PhD THESIS

prepared at
Inria Sophia Antipolis - Méditerranée

and presented at the
University of Nice-Sophia Antipolis
Graduate School of Information and Communication Sciences

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF SCIENCE
Specialized in Control, Signal and Image Processing

Compressive Sensing in diffusion MRI

Sylvain MERLET

Advisor	Dr. Rachid Deriche, Inria Sophia Antipolis-Méditerranée, France
Reviewers	Dr. Christian Barillot, CNRS/Inria Rennes Bretagne-Atlantique, France
	Pr. Ragini Verma, University of Pennsylvania, USA
Examiners	Pr. Maxime Descoteaux, Université de Sherbooke, CA
	Dr. Laure Blanc-Féraud, CNRS/Inria Sophia Antipolis-Méditerranée, France
	Pr. Nikos Paragios, Ecole Centrale de Paris, France

UNIVERSITÉ NICE-SOPHIA ANTIPOLIS - UFR Sciences

École Doctorale STIC
(Sciences et Technologies de l'Information et de la Communication)

THÈSE

pour obtenir le titre de
DOCTEUR EN SCIENCES
de l'UNIVERSITÉ de Nice-Sophia Antipolis
Discipline: Automatique, Traitement du Signal et des Images

présentée et soutenue par

Sylvain MERLET

Compressive Sensing en IRM de diffusion

Thèse dirigée par Rachid DERICHE

Soutenue le 11 septembre 2013

Composition du jury:

Directeur	Dr. Rachid Deriche, Inria Sophia Antipolis-Méditerranée, France
Rapporteurs	Dr. Christian Barillot, CNRS/Inria Rennes Bretagne-Atlantique, France
	Pr. Ragini Verma, University of Pennsylvania, USA
Examinateurs	Pr. Maxime Descoteaux, Université de Sherbooke, CA
	Dr. Laure Blanc-Féraud, CNRS/Inria Sophia Antipolis-Méditerranée, France
	Pr. Nikos Paragios, Ecole Centrale de Paris, France

Abstract

This thesis is dedicated to the development of new acquisition and processing methods in diffusion MRI (dMRI) to characterize the diffusion of water molecules in white matter fiber bundles at the scale of a voxel. In particular, we focus our attention on the accurate recovery of the Ensemble Average Propagator (EAP), which represents the full 3D displacement of water molecule diffusion. Diffusion models such that the Diffusion Tensor or the Orientation Distribution Function (ODF) are largely used in the dMRI community in order to quantify water molecule diffusion. These models are partial EAP representations and have been developed due to the small number of measurement required for their estimations. It is thus of utmost importance to be able to accurately compute the EAP and order to acquire a better understanding of the brain mechanisms and to improve the diagnosis of neurological disorders. Estimating the full 3D EAP requires the acquisition of many diffusion images sensitized to different orientations in the q-space, which render the estimation of the EAP impossible in most of the clinical dMRI scanner. A surge of interest has been seen in order to decrease this time for acquisition. Some works focus on the development of new and efficient acquisition sequences. In this thesis, we use sparse coding techniques, and in particular Compressive Sensing (CS) to accelerate the computation of the EAP. Multiple aspects of the CS theory and its application to dMRI are presented in this thesis.

Keywords diffusion MRI; compressive sensing; sparse coding; dictionary learning; q-space sampling; q-ball imaging; ensemble average propagator; diffusion spectrum imaging;

Résumé (en français)

Cette thèse est consacrée à l'élaboration de nouvelles méthodes d'acquisition et de traitement de données en IRM de diffusion (IRMd) afin de caractériser la diffusion des molécules d'eau dans les fibres de matière blanche à l'échelle d'un voxel. Plus particulièrement, nous travaillons sur un moyen de reconstruction précis de l'Ensemble Average Propagator (EAP), qui représente la fonction de probabilité de diffusion des molécules d'eau. Plusieurs modèles de diffusion tels que le tenseur de diffusion ou la fonction de distribution d'orientation sont très utilisés dans la communauté de l'IRMd afin de quantifier la diffusion des molécules d'eau dans le cerveau. Ces modèles sont des représentations partielles de l'EAP et ont été développés en raison du petit nombre de mesures nécessaires à leurs estimations. Cependant, il est important de pouvoir reconstruire précisément l'EAP afin d'acquérir une meilleure compréhension des mécanismes du cerveau et d'améliorer le diagnostique des troubles neurologiques. Une estimation correcte de l'EAP nécessite l'acquisition de nombreuses images de diffusion sensibilisées à des orientations différentes dans le q-space. Ceci rend son estimation trop longue pour être utilisée dans la plupart des scanners cliniques. Dans cette thèse, nous utilisons des techniques de reconstruction parcimonieuses et en particulier la technique connue sous le nom de Compressive Sensing (CS) afin d'accélérer le calcul de l'EAP. Les multiples aspects de la théorie du CS et de son application à l'IRMd sont présentés dans cette thèse.

Keywords IRM de diffusion; acquisition compressée; reconstruction parcimonieuse; apprentissage de dictionnaire; propagateur de diffusion;

Contents

I Introduction **15**

1 Introduction **17**
1.1 Context . . . 18
1.2 Organization of this manuscript . . . 19
1.2.1 Part 1: Background . . . 19
1.2.2 Part 2: Contributions . . . 20
1.2.3 Appendix . . . 21

2 Introduction (en français) **23**
2.1 Contexte . . . 24
2.2 Organisation du manuscrit . . . 25
2.2.1 Partie 1: Background . . . 25
2.2.2 Partie 2: Contributions . . . 26
2.2.3 Annexe . . . 27

II Background **29**

3 Diffusion Weighted MRI **31**
3.1 Physics in Magnetic Resonnance Imaging . . . 32
3.1.1 Basics understanding of nuclear magnetization . . . 32
3.1.2 Image formation . . . 34
3.1.3 Adding diffusion weighted and the gradient echo pulse sequence . . . 37
3.2 Modeling the diffusion MRI signal and its features . . . 38
3.2.1 Principles of diffusion and scalar quantification of the phenomenon . . . 39
3.2.2 3D probability distribution of water molecule diffusion . . . 41
3.3 Conclusion . . . 52

4 Compressive Sensing and dMRI **53**
4.1 Beyond the Shannon theory . . . 54
4.1.1 The Shannon theorem . . . 54
4.1.2 A surprising result . . . 55
4.2 The discrete Compressive sensing theory . . . 57
4.2.1 CS and Sparsity . . . 58
4.2.2 CS and incoherence . . . 60
4.2.3 CS recovery . . . 61
4.3 Generalization to a broader class of representations . . . 62
4.3.1 CS with overcomplete and redundant dictionary . . . 62
4.3.2 Toward extending CS to recover continuous signals . . . 63
4.4 A state of the art of CS in dMRI . . . 65
4.4.1 CS to reconstruct the diffusion tensor . . . 67
4.4.2 CS to reconstruct the ODF . . . 67
4.4.3 CS to reconstruct the 3D diffusion signal and underlying features . . . 68

4.5 Summary 70

III Contributions 71

5 Accelerated DSI via Compressive Sensing 73
5.1 Motivations 74
5.2 Method 75
5.2.1 Sparse representation of the EAP 75
5.2.2 Features used to characterize the EAP in this chapter 77
5.2.3 Synthetic data simulation 77
5.2.4 Real data acquisition 78
5.2.5 Validation 78
5.3 Results 78
5.3.1 Choice of sampling schemes 79
5.3.2 Experiments on synthetic data 80
5.3.3 Experiments on real data 84
5.4 Discussion 85
5.5 Conclusion 89

6 Continuous diffusion signal estimation via Compressive Sensing 91
6.1 Motivations 92
6.2 Bases for Diffusion Signal Modeling 94
6.2.1 Diffusion signal modeling 94
6.2.2 EAP and ODF modeling 95
6.3 CS properties 97
6.3.1 On the incoherence property of CS bases 97
6.3.2 On the sparsity of CS bases 99
6.3.3 CS recovery 101
6.4 Experimental results 102
6.4.1 Sampling protocol 104
6.4.2 Synthetic and noisy reconstruction 108
6.4.3 Real human brain data 112
6.4.4 Phantom data 113
6.5 Conclusion 116
6.6 Appendix 117
6.6.1 Analytical ODF solution when the signal is modeled in SPF dual basis 117
6.6.2 Analytical ODF solution when the signal is modeled in SHORE basis 118

7 Parametric and sparse dictionary learning 121
7.1 Motivations 122
7.2 A framework for the recovery of the complete diffusion MRI process 123
7.2.1 Continuous diffusion modeling with a constrained dictionary 123
7.2.2 Closed formulae for diffusion features 125
7.3 A parametric dictionary learning for sparse dMRI 126
7.3.1 Dictionary learning algorithm 127
7.3.2 Signal estimation via the learned dictionary 129

7.4 Experiments on synthetic data . . . 129
7.4.1 Which radial and angular order for the dictionary ? . . . 130
7.4.2 Training phase . . . 131
7.4.3 Validation . . . 132
7.4.4 Discussion on experiments with synthetic data . . . 138
7.5 Experiments on real data . . . 139
7.5.1 Learning and reconstruction on a 7T scanner data . . . 141
7.5.2 Learning and reconstruction on a 3T scanner data . . . 141
7.6 Conclusion . . . 144
7.7 Appendix . . . 144
7.7.1 Derivation of the normalization constant . . . 144
7.7.2 Derivation of the Ensemble Average Propagator . . . 145
7.7.3 Derivation of the Orientation Distribution Function . . . 146

IV Conclusion 149

8 Conclusions and Perspectives 151
8.1 General conclusion . . . 152
8.1.1 Local diffusion estimation . . . 152
8.1.2 The multiple aspects of the Compressive Sensing theory . . . 152
8.2 Applications and Collaborations . . . 152
8.2.1 Fiber tractography via the Ensemble Average Propagator . . . 153
8.2.2 Software . . . 153

9 Conclusions (en français) 155
9.1 Conclusion générale . . . 156
9.1.1 L'estimation de la diffusion locale . . . 156
9.1.2 Les multiples aspects de la théorie du Compressive Sensing . . . 156
9.2 Applications et collaborations . . . 157
9.2.1 Tractographie de fibres via l'EAP . . . 157
9.2.2 Logiciel . . . 157

10 Appendix 159
10.1 Fourier transform . . . 160
10.2 A world of wavelets . . . 160
10.3 Inner product space . . . 161
10.4 ℓ_p-norm in an n-dimensional Euclidean space . . . 162
10.5 Orthonormal basis . . . 162
10.6 Overcomplete and redundant dictionary . . . 163
10.7 Analytical diffusion signal, EAP and ODF from the multiple tensor model . 163

11 Publications of the author 165

Bibliography 169

Remerciements

En premier lieu, je tiens à remercier mon directeur de thèse, Rachid Deriche, pour toute l'attention qu'il a pu accorder à mon travail. Je suis ravi d'avoir travaillé en sa compagnie car en dehors de son appui scientifique, il a toujours été là pour me soutenir et me conseiller au cours de l'élaboration de cette thèse.

Je souhaiterais exprimer ma gratitude à Christian Barillot et Ragini Verma pour leurs commentaires constructifs et leurs encouragements, ainsi qu'à Maxime Descoteaux, Laure Blanc-Féraud et Nikos Paragios pour avoir accepté de participer à ce jury de thèse

Je remercie également tous les membres de l'équipe-projet ATHENA avec qui j'ai passe des moments formidables. Je pense à ces après midi piscine chez Rachid, aux sessions d'escalade, aux sorties vélo, et aussi à toutes ces parties de contrée après le repas de midi.

Un grand merci aussi à nos amis Canadiens de l'université de Sherbrooke, Maxime Descoteaux, Michaël Paquette, Gabriel Girard, Marc-Alexandre Côté et Jean-Christophe Houde, qui ont donné un nouveau souffle à mes recherches grâce à leurs volonté d'aller toujours plus loin dans l'application de l'IRMd dans le domaine clinique.

J'aimerais aussi remercier Emmanuel Caruyer, ancien membre de l'équipe-projet ATHENA, pour toute l'aide qu'il m'a apporté lorsque j'en avais besoin.

Enfin, les dernières lignes de ces remerciements sont destinées à remercier Mélissa, toute ma famille et mes amis qui me soutiennent depuis tant d'années.

Sylvain Merlet.

Acronyms

AE	Angular Error
CDF	Cohen-Daubechies-Feauveau
CS	Compressive Sensing
CV	Cross Validation
DNC	Difference in the Number of Compartments
DSI	Diffusion Spectrum Imaging
DTI	Diffusion Tensor Imaging
DTWT	Dual Tree Wavelet Transform
DWI	Diffusion Weighted Imaging
DWT	Discrete Wavelet Transform
EAP	Ensemble Average Propagator
EPI	Echo Planar Imaging
FA	Fractional Anisotropy
FT	Fourier Transform
HARDI	High Angular Resolution Diffusion Imaging
NMSE	Normalized Mean Square Error
ODF	Orientation Distribution Function
TV	Total Variation
RIP	Restricted Isometry Property
SC	Sampling Scheme
SNR	Signal to Noise Ratio
SWT	Stationary Wavelet Transform

Part I

Introduction

Chapter 1

Introduction

1.1 Context

Magnetic Resonance Imaging (MRI) is a medical imaging technique and an application of nuclear magnetic resonance (NMR) that allows one to reconstruct and visualize the internal structure of biological tissues, non invasively. MRI scans measure the current induced in a receiver coil from the magnetic field generated by the rotational movement of nuclei in the body part under examination. Diffusion MRI (dMRI) is a specific MRI acquisition process, which assesses the integrity of brain anatomical connectivity and is very useful for examining and quantifying white matter (WM) microstructure and organization not available with other imaging modalities. dMRI is also crucial for the diagnosis of neurological disorders.

dMRI determines the WM structure by exploiting the way the water molecules diffuse. The first diffusion images were obtained in the mid-1980s by Le Bihan et al. [73, 74], Merboldt et al. [83] , and was based on the pioneering work of [117], who introduced the pulsed gradient spin-echo (PGSE) sequence. It allows the quantification of the water diffusion by estimating the displacement of water particles from the phase change that occurs during the acquisition process. This average displacement can be accurately described at the scale of a voxel by the Ensemble Average Propagator (EAP), which represents the full 3D displacement of water molecule diffusion. Using dMRI to infer the EAP requires the acquisition of many diffusion images sensitized to different orientations in the sampling space. The number of diffusion weighted images (DWI) required depends on how the diffusion is modeled. For instance, the well known Diffusion Tensor (DT) model [8, 10, 9, 32, 125, 31, 4, 75, 54, 98] assumes the EAP is Gaussian and requires at least 6 DWIs plus an additional unweighted image. However, the Gaussian assumption of Diffusion Tensor Imaging (DTI) over-simplifies the diffusion of water molecules in the brain and, thus, has some limitations for voxels in which there are more complicated internal structures. Therefore, it is of utmost importance to develop techniques that go beyond the limitations of DTI. For this purpose, High Angular Resolution Diffusion Imaging (HARDI) has been proposed to measure the diffusion of water molecules along more directions than DTI. Among HARDI techniques, there is Q-Ball Imaging (QBI) [126, 130, 44, 1, 96], which estimates the Orientation Distribution Function (ODF) from measurements taken at the same radii. However, the ODF only captures angular information of the diffusion process. Another HARDI technique has been proposed in [66], where the authors characterize the diffusion signal by a superposition of Gaussians, resulting in a Wishart distribution. They show improvements over the classical DTI technique and presents an estimation scheme for the fiber orientation and EAP.

This thesis is dedicated to the development of new acquisitions and processing methods in order that the diffusion may be completely characterized. In particular, we focus our attention on the accurate recovery of the EAP, which can be considered as a starting point to estimate many diffusion features as the DT and the ODF. In Diffusion Spectrum Imaging (DSI) [131, 130, 127], the EAP is obtained by directly taking the inverse Fourier transform of the normalized signal measured in the q-space. It aims to reconstruct the EAP in a numerical way without any prior knowledge. This results in estimating the EAP in a very accurate fashion. However, DSI provides a discrete EAP and the estimation of diffusion features requires numerical calculation, which can be long and computationally expensive. Recently, multiple shells HARDI techniques were introduced to model the EAP in a continuous and analytical manner [100, 5, 99, 34, 45, 135, 63]. These techniques aim to catch both radial and angular information about the water diffusion process as DSI does,

and they provide analytical formulae to extrapolate/interpolate the dMRI signal as well as to estimate important diffusion features.

In this thesis we tackle the problem of estimating the EAP with a reduced number of measurements. For this purpose, we consider sparse coding techniques, and in particular Compressive Sensing (CS) [23, 46, 52, 22, 49] to accelerate the computation of common water molecule diffusion models. The CS theory is thus the central point of this thesis. It has already been used in numerous applications [6, 77, 17, 48, 25, 61, 133, 13, 118, 134] and, in particular has been proved useful in recovering magnetic resonance (MR) images by significantly undersampling their k-spaces [77, 60, 55, 30, 124]. Many research groups have also handle the problem of estimating the dMRI signal via CS starting from the idea of accelerating DSI [91, 82, 15] to a continuous and analytical modeling of the dMRI signal from sparse basis or dictionary [96, 109, 36, 89, 122, 95]. An interest in dictionary learning has also been seen in order to construct highly sparse representation of the dMRI signal [15, 57, 136, 88]. This thesis contributes to these three topics, i.e., the acceleration of DSI via CS in **chapter 5**, the continuous and analytical modeling of the dMRI signal via CS in **chapter 6** and sparse dictionary learning from dMRI data in **chapter 7**.

1.2 Organization of this manuscript

This manuscript is organized in two parts: a background part and a part containing the contributions.

- The background part presents the context of this thesis, i.e an introduction to diffusion MRI in **chapter 3** and a description of Compressive Sensing (CS) in **chapter 4**.
- The contributions part consists in three chapters. **Chapter 5** describes the method to accelerate DSI via CS. **Chapter 6** deals with the continuous and analytical modeling of dMRI signals with respect to sparse and continuous orthonormal bases. **Chapter 7** tackles the problem of sparse and parametric dictionary learning in diffusion MRI.

Sketch of the chapters in this thesis is presented Fig. 2.1.

1.2.1 Part 1: Background

Chapter 3 introduces the principles of Magnetic Resonance Imaging, with a particular emphasis on diffusion MRI. This chapter provides the knowledge necessary to understand the biological and physical bases of the diffusion MR signal. It also introduces mathematical models that characterize the diffusion phenomenon, namely the Diffusion Tensor (DT), the Orientation Distribution Function (ODF) and the full probability distribution of water molecules diffusion also known as the Ensemble Average Propagator (EAP). We also present state-of-the-art methods to estimate these models.

Chapter 4 introduces the sampling theory of Shannon and describes the Compressive Sensing (CS) theory introduced by Candes and Donoho [23, 46]. The classical CS theory was originally developed to recover discrete and finite-length signals under the assumption that this signal admits a sparse representation with respect to an orthonormal basis. This classical CS theory was found limited in two points. First, in many applications, the

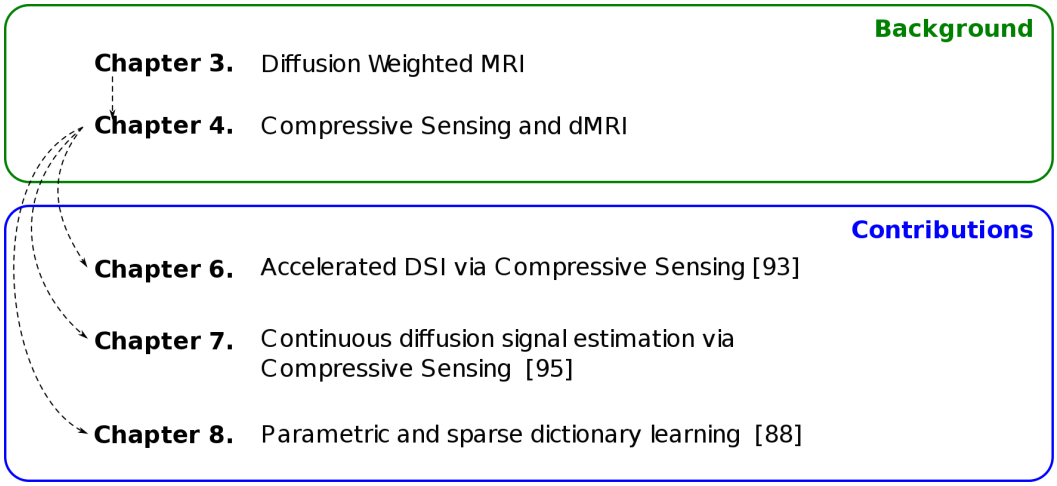


Figure 1.1: Sketch of the chapters in this thesis.

signal is likely to admit a sparse representation with respect to an arbitrary dictionary. In addition, most of the natural phenomena are analog and not discrete. For these two reasons, we present an extension of the classical CS framework from a discrete and finite length signal reconstruction to the recovery of continuous signal representation with respect to orthonormal bases or dictionaries. This chapter is the central point of this thesis and can be used as a link between the different contributions presented in the subsequent chapters.

1.2.2 Part 2: Contributions

Chapter 5 presents an application of the classical CS theory to accelerate the technique known as Diffusion Spectrum Imaging (DSI). DSI is able to accurately recover a discrete EAP defined in a Cartesian grid at the expense of having to acquire a large number of measurements. Compressive Sensing (CS) offers an efficient way to decrease this number of measurements. This chapter covers two main topics of CS, namely the sparsity of the dMRI signal and the q-space sampling protocol. The robustness and the efficiency of several sampling protocols and sparse transforms are thoroughly compared to identify the ingredients of an optimal CS reconstruction. Experiments on synthetic and human brain data have been carried out, which demonstrate that approximately 64 measurements are sufficient to recover significant information regarding the EAP and to accurately estimate the orientation distribution function (ODF) and the diffusion kurtosis. We finally show that an appropriate sparse transform and sampling protocol considerably improves the quality of the CS reconstruction of the EAP, the ODF and the diffusion kurtosis. This contribution has been submitted and is under revision in [93].

Chapter 6 studies the impact of the sparsity, the incoherence and the Restricted Isometry Property (RIP) when the dMRI signal is modeled as a linear combination of an orthonormal basis atoms. This provides a continuous representation of the dMRI signal and extends the classical CS framework presented in **chapter 5**. This continuous framework allows one to interpolate/extrapolate the dMRI signal at every q-space locations and renders the

computation of some diffusion features such that the EAP and the ODF very fast thanks to analytical and closed-form formulae. We show that an efficient use of the CS theory enables the drastic reduction of the number of measurements commonly used in dMRI acquisitions. Only 20-30 measurements, optimally spread on several b-value shells, are shown to be necessary, which is less than previous attempts to recover the diffusion signal using CS. This opens an attractive perspective to measure the diffusion signals in white matter within a reduced acquisition time and shows that CS holds great promise and opens new and exciting perspectives in diffusion MRI. This contribution has been published in [95].

Chapter 7 introduces a new dictionary learning algorithm to design parametric and sparse dictionaries from a set of diffusion MRI signals. This original framework provides a way to obtain a faithful and continuous representation of the dMRI signal while highly decreasing the number of measurements which is required with the orthonormal bases seen in chapter **6**. Analytical formulae are also provided to estimate the EAP and the ODF. The properties and potentials of the technique are demonstrated using various simulations on synthetic data and on human brain data acquired from 7-T and 3-T scanners. It is shown that the technique can clearly recover the dMRI signal and its features with a much better accuracy compared to state-of-the-art approaches, even with a small and reduced number of measurements. In particular, we can accurately recover the ODF in regions of multiple fiber crossing, which could open new perspectives for some dMRI applications such as fiber tractography. This chapter can be considered as an extension of **chapter 6** where the orthonormal basis is replaced by an overcomplete and redundant dictionary learned from a training data set. This contribution has been published in [88].

1.2.3 Appendix

The appendix presents the tools required for a complete understanding of the manuscript. It describes mathematical tools that have to be well understood to handle the theory underlying this thesis. In particular, we describe in Sec. 10.7 the mathematical framework commonly used in dMRI to synthesize signals, namely the multi-Gaussian model.

Chapter 2

Introduction (en français)

2.1 Contexte

L'imagerie par Résonance Magnétique (IRM) est une application médicale de la résonance magnétique nucléaire (RMN) et permet de reconstruire et de visualiser la structure interne des tissus biologiques de façon non invasive. L'IRM mesure le courant induit dans une bobine à partir du champ magnétique généré par le mouvement de rotation de certains noyaux atomiques possédant un spin nucléaire. L'IRM de diffusion (IRMd) est une modalité d'acquisition d'IRM, qui évalue la connectivité cérébrale anatomique et est très utile pour examiner et quantifier la microstructure et l'organisation des fibres de matière blanche (MB). L'IRMd est également cruciale pour le diagnostic des troubles neurologiques.

L'IRMd détermine la structure de la MB, en exploitant la façon dont les molécules d'eau se diffusent. Les premières images de diffusion ont été obtenus dans le milieu des années 80 par Le Bihan et al. [73, 74], Merboldt et al. [83] grâce au travail pionnier de [117], qui a présenté une séquence d'acquisitions basée sur des gradients de très courte durée. Cette découverte permet la quantification de la diffusion de l'eau en estimant le déplacement des particules d'eau à partir du changement de phase qui se produit pendant l'acquisition. Ce déplacement moyen peut être décrit avec précision à l'échelle d'un voxel par le propagateur moyen de diffusion (PMD ou EAP), qui représente le déplacement en 3D de la diffusion des molécules d'eau. Estimer le PMD nécessite l'acquisition de nombreuses images de diffusion sensibilisées à des orientations différentes dans l'espace d'échantillonnage. Le nombre d' images pondérés en diffusion (IPD ou DWI) requis dépend de la façon dont la diffusion est modélisée. Par exemple, le tenseur de diffusion (TD) [8, 10, 9, 32, 125, 31, 4, 75, 54, 98] assume que le PMD est gaussien et nécessite au moins 6 IPDs ainsi qu'une image supplémentaire non pondérée. Cependant, l'hypothèse gaussienne simplifie à l'excès la diffusion des molécules d'eau dans le cerveau. Par conséquent, ce modèle possède quelques limitations pour les voxels dans lesquels ils existent des structures internes plus complexes. Il est donc primordial de développer des modèles qui vont au-delà des limites du tenseur de diffusion. À cette fin, l'imagerie par haute résolution angulaire (HARDI) a été proposé pour mesurer la diffusion de molécules d'eau dans plus de directions que DTI. Parmi les techniques HARDI, nous pouvons citer le Q-Ball Imaging (QBI) [126, 130, 44, 1, 96] qui estime la fonction de distribution des orientations (ODF) à partir de mesures prises à un même rayon. Cependant, l'ODF ne capture que l'information angulaire.

Cette thèse est consacrée à l'élaboration de nouvelles techniques acquisitions et de traitement des données en IRM de diffusion. En particulier, nous concentrons notre attention sur la reconstruction exacte du propagateur de diffusion. Celui ci peut être considéré comme un point de départ afin d'évaluer de nombreuses fonctionnalités de diffusion comme le tenseur de diffusion et la fonction de distribution des orientations. Avec la technique connue sous le nom de Diffusion Spectrum Imaging (DSI), le PMD est obtenu en prenant directement la transformée de Fourier inverse du signal normalisé. Il vise à reconstruire le PMD numériquement sans aucune connaissance préalable. Il en résulte une estimation très précise du PMD. Cependant, le DSI fournit un PMD discret et l'estimation des caractéristiques de diffusion nécessite des calculs numériques, qui peut être long et coûteux en ressources informatiques. Récemment, plusieurs techniques HARDI ont été introduites pour modéliser le PMD de façon continu et d'une manière analytique [100, 5, 99, 34, 45, 135, 63]. Ces techniques visent à approximer l'aspect radial et angulaire du processus de diffusion comme DSI. Elles fournissent également en plus des formules analytiques pour extrapoler

/ interpoler le signal de diffusion et estimer des caractéristiques de diffusion importantes.

Dans cette thèse, nous abordons le problème de l'estimation du PMD avec un nombre réduit de mesures. A cet effet, nous utilisons des techniques de codage parcimonieuses, et en particulier la technique connue sous le nom de Compressive Sensing (CS) [23, 46, 52, 22, 49] afin d'accélérer le calcul des modèles de diffusion. La théorie du CS est donc le point central de cette thèse. Elle a déjà été utilisée dans de nombreuses applications [6, 77, 17, 48, 25, 61, 133, 13, 118, 134] et, en particulier, s'est révélée utile en IRM [77, 60, 55, 30, 124]. De nombreux groupes de recherche ont aussi traité le problème d'estimation du signal de diffusion avec le CS à partir du DSI [91, 82, 15] ou pour modeliser de façon continu le signal dans une base orthonormal ou un dictionnaire [96, 109, 36, 89, 122, 95]. Un intérêt pour l'apprentissage de dictionnaire a également été vu afin de trouver des representations très sparses du signal de diffusion [15, 57, 136, 88]. Cette thèse contribue à ces trois objectifs, à savoir l'accélération du DSI via CS dans le **chapter 5**, la modélisation continue du signal de diffusion via CS dans le **chapter 6** et l'apprentissage de dictionnaire à partir de données pondérées en diffusion dans le **chapter 7**.

2.2 Organisation du manuscrit

Ce manuscrit est composé de deux parties : une partie définissant le contexte et les outils utilisés, et une partie présentant les contributions.

- La première partie présente le contexte de cette thèse, à savoir une introduction à l'IRM de diffusion dans le **chapitre 3** et une description du Compressive Sensing dans le **chapitre 4**. Le
- La partie contributions se compose de trois chapitres. Le **chapitre 5** décrit la méthode pour accélérer le DSI via CS. Le **chapitre 6** traite de la modélisation continue et des signaux de diffusion dans des bases orthonormales. **chapitre 7** aborde le problème de l'apprentissage d'un dictionnaire paramétrique en IRM de diffusion.

Un plan de ces chapitres est presenté Fig. 2.1.

2.2.1 Partie 1: Background

Chapitre 3 présente les principes de l'imagerie par résonance magnétique (IRM), avec un accent particulier sur l'IRM de diffusion. Ce chapitre fournit les connaissances de base nécessaires à la compréhension de la biologie et de la physique de l'IRM de diffusion. Il introduit également les modèles mathématiques qui caractérisent le phénomène de diffusion, à savoir le tenseur de diffusion , la fonction de distribution d'orientation et la distribution de probabilité de la diffusion des molécules d'eau également connu comme l' Ensemble Average Propagator (EAP). Nous présentons également l'état de l'art des méthodes d'estimation de ces modèles.

Chapitre 4 introduit la théorie de l'échantillonnage de Shannon et décrit la théorie Compressive Sensing (CS) introduite par Candes et Donoho [23, 46]. La théorie classique du CS a été initialement développée pour reconstruire des signaux discrets et de longueur finie sous l'hypothèse que le signal admet une représentation parcimonieuse dans une base orthonormale. Cette théorie classique du CS est limitée à deux points. Tout d'abord, dans

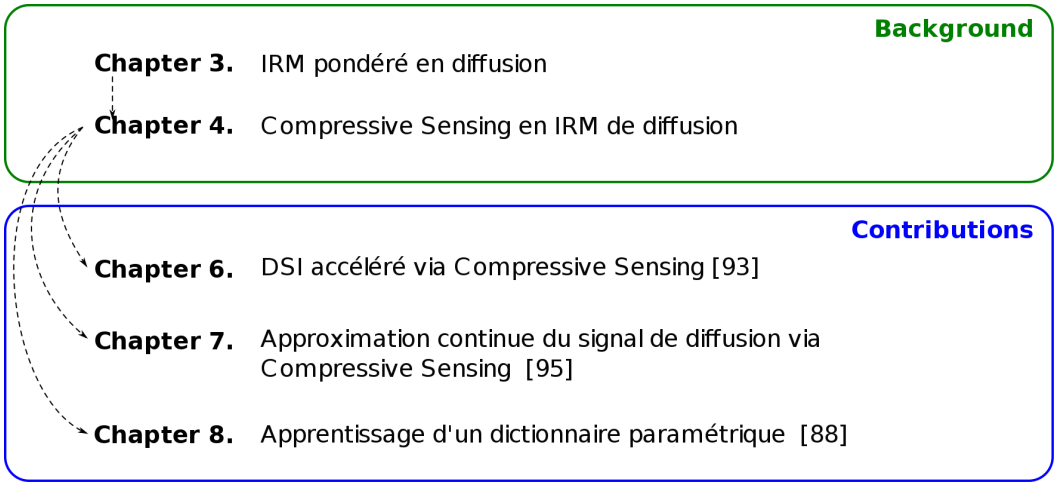


Figure 2.1: Plan des chapitres de ce manuscrit.

de nombreuses applications, le signal est susceptible d'admettre une représentation parcimonieuse dans un dictionnaire quelconque. En outre, la plupart des phénomènes naturels sont analogiques et non discrets. Pour ces deux raisons, nous présentons une prolongation du CS classique pour reconstruire des signaux continus modélisés dans des bases orthonormales et des dictionnaires. Ce chapitre est le point central de cette thèse et peut être utilisé comme un lien entre les différentes contributions présentées dans les chapitres suivants.

2.2.2 Partie 2: Contributions

Chapitre 5 présente une application de la théorie classique du CS pour accélérer la technique appelée Diffusion Spectrum Imaging (DSI). DSI est capable de reconstruire précisément l'EAP discrétisé sur une grille cartésienne au détriment de l'acquisition d'un grand nombre de mesures. CS offre un moyen efficace de réduire ce nombre de mesures. Ce chapitre couvre deux thèmes principaux du CS, à savoir le choix d'une représentation parcimonieuse et le choix d'un protocole d'échantillonnage. La robustesse et l'efficacité de plusieurs protocoles d'échantillonnage et de plusieurs transformées parcimonieuses sont soigneusement comparées.. Des expériences sur des données synthétiques et sur le cerveau humain sont réalisées, et démontrent qu'environ 64 mesures sont suffisantes pour récupérer l'information essentielle de l'EAP et permet d'estimer avec précision l'ODF et le kurtosis de diffusion. Cette contribution a été soumise dans [93].

Chapitre 6 étudie l'impact de la parcimonie, de l'incohérence et de la propriété RIP lorsque le signal est modélisé comme une combinaison linéaire d'atomes d'une base orthonormale. Cela donne une représentation continue du signal de diffusion et étend le cadre du CS classique présentée dans le **chapitre 5**. Ce cadre continue permet d'interpoler / extrapoler le signal de diffusion dans tout le q-space et de rendre le calcul de certaines caractéristiques de diffusion tels que l'EAP et la fonction de distribution des orientations très rapide grâce à des formules analytiques. Nous montrons qu'une utilisation efficace du CS permet de réduire considérablement le nombre de mesures couramment utilisées lors

d'acquisitions en IRM de diffusion. Seulement 20-30 mesures réparties de façon optimale sur plusieurs valeurs de b sont nécessaires. Cela ouvre des perspectives nouvelles afin d'acquérir les signaux en IRM de diffusion. Cette contribution est publiée dans [95].

Chapitre 7 introduit un nouvel algorithme d'apprentissage de dictionnaire afin de concevoir des dictionnaires paramétriques et parcimonieux à partir d'un ensemble de signaux d'IRM de diffusion. Les dictionnaires appris fournissent une représentation fidèle et continu du signal de diffusion tout en diminuant fortement le nombre de mesures. Des formules analytiques sont également fournies pour estimer l'EAP et la fonction de distribution des orientations (ODF). En particulier, nous pouvons récupérer avec précision l'ODF dans les régions de croisement de fibres multiples, ce qui pourrait ouvrir de nouvelles perspectives pour une application telle que la tractographie de fibres. Ce chapitre peut être considéré comme une extension du **chapitre 6** où la base orthonormale est remplacée par un dictionnaire redondant appris à partir d'un ensemble de données de diffusion. Les propriétés et les potentialités de la technique sont mises en évidence via diverses simulations sur des données synthétiques et via des données acquises sur le cerveau humain. Il est démontré que la technique peut clairement récupérer le signal de diffusion et certaines de ses caractéristiques avec une très bonne précision et un nombre réduit de mesures. En particulier, nous pouvons récupérer avec précision l'ODF dans les régions de croisement de fibres multiples. Cette contribution a été publiée dans [88].

2.2.3 Annexe

L'annexe présente les outils nécessaires à une compréhension complète du manuscrit. Il décrit les outils mathématiques, qui doivent être bien compris pour gérer la théorie sous-jacente à cette thèse. En particulier, nous décrivons dans la Sec. 10.7 le cadre mathématique couramment utilisé en IRMd afin de synthétiser les signaux, à savoir le modèle multi-gaussien.

Part II

Background

Chapter 3

Diffusion Weighted MRI

Contents

3.1 Physics in Magnetic Resonnance Imaging 32
3.1.1 Basics understanding of nuclear magnetization 32
3.1.2 Image formation 34
3.1.3 Adding diffusion weighted and the gradient echo pulse sequence . 37
3.2 Modeling the diffusion MRI signal and its features 38
3.2.1 Principles of diffusion and scalar quantification of the phenomenon 39
3.2.2 3D probability distribution of water molecule diffusion 41
3.3 Conclusion 52

Overview

Diffusion MRI (dMRI) is a medical imaging technique that is part of the nuclear magnetic resonance (NMR) applications. In dMRI, NMR provides a way to quantify the water molecule diffusion in biological tissues. NMR was described by Bloch [16] and Purcel [107] in 1946. NMR is based on the observation that magnetic nuclei, often referred to as spins, align themselves to a strong and external magnetic field $\mathbf{B_0}$. After the application of this magnetic field, the MRI scan start the excitation of the nuclei with a 90° radio frequency (RF) pulse that tilts the spins to the plane whose normal is the direction of $\mathbf{B_0}$. Then the spins start to precess around $\mathbf{B_0}$ with an angular frequency depending on the nuclei under examination. The generated echo produced by the precession of all the nuclei induces a current in a receiver coil and is used to produce the MR signal. However, since the magnetic field is not homogeneous, the spins lose their phase coherence, which causes the MR signal to decrease. In 1950, Hahn [59] proposed to reverse the dephasing by applying a 180° RF pulse. In this sequence, the time between the excitation of the spins and the reception of the echo by the MR coil is twice the time between the two RF pulses. The generated MR signal is often called spin-echo MR signal. Hahn also observed that the reduction of the spin-echo signal could be explained by the spin dephasing due to the translational diffusion of the nuclei within an inhomogeneous magnetic field. In 1954, Carr and Purcell [24] proposed a way to quantify the diffusion of nuclei. They modify the Hahn spin-echo sequence and introduce a constant magnetic field gradient, in addition to $\mathbf{B_0}$. The spins that are at different locations in this magnetic field gradient precess at different angular frequencies. From this sequence, Carr and Purcell [24] developed a mathematical framework to quantify the diffusion. In 1963, Stejskal and Tanner [117] introduced the pulsed gradient spin-echo (PGSE) sequence, which replaced the constant magnetic field gradient of the Carr and Purcell sequence by short duration gradient pulses. The PGSE sequence simplified the

mathematical framework underlying the MR signal interpretation and is used in modern MRI systems.

