comparisons and to test regional or local correlations with behavioral and cognitive measures.

Many clustering methods have been developed to analyze tractography results, such as spectral clustering [3], Dirichlet distributions [4], and a Gaussian process framework [5]. Relatively few methods are designed to analyze shape differences among white matter tracts. Batchelor et al. [6] studied tract shapes using rigid body invariants (curvature and torsions) and compared them, with Fourier descriptors. They also explored the tract registration problem, via principal component analysis. Corouge et al. [7] proposed a framework for tract-oriented DTI analysis based on tensor interpolation and averaging, using a Riemannian metric to perform statistics on the Lie group of positive definite symmetric tensors. In a similar fashion to us, they calculate average and variance based on the geodesic distance, but at the tensor level and for a single bundle of fibers. Savadjiev et al. [8] also modeled white matter tracts as 3D curves by using co-helicity and the estimated local tangent, curvature and torsion for single curves. Yushkevich et al. [9] proposed to use continuous medial representations to model tracts with a sheet-like structure as parametric surfaces (such as the corpus callosum and corticospinal tract). They compared typically developing children with a group of children affected by the neurogenetic syndrome 22q11.2DS, evaluating differences in the apparent diffusion coefficients along the tracts. In contrast, our method analyzes deformations derived from paths along the bundles as a basis for statistical analysis. Recently, Prasad et al. [10] proposed a method for white matter tract registration based on maximum density paths. They cluster fibers based on a statistical atlas, represent them using a path following points of highest density in the bundle, and use geodesic registration to find correspondences in these paths across a population of subjects [10]. However their tract based analysis only compares standard DTI-based measures such as fractional anisotropy and mean diffusivity between two populations; with models of tract geometry, additional shape-related features are available for analysis.

In this work we propose a statistical analysis framework that analyzes the geometric variation in the maximum density paths [10] of two specific fiber tracts, through an analysis of their deformation fields derived from registration to a population mean path. We use the

same compact representation of white matter tracts proposed in [10] but we additionally propose a method to analyze the shape and geometry of the white matter tracts. We analyze fiber geometry by analyzing tract-based deformation fields along paths, and, in an example application, we assess how these measures may relate to behavioral and cognitive assessments. We first describe the method mentioned above to represent the white matter tracts - we cluster the fibers, represent them as paths, and register these paths into the same space [10,11]. We then outline our method to analyze the deformation fields derived from this registration, which allows localized comparisons across populations.

## 2 Methods

### 2.1 Data

Our analysis included children with a neurogenetic syndrome known as 22q11.2DS, as well as a group of age-matched, typically developing (TD) children. The subjects consisted of 22 children with 22q11.2DS and 21 TD controls (mean age: 10.2 $\pm$ 2.9SD; 28 males/15 females). We also examined visual and verbal memory abilities, as assessed with the Children’s Memory Scale (CMS) and the California Verbal Learning Test (CVLT). All the subjects underwent high resolution MRI scanning (on a 3T Siemens Trio MRI scanner). T1-weighted anatomical images were acquired with an MP-RAGE sequence. Acquisition parameters were: TR/TE/TI = 2300/2.91/900 msec; flip angle = 9 degrees; slice thickness = 1.20 mm, with a 240x256 acquisition matrix). Diffusion-weighted images (DWI) were also acquired using diffusion-encoded, spin-echo echo planar imaging. Acquisition parameters were optimized to provide the best signal-to-noise ratio for estimating diffusion tensors [13]. Imaging parameters were: 190 mm FOV, TR/TE 7100 ms/93 ms, with a 96x96 acquisition matrix. Each 3D volume consisted of 50 2-mm thick axial slices with no gap and 1.97x1.97 mm<sup>2</sup> in-plane resolution. 65 images were acquired per subject: 1 with no diffusion sensitization (i.e., T2-weighted  $b_0$  images) and 64 diffusion-weighted (DW) images ( $b = 1000$  s/mm<sup>2</sup>) with gradient directions evenly distributed on the hemisphere.

We utilized FSL’s “BET” tool (FMRIB Software Library, <http://fsl.fmrib.ox.ac.uk/fsl/>) to skull-strip the T1-weighted images. After being corrected for field inhomogeneities with N3, all T1-weighted images were linearly aligned using FSL’s “FLIRT” (with 6 degrees of freedom) to a common space (Colin27) with 1mm isotropic voxels and a 220x220x220 voxel matrix. Non-brain regions were automatically removed from T2-weighted  $b_0$  image from the DWI set using “BET” and the mask was applied to the rest of 64 volumes. Raw diffusion-weighted images were corrected for eddy current distortions using the FSL tool, “eddy\_correct” (<http://fsl.fmrib.ox.ac.uk/fsl/>). For each subject, the eddy-corrected  $b_0$  image was linearly aligned (9 degrees of freedom) and downsampled to a version of their corresponding T1 image (110x110x110, 2x2x2mm) that was previously aligned to the Colin27-MNI space. The  $b_0$  map was also elastically registered to the T1 structural scan to compen-

sate for EPI-induced susceptibility artifacts. The transformation matrix from the linear alignment of the mean  $b_0$  image to the T1-weighted volume was applied to each of the diffusion sensitized 64 volumes. We then rotated the original gradient vectors using the rotation matrix from the linear transformation. We computed diffusion tensors with Diffusion Toolkit (<http://trackvis.org/dtk/>) as well as corresponding diffusivity measure volumes (i.e., fractional anisotropy and mean diffusivity). We performed whole-brain tractography, for all subjects, using an interpolated streamline method, with a 1mm step-length, and a turning angle threshold of 35 degrees. We used a spline filter to smooth and ‘clean up’ the fibers. After calculating the fibers, elastic deformations were obtained from the EPI distortion correction by registering the  $b_0$  image to the T1-weighted image. These deformations were then applied to the fibers’ 3D coordinates for accurate alignment of the anatomy.

## 2.2 Cluster Representation using Maximum Density Path

We used the previously described method in [10,11] to cluster the fibers in the image space, represent them using a path following points of highest density, and register these paths into the same space to compute correspondences across a cohort of subjects.

This method incorporates prior anatomical knowledge by incorporating the Johns Hopkins University (JHU) white matter atlas [12], which delineates 50 probabilistic white matter regions of interest (ROI). The atlas designates 3 types of white matter tracts: commissural (interhemispheric), association (antero-posterior), and projection (thalamocortical) fibers. We performed our analysis on the commissural and association fibers, as we hypothesized that the behavioral abnormalities would be associated with measurable variations in these types of fibers.

The FA volume of JHU atlas, which is in the ICBM-152 space, was aligned to the Colin27-MNI template (110x110x110, 2mm x 2mm x 2mm), used previously to align the diffusion-weighted volumes. We then elastically registered the resampled FA image of the JHU atlas to each subject’s FA map. This transformation was also applied to each of the 50 3D probabilistic white matter ROIs. In this way, we ensured that each of the probabilistic ROIs was registered to each subject’s space.

To extract fibers corresponding to a white matter tract of interest, a 3D cubic spline curve was fitted to each fiber, and it was regularly re-sampled. We found the set of voxels that it intersects in the encompassing image space, and then we computed their overlap with the ROI. The number of voxels in the intersection measures how well a tract fits into an atlas region. We then selected all tracts that intersected the atlas.