In this chapter, we describe the MR signal and image acquisition. The first section of this chapter aims to understand the physics, which occurs during the measurement process. We will cover the acquisition principles of basic MRI, namely the T1 and T2 weighted MRI. Then, we will see a more specific MRI process: the diffusion MRI (dMRI). In the second section, we present several ways to model the diffusion starting from scalar quantification of the phenomenon to more complex model that enable the estimation of the 3D dMRI signal.

3.1 Physics in Magnetic Resonnance Imaging

3.1.1 Basics understanding of nuclear magnetization

The human body is mostly made of water and the proportion of hydrogen nuclei contained in the water gives relevant information regarding to the internal structure of a body part. Magnetic Resonance Imaging (MRI) systems observe the hydrogen nuclei behavior in order to reflect the water concentration. In this section, we explain the underlying physics necessary to understand MRI acquisition.

Let us firstly explain what is a nucleus. As it is represented in the figure 3.1, a nucleus is a magnetic dipole, also called a spin, which is defined by its rotating axis and the rotational movement around this axis.

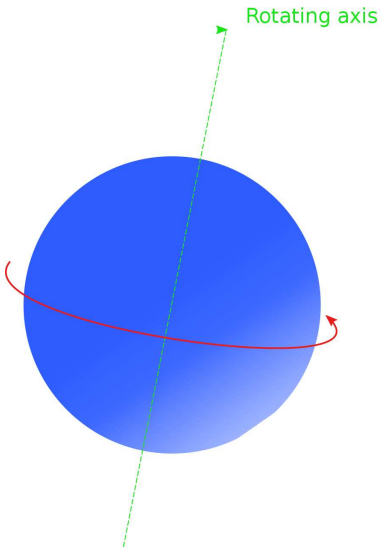


Figure 3.1: Dipole, often referred as a spin. The spin precesses around its rotating axis.

Without external magnetic field (see fig.3.2), every spin has a different direction and the sum of all magnetic moments results in a cancellation of the overall magnetic field. However, the dipole behaviors change when they are prone to an external magnetic field (see fig.3.3). At this moment, every spin aligns with this magnetic field and start precessing about the external field axis with an angular frequency known as the Larmor frequency.

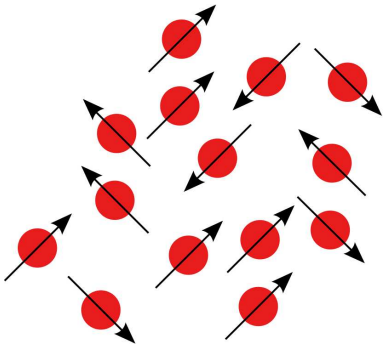


Figure 3.2: Magnetic dipoles, or spins, with different directions.

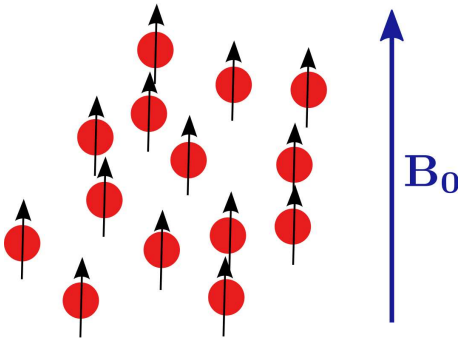


Figure 3.3: Magnetic dipoles, or spins, prone to an external magnetic field $\mathbf{B_0}$. Every spin aligns with this magnetic field.

Therefore, the MRI machine first applies a magnetic field $\mathbf{B_0}$ in order to align the spins. Afterward, a radio frequency pulse, often referred as a RF pulse, is applied in order to tilt the spins in the plan perpendicular to $\mathbf{B_0}$ (3.4 (a)). When the RF pulse frequency is the same as the Larmor frequency of the spins, it tilts the latter in the plan perpendicular to the magnetic field $\mathbf{B_0}$. The particles continue to precess about $\mathbf{B_0}$ at the frequency ν given by the Larmor equation $\nu = \gamma B_0$ where γ is the gyromagnetic ratio. We represent this phenomenon in the Fig. 3.4 (b).

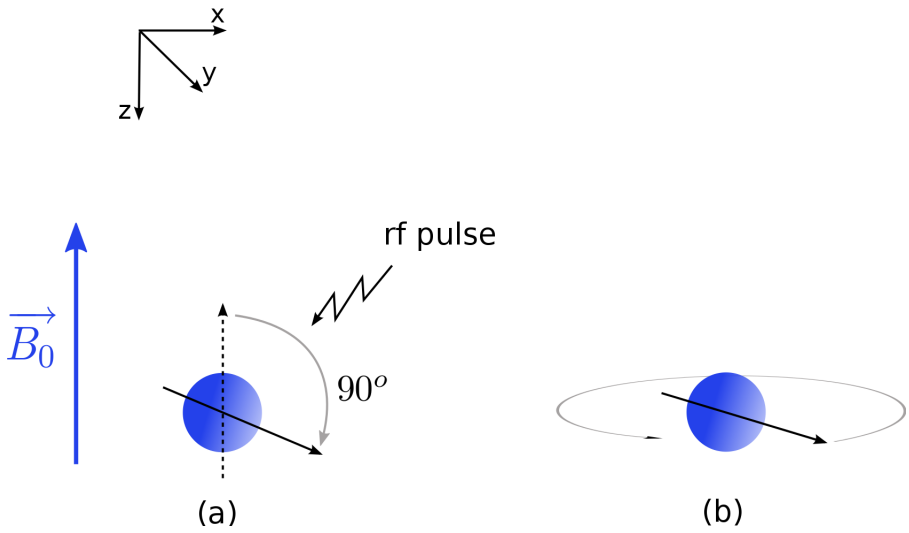


Figure 3.4: (a) The spin tilts in the plan perpendicular to $\mathbf{B_0}$. Then the spin continue to precess about $\mathbf{B_0}$ (b).

After the RF pulse is applied, the spins begin to return to their stable states, i.e. aligned to the main magnetic field $\mathbf{B_0}$. While the spins are returning to their stable states, they continue to precess about $\mathbf{B_0}$ and a magnetic field is produced. The magnetic field during

which the protons are returning to their stable states is characteristic of the biological tissue under study and, thereby, enables the distinction between different structures.

To summarize, an RF pulse is used to destabilize the spins. Afterwards, while returning to their stable states, they generate a magnetic field whose the intensity depends on a given tissue.

The next sections describe more precisely the relevant information useful in three kind of MR imaging: The T_2 and T_1 weighted MRI and the diffusion weighted MRI.

3.1.2 Image formation

In this section, we refer to the rotating magnetic field induced by the spins as the net magnetization vector **N**.

As it is said in the previous section, there are three phases in the acquisition process:

1. An alignment phase: A magnetic field $\mathbf{B_0}$ is applied to align **N** with this field.

2. An excitation phase: A RF pulse is applied to tilt **N** away from $\mathbf{B_0}$.

3. A relaxation phase: **N** returns to its initial state, i.e., aligned with $\mathbf{B_0}$.

During the relaxation phase, the receiver coil of the MRI system is prone to the net magnetization vector **N** and induces a current that will be interpreted to generate the final image. **N** consists in two components: a longitudinal component $\mathbf{N_L}$ and a transverse component $\mathbf{N_T}$. The figure 3.5 illustrates the decomposition of **N** in these two components at a moment during the relaxation phase.

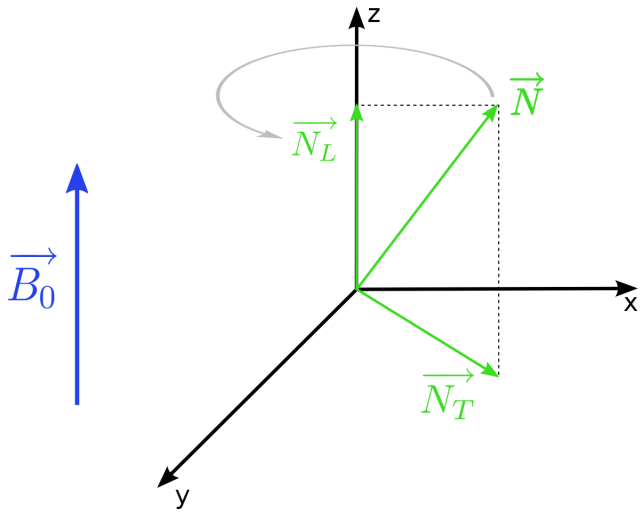


Figure 3.5: The net magnetization N and its two components

During the relaxation phase, the longitudinal component increases and the transverse component decreases. The T_2 and T_1 weighted MRI depend on the evolution of these two components. In particular,

- **The T_1-weighted MR imaging** follows the evolution of $\mathbf{N_L}$. It is based on the current induced in the coil from the restoration of the longitudinal component and is completely characterized by the constant time T_1. The constant time is the time for $|\mathbf{N_L}|$ to recover 63% of its original value.

- **The T_2-weighted MR imaging** follows the evolution of $\mathbf{N_T}$. The constant time T_2 is the time for $|\mathbf{N_T}|$ to loose 63% of its value just after the RF pulse is applied. The transverse component of the net magnetization vector depends on the phase coherence between spins. As the magnetic field $\mathbf{B_0}$ is not perfectly homogeneous, the spins loose their phase coherence making the transverse component decrease faster. An 180^o RF pulse is applied in order to eliminate spin dephasing. The process is referred to as a T_2^*-weighted MR imaging.

The intensity of the current induced in the receiver coil depends on the values of the constant time (T_1 or T_2). Since each tissue is characterized by a constant time, it is possible to distinguish different biological tissues.

Spatial selection The net magnetisation vector $\mathbf{N}$ precesses about the main magnetic field $\mathbf{B_0}$. The angular frequency of the precession (The Larmor frequency) depends of the $\mathbf{B_0}$ value. In order to tilt $\mathbf{N}$, a 90^o RF pulse is applied at the Larmor frequency of the spins. In other words, only the spins precessing at the pulse frequency are tilted. A way, to target a specific part among the entire scanned tissue, is to expose the target to a magnetic field that makes the spin to precess at the RF pulse frequency. As it is illustrated in figure 3.6, three other gradients $\mathbf{G_x}$, $\mathbf{G_y}$ and $\mathbf{G_z}$ are applied in order to enable spatial localisation in MRI:

- A first magnetic field gradient $\mathbf{G_z}$ is applied to make the spins of an entire slice match with the Larmor frequency of the RF pulse.

- Two other gradients $\mathbf{G_x}$ and $\mathbf{G_y}$ enable the excitation of the spins in the x-y plan of the slice selected by $\mathbf{G_z}$. As the application of these two gradients are not done simultaneously in the entire slice, it involves a phase difference θ_{xy} between every spin based on their locations. Finally, the signal acquired by the receiver coil is the integration of all spins. It is given by:

$$s(t) = \int f(x,y) \exp(2i\pi\theta_{xy}) dxdy \tag{3.1}$$

  $f(x,y)$ is the current value induced by the net magnetization at the x-y localization. Therefore, $s(t)$ is the Fourier transform of the selected slice image, referred to as a k-space.

The MR image is thus obtained by applying an inverse Fourier transform to the corresponding k-space. The figure 3.7 and 3.8 show a k-space and its inverse Fourier transform (the MR image).

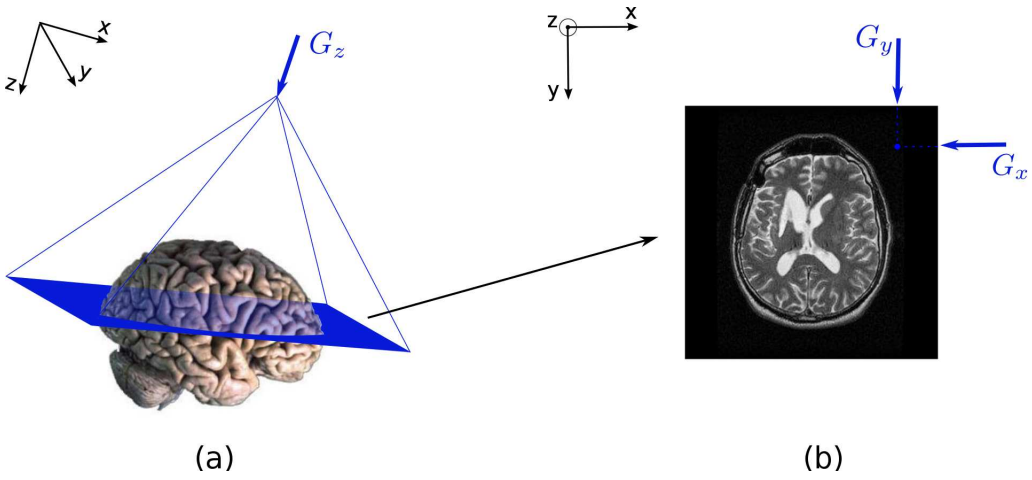


Figure 3.6: (a) slice selection (b) excitation of the spins in the x-y plan

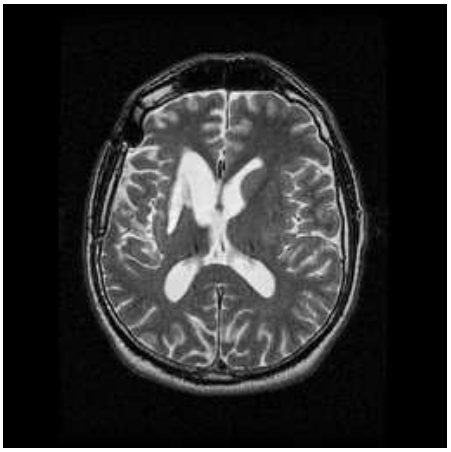


Figure 3.7: MRI image in its original space

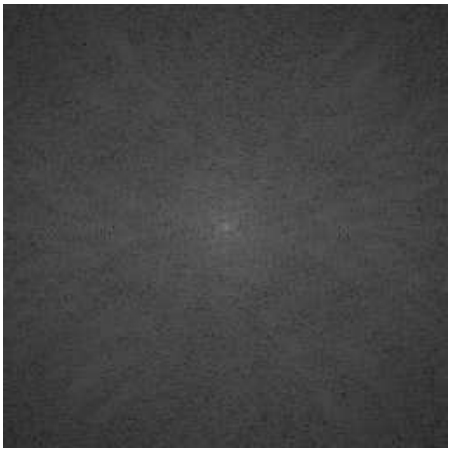


Figure 3.8: MRI image in the Fourier space

3.1.3 Adding diffusion weighted and the gradient echo pulse sequence

In the following, we call $\mathbf{g}$ the diffusion gradient which is the magnetic field applied by the MRI scanner given by,

$$\mathbf{g} = \mathbf{B_0} + \mathbf{G_x} + \mathbf{G_y} + \mathbf{G_z}. \tag{3.2}$$

$\mathbf{g}$ is often referred as a bipolar gradient, which consists in two gradients ($\mathbf{g}_+$ and $\mathbf{g}_-$) with the same magnitude but opposite directions.

T_2-weighted MR images are characterized by the phase coherence between spins. Water molecule diffusion is a factor that increase the incoherence between spins and reduce the MR signal intensity. In this section we explain how to modify the MR acquisition to become sensitive to diffusion.

After the excitation of the spins by a 90^o RF pulse, a bipolar gradient is applied in the direction of interest. The application of the first gradient $\mathbf{g}_+$ adds a positive phase to each spin depending on their positions along the gradient direction (see Fig. 3.9 (a),(b)). The second gradient $\mathbf{g}_-$ is applied in the opposite direction and add the opposite phase to each spin. After the application of $\mathbf{g}_-$, two cases are usually encountered:

- In the first case, there was no displacement between the application of $\mathbf{g}_+$ and $\mathbf{g}_-$ (Fig. 3.9 (c)). Hence, the spins return to their previous positions (fig 3.9 (e)) and are completely in phase.
- In the second case, the spins have been displaced due to the Brownian motion (fig 3.9 (d)) and are subjected to a different field strength during the application of $\mathbf{g}_-$. The spins do not return to their initial states and are not in phase anymore. As a result, the transversal component of the net magnetization field $\mathbf{N_T}$ decreases (fig 3.9 (f)).

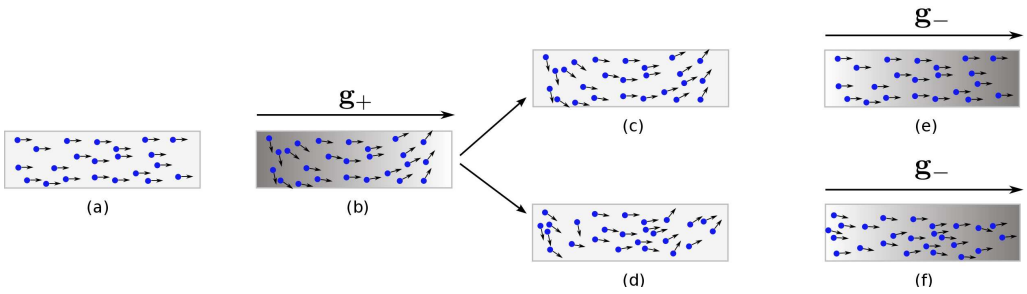


Figure 3.9: Molecules diffusion after the application of a magnetic field.

The diffusion is thus characterized by a loss of coherence in the spin phases resulting in a decrease of the transverse component of the net magnetization field.

The pulse gradient spin echo sequence Stejskal and Tanner [117] introduced in 1965 the pulse gradient spin echo (PGSE) sequence. This particular sequence apply short duration diffusion gradients ($\mathbf{G_x}$, $\mathbf{G_y}$ and $\mathbf{G_z}$), so that the diffusion that occurs during the

application of a gradient is negligible with respect to the diffusion that occurs between the application of these gradients. For this reason, the PGSE is often referred as the narrow pulse sequence. We describe the PGSE sequence in the schematic of Fig. 3.10

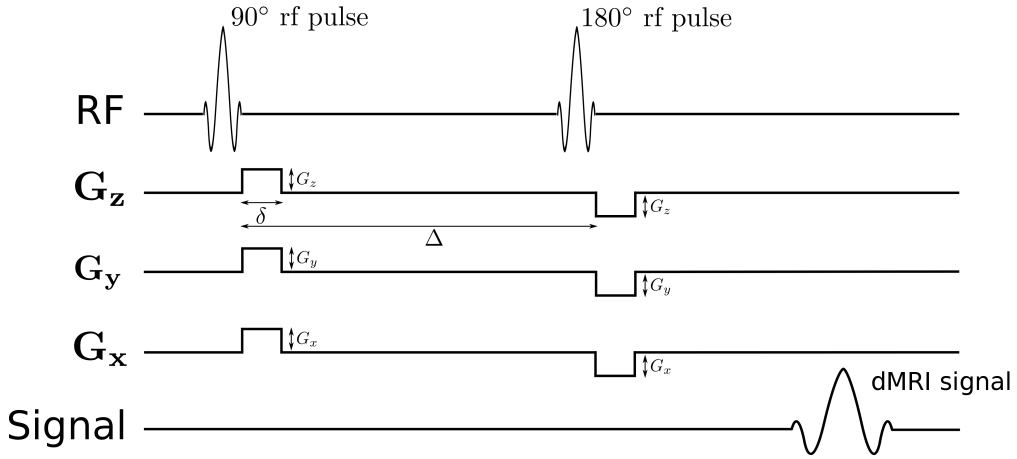


Figure 3.10: A schematic of the PGSE sequence. δ is the diffusion weighted gradient duration and $\Delta \gg \delta$ is the time between two gradient pulses. The intensity of the three elementary diffusion gradients, respectively G_x, G_y and G_z, define the diffusion gradient direction applied by the MRI system.

The PGSE sequence is also characterized by the echo time T_E and the repetition time T_R, which respectively define the time between the first RF pulse and the maximum of the received dMRI signal , and the delay between two similar RF pulses. Then the MR (f in 3.1) signal from the PGSE sequence is as follows,

$$f(x, y) = M_0(1 - \exp(-T_R/T_1)\exp(-T_E/T_2)\exp(-bD)) \tag{3.3}$$

i.e. f is attenuated exponentially by the product of the diffusion coefficient D and a factor b, which is a function of the diffusion weighted gradients. For the PGSE sequence, b is defined as

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right), \tag{3.4}$$

where G is the diffusion weighted gradient amplitude, i.e., $G = |\mathbf{g}|$ and M_0 is the initial spin density.

3.2 Modeling the diffusion MRI signal and its features

As it has been described in the previous section, diffusion MRI quantifies the water molecule diffusion at the scale of a voxel thanks to the local phase variation of the spin that occurs during diffusion.

Modeling the water diffusion in a free medium is easily handled since the water molecules have the same probability to displace a certain distance, which results in a scalar quantification of the whole phenomenon. However, what happens if the diffusion is constrained in certain directions? Modeling it becomes difficult since the water molecule diffusion is different from one direction to another. Scalar quantification of diffusion at every direction is still possible but is not adapted for analysis. Then, recent works have been focused on the development of more complex models that enable a better interpretation of the phenomenon.

The first part of this section describes the free and constrained diffusion together with a scalar analysis of the phenomenon. In particular, we introduce the Apparent Diffusion Coefficient (ADC). The second part introduces the Ensemble Average Propagator (EAP), which represents the full 3D displacement of water molecule diffusion. Then, we present a couple of model that leads to interesting features related to the EAP, namely the Diffusion Tensor (DT) and the Orientation Distribution Function (ODF).

3.2.1 Principles of diffusion and scalar quantification of the phenomenon

Let us consider a water molecule with initial position $\mathbf{R_0} \in \mathbb{R}^3$. After a time τ, it moves to a position $\mathbf{R}_\tau \in \mathbb{R}^3$. We define the displacement vector $\mathbf{R} \in \mathbb{R}^3$ as

$$\mathbf{R} = \mathbf{R_0} - \mathbf{R}_\tau. \tag{3.5}$$

In this section, we introduce the mathematical formulation used to characterize the distribution of water molecules displacement $\mathbf{R}$ starting by the simplest scenario, namely the free diffusion, to more complex phenomena.

Free diffusion In a free environment, water molecules follow a random trajectory, referred to as a random walk. A random walk is a mathematical model consisting in a sequence of random steps which are completely uncorrelated from one another. We show in Fig. 3.11 a 2-dimensional random walk corresponding to one water molecule. Because the molecule undergoes a random trajectory, it is impossible to predict its final location. In addition, there is an uncountable number of water molecules which makes impossible to follow each molecule independently.

Hence, instead of considering only one molecule, Einstein proposed in 1905 [51] to observe the mean square displacement $\langle \mathbf{R}\mathbf{R}^T \rangle$ of the whole set of molecules. He found that $\langle \mathbf{R}\mathbf{R}^T \rangle$ is proportional to the observation time τ, i.e.,

$$\langle \mathbf{R}\mathbf{R}^T \rangle = 6D\tau, \tag{3.6}$$

where D is often called the diffusion constant. It means that each molecule has the same probability of displacing a given distance. In a free medium, (3.6) describes an isotropic diffusion where the water molecule distribution of squared displacement is characterized by a Gaussian.

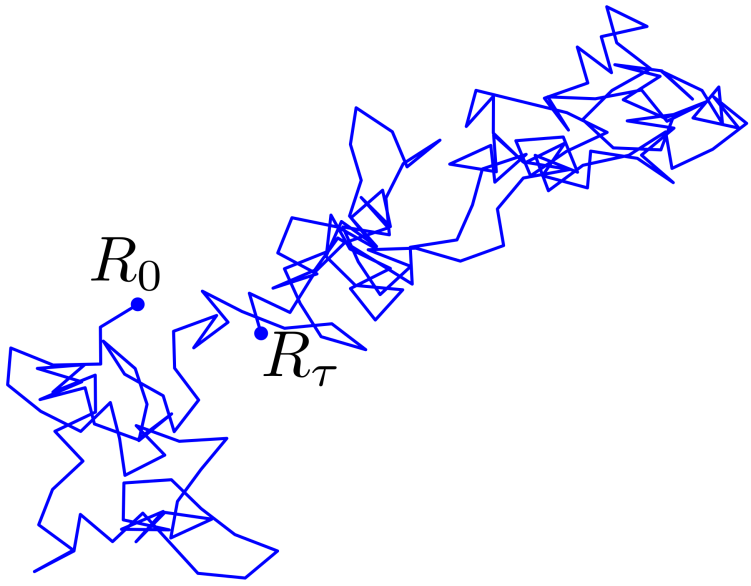


Figure 3.11: 2-dimensional random walk of a water molecule. $\mathbf{R_0}$ is its initial position and $\mathbf{R}_\tau$ its position after a time τ.

Constrained diffusion In biological tissues, such as the white matter, the water molecules are constrained by their surrounding environment. In this case, they are confined within such hindrance meaning that the diffusion has directionality (see Fig. 3.12 for a 2-dimensional example). In addition, biological barriers have pores which make them permeable. Then, in a direction where the diffusion is constrained the diffusion coefficient D is lower than the value of D in a free diffusion scenario. In this case, D is referred as the Apparent Diffusion Coefficient (ADC) [74], denoted by D_{ADC} in the following.

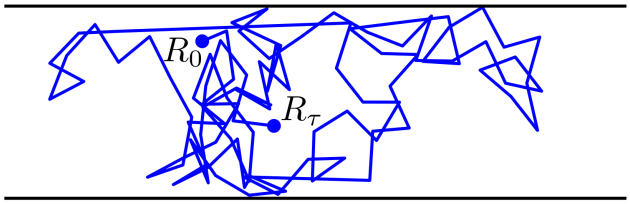


Figure 3.12: 2-dimensional random displacement of a water molecule constrained in a rectangle. $\mathbf{R_0}$ is its initial position and $\mathbf{R}_\tau$ its position after a time τ.

The dMRI signal acquired by the MRI system with the PGSE sequence presented in Sec. 3.1.3 directly depends on the mean squared displacement $\langle \mathbf{R}\mathbf{R}^T \rangle$ of the water molecule in the applied diffusion gradient direction. The higher the $\langle \mathbf{R}\mathbf{R}^T \rangle$, the lower is the dMRI signal in the considered direction. Thus, at a given diffusion time τ, we can infer the value of the ADC in a particular direction, from the acquired signal.

Strictly speaking, the dMRI signal S for a diffusion weighted gradient $\mathbf{g}$ is attenuated exponentially by the product of the ADC value D_{ADC} and a factor b which is a function of the diffusion-weighted gradient described in Fig. 3.10, i.e.,

$$S(b) = S_0 \exp(-bD_{\text{ADC}}), \tag{3.7}$$

where S_0 is the signal in the absence of any gradients and is given by the constant part of (3.3), i.e $S_0 = M_0(1 - \exp(-T_R/T_1)\exp(-T_E/T_2)$.

One can estimate D_{ADC} by rearranging (3.7) and, then, compute a quantitative map of the apparent diffusion coefficient [74], thus giving a scalar approximation of the diffusion in a given direction. Fig. 3.13 illustrates an ADC map at a given direction corresponding to an axial view of a brain.

3.2.2 Characterizing the 3D probability distribution of water molecule diffusion

The ADC is a scalar quantification of the dMRI signal in a given diffusion direction and is only true for Gaussian diffusion. This hypothesis excludes many phenomena such as

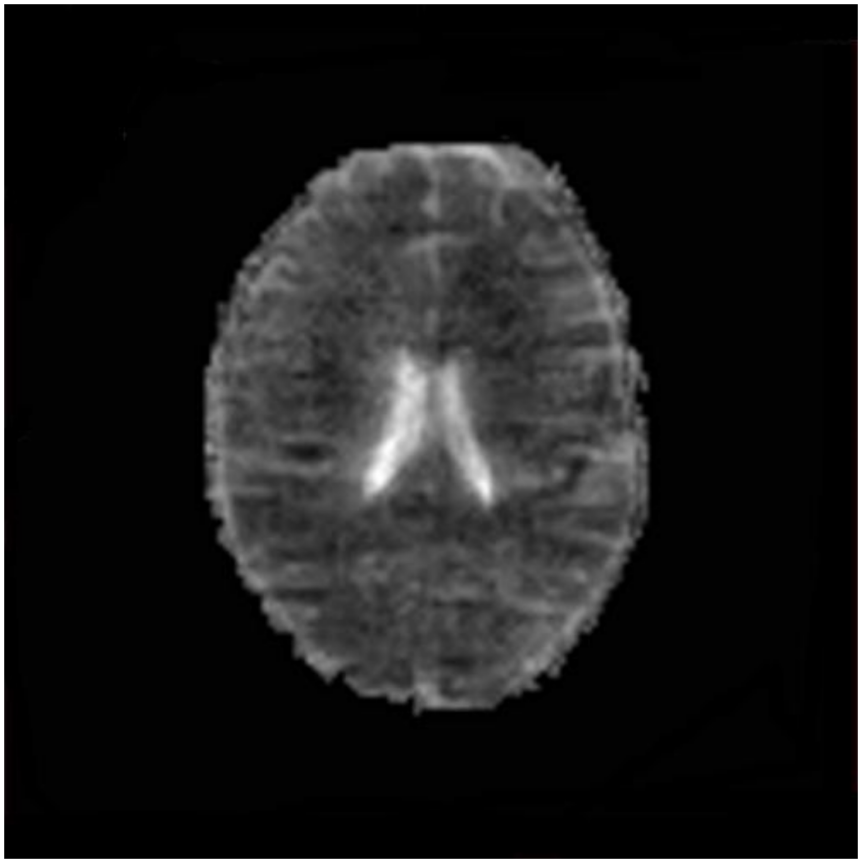


Figure 3.13: ADC map at a given direction corresponding to an axial view of a brain.

permeability and heterogeneity. In addition, this scalar analysis is not sufficient to interpret the 3D diffusion phenomenon in its entirety. One may want to compute the probability distribution of the water molecule diffusion in a voxel in order to characterize the full 3D diffusion profile.

In this section, we present the general formalism to characterize the so-called Ensemble Average Propagator (EAP), which is a full description of the water molecule diffusion. We also present some techniques to approximate the EAP. Then, we introduce other models of diffusion, which simplify the EAP formalism. In particular, we describe the Diffusion Tensor (DT), which is the most common simplification of the EAP. We also present the High Angular Resolution Imaging models, which tend to a more accurate approximation of the EAP than the DT.

The Ensemble Average Propagator (EAP)

When the tissue's microstructural features are not known *a priori*, a mathematical model that does not take into account any hypothesis is appropriate. For this reason it would be desirable to obtain the quantity $P(\mathbf{R}_\tau|\mathbf{R_0},\tau)$, which gives the probability that a water molecule displaces from $\mathbf{R_0}$ to $\mathbf{R}_\tau$ after a time τ.

However, the voxel resolution in dMRI does not permit the inferring of $P(\mathbf{R}_\tau|\mathbf{R_0},\tau)$ for only one molecule. Instead, one may want to compute this quantity averaged on all the molecules contained in a voxel, often referred as the Ensemble Average Propagator (EAP). We define the EAP as

$$P(\mathbf{R},\tau) = \int_{\mathbf{R_0}\in\Omega_\nu} P(\mathbf{R}_\tau|\mathbf{R_0},\tau)\rho(\mathbf{R_0})\mathrm{d}\mathbf{R_0}, \tag{3.8}$$

where $\rho(\mathbf{R_0})$ is the initial density of molecules in a voxel such that $\int_{\mathbf{R_0}\in\Omega_\nu}\rho(\mathbf{R_0})\mathrm{d}\mathbf{R_0}=1$ and Ω_ν is the domain where the voxel is defined. In the following, because the obervation time τ is constant for every voxel, we denote the EAP as $P(\mathbf{R})$.

Under the narrow pulse condition (see Sec. 3.1.3), i.e., when $\delta \ll \Delta$, the normalized dMRI signal $E(\mathbf{q}) = \frac{S(\mathbf{q})}{S_0}$ [18, 39, 117] is written as the Fourier transform of the Ensemble Average Propagator (EAP) $P(\mathbf{R})$,

$$E(\mathbf{q}) = \int_{\mathbf{R}\in\mathbb{R}^3} P(\mathbf{R})\exp(-2\pi i\mathbf{q}\cdot\mathbf{R})d\mathbf{R}, \tag{3.9}$$

where $\mathbf{q}$ represents the effective gradient direction given by $\mathbf{q}=\gamma\delta\mathbf{g}$. The b-value introduced in 3.4 is related to the effective gradient direction as $b=(\Delta-\delta/3)\|\mathbf{q}\|^2$. Note that in the following, $\mathbf{q}$ and $\mathbf{R}$ may be decomposed as $\mathbf{q}=q\mathbf{u}$ and $\mathbf{R}=R\mathbf{r}$, where $\mathbf{u}$ and $\mathbf{r}$ are 3D unit vectors.

Then, the EAP is the inverse Fourier transform of the dMRI signal,

$$P(\mathbf{R}) = \int_{\mathbf{q}\in\mathbb{R}^3} E(\mathbf{q})\exp(2\pi i\mathbf{q}\cdot\mathbf{R})\mathrm{d}\mathbf{q}. \tag{3.10}$$

Estimating the Ensemble Average Propagator

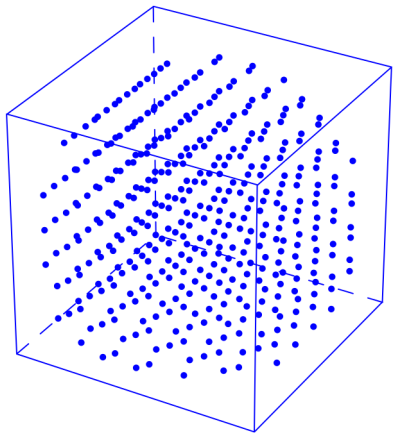


Figure 3.14: q-space sampling in a $8 \times 8 \times 8$ Cartesian grid where only the samples within the sphere centered in the grid are displayed.

- In **Diffusion Spectrum Imaging (DSI)** [131, 130, 127] the EAP $P(\mathbf{R})$ is directly obtained by taking the inverse Fourier transform of the normalized signal $E(\mathbf{q})$ measured in the q-space on a cartesian grid.

  In [130] the size of this grid is $11 \times 11 \times 11$. They do not consider every q-space sample on this grid and take only the samples comprised within the sphere of five lattice unit radius, which correspond to 514 measurements. Considering the measurement at $b = 0$, DSI as described in [130] requires 515 measurements. The cartesian grid is not fixed and can take any dimension. We show in Fig. 3.14 a q-space sampling in a $8 \times 8 \times 8$ Cartesian grid where we keep only the samples within the sphere centered in the grid.

  The inverse Fourier transform is performed via a inverse Fast Fourier Transform (iFFT). In practice, we apply a 3-dimensional Hamming window on $E(\mathbf{q})$ before applying the iFFT. The Hamming window in 1 dimension is defined as,

$$w(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right)\right) \tag{3.11}$$

  where N represents the width of the grid. The 3-dimensional Hamming window is obtained by applying this 1-dimensional window in every dimension of the cartesian grid.

  The resulting EAP is evaluated on a discrete dual grid, and can be interpolated to evaluate characteristics such as the orientation distribution function (ODF), the mean square displacement (MSD) or the return to origin probability (RTO). This method was applied successfully in brain imaging, for the reconstruction of complex configurations of fibers, including fiber crossing [130].

  DSI is known to be a greedy technique and researchers have recently started to work in a way to accelerate it [91, 82, 70]. We have also addressed this problem and proposed

in chapter 5 a method based on Compressive Sensing in order to accelerate the DSI acquisition (A complete description of Compressive Sensing is given in **chapter 4**).

- **Analytical and continuous modeling (ACM)** of the EAP with multiple shell acquisition. It consists in modeling the EAP with a linear combination of 3D continuous functions from measurements acquired in a discrete range of b-values. We can refer the ACM techniques as multiple shells HARDI methods. All the ACM techniques start by fitting the dMRI signals and then use an analytical Fourier relation to estimate the EAP. A large panel of bases (or dictionaries) were introduced for this purpose. We can cite the Spherical Polar Fourier (SPF) [5, 34], the SPF dual basis [89], the SHORE [99, 35] and the Solid Harmonic (SoH) [45] bases. These four bases are described in chapter 6. We also present in **chapter 6** an acceleration of the ACM techniques using the Compressive Sensing framework. More recently the Bessel Fourier Orientation Reconstruction (BFOR) [62] was proposed, whose solution is based on the estimation heat equation of the diffusion signal for each shell acquisition and a generalization of the spherical ridgelets to a radial fitting [109].

- **Generalized Diffusion Tensors Imaging (GDTI)** [76] uses high-order tensor (HOT) to estimate the radial and angular part of the 3D dMRI signal. Then the authors propose an analytical formula to compute the EAP.

The Diffusion Tensor (DT)

Estimating the 3D EAP with the methods described above requires a large amount of data. In addition, these methods are not often computationally efficient. Then, instead of trying to compute a complete EAP, one may want to approximate it by a simpler model which requires less measurements.

A common assumption is to consider that the water molecule diffusion has a Gaussian behavior. Under this hypothesis, one model to characterize the diffusion is the diffusion tensor (DT) [117, 8, 10]. The DT is a second order tensor, which can be represented by a 3×3 symmetric and positive definite matrix $\mathbf{D}$, i.e., $\mathbf{R}^T\mathbf{D}\mathbf{R} > 0$ whose components characterize displacements in 3D, i.e.,

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{bmatrix}. \tag{3.12}$$

The three diagonal elements represent the diffusivity along the axes $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ and the off diagonal elements represent the correlation between the displacement along these axes. The DT is well represented by an ellipsoid with its main axis corresponding to the direction where the diffusion is the largest. The surface of the ellipsoid represents the distance where a molecule will diffuse after a given observation time τ. In particular, when the diffusion is isotropic, i.e., only the diagonal elements of $\mathbf{D}$ are non zero and are equal, then this surface becomes a sphere. We show in Fig. 3.15 some examples of diffusion tensors and their corresponding ellipsoids.

The DT can be viewed as the covariance matrix of the three dimensional Gaussian distribution of the water molecule diffusion. In this framework, we can characterize the

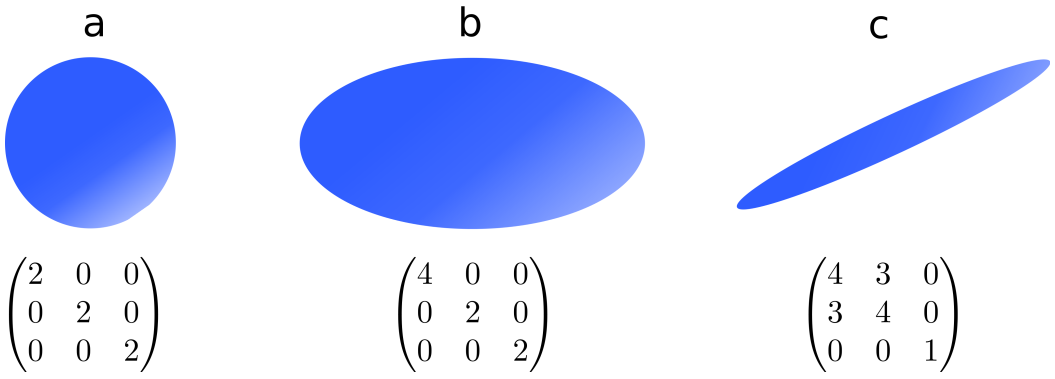


Figure 3.15: Diffusion tensor (DT) matrices with the corresponding ellipsoids. (a) is a isotropic DT. (b) and (c) are anisotropic DTs.

EAP as

$$P(\mathbf{R}) = \frac{1}{\sqrt{(4\pi\tau)^3|\mathbf{D}|}} \exp\left(\frac{\mathbf{R}^T\mathbf{D}^{-1}\mathbf{R}}{4\tau}\right), \tag{3.13}$$

We show in Fig. 3.16 (a) a tensorial field. Because of the ellipsoidal nature of the diffusion tensor, we see on this figure that the DT can not resolve crossing fibers.

Estimating the Diffusion Tensor Basser [8, 10] introduced Diffusion Tensor Imaging (DTI), which describes a framework to estimate the diffusion tensor $\mathbf{D}$. It defines the three dimensional Gaussian distribution of the water molecule diffusion. In [117, 8, 10], the authors propose the following relation between the diffusion tensor and the normalized diffusion signal $E(\mathbf{q})$

$$E(q) = \frac{S(q)}{S(0)} = \exp\left(-4\pi^2\tau\mathbf{q}^T\mathbf{D}\mathbf{q}\right) \tag{3.14}$$

Since $\mathbf{D}$ is symmetric, the DT is completely defined by 6 components. D can be estimated by means of a least square fitting [8, 10]. Other approaches have been developed since the first work of [8, 10]. These include the weighted linear least squares approach [9] that takes into account the effect of the log transform of the signal, the tensor estimation with a positive constraint [32, 125, 31, 98], or riemanian based estimations [75, 4, 54].

Scalar measures derived from the Diffusion Tensor The ellipsoid is completely defined by the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of the corresponding tensor matrix (see Fig. 3.17 for an illustration).

From this system of eigenvectors and eigenvalues, one can derive significant diffusion parameters. These measures are commonly used in a clinical setting to localize white matter lesions that do not show up on other forms of clinical MRI. Here is a list of important parameters, which can be derived from the DT:

- The simplest DT quantity that one can compute is the trace of $\mathbf{D}$, which comes to

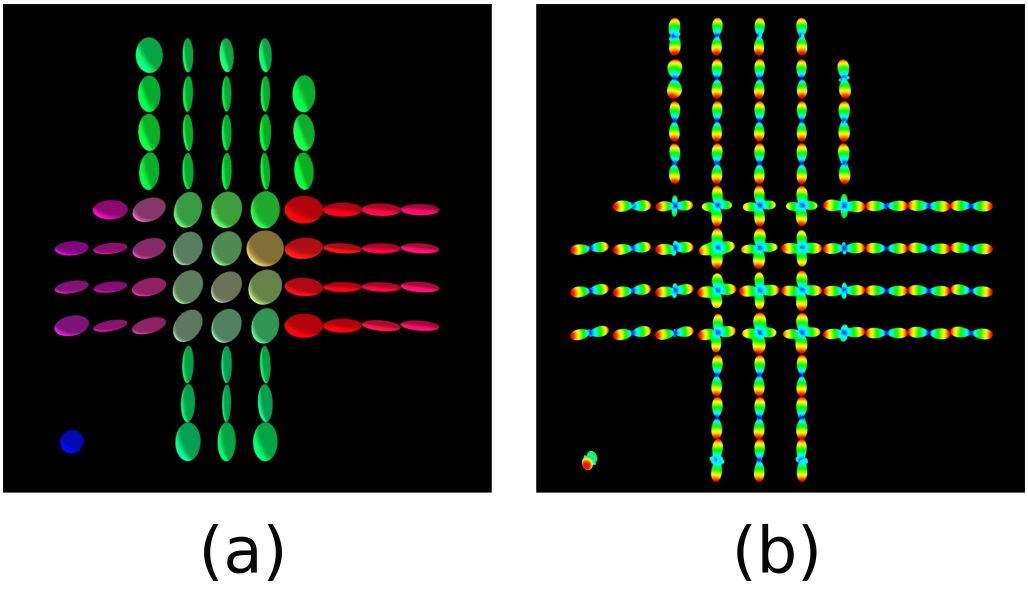


Figure 3.16: (a) is a tensorial field representing two crossing fibers. (b) is a field of ODF representing two crossing fibers. The underlying dMRI signal comes from the HARDI challenge at ISBI 2013 conference.

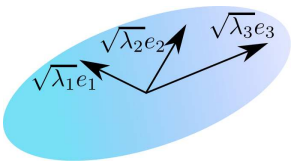


Figure 3.17: Ellipsoidal representation of a diffusion tensor. The diagonalisation of the diffusion tensor matrix is written as $D = V^T D_\triangle V$ where $D_\triangle$ is a diagonal matrix whose the diagonal elements are the eigenvalues $\sqrt{\lambda_i}$ and V consists in the eigenvectors $\mathbf{e}_i$.

sum its eigenvalues, i.e.,

$$trace(\mathbf{D}) = D_{xx} + D_{yy} + D_{zz} = \lambda_1 + \lambda_2 + \lambda_3. \tag{3.15}$$

$trace(\mathbf{D})$ can also be computed by taking the average of ADCs measured in three orthogonal directions [101]. This quantity is independent of fiber orientation and has been shown to be very helpful in detecting neurological disorders such as the acute ischemic lesions [78].

- An important quantity, which account for the diffusion anisotropy is the fractional anisotropy (FA) [104, 11]. The FA is the normalized variance of the eigenvalues express as,

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \langle\lambda\rangle)^2 + (\lambda_2 - \langle\lambda\rangle)^2 + (\lambda_3 - \langle\lambda\rangle)^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{3.16}$$

where $\langle\lambda\rangle$ is the mean of the three eigenvalues. The FA takes values between 0 and 1. Larger the FA, more constrained the diffusion along an axis. Then, the FA gives very valuable information regarding the diffusion anisotropy.

- The trace and the FA respectively characterize the mean diffusivity and the diffusion anisotropy, but do not indicate the shape of the ellipsoid underlying the DT. Different measures considering the tensor shape were introduced. One of them is the skewness, S_{kew}, of the three eigenvalues [7], i.e.,

$$S_{kew} = \frac{(\lambda_1 - \langle\lambda\rangle)^3 + (\lambda_2 - \langle\lambda\rangle)^3 + (\lambda_3 - \langle\lambda\rangle)^3}{3} \tag{3.17}$$

The skewness is positive for prolate tensors and negative for oblate tensors. One has to be careful when computing S_{kew} since this measure is highly sensitive to noise.

Alternatively the authors of [132] propose three indices. Each of them quantifies a particular tensor shape : the sphericity C_s, the linearity C_l and the planarity C_p defined as

$$C_s = \frac{\lambda_3}{\langle\lambda\rangle} \tag{3.18}$$

$$C_l = \frac{\lambda_1 - \lambda_2}{3\langle\lambda\rangle} \tag{3.19}$$

$$C_p = \frac{2(\lambda_2 - \lambda_3)}{3\langle\lambda\rangle} \tag{3.20}$$

The Orientation Distribution Function (ODF)

The Gaussian assumption of the DT model is an over-simplification of the diffusion of water molecules and, thus, has some limitations for voxels in which there are more complicated

internal structures. In particular the DT can not resolve crossing fibers. Therefore, we have to look at a more complex model to characterize the diffusion. Another model is the Orientation Distribution Function (ODF), which is the probability density function that a water molecule diffuses in a given direction. On can express the ODF $\Upsilon(\mathbf{r})$ as the integration of the EAP over a solid angle [130], i.e.,

$$\Upsilon(\mathbf{r}) = \int_0^\infty P(R.\mathbf{r})R^2 dR. \tag{3.21}$$

Some groups [126, 44] have also considered the ODF as the integration of the EAP over its radius. However this definition of the ODF does not lead to a true probability density function.

The ODF is an efficient model to describe the diffusion orientation in a voxel, and is often associated to the tractography of the white matter bundles of fibers in order to infer the fiber orientation. The Fig. 3.18 illustrates two examples of ODF.



Figure 3.18: ODF. On the left one fiber. On the right two crossing fibers

Fig. 3.16 (b) shows a field of ODF estimated with the data used for the DT estimation in Fig. 3.16 (a). We observe that the ODF describes the crossing region well whereas the DT is not able to resolve it.

Estimating the Orientation Distribution Function

- The ODF can be computed by integrating over a solid angle the EAP in a discrete fashion as it is illustrated on Fig. 3.19. This is always the case, when the EAP is inferred from the DSI technique where the measurements are acquired on a Cartesian grid. Practical details regarding the ODF computation from a discrete EAP are given in **chapter 5**.

- Q-ball imaging ([126, 130, 44, 1]) is part of the Single Shell HARDI methods. The idea in QBI is to sample the q-space on a single sphere at a given b-value in several directions instead of attempting to get the entire q-space. It means every diffusion gradient has the same magnitude. This method has proven useful to significantly reduce the number of measurements and the time for acquisition.

  At first the Q-ball imaging was implemented via numerical method [126]. However, the numerical approach was computationally intensive. Recent papers [44, 1, 96] have proposed an analytical solution of the QBI problem, where the ODF is modeled as a linear combination of spherical functions such that the modified spherical harmonic

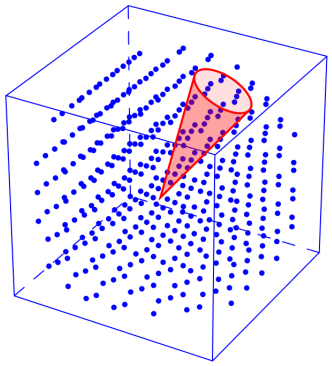


Figure 3.19: Integration of a discrete EAP over a solid angle. The red cone represents the space where the measurements are integrated.

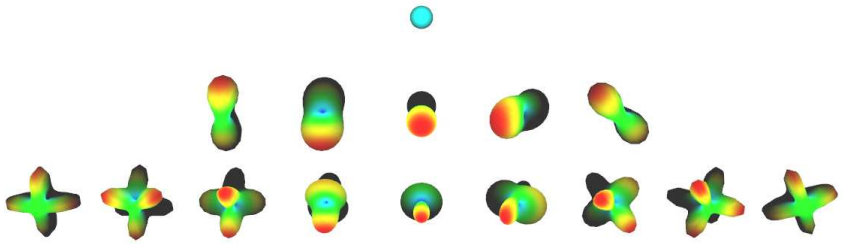


Figure 3.20: The spherical harmonic basis functions. Each line corresponds to orders $\ell = 0, 2, 4$. The parameter m ranges from $-\ell$ to ℓ.

(SH) basis, adapted to real, symmetric functions [44]. The SH functions are depicted on Fig. 3.20, for a truncation degree $L = 4$.

The mathematical expression of the real and symmetric SH functions is

$$Y_{\ell,m}(\theta,\phi) = \begin{cases} \sqrt{\dfrac{2(2\ell+1)(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta)\cos(m\phi) & \text{for } m < 0 \\ \sqrt{(2\ell+1)} P_\ell^m(\cos\theta) & \text{for } m = 0 \\ \sqrt{\dfrac{2(2\ell+1)(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta)\sin(m\phi) & \text{for } m > 0 \end{cases}$$

where P_ℓ^m is an associated Legendre Polynomial.

When truncated up to order L, this SH basis has dimension $J = (L+1)\cdot(L+2)/2$. To

ease matrix representation, we introduce a single index j to designate the spherical harmonic function $Y_j = Y_{\ell,m}$ [44]. The correspondence is given by $j \equiv j(l,m) = (\ell^2 + \ell + 2)/2 + m$.

The dMRI signal $E(q\mathbf{u})$ acquired at a unique q-value q_0 can be described as,

$$E(q_0\mathbf{u}) = \sum_{j=1}^{J} c_j Y_j(\mathbf{u}), \tag{3.22}$$

where $\mathbf{u}$ is a unit vector defining the spherical coordinate (θ, ϕ), $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$, with

$$\mathbf{u} = \begin{bmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{bmatrix}. \tag{3.23}$$

Then, from 3.22 one can express the ODF as a linear combination of SH,

$$\Upsilon(\mathbf{r}) = \sum_{j=1}^{J} c_j^{'} Y_j(\mathbf{u}), \tag{3.24}$$

where $c_j^{'}$ are the SH coefficients.

When the ODF is defined as the radial integration of the EAP [126, 44], Descoteaux et al. [44] expressed the SH coefficients as

$$c_j^{'} = 2\pi P_l(0) c_j \tag{3.25}$$

where P_l is the Legendre polynomial.

Moreover, when the ODF is defined as the integration of the EAP over a solid angle (see (3.21)), Aganj et al. [1] showed that it can be expressed as

$$c_j^{'} = \begin{cases} \dfrac{1}{2\sqrt{\pi}} & j = 1 \\ -\dfrac{1}{8\pi}(-1)^{\ell(j)/2} \dfrac{1 \times 3 \times \cdots \times (\ell(j)+1)}{2 \times 4 \times \cdots \times (\ell(j)-2)} c_j & j > 1 \end{cases}$$

Note that other spherical functions were used in the literature to estimate the ODF, namely the spherical Ridgelets [96] or the spherical wavelets [123]. They were developed to sparsely represent the angular profile of the diffusion signal and are used in sparse coding techniques.

- In [3, 120, 121], the authors propose to estimate the fiber orientation distribution called the fiber ODF (fODF). As in QBI, the estimation method requires samples taken at a unique b-value. Then, the signal is deconvoluted with a Gaussian kernel and Fourier-tranformed to get the fODF. The solution is analytical and enables to model the ODF using the Spherical Harmonics. The fODF is able to resolve up to 30 degrees crossings consistently [121], which makes this model a promising resource to estimate fiber orientation. A review of methods for reconstructing the fODF can be found in [67].

- Another way to estimate the ODF is based on the ACM techniques presented above. In this continuous framework, works propose analytical formulae to estimate the ODF [99, 34, 45, 95], which come for free after having modeled the 3D dMRI signal. We describe these formulae in **chapter 6**.

3.3 Conclusion

In this chapter, we presented different ways to characterize the diffusion of water molecules, starting from the simplest scenario where the diffusion is isotropic to cases where the diffusion is constrained by biological tissues.

We have seen that scalar quantification of the diffusion with the ADC coefficients is not convenient for analysis. Other models have been presented in order to describe the diffusion, namely the Diffusion Tensor (DT) and the Orientation Distribution Function (ODF). However, the full 3D characterization of the diffusion is done via the computation of the Ensemble Average Propagator (EAP), which is the probability distribution of water molecule diffusion at the scale of a voxel.

The EAP estimation often requires a large amount of data. This thesis is focused on how to decrease the number of measurements usually required to estimate the EAP. For this purpose, we use sparse coding techniques, and in particular the Compressive Sensing technique.

Chapter 4

Compressive Sensing and dMRI

Contents

4.1 Beyond the Shannon theory 54
4.1.1 The Shannon theorem 54
4.1.2 A surprising result 55
4.2 The discrete Compressive sensing theory 57
4.2.1 CS and Sparsity 58
4.2.2 CS and incoherence 60
4.2.3 CS recovery 61
4.3 Generalization to a broader class of representations 62
4.3.1 CS with overcomplete and redundant dictionary 62
4.3.2 Toward extending CS to recover continuous signals 63
4.4 A state of the art of CS in dMRI 65
4.4.1 CS to reconstruct the diffusion tensor 67
4.4.2 CS to reconstruct the ODF 67
4.4.3 CS to reconstruct the 3D diffusion signal and underlying features 68
4.5 Summary 70

Overview

Most modern signal processing systems are digital, hence the development of the sampling theory. Sampling consists in converting an analog signal into a digital or discrete signal. In order to ensure perfect reconstruction of the analog signal from its samples, the Shannon sampling theorem provides a sufficient condition for the sampling and recovery of a band-limited signal. However, the sampling rate increases with the band-width of the signal and the Shannon theorem does not consider infinite band signals. The CS theory was introduced with the desire to sample wideband signals at a sampling rate far lower from the Shannon rate. CS [23, 46, 52, 22, 49] has already been used in numerous applications [6, 77, 17, 48, 25, 61, 133, 13, 118, 134] and, in particular has been proved useful in recovering magnetic resonance (MR) images by significantly undersampling their k-spaces [77, 60, 55, 30, 124]. At first, the technique was designed to recover discrete and finite-length signals under the assumption that this signal admits a sparse representation with respect to an orthonormal basis [23, 46]. The standard theory was found limited in many applications where the signal is likely to admit a sparse representation with respect to an arbitrary dictionary. For this purpose, the authors in [22] generalize the classical theory to signals having sparse representations in overcomplete and redundant dictionaries. All these works [23, 46, 22] on CS have focused on reconstructing discrete and finite-length vectors, whereas most of the natural phenomena are analog. Then, [52, 49] extend the classical CS framework to allow for analog sampling and, thus, include the case of continuous time signals.

This part of the manuscript is organized as follows. Sec. 4.1 introduces the sampling theory of Shannon and its limitations. Sec. 4.2 describes the classical CS theory. The extension to the case of overcomplete and redundant dictionary, as well as continuous signals recovery are presented in Sec. 4.3. State of the art of CS in dMRI is presented in Sec. 4.4.

4.1 Beyond the Shannon theory

4.1.1 The Shannon theorem

Natural phenomena, such as diffusion in white matter bundles, are analog. In most of modern computational systems, we do not directly process these analog signals. Instead, we sample them, which means that we observe these analog signals at different instants. Such a sampled signal is called a digital or discrete signal.

Converting analog signals to digital signals is done via a sampling procedure. Sampling consists in representing an analog signal $x(t)$ by its values $x(kT_s)$ with $k \in \mathbb{Z}$, $T_s = \frac{1}{F_s}$ is the sampling period and F_s the sampling frequency, i.e.,

$$x_s(k) = x(kT_s) = x(t) \sum_{k=-\infty}^{+\infty} \delta(t - kT_s) \tag{4.1}$$

Recovering $x(t)$ from its sample values $x_s(k)$, requires the use of the Shannon sampling theory. The Shannon sampling theorem [116], also called the Nyquist sampling theorem, is a fundamental tool in digital signal processing, which allows one to recover a band limited signal from correctly sampled measurements.

Before defining the Shannon sampling theorem, let us see the effect of sampling in the Fourier domain. $X_s(f)$ the spectrum of $x_s(k)$ is defined as

$$X_s(f) = X(f) * F_s \sum_{k=-\infty}^{+\infty} \delta(f - kF_s) = F_s \sum_{k=-\infty}^{+\infty} X(f - kF_s), \tag{4.2}$$

which means that $X_s(f)$ is $X(f)$ periodized with a sampling period equal to F_s. If $x(t)$ is a band limited signal with f_{max} the highest frequency contained in its spectrum $X(f)$, then the choice of the sampling frequency F_s is limited by the Shannon sampling theorem:

Definition 1. *If a signal $x(t)$ is band limited, the minimum sampling rate is twice the highest frequency component of the signal.*

Specifically, $x(t)$ can be recovered by a sinus cardinal interpolation, which is equivalent to applying a low pass filtering in the frequency domain. We call this minimum sampling rate, the Shannon rate.

In the following we give an example of continuous signal reconstruction from equally spaced samples. Let us consider a sum of two sines at frequencies $f_1 < f_2$, i.e., $x(t) = 2\sin(2\pi f_1 t) + \sin(2\pi f_2 t)$ ((a) in Fig. 4.1). Ideally, when t is a continuous variable, the magnitude of its Fourier transform $|X(f)|$ is two pairs of dirac $\delta(f)$ at frequencies $[-f_2, -f_1, f_1, f_2]$, i.e., $|X(f)| = \delta(f - f_1) + \delta(f + f_1) + \frac{1}{2}\delta(f - f_2) + \frac{1}{2}\delta(f + f_2)$ ((b) in Fig. 4.1).

We sample $x(t)$ at two different frequencies F_s such that,

- $F_s > 2f_2$, i.e., above the Shannon rate.
- $F_s < 2f_2$, i.e., below the Shannon rate.

We get two digital signals $x(kT_s) = 2\sin(2\pi f_1 kT_s) + \sin(2\pi f_2 kT_s)$ with $T_s = \frac{1}{F_s}$ ((c) and (d) in Fig. 4.1). Afterwards, we compute the discrete Fourier transform of the two undersampled signals ((e) and (f) in Fig. 4.1). The consequence of signal undersampling, in the frequency domain, is the repetition of $|X(f)|$ at frequencies kF_s with $k \in \mathbb{Z}$. In the case where the Shannon recovery condition is respected ((e) in Fig. 4.1), we can easily isolate the original Fourier transform $|X(f)|$ by applying a low pass filter and, then, recover $x(t)$ with an inverse Fourier transform ((g) in Fig. 4.1). When the Shannon condition is not respected ((f) in Fig. 4.1), we observe that the repetition pattern $|X(f + kF_s)|$ cross the original Fourier signal $X(f)$. This phenomenon is called aliasing. Low pass filtering is not sufficient to recover the original Fourier signal since we are not able to distinguish $X(f)$ from the aliased pattern $|X(f + kF_s)|$.

4.1.2 A surprising result

We have just seen that the recovery of the original signal from equally spaced samples is hopeless when the sampling condition of the Shannon theorem is not respected. However, [77] observes that random sampling results in incoherent artefact instead of aliasing and equally spaced patterns. We propose to illustrate this phenomenon by randomly sampling the signal $x(t)$ of the precedent section at a frequency $F_s < f_2$, i.e., below the Shannon rate. In Fig. 4.2 we show $x(t)$ equally spaced sampled (a) and randomly spaced sampled (b) , as well as their respective Fourier transforms (c) and (d) . We observe that random

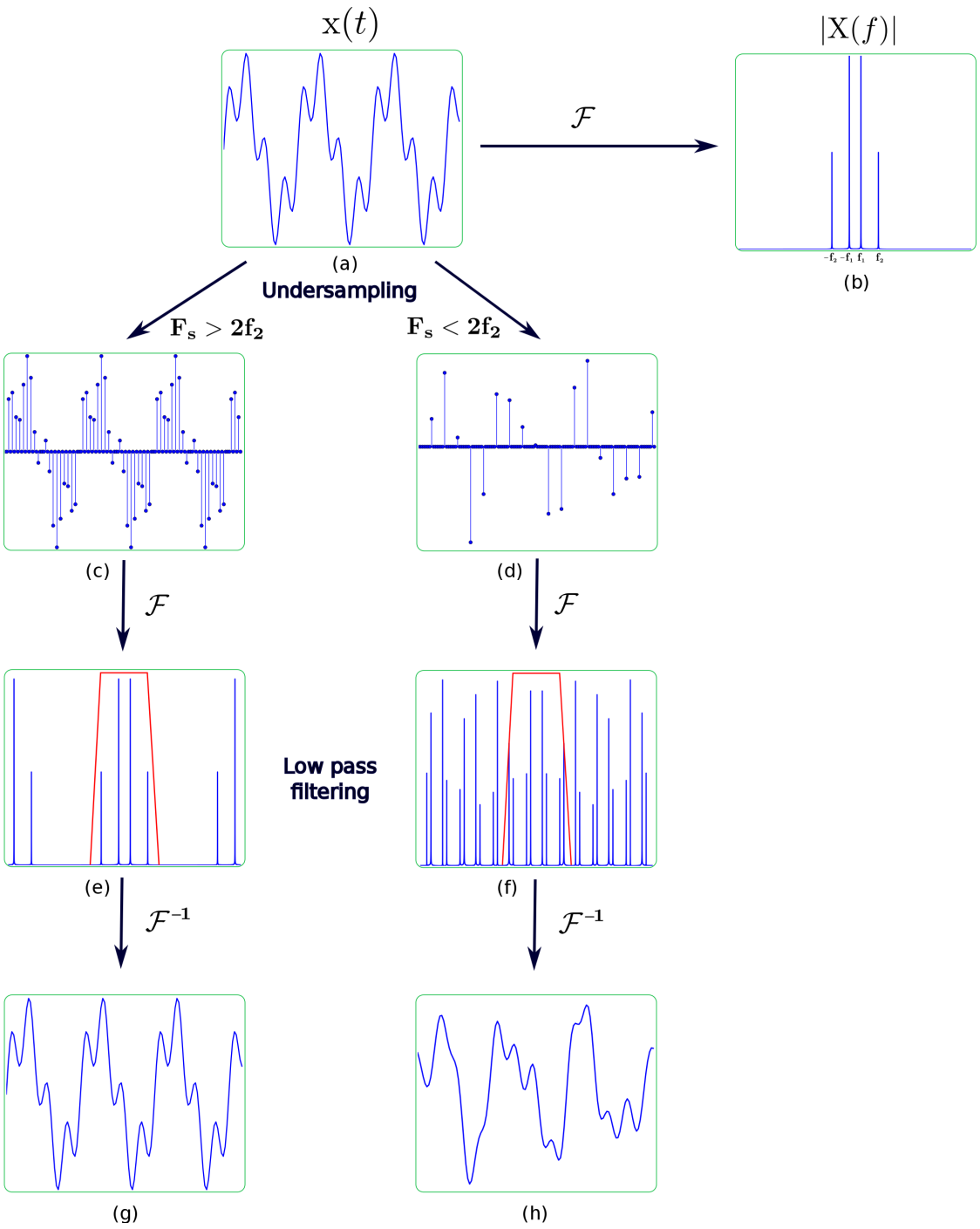


Figure 4.1: Illustration of the Shannon sampling theorem. (a) is a continuous signal $x(t) = 2\sin(2\pi f_1 t) + \sin(2\pi f_2 t)$ and (b) its Fourier transform $|X(f)|$. (c) and (d) are $x(t)$ undersampled at two different sampling rates: one above the Shannon sampling rate (c) and the other below the Shannon sampling rate (d). (e) and (f) are their respective Fourier transforms. A low pass filtering and inverse Fourier transform are applied resulting in (g) and (h).

sampling causes leakage of the aliasing components in the Fourier domain (d). Artefacts occur but, this time, it is spread all over the Fourier signal and looks like random noise instead of equispaced patterns. Then, it is possible to recover the original signal by simply retaining the Fourier values above a certain threshold. This straightforward example shows that one can sample a signal below the Shannon rate and hope to recover it.

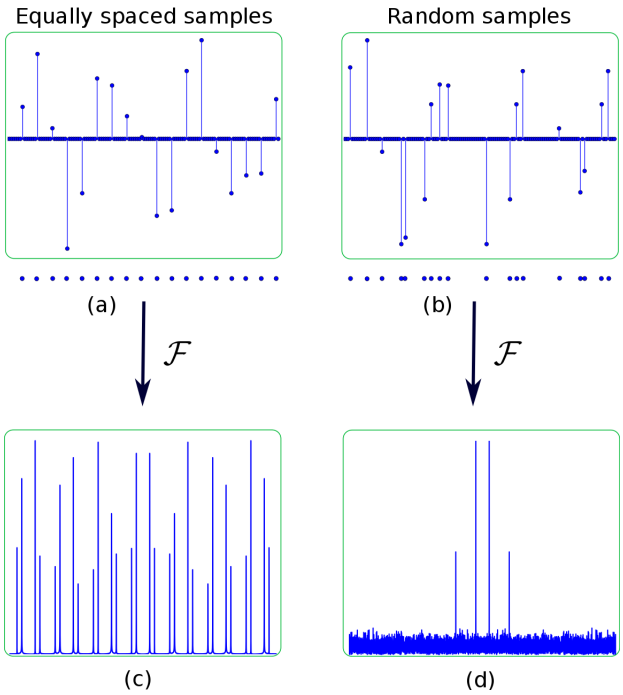


Figure 4.2: Two signals are undersampled at the same rate. (a) is equally spaced sampled and (b) is randomly spaced sampled. (c) and (d) are their respective Fourier transforms.

4.2 The discrete Compressive sensing theory

In this section, we restrict our attention to the classical finite-dimensional CS framework, i.e., we consider the reconstruction of a finite-length vector $\mathbf{x}$, which corresponds to a signal sampled at the Nyquist rate. We use this classical CS framework to accelerate the DSI acquisition in **chapter 5**.

Mathematically, we consider the problem of recovering a signal $\mathbf{x} \in \mathbb{R}^n$ from an observation vector $\mathbf{y} \in \mathbb{R}^m$ as

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{4.3}$$

$\mathbf{A} \in \mathbb{R}^{m \times n}$ is the, so called, sensing matrix, which represents the sensing system.

CS aims to infer the coefficient vector $\mathbf{x}$ from m measurements such that $m \ll n$. Accordingly, the problem (4.3) is ill-posed, i.e., there is an infinite number of different $\mathbf{x}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$.

CS asserts that one can recover $\mathbf{x}$ from fewer measurements than traditional methods. In order to make possible the signal recovering, CS relies on several properties:

- The **sparsity**. Sparsity expresses the idea that a signal can be approximated with just a few elements.
- The **incoherence**. It expresses the idea that the signal of interest must be spread out in the acquisition domain.
- An ℓ_1**-minimization** recovery combined with an sensing matrix respecting the RIP.

In this section, we present the theoretical aspect of the CS theory described in [46, 20, 21, 23, 47, 19].

4.2.1 CS and Sparsity

Definition of sparsity

Sparsity expresses the idea that a signal has a compact representation, i.e. one can exactly approximate it with just a few elements. If the signal $\mathbf{x}$ has exactly k nonzeros elements, i.e. $\|\mathbf{x}\|_0 = k$, we say $\mathbf{x}$ is k-sparse. Let $\mathbf{x}_k$ be the vector $\mathbf{x}$ with all but the largest k components set to zero. If $\mathbf{x}$ is strictly k-sparse, then

$$\|\mathbf{x} - \mathbf{x}_k\| = 0. \tag{4.4}$$

In practice, real-world signals are not strictly sparse. Instead, they may be approximately sparse or compressible in the sense that they can be represented accurately by $k \ll n$ coefficients, i.e.,

$$\|\mathbf{x} - \mathbf{x}_k\| = \epsilon. \tag{4.5}$$

ϵ quantify the compressibility of $\mathbf{x}$. In other words, one can neglect a few number of coefficients without much loss in terms of accuracy. In the restant of this manuscript, we will name sparse signals both those that are strictly and approximately sparse.

The utilization of sparsity in CS relies on the fact that the number of freedom in high dimensional signals is often small compare to their ambient dimensionality. If the signal is not sparse, it may be sparsified in an appropriate transform domain. For instance, a sinusoid is clearly not sparse, but its Fourier transform is extremely sparse (actually, the sum of two delta functions) (see Sec. 4.1). Then this signal, having originally an infinite bandwidth, can be in fact represented by only two elements, i.e, the sinuoid is exactly a 2-sparse signal in the Fourier domain. More generally, let consider an orthonormal basis $\mathbf{\Psi} = [\boldsymbol{\psi}_1 \boldsymbol{\psi}_2 \cdots \boldsymbol{\psi}_n]$ with $\boldsymbol{\psi}_i \in \mathbb{R}^n, \quad i \in [1, n]$ ($\mathbf{\Psi}$ is a $n \times n$ orthonormal matrix), we can express $\mathbf{x}$ as

$$\mathbf{x} = \mathbf{\Psi}\mathbf{c} = \sum_{i=1}^{n} c_i \boldsymbol{\psi}_i, \tag{4.6}$$

where $\mathbf{c}_i = \langle \mathbf{x}, \boldsymbol{\psi}_i \rangle$ are called the transform coefficients of $\mathbf{x}$ with respect to the orthonormal basis $\mathbf{\Psi}$. In the case where $\mathbf{x}$ is a linear combination of few basis functions $\boldsymbol{\psi}_i$, we still refer to $\mathbf{x}$ as being k-sparse, with the understanding that $\|\mathbf{c}\|_0 = k$. The classical finite-dimensional

CS framework considers that $\mathbf{\Psi}$ is a $n \times n$ orthonormal matrix, but we will see shortly that $\mathbf{\Psi}$ can be extended to a broader class of representation. In particular, $\mathbf{\Psi}$ can be represented by an over-complete and redundant dictionary (Sec. 4.3.1) or a more general orthonormal basis matrix (Sec. 4.3.2). In the rest of this section, $\mathbf{\Psi}$ will be characterized as a $n \times n$ orthonormal matrix.

Examples of application

Sparse approximation is at the heart of a wide range of signal processing applications (compression, restorations, feature extraction, source separations etc...) where one wants to reduce the dimensionality of a signal by approximating it as a linear combination of few elements from a known basis, which leads to a sparse representation of this signal.

There has been a surge of interest in recent years with the development of fast sparse transform exploiting the underlined sparse property of a signal. For instance, the JPEG image compression standard performs a two-dimensional discrete cosine transform (DCT) over each 8 by 8 pixels block of an image. In the mid 80's fast discrete wavelet transform (DWT) have been exploited in signal processing such as signal denoising for rejecting certain type of noise. Since the introduction of the DWT, the wavelets have been largely used to reduce the dimensionality of a wide range of signals. In particular, the DWT is performed in the image compression standard JPEG2000 since it enables a highly sparse representation of natural images. We show in Fig. 4.3 an image (a) together with its wavelet coefficients (b). As it is shown in the histogram of these coefficients (c), the wavelet representation of this natural image is sparse. The significant information is gathered in few coefficients.

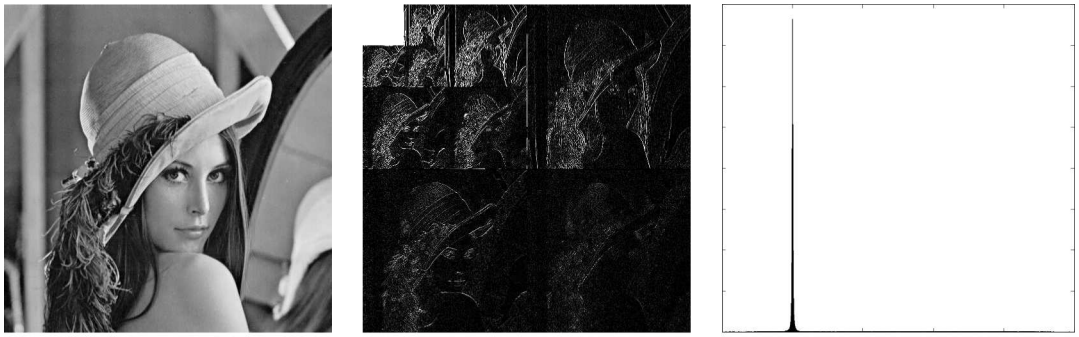


Figure 4.3: (a) Natural image. (b) Wavelet coefficients of the image. The largest values are in white color. (c) The histogram of the wavelet coefficients.

We estimate the original image by keeping 10% of the largest coefficients of its wavelet coefficients, the others being replaced by zero values. Fig. 4.4(b) shows that the estimated image is still a very good representation of the original one.

Sparsity is a key ingredient in CS recovery [46, 23]. Under certain conditions, CS gives strong theoretical results to reconstruct sparse signals using non-linear reconstruction methods.



Figure 4.4: (a) Natural image and its estimation keeping only 10% of its wavelet coefficients (b).

4.2.2 CS and incoherence

Definition of incoherence

Let us consider the orthonormal matrix $\bar{\mathbf{A}}$, which can be seen as a "non-undersampled" version of the sensing matrix $\mathbf{A}$ in (4.3), i.e., $\mathbf{A}$ is obtained by extracting m rows from $\bar{\mathbf{A}}$. [20] defines the coherence of $\bar{\mathbf{A}}$ as

$$\mu(\bar{\mathbf{A}}) = \sqrt{n} \cdot \max_{1 \leq k,j \leq n} |\bar{\mathbf{A}}_{k,j}|. \tag{4.7}$$

The lower $\mu(\bar{\mathbf{A}})$, the more incoherent $\bar{\mathbf{A}}$. The coherence parameter $\mu(\bar{\mathbf{A}})$ can take values between 1 and $\sqrt{n}$. The upper bound is easily derived using the fact that a component value of a column vector cannot exceed the norm of this vector (equal to $\sqrt{n}$ since $\bar{\mathbf{A}}$ is a orthonormal matrix). The lower bound is the special case where $\bar{\mathbf{A}}$ is flat, i.e., all column vector components have the same values $\bar{\mathbf{A}}_{1j} = ... = \bar{\mathbf{A}}_{nj}$ for $j \in [1, n]$. Hence, $\sqrt{\sum_{i=1}^{n} \bar{\mathbf{A}}_{ij}^2} = \sqrt{n\bar{\mathbf{A}}_{1j}^2} = \sqrt{n}$, which comes to $\bar{\mathbf{A}}_{1j} = ... = \bar{\mathbf{A}}_{nj} = 1$.