Once the final set of fibers that represent a specific tract was selected, we were able to compute a more compact representation of the tract. We constructed a graph (a set of nodes and

undirected edges) to represent the voxel-wise density of our fiber bundle. We computed a density volume of our fiber bundle to characterize our search space. Nodes in the graph represent voxel locations that have a non-zero density value, and the edges connect each voxel to its surrounding 26 neighboring voxels. Each edge was weighted by the negative sum of the voxel densities it connected, as follows:

$$-(d_i + d_j) + e_m$$

Here  $d_i$  and  $d_j$  denote node  $i$  and  $j$ 's corresponding voxel density value.  $e_m$  is the maximum absolute value of the edge costs; this term is added to ensure that the edges are positive.

We then select a start and end node in the graph by first assigning a start and end voxel location in the probabilistic ROI. These start and end voxels are located within the highest probability section of the ROI. By this we ensure that as many fibers as possible of the corresponding ROI are captured. The closest non-zero voxel locations in the density image (by Euclidean distance) to the start and end voxels in the atlas are used as the voxels that correspond to the start and end nodes in the graph. Dijkstra's algorithm [13] is then run on the graph to find the shortest path connecting the start and end nodes. Dijkstra's algorithm is a graph search algorithm that efficiently finds the shortest path from a start node to every other node in the graph. In this case, the shortest path will include edges connecting nodes with high-density values. As a result, it follows the path with the highest density of fibers in the image.

The centers of the voxels in the path are then used as the coordinates that trace the path, and a Gaussian filter is used to smooth the curve. The resulting path is used to compactly represent the fiber bundle; we refer to it as the maximum density path (MDP).

### 2.3 Maximum Density Path Registration

We register the maximum density paths (MDPs) together to find correspondences across paths representing the same white matter regions across subjects. In this pursuit, we seek a unique invariant representation by considering an equivalence class of all transformations such as scaling, translations, rotations and reparameterizations that leave the shape of the MDP curve unchanged. The underlying premise is that shapes of white matter tracts encode the reduced dimensional geometry of the white matter tracts [14,15].

The MDP path is represented by a parameterized curve in  $\mathfrak{R}^3$ . Assuming a standard unit interval, the coordinate function of this path is denoted by  $\beta(s) : [0,1] \rightarrow \mathfrak{R}^3$ . The coordinate function is not a good choice to analyze the shape of the path, since it is confounded by

global location, scale, and orientation. We represent the shape of the curve  $\beta$  by the function  $q: [0,1] \rightarrow \mathfrak{R}^3$  as:

$$q(s) = \frac{\dot{\beta}(s)}{\sqrt{\|\dot{\beta}(s)\|}} \in \mathfrak{R}^3 \quad (1)$$

Here,  $s \in [0,1]$ ,  $\|\cdot\| \equiv \sqrt{(\cdot, \cdot)_{\mathfrak{R}^3}}$ , and  $(\cdot, \cdot)_{\mathfrak{R}^3}$  is the standard Euclidean inner product in  $\mathfrak{R}^3$ . To make the representation scale invariant, we normalize the function  $q$  by dividing it by its magnitude.

$$\tilde{q} = \frac{q}{\sqrt{\int_0^1 (q(s), q(s))_{\mathfrak{R}^3} ds}} \quad (2)$$

This is the projection into the space of all unit-length curves that are scale-invariant. The shape of a curve will be assumed to have this scale-invariant form given by  $\tilde{q}$ , and with a slight abuse of notation, it will be subsequently referred to as  $q$ . We denote the space of all elastic scale-invariant curves by:

$$\mathcal{Q} \equiv \left\{ q | q(s) : [0,1] \rightarrow \mathfrak{R}^3 \mid \int_0^1 (q(s), q(s))_{\mathfrak{R}^3} ds = 1 \right\} \quad (3)$$

Due to scale invariance, the space  $\mathcal{Q}$  becomes an infinite-dimensional Hilbert unit-sphere of functions, and represents all open elastic curves invariant to translation and uniform scaling. We denote this as the “pre-shape” space. Each curve is projected as a single element of the pre-shape space.

As our goal is to find a shape representation that is also invariant to translations and reparameterizations, we also define a *quotient space* of shapes, where the computation of geodesics, and thus the matching between curves is achieved by constructing the space of elastic shapes. Therefore, we define the elastic shape space as the quotient space  $S = \mathcal{Q}/(SO(3) \times D)$ . Given a pair of shapes, the corresponding distance is then calculated as the length of the shortest path - i.e., of a geodesic - between the respective equivalence classes on the shape space. Additionally, curves can assume arbitrary speeds without changing the shape, giving rise to multiple parameterizations that represent the same tract shape. This ambiguity of representation can be denoted by a new group action constituting a reparameterization by a nonlinear map that changes the speed of the tract. In order to ensure that the ordering of the points on the MDPs remain the same after reparameterization and that the

speed function  $\gamma$  does not exhibit sharp jumps and discontinuities, we need  $\gamma$  to be differentiable and inverse-differentiable as well. Thus, we define the space

$$D = \{\gamma : Q \rightarrow Q\} \quad (4)$$

as the space of all orientation-preserving diffeomorphisms.

We will find a geodesic between the equivalence classes of  $q_1$  and  $q_2$  by fixing the parameterization of  $q_1$  and iteratively reparameterizing and reorienting  $q_2$  according to  $(O_3 q_2) \cdot \gamma$ , where  $O_3 \in SO(3)$ ,  $\gamma \in D$ , such that the length of the geodesic path given by

$$d_e(q_1, q_2) = \min_{O_3 \in SO(3), \gamma \in D} d(q_1, (O_3 q_2) \cdot \gamma) \quad (5)$$

is minimized. Here,  $d$  is the geodesic distance given by [16]. With this optimization procedure, we make sure that the shapes are also represented in a space that is invariant to reparameterization and rotation.

#### 2.4 Statistical Analysis of Maximum Density Path Deformation Fields

For a collection of tract landmark curves, we now construct a statistical shape average for the entire set of those curves calculated for the TD group (healthy control children). The reason for this is that we want to find shape differences between 22q11.2DS children and TD children. We used the intrinsic average, which is computed directly on the shape space, and makes use of distances and lengths that are defined strictly on the shape space. We use the well-known definition of the Karcher mean [17] to denote this intrinsic average.

The Karcher mean for a set of shapes  $\{q_i\}$ ,  $i = 1, \dots, N$  is given by

$$\mu = \arg \min_{q_\mu} \sum_{i=1}^N d_e(q_\mu, q_i)^2, i = 1, \dots, N. \quad (6)$$

Unlike the extrinsic mean, the Karcher mean is calculated by an iterative optimization procedure that involves repeated computations of geodesics from each of the shapes of the population to the current estimate of the mean. In other words the algorithm for Karcher mean minimizes the geodesic variance instead of the Euclidean variance in case of the arithmetic mean. This Karcher mean of the normal control population will be referred to

from now on as the *mean MDP*, while still bearing in mind that a geodesic method was used to compute it, rather than a pointwise Euclidean average.

To derive the deformation field from the mean MDP to the individual subject, we denote the geodesic path by  $\alpha_t : [0,1] \rightarrow \mathcal{S}$ , such that at any time  $t \in [0,1]$ , we have the intermediate MDP shape along the geodesic lying in the shape space  $\mathcal{S}$ . Then Eqn. (5) may also be written in terms of the velocity vector along the geodesic path as

$$d_e(q_1, q_2) = \int_0^1 \sqrt{\left\langle \dot{\alpha}_t, \dot{\alpha}_t \right\rangle} dt \quad (7)$$

The velocity vector represents the first-order variation of the deformation of one MDP to another, and enables one to reconstruct the entire geodesic if the starting MDP path is known. For the purpose of statistical analysis, the starting MDP is always given by the mean MDP, to which all the subjects in the population are registered. Given the mean MDP  $q_\mu$  and a subject MDP  $q_i$ , and a geodesic  $\alpha_t$  between them, we can calculate the deformation fields as

$$\psi(s) = \left[ \int_0^s q_\mu(\tau) \parallel q_\mu(\tau) \parallel d\tau - \int_0^s O_3 \sqrt{\dot{\gamma} q_i(\gamma(\tau))} \parallel O_3 \sqrt{\dot{\gamma} q_i(\gamma(\tau))} \parallel d\tau \right] \quad (8)$$

where  $O_3$  and  $\gamma$  are given by solving Eqn. (5).

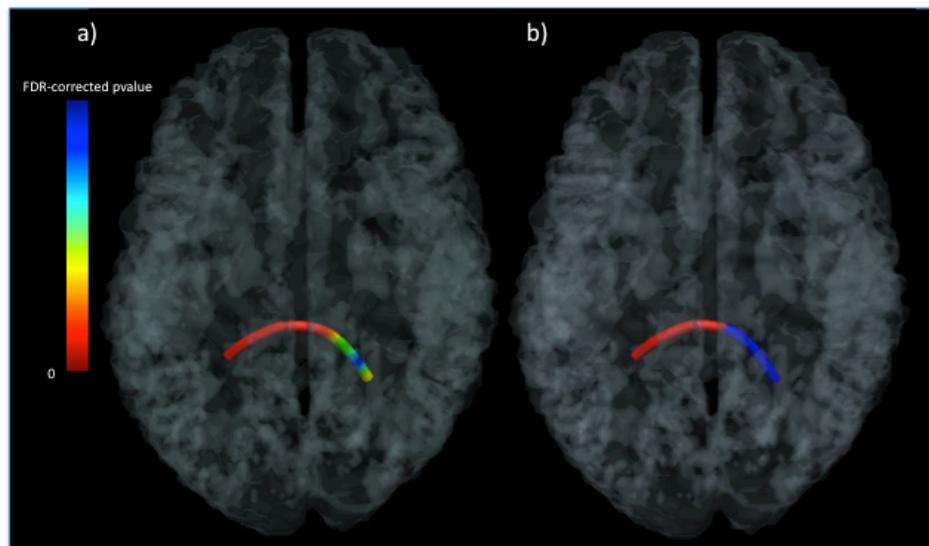
At this point, we now have the deformation fields (magnitude of the velocity vector) for matching each subject's tract to the corresponding tract's mean MDP (which was calculated with TD children only). We then proceeded to compute a linear regression, where we considered the deformation fields as the dependent variable and the diagnosis (controls vs 22q11.2DS) and the cognitive scores as regressors, with age and sex as covariates.

### 3 Results

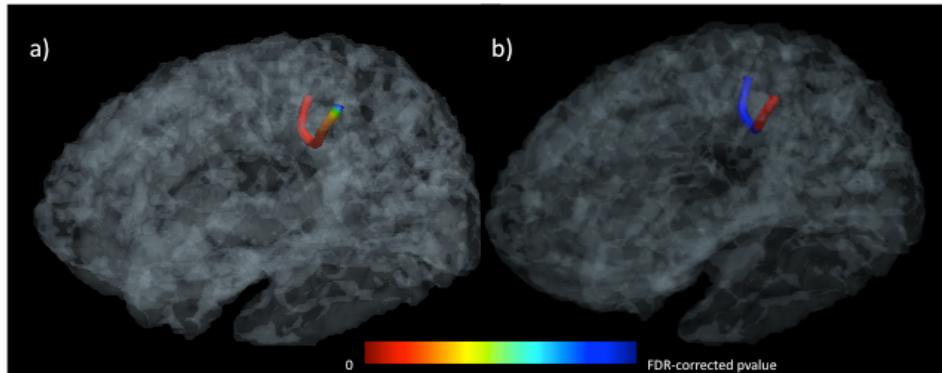
Neuroanatomical studies of children with 22q11.2DS have consistently shown structural abnormalities in midline cortical regions. There is a myelin-related gene (*PIK4CA*) in the 22q11.2 locus that is hemizygotously deleted in these patients [18], suggesting disrupted axonal myelination and impaired white matter integrity. Due to this, we decided to focus our analysis on two types of major white matter tracts that are critical for integrated cognitive

functioning - the corpus callosum and the long association fibers. The probabilistic JHU atlas divides the corpus callosum into 12 sections from anterior to posterior, and the long association fibers are divided into 6 different tracts. We studied the long association tracts due to their key role in fronto-parieto-temporal connectivity. This function is markedly impaired in 22q11.2DS [19,20]. We also focused on the corpus callosum, which carries the majority of the brain's interhemispheric fibers and also shows altered morphology in 22q11.2DS [21,22].

In regression analyses assessing relationships with the Children's Memory Scale (CMS) visuo-spatial memory test (**Figure 1**), we found a positive correlation between the shape variation of fibers of the corpus callosum that project to the occipital lingual gyrus (visual areas 1 and 2) and better task performance. This effect was mainly found in the medial and right lateral sections of the posterior fibers of the corpus callosum. When examining the relationship with the California Verbal Learning Test (CVLT), a measure of declarative verbal memory, we also found a positive correlation between the geometric variation of the fibers of the corpus callosum connecting the superior parietal gyrus in both hemispheres (**Figure 2**).



**Figure 1.** (a) Shows the mean MDP of the corpus callosum fibers that project to the lingual gyrus in the occipital lobe. Significant p-values (in red) represent the sections of this bundle that showed the highest geometric variation between normal controls and 22q11.2DS (p-value controlling the FDR at a q-value of 0.05 is 0.022901). (b) Shows the  $\beta$ -values along the same corpus callosum fiber bundle. Areas shown in red are areas where the score of the CMS short-term visual test were positively correlated with the deformation fields.



**Figure 2** (a) Shows the mean MDP of the corpus callosum fibers that project to the superior parietal gyrus. Significant p-values (in red) represent the sections of this bundle that showed the highest geometric variation between normal controls and 22q11.2DS (p-value controlling the FDR at a q-value of 0.05 is 0.03113). (b) Shows the  $\beta$ -values along the same corpus callosum fiber bundle. Areas shown in red are areas where the score of the CVLT short-term verbal memory test were positively correlated with the deformation fields.

#### 4 Discussion

In this paper, we introduced a method for population statistical analysis of DTI, which derives tracts from whole-brain tractography, and distills them into meaningful bundles. Rather than use pointwise averaging, we used a fully intrinsic method that allows the assessment of differences in both the shapes and integrity of the tracts. Both of these are commonly altered in neurogenetic syndromes, which tend to alter the gross morphology of the brain as well as its microstructure.

We applied our method to a cohort of children with 22q11.2 deletion syndrome (22q11.2DS) and a matched group of typically developing children (TD). 22q11.2DS is a neurogenetic syndrome, for which a wide range of brain morphologic abnormalities have been described, such as generalized reduced brain volume, larger ventricles, larger corpus callosum, and larger subcortical grey matter nuclei amongst others [19,20]. In this work, we show the beginnings of a method that may be expanded to survey tract-based geometric variation in white matter tracts, to find differences that may be relevant for behavior and cognition in 22q11.2DS and in typically developing children.