We can interpret $\mu(\bar{\mathbf{A}})$ as how the energy of $\bar{\mathbf{A}}$ is concentrated. The higher $\mu(\bar{\mathbf{A}})$, the more concentrated the energy of the columns of $\bar{\mathbf{A}}$. A high concentration results in constraining the significant information on few rows. On the contrary, if the columns energy is little concentrated, i.e., the columns values are nearly constant, then the information in the signal is spread out in the observation domain.

For a particular application, the sensing matrix $\bar{\mathbf{A}}$ can be decomposed as a product of an orthonormal sparsity basis $\mathbf{\Psi}$ and an orthogonal measurement system $\bar{\mathbf{\Phi}}$ ([20],[46]), i.e. $\bar{\mathbf{A}} = \bar{\mathbf{\Phi}}\mathbf{\Psi}$. In this case, we can write μ as

$$\mu(\bar{\mathbf{\Phi}}, \mathbf{\Psi}) = \sqrt{n} \cdot \max_{1 \leq k,j \leq n} \langle \Phi_k, \Psi_j \rangle . \tag{4.8}$$

where $\boldsymbol{\phi_k}$ and $\boldsymbol{\psi_j}$ are the columns of $\bar{\mathbf{\Phi}}$ and $\mathbf{\Psi}$. This is sometime referred to as mutual coherence [23].

A small μ indicates a low coherence system, i.e. $\bar{\mathbf{\Phi}}$ and $\mathbf{\Psi}$ are highly incoherent. It can be interpreted as a rough characterization of the degree of similarity between $\bar{\mathbf{\Phi}}$ and $\mathbf{\Psi}$. A example of maximally incoherent system is the pair Fourier basis and canonical basis.

4.2.3 CS recovery

Signal recovery via ℓ_1 minimization

Given an observation vector $\mathbf{y}$ and signal $\mathbf{x}$, which is known to be sparse, we search for the sparsest solution such that $\mathbf{y} = \mathbf{A}\mathbf{x}$. Denoting $\|.\|_{\ell_0}$ as the ℓ_0 pseudo-norm, which is the number of non zero elements in a vector, it is natural to recover $\mathbf{x}$ by solving an optimization problem of the form

$$\min_{\tilde{\mathbf{x}} \in \mathcal{R}^n} \|\tilde{\mathbf{x}}\|_{\ell_0} \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\tilde{\mathbf{x}}. \tag{4.9}$$

Nevertheless this problem, involving the ℓ_0 norm, requires combinatorial optimization. Fortunately, replacing the ℓ_0 norm by the ℓ_1 norm, which is the sum of absolute values of every element in a vector, has been shown to find the solution of (4.9) equivalently if $\mathbf{x}$ is sufficiently sparse [46]. It comes to solve this ℓ_1 minimization problem

$$\min_{\tilde{\mathbf{x}} \in \mathcal{R}^n} \|\tilde{\mathbf{x}}\|_{\ell_1} \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\tilde{\mathbf{x}}. \tag{4.10}$$

This is often called LASSO [119]. Alternatively, the measurements might be contaminated with a certain amount of noise. In this case the data consistency constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$ could be relax and instead defined as $\|\mathbf{y} = \mathbf{A}\mathbf{x}\|_{\ell_2} < \epsilon$. It comes to solve this relaxed version of (4.10)

$$\min_{\tilde{\mathbf{x}} \in \mathcal{R}^n} \|\tilde{\mathbf{x}}\|_{\ell_1} \quad \text{subject to} \quad \|\mathbf{y} = \mathbf{A}\mathbf{x}\|_{\ell_2} < \epsilon. \tag{4.11}$$

Many algorithms exist to solve (4.11). We can cite the family of soft-thresholding algorithm as ISTA and FISTA [12] or the coordinate descent method [14]. A more complete list of convex solver is presented in [139].

Bound on the number of measurements

The following result asserts that when $\mathbf{x}$ is S-sparse, the recovery via ℓ_1 minimization is exact if the number of measurements is above a certain bound [20]. Indeed, select m measurements uniformly at random, then the recovery is exact if

$$m \geq C \cdot \mu(\bar{\mathbf{A}})^2 \cdot S \cdot \log n, \tag{4.12}$$

where C is a constant, S the signal sparsity, $\mu(\bar{\mathbf{A}})$ the coherence of $\bar{\mathbf{A}}$ and n the length of the sparse signal $\mathbf{x}$. Note that the bound is given up to a constant multiplier. This result shows that smaller the coherence of the system, fewer measurements are required.

Stable recovery and the Restricted Isometry Property

Reference [21] established some results about the accuracy of the CS reconstruction of a sparse signal $\mathbf{x}$ at the condition that the sensing matrix $\mathbf{A}$ obeys the RIP, with an isometry

constant defined as :

Definition 2. *For each integer S=1,2,..., we define the isometry constant δ_s of a matrix* $\mathbf{A}$ *as the smallest number such that*

$$(1-\delta_s)\|\mathbf{x}\|_{\ell_2} \leq \|\mathbf{A}\mathbf{x}\|_{\ell_2} \leq (1+\delta_s)\|\mathbf{x}\|_{\ell_2} \tag{4.13}$$

holds for all S-sparse vectors, that is vectors that have at most S nonzero entries.

In particular, [21] proved, for $\delta_{2S} < \sqrt{2}-1$ and noiseless recovery, that the solution x^* of (4.10) obeys :

$$\|\mathbf{x}^* - \mathbf{x}\|_{\ell_2} \leq C_0 S^{-1/2}\|\mathbf{x}_s - \mathbf{x}\|_{\ell_1}, \tag{4.14}$$

with C_0 a constant and x_S the best sparse approximation knowing exactly the locations and amplitudes of the S-largest entries of $\mathbf{x}$. Then, if $\mathbf{x}$ is exactly S-sparse, the solution is exact. This bound is given for noiseless recovery.

When $\mathbf{x}$ is corrupted by noise, the error between $\mathbf{x}^*$ (the solution of (4.11)) and $\mathbf{x}$ is

$$\|\mathbf{x}^* - \mathbf{x}\|_{\ell_2} \leq C_0 S^{-1/2}\|\mathbf{x}_S - \mathbf{x}\|_{\ell_1} + C_1\epsilon, \tag{4.15}$$

with C_0 and C_1 two constants and ϵ the noise level such that $\|\mathbf{z}\|_2 \leq \epsilon$. [21] still assumes that result in (4.15) holds for $\delta_{2s} < \sqrt{2}-1$. (4.15) clearly demonstrates that if x is exactly k-sparse than the accuracy of the solution only depends on the noise.

4.3 Generalization to a broader class of representations

4.3.1 CS with overcomplete and redundant dictionary

Before beginning this section, we introduce a more convenient formulation of the sensing matrix $\mathbf{A}$. We decompose it as a product of a dictionary $\mathbf{D}$ and a measurement system $\mathbf{\Phi}$ such that $\mathbf{A} = \mathbf{\Phi}\mathbf{D}$.

The discrete and finite length CS theory described in Sec. 4.2 and [46, 20, 21, 23] holds when $\mathbf{D}$ (called $\mathbf{\Psi}$ in the precedent section) is an orthonormal matrix. A recent work [22] was published to generalize this CS theory and to recover $\mathbf{x}$ which is sparse in a redundant and overcomplete dictionary. The motivation of this work is to enlarge the scope of applications of the CS theory where signals are not sparse in an orthonormal basis but in overcomplete dictionaries. Our contribution in **chapter 7** is based on this framework.

In this section we present the results regarding this generalization of the CS theory. In particular, we focus our attention on the two main differences with the classical CS theory presented in Sec. 4.2, namely

- The orthonormality of the dictionary $\mathbf{D}$ is no longer required.
- We use a RIP adapted to the dictionary $\mathbf{D}$, called the D-RIP, instead of the RIP described in definition 2.

Is an orthogonal sparse system a necessary recovery condition?

In the previous section, we assume that the columns of the sparse basis are orthogonal. This assumption leads to well established results regarding the accuracy and the robustness of the CS recovery. If the columns of $\mathbf{D}$ are correlated then there would be no hope to reconstruct a unique sparse vector $\mathbf{c}$ from measurements $\mathbf{y} = \mathbf{\Phi D c}$. However, in many applications, we may want to recover the signal $\mathbf{x} = \mathbf{Dc}$ rather than the sparse vector $\mathbf{c}$. Then [22] suggests that the orthogonality is not necessary but, instead, prove that the solution of (4.11) is very accurate provided that $\mathbf{c}$ has rapidly decreasing coefficients.

The D-RIP

Equivalently to the RIP, [22] gives a broader condition on the property of the measurement matrix $\mathbf{\Phi}$ called the D-RIP.

Definition 3. *(D-RIP) Let Σ_s be the union of all subspaces spanned by all subsets of s columns of $\mathbf{D}$. We say that the measurement matrix $\mathbf{\Phi}$ obeys the restricted isometry property adapted to $\mathbf{D}$ (the D-RIP) with constant δ_s if*

$$(1 - \delta_s)\|\mathbf{v}\|_{\ell_2} \leq \|\mathbf{\Phi v}\|_{\ell_2} \leq (1 + \delta_s)\|\mathbf{v}\|_{\ell_2} \tag{4.16}$$

holds for all $\mathbf{v} \in \Sigma_s$.

In particular, if the D-RIP is satisfied with $\delta_{2s} < 0.08$, then the solution $\mathbf{x}^*$ of (4.11) satisfies

$$\|\mathbf{x}^* - \mathbf{x}\|_{\ell_2} \leq C_0 s^{-1/2} \|(\mathbf{D}^*\mathbf{x})_s - \mathbf{D}^*\mathbf{x}\|_{\ell_1} + C_1 \epsilon, \tag{4.17}$$

where C_0 and C_1 are constants depending on δ_{2s}. (4.17) demonstrates that we may be able to recover a signal which is sparse with respect to an arbitrary dictionary. Note that the classical CS framework described in Sec. 4.2 provides analogous bound on the accuracy of the reconstruction with an isometry constant $\delta_{2s} < \sqrt{2} - 1$, which is less constrained and more easily satisfied than $\delta_{2s} < 0.08$.

4.3.2 Toward extending CS to recover continuous signals

We start by a reminder on precedent notations as to make easier the understanding of this section. We called $\mathbf{A} \in \mathbb{R}^{m \times n}$ the sensing matrix and we try to recover a signal $\mathbf{x} \in \mathbb{R}^n$ from a measurement vector $\mathbf{y} \in \mathbb{R}^m$, such that $\mathbf{y} = \mathbf{Ax}$. In Sec. 4.2.2, we also introduced $\bar{\mathbf{A}} \in \mathbb{R}^{n \times n}$, which is an "non-undersampled" version of $\mathbf{A}$.

The framework introduced by [46, 23] and described in Sec. 4.2 is focused on the recovery of discrete and finite length vectors and the generalization to the recovery of continuous signal is challenging. Some works have handled the case of the extension of the CS theory to a continuous signal reconstruction [52, 49], even though there is not a clearly defined underlying theory. In dMRI, we find several works toward extending this notion of continuous diffusion signal reconstruction via CS [36, 89, 97, 122]. These works are mainly experimental and show promising results regarding continuous signal recovery via CS. Although the

application of this discrete theory to continuous signals is not clearly defined, we feel from these works [36, 89, 97, 122] that the conditions inherent in the classical and discrete CS theory can be applied to recover continuous signals. This extension to the continuous case is made possible considering two important aspects of our continuous CS-dMRI framework:

- $\bar{\mathbf{A}}$ is a matrix whose the columns form an orthonormal basis of $\mathbb{R}^3$.
- The signal of interest is band limited, which allows one to to sample it in a specific range in the q-space.

A first step toward extending this theory in dMRI has been done in our recent work [95] and is presented in **chapter 6**. In this continuous dMRI framework, we show from very straightforward examples that some important conditions described in the classical CS are still required to obtain accurate and robust reconstruction of continuous dMRI signals, namely the sparsity, the incoherence and the RIP. In the following, we briefly describe these notions in the continuous case and make a link with the discrete case. The complete contribution, which has been published in [95], is presented in **chapter 6**.

Sparsity in continuous signals

We aim to reconstruct a continuous signal instead of a discret and finite length vector. For this purpose, we model the diffusion signal as a linear combination of a small number of continuous and orthonormal basis functions, leading to a continuous representation of the signal. We take advantage of the low degree of freedom inherent in the dMRI signal to decompose it into a sparse expansion of a finite dimensional orthonormal basis. Then, we speak about structured sparsity, i.e. the sparsity operator is built under the signal consideration so as to best fit it. We characterize the sparsity of a continuous signal by the number of continuous functions required to accurately model it. Moreover, the fact that the orthonormal basis is truncated is a very important point ensuring that the sparse decomposition is finite. If the orthonormal basis is composed of an infinite number of elements, it would not be clear how to define the sparse properties of these bases. In [52] the formulation of the problem is more complicated since the sparse expansion has an infinite length . Since we deal with truncated orthonormal bases, our study is focused on the case where the sparse expansion is finite. We present results on the sparsity of the dMRI signal with respect to different orthonormal bases in Sec. 6.3.2.

Incoherence of the sensing matrix

In the discrete CS theory, the coherence parameter μ in (4.7) is defined for a sensing matrix $\bar{\mathbf{A}}$ of finite size $n \times n$. At first glance, it appears impossible to extend this notion of incoherence to continuous functions defined on an infinite range of values, i.e., when the number of lines in $\bar{\mathbf{A}}$ is theoretically infinite. However, since the diffusion signal and the basis functions used to model it are band limited, we are able to compute the coherence parameter in an appropriate bandwidth. We further describe the coherence parameter and compute it for several orthonormal bases in Sec. 6.3.1. Obviously, when dealing with orthonormal bases, the maximal value of the coherence parameter μ is 1, but we cannot define a lower bound for μ in general since is depends on the basis function itself. In [52], they cleverly compute the largest spectral value of the sampled cross correlation between

basis elements because they deal with not band limited signal. Again, the fact that the signal is band limited facilitates the computation of the coherence and does not involve such derivation.

CS recovery

Since the problem is discretized and the orthonormal basis is truncated, it leads to a finite-dimensional convex problem, where one can use conventional algorithm to solve it and, then find a sparse signal decomposition $\mathbf{x}$ with respect to continuous and orthonormal basis functions. Globally, we operate in three steps. First, we acquire a finite number of measurement in appropriate locations (We see in **chapter 6** the best acquisition strategy). This results in a discrete and finite length measurement vector $\mathbf{y}$. We also discretize the sensing matrix $\mathbf{A}$ according to the acquisition protocol. Afterward, we reformulate the problem as a finite-dimensional problem and solve the problem (4.11). Finally, from this sparse decomposition we can estimate the corresponding continuous signal. We will see in **chapter 6** that the RIP still provide useful information regarding the robustness of the reconstruction. .

4.4 A state of the art of CS in dMRI

CS has brought an opportunity to reduce the amount of data usually required to recover signals for applications in many fields ranging from radar imaging [6, 61], through astronomy [17, 133] and communication [13] to medical imaging [77, 25, 118]. In particular the CS technique was found successful to recover MR images from highly undersampled k-space [77, 60, 55, 30, 124]. Diffusion MRI is no exception. Recently, there have been a high interest in developing dMRI application using the CS technique [108, 91, 36, 97, 82, 89, 106, 112, 122, 15, 71, 15, 57, 81, 92, 136, 138, 69, 28, 95] (This list of papers is non exhaustive).

Some are focused on the sparsity of the dMRI signal, others on the sampling protocol, or even both these aspects. Zhu et al [138] present a method that takes advantage of the joint sparsity between DW measurements sensitized to different diffusion gradients but acquired at the the same spatial location. In the same idea, Mania et al [81] propose to jointly undersample the k-space and the q-space in order to achieve high spatial and angular resolution. Specifically, a different sampling strategy for undersampling the k-space is involved for each diffusion direction. In our recent work [85], we investigate the choice of q-space sampling when one wants to acquire the DW images on multiple b-values.

We can not compare all these works since these different approaches do not aim to recover the same diffusion features. If one wants to estimate micro structure information, such as the FA, MD, he would use CS to recover the diffusion tensor. Landman et al [71] handled the CS-based recovery of the diffusion tensor. They characterize the dMRI signal as a sparse combination of diffusion tensors and use the underlying theory described in [9, 10] to get a diffusion tensor based propagator. In [138], the authors compute the FA and the MD from an estimated diffusion tensor. When directional diffusion features is asked then a CS-based estimation of the ODF is appropriate. For instance, [122, 97] focus on the reconstruction of the dMRI signal acquired at a unique b-value, in which the single shell measurements are modeled as a sparse and linear combination of spherical functions from

an overcomplete and redundant dictionary. These two works provide analytical formulae to estimate the ODF. In addition, a comparison study has been published [69] between a DTI-based fiber tractography and a CS+HARDI-based tractography using the spherical ridgelets of [96]. The comparison is performed at the same number of measurement, and shows promising results concerning the use of CS-based ODF in fiber tractography.

Recently published works [91, 82] describe a CS-based method to accelerate the DSI acquisitions and recover the EAP constrained on a cartesian grid of size $11 \times 11 \times 11$. This method is discrete and requires numerical computation to estimate features related to dMRI. Still in the discrete framework, Paulsen et al [102] uses CS to recover a very high resolution EAP on a cartesian grid of size $128 \times 128 \times 128$, which comes to nearly 1576 times larger than the classical grid size used in DSI. From here, they manage to obtain an acceleration factor of 32. However, even with this huge acceleration factor, this technique still requires nearly 49 times more measurements than DSI does.

More recently, continuous formulations of the CS problem have emerged in order to analytically estimate the 3D dMRI signal [36, 89, 95]. It allows one to interpolate and extrapolate the observed data, to analytically estimate the EAP and several diffusion features. Some of these works are described in the following. We published a paper regarding the continuous recovery of dMRI signal via CS [95]. **Chapter 6** presents this work.

It is worth noting that sparse dictionary learning techniques have recently become popular in dMRI. Some of these approaches are non-parametric and aim to learn a collection of discrete atoms. For instance, [15, 57] learn dictionaries from DSI like acquisitions and use it to either denoise full DSI data or to perform undersampled DSI acquisitions and reconstructions. In particular, Gramfort et al [57] nicely exploit the symmetry of the signal in order to assess free parameters of the dictionary learning problem. Other dictionary learning approaches are parametric and provide continuous representations of the atom. A work regarding parametric dictionary learning was published in [136], in which the dictionary atoms are formed by a weighted combination of 3^{rd} order B-splines. The work of [136] appears promising in reconstructing the diffusion signals, and further enhancement could be done regarding the development of analytical formulae to estimate other diffusion features. More recently, we proposed in [86] to learn a dictionary where each atom is constrained to be a parametric function. In [86], this parametric function is a combination of a radial part and an angular part represented by the symmetric and real Spherical Harmonics (SH). The radial part is a polynomial weighted by an exponential. However, this approach essentially handles the learning of the radial part, i.e. the polynomial coefficients and a scale parameter in the exponential. One of our recent contribution, which is presented in chapter 7 and published in [28], propose a more complete model where the angular part is also learned.

In this section we propose to review a non exhaustive list of papers regarding CS recovery in dMRI including also sparse reconstruction techniques which are closed in spirit to CS and we try as best as we can to insert them within the different frameworks described in this chapter. We chose these papers because they present similarities with our own contributions regarding the application of CS in dMRI. These contributions are described in **chapters 5, 6 and 7**.

4.4.1 CS to reconstruct the diffusion tensor

Resolution of crossing fibers with constrained compressed sensing using diffusion tensor MRI [71]

Landman et al [71] use CS to fit a multi-tensor model to the observed data $\mathbf{y}$. In this setting, the sensing matrix $\mathbf{A}$ is a finite and large collection of tensor and $\mathbf{x} = [x_0, ..., x_N]$ (in Eq. 4.3) corresponds to the weighting coefficients in the multi-tensor model such as $y = \sum_i^N x_i \exp -bg_k^T \mathbf{D}_i g_k + \eta$. Then, each atom (or column) of $\mathbf{A}$ is defined by a tensor model characterized by $\mathbf{D}_i$ (Through $\mathbf{D}_i$ we can change the FA, the orientation etc..) and they constrain the weighting parameters vector $\mathbf{x}$ to be sparse. Moreover, they force $\mathbf{x}$ to be positive which enforce the sparsity of the solution. This work is part of the overcomplete and redundant dictionary since elements of $\mathbf{A}$ may be correlated. Furthermore, the dMRI signal representation using $\mathbf{A}$ is parametric, thus, enables a continuous representation of a given dMRI signal.

Here is listed the negative and positive aspects of the method:

+ Accurate estimation of the multi-tensor model without any prior knowledge regarding the number of compartments.

+ Outperform the conventional DTI technique in estimating fiber crossing, with the same number of samples (30 samples).

+ The positivity constraint enforce the sparsity of the solution.

- Finite basis with arbitrary chosen and fixed parameters.

- Sensitive to the parameter used to build the dictionary (Not robust to a variation of the FA parameter for instance)

- The technique is limited to the estimation of the multi-tensor model

4.4.2 CS to reconstruct the ODF

Spatially regularized Compressed Sensing for High Angular Resolution Diffusion Imaging [97]

Michailovich et al [97] model the attenuation signal $\mathbf{y}$ acquired for an unique b-value (Single shell acquisition) and provide a framework to estimate the ODF. They represent the attenuation signal as a sparse linear combination of spherical ridgelet Ψ_i, i.e. $y(\mathbf{u}) = \sum_i^N x_i \Psi_i(\mathbf{u}) + \eta$. The sensing matrix $\mathbf{A}$ is, thus, an overcomplete dictionary composed of a finite set of spherical ridgelet evaluated at the corresponding measurement positions. In addition to the sparse regularization, they propose to use a spatial regularization between adjacent signals. They use a modified version of the Bregman algorithm to solve this problem. Michailovich et al [97] also propose an analytical and continuous formula to estimate the ODF.

Here is listed the negative and positive aspects of the method:

+ Low coherence between the Dirac sampling basis and the spherical ridgelets.

+ Multi-fibre analysis (via the ODF) with only 16-24 measurements (typical number of measurements for DTI, for instance).

+ Take into account the dependencies between spatially adjacent signals.

- Difficulty in setting the two free parameters in the reconstruction problem (for the sparse regularization and the spatial regularization).

Probabilistic ODF Estimation from Reduced HARDI data with Sparse Regularization [122]

In this work, Tristan-Vega et al [122] starts to consider an appropriate framework to sparsely represent the ODF, named the Spherical Wavelet (SW). Then they derive a dual framework to model the diffusion signal in the q-space. They solve the CS problem (in Eq. 4.11) using a discrete set of this dual framework as columns of the sensing matrix $\mathbf{A}$, which seems to provide a sparse representation of the diffusion signal in the q-space. They also propose to relax the ℓ_1 regularization in Eq. 4.11 to an ℓ_2 regularization (using a non negative least square (NNLS) algorithm). This ℓ_2 problem is resolved by mean of a closed form. However, an NNLS algorithm is not sufficient to ensure that the solution of problem in 4.11 get closer to the solution of ℓ_2 approach. For this reason, and because the ODF is a probability law, Tristan-Vega et al constrain the ODF to be positive in a discrete set of directions. The minimum number of measurements to obtain reliable ODF is approximately 14-16, but all these synthetic experiments are performed with a SNR equal to 40 and higher, thus making this conclusion unfair. Because [122] provide a continuous representation of the ODF from a sparse expansion with respect to an overcomplete and redundant dictionary, it enters both CS categories.

Here is listed the negative and positive aspects of the method:

+ Original and efficient framework to sparsely represent the ODF.

+ Relaxation of the problem in Eq. 4.11 to a ℓ_2 approach, making the reconstruction time attractive.

+ Only 14-16 measurements are needed to obtain reliable ODF reconstruction.

- All the synthetic experiments are performed with a SNR equal to 40 and higher, which roughly corresponds to a noiseless case.

4.4.3 CS to reconstruct the 3D diffusion signal and underlying features

Accelerated Diffusion Spectrum imaging in the human Brain Using Compressed Sensing [82]

Menzel et al [82] handle the case of reconstructing the EAP on a cartesian grid, which offers an original way to accelerate the DSI acquisition. It is a direct extension to the work of Lustig et al [77] with the difference that [82] apply the CS theory on 3D signals. [82] randomly undersamples the diffusion signal following a Gaussian distribution and constrains the gradient of the EAP to be sparse. Then, the measurement system $\mathbf{\Phi}$ is the Fourier matrix and the sparse system $\mathbf{\Psi}$ is characterized by a gradient transform. This formulation (included the gradient transform) is also known as a total variation (TV) reconstruction. This work mainly uses results from the classical and discrete CS theory, excepted that the gradient transform is not orthonormal.

Here is listed the negative and positive aspects of the method:

+ Original and efficient way to accelerate the DSI acquisition.

- The gradient transform is not sparse enough to represent the EAP and another sparse representations should be used intead.

- Discrete approach, which leads to discrete computation of diffusion features.

- High b-values required for the acquisition.

Spherical Polar Fourier EAP and ODF reconstruction via Compressed Sensing in diffusion MRI [89]

In this work [89], we describe a method to model the EAP in terms of the Spherical Polar Fourier (SPF) basis using CS. It exploits the duality between the SPF basis and its counterpart in the q-space, called the SPF dual (SPFd) basis, in order to sparsely represent the diffusion signal. Thus, the sensing matrix $\mathbf{A}$ is the orthonormal basis matrix constructed from the SPFd orthonormal functions. In this approach, $\mathbf{x}$ contains the coefficients representing the diffusion signal with respect to the SPFd basis. This approach enables a continuous modeling of the EAP at any radius (with respect to the SPF basis functions) and provides an analytical formula to estimate the ODF. In **chapter 6** of this thesis manuscript, we further describe this work while adding comparisons with other orthonormal bases.

Here is listed the negative and positive aspects of the method:

+ Continuous modeling of the diffusion signal, EAP and ODF.

+ Low to medium b-values required for the acquisition (contrary to DSI).

- Multiple shells sampling is needed (sometime not well known from the community).

An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators [136]

Ye et al [136] published a work regarding parametric dictionary learning combined with a spatial regularization. The sensing matrix $\mathbf{A}$ is an overcomplete and redundant dictionary, in which the dictionary atoms are formed by a weighted combination of 3^{rd} order B-splines. It provides a continuous modeling of the diffusion signal. It proved that the method is efficient on synthetic data simulated with 81 gradient directions at a single b-value. This work appears promising in reconstructing the diffusion signals, and further enhancement could be done regarding the development of analytical formulae to estimate other diffusion features.

Here is listed the negative and positive aspects of the method:

+ Continuous modeling of the diffusion signal and the EAP.

+ A single shell sampling is sufficient to recover the diffusion signal.

+ Spatial regularization.

- An analytical formula to estimate the ODF lacks.

Sparse DSI: Learning DSI structure for denoising and fast imaging [57]

Gramfort et al [57] learn dictionaries from DSI like acquisitions and use it to either denoise full DSI data or to perform undersampled DSI acquisitions and reconstructions. In particular, they nicely exploit the symmetry of the signal in order to assess free parameters of the dictionary learning problem. Each atom of **A** represents a particular structure of the DSI signal in a cartesian grid. They manage to recover a full DSI signal from an undersampled q-space of 40 measurements.

Here is listed the negative and positive aspects of the method:

\+ They exploit the positivity and the summetry inherent to the diffusion signal

\- Numerical computation of diffusion features is needed.

4.5 Summary

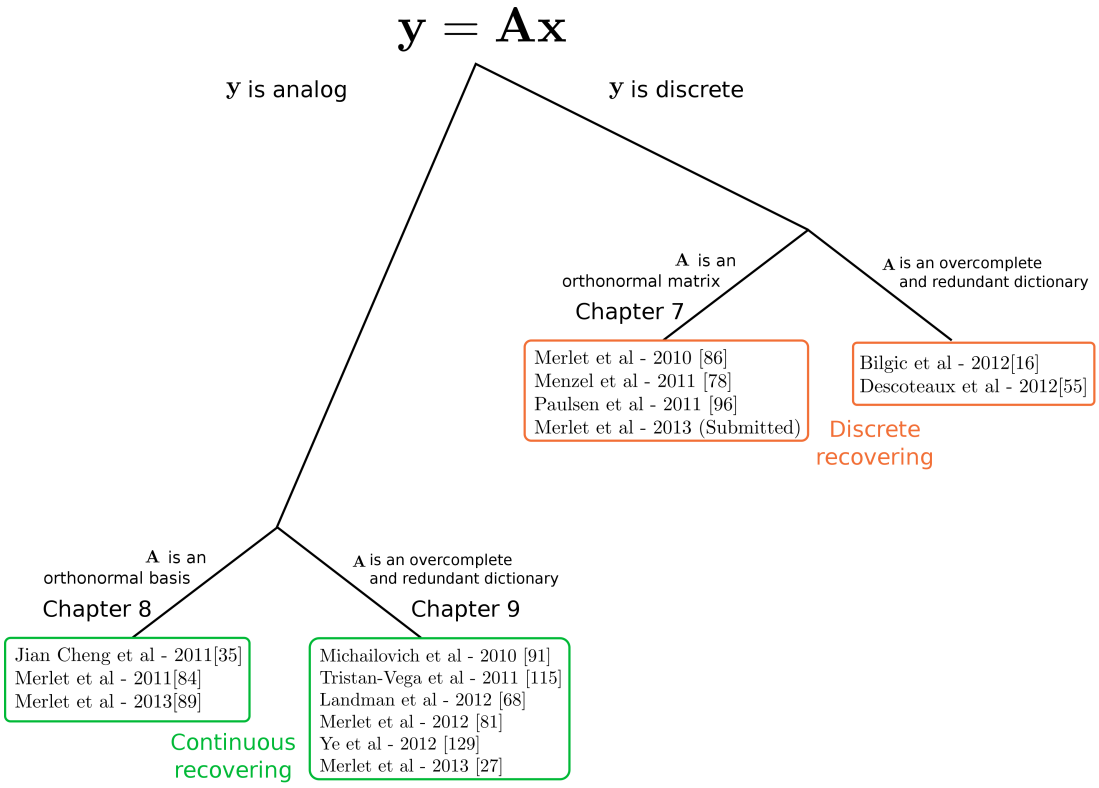


Figure 4.5: Contribution of CS in dMRI and related work.

Part III

Contributions

Chapter 5

Accelerated DSI with efficient sampling and sparse discrete representation

Contents

5.1 Motivations . . . 74
5.2 Method . . . 75
5.2.1 Sparse representation of the EAP . . . 75
5.2.2 Features used to characterize the EAP in this chapter . . . 77
5.2.3 Synthetic data simulation . . . 77
5.2.4 Real data acquisition . . . 78
5.2.5 Validation . . . 78
5.3 Results . . . 78
5.3.1 Choice of sampling schemes . . . 79
5.3.2 Experiments on synthetic data . . . 80
5.3.3 Experiments on real data . . . 84
5.4 Discussion . . . 85
5.5 Conclusion . . . 89

Overview

Diffusion Spectrum Imaging (DSI) enables one to accurately reconstruct the Ensemble Average Propagator (EAP) at the expense of having to acquire a large number of measurements. Compressive Sensing (CS) offers an efficient way to decrease the required number of measurements. The purpose of this work is to describe a way of performing an accurate and robust EAP reconstruction using the CS technique to accelerate the DSI acquisition. Then, this work is part of the discrete and finite-dimensional CS framework described in Sec. 4.2.

This study is focused on two key ingredients required for an accurate and robust CS-recovery of a discrete EAP, i.e. the sparsity and the sampling protocol. Firstly, several sparse representations of the EAP are studied. Then, a new and original sampling scheme is proposed to obtain significant information from the q-space.

The robustness and the efficiency of several sampling protocols and sparse transforms are thoroughly compared to identify the ingredients of an optimal CS reconstruction. Experiments on synthetic and human brain data have been carried out, which demonstrate that approximately 64 measurements are sufficient to recover significant information regarding the EAP and to accurately estimate the orientation distribution function (ODF) and the kurtosis.

We conclude that an appropriate sparse transform and sampling protocol considerably improves the quality of the CS reconstruction of the EAP, the ODF and the kurtosis.

5.1 Motivations

In Diffusion Spectrum Imaging (DSI) [130], we obtain the EAP $P(\mathbf{R})$ by directly taking the inverse Fourier transform of $E(\mathbf{q})$. However, DSI requires the acquisition of many diffusion weighted images (DWI) sensitized to different orientations in the sampling space, to obtain a high-resolution EAP. In brief, while this technique has the advantage of being a very good approximation of the water diffusion phenomenon, it is limited by the long acquisition time due to the large number of required samples. Recent improvements have been shown to decrease the scan time by at least a factor of three using fast acquisition sequences and parallel imaging [115, 114]. Further improvements can be achieved using signal processing techniques, in order to decrease the acquisition time. That is where the Compressive Sensing (CS) technique comes in to accelerate DSI.

As described in **chapter 4**, CS aims to accurately reconstruct sparse signals from undersampled measurements acquired below the Shannon-Nyquist rate. Many contributions regarding the application of CS and related topic in dMRI have emerged [108, 91, 36, 97, 82, 89, 112, 122, 128, 15, 72, 15, 57, 92, 86, 136, 95]. We give a review of these papers in **chapter 4**. Some of these works consider modeling the signal in orthonormal bases as in [36, 89, 95] or in overcomplete and redundant dictionaries as in [108, 97, 122, 72, 15, 57, 86, 136]. The contribution of this chapter is confined on the acceleration of DSI via the CS technique (CS-DSI). This problem has already been adressed by [91, 82, 112, 15, 92] and consists in reconstructing a discrete EAP, which is sparse with respect to a prespecified discrete transform. For instance, Menzel et al [82] solve the CS-DSI problem by considering the gradient operation as sparse transform, also known as total variation (TV) regularization. Bilgic et al [15] combine a TV regularization and a sparse

constraint using the Haar based discrete wavelet transform (DWT). In [91], we take the identity operator as sparse transform. However, these works were conducted independently and there has not been an analysis and comparison of the differences between them, which makes difficult the choice of an optimal sparse transform. Moreover, some improvements can be made to enhance the accuracy and the robustness of the reconstruction. For instance, the Haar wavelet (used in [15]) is not a good sparse representation of the EAP and other wavelet bases appear more appropriate for this task [92]. The sampling protocol is, as well, not deeply studied.

In this work, we describe how to accurately and efficiently reconstruct a discrete EAP using the CS technique, and we show that a CS-based method using an adequate sparse transform and sampling scheme accelerates DSI by significantly reducing the common number of measurements required in this technique. For this purpose, we highlight two important points of the CS theory, i.e. the sparse representation of the propagator and the sampling protocol. At first, we study and compare five sparse and discrete representations of the EAP, i.e. the DWT, the stationary wavelet transform (SWT), the dual tree wavelet transform (DTWT), the gradient of the EAP (also known as total variation (TV) regularization) and the identity (I), i.e. the canonical basis. Then, we propose a new and original sampling scheme to use with our best sparse transform in order to obtain significant information regarding the EAP. This work primarily handles the reconstruction of the EAP, which is a radial and angular characterization of the diffusion phenomenon, and the starting point to estimate many other diffusion features such as the orientation distribution function (ODF) [127, 130, 42], its corresponding GFA, and the Kurtosis [65] as shown in this chapter. Moreover other features could be derived from the EAP and a tractography algorithm directly based on the EAP was recently developed [94], which make the EAP a key tool for a good understanding of the diffusion phenomenon.

The chapter is structured as follows. At first, we briefly describe the CS technique. Then, we present the five sparse representations described above. The last section is divided into three parts: 1) we propose an efficient and robust sampling protocol; 2) we review the capabilities of each of the sparse representations on the reconstruction of synthetic data; 3) we validate the CS reconstruction on in vivo human brain data.

5.2 Method

The CS theory described in Sec. 4.2 is well established considering discrete signals [46, 20, 21, 23, 22] and is adapted to the reconstruction of a discrete EAP. In this application, the sensing matrix is the product of a sparse system $\mathbf{\Psi}$ and an orthogonal measurement system characterized by the Fourier basis. We saw that there are two important parameters to obtain accurate and robust CS reconstructions: the sparse system $\mathbf{\Psi}$, which has to be incoherent with the measurement system, and the sampling protocol inherent in the RIP. In the following section, we review five potential choices for $\mathbf{\Psi}$. The choice of the sampling scheme is discussed in the experimental section.

5.2.1 Sparse representation of the EAP

In this section, we present five candidates of EAP representation to be used in the CS context. Two of them are part of the orthonormal basis framework, i.e. the canonical basis

and the discrete wavelet transform (DWT). Two others are overcomplete and redundant representations, namely the discrete stationary wavelet transform (SWT) and the real dual-tree discrete wavelet transform (DTWT). Then, we propose to study the gradient operation as a sparse transform, commonly known as a total variation constraint when combined with the ℓ_1 regularization.

The canonical basis

Because of the quasi-Gaussian nature of the EAP attenuation, we assume in [91] that the EAP is sparse in its original domain and do not use any sparse transform, i.e. $\mathbf{\Psi} = I$ where I represent the canonical basis. This assumption is true in certain cases. For example, when the voxel configuration corresponds to diffusion in only one fiber direction, the EAP values decay rapidly to zero in the direction orthogonal to the fiber. However, when diffusion occurs in several fiber directions or in the gray matter and CSF areas where the EAP is isotropic, the sparsity assumption is not correct anymore. In [91] we have still found experimentally that solving (4.11) in this setting leads to qualitatively good EAP reconstructions. This is explained by the fact that the pair canonical and Fourier bases provide a maximally incoherent system. Then, the ℓ_1 term in (4.11) plays a role of denoiser. Even if the sparsity assumption is weak in the sense that the canonical basis does not always provide a sparse EAP representation, we still call it Compressive Sensing since the coherence condition is met.

The discrete wavelet transform (DWT)

The DWT is an extremely well established tool in the image processing community and is used, for instance, in image compression because it provides highly sparse representation of natural images. In dMRI, Bilgic et al [15] used a Haar-based DWT to solve the CS-DSI problem, whereas the Haar wavelet was found to not provide a sparse EAP representation [112, 92]. Based on the work of [112, 92], we consider a DWT based on the biorthogonal and symmetric Cohen-Daubechies-Feauveau (CDF) wavelet. In particular, we use the CDF 9/7, which has four vanishing moments. Moreover, the system built from the pair wavelet and Fourier bases has a large incoherence as long as we select the measurements at random [23]. Several public and open-source libraries such as WaveLab [1] in MATLAB and PyWavelets [2] in Python propose efficient DWT implementations.

The discrete stationary wavelet transform (SWT)

The classical DWT is not a shift-invariant transform, which is a useful property to avoid visual artifacts around the discontinuities of a signal [37]. The SWT overcomes this limitation at the expense of providing more coefficients than the size of the signal itself, leading to an overcomplete and redundant transform. The mother wavelet used for the SWT is still the CDF 9/7.

[1] http://www-stat.stanford.edu/∼wavelab/Wavelab_850/index_wavelab850.html
[2] http://www.pybytes.com/pywavelets/

The real 3D dual-tree discrete wavelet transform (DTWT)

The DTWT is implemented using four separable 3D DWTs in parallel. Then, the subbands of the four DWTs are combined appropriately [113]. This transform has the benefit of being oriented, which is appreciated to sparsely represent 3D signals with directional properties such as the EAP. However, because the DTWT is 4-times more expensive compared to the DWT, it comes under the framework of redundant and overcomplete transforms.

The gradient transform

Setting Ψ as a gradient transform comes to replace the sparsity constraint by a total variation (TV) regularization [110]. Strictly speaking, this method is known as TV reconstruction. The TV reconstruction involves non linear optimization and the use of the conjugate gradient algorithm to find the corresponding solution. [82] was the first to use the TV regularization as a diffusion-domain constraint in the CS-DSI problem. Note that the TV regularization was also applied in the spatial domain in combination with a sparse constraint in the diffusion domain by [97].

5.2.2 Features used to characterize the EAP in this chapter

In addition to the EAP, we can compute the Orientation Distribution Function (ODF) as defined in (3.21). Note that in our setting the maximal bound for the ODF integration is limit to $R_{\max}$ (the maximum radius of the EAP in the grid), since we only have a discrete and finite EAP.

Another diffusion feature, which account for both the angular and radial information of the diffusion phenomenon is the kurtosis. The kurtosis aims to quantify the non-Gaussianity of a probability density function. The diffusion kurtosis K in the direction $\mathbf{n}$ is defined by

$$K(\mathbf{n}) = \frac{\langle (\mathbf{R}\cdot\mathbf{n})^4 \rangle}{\langle (\mathbf{R}\cdot\mathbf{n})^2 \rangle^2} - 3, \tag{5.1}$$

with $\mathbf{n}$ is a 3D unit vector, $\mathbf{R} = R\mathbf{r}$ and $\langle (\mathbf{R}\cdot\mathbf{n})^n \rangle = \int P(\mathbf{R})(\mathbf{R}\cdot\mathbf{n})^n d^3\mathbf{R}$ the n^{th} order moment of $P(\mathbf{R})$ about its mean value. In our experiments, we use the method described in [64] to estimate $K(\mathbf{n})$.

5.2.3 Synthetic data simulation

We validate the reconstruction on synthetic data (Sec. 5.3.1 and 5.3.2) using the data set provided by the HARDI contest at ISBI 2012 [3]. The contest was organized with the aim of providing a way for different groups to propose their own reconstruction algorithms and to compare fairly their methods against each other on a common set of ground-truth data. In this contest, the normalized diffusion signal $E(\mathbf{q})$ is generated from the multi-tensor model (see Sec. 10.7).

From this synthetic model, other ground-truth diffusion features have been derived to estimate other diffusion features such as the orientation distribution function (ODF) [127, 130, 42] and the diffusion Kurtosis [65]. Note that, for the synthetic experiments, we

[3] http://hardi.epfl.ch/

consider the dMRI signals contained in the file Training_IV, with 2-fibers crossing at equally represented angles from 30° to 90° and different volume fractions, yielding 610 synthetic signals.

Some experiments are performed in a noisy scenario where Rician noise is added. For these experiments, the noisy signal is computed as $E_{\text{noisy}} = \sqrt{(E+\epsilon_1)^2 + \epsilon_2^2}$, where $\epsilon_1, \epsilon_2 \sim \mathcal{N}(0, \sigma)$ with $\sigma = 1/\text{SNR}$.

5.2.4 Real data acquisition

A standard DSI acquisition mimicking the original DSI protocol [130] was done on a 3 T system (Philips Achieva X, Best, The Netherlands), equipped with a whole body gradient (40 mT/m and 200 T/m/s) and a 8-channel head coil. Single-shot spin-echo EPI measurements with isotropic 2 mm spatial resolution and 515 DW measurements were acquired comprising q-space points of a cubic lattice within the sphere of five lattice units in radius. TE/TR= 116 ms/14.9 s (including time for dynamic B0 stabilization), bandwidth in EPI direction= 1101 Hz, 128x128 matrix, 60 axial slices with a parallel imaging (SENSE) factor of 2, delta and Delta were 45.4 and 57.7 ms and maximal b-value of bmax = 6000 s/mm^2. We compute the SNR of the data as done in [45] and find a value of 38 at a b-value $b = 0$ s/mm^2 and 6.5 at $b = 6000$ s/mm^2. The SNR remains higher than 4-5, which is known to be the critical limit under which the noise profile becomes Rician. Hence, in our data, we make the coarse hypothesis that the noise nature is Gaussian.

5.2.5 Validation

We validate the reconstruction on two metrics derived from the EAP, i.e. the ODF [127, 130, 42] and the kurtosis [65]. From each estimated ODF, we assess the quality of the angular information contained in the EAP by computing the difference in the number of fiber compartments (DNC) and the angular error (AE) with respect to the known ground truth. To compute the DNC and the AE, we extract the maxima on the estimated ODFs and compare them to the ground truth maxima. Then, the DNC becomes the mean difference between the number of maxima extracted on the estimated ODFs and the true number of maxima, and the AE is computed between the maxima extracted on the estimated ODFs and the respective maxima within the ground truth. Concerning the kurtosis, it contains both radial and angular diffusion information. We validate the kurtosis by computing the normalized mean square error (NMSE) between the ground truth kurtosis and the estimated kurtosis evaluated on 100 directions generated using the static repulsion algorithm [68]. In the experimental part on real data, we also validate the EAP itself by computing the NMSE between the discrete EAP from the full DSI and the discrete EAP estimated via CS. The NMSE between a signal x and its estimation x_e is given by $NMSE = \frac{\|x-x_e\|_2^2}{\|x\|_2^2}$.

5.3 Results

In this section, we first present and study the efficiency and the robustness of the sampling protocol on synthetic data (Sec. 5.3.1). Then, we validate the reconstruction using various metrics while considering both synthetic (Sec. 5.3.2) and real data (Sec. 5.3.3).

5.3.1 Choice of sampling schemes

We saw in Sec. 4.2 that random sampling allows one to stably and accurately recover the EAP from noisy measurements. Moreover, [82] shows the importance of the sampling pattern to reconstruct oriented structures. In particular, [82] observes that a sampling scheme (SC) performs well with samples randomly distributed according to a Gaussian distribution with its mean being the center of the grid. This can be explained by the fact that most of the signal energy is focused around its center. However, when considering 3D signals of size $11 \times 11 \times 11$ (common size in DSI), such random undersampling schemes do not always ensure an homogeneous angular distribution of the samples. This is especially a matter while dealing with directional signal as dMRI signal. For this purpose, we start building our SC by generating N samples uniformly distributed on the sphere using the static repulsion algorithm [68]. Then, for each sample, we draw a random radius between the origin and the maximum radius of a sphere comprised in our acquisition grid. Finally, we match each of those samples in our grid as a way to obtain N samples in our DSI grid. This SC both ensures a homogeneous angular covering of the q-space and meets the CS requirement concerning the random aspect of the sampling protocol inherent in the RIP since it is radially random. We illustrate, in 2D, the way we build this SC in Fig. 5.1.

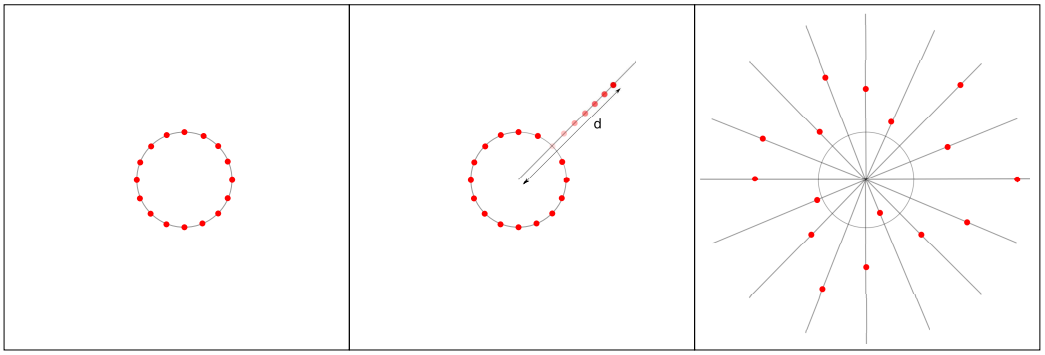


Figure 5.1: Sampling scheme construction in 2D. (left) We generate samples uniformly distributed on the sphere using the static repulsion algorithm [68]. (middle) Considering one sample, we draw it at a random distance d from the the origin and the maximum radius of a sphere comprised in our acquisition grid. (right) We repeat until all the samples are drawn. Note that, in the 3D case, we generate the sampling scheme on one hemisphere and then symmetrize it.

In this section we compare 3 sampling schemes and show the importance of homogeneous angular covering of the q-space :

- A uniform random sampling scheme (RU-SC), i.e. the samples are randomly distributed according to a uniform distribution.
- A Gaussian random sampling scheme (RG-SC), i.e. the samples are randomly distributed according to a Gaussian distribution with its mean corresponding to the center of the grid. This sampling scheme was used in [82]
- A random sampling scheme ensuring a homogeneous angular covering of the q-space (HA-SC) as described above.

To have an overview of these three SCs, we show in Fig.5.2 what would be given when dealing with 2D signals.

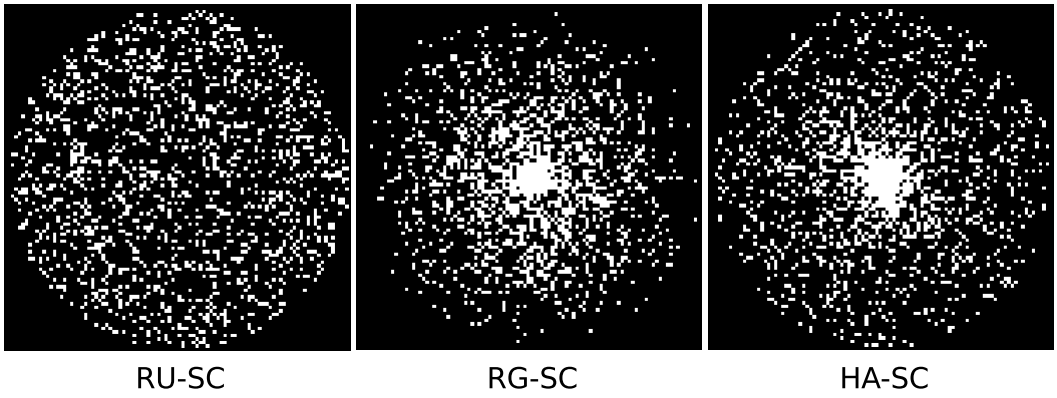


Figure 5.2: 2D sampling schemes corresponding to an uniform random sampling scheme (RU-SC), a Gaussian random sampling scheme (RG-SC), a random sampling scheme ensuring a homogeneous angular covering of the q-space (HA-SC). 2000 samples are used to build these three sampling scheme.

We propose to study the efficiency and the robustness of these SCs. For this purpose, we generate 100 instances of each of the three sampling schemes described above, i.e. the RU-SC, the RG-SC and the HA-SC. Note that, because the dMRI signal is symmetric, we generate each SC on a half hemisphere and then symmetrize it. Afterward, we compute the corresponding diffusion signals from the data set provided by the HARDI contest at ISBI 2012. From these diffusion signals, we estimate the EAPs via an inverse Fourier transform, and the ODFs via a radial integration of the EAPs. We then compute the variance and the mean values of the DNC and AE over the 100 experiments and repeat the process with a number of samples $N = 32, 64, 96, 128, 160, 192$. The results, shown in Fig. 5.3, give an overview of the angular information obtained before applying a reconstruction method on the signal.

Fig. 5.3 shows that the HA-SC leads to more accurate (lower mean values) and more robust (lower variance values) results than the two others schemes. These results seem to confirm that a higher degree of angular information is contained in the HA-SC. Fig. 5.3 also confirms the finding of [82] regarding that RG-SC is a better choice than RU-SC. In conclusion, even if the CS theory is usually based on uniform random sampling, we observe from these experiments and previous work [82] that random sampling protocols are more appropriate when we control the distribution of points. In particular, the HA-SC appears efficient and robust in recovering directional information, which is an important aspect in fiber tractography and other applications. For this reason, the following experiments are performed using the HA-SC. We validate, in Sec. 5.3.2, these preliminary results via a comparison of these three SCs on the CS reconstruction of synthetic data (see Fig. 5.5).

5.3.2 Experiments on synthetic data

Choice of the sparse representation. In this section, we compare the CS reconstruction using the five sparse representations presented in section 5.2.1, i.e.

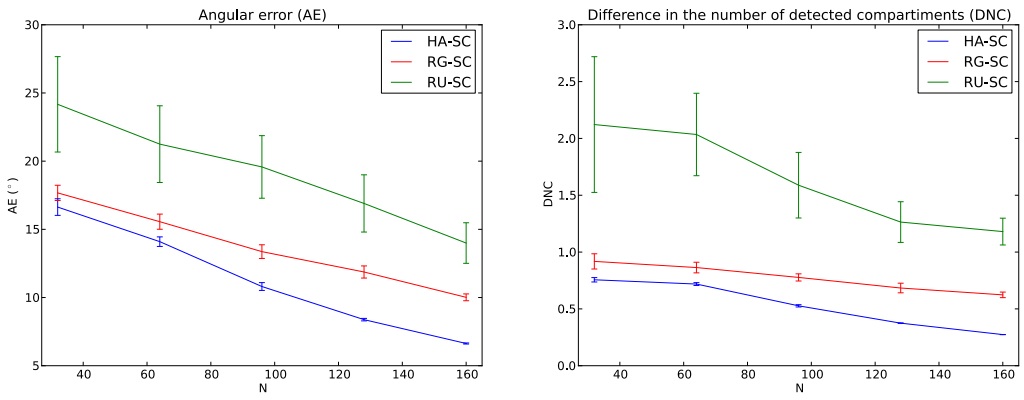


Figure 5.3: Mean and variance (vertical bar) of the angular error (AE) and the difference in the number of fiber compartments (DNC) when using an uniform random sampling scheme (RU-SC), a Gaussian random sampling scheme (RG-SC), a random sampling scheme ensuring a homogeneous angular covering of the q-space (HA-SC). These results are drawn from 100 trials.

- A CS reconstruction without applying any sparse transform on the EAP, i.e. $\Psi = I$ (CS-I).
- A CS reconstruction while applying a discrete wavelet transform (DWT) on the EAP (CS-DWT).
- A CS reconstruction while applying a stationary wavelet transform (SWT) on the EAP (CS-SWT).
- A CS reconstruction while applying a dual tree wavelet transform (DTWT) on the EAP (CS-DTWT).
- A total variation reconstruction (CS-TV)

We validate these five approaches on the reconstruction of synthetic data provided by the HARDI contest at ISBI 2012. We add Rician noise with SNR=20 and SNR=10 to the data. We generate several HA-SC and choose one given satisfactory results. We also compare the CS approaches with the full DSI reconstruction, i.e. when 257 samples are used, and the half sphere undersampled DSI presented in [70], i.e. with 102 samples (We call it DSI102 in the following). From the reconstructed EAPs, we estimate both the ODFs and the kurtosis. The resulting DNC, AE and kurtosis NMSE are shown in Fig. 5.4. Note that, to obtain the ODFs, we do not integrate between 0 and and $R_{\max}$ (the maximum radius of the EAP in the grid). Instead, we integrate the EAPs between the range $[R_{\max} \times \alpha, R_{\max} \times \beta]$ with α, $\beta \in [0, 1]$, because we had observed that well chosen bounds for the radial integration lead to better results in terms of DNC and AE. The optimal values of α and β are presented in Tab. 5.1. These values will be used for future ODF estimation.

Firstly, we observe that the bounds in Tab. 5.1 do not vary much for a given sparse transform and SNR. Moreover, the values of α and β are lower at SNR=10, since the signal

SNR=10 (α/β)	CS-I	CS-DWT	CS-SWT	CS-DTWT	TV
32 samples	0.2/0.6	0.2/0.7	0.3/0.8	0.3/0.8	0.2/0.8
64 samples	0.2/0.6	0.2/0.7	0.3/0.8	0.3/0.7	0.3/0.8
96 samples	0.2/0.6	0.3/0.7	0.3/0.7	0.3/0.6	0.2/0.6
128 samples	0.2/0.6	0.2/0.7	0.3/0.7	0.3/0.6	0.2/0.6
160 samples	0.2/0.6	0.3/0.7	0.3/0.7	0.3/0.7	0.2/0.6
192 samples	0.2/0.6	0.3/0.7	0.2/0.7	0.2/0.6	0.2/0.6

SNR=20 (α/β)	CS-I	CS-DWT	CS-SWT	CS-DTWT	TV
32 samples	0.3/0.8	0.2/1.0	0.3/0.9	0.3/0.8	0.3/0.8
64 samples	0.3/0.8	0.4/0.8	0.4/0.8	0.3/0.8	0.3/0.8
96 samples	0.3/0.8	0.3/0.8	0.4/0.8	0.4/0.8	0.3/0.8
128 samples	0.4/0.7	0.4/0.8	0.4/0.8	0.4/0.7	0.3/0.8
160 samples	0.4/0.7	0.4/0.8	0.4/0.8	0.3/0.8	0.3/0.8
192 samples	0.3/0.8	0.4/0.8	0.4/0.8	0.4/0.8	0.3/0.8

Table 5.1: Optimal α and β parameters (regarding the DNC and AE) to integrate the EAP for a given number of samples and a CS reconstruction method. The signals are contaminated by Rician noise at SNR=10 (top table) and SNR=20 (bottom table). For DSI and DSI102, $\alpha = 0.2$ and $\beta = 0.7$ for SNR=10 and $\alpha = 0.3$ and $\beta = 0.8$ for SNR=20.

is likely to be more noisy when going away from the origin and, thus, provide erroneous information (especially considering the r^2 factor in the ODF computation, amplifying these errors even more).

We first compare the CS-DWT and the CS-TV methods used in [15, 82]. Note that we use the CDF 9/7 wavelet in the CS-DWT method instead of the Haar wavelet used in [15] because experiments have shown that the Haar wavelet is not appropriate in sparsely representing the EAP [112, 92]. In Fig. 5.4, we observe that the CS-DWT method gives lower DNC values than the other methods. At SNR=10, the CS-TV method is not able to provide correct directional information. We deduce, from this observation, that the CS-TV method is too sensitive to noise. Regarding the angular error (AE), the CS-TV method gives slightly more accurate directions, but we cannot rely on this information since the number of maxima is not well approximated (high DNC values). The AE is, thus, underestimated in voxels where the number of detected compartments is higher than the true number of compartments. In terms of kurtosis NMSE, the CS-DWT globally leads to a lower NMSE than the other methods considering both SNR=20 and SNR=10. We conclude that the CS-DWT method remains robust to noise (even for a low SNR), and that very good results are observed for this technique. The CS-TV method gives satisfactory results in terms of DNC, AE and kurtosis at a SNR=20, but is not able to provide correct directional information at a SNR=10.

The CS-I method gives satisfactory results even if no sparse transform is applied on the EAP. This can be explained by the high incoherence between the real and Fourier space. Note that in our work the CS-I method is not better than the CS-DWT method, whereas the author in [15] observed that the application of a DWT leads to worse results compared

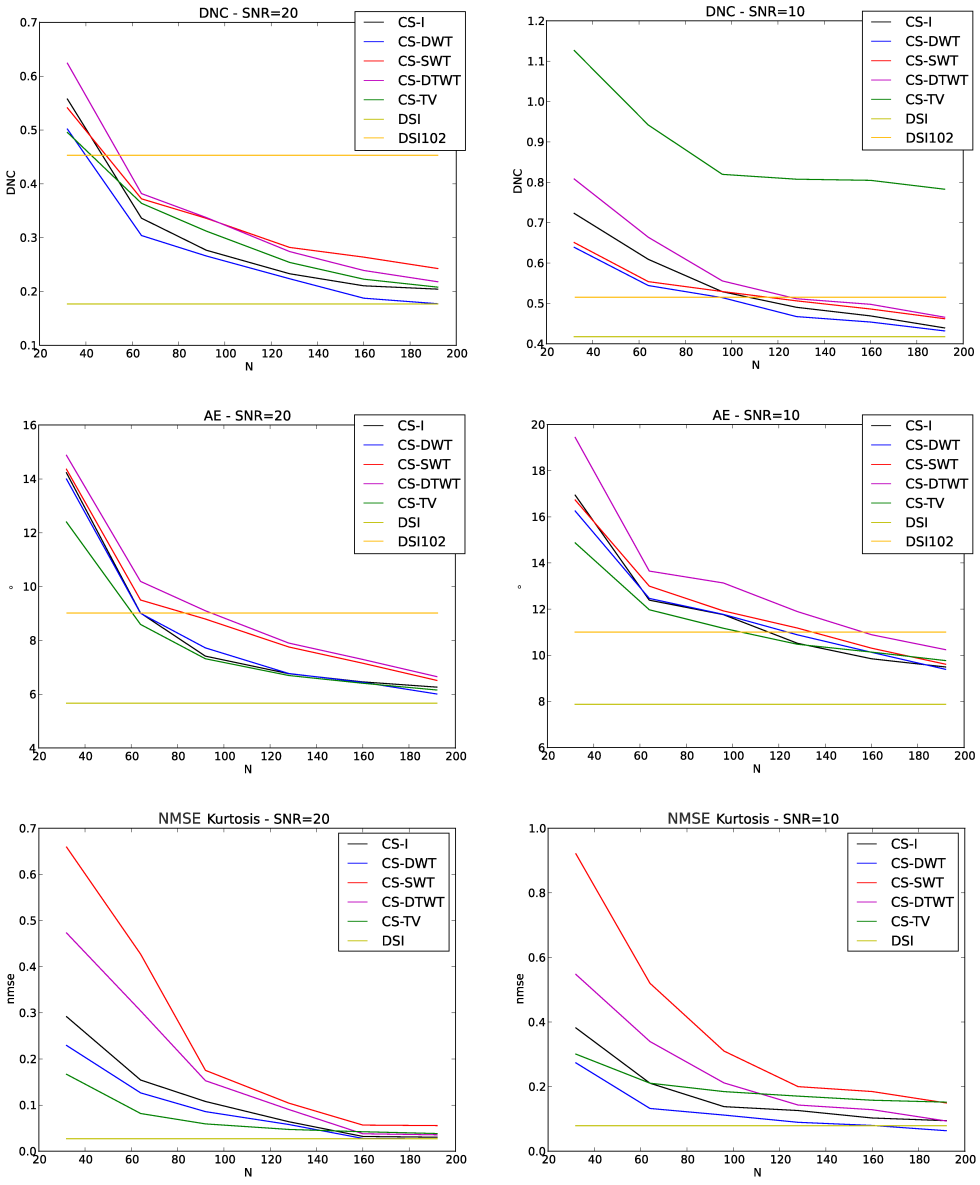


Figure 5.4: Evolution of the DNC, AE and kurtosis NMSE in function of the number of samples for the five reconstruction methods, i.e. CS-I (black curve), CS-DWT (blue curve), CS-SWT (red curve), CS-DTWT (pink curve), TV (green curve). We also show the results of the full DSI reconstruction (yellow curve), i.e. when 257 samples are used, and the DSI102, i.e. when 102 samples are used according to [70] Rician noise is added with SNR=10 (top) and SNR=20 (bottom).

to the case where no transform is applied. This difference is due to the use of an appropriate wavelet basis, i.e. the CDF 9/7 instead of the Haar wavelet basis. Indeed, Haar wavelet is a step function that obviously cannot represent well signals other than piece-wise signals, whereas the CDF 9/7 wavelet adequately describes the quasi-Gaussian attenuation of the EAP [112, 92].

We also remark that the application of the two overcomplete and redundant wavelet transforms does not improve the results, even if these two transforms benefit from a shift-invariance property (the SWT) or an orientation property (the DTWT).

Regarding the five CS approaches, we conclude that the DWT based on the CDF 9/7 mother wavelet remains the most appropriate choice to reconstruct a discrete EAP using the CS technique among the various transform investigated. We also observe that $N \approx 64$ measurements and above lead to satisfactory results in terms of accuracy. We thus perform the CS reconstruction on real data using the DWT as sparse transform.

We also compare the CS-DWT method with the full DSI reconstruction and the DSI102 reconstruction presented in [70]. Overall, it is not surprising to see that the DNC and the AE obtained with the CS-DWT method come closer to the respective values obtained with the full DSI when the number of measurement increases. As for the DSI102, we observe that this method gives nearly the same results as the CS-DWT method with respect to the same number of measurements (i.e. 102 measurements) at a SNR equal to 10. However, the CS-DWT method outperforms the DSI102 when considering a SNR equal to 20 at the same number of measurements.

Validation of the HA-SC. We also validate the results shown in Fig. 5.3 regarding the choice of the HA-SC as our favorite sampling scheme. For this purpose, we repeat the experiments performed in Sec. 5.3.1 but by replacing the inverse Fourier transform by the CS-DWT reconstruction. It aims to demonstrate that the HA-SC is the most efficient and robust sampling scheme when performing a CS reconstruction of the EAP. The mean and the variance of the DNC and AE values are shown in Fig. 5.5, and confirms our finding regarding the higher degree of angular information obtained with the HA-SC.

5.3.3 Experiments on real data

We estimate the EAPs from the human brain data described in Sec. 5.2.4 using the DSI method with the full set of measurements (i.e. $N = 257$ measurements) and the CS-DWT method with $N = 128, 64, 32$ measurements (Note that an additional DW image without any diffusion sensitization is also acquired and used by all these methods). Then, for the ODF estimation (in Fig. 5.6), we choose the integration parameters as $\alpha = 0.4$ and $\beta = 0.8$. Note that β could be higher due to the low level of noise. We also show on Fig. 5.7 an EAP from a chosen voxel in the human brain data. We represent this EAP as two isosurfaces respectively corresponding to radii 0.2 and 0.5 (with 1.0 the distance between the origin and the side of the grid).

The ODFs computed from the EAPs with $N = 128$ and $N = 64$ measurements describe well the underlying fiber structure shown by the ODFs estimated with the full DSI technique (i.e. with $N = 257$ measurements). In the same way, we see in Fig. 5.7 that the main shape of the EAP are well preserved with $N = 128$ and $N = 64$ measurements. With 64 measurements, we come closer to what is obtained using single shell HARDI techniques

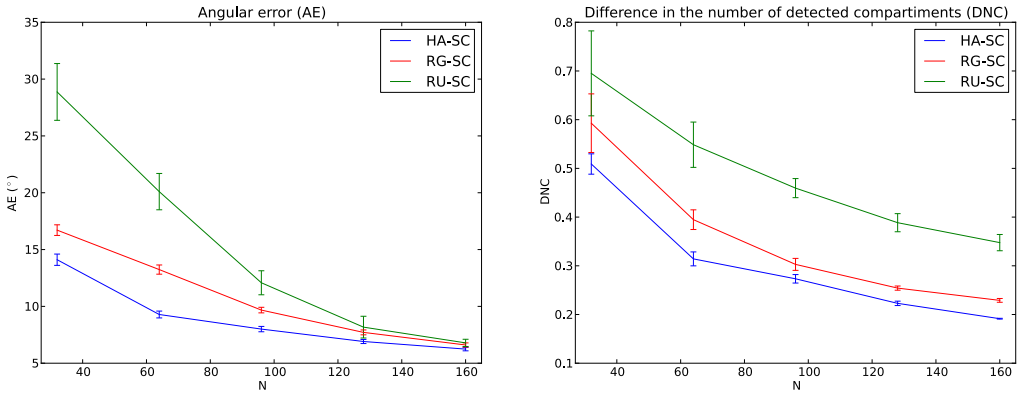


Figure 5.5: Mean and variance (vertical bar) of the angular error (AE) and the difference in the number of fiber compartments (DNC) when using an uniform random sampling scheme (RU-SC), a Gaussian random sampling scheme (RG-SC), a random sampling scheme ensuring a homogeneous angular covering of the q-space (HA-SC) associated with a CS-DWT reconstruction. These results are drawn from 100 trials.

[44, 1]. In addition, the whole EAP is recovered whereas single shell HARDI techniques only estimate the ODF, which is a small part of the information provided by the EAP.

At $N = 32$, only single fiber structures are correctly estimated via the ODFs (see Fig. 5.6). However, we can not rely on the ODFs corresponding to more complex configurations, such that crossing fibers. Moreover, in Fig. 5.7 (right), the two isosurfaces of the EAP do not correctly represent the information exhibited by the full DSI EAP (left of Fig. 5.7)

We also show quantitative results comparing the EAPs estimated with the CS-DWT method and the EAPs estimated with the classical DSI (i.e. with $N = 257$ measurements) on the human brain data. For this purpose, we compute the NMSE between the CS-DWT based EAPs with a number of measurements $N = 32, 48, 64, 80, 96, 112, 128, 144, 160$ and the DSI-based EAPs. Fig. 5.8 shows the resulting values. Based on this figure, we see that the minimum number of measurements to get a reliable EAP reconstruction is approximately 64.

5.4 Discussion

CS offers an efficient way to accelerate the DSI acquisition, as has been demonstrated by several recent papers [15, 82, 92, 112]. However, these works were conducted independently and lack a thorough comparison. Moreover some improvements can be made to enhance these preliminary results. In this chapter, we described a way of performing an accurate and robust CS reconstruction of a discrete EAP. In particular, we studied two important points, namely the choice of an efficient sparse representation of the EAP and the choice of a robust sampling protocol.

Five EAP representations were considered in the CS reconstruction, i.e. the canonical basis (CS-I method), the discrete wavelet transform (CS-DWT method), the discrete sta-

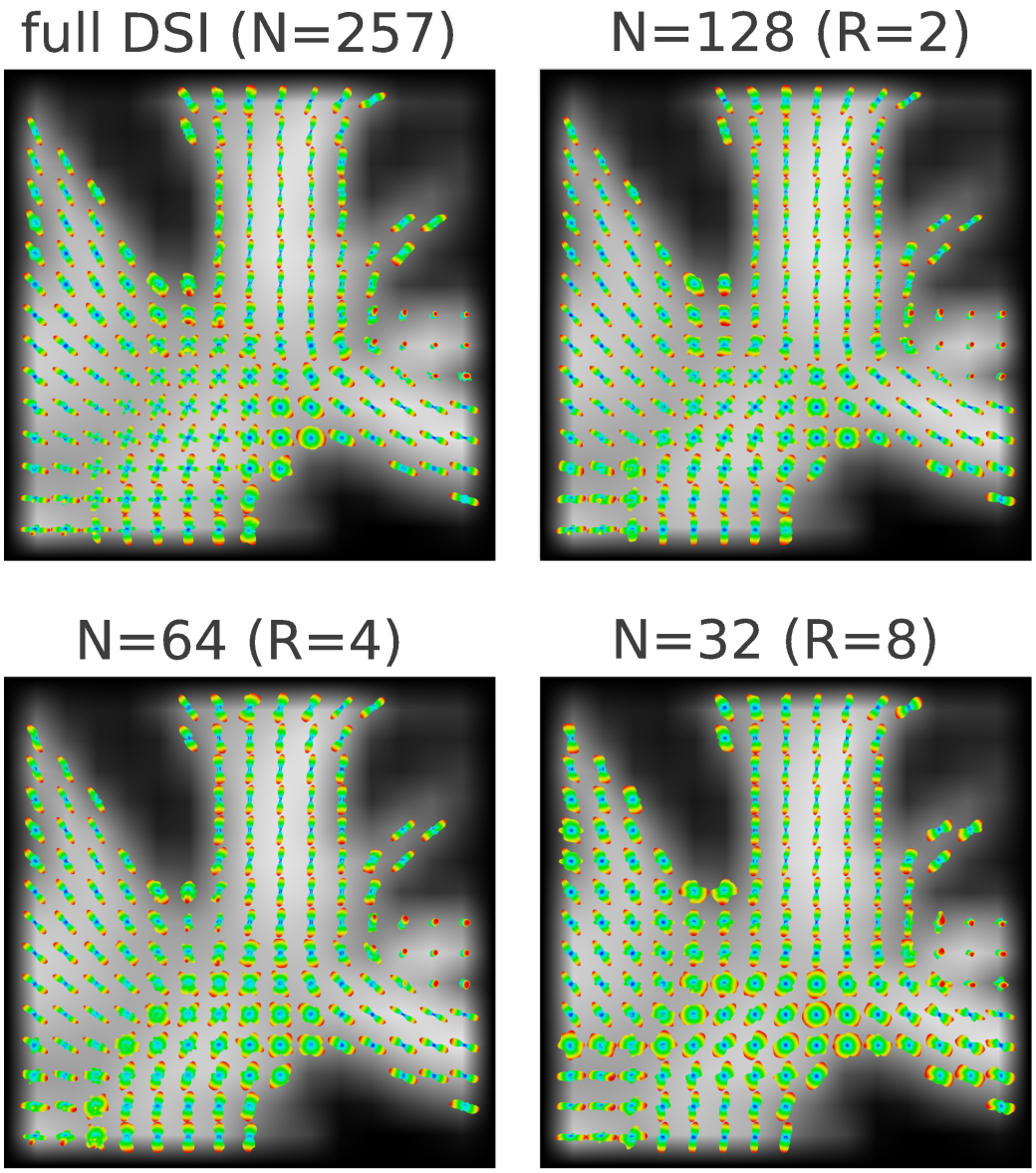


Figure 5.6: ODFs estimated from a humain brain data, together with the extracted maxima, using the DSI method (top) and the CS-DWT method with a number of samples $N = 128, 64, 32$. We compute the pair (DNC,AE) between the CS-DWT and the full DSI estimation of the ODFs : (DNC=0.1555, AE=4.7364) for $N = 128$, (DNC=0.2, AE=6.3011) for $N = 64$, (DNC=0.3111, AE=10.2349) for $N = 32$.

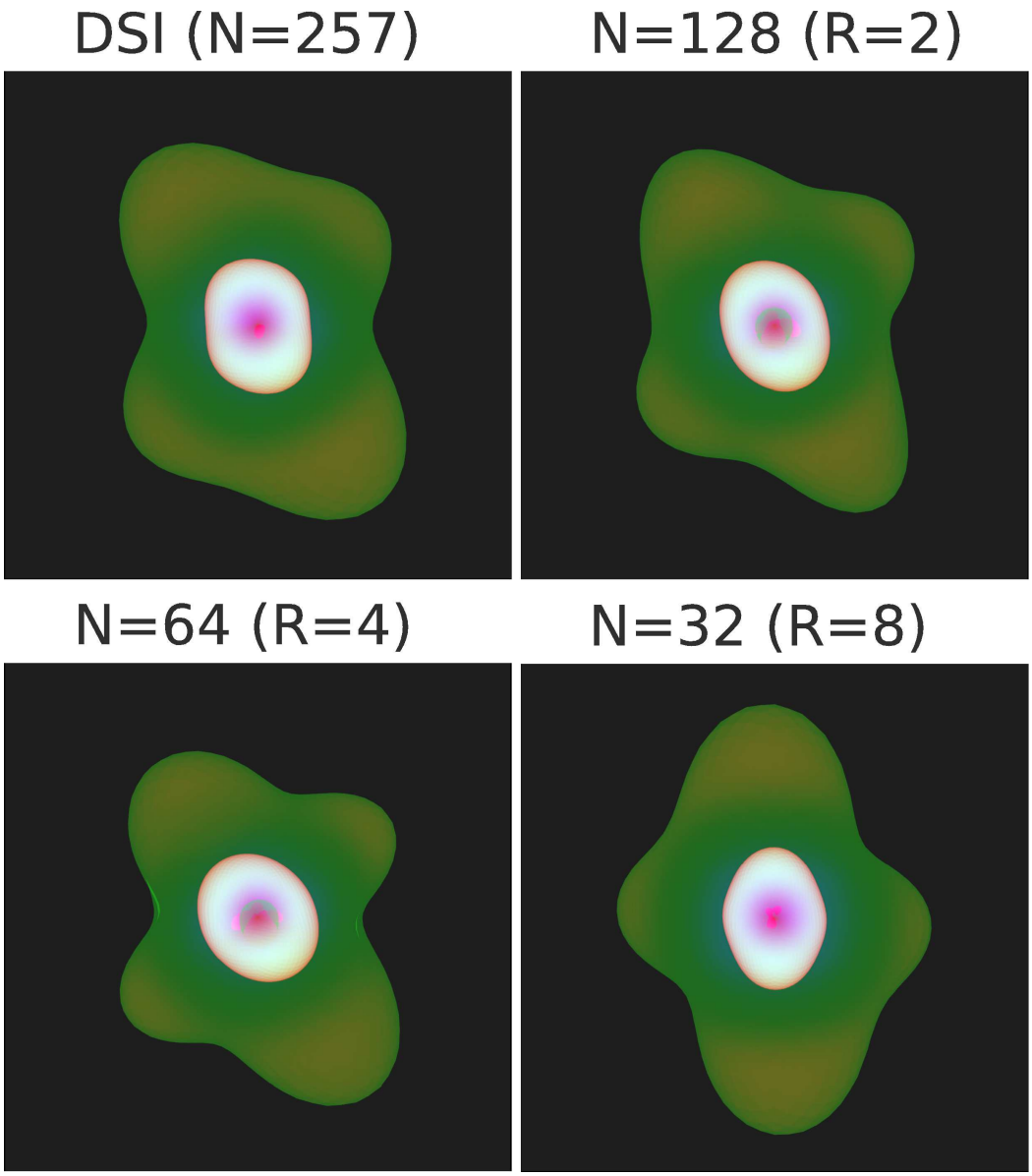


Figure 5.7: EAP estimated from a voxel in humain brain data, using the DSI method and the CS-DWT method with a number of samples $N = 128, 64, 32$.