## References

1. Leporé, N. et al. (2010). Brain structure changes visualized in early- and late-onset blind subjects. *NeuroImage*, 49(1), 134-140.
2. Hua, X. et al. (2008). Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. *NeuroImage*. 43(3). 458-469.
3. O'Donnell, L. et al. (2005). White matter tract clustering and correspondence in populations. In: *Lect Notes Comput Sc. Medical Image Computing and Computer-Assisted Intervention– MICCAI 2005*, 140–147.
4. Maddah, M. et al. (2008). Modeling of anatomical information in clustering of white matter fiber trajectories using Dirichlet distribution. In: *PROC CVPR IEEE*. 1–7.
5. Wassermann, D. et al. (2010). Diffusion-Based Population Statistics Using Tract Probability Maps. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, 631–639.
6. Batchelor, P. G. et al. (2006). Quantification of the shape of fiber tracts. *Magn Reson Med*. 55(4). 894 – 903.
7. Corouge, I. et al. (2006). Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Med Image Anal*. 10(5), 786–798.
8. Savadjiev, P. et al. (2006). 3D curve inference for diffusion MRI regularization and fibre tractography, *Med Im Anal*. 10(5). 799-813.
9. Yushkevich, P. et al. (2008). Structure-specific statistical mapping of white matter tracts. *NeuroImage*, 41(2). 448-461.
10. Prasad, G. et al. (2011). White Matter Tract Analysis in 454 Adults using Maximum Density Paths. In: *CDMRI 2011 Workshop proceedings, MICCAI 2011*.
11. Prasad, G. et al. (2011). Atlas-based fiber clustering for multi-subject analysis of high angular resolution diffusion imaging tractography. In: *Proceedings of the 8th IEEE International Symposium on Biomedical Imaging, ISBI 2011*, 276-280.
12. Oishi, K. et al. (2008). Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. *NeuroImage*. 43(3), 447-57.
13. Dijkstra, E. (1959). A note on two problems in connexion with graphs. *Numerische mathematik* 1(1), 269–271.
14. Joshi, S.H. et al. (2007). A novel representation for Riemannian analysis of elastic curves in  $R^n$ . In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–7.
15. Joshi, S.H. et al. (2007). Removing Shape-Preserving Transformations in Square-Root Elastic (SRE) Framework for Shape Analysis of Curves, In *Proceedings of Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science 4679, Springer*, pp. 387-398.
16. Mangin, J. et al. (2003). 3-D moment invariant based morphometry. In *Proc. Med. Image Computing Computer-Assisted Intervent. MICCAI 2003*. 505–512.
17. Karcher H (1977) Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics* 30:509-541.
18. Vorstman J.A. et al. (2009). Association of the PIK4CA schizophrenia-susceptibility gene in adults with the 22q11.2 deletion syndrome. *Am J Med Genet B Neuropsychiatr Genet*. 150B(3). 430-433.
19. Simon T, et al. (2005) Volumetric, connective, and morphologic changes in the brains of children with chromosome 22q11.2 deletion syndrome: an integrative study. *NeuroImage* 25, 169-180.

20. Kates WR, et al. (2001) Regional cortical white matter reductions in velocardiofacial syndrome: a volumetric MRI analysis. *Biol Psychiatry* 49, 677-684.
21. Bearden C.E. et al. (2009). Alterations in Midline Cortical Thickness and Gyrfication Patterns Mapped in Children with 22q11.2 Deletions. *Cereb Cortex*. 19(1): 115-126.
22. Machado AM, et al. (2007) Corpus callosum morphology and ventricular size in chromosome 22q11.2 deletion syndrome. *Brain Res* 1131.197-210.

# A Minimum Cost Approach to Connectivity from Orientation Distribution Functions via Efficient Multi-directional Graph Propagation

Ipek Oguz, Alexis Boucharin, Wenyu Lu, Clement Vachet, Francois Budin, Yundi Shi, and Martin Styner

Depts. of Psychiatry and Computer Science, University of North Carolina

**Abstract.** Regional connectivity measurements derived from diffusion imaging datasets are of considerable interest in the neuroimaging community for better understanding white matter connectivity. Current connectivity measurements are usually either based on fiber tractography applied in Monte-Carlo fashion, or variations of the Hamilton-Jacobi approach, or are graph-based. We propose a novel, graph-based algorithm that provides a fully deterministic, efficient and stable connectivity measure. This method handles crossing fibers and deals well with multiple seed regions. The computation is based on a multi-directional graph propagation algorithm applied to sampled orientation distribution functions computed directly from the original diffusion imaging data. While a maximum probability approach is possible, here we focus on a minimum cost formulation. We present results on synthetic and real datasets to illustrate the potential of our method towards subject-specific connectivity measurements performed in an efficient, stable and reproducible manner. Such individual connectivity measurements would be well suited for neuroimaging studies.

**Keywords:** DTI, MRI, ODF, connectivity, graph-based, F\*

## 1 Introduction

In recent years, the use of regional connectivity measurements computed from diffusion MRI data have become of considerable interest in the neuroimaging community in an effort to better understand cortical and subcortical white matter connectivity. While measuring cortical and subcortical white matter connectivity is an important endpoint in and of itself, efficient and stable methods for quantifying connectivity is particularly important for network analysis studies. A connectivity matrix designating the connection strength between all brain regions is typically the starting point of such studies.

The most elementary measurement of connectivity strength is the mean Fractional Anisotropy (FA) along the path between two regions. While this is a straightforward metric to compute, it is far from ideal as it cannot effectively handle crossing fibers and also because FA has a non-uniform and non-linear distribution along the fiber, which simple averaging completely disregards.

Stochastic tractography methods overcome some of these issues by repeatedly applying streamline tractography in a Monte-Carlo fashion [1, 5]. The connectivity strength at a given voxel is then defined as the number of paths reaching that voxel divided by the total number of generated paths. While this powerful method can overcome the problem of crossing fibers with an appropriate local diffusion model, it is not deterministic and can be very inefficient due to the large number of tracts that must be generated to converge to a stable result.

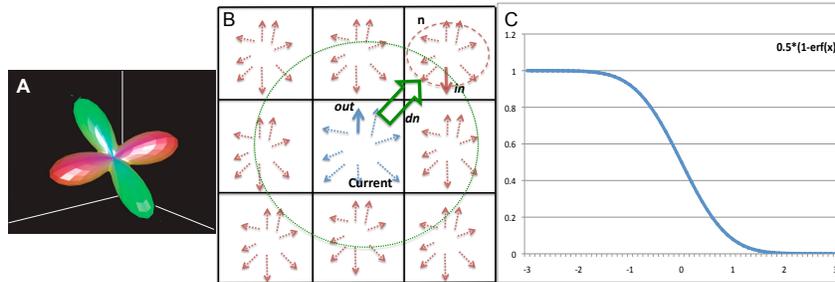
Various methods based on the Hamilton-Jacobi approach have been proposed to overcome some of the difficulties arising in tractography [7, 12, 10]. The main idea for these methods is to compute the shortest path where the cost associated with each path is an integral dependent on position, path orientation and local diffusion anisotropy/strength. These formulations result in first- or higher-order partial differential equations which model evolving fronts whose speeds are determined by information from the diffusion tensor [7]. These methods are inherently more robust to noise than tractography methods and also have the advantage of being fully deterministic and computationally efficient. However, these methods cannot handle crossing fibers effectively, whereas this can be incorporated into explicit tractography models. Moreover, these methods can not take into account the consistency of path orientation along the minimal path.