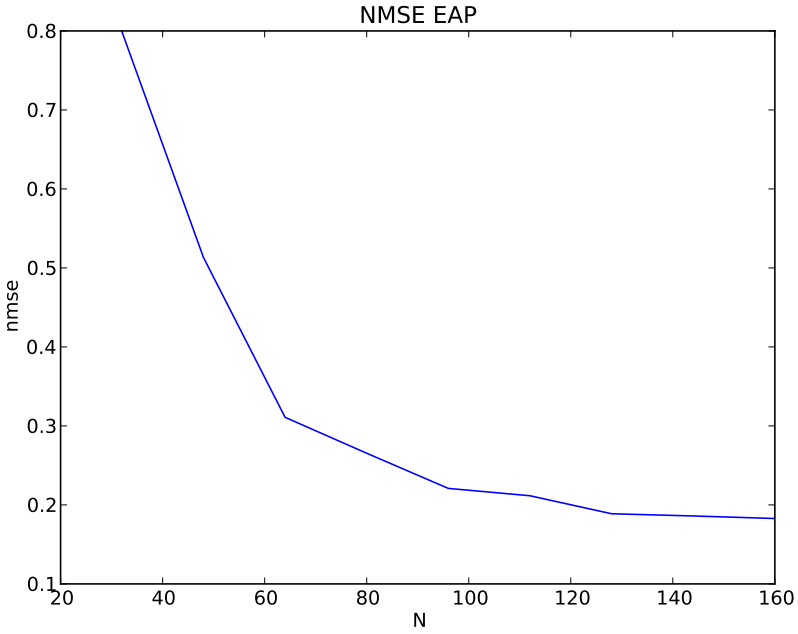


Figure 5.8: NMSE bewtween the CS-DWT based EAPs with a number of measurements N=32, 48, 64, 80, 96, 112, 128, 144, 160 and the DSI-based EAPs.

tionary wavelet transform (CS-DWT method), the real dual tree discrete wavelet transform (CS-DTWT method), the gradient transform (CS-TV method). Furthermore, we studied and compared three sampling protocols, i.e. an uniform random sampling scheme (RU-SC), a Gaussian random sampling scheme (RG-SC) and a random sampling scheme, which ensures a homogeneous angular covering of the q-space (HA-SC). We also compare the CS approaches with the full DSI reconstruction with 257 samples, and a half sphere under-sampled DSI with 102 samples All the EAP representations and sampling protocols were studied and compared through synthetic and human brain data experiments.

From the results, one transform performs best: the DWT. The CS-DWT method leads to accurate estimation of the EAP and some of its derived features, such as the ODF and the kurtosis. However, one has to be careful regarding the mother wavelet used in the DWT. In particular, our recent studies [112, 92], observed that the CDF 9-7 was the most appropriate wavelet to sparsely represent the EAP. We also found a robust and efficient sampling protocol exploiting the fact that the angular coverage of the q-space should be homogeneous. From these studies, we finally found that the combination of the DWT with the HA-SC was the best way to solve the CS-DSI problem with approximately 64 measurements.

We have seen that the orthonormal property of the DWT was found to be more important than the particular and attractive property of the SWT (shift invariance) and the DTWT (directional property). Future works could be focused on an orthonormal and fast transform, which allows additional properties such as shift-invariance or directionality. As far as we know, there are no such transforms in the literature. Thus, at the moment, the DWT remains the most appropriate choice of sparse transform for the CS-DSI problem. Moreover, the CS-DWT method leads to more accurate angular information than the DSI102 reconstruction considering the same number of measurements, especially at a SNR equal to 20.

5.5 Conclusion

We have described a way of performing an accurate and robust Compressive Sensing recovery, which allows the reconstruction of a discrete EAP with approximately 64 measurements and enables the acceleration of the DSI acquisition, without losing much information concerning the diffusion phenomenon. Extensive experiments on synthetic and human brain data demonstrate that both the directional information and the kurtosis are preserved, in addition to the EAP.

In particular, we found an appropriate sparse transform (the CDF 9-7 based DWT) and an efficient and robust sampling protocol, which ensures an homogeneous angular coverage of the q-space.

This chapter was focused in the CS reconstruction of a discrete EAP from a diffusion signal acquired in a Cartesian grid. The sampling protocol is fixed and require measurements corresponding to large b-values. In the next chapter, we present a CS reconstruction of a continuous and analytical signal which, among other advantages, allows more freedom in the acquisition process.

Chapter 6

Continuous diffusion signal, EAP and ODF estimation via Compressive Sensing in diffusion MRI

Contents

6.1 Motivations 92
6.2 Bases for Diffusion Signal Modeling 94
6.2.1 Diffusion signal modeling 94
6.2.2 EAP and ODF modeling 95
6.3 CS properties 97
6.3.1 On the incoherence property of CS bases 97
6.3.2 On the sparsity of CS bases 99
6.3.3 CS recovery 101
6.4 Experimental results 102
6.4.1 Sampling protocol 104
6.4.2 Synthetic and noisy reconstruction 108
6.4.3 Real human brain data 112
6.4.4 Phantom data 113
6.5 Conclusion 116
6.6 Appendix 117
6.6.1 Analytical ODF solution when the signal is modeled in SPF dual basis 117
6.6.2 Analytical ODF solution when the signal is modeled in SHORE basis 118

Overview

In this chapter, we exploit the ability of Compressive Sensing (CS) in recovering the whole 3D Diffusion MRI (dMRI) signal from a limited number of samples while efficiently recovering important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Contrary to the previous chapter, we recover a continuous and analytical signal. Some attempts to estimate continuous diffusion signals via CS have been done recently. However, this was mainly an experimental insight of CS capabilities in dMRI and the CS theory has not been fully exploited. In this work, we also propose to study the impact of the sparsity, the incoherence and the RIP property on the reconstruction of diffusion signals. We show that an efficient use of the CS theory enables the drastic reducing of the number of measurements commonly used in dMRI acquisitions. Only 20-30 measurements, optimally spread among several b-value shells, are shown to be necessary, which is less than previous attempts to recover the diffusion signal using CS. This opens an attractive perspective for measuring the diffusion signals in white matter within a reduced acquisition time and shows that CS holds great promise and opens new and exciting perspectives in diffusion MRI (dMRI). This contribution has been published in [95].

6.1 Motivations

Many measurements and a long acquisition time are necessary to obtain high-resolution EAP. Therefore, it's clear that there is a strong need for new techniques to estimate the whole EAP with fewer measurements. In the previous chapter, we presented a method to accelerate the DSI acquisition. In this chapter, we focus our attention on Analytical and Continuous Modeling (ACM) of the EAP. We can also refer these methods as multiple shells HARDI techniques. ACM methods have been used in [100, 5, 99, 34, 45, 135, 63]. They consist in acquiring the signal following multiple shells schemes and then, modeling it with an adequate basis. These techniques aim to catch both radial and angular information about the water diffusion process. However, an increase of the number of measurements is expected over methods as DTI or QBI. An important problem is to accurately estimate the diffusion signal and the underlying EAP with a small number of samples. A first answer has been given while using suitable bases as the Spherical Polar Fourier (SPF) basis [5] , the SPF dual (SPFdual) basis [89], the Solid Harmonic (SoH) basis [45] or the SHORE basis [99]. We give a complementary solution by using a new acquisition and reconstruction technique called Compressive Sensing (CS).

CS aims to accurately reconstruct signals from under sampled measurements (details on the CS theory is given in **chapter 4**). We saw in **chapter 4** that the application of CS in diffusion MRI can be separated in several categories. In **chapter 5**, a discrete CS based reconstruction has been used to accelerate the DSI technique. In this chapter, we present a continuous CS recovering. Continuous CS recovering consists in modeling a signal with a continuous framework from few measurements via a CS reconstruction. A continuous signal modeling is advantageous because it is is not acquisition dependent and enables data interpolation and extrapolation. Some works have been published toward this [96, 109, 36, 89, 122]. In [96, 122], the authors work with measurements taken at same radii and only estimate the ODF. [109] generalizes the single shell spherical ridgelets basis of [96]

to a multiple shells framework for a sparse and continuous representation of the diffusion signal. In [109], a total number of 64 measurements are used to estimate the diffusion signal well. However, it does not provide any analytical formula to estimate diffusion features. [36, 89] consider a CS reconstruction combined with a continuous representation of the diffusion signal and available closed formulae to estimate the EAP and the ODF. [89] is about CS recovering in SPFdual basis with 80 measurements and [36] is about CS recovering in SPF basis from a minimum number of 60 measurements. These two papers give a first experimental insight of CS capabilities in dMRI, where analytical formulae are available to estimate the EAP and the ODF.

It is also worthwhile to note that very recent works started to handle the learning of dictionaries from a training data set [137, 84, 58, 15]. These techniques lead to very sparse representations of diffusion signals and are worth to be minutely investigated. The analysis done in this chapter considers predefined sets of functions that form orthonormal bases commonly used in the dMRI field. The question of dictionary learning is addressed in **chapter 7**.

In this chapter, we investigate the Compressive Sensing technique in order to accurately and continuously estimate the full 3D diffusion phenomenon as well as some of its features with a very small number of samples. More precisely, we show that only 20/30 measurements are necessary to well estimate the diffusion signal. It is nearly three times less than previous studies encountered in [89, 36, 109]. This significant improvement over these previous works is due to a correct use and consideration of every point of the CS theory. Then, we demonstrate that it is worth using CS recovery, when CS requirements are fulfilled and we also demonstrate how to take advantage of this technique. Before starting with the central point of this chapter, i.e. the CS technique, we describe, in the first section, four bases used to model the diffusion signal. Our approach considers common and continuous representations of the diffusion signal, which enable to obtain various diffusion features such the EAP and the ODF (as in [89, 36]). Then, we give in the second part (Sec. 6.3) a description of the CS properties, which complete Sec. 4.3.2. In Sec. 6.3.1 and 6.3.2 respectively, we study the incoherence and sparse properties of the bases described in section 1. Sec. 6.3.3 describes the CS recovery. We also, handle in Sec. 6.3.3, the acquisition point by describing some theoretical tools to validate sampling protocols, i.e a partial evaluation of the Restricted Isometry Property (RIP). These points are studied both in a theoretical and experimental ways. In the last part, we present some experimental results confronting CS recovery and state of the art recovery. In this experimental part, we begin with synthetic data and focus our attention on several points : 1) the sampling protocol (Sec. 6.4.1), where a powerful technique is described to build robust sampling schemes and 2) the quality of reconstruction on noisy synthetic data (Sec. 6.4.2). Especially in Sec. 6.4.2, we demonstrate the efficiency of CS recovery in reconstructing the diffusion signal and the ODF. These synthetic experiments also enable us to compare our CS-based EAP recovery with the EAP obtained via the DSI technique. Then, we give in Sec. 6.4.3 and 6.4.4 some results, respectively on real monkey brain data and phantom data.

From these experiments, we finally show that CS enables the accurate handling of the whole diffusion process with a lesser number of samples than state-of-the-art methods ($\sim$ 20/30 measurements), while modeling the diffusion signal in one of the bases described in the following.

6.2 Bases for Diffusion Signal Modeling

In this section we describe four bases used to model the diffusion signal, known as

- The Spherical Polar Fourier (SPF) basis [5, 34]
- The SPF dual (SPFdual) basis [89] (6.6.1)
- The Solid Harmonic (SoH) basis, which is part of the DPI method [45]
- The SHORE basis [99, 35] (6.6.2)

We consider these bases because they are the most commonly used in dMRI for a continuous modeling of the diffusion signal. Furthermore, closed formulae for some diffusion features, such that the EAP and the ODF, have been analytically derived in [45, 34, 33, 35] and (6.6.1, 6.6.2).

Our presentation is structured as follows : We first give the analytical formulation for the bases functions (Sec. 6.2.1) and, then, give the closed formulae for the EAP and ODF (Sec. 6.2.2).

Notations : Before exploring this topic any further, we need to set some notations. We call $E(\mathbf{q}) = S(\mathbf{q})/S_0$ the normalized diffusion signal with $S(\mathbf{q})$ the diffusion signal acquired at $\mathbf{q}$ in the q-space and S_0 the signal without any gradient applied. $\mathbf{q}$ is a 3D-vector, which can be decomposed in $\mathbf{q} = q\mathbf{u}$,where $\mathbf{u}$ is a 3D unit vector. $P(\mathbf{R})$ is the EAP at the 3D space location $\mathbf{R} = R\mathbf{r}$, with $\mathbf{r}$ is a 3D unit vector.

6.2.1 Diffusion signal modeling

We represent $E(q\mathbf{u})$ as a truncated linear combination of $\Phi_{n\ell m}(q\mathbf{u})$ with n the radial order, ℓ and m the angular order and degree,

$$E(q\mathbf{u}) = \sum_{n=0}^{N} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_{n\ell m} \Phi_{n\ell m}(q\mathbf{u}), \tag{6.1}$$

where the $c_{n\ell m} = \langle E, \Phi_{n\ell m} \rangle$ are the transform coefficients.

Diffusion signal modeling in the SPF basis : The orthonormal Spherical Polar Fourier basis was used by [5], in order to describe the diffusion signal E. [5] reports that the SPF basis is appropriate for sparsely representing multiple configurations of the water diffusion including isotropy and crossing fibers aspects as well as the multiple compartments profile. [5] describes the SPF basis function $\Phi^{(\mathrm{SPF})}_{n\ell m}(q\mathbf{u})$ as

$$\Phi^{(\mathrm{SPF})}_{n\ell m}(q\mathbf{u}) = \left[\frac{2n!}{\zeta^{3/2}\Gamma(n+3/2)} \right]^{1/2} exp(\frac{-q^2}{2\zeta}) L_n^{1/2}(\frac{q^2}{\zeta}) Y_\ell^m(\mathbf{u}),, \tag{6.2}$$

where ζ is a scale factor, $L_n^{(1/2)}$ is the generalized Laguerre polynomial $L_n^{(\alpha)}$, of order n with $\alpha = 1/2$, and $Y_\ell^m(\mathbf{u})$ is the Spherical Harmonic (SH) function of order ℓ and degree m. Note that we use the real and symmetric version of the Spherical Harmonic basis, i.e. we consider only the SH of even degree.

Diffusion signal modeling in the SPFdual basis : Here, the SPF basis is not used to model the diffusion signal but the EAP. Then, [89] derives a dual basis to model E.

This formulation was first used in a reconstruction in [89]. [89] expresses the SPFdual basis function $\Phi_{n\ell m}^{(\text{SPFdual})}$ as

$$\Phi_{n\ell m}^{(\text{SPFdual})}(q\mathbf{u}) = 4(-1)^{\ell/2}\frac{\zeta^{\ell/2+3/2}\pi^{\ell+3/2}q^{\ell}}{\Gamma(l+3/2)}\gamma(\zeta,q)Y_{\ell}^{m}(\mathbf{u}), \tag{6.3}$$

$$\text{with } \gamma(\zeta,q) = \left[\frac{2^{\ell}n!}{\zeta^{\frac{3}{2}}\Gamma(n+\frac{3}{2})}\right]^{\frac{1}{2}}\sum_{k=0}^{n}\frac{(-1)^k}{k!}\binom{n+\frac{1}{2}}{n-k}2^k \Gamma\left(\frac{\ell}{2}+k+\frac{3}{2}\right){}_1F_1\left(\frac{2k+l+3}{2}, l+\frac{3}{2}, -2(\pi q)^2\zeta\right),$$

where ${}_1F_1$ is the confluent hypergeometric function and Γ the Gamma function. Note that $\Phi_{n\ell m}^{(\text{SPFdual})}$ and $\Phi_{n\ell m}^{(\text{SPF})}$ (in (6.2)) are related by a Fourier transform.

Diffusion signal modeling in the SHORE basis : The SHORE basis has been introduced by [99] but the basis was only orthogonal. [34] proposed a new formulation where the basis functions have a l_2 norm equal to one. This orthonormal basis is a generalization of the SPF basis and $\Phi_{n\ell m}^{(\text{SHORE})}$ is expressed as,

$$\Phi_{n\ell m}^{(\text{SHORE})}(q\mathbf{u}) = \left[\frac{2(n-\ell)!}{\zeta^{3/2}\Gamma(n+3/2)}\right]^{1/2}\left(\frac{q^2}{\zeta}\right)^{l/2}\exp\left(\frac{-q^2}{2\zeta}\right)L_{n-\ell}^{l+1/2}\left(\frac{q^2}{\zeta}\right)Y_{\ell}^{m}(\mathbf{u}). \tag{6.4}$$

Note that, because the Laguerre polynomial order ($n-\ell$ in (6.4)) must be positive, the angular order ℓ depends on the radial order n, such that $n \geq \ell$. Then, the angular order ℓ is bounded by n.

Diffusion signal modeling in the SoH basis : [45] considers the total 3D solution to the Laplace equation in spherical coordinates, i.e. the angular and radial part. It results in the Solid Harmonic (SoH) basis. In this framework, the diffusion E is approximated by two series of SH coefficients $c_{\ell m}$ and $d_{\ell m}$, i.e. the radial order is fixed to 2. We have

$$\begin{aligned} E(q\mathbf{u}) &= \sum_{\ell=0}^{L}\sum_{m=-\ell}^{\ell}\underbrace{\left[\frac{c_{\ell m}}{q^{l+1}} + d_{\ell m}q^{l}\right]}_{radial part}Y_{\ell}^{m}(\mathbf{u}) \\ &= \sum_{\ell=0}^{L}\sum_{m=-\ell}^{\ell}\Phi_{\ell m}^{(\text{SoH})}(q\mathbf{u}). \end{aligned} \tag{6.5}$$

6.2.2 EAP and ODF modeling

As described in (3.9), the EAP P and the diffusion signal E are related by a Fourier transform. Following this relation, we can obtain a closed formula to estimate P from E as follows:

$$P(R\mathbf{r}) = \sum_{n=0}^{N} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_{n\ell m} \Psi_{n\ell m}(R\mathbf{r}). \tag{6.6}$$

where the $c_{n\ell m}$ are the transform coefficients of E in (6.1), and the the $\Psi_{n\ell m}$ are the functions given in Table 6.1.

Moreover, defining the ODF Υ as the integration of the EAP over a solid angle as given by (3.21), we can also derive a closed formula for the ODF in terms of real and symmetric Spherical Harmonic (SH) basis functions :

$$\Upsilon(\mathbf{r}) = \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} \upsilon_{\ell m} Y_{\ell}^{m}(\mathbf{r}), \tag{6.7}$$

where $Y_{\ell}^{m}(\mathbf{u})$ is the SH function of order ℓ and degree m and the coefficients $\upsilon_{\ell m}$ are those given in Table 6.2

SPF	$\Phi_{n\ell m}^{(\mathrm{SPFdual})}(R\mathbf{r})$ [34]
SPFdual	$\Phi_{n\ell m}^{(\mathrm{SPF})}(R\mathbf{r})$ [89]
SoH	$\frac{1}{Z}\left[\frac{(-1)^{l/2}}{R^{3/2}}\left[\frac{2^l R^{l-1/2}\pi^{l-1}}{(2l-1)!!} - \frac{J_{l-1/2}(2\pi q_{max}R)}{q_{max}^{l-1/2}}\right] c_{lm} + (-1)^{l/2} q_{max}^{l+3/2} R^{-3/2} J_{l+3/2}(2\pi q_{max}R) d_{lm}\right] Y_{\ell}^{m}(\mathbf{r})$ [45]
SHORE	$(-1)^{n-l/2}\left[\frac{2(4\pi^2\zeta)^{3/2}(n-\ell)!}{\Gamma(n+3/2)}\right]^{1/2}\left(4\pi^2\zeta R^2\right)^{l/2}\exp\left(-2\pi^2\zeta R^2\right) L_{n-\ell}^{l+1/2}\left(4\pi^2\zeta R^2\right) Y_{\ell}^{m}(\mathbf{r})$ [35]

Table 6.1: $\Psi_{n\ell m}(R\mathbf{r})$ coefficients for the EAP closed formulae (see Eq 6.6). For SoH basis : $n!! = (n)(n-2)\cdots(4)(2)$, $J_n(x)$ is the Bessel function of order n, Z a normalization constant and q_{max} the maximum q-value used during the acquisition.

SPF	$\frac{1}{\sqrt{4\pi}}\delta(l)\delta(m) - \frac{1}{8\pi}\sum_{n=1}^{N}\sum_{i=1}^{n}(-1)^i\left[\frac{2n!}{\zeta^{\frac{3}{2}}\Gamma(n+\frac{3}{2})}\right]^{\frac{1}{2}}\binom{n+\frac{1}{2}}{n-i}\frac{2^i}{i!}P_\ell(0)(-\ell)(\ell+1)c_{n\ell m}$ [33]
SPFdual	$\sum_{n=0}^{N} 2\zeta^{3/4}(-1)^n c_{nlm}\left[\frac{\Gamma(n+3/2)}{n!}\right]^{1/2}$ (6.6.1)
SoH	$(-1)^{l/2}(\frac{2^l\pi^{l-1}}{(2l-1)!!}\left(\frac{R_{max}^{l+1}}{l+1}\right)c_{\ell m} - \frac{(l-1)!!}{2\pi q_{max}^{l+1}2^{l/2}(l/2-1)!}c_{\ell m} + \frac{1}{2\pi}\frac{(l+1)!!}{2^{l/2}(l/2)!}d_{\ell m})$ [45]
SHORE	$\sum_{n=0}^{N} c_{n\ell m}\frac{(-1)^{n-l/2}}{2(4\pi^2\zeta)^{3/2}}\left[\frac{2(4\pi^2\zeta)^{3/2}(n-l)!}{\Gamma(n+3/2)}\right]^{1/2}\frac{\Gamma(\ell/2+3/2)\Gamma(3/2+n)}{\Gamma(l+3/2)(n-l)!}\left(\frac{1}{2}\right)^{-\ell/2-3/2}{}_2F_1(-n+l,l/2+3/2;l+3/2;2)$ (6.6

Table 6.2: $\upsilon_{\ell m}$ coefficients for the ODF closed formulae (see (6.7)). For SPF basis: P_l is the Legendre polynomial of order l. For SoH basis: ${}_2F_1$ is the Gauss hypergeometric function. For SoH basis : $n!! = (n)(n-2)\cdots(4)(2)$, $J_n(x)$ is the Bessel function of order n, Z a normalization constant and q_{max} the maximum q-value used during the acquisition.

6.3 CS properties

In this section, we emphasize the Sec. 4.3.2 and describe why some important conditions described in the classical CS are still required to obtain accurate and robust reconstruction of continuous dMRI signals, namely the sparsity, the incoherence and the RIP. We also highlight theoretical and experimental comparisons between the bases introduced in Sec. 6.2. We choose these four bases mainly because : 1- They are the most commonly used in the dMRI field; 2- They enable a continuous representation of the diffusion signal; 3- They provide analytical formulae for estimating the EAP and the ODF. Note that this work is by no means a theoretical CS framework for continuous signal. We only show that some results obtained in the classical and discrete CS theory can be extended to the reconstruction of continuous signal.

We start by a reminder on previous notations to facilitate the understanding of this section. We called $\mathbf{A} \in \mathbb{R}^{m \times n}$ the sensing matrix and we try to recover a signal $\mathbf{x} \in \mathbb{R}^n$ from a measurement vector $\mathbf{y} \in \mathbb{R}^m$, such that $\mathbf{y} = \mathbf{A}\mathbf{x}$. In Sec. 6.3.1, we also introduced $\bar{\mathbf{A}} \in \mathbb{R}^{n \times n}$, which is an "non-undersampled" version of $\mathbf{A}$. In this chapter $\bar{\mathbf{A}}$ is an orthonormal basis matrix, i.e., a matrix whose columns form an orthonormal basis. $\bar{\mathbf{A}}$, $\mathbf{y}$ and $\mathbf{x}$ are expressed as,

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{6.8}$$

6.3.1 On the incoherence property of CS bases

Because $\mathbf{x}$ is S-sparse, we need $\mathbf{A}$ to admit a certain amount of incoherence to guarantee its recovery. In this section, we focus our attention on this incoherence property.

In order to understand the importance of an incoherent system, let us see an example. Suppose $\mathbf{x} \in \mathcal{R}^n$ is S-sparse with $S = 1$. The non zero coefficients is the ℓ^{th} with $\ell \in [1, n]$. Then

$$\begin{pmatrix} y_1 \\ . \\ . \\ . \\ y_M \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ . & & . \\ . & & . \\ . & & . \\ a_{M1} & \cdots & a_{Mn} \end{pmatrix} \begin{pmatrix} 0 \\ . \\ x_\ell \\ . \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1\ell}x_\ell \\ . \\ . \\ . \\ a_{M\ell}x_\ell \end{pmatrix}. \tag{6.9}$$

Suppose that $a_{k\ell} = 1$ with $k \in [1, M]$. It means that $\mu(\mathbf{A})$ reaches its upper bound. Because the column of $\mathbf{A}$ has an ℓ_2 norm equal to 1, $a_{i\ell} = 0$ for any $i \neq k$. Then

$$\begin{pmatrix} y_1 \\ . \\ y_k \\ . \\ y_M \end{pmatrix} = \begin{pmatrix} 0 \\ . \\ a_{k\ell}x_\ell \\ . \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ . \\ x_\ell \\ . \\ 0 \end{pmatrix}. \tag{6.10}$$

In this case, we are constrained to observe y_k, otherwise $\mathbf{x}$ disappears.

Now, suppose $\mu(\mathbf{A})$ reaches its lower bound. It means $\mathbf{A}$ is a flat matrix with every component equal to $\frac{1}{\sqrt{M}}$. Then

$$\begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix} = \frac{1}{\sqrt{M}} \begin{pmatrix} x_\ell \\ \vdots \\ x_\ell \end{pmatrix}. \tag{6.11}$$

In this case, we can sample at any location and still have information about x_l. The information of $\mathbf{x}$ is spread out on the observation.

Due to the small number of measurements m, it is obvious that a low coherence is crucial to recover $\mathbf{x}$. We prefer the information to be spread out in the observable domain. We cannot take the risk to see $\mathbf{y_\Omega}$ disappears because the measurements are not properly localized (see case where μ reaches its upper bound above). Hence, $\mu(\mathbf{A})$ directly influences the number of acquisitions.

Evaluation of incoherence. (4.7) comes to search for the maximum of all the CS basis elements. Before going any further, we need to fix 1) the scale parameter ζ in the SPF (see (6.2)), SPFdual (see (6.3)) and SHORE (see (6.4)) bases and 2) the bases orders n, ℓ. First, the zeroth order radial function of these bases should have a Gaussian shape dependent on the signal attenuation $E(q) = \exp(-4\pi^2\tau q^2 D)$ with τ the time diffusion constant characteristic of the acquisition process and D the mean diffusivity constant. Considering the SPF and SHORE bases, we obtain the same zeroth order radial function,

$$\begin{aligned} K_{000}(q\mathbf{u}) &= \left[\frac{2}{\zeta^{3/2}\Gamma(3/2)}\right]^{1/2} \exp\left(\frac{-q^2}{2\zeta}\right) L_0^{1/2}\left(\frac{q^2}{\zeta}\right) \\ &= Const. \exp\left(\frac{-q^2}{2\zeta}\right). \end{aligned} \tag{6.12}$$

Then, the scale parameter is given by setting $\exp\left(\frac{-q^2}{2\zeta}\right) = \exp(-4\pi^2\tau q^2 D)$, i.e $\zeta = \frac{1}{8\pi^2\tau D}$. Now considering the SPFdual basis, we get the zeroth order radial function,

$$\begin{aligned} K_{000}(q\mathbf{u}) &= 4\frac{\zeta^{3/2}\pi^{3/2}}{\Gamma(3/2)} \left[\frac{1}{\zeta^{\frac{3}{2}}\Gamma(\frac{3}{2})}\right]^{\frac{1}{2}} \Gamma\left(\frac{3}{2}\right) {}_1F_1\left(\frac{3}{2}, \frac{3}{2}, -2(\pi q)^2\zeta\right) \\ &= Const.{}_1F_1\left(\frac{3}{2}, \frac{3}{2}, -2(\pi q)^2\zeta\right) = Const. \exp(-2(\pi q)^2\zeta). \end{aligned} \tag{6.13}$$

We get $\zeta = 2\tau D$. Secondly, we fix each basis order in such way they enable to model many diffusion profiles. Then, for SPF and SPF dual bases we set an angular order $\ell = 6$ and a radial order $n = 5$, which give 140 atoms for each basis respectively. For the SoH basis, we set an angular order $\ell = 6$, which gives 56 atoms (the radial order is fixed and equal to 2). For SHORE basis, the radial order is $n = 6$ and because ℓ is bounded by n, this gives 72 atoms. On the whole, experiments show that we don't need a higher angular order to accurately model the angular diffusion profiles.

Now, we are able to measure the coherence μ (4.7) for each basis. You should note that

this measure depends directly on the scale parameter ζ in (6.2) and (6.4). ζ itself depends on the diffusion constant τ and the mean diffusivity constant D. We use common values for these parameters: $\tau = 1/4\pi^2$ and D is in range $[0.1 \cdot 10^{-3}, 2 \cdot 10^{-3}]$. Recall that μ is computed on orthonormal bases, so we need to numerically normalize the atom of the SoH basis using the ℓ_2 norm.

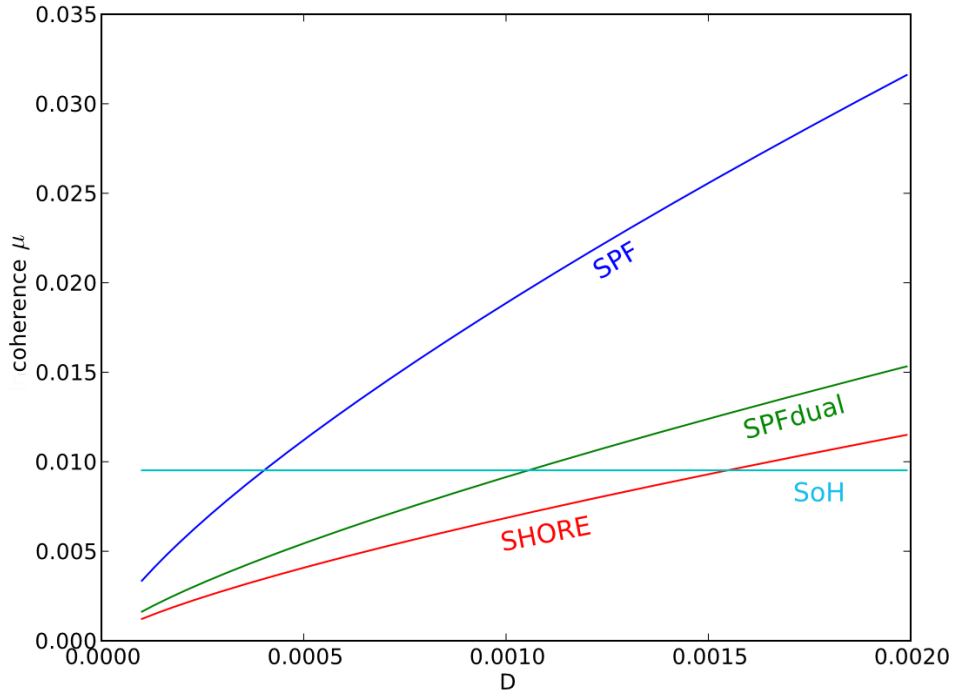


Figure 6.1: Coherence μ of the SPF, SPF dual, SoH and SHORE basis

Fig. 6.1 gives the evaluation of the coherence μ while choosing the CS matrix as the SPF, SPFd, SoH and SHORE basis with the parameters described above. We see that SHORE basis (red curve) and SPFdual basis (green curve) have the lowest coherence values (then the highest incoherence). The coherence of the SPF basis (blue curve) is approximately twice the coherence of the latter bases and the coherence of the SoH basis is constant because it does not depend on D. For the following experiments (in Sec. 6.4), we fix ζ with respect to both the mean diffusivity D (which is peculiar to a given diffusion signal) and the diffusion constant τ (which is used during the acquisition).

6.3.2 On the sparsity of CS bases

As described in chapter 4, we characterize the sparsity of a continuous signal by the number of continuous functions required to accurately model it. In this section, we evaluate the sparsity of the CS bases introduced in Sec. 6.2.

Evaluation of the sparsity: In practice, we have an observation vector $\mathbf{y} \in \mathbb{R}^M$

(In our case the diffusion signal) and we look for its sparse representation $\mathbf{x} \in \mathbb{R}^n$, in a given basis represented by an orthonormal basis matrix $\mathbf{A} \in \mathcal{R}^{M \times n}$, such that $\mathbf{y} = \mathbf{A}\mathbf{x}$. Then, we define the sparsity of $\mathbf{x}$ as the number of basis atoms, associated with the largest coefficient of $\mathbf{x}$, which enables to correctly estimate the observation signal $\mathbf{y}$. By correct representation, we mean the normalized mean square error (NMSE) between the original observation signal $\mathbf{y}$ and the estimated observation signal $\tilde{\mathbf{y}}$ is less or equal to $1 \cdot 10^{-2}$.

$$NMSE = \frac{\|\mathbf{y} - \tilde{\mathbf{y}}\|^2_{\ell 2}}{\|\mathbf{y}\|^2_{\ell 2}} \tag{6.14}$$

Then, we progressively estimate $\tilde{\mathbf{y}}$ with the atoms of the given basis associated with the most important coefficients until we reach a NMSE of $1e^{-2}$. For this purpose, we need to get the coefficients vector $\mathbf{x}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$. The question in this section is not which basis or which reconstruction method gives the best evaluation of the diffusion signal. We only want to fairly compare the sparse properties of the four bases (SPF, SPFdual, SoH and SHORE). To do that, we evaluate $\mathbf{x}$ in the least square sense by considering an overdetermined system, i.e. the number of measurements $M \gg n$. In this case, and only because the system is overdetermined, we do not need any regularization (l_1 or l_2 regularization) for the evaluation of $\mathbf{x}$. It comes to compute $\mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y}$, where $\mathbf{A} \in \mathbb{R}^{n \times M}$ represents one of the four bases. We set $M = 10000$ measurements uniformly spread between $b_{min} = 0$ and $b_{max} = 10000\ \mathrm{s \cdot mm^{-2}}$ (beyond this value the signal is considered to be equal to zero).

Tab. 6.3 gives the average number of atoms necessary to correctly estimate Gaussian-based diffusion signals (see 10.7) in the SPF, SPFdual, SoH and SHORE bases with the order previously defined in section 6.3.1. The number of atoms is averaged for several diffusion profiles (one fiber, 90°-crossing fibers case and 60°-crossing fibers). We have repeated the experiments, while increasing the radial and angular order and similar results were obtained.

	SPF	SPFdual	SoH	SHORE
Number of atoms	12.67	17	34.67	7.33

Table 6.3: Number of atoms necessary to correctly estimate Gaussian-based diffusion signal in the SPF, SPFdual, SoH and SHORE bases. These results are averaged for several diffusion profiles (one fiber, 90°-crossing fibers case and 60°-crossing fibers).

For SPF basis, approximately 12.67 atoms are necessary to correctly reconstruct a diffusion signal with a NMSE less or equal to $1e^{-2}$. SPFdual basis is not as sparse as SPF basis. However, it doesn't mean that the quality of the diffusion signal modeled in the SPFdual basis will be worse than the one modeled in SPF basis. The quality of the estimation depends on other parameters (Recovery method, coherence of the basis). SoH is not sparse according to our criteria. SHORE basis obviously gives the best sparse property in modeling diffusion signal. Nearly half the number of coefficients are necessary as compared to SPF basis. We have repeated the experiments, with a NMSE less or equal to $1e^{-2}$, which lead to the same conclusion..

6.3.3 CS recovery

In this section, we explain the importance of the Restricted Isometry Property (RIP) for an accurate CS recovery and we propose a method to evaluate the sampling scheme via a partial computation of the RIP.

The RIP sets a condition in order to well define the sensing matrix $\mathbf{A_\Omega}$. In other words, it allows us to evaluate the efficiency of our sampling scheme to obtain accurate reconstruction (together with the condition that the signal of interest is sparse). The RIP leads to the fact that all subsets of s columns taken from $\mathbf{A}$ are nearly orthogonals. In order to clarify this point, we present an example for reconstructing a s-sparse signal with $s = 2$, where the non-zero entries are the l^{th} and k^{th}. We have

$$\mathbf{Ax} = \begin{pmatrix} a_0(0) & \cdots & a_n(0) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_0(m) & \cdots & a_n(m) \end{pmatrix} \begin{pmatrix} 0 \\ x_l \\ 0 \\ x_k \\ 0 \end{pmatrix} = \begin{pmatrix} a_l(0)x_l + a_k(0)x_k \\ \cdot \\ \cdot \\ \cdot \\ a_l(m)x_l + a_k(m)x_k \end{pmatrix}. \tag{6.15}$$

If a_l and a_k are orthogonals, i.e $\langle \mathbf{a}_l \cdot \mathbf{a}_k \rangle = 0$, then $\mathbf{a}_l$ and $\mathbf{a}_k$ are linearly independent, i.e $\lambda_l \mathbf{a}_l + \lambda_k \mathbf{a}_k \neq 0, \quad \forall (\lambda_l, \lambda_k) \in \mathcal{R}^2$. If $\langle \mathbf{a}_l \cdot \mathbf{a}_k \rangle \neq 0$, then it may be a sparse vector $\mathbf{x}$ such that $x_l \mathbf{a}_l + x_k \mathbf{a}_k = 0$, i.e. x lies in the nullspace of $\mathbf{A_\Omega}$. To avoid this, we want every subset of columns of $\mathbf{A_\Omega}$ to be orthogonal. By generalizing this example, it is easy to understand that larger are the groups of orthogonal columns in $\mathbf{A_\Omega}$, lower the risk that $\mathbf{x}$ lies in the nullspace of the sensing matrix $\mathbf{A_\Omega}$ Hence, the RIP property ensures that $\mathbf{x}$ lies away from the nullspace of the sensing matrix.

Unfortunately, the columns of the sensing matrix are never truly orthogonal and it would be very time consuming to evaluate the orthogonality of every subset of columns taken from $\mathbf{A_\Omega}$. Instead of doing that, we propose to compute what we call a partial RIP (pRIP) distance. For this purpose, we randomly choose columns from $\mathbf{A_\Omega}$. Let call $\mathcal{A}_{\tau_r} = \{a_{rand(2,n)}\}$ (n is the number of columns of $\mathbf{A_\Omega}$) one subset of columns randomly taken from $\mathbf{A_\Omega}$. The number of column vectors in $\mathcal{A}_{\tau_r}$, i.e card$(\tau_r) = T$, is also set as random such that $1 < T \leq n$. Then, we define the quantity $\delta_{\tau_r} = \frac{1}{T} \sum_{\substack{i,j \subset \tau_r \\ i \neq j}} \langle a_i, a_j \rangle$, which represents the mean of the dot product between every vector in $\mathcal{A}_{\tau_r}$. The pRIP distance is given as $pRIP(\mathbf{A_\Omega}) = \frac{1}{R} \sum_{r=0}^{R} \delta_{\tau_r}$ with R the number of subset to evaluate. It is easy to see that the pRIP can take its value between 0 (when all the columns of $\mathbf{A_\Omega}$ are pairwise orthogonal) and 1 (when all the columns of $\mathbf{A_\Omega}$ are identical). In practice and for our experiments, we have found that the pRIP distance is well computed for $R = 1000$.

Formally, we can state the problem as follows: We look for the subset of samples Ω such that,

$$\Omega = \arg\min_{\Omega} pRIP(\mathbf{A_\Omega}). \tag{6.16}$$

We use this property to evaluate the efficiency and the robustness of our sampling scheme and the underlying sensing matrix.

6.4 Experimental results

The first part of the experiments are concerned with the reconstruction of synthetic data. Synthetic data allows us to evaluate 1) a sampling protocol (Sec. 6.4.1) and 2) the robustness to noise (Sec. 6.4.2). More specifically in Sec. 6.4.2, we present experiments on diffusion signal and ODF reconstruction (Sec. 6.4.2 and Sec. 6.4.2), and compare the EAP based CS recovery with the EAP computed via the DSI technique (Sec. 6.4.2). Then, we show results on a monkey brain (Sec. 6.4.3) data and phantom data (Sec. 6.4.4). In all the experiments, we set the scale parameter ζ and the bases order as in Sec. 6.3.1 , i.e. an angular order $\ell = 6$ and a radial order $n = 5$ for SPF and SPF dual bases, a radial order of $n = 6$ for SHORE basis and an angular order of $\ell = 6$ for SoH basis. ζ changes along with the mean diffusivity constant D, peculiar to a given diffusion signal. For the synthetic experiments we set $D = 0.7e-3$ and ζ is respectively set to 700, $3.5462e-5$, 700 for SPF, SPFdual and SHORE bases (For details on these values, see Sec. 6.3.1). For the experiments on the monkey and phantom data, we will fix ζ with the mean diffusivity constant averaged on voxels in the region of interest.

Before going any further, we give the notations and materials used in the experimental part. We also describe the ℓ_1 minimization problem (called ℓ_1 recovery in the following), which is used in CS recovery. Then, we describe the least square recovery with an ℓ_2 regularization (called ℓ_2 recovery in the following), which is used for comparison.

Notations :

We represent the normalized diffusion signal $E(q\mathbf{u})$ as a truncated linear combination of $\Phi_{n\ell m}(q\mathbf{u})$ with n the radial order, ℓ and m the angular order and degree, i.e.

$$E(q\mathbf{u}) = \sum_{n=0}^{N} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_{n\ell m} \Phi_{n\ell m}(q\mathbf{u}), \tag{6.17}$$

where the $c_{n\ell m} = \langle E, \Phi_{n\ell m} \rangle$ are the transform coefficients. We call n_c the total number of atoms.

Suppose n_q is the number of measurements, $\mathbf{E} \in \mathcal{R}^{n_q}$ a vector representing the diffusion signal, $\mathbf{c} \in \mathcal{R}^{n_c}$ the vector of transform coefficients $c_{n\ell m}$ and $\mathbf{\Phi} \in \mathcal{R}^{n_q \times n_c}$ the matrix constructed with the atoms of the given basis as

$$\mathbf{\Phi} = \begin{pmatrix} \Phi_{000}(q_1\mathbf{u_1}) & \cdots & \Phi_{NLL}(q_1\mathbf{u_1}) \\ \vdots & \ddots & \vdots \\ \Phi_{000}(q_{n_q}\mathbf{u_{n_q}}) & \cdots & \Phi_{NLL}(q_{n_q}\mathbf{u_{n_q}}) \end{pmatrix}, \tag{6.18}$$

We can write equation (6.17) as a linear system of overdetermined equations,

$$\mathbf{E} = \mathbf{\Phi c}. \tag{6.19}$$

The aim is to recover $\mathbf{c}$ given the normalized diffusion signal $\mathbf{E}$ and the matrix $\mathbf{\Phi}$. Then, the EAP and ODF can be analytically computed from $\mathbf{c}$ using the closed forms given in Sec. 6.2.

ℓ_1 recovery

In order to find $\mathbf{c}$, we solve the convex optimization problem, specified in (4.11), i.e. with the above notations,

$$\min_{\mathbf{c}\in\mathcal{R}^{n_c}} \|\mathbf{c}\|_{\ell_1} \text{ subject to } \|\mathbf{E}-\mathbf{\Phi c}\|_{\ell_2} \leq \epsilon. \tag{6.20}$$

For convenience, we replace the constraint with a penalty. It comes to solve

$$\arg\min_{\mathbf{c}\in\mathcal{R}^{n_c}} \|\mathbf{E}-\mathbf{\Phi c}\|_{\ell_2} + \lambda\|\mathbf{c}\|_{\ell_1}, \tag{6.21}$$

where the regularization parameter λ replaces the noise level ϵ in (6.20), in governing the trade-off between the data consistency and its sparsity. We solve the problem by means of a Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [12, 139], an iterative algorithm where each iteration involves a shrinkage step. We chose this algorithm because of its efficiency, its speed and its convergence.

Appropriate choice of the regularization parameter λ is a critical issue in ℓ_1 recovery. The regularization parameter selection in (6.21) depends essentially on the level of noise, the data sparsity and the number of measurements. Thus, we need a technique that adaptively chooses λ , i.e. for each instance of diffusion signal contained in a whole data set. We use cross validation to assess the reconstruction parameters [129]. In particular, we use a K-fold Cross Validation, which consists of splitting the entire data set in K subsets. $K-1$ subsets are used to reconstruct the signal whereas the K^{th} left apart subset enables an estimation of the regularization parameter λ_K via the evaluation of a cross validation distance. This operation is repeated K times by considering the other subsets. Then, we keep an average value, $\lambda = \frac{1}{K}\sum_{k=1}^{K}\lambda_k$. This procedure is computationally extensive but experiments show that it enables a close approximation of the optimal regularization parameter. To avoid a drastic increase of the computational effort, we split our data set in 5 partitions, i.e. $K=5$. Our experiments showed that it is sufficient to well evaluate λ in (6.21). See [129] for details regarding the use of cross validation in CS.

ℓ_2 recovery

Least square recovery with an ℓ_2 regularization is the common method used so far in order to solve a system of overdetermined equations involving the bases in section (6.2). The ℓ_2 recovery problem that is to be solved is

$$\arg\min_{\mathbf{c}\in\mathcal{R}^{n_c}} \|\mathbf{E}-\mathbf{\Phi c}\|_{\ell_2} + \lambda_l\|\mathbf{L}\|_{\ell_2} + \lambda_n\|\mathbf{N}\|_{\ell_2}, \tag{6.22}$$

where $\mathbf{N}\in\mathcal{R}^{n_q\times n_c}$ and $\mathbf{L}\in\mathcal{R}^{n_q\times n_c}$ are two diagonal matrices such that $diag(\mathbf{N}) = n(n+1)$ and $diag(\mathbf{L}) = l(l+1)$. $\mathbf{L}$ is called the Laplace-Beltrami operator. These two matrices penalize the high frequencies of the radial and angular part, respectively. λ_n and λ_l enable to weight the high frequency penalization. The least square solution has a closed form given by [5, 99],

$$\mathbf{c} = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda_l\mathbf{L}^T\mathbf{L} + \lambda_n\mathbf{N}^T\mathbf{N})^{-1}\mathbf{\Phi E}, \tag{6.23}$$

For the SoH basis [45], there is only an angular regularization. Then, the least square solution is given by,

$$\mathbf{c} = (\mathbf{\Phi}^T\mathbf{\Phi} + \lambda_l\mathbf{L}^T\mathbf{L})^{-1}\mathbf{\Phi E}. \tag{6.24}$$

To enable a fair comparison with the ℓ_1 recovery, we apply a variant of cross validation, the Generalized Cross Validation (GCV) algorithm [40], to assess the regularization parameters λ_l and λ_n. The GCV method is based on a K-fold cross validation where K is the number of measurements. Fortunately, we have a closed form expression, which enables the quick estimation of λ_l and λ_n. We simply take λ_l and λ_n as the minimum argument of the GCV function, expressed as

$$\mathrm{GCV}(\lambda; \mathbf{E}) = \frac{\|\mathbf{E} - \hat{\mathbf{E}}_\lambda\|^2}{K - \mathrm{Tr}(\mathbf{S}_\lambda)}, \tag{6.25}$$

which makes this method very efficient. The matrix $\mathbf{S}_\lambda = \mathbf{\Phi}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} + \lambda\mathbf{\Lambda})^{-1}\mathbf{\Phi}^{\mathrm{T}}$ is the smoother matrix, and $\hat{\mathbf{E}}_\lambda = \mathbf{S}_\lambda\mathbf{E}$. Λ is either $\mathbf{L}^T\mathbf{L}$ or $\mathbf{N}^T\mathbf{N}$ depending on which regularization parameter we want to assess (λ_l or λ_n). With the GCV method, it is possible to adapt the regularization parameters to the data.

6.4.1 Sampling protocol

The quality of reconstruction is sensitive to the acquisition scheme [85, 26]. In particular, the random sampling protocol usually used in CS recovery can miss information and does not ensure good reconstruction in every case. Hence, in order to remove the variance of the results due to the random aspect of the sampling scheme, it is necessary to find a robust and efficient way to acquire DW-MRIs. For this purpose, we evaluate and compare several sampling protocols. We begin by presenting some methods to distribute points on one or several spheres. It follows a presentation of the sampling schemes to be evaluated. Then, we give a first appreciation of the quality of each sampling scheme using the partial RIP (pRIP) distance presented in Sec. 6.3.3. The pRIP distance evaluates the capabilities of the sensing matrix $\mathbf{\Phi}$ to give a robust and accurate reconstruction but doesn't consider the signal itself. Hence, we also evaluate experimentally each sampling scheme on the reconstruction of synthetic diffusion data.

Jones algorithm

References [68, 41] give an algorithm to uniformly distribute N points $q_n \in R^3$ on a sphere by considering each point as an antipodal pair of electrical charges. The method involves the minimization of the electrostatic force of repulsion between each couple of charges. The electrostatic repulsion between two points q_i and q_j is given by

$$E(q_i, q_j) = \frac{1}{\|q_i + q_j\|} + \frac{1}{\|q_i - q_j\|} \tag{6.26}$$

For a set of N points, the energy to minimize becomes

$$J_J = \sum_{i \neq j} E(q_i, q_j) \tag{6.27}$$

Reference [38] provides Camino, an Open-Source Diffusion-MRI Reconstruction and Processing software. They include several sets of directions, from N=3 to 150 points, computed by electrostatic energy minimization.

Generalized Jones algorithm

This method is proposed by [29] as a generalization of [68] to multiple shells (MS) acquisition. It enables the distribution of N points $q_n \in R^3$ on K shells of radius r_k. The points from each shell have staggered directions and follow a near-optimal uniform distribution. Another important point in this method is the possibility to balance the proportion α_k of samples between shells. We will take advantage of this feature in order to test out different radial distributions.

Firstly, the method consists in minimizing the electrostatic repulsion between every point for each shell independently, that is

$$E_1 = \sum_k r_k \alpha_k \sum_{i \neq j \ s.t \ \|q_i\| = \|q_j\| = r_k} E(q_i, q_j) \tag{6.28}$$

Then, in order to have staggered directions between shells, [29] introduces a new term that minimizes the electrostatic repulsion of the N points projected on the unit sphere. It comes to minimize

$$E_2 = \sum_{i \neq j} \frac{1}{\left\| \frac{q_i}{\|q_i\|} - \frac{q_j}{\|q_j\|} \right\|} + \frac{1}{\left\| \frac{q_i}{\|q_i\|} + \frac{q_j}{\|q_j\|} \right\|} \tag{6.29}$$

Finally, the energy to minimize is $J_{GJ} = (1-\mu)E_1 + \mu E_2$, where μ is a weighting factor.

4 sampling schemes

We perform our experiments on four MS sampling schemes. [85, 26] showed the advantage of taking staggered directions between shells in order to well approximate the angular profile of the diffusion process. Hence, each sampling scheme will be computed via the Generalized Jones algorithm in order to obtain staggered directions between shells. Here we, evaluate the impact of the radial distribution of the samples. Then, we test four sampling schemes where the number of samples on each shell is proportional to q^γ with $\gamma = 1, 2, 3, 4$.

Results on sampling protocol

We first provide a critical appreciation of the quality of each sampling scheme using the partial RIP (pRIP) distance as described in Sec. 6.3.3. The pRIP distance roughly evaluates the efficiency of a sampling scheme to obtain accurate reconstruction. We compute it with $N = 10$ to 90 for each basis previously presented and for a sampling scheme following radial distributions in q^γ with $\gamma = 1, 2, 3, 4$. Overall, and for the SPF, SPFdual and SHORE bases, we see that the pRIP distance is higher for the schemes whose the radial distribution follow a law in q^2 and q^3 than those in q^0 and q^1. It means the schemes in q^2 and q^3 are more sensitive to noise and lead to less accurate reconstruction. The schemes in q^0 and q^1 have similar pRIP values for different number of samples N. Considering the SoH basis, the isotropy distance does not vary a lot between sampling schemes. It is also quite stable along N. It means all the sampling schemes have similar effect on the accuracy of the reconstruction while using the SoH basis.

Now, we experimentally review the outcome of multiple shells (MS) sampling on the reconstruction of synthetic data. The performance of each sampling scheme is determined on diffusion signal reconstruction while the signal is generated from a multi-tensor model

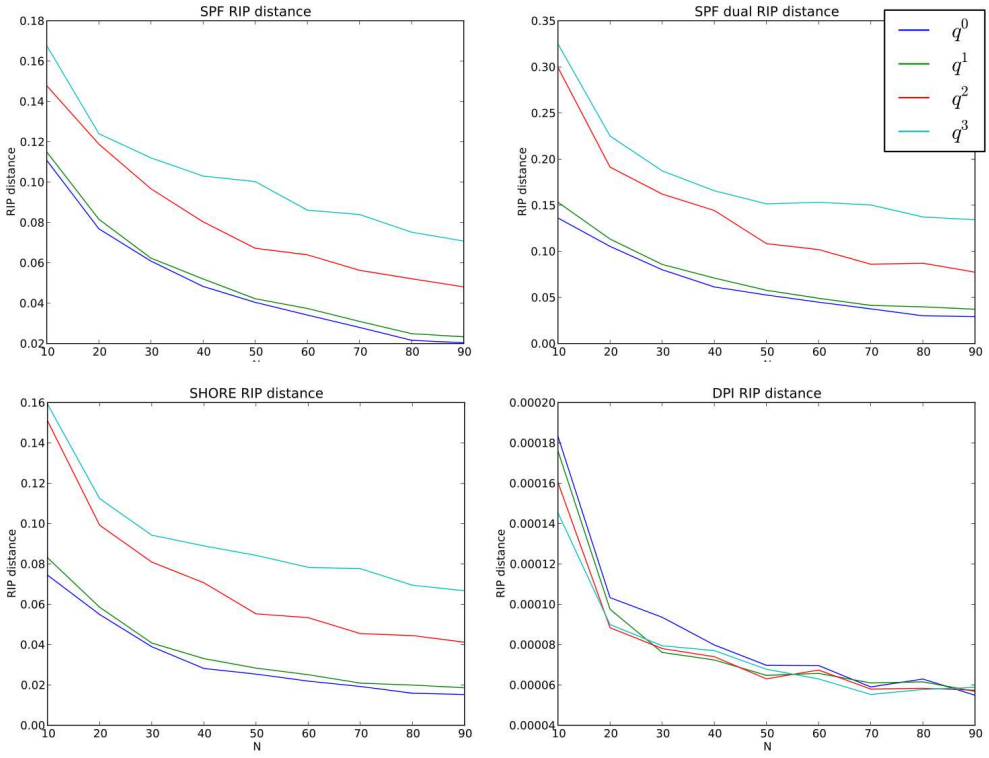


Figure 6.2: pRIP distance evaluated for the SPF, SPFdual, SoH and SHORE bases. We considered four sampling schemes following radial distributions in q^γ with $\gamma = 1, 2, 3, 4$ and compute δ while taking $N = 10$ to 90 measurements.

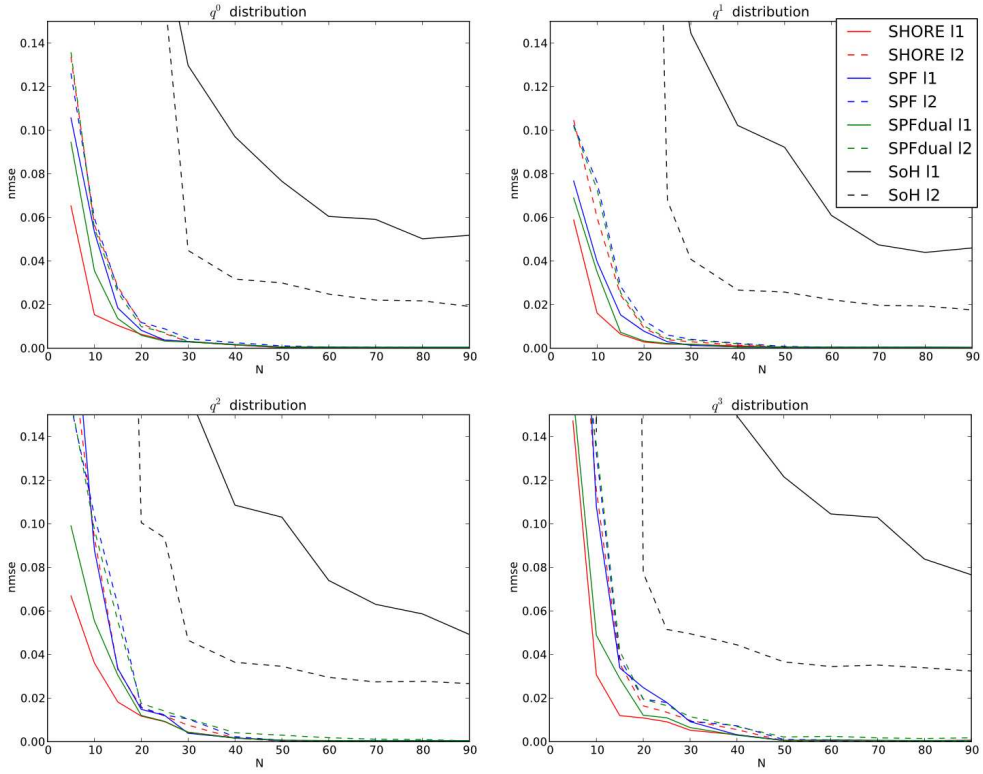


Figure 6.3: NMSE evaluated for the reconstruction of synthetic data using ℓ_1 recovery (plain lines) ℓ_2 recovery (dash lines) for the SPF, SPFdual, SoH and SHORE bases. We considered four sampling schemes following radial distributions in q^γ with $\gamma = 1, 2, 3, 4$ and compute δ while taking $N = 5$ to 90 measurements.

(see 10.7) through three scenarios : One fiber, two 60°-crossing fibers, two 90°-crossing fibers evaluated at b values, $b = [1000, 2000, 3000]\text{s}\cdot\text{mm}^{-2}$. We evaluate the reconstruction for a number of samples $N = 5$ to 90. The quality of the diffusion signal estimation $\tilde{E}$ is given by the normalized mean square error (NMSE) (see (6.14)), evaluated on sample at b-value between 0 and 10000, i.e. different from the measurements used for the reconstruction. We average the NMSE obtained on 1000 independent trials and for each scenario. Because this section aims to evaluate sampling schemes, and not the robustness to noise, we generate a noise-free diffusion signal. The robustness to noise was already evaluated via the computation of the isotropy distance. The results are given in Fig 6.3. In this figure, we compare the ℓ_1 and ℓ_2 reconstruction methods. The curves corresponding to ℓ_1 recovery are in plain lines and the curves corresponding to ℓ_2 recovery are in dash lines. We first consider the SHORE, SPF and SPFdual bases. NMSE values for the schemes following a radial distribution in q^2 and q^3 are in line with our previous results on the pRIP distance, i.e. we get a higher NMSE for these schemes than the schemes in q^0 and q^1. Within the schemes in q^0 and q^1, we see that the NMSE for the one in q^1 remains lower than the NMSE for the one in q^0 while decreasing the number of samples N. These observations remain valid for ℓ_1 and ℓ_2 recovery. Besides, [26] comes to the same conclusions regarding ℓ_2 recovery. Considering the SoH basis we have two remarks. First, the NMSE is quite the same for each radial distribution either for ℓ_2 recovery or ℓ_1 recovery. Next we observe, that ℓ_2 recovery gives lower NMSE values than ℓ_1 recovery using the SoH basis.

To summarize this section, we have seen that a sampling scheme stands out for both ℓ_1 recovery and ℓ_2 recovery. This scheme uses the generalized Jones algorithm presented in Sec. 6.4.1, where the samples follow a radial distribution in q^1. The pRIP distance evaluated for this scheme indicates that it is robust to noise and the experiments on synthetic data show that it gives the lowest NMSE. Hence, we keep this scheme for the rest of the experiments.

6.4.2 Synthetic and noisy reconstruction

This section presents experiments on the reconstruction of synthetic data. We describe results from signal, ODF and EAP reconstructions, followed by a conclusion where we highlight the important points of these experiments.

Diffusion signal recovery

We first present results on diffusion signal recovery. Based on the previous results in Sec. 6.4.1, we choose a sampling scheme with a radial distribution following a law in q^1. Here, we perform experiments on diffusion signal recovery where the data is contaminated with Rician noise. As previously, we compare the ℓ_1 recovery and ℓ_2 recovery. We keep the NMSE criteria to evaluate the reconstruction of diffusion signal from a multi-tensor model (10.7) through three scenarios : One fiber, two 60°-crossing fibers, two 90°-crossing fibers at b values $b = [1000, 2000, 3000]\text{s}\cdot\text{mm}^{-2}$. We evaluate the reconstruction for a number of samples $N = 5$ to 90. We add Rician noise with SNR between 30 for data contaminated with low noise level and 10 for very noisy data. The Rician noise is added to the normalized diffusion signal E in the following way : $E_{noisy} = \sqrt{(E + \epsilon_1)^2 + \epsilon_2^2}$ where $\epsilon_1, \epsilon_2 \sim \mathcal{N}(0, \sigma)$ with $\sigma = \frac{1}{SNR}$. Again, all the results are obtained on 1000 independent trials. Then,

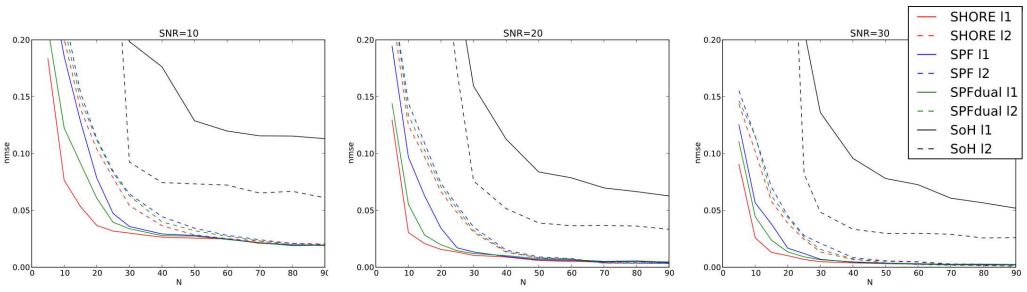


Figure 6.4: NMSE for noisy and synthetic reconstruction with SNR=10 to SNR=30 (left to right). Results for ℓ_1 recovery is represented in plain line and results for ℓ_2 recovery is represented in dash line. We consider the SPF (in blue), SPFdual (in green) , SoH (in black) and SHORE (in red) bases for the reconstruction.

we average the results of one fiber, two 60°-crossing fibers, two 90°-crossing fibers. The resulting NMSE for ℓ_1 recovery (plain line) and ℓ_2 recovery (dash line) are presented in Fig. 6.4.

As expected from the preliminary results in Sec. 6.3, SPF, SPFdual and SHORE based reconstruction broadly give better results for ℓ_1 recovery than ℓ_2 recovery, and inversely for the SoH basis. This is not surprising while looking at the sparsity level of SoH basis (see Tab. 6.3 in Sec. 6.3.2). It gives the worst sparsity level with respect to the others bases. More precisely, for ℓ_2 recovery, we see a global increase of the NMSE below $N \approx 40/50$ samples at any SNR, whereas the NMSE for ℓ_1 recovery globally increases below $N = 30$ samples. For $N = 50$ and higher, the two methods lead to similar NMSE values. These results show that SHORE, SPF and SPFdual bases are appropriate for CS recovery, and enable to reconstruct diffusion signals as accurately as with ℓ_2 recovery with a smaller number of measurements. The SoH basis, however, does not enter in the CS framework and the results show that it is preferable to use the ℓ_2 recovery instead of the ℓ_1 recovery with the SoH basis.

Regarding the ℓ_1 recovery, we give a comparison on the number of samples required to correctly approximate the diffusion signal while considering the SPF, SPFdual and SHORE bases. For the same NMSE, we see that SHORE based ℓ_1 recovery needs less samples than SPF and SPFdual based ℓ_1 recovery, and that SPFdual based ℓ_1 recovery needs less samples than SPF based ℓ_1 recovery at same NMSE (see Fig. 6.4). For instance, to obtain a NMSE equal to 0.03 at $SNR = 30$, we need nearly 10, 16 and 18 measurements respectively for SHORE, SPFdual and SPF based reconstructions. To summarize, SHORE based ℓ_1 recovery gives the best signal estimation in terms of NMSE. This results was expected from the studies regarding the coherence and sparsity of these bases in Sec. 6.3.2 and 6.3.1. To quantify this result, we give the number of measurements required before a too large increase of the NMSE. We notice that approximately 20/30 measurements are needed (depending on the underlying SNR) for a SHORE based ℓ_1 recovery. This leads to approximately three times less number of measurements than in [89, 36, 109].

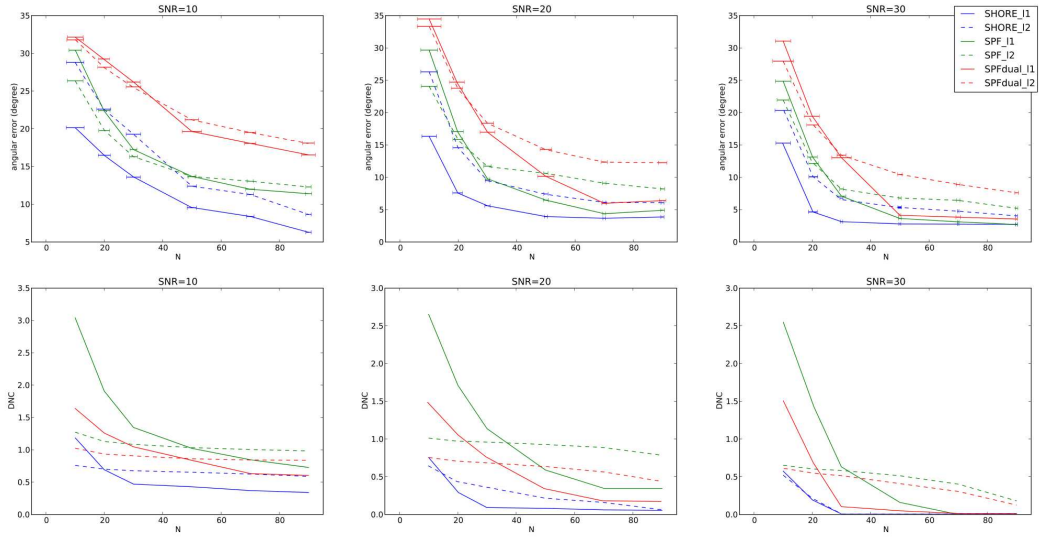


Figure 6.5: Angular error (AE) and difference on the number of compartments (DNC) from noisy and synthetic ODF reconstruction with SNR=10 to SNR=30 (left to right). Results for ℓ_1 recovery are represented in plain line and results for ℓ_2 recovery are represented in dash line. We consider the SPF (in green), SPFdual (in red) and SHORE (in blue) bases for the reconstruction.

Orientation Distribution Function recovery

In this section, we describe results on the angular information provided by the Orientation Distribution Function (ODF) estimated from the SPF, SPFdual and SHORE bases via the closed form presented in Sec. 6.2.2. We set apart the SoH basis due to the bad results concerning this basis in the previous section. We set up the experiments as in the previous section on diffusion signal recovery, except for the number of samples used for the evaluation. Here, we consider $N = 10$ to 90. We evaluate the angular information via two metrics : the angular error (AE) between the directions of the extracted ODF maxima and the ground truth directions, and the corresponding difference on the number of compartments (DNC) i.e. the difference between the number of maxima detected and the true number of maxima. These two metrics are computed for each trial (1000 trials in all) and, then, averaged. We also compute the standard deviation of the AE. Note that the AE is computed between the ground truth direction and the closest extracted ODF maxima until every maxima or ground truth direction are scanned. Hence, the DNC metric shows its importance since it gives information regarding the reliability of the AE. All the results are shown in Fig. 6.5.

Regarding the AE, we observe that the ℓ_1 recovery gives globally more accurate estimation of the directions than the ℓ_2 recovery. One explanation comes from the underlying property of the ℓ_1 recovery, which provides sharper ODF estimation, while the ℓ_2 constraint is known to favor low frequency components and thus provides smoother ODF estimation.

We, now, consider the DNC values of Fig. 6.5 (bottom curves). We observe, when a small number of samples is used, that the ℓ_2 recovery gives a more accurate DNC estimation than the ℓ_1 recovery. This observation is especially true when the SPF and the SPFdual

bases are used to model the signal. Moreover, this phenomenon is emphasized for a small SNR. It may be due to the smoothing consequence of the ℓ_2 constraint, which somehow acts as a signal denoiser. However, this advantage of the ℓ_2 recovery is hampered by the very bad evaluation of the maxima directions extracted from the ODFs. For instance, at SNR=10 and below $N = 30$, the AE exceeds 15 degrees. Then, we can not rely on the ℓ_2 recovery for diffusion directions evaluation while considering a small number of measurements in the reconstruction. To conclude on the DNC, only SHORE based ℓ_1 recovery is able to provide reliable ODF maxima, except for $N = 10$.

Overall, for a minimum of 20/30 measurements and above $SNR = 20$, SHORE based ℓ_1 recovery ensures reasonably low AE and DNC values, especially at SNR=30 where the DNC is zero above $N = 30$ measurements. At SNR=10 and above $N = 30$ measurements, we still have acceptable AE and DNC ($AE < 15°$ and $DNC < 0.5$) with the SHORE based ℓ_1 recovery. Then, we conclude that 20/30 measurements required are enough to obtain reliable and accurate diffusion directions from the estimated ODFs using the SHORE based ℓ_1 recovery (as concluded in Sec. 6.4.2). Finally, regarding the standard deviation of the error, we observe, as expected, an increase of these values for decreasing SNRs.

Comparison with DSI

Here, we compare the EAP estimated via a SHORE based ℓ_1 recovery and via the Diffusion Spectrum Imaging (DSI) technique [130].

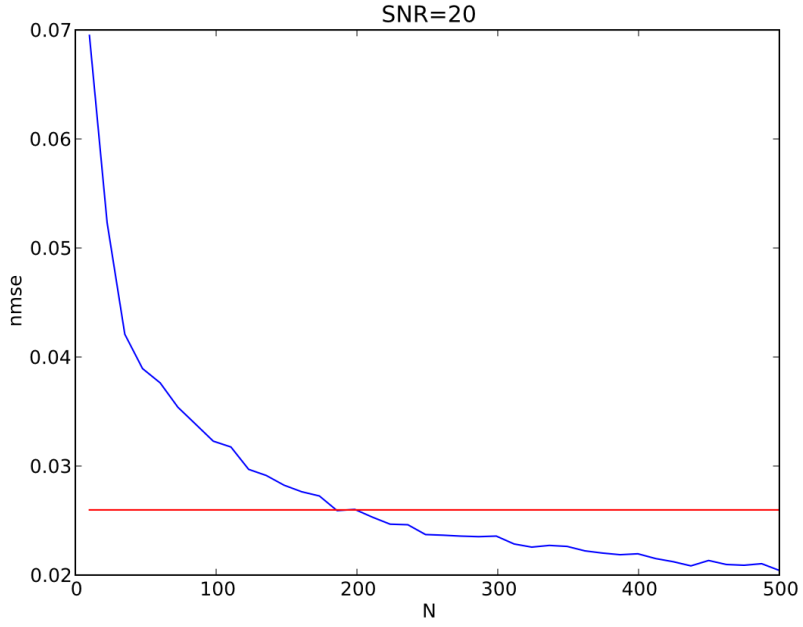


Figure 6.6: Comparison between the NMSE obtained for DSI reconstruction (red line) and SHORE based ℓ_1 recovery (blue curve).

We show quantitative results on the reconstruction of synthetic data, described in 10.7.

Again, we consider three cases : one fiber, two fibers crossing at 90° and two fibers crossing at 60°. Then, we generate the normalized diffusion signal E on a $11 \times 11 \times 11$ cartesian grid (1331 samples). The ground truth EAP P is the inverse Fourier Transform (IFT) of E. For DSI, we generate 514 DWIs in a cube plus an additional unweighted image. The number of 514 corresponds to the number of samples comprised within the sphere of five lattice unit radius. It is the common sampling protocol with the DSI technique. The EAP is, then, obtained by applying an hamming window and an inverse Fourier Transform. We choose the SHORE basis combined with ℓ_1 recovery for comparison. In this case, we reconstruct the diffusion signal for a number of samples $N = 10$ to 500. These samples are spread on 3 shells at b values $b = [1000, 2000, 3000]$ using the generalized Jones algorithm where the radial distribution follow a law in q^1. Then, we use the closed form in Tab. 6.1 to estimate the EAP $\tilde{P}$. For all the experiments, we add a Rician noise to the signal with $SNR = 20$. We evaluate the NMSE between $\tilde{P}$ and the ground truth EAP P on the $11 \times 11 \times 11$ cartesian grid. For DSI this NMSE is fixed and equal to $NMSE_{DSI} = 0.025976$. For the SHORE based EAP reconstruction using the ℓ_1 recovery, we show in Fig. 6.6 the NMSE while taking $N = 10$ to 500 samples (blue curve). In the same figure, we plot a horizontal line (red line) whose the ordinate value is $NMSE_{DSI}$. We see in Fig. 6.6, that 180 measurements are necessary, with SHORE based ℓ_1 reconstruction, to attain the NMSE of DSI, which was obtained with 514 measurements plus an additional unweighted image. Therefore, an acceleration factor of $515/180 \sim 2.86$ is observed. Besides this improvement, our SHORE based ℓ_1 reconstruction has the advantage to give a continuous modeling of the EAP whereas the EAP based DSI reconstruction is dependent to the cartesian grid.

Conclusion on synthetic results

The conclusion to the noisy synthetic experiments is fourfold. First, these results show that the SHORE, SPF and SPFdual bases are appropriate for ℓ_1 recovery, and enables the reconstruction of diffusion signals as accurately as with ℓ_2 recovery with a smaller number of measurements, whereas it is preferable to use an ℓ_2 recovery with the SoH basis. Next, we know that SHORE perfectly enter in the CS framework and is the most suitable basis to use for diffusion signal reconstruction via ℓ_1 recovery, where nearly 20/30 measurements are necessary to well estimate the diffusion signal. Thus, considering the ODF experiments, we see that the SHORE based ℓ_1 recovery gives the best description of the angular information of the underlying synthetic signals. Overall, it provides lower DNC and AE values than the other reconstructions. In particular, we see that 30 measurements are sufficient to obtain a reliable and accurate estimation of the diffusion directions. Finally, we have performed a comparison with DSI and our experiments show that the association SHORE/ℓ_1 recovery enables us to attain the accuracy of DSI with less than half the number of samples used in this high resolution acquisition technique (an acceleration factor of 2.86).