A final class of methods for computing connectivity consists of graph-based approaches [8, 15, 13]. These approaches treat the image as a graph by placing a vertex at each voxel and edges between neighboring voxels. The edges have weights, which can be considered as costs or probabilities of their two end voxels being connected to each other. These methods can take advantage of the vast graph processing literature for efficient computations. However, these methods are typically prone to taking “shortcuts” in the graph (or creating “phantom bridges”, as Zalesky et al. calls them), i.e. preferring a shorter path that doesn’t fit the data over a longer path that fits the data.

We propose a novel multi-directional graph-propagation based algorithm that computes connectivity between brain regions in a fully deterministic and efficient way, much like the Hamilton-Jacobi approach, while allowing crossing fibers. Furthermore, our method respects the local connectivity patterns in the data, which is not the case for the global-level optimization of Hamilton-Jacobi methods. This further allows our method to avoid the phantom bridge problem.

## 2 Methods

Motivated by the well-known  $F^*$  (pronounced “f-star”) graph traversal algorithm [6] and previous work using  $F^*$  to compute non-linear distances within the brain [14], we extended  $F^*$  to incorporate multiple incoming and outgoing directions for computing the overall probability or cost of a voxel. While we designed two formulations of this algorithm, one for propagating costs and one for propagating probabilities, here we will focus purely on the former. For details on the probability formulation, please refer to [2]. In addition to being more computationally expensive (probabilities are multiplicative whereas cost is ad-



**Fig. 1.** An ODF representation at a single voxel (A). Multi-directional cost function diagram (B) shows the variables that affect the cost of traversing an edge in 2D. For the voxel *Current*, for each direction *out*, we consider all neighboring voxels (green) and all the possible directions (red) at these voxels in order to identify the neighbor *n*, which lies in the direction  $\mathbf{d}_n$  from the *Current* voxel, and the direction *in* at the neighbor *n*, such that the cost is minimal. (C) Angle penalty function.

ditive), one of the major issues with the probabilistic approach is its proneness to cyclic connections, whereby a small loop of strongly connected voxels can be revisited many times rather than leaving the loop to follow the fiber path. The cost approach, as discussed below, allows for a length-independent formulation, which makes such loops unadvantageous and therefore avoids this problem.

## 2.1 Scalar $F^*$ graph traversal

The basic  $F^*$  algorithm [6] uses an adapted TV-scan processing of the image: starting from one corner of the three dimensional input image, the algorithm will visit each voxel in consecutive order. The processing will first visit each voxel of the current line until its end, then it will visit each voxel of the same line again in reverse order. This is performed for each line until the end of the current slice. At the end of the current slice, it will traverse that slice in reverse order. This interweaving of forward and backward iterations greatly reduces the overall computation time: during backward iterations, only the neighbors visited in the preceding forward iteration need to be considered; similarly, during forward iterations, only the neighbors visited in the previous backward iteration need to be considered. In the experiments presented in the Results section, the number of iterations were typically in the range of 5 to 10 before convergence.

The  $F^*$  algorithm keeps track of the current optimal cost for each voxel as well as the length and the origin node of the path associated with this optimal cost. Note that this basic  $F^*$  algorithm is only suitable for graph traversal problems where the cost of traversing an edge is a single scalar.

## 2.2 Local diffusion model

We use orientation distribution functions (ODF) for our local diffusion model. The ODF image can be computed directly from a diffusion-weighted image

(DWI) using the technique presented by Descoteaux et al. [4]. Briefly, this method provides an analytical solution by modeling the diffusion imaging signal with a spherical harmonic basis that incorporates a regularization term based on the Laplace-Beltrami operator. The ODF image thus provides a continuous diffusion distribution function at each voxel of a 3D image (Fig. 1A). This ODF image is sampled at each voxel into a spherical sampled distribution using a standard electrostatic repulsion scheme[9], which allows for a high number of directions to achieve the desired level of accuracy.

### 2.3 Multi-directional F\* graph traversal

In order to accommodate the multi-directional local diffusion model, we extend the basic scalar F\* algorithm to keep track of multiple directions at each voxel (represented by a graph vertex). Each voxel visit consists of updating the cost for each sampled ODF direction of the current voxel based on the costs of its neighbors (graph vertices connected by an edge to the current vertex). To this end, for each sampled ODF direction *out*, for each neighbor *n*, the cost of arriving to *n* in direction *in* and traversing an additional edge along *out* is computed. The minimum cost thus obtained is:

$$C_{current}^{out} = \min_{\substack{n=neighbor \\ in=direction}} \left[ C_n^{in} + cost(in, out, d_n, ODF_{current}, ODF_n) \right] \quad (1)$$

where  $d_n$  represents the vector from the voxel-center of *current* to the voxel-center of *n*. The  $ODF_{voxel}^{dir}$  terms refer to the value of the input orientation distribution function sampled at direction *dir* for location *voxel*. Figure 1B illustrates each of the voxels and vectors that factor into this cost function. *C* is initialized to 0 for the seed voxels and to  $\infty$  for all other voxels.

In contrast with the scalar F\* algorithm, the multi-directional F\* algorithm keeps track of the current optimal cost for each voxel *along each sampled ODF direction*, as well as the length and the starting vertex (i.e. origin) of the paths associated with each of these optimal costs at each voxel along each sampled ODF direction. This results in significantly larger memory costs than the scalar F\* algorithm; however, the computation time is nonetheless considerably smaller than Monte Carlo approaches.

### 2.4 Length-independent connectivity cost

A major shortcoming of current connectivity assessment methods is the length-dependency of the resulting connectivity strengths; longer tracts are biased against in stochastic tractography, Hamilton-Jacobian and most graph-based methods. This essentially means that especially the Hamilton-Jacobi and graph-based algorithms may favor unrealistic “shortcuts” rather than longer paths that may be more true to the anatomy and local tensor alignment. The same weakness also applies to the probability formulation of our algorithm presented in

[2]. However, the cost formulation we present here lends itself easily to length-independent connectivity computation. This is achieved by considering the minimum *average* cost of a path rather than the minimum value of the accumulated cost itself (see Fig. 2).

$$\hat{C}_{current}^{out} = \min_{\substack{n=neighbor \\ in=direction}} \frac{C_n^{in} + cost(in, out, \mathbf{d}_n, ODF_{current}, ODF_n)}{L_n^{in} + |\mathbf{d}_n|} \quad (2)$$

where  $L_n^{in}$  represents the length of the current best path to neighbor  $n$  at direction  $in$ , which is the sum of the lengths of each edge along the path from the seed to  $n$ . The length of a graph edge will be equal to 1 for an edge between voxels with a common face,  $\sqrt{2}$  for an edge between voxels with only one common edge, and  $\sqrt{3}$  for an edge between voxels with only one common vertex.

With this modification, the cost of a node will only be updated if the necessary edge traversal cost is less than the average cost of the path, thus reducing the average cost of the path. This allows for longer paths to remain possible even if less anatomically accurate “shortcuts” exist.

## 2.5 Cost of traversing a graph edge

The cost function used in Eqn. 2 is based on the diffusion in the current and neighboring voxel as well as the consistency of the incoming direction  $in$ , outgoing direction  $out$  and the direction and distance towards the neighbor  $n$ :

$$cost(in, out, \mathbf{d}_n, ODF_{current}, ODF_n) = |\mathbf{d}_n| * \left[ \alpha * [f(ODF_{current}, out) + f(ODF_n, in)] + [Penalty(in, out) + Penalty(in, \mathbf{d}_n) + Penalty(out, \mathbf{d}_n)] \right] \quad (3)$$

where  $Penalty$  is an angle penalty function between two given directions and  $f$  is a diffusion weighting function, both of which are further discussed below.  $\alpha$  is a weight that can be set based on the chosen  $f$  and the parameters of the  $Penalty$  function. In most of our experiments, we have simply opted for equal weighting of the two terms, i.e.  $\alpha = 1$  which we empirically found to yield satisfactory results. A more thorough empirical evaluation of this parameter was conducted for the synthetic datasets where ground truth was known. As can be expected from studying the above equation, we found that lower values of  $\alpha$  rely heavily on path consistency and may be appropriate for processing noisy images where this can serve as a regularization term. Higher values of  $\alpha$ , on the other hand, assign more importance to the raw data than to path consistency; this may be an appropriate setting for images with higher SNR ratio and/or when curvier fiber bundles are being investigated.

**Angle penalty function.** We use an angle penalty function to allow the propagation of the diffusion only between two directions that are close to each other, that is to say, that have a small angle between them. This is to prevent the propagation from going backward and less likely along orthogonal paths. This function

penalizes paths that are highly curved and reduces the likelihood of ‘switching’ paths at crossing fiber voxels. This penalty metric is based on the angle between the two directions and has a value between 0 (no loss in propagation) and 1 (no propagation allowed):

$$Penalty(\mathbf{v}_1, \mathbf{v}_2) = \frac{1}{2} * (1 - \text{erf}(\frac{\mathbf{v}_1 \cdot \mathbf{v}_2 - \mu}{\sigma})) \quad (4)$$

The  $\mu$  and  $\sigma$  parameters determine the mean and spread of the error function. We choose  $\mu = 0$  since we want the penalty function to be centered around a 90 degree angle, with going “backward” having the highest penalty. Figure 1C illustrates the angle penalty for  $\sigma = 1$ , as used for all results presented in the next section. The angle penalty function may allow for a tighter or a wider propagation along crossing fibers based on  $\sigma$ . If the function allows for a too wide angle of propagation, there will be almost no loss of connectivity when crossing over to other fibers. Our tests so far showed that the results are rather stable with respect to the  $\sigma$  parameter. Further testing and comparison with alternative penalty functions remain for future work.

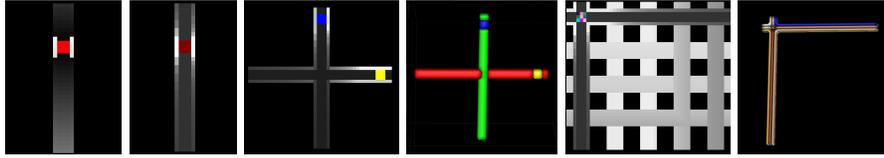
**Diffusion weighting function  $f(ODF)$ .** Since we want the cost at a voxel in a given direction to be inversely proportional to the amount of diffusion at that voxel along that direction, we need to define a diffusion weighting function that will transform the input ODF data. Given the ODF  $\phi$ , and the direction  $d$ , we have initially tried the simplest possible choice by using  $f(\phi, \mathbf{d}) = 1 - \phi[\mathbf{d}]$ , which has, as expected, yielded great computational efficiency. However, we found that this diffusion weighting function was too simplistic and did not accurately capture the underlying data. In particular, consider two ODF’s, one with two peaks,  $O_1$ , and one with a single peak,  $O_2$  (on the unit hemisphere). The above weighting function would assign the same value to the direction along one of the peaks in  $O_1$  and a direction that is relatively separated from the single peak in  $O_2$ . The desired behavior, clearly, is to assign a smaller cost to directions that are the preferred directions of diffusion according to the ODF. For this reason, we have used a Finsler-type weighting function modeled after Melonakos et al. [10]. Given the total number of sampled directions  $nDirs$ , we define the following weighting function:

$$f(\phi, \mathbf{d}) = 1 - \frac{nDirs * \phi[\mathbf{d}]}{\sum_{dir} \phi[dir] \cdot \mathbf{d}_{\perp}} \quad (5)$$

where  $\mathbf{d}_{\perp}$  represents the normal to the ODF surface at  $\mathbf{d}$ . This function assigns weights based on the ratio of diffusion along the current direction  $\mathbf{d}$  to diffusion along all other sampled directions projected to the plane whose normal  $\mathbf{d}$  represents.

### 3 Results

**Simple Synthetic Data.** We first evaluate our algorithm on three synthetic 3D examples (Fig. 2): a) a single fiber tract, b) two crossing fiber tracts with



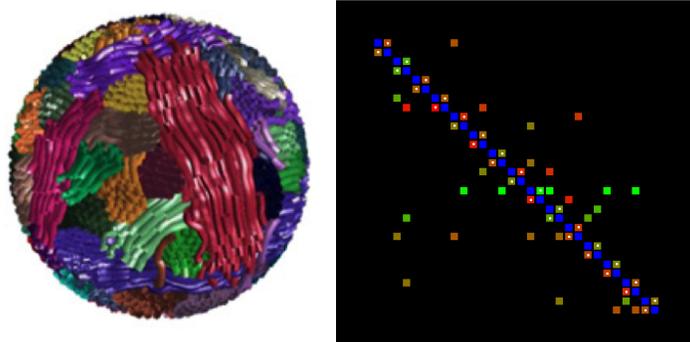
**Fig. 2.** *Single fiber:* minimum cost (a), average cost (b). Minimum cost increases away from the source, while the average cost remains constant. *Crossing fibers* with two seed regions. The average cost map(c) and the volumetric rendering of the paths originating from each seed region (d) clearly demonstrate the algorithm can resolve the crossing fibers. Note that in (b) and (c), the high cost of taking the expensive low-diffusion path is distributed over longer fibers, which explains the gradient along the outer band. *Grid dataset.* The average cost map (e) shows only the fibers that go through the source can be cheaply reached; the other fibers, although accessible, require multiple sharp turns that do not agree with the diffusion data. The 3D rendering of the origins of the “cheapest” paths (f) shows each path can be traced back to the originating voxel.

the source in each one of the two tracts and c) a grid of fiber tracts with a single source. Computation time for all synthetic datasets was under 2 minutes on a standard linux 64 workstation for 92 sampled directions for the ODF.

For the single source and tract settings, the diffusion propagates along the synthetic fibers; while the minimum cost of a path increases with distance to the source, the average cost remains the same. The crossing fiber synthetic example illustrates how well our method handles/propagates along crossing fibers as well as the ease of obtaining volumetric renderings of the paths as a side product of the algorithm. Here, we thresholded the average cost map to create a “propagation map”. Fig. 2d shows a volumetric rendering of the seed labels associated with each voxel inside this propagation map. Clearly, the algorithm was able to successfully resolve the crossing fibers, while keeping track of the origin of each track, its length, as well as its average cost, all in a very efficient manner.

The grid example tests a multiple crossing fibers situation (Fig. 2e and 2f). The 3D rendering clearly shows that the algorithm was able to resolve the crossing fibers successfully, as well as distinguish between each voxel in the source. The average cost map shows that the cost propagation was effectively carried out to the entire grid, even though the cost was high to reach the regions where multiple sharp turns that deviate from the local diffusion model are necessary.

**Numerical Fiber Generator.** While the simple synthetic data presented above illustrates the basic function of our algorithm and is useful for demonstrating the key concepts, it is much too simplistic to be used for validation purposes. In real life, fibers seldom align perfectly with grid axis and images have considerable noise. For these reasons, we have tested our algorithm using a publicly available synthetic dataset created by a software, Numerical Fiber Generator (NFG)[3] for generating random numerical structures consisting of densely packed bundles of fibers, including kissing and crossing fibers, fibers of

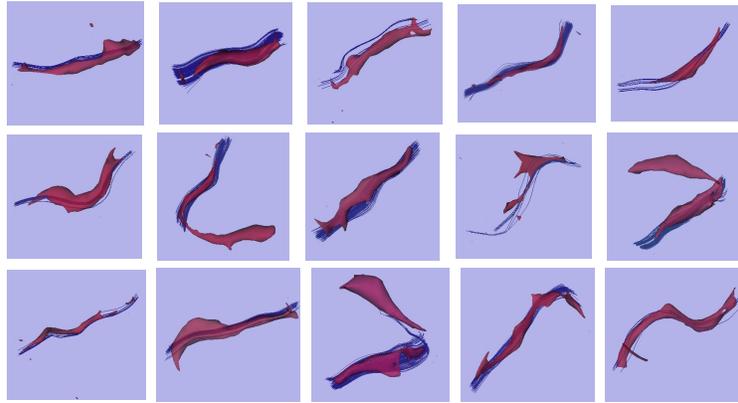


**Fig. 3.** NFG data. Left, a representation of the phantom A (figure courtesy of Close et al.) Right, connectivity matrix between the 30 seed regions. The color of each square represents the computed cost (red=high, green=low); only connections cheaper than the true connection are shown for each seed. Blue squares represent self-connections (0 cost); squares marked with white dots are true connections.

various radii and realistic noise. For our experiment, we have used the phantom A distributed with the NFG software, which consists of a matrix size  $20 \times 20 \times 20$ , packing 15 fiber bundles in various configurations (Fig. 3).

In this experiment, we have computed connectivity from each one of the 30 seed regions using 160 sampled directions. To allow for the recovery of the curvier paths, we set  $\alpha = 8$ . The computation time was under 5 minutes. We have also computed streamline tractography (using Slicer3, [www.slicer.org](http://www.slicer.org)) for illustrating the ‘true’ paths. We have filtered the computed streamlines to only show the strands that pass through the target region. For fair comparison, we have added the two cost maps starting from the source and target regions. Figure 4 shows the volume rendering of the cost maps overlaid with the streamline results for all 15 fiber bundles. The cost maps were thresholded at 8 percent of the maximum cost value, an empirically set value that needs to be further investigated. We further constructed a connectivity matrix by computing connectivity cost from each source region and averaging the cost at each target region. The resulting connectivity matrix (Fig. 3 right), indicates that all true connections are recovered, with very few “false positives”.

**Primate Data.** We applied our method to *in-vivo* primate data as a first application on real datasets. Ten rhesus monkeys of age 12 months were scanned on a 3T Siemens Trio scanner with 8-channel phase array trans-receiving volume coil. DWI were acquired along 60 directions with voxel size of  $1.3 \times 1.3 \times 1.3 \text{mm}^3$ , with a matrix size of  $128 \times 128 \times 80$ . A population average was computed via the DWI atlas building method described in [11]. The ODF image was computed on this DWI atlas. Source regions were manually placed in the internal capsule (3 consecutive axial slices), in the genu of the corpus callosum (CC) (single sagittal slice) and in the whole length of the CC (single sagittal slice). Fig. 5 shows the



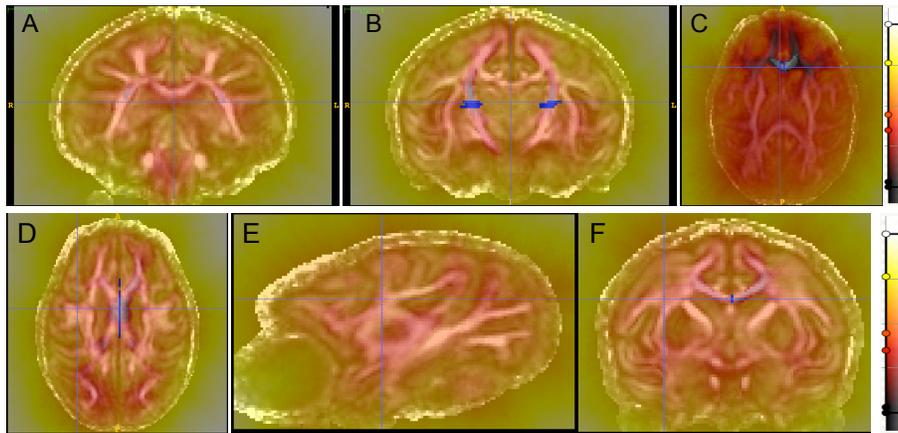
**Fig. 4.** Individual fiber bundle paths recovered from the NFG phantom. The cost maps were thresholded and rendered using marching cubes (red). For each seed region, we also show the streamline tractography results for illustration purposes (blue).

computed average costs, using 40 sampled directions on the unit sphere. These computations took approximately 80 minutes. Note that all the primate cost maps were computed with the same parameters and resulted in comparable cost values despite the differences in ROI size as well as tract length and shape. It should be noted that for this dataset, the atlas building process was done using co-registration based on a single tensor representation. This leads to a rather strong smoothing of the ODF data and the derived connectivity maps.

Importantly, note the lateral projections of the genu into the frontal lobe; this clearly demonstrates that our method is able to successfully resolve *crossing fibers*. Similarly, the internal capsule connectivity maps show large lateral sections of the cerebrospinal tract, which are only accessible via crossing fibers.

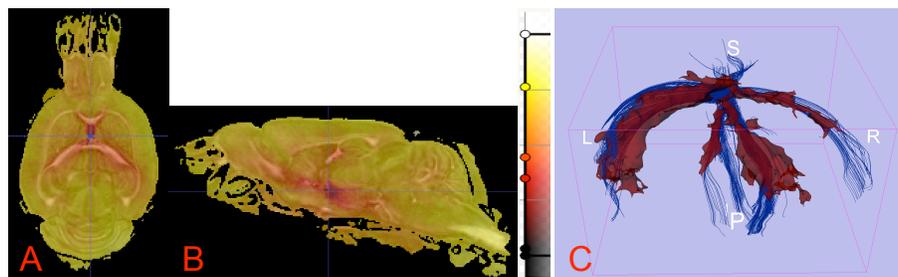
**Mouse Data.** We next applied our connectivity method to *ex-vivo* mouse data to test its performance at very high resolution settings. The DWI were acquired along 42 directions with a voxel size of  $59 \times 71 \times 75$  microns, leading to a matrix size of  $159 \times 288 \times 188$ . A seed region was placed on a single coronal slice for the fornix region. We computed connectivity using 20 sampled directions, which took approximately 4 hours. Comparison of this with the performance on the monkey dataset shows that the algorithm’s run time is mainly driven by the number of sampled directions (approximately quadratic), followed by spatial resolution (approximately linear), which is consistent with the theoretical analysis. Fig. 6 shows the computed average cost, overlaid with the streamline tractography results for comparison purposes. Note the left-right asymmetry in the streamline tractography, which is accurately captured by the cost map.

**Human data.** The final dataset consists of a single healthy adult volunteer scanned with the same protocol at 3 different b-values (1000, 2000, 3000) with

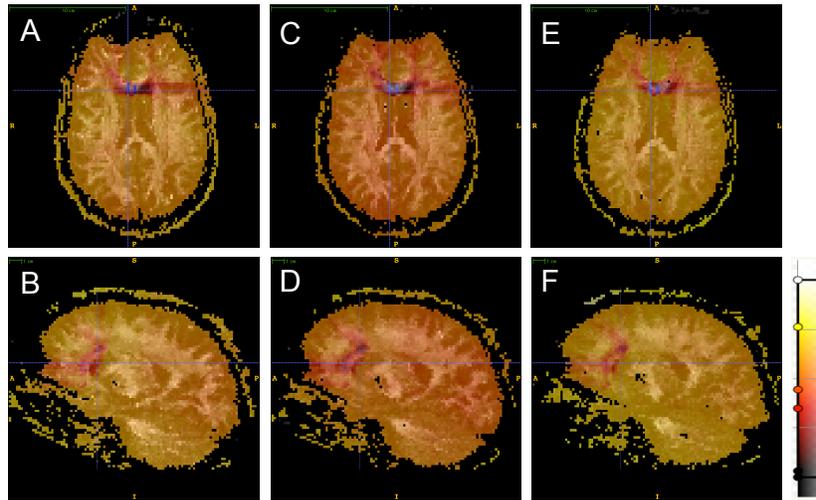


**Fig. 5.** Average cost map computed on the rhesus monkey DWI atlas with the source located in the internal capsule (A, B); the genu of the CC (C); and the full length of the CC (D, E, F). The grayscale overlays show the FA map, the heat map shows average cost. The source regions are shown in blue.

a Siemens Tim Trio 3T scanner. DWI were acquired along 64 directions with an isotropic voxel size of 2mm, with a matrix size of  $106 \times 106 \times 76$ . After rigidly registering the three scans, a manually placed ROI in the genu of the corpus callosum was used as source. The average cost maps, computed using the exact same parameters (20 sampled directions on the unit sphere, computation time 10-12 minutes) as well as the same ROI, are visualized in Figure 7 using the same color map. As expected, the connectivity maps appear less smoothed and we find sharper connectivity as the b-value increases. Additionally, this dataset further illustrates both the robustness of our algorithm to noise and the comparability of our connectivity metric across different scan parameters, ROI and brain sizes, tract properties, and species.



**Fig. 6.** Average cost map for fornix on a mouse. (A) axial view, (B) sagittal view (C) 3D rendering of the thresholded cost map (red) and streamline tractography (blue). The grayscale overlays show the FA map, the heat map shows the average cost.



**Fig. 7.** The average cost maps for the genu tract on the same human subject for different b-values: 1000, 2000, 3000 (left to right; top: axial view, bottom: sagittal view). Note that the connectivity map becomes sharper with higher b-values, in addition to more robust tracking.

## 4 Discussion

We present a new method for the computation of diffusion imaging based white matter connectivity. This method is efficient and resilient to noise, handles multiple seed regions straightforwardly and works well in presence of crossing fiber tracts. The proposed method is generic and could be easily applied to non-diffusion data as long as local directional data can be derived.

A significant contribution of this method is the ability to recover longer fiber tracts, which are typically biased against due to accumulated cost or loss of probability. Our average cost metric overcomes these difficulties. While in theory the average cost metric will in fact favor longer tracts (if the propagation cost is infinitesimally small while the tract length increases, the average cost will be reduced by following a long tract), these situations will only occur in synthetic data since the propagation cost is not negligible in real data where perfect grids of tracts colinear with the diffusion axes do not exist. Therefore, our method in practice offers a length-independent connectivity metric as demonstrated by the presented datasets. Such a metric is far more suitable than length-dependent metrics for subject-specific studies of neuropathology as well as network studies where the connectivity strengths of various ROI's need to be compared. Additionally, it should be noted that the connectivity maps, rather than thresholded binary fiber reconstructions (which are used here for visualization purposes), are the primary product of this method targeting network studies.

Currently we are performing more thorough evaluation studies of the method using human, non-human primate as well as rodent imaging data. Furthermore,

we are working on the empirical evaluation of parameters and the improvement of the ODF weight function and the angle penalty function.

**Acknowledgments.** This work is funded by the following funding sources: R41 NS059095, R21 MH084132, P50 MH078105, P50 MH078105-01A2S1, RC1-AA019211, P01-DA022446, U01-AA020022, NAMIC U54 EB005149, R01 MH091645, ACE R01 HD055741, IDDRRC P30 HD03110.

## References

1. Behrens, T., Woolrich, M., Jenkinson, M., Johansen-Berg, H., Nunes, R., Clare, S., Matthews, P., Bray, J., Smith, S.: Characterization and propagation of uncertainty in DW-MRI. *Magnetic Resonance in Medicine* 50(5), 1077–88 (Jan 2003)
2. Boucharin, A., Oguz, I., Vachet, C., Shi, Y., Sanchez, M., Styner, M.: Efficient, graph-based white matter connectivity from orientation distribution functions via multi-directional graph propagation. *Proc. SPIE* 7962, 79620S (2011)
3. Close, T.G., Tournier, J.D., Calamante, F., Johnston, L.A., Mareels, I., Connelly, A.: A software tool to generate simulated white matter structures for the assessment of fibre-tracking algorithms. *NeuroImage* 47(4), 1288–300 (Oct 2009)
4. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical Q-ball imaging. *Magn Reson Med* 58(3), 497–510 (Sep 2007)
5. Draganski, B., Kherif, F., Klöppel, Cook, P., Alexander, D., Parker, G., Deichmann, Ashburner, Frackowiak, R.: Evidence for segregated and integrative connectivity patterns in the human basal ganglia. *J Neuroscience* 28, 7143–52 (2008)
6. Fischler, M., Tenenbaum, J., Wolf, H.: Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Readings in computer vision: issues, problems, principles, and paradigms* (1987)
7. Fletcher, T., Tao, R., Jeong, W., Whitaker, R.: A volumetric approach to quantifying region-to-region WM connectivity in DT-MRI. *IPMI* pp. 346–58 (2007)
8. Iturria-Medina, Y., Canales-Rodríguez, E.J., Melie-García, L., Valdés-Hernández, P.A., Martínez-Montes, E., Alemán-Gómez, Y., Sánchez-Bornot, J.M.: Characterizing brain anatomical connections using DWMRI and graph theory. *NeuroImage* 36(3), 645–60 (2007)
9. Jones, D.K., Horsfield, M.A., Simmons, A.: Optimal strategies for measuring diffusion in anisotropic systems by MRI. *Magn Reson Med* 42(3), 515–25 (1999)
10. Melonakos, J., Mohan, V., Niethammer, M.: Finsler tractography for white matter connectivity analysis of the cingulum bundle. *MICCAI* (Jan 2007)
11. Niethammer, M., Shi, Y., Benzaid, S., Sanchez, M., Styner, M.: Robust model-based transformation and averaging of DWI applied to DW atlas construction. *MICCAI-CDMRI* (2010)
12. Pichon, E., Westin, C.F., Tannenbaum, A.R.: A Hamilton-Jacobi-Bellman approach to high angular resolution diffusion tractography. *MICCAI* 8, 180–7 (2005)
13. Sotiropoulos, S.N., Bai, L., Tench, C.R.: Fuzzy anatomical connectedness of the brain using single and multiple fibre orientations estimated from diffusion MRI. *Comput Med Imaging Graph* 34(6), 504–13 (Sep 2010)
14. Styner, M., Coradi, T., Gerig, G.: Brain morphometry by distance measurement in a non-Euclidean, curvilinear space. *IPMI* (Jun 1999)
15. Zalesky, A.: DT-MRI fiber tracking: a shortest paths approach. *IEEE Trans Med Imaging* 27(10), 1458–71 (Oct 2008)