6.4.3 Real human brain data

We also performed experiments on a in-vivo human brain. We acquired the data over 3 shells and 67 directions uniformly distributed on the sphere (A total of 201 directions). We used three b-values (500, 1000, and 2000 $\mathrm{s \cdot mm^{-2}}$), and an imaging matrix of $93 \times 116 \times 93$ with isotropic voxels of 1.7 mm^3. The large number of samples enables the subsampling of the data and apply various sampling schemes. Based on our previous results, we take

samples according to a radial distribution in q^1. All the experiments in this part are performed with 20 samples in both recovery methods (ℓ_1 and ℓ_2). The mean diffusivity constant averaged on all the voxels contained in the region of interest is $D = 0.00054$. We have $\tau = 0.0257$, then $\zeta = 925.93$ for SPF and SHORE bases , and $\zeta = 2.7357e - 05$ (see Sec. 6.3.1 for details about ζ computation). We reconstruct the diffusion signal and, then, use the closed forms defined in Sec. 6.2 to compute the ODF. Fig. 6.7 shows the estimated ODFs and their extracted maxima while using SPF, SPFdual and SHORE in both ℓ_1 and ℓ_2 recovery methods. Regarding our previous experiments, we put aside the SoH basis. We put the fractional anisotropy in background of the region of interest. We also compute a ground truth by considering all the measurements. Because synthetic experiments show that every combination of basis and recovery method gives similar results when considering many measurements, we compute this near ground truth using the SHORE basis and the ℓ_2 recovery method.

First at all, we focus our attention on the advantages/disadvantages of ℓ_1 over ℓ_2 recovery. We see that the ODF reconstructed using ℓ_1 recovery look sharper than in case of ℓ_2 recovery. This particularity is due to the nature of the ℓ_1 minimization problem. Indeed, the ℓ_1 regularization equally penalizes the low frequency components and the high frequency components, whereas the ℓ_2 regularization penalizes more the high frequency components. Hence, ℓ_1 recovery improves the angular resolution and produces sharper diffusion signal estimation than ℓ_2 recovery, and the ODFs appearance directly results from this phenomenon. This sharpening enables the extraction of more maxima. This was already observed from the synthetic experiments. With SPF and SPF dual bases, the CS based ODFs give also some maxima with noisy peaks. ODFs estimated with SHORE basis and ℓ_1 recovery better agree the underlying structure and are not as perturbed as with SPF and SPFdual bases. Then, SHORE basis better resolves the crossing area using ℓ_1 recovery and 20 measurements without catching as much noise.

In conclusion, SHORE appears to be the best basis, which especially gives good results in case of ℓ_1 recovery with only 20 samples. Overall, ℓ_1 recovery gives sharper ODF than ℓ_2 recovery.

6.4.4 Phantom data

Here, we perform our experiments on a phantom data provided by LNAO, which was used in a fiber cup contest in MICCAI 2009 [105, 53]. Diffusion-weighted data of the phantoms were acquired at a spatial resolution of 3 mm^3 isotropic and an imaging matrix of 64x64x3. The data were obtained for three different b-values b=650/1500/2000 $s \cdot mm^{-2}$ along a set of 64 orientations, uniformly distributed over the sphere (A total of 192 directions).

In this experiment, we consider a crossing region and estimate the diffusion signal in three bases : SPF, SPFdual and SHORE. Again, we put aside the SoH basis and take only 20 samples. The mean diffusivity constant averaged on the crossing region is $D = 0.00081$. We have $\tau = 0.0253$, then $\zeta = 617.28$ for SPF and SHORE bases , and $\zeta = 4.1035e - 05$ for SPFdual basis (see Sec. 6.3.1 for details about ζ computation). Figure 6.8 shows the EAP at radii $15\mu m$ and $20\mu m$.

We analytically compute the EAP from the closed formula described in Sec. 6.2. Again, the EAPs estimated with ℓ_1 recovery are sharper that the EAPs estimated with ℓ_2 recovery Based on the known ground truth, we can say that our best estimation are again the one based on the association ℓ_1 recovery/SHORE basis, because its EAP estimation agrees more

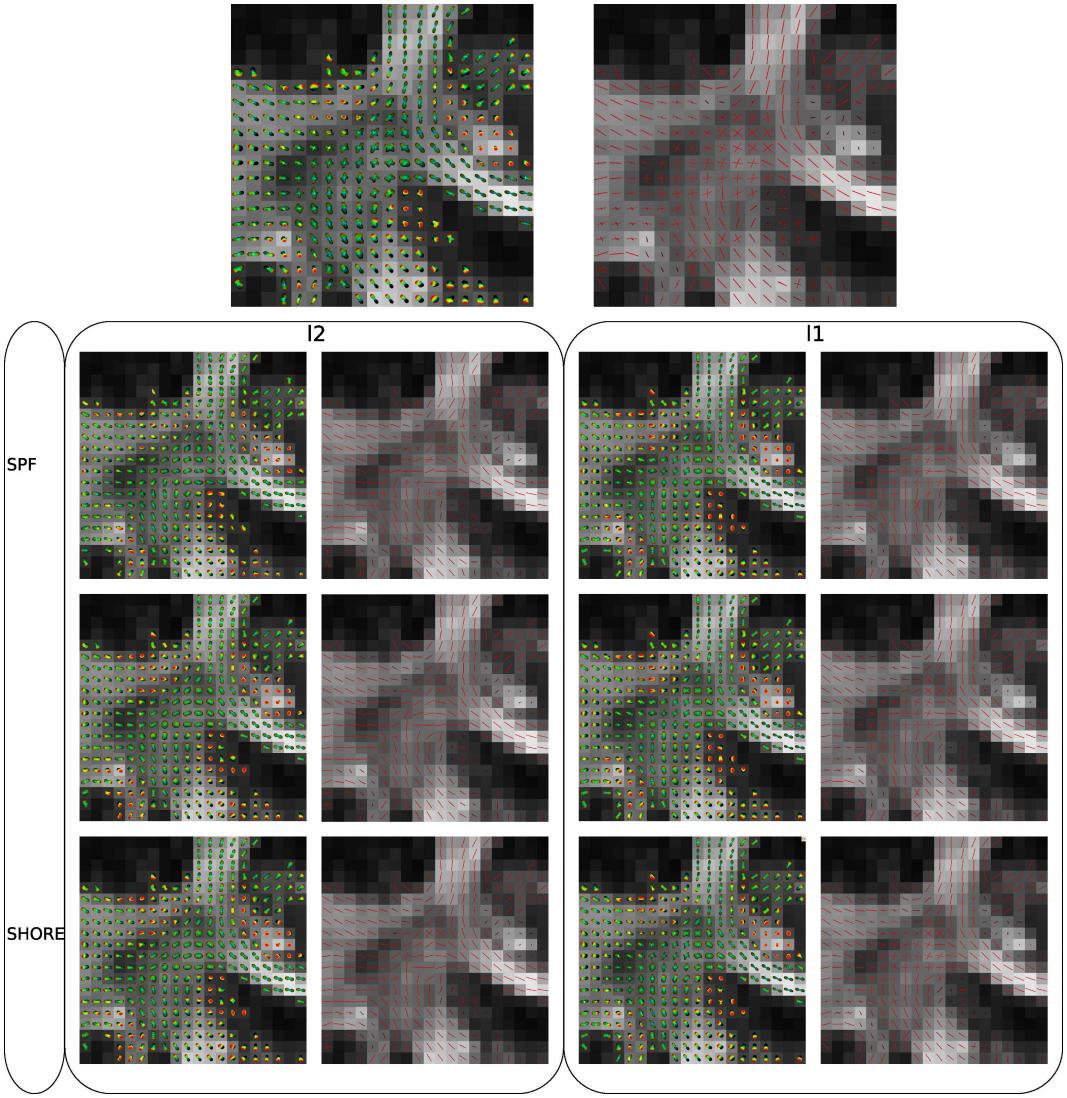


Figure 6.7: ODF and extracted maxima. The underlying diffusion signal is estimated via ℓ_2 recovery (left part) and ℓ_1 recovery (right part). On the top is shown the ground truth of the region of interest. The array below shows the estimated ODF and extracted maxima. Three bases are used for the reconstruction : the Spherical Polar Fourier basis (Top line of the array), the Spherical Polar Fourier dual basis (Second line of the array), the Spherical Polar Fourier basis (Bottom line of the array).

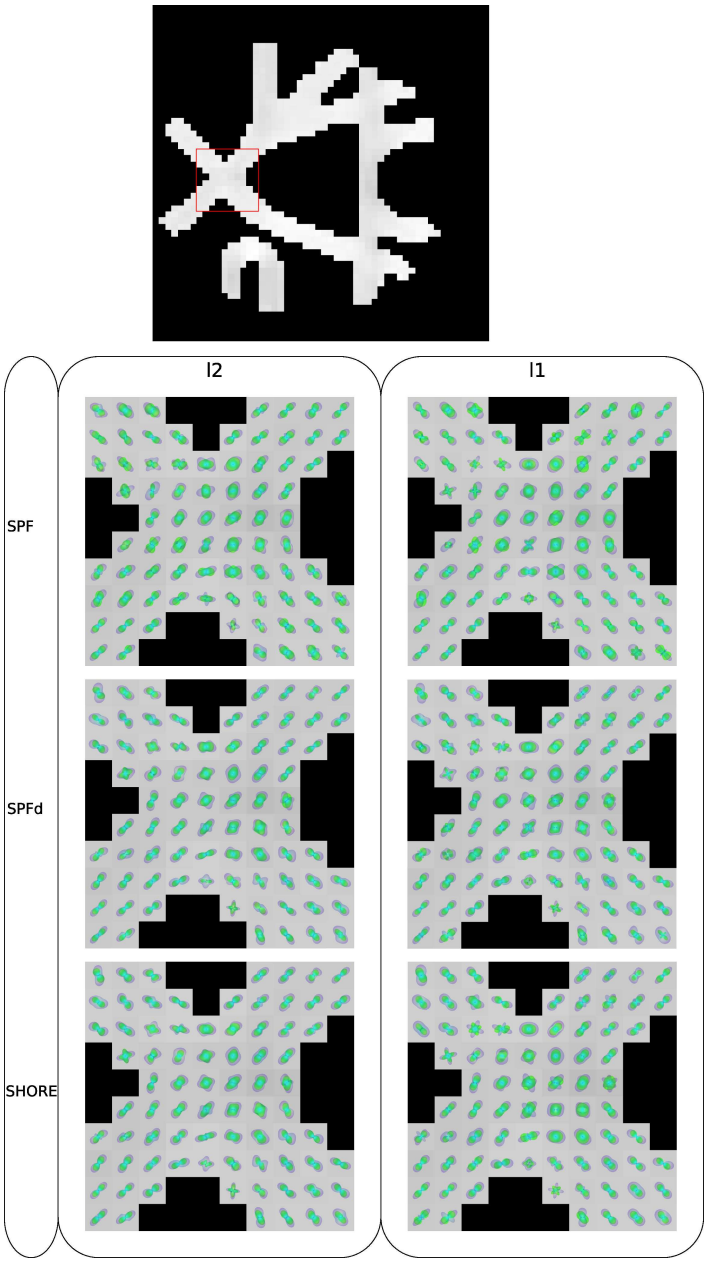


Figure 6.8: Estimated EAP at radii $15\mu m$ (green surface) and $20\mu m$ (blue surface). The underlying diffusion signal is estimated via ℓ_2 recovery (left part) and CS recovery (right part). Three bases are used for the reconstruction : the Spherical Polar Fourier basis (top line), the Spherical Polar Fourier dual basis (second line), the Spherical Polar Fourier dual basis (bottom line).

the underlying crossing area than the other estimations.

In figure 6.9, we compare estimated EAP modeled with SPF and SHORE bases for different numbers of samples. For this purpose, we take a voxel in the phantom where fibers are crossing. We set a 'pseudo' ground truth (GT) as the reconstructed signal for both bases using the whole set of acquisitions (N=300 samples). Then, we estimate the EAP at radii $15\mu m$ (green surface) and $20\mu m$ (blue surface) and show it in Fig. 6.9 (first column). Afterward, we repeat the experiment for $N = 120$, $N = 90$, $N = 60$, $N = 30$ samples. Considering SHORE basis, we note that the estimated EAP is qualitatively similar to the 'pseudo' GT for both radii from $N = 120$ to $N = 60$. For a smaller N the angular information at radius $15\mu m$ becomes too smooth to correctly estimate the diffusion directions. For SPF basis, we remark that this phenomenon (the bad estimation of the diffusion directions at a low radius) appears at $N = 90$. Indeed, at radius $15\mu m$ the diffusion directions given are staggered with the ground truth. On the whole, we see a better angular resolution for higher radii.

This example shows the efficiency of SHORE basis to correctly model the signal (and so the EAP) with less samples than needed with the SPF basis. These remarks hold for a ℓ_1 recovery scheme.

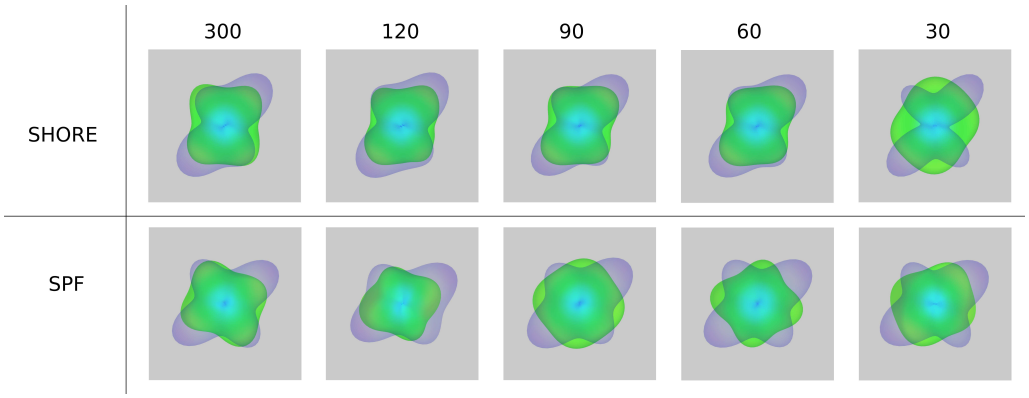


Figure 6.9: Estimated EAP at radii $15\mu m$ (green surface) and $20\mu m$ (blue surface) for different number of samples, while the diffusion signal is modeled with SHORE and SPF bases (respectively top and bottom lines). We used $N = 300$, $N = 120$, $N = 90$, $N = 60$, $N = 30$ samples (from left to right) for the reconstruction.

6.5 Conclusion

We have proposed a new solution to characterize the complete water diffusion process in the white matter, with a very small number of measurements (only 20/30 measurements). The main contribution of this chapter is the investigation of Compressive Sensing to estimate the whole 3D diffusion signal in diffusion MRI. We reviewed every point of CS both in a theoretical and experimental ways for the dMRI purpose. We showed the importance of the sparsity, the incoherence and the RIP in estimating the diffusion signal using an ℓ_1 minimization scheme. Moreover, in this approach, we considered only continuous frameworks, which enable us to compute analytical solutions of the EAP and ODF. The EAP captures

both radial and angular information and completely describes the diffusion process. The ODF is the diffusion feature commonly used to perform tractography. Derivation of more features seems conceivable in these frameworks.

Our extensive review of CS for modeling the diffusion signal using SPF, SPFdual, SoH and SHORE bases leads to two important conclusions about : 1) the efficiency of a ℓ_1 recovery compared to a ℓ_2 recovery, 2) the conditions for an efficient CS recovery. At first, why should we use a ℓ_1 recovery rather than a ℓ_2 recovery ? We showed, when the CS requirements are fulfilled, that a ℓ_1 recovery produces better diffusion signal estimation than the ℓ_2 recovery method does, with not more than 20/30 samples optimally spread. Moreover, ODFs estimated from a ℓ_1 recovery are sharper than with a ℓ_2 recovery and, thus, could be efficiently used for tractography. Nevertheless, we found that the CS technique is quite constrained. Even if we can fully choose our samples to nearly respect the RIP property, two other conditions depend directly on the signal to recover and the basis used for modeling it. These conditions, the sparsity and the incoherence, are not fulfilled for every case. For instance, the SoH basis is not theoretically appropriate for the CS technique. Inversely, the theory pointed out the efficiency of SHORE basis in CS recovery, which was confirmed by experimental results.

To conclude, we showed that the Compressive Sensing technique enables to accurately estimate the diffusion signal along with the EAP and the ODF with only 20/30 samples. However, the Compressive Sensing theory is not easy to handle and is more appropriate, in some cases, than the commonly used ℓ_2 recovery method. In this investigation, we also found an adequate association which fulfills the CS requirements. High sparsity and incoherence are given by SHORE basis and the RIP property is approximately satisfied with a robust multiple shells sampling protocol, which is performed by the generalized Jones algorithm by setting a number of samples by shells proportional to q^1. Moreover, we get rid of the problem of sensitivity towards the regularization parameter in ℓ_1 recovery by using cross validation.

To summarize, we showed that CS Diffusion MRI holds great promise for reconstructing the full water diffusion process within a clinically feasible acquisition time. This opens new and exciting perspectives in diffusion MRI.

The next chapter is close in spirit to this one, except that the sensing matrix is not an orthonormal basis but an overcomplete and redundant dictionary learned from a training data set.

6.6 Appendix

6.6.1 Proof of the analytical ODF solution when the signal is modeled in SPF dual basis

In the SPF formalism, one can express the ODF as,

$$\Upsilon(\mathbf{r}) = \int_{R=0}^{\infty} \sum_{n=0}^{N} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_{nlm} K_n(R) Y_\ell^m(\mathbf{r}) R^2 dR$$
$$= \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} Y_\ell^m((r)) \underbrace{\sum_{n=0}^{N} c_{nlm} \int_{R=0}^{\infty} K_n(R) R^2 dR}_{\upsilon_{\ell m}}.$$
$$\text{with } K_n(R) = \left[\frac{2n!}{\zeta^{3/2}\Gamma(n+3/2)}\right]^{1/2} exp(\frac{-R^2}{2\zeta}) L_n^{1/2}(\frac{R^2}{\zeta}). \tag{6.30}$$

Using the formula $\int_0^\infty exp(-sx)x^\alpha L_n^\alpha(x)dx = \frac{\Gamma(\alpha+n+1)(s-1)^n}{n!s^{\alpha+n+1}}$ [111] and this substitution : $R = \sqrt{x\zeta}$ and $dR = \frac{1}{2}\sqrt{\frac{\zeta}{x}}dx$, we get,

$$\upsilon_{\ell m} = \sum_{n=0}^{N} c_{nlm} \frac{\zeta^{3/2}}{2} \left[\frac{2n!}{\zeta^{3/2}\Gamma(n+\frac{3}{2})}\right]^{1/2} \frac{\Gamma(\frac{3}{2}+n)(-\frac{1}{2})^n}{n!\frac{1}{2}^{(3/2+n)}}$$
$$= \sum_{n=0}^{N} c_{nlm} 2(-1)^n \zeta^{3/4} \left[\frac{\Gamma(n+\frac{3}{2})}{n!}\right]^{1/2}.$$

6.6.2 Proof of the analytical ODF solution when the signal is modeled in SHORE basis

In the SHORE formalism, one can express the ODF as,

$$\Upsilon(\mathbf{r}) = \int_{R=0}^{\infty} \sum_{n=0}^{N} \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} c_{nlm} K_n(R) Y_\ell^m(\mathbf{r}) R^2 dR$$
$$= \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} Y_\ell^m((r)) \underbrace{\sum_{n=0}^{N} c_{nlm} \int_{R=0}^{\infty} K_n(R) R^2 dR}_{\upsilon_{\ell m}}$$
$$\text{with } K_n(R) = (-1)^{n-l/2} \left[\frac{2(4\pi^2\zeta)^{3/2}(n-\ell)!}{\Gamma(n+3/2)}\right]^{1/2} \left(4\pi^2\zeta R^2\right)^{l/2}$$
$$\exp\left(-2\pi^2\zeta R^2\right) L_{n-l}^{l+1/2}\left(4\pi^2\zeta R^2\right) \tag{6.31}$$

Using the formula $\int_0^\infty exp(-sx)x^\beta L_n^\alpha(x)dx = \frac{\Gamma(\beta+1)\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+1)} s^{-\beta-1} {}_2F_1(-n, \beta+1; \alpha+1, 1/s)$ [111] and this substitution : $R = \sqrt{\frac{x}{4\pi^2\zeta}}$ and $dR = \frac{1}{2}\sqrt{\frac{1}{4\pi^2\zeta x}}dx$, we get,

$$v_{\ell m} = \sum_{n=0}^{N} c_{n\ell m} \frac{(-1)^{n-l/2}}{2(4\pi^2\zeta)^{3/2}} \left[\frac{2(4\pi^2\zeta)^{3/2}(n-l)!}{\Gamma(n+3/2)} \right]^{1/2} \tag{6.32}$$

$$\frac{\Gamma(\ell/2+3/2)\Gamma(3/2+n)}{\Gamma(l+3/2)(n-l)!} \left(\frac{1}{2}\right)^{-\ell/2-3/2} \tag{6.33}$$

$${}_2F_1(-n+l, l/2+3/2; l+3/2; 2) \tag{6.34}$$

Chapter 7

Parametric and sparse dictionary learning

Contents

7.1 Motivations 122
7.2 A framework for the recovery of the complete diffusion MRI process 123
7.2.1 Continuous diffusion modeling with a constrained dictionary . . 123
7.2.2 Closed formulae for diffusion features 125
7.3 A parametric dictionary learning for sparse dMRI 126
7.3.1 Dictionary learning algorithm 127
7.3.2 Signal estimation via the learned dictionary 129
7.4 Experiments on synthetic data 129
7.4.1 Which radial and angular order for the dictionary ? 130
7.4.2 Training phase 131
7.4.3 Validation 132
7.4.4 Discussion on experiments with synthetic data 138
7.5 Experiments on real data 139
7.5.1 Learning and reconstruction on a 7T scanner data 141
7.5.2 Learning and reconstruction on a 3T scanner data 141
7.6 Conclusion 144
7.7 Appendix 144
7.7.1 Derivation of the normalization constant 144
7.7.2 Derivation of the Ensemble Average Propagator 145
7.7.3 Derivation of the Orientation Distribution Function 146

Overview

In this chapter, we first propose an original and efficient computational framework to model continuous diffusion MRI (dMRI) signals and analytically recover important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Then, we develop an efficient parametric dictionary learning algorithm and exploit the sparse property of a well-designed dictionary to recover the diffusion signal and its features with a reduced number of measurements. The properties and potentials of the technique are demonstrated using various simulations on synthetic data and on human brain data acquired from 7-T and 3-T scanners. It is shown that the technique can clearly recover the dMRI signal and its features with a much better accuracy compared to state-of-the-art approaches, even with a small and reduced number of measurements. In particular, we can accurately recover the ODF in regions of multiple fiber crossing, which could open new perspectives for some dMRI applications such as fiber tractography. This chapter can be seen as an extension of **chapter 6** where the orthonormal basis is replaced by an overcomplete and redundant dictionary learned from a training data set. This contribution have been published in [88].

7.1 Motivations

So far, we have seen that sparse reconstruction approaches, in particular Compressive Sensing (CS), were found to successfully reduce the number of acquisitions in dMRI . In **chapter 5** we used CS to accelerate the DSI acquisition. In **chapter 6**, we combined CS with a continuous modeling of the diffusion signal using orthonormal bases. In [97] and in [122], the authors elegantly designed dictionaries for sparse modeling in dMRI. They provide an overcomplete dictionary computed from a discretized version of predefined functions, i.e. the Spherical Ridgelets in [97] (see [109] for the multiple shells version) and the Spherical Wavelets in [122]. Learning a dictionary provides an alternative way to design sparse dictionaries [15, 57, 86, 136].

Some approaches have been recently proposed in order to design dictionaries that enable sparse representations (A good overview can be found in [2]). For instance, Bilgic et al [15] and Gramfort et al [57] learn dictionaries from DSI like acquisitions and use it to either denoise full DSI data or to perform undersampled DSI acquisitions and reconstructions. In particular, the authors in [57] nicely exploit the symmetry of the signal in order to assess free parameters of the dictionary learning problem. However, these two latter works lead to non-parametric dictionaries, which does not provide continuous representations of the diffusion signal nor allow the determination of analytical formulae for diffusion features. The strength of the parametric dictionary learning approach, as the one we propose in this article, lies in its ability to address these weaknesses. A work regarding parametric dictionary learning was published in [136], in which the dictionary atoms are formed by a weighted combination of 3^{rd} order B-splines. It proved that the method is efficient on synthetic data simulated with 81 gradient directions. The work of [136] appears promising in reconstructing the diffusion signals, and further enhancement could be done regarding the development of analytical formulae to estimate other diffusion features. This would make this work a good resource in the context of dictionary learning. More recently, we proposed in [86] to learn a dictionary where each atom is constrained to be a parametric

function. In [86], this parametric function is a combination of a radial part and an angular part represented by the symmetric and real Spherical Harmonics (SH) [44]. The radial part is a polynomial weighted by an exponential. 50 measurements were sufficient to reconstruct very good quality diffusion signals, ODFs and EAPs. However, this approach essentially handles the learning of the radial part, i.e. the polynomial coefficients and a scale parameter in the exponential, whereas we observed (see [87]) that the angular part could make the dictionary much sparser if we adequately combine several SH functions instead of only one.

In this chapter, we present a method, which exploits the sparse property of a well designed dictionary based on a computational dMRI framework, in order to recover the diffusion signal with a reduced number of measurements. This framework enables a continuous modeling of the diffusion signal and leads to analytical formulae to estimate important diffusion features, namely the ODF and the EAP. To improve our previous work in [86], we modify the parametric function, describing the atoms, to learn both the radial and the angular part, which provide a very sparse representation of diffusion signals and further reduce the number of measurements (15 measurements are found to be sufficient to start recovering the EAP and some derived diffusion features whereas 50 measurements are used in [86]). Furthermore, we extend the experimental part of [86] by learning and validating our approach on the synthetic data proposed in the HARDI contest at ISBI 2012 [1], and on real data acquired from both 3T and 7T scanners. A preliminary work [87] regarding the learning of both the radial part and the angular part of the diffusion signal was published in the proceedings of the HARDI contest at ISBI 2012 and we obtained the best results in our category. Our approach presented in this paper indicates an increase in terms of reconstruction accuracy compared to the results presented in [87].

This chapter is structured as follows : we start by introducing the dMRI framework together with the proposed dictionary, then we focus on the parametric dictionary learning algorithm and finally we conclude with an experimental part illustrating the added-value of our approach with promising results showing how our approach allows the accurate reconstruction of the diffusion signal and some of its features. This experimental part is completed by a comparison with state of the art approaches, and is performed on synthetic and real data from 3T and 7T scanners.

7.2 A computational framework for the recovery of the complete diffusion MRI process

In this section, we introduce a new dMRI framework for modeling the diffusion signal. From this continuous representation, we derive analytical formulae that enable the estimation of important diffusion features such as the Ensemble Average Propagator (EAP), the Orientation Distribution Function (ODF). We give full derivations for these formulae in the appendices.

7.2.1 Continuous diffusion modeling with a constrained dictionary

We propose to design an overcomplete dictionary $\Psi = \{\Psi_k\}_{k=0,\dots,K}$, such that the diffusion signal E is expressed as a truncated linear combination of K 3D atoms Ψ_k, i.e

[1] http://hardi.epfl.ch/

$$E(q\mathbf{u}) = \sum_{k=0}^{K} c_k \Psi_k(q\mathbf{u}), \tag{7.1}$$

where c_k are the transform coefficients, q the norm of the effective gradient and $\mathbf{u}$ an unitary vector.

In diffusion multiple shells imaging [5, 99, 45], previous works have proposed the modeling of the diffusion signal in bases where each atom Ψ_k is a combination of a radial part and an angular part. These works primarily handle the question of the radial part. For instance, Descoteaux et al [45] describe the radial part of the atom as a combination of two rational functions coming from the total solution of the Laplace equation, whereas the authors in [99, 5] use polynomials weighted by an exponential. Merlet et al [86] increased the sparsity of the representation in [99] by designing a dictionary where the polynomial coefficients and the scale parameters in the exponentials are learned from a training data set [86]. Although the radial attenuation of the diffusion signal is more or less well fitted with these bases/dictionaries, the accurate and sparse estimation of the full diffusion signal is still limited by the angular part described by the real and symmetric Spherical Harmonic basis functions (SH). The SH have been proved useful in many settings but are not sparse enough in modeling the directional features of the diffusion process [96]. However, our preliminary work published in [87] shows that a well-chosen combination of SH could sparsely represent the angular part. Therefore, we propose to model this angular part with such a combination of SH. As for the radial part, we model it with a combination of exponential functions weighted by a monomial in order to ensure the continuity of the function at zero [27]. The complete description of each atom Ψ_k of the dictionary is given by :

$$\begin{aligned}
\Psi_k(\mathbf{q}) = \Psi_k(q\mathbf{u}) &= \frac{1}{\sqrt{\chi_k}} \sum_{i=0}^{I} \alpha_{ki} \exp\left(-\nu_{ki} q^2\right) \sum_{j=0}^{J} \beta_{kij} q^{l(j)} Y_j(\mathbf{u}) \\
&= \frac{1}{\sqrt{\chi_k}} \sum_{i=0}^{I} \sum_{j=0}^{J} \gamma_{kij} \exp\left(-\nu_{ki} q^2\right) q^{l(j)} Y_j(\mathbf{u}) \\
&= \Psi_k(\boldsymbol{\gamma_k}, \boldsymbol{\nu_k}, q\mathbf{u}),
\end{aligned} \tag{7.2}$$

with $\mathbf{q}$ the 3D effective gradient, $\mathbf{u}$ a unitary vector and q the norm of the effective gradient such that $\mathbf{q} = q\mathbf{u}$. I and J are, respectively, the radial order and the angular order of the dictionary. J also corresponds the total number of SH taken into account in the modeling not to be confused with the maximal SH order L. Indeed, J is directly related to the maximal SH order L as $J = (L+1)(L+2)/2$. $Y_j(\mathbf{u})$ is the SH of order $l(j) = 0$ for $j = 1$, $l(j) = 2$ for $j \in \{2, ..., 6\}$, $l(j) = 4$ for $j \in \{7, ..., 15\}$ ... . $\boldsymbol{\gamma_k} = \{\gamma_{kij}\}_{i=0...I, j=0...J}$ and $\boldsymbol{\nu_k} = \{\nu_{ki}\}_{i=0...I}$ are two vectors of parameters, which will be set during the learning process. The term $q^{l(j)}$ ensures the continuity of Ψ_k at zero. χ_k is a constant, which ensures the normalization of Ψ_k for the ℓ_2 norm, i.e $\sqrt{\int_{\mathcal{R}^3} \Psi_k^2(\mathbf{q}) d\mathbf{q}} = 1$, and is expressed as

$$\chi_k = \sum_{i'=0}^{I}\sum_{i=0}^{I}\sum_{j=0}^{J} \frac{\gamma_{kij}\gamma_{ki'j}}{2(\nu_{ki}+\nu_{ki'})^{\ell(j)+3/2}}\Gamma\left(\ell(j)+\frac{3}{2}\right), \tag{7.3}$$

with Γ the gamma function. We derive (7.3) in appendix 7.7.1.

If we consider the simple case where all the coefficients γ_{kij} and ν_{kj} are zero except γ_{k00} and ν_{k0}, we get $\Psi_k(q\mathbf{u}) = \frac{1}{\sqrt{\chi_k}}\frac{\gamma_{k00}}{2\sqrt{\pi}}\exp(-\nu_{k0}q^2)$. This mono-exponential representation of the atom Ψ_k does not depend on the angular direction $\mathbf{u}$ and is, for instance, appropriate to describe isotropic diffusion configuration.

Note that in [87] each atom is described by a combination of SHORE basis functions [99] with a predefined scale parameter. The main advantage of the atom description in (7.2) lies in the possibility of learning the scale parameters ν_{ki}, which provides a sparser signal estimation than in [87]. We see in the experimental part that this new framework leads to an increase in terms of reconstruction accuracy compared to the results presented in [87], which were already the best in their category.

7.2.2 Closed formulae for diffusion features

Using the dictionary, $\Psi = \{\Psi_k\}_{k=0,\dots,K}$, proposed in the previous section to reconstruct the diffusion signal (see (7.2)), we derive important and analytical closed formulae for estimating the EAP and the ODF. We describe these formulae in the following.

The Ensemble Average Propagator

The EAP, denoted by $P(R\mathbf{r})$, represents the full 3D displacement probability function of water molecules in every voxel and underlies the derivation of the ODF. It is the inverse Fourier transform of the normalized diffusion signal, denoted by$E(q\mathbf{u})$. If we decompose (3.10) in two integrals we can express $E(q\mathbf{u})$ as

$$P(R\mathbf{r}) = \int_{q=0}^{\infty}\int_{\mathbf{u}\in S^2} E(q\mathbf{u})\exp(+2\pi i q R\mathbf{u}.\mathbf{r})\mathrm{d}\mathbf{u}q^2\mathrm{d}q. \tag{7.4}$$

q and R are, respectively, the norm of the effective gradient and the radius of the 3D location in every voxel, $\mathbf{u}$ and $\mathbf{r}$ are unit vectors. From (7.1) and (7.4), we derive in 7.7.2 the following expression for the EAP :

$$P(R\mathbf{r}) = \sum_{k=0}^{K}\frac{c_k}{\sqrt{\chi_k}}\sum_{i=0}^{I}\sum_{j=0}^{J}\gamma_{kij}(-1)^{l(j)/2}\left(\frac{\pi}{\nu_{ki}}\right)^{l(j)+3/2} R^{l(j)}\exp\left(\frac{-(\pi R)^2}{\nu_{ki}}\right)Y_j(\mathbf{r}) \tag{7.5}$$

If we consider the special case of the mono-exponential representation of the atom Ψ_k, we get $P(R\mathbf{r}) = \sum_{k=0}^{K}\frac{c_k}{\sqrt{\chi_k}}\frac{\gamma_{k00}}{2\sqrt{\pi}}\left(\frac{\pi}{\nu_{k0}}\right)^{3/2}\exp\left(\frac{-(\pi R)^2}{\nu_{k0}}\right)$, i.e an isotropic propagator described by an mono-exponential decay similar in every direction.

Solid angle ODF

The ODF represents the full angular distribution of $P(R\mathbf{r})$. One relies on the ODF to perform fiber tractography [43], then an accurate and fast computation of this diffusion feature is very appreciated. From (3.21), we derive in 7.7.3 the following closed form for the ODF :

$$\Upsilon(\mathbf{r}) = \sum_{k=0}^{K} \frac{c_k}{\sqrt{\chi_k}} \sum_{i=0}^{I} \sum_{j=0}^{J} \gamma_{kij} (-1)^{l(j)/2} \left(\frac{\pi}{\nu_{ki}}\right)^{l(j)+1} \frac{(l(j)+1)!!}{2\left(\frac{2\pi^2}{v_{ki}}\right)^{l(j)/2+1}} Y_j(\mathbf{r}) \tag{7.6}$$

Considering the precedent example, the ODF is represented by a scalar, i.e. $\Upsilon(\mathbf{r}) = \sum_{k=0}^{K} \frac{c_k}{\sqrt{\chi_k}} \frac{\gamma_{k00}}{8\pi^{3/2}}$, meaning that the ODF has the same value for every direction.

7.3 A parametric dictionary learning for sparse dMRI

Here, we introduce a parametric dictionary learning (PDL) method that enables a sparse representation of any diffusion signal from continuous and parametric functions. There are four advantages to consider a parametric approach for dictionary learning:

- A parametric dictionary is defined by a set of parameters ($\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$ in Sec. 7.2), which gives a continuous representation of each atom and, thus, enables a continuous modeling of the diffusion signal. This is suitable for data interpolation and extrapolation.
- Analytical formulae can be derived to estimate important diffusion features as the EAP and the ODF.
- PDL is acquisition independent, i.e. the sampling scheme used for learning the dictionary does not have to be the same as the sampling scheme used for reconstructing the signals.
- PDL enables one to reduce the dimensionality of the dictionary atoms.

These four advantages, together with the quality of the obtained results (see Sec. 7.4), makes our parametric dictionary approach very attractive compared to non parametric methods [15, 57].

Concerning the development of our algorithm, we started by considering the K-SVD algorithm [2] as a model for our own method. Although the K-SVD method appears powerful in designing sparse dictionaries, this technique as described in [2] designs only non-parametric dictionaries, which do not present the advantages described above. Hence, we developed our own algorithm, which overcomes the limitation of the K-SVD algorithm. Our algorithm alternates between 2 steps: a sparse coding step and a dictionary update step, where the vectors of parameters $\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$ (see Sec. 7.2) are estimated for every atom d_k of the dictionary, using the non linear Levenberg-Marquardt (LM) algorithm. The section 7.3.1 presents our dictionary learning algorithm and the section 7.3.2 describes the method we use to reconstruct any diffusion signal using the dictionary previously learned.

7.3.1 Dictionary learning algorithm

Notation : Suppose the training data set consists of M observations $\{\mathbf{s}_i\}_{i=1}^{M}$ (i.e. M voxels). For each observation s_i we have m_s samples in the q-space, i.e. $\mathbf{s}_{1..M} \in \mathbb{R}^{m_s}$. We represent $\{\mathbf{s}_i\}_{i=1}^{M}$ in matrix form $\mathbf{S} \in \mathbb{R}^{m_s \times M}$ where $\mathbf{s}_i$ is the i^{th} column. The algorithm searches for the dictionary $\mathbf{D} \in \mathbb{R}^{m_s \times K}$, that enables the sparsest representation for every column of $\mathbf{S}$. The dictionary consists in K atoms $\{\mathbf{d}_k\}_{k=1}^{K}$ with $\mathbf{d}_k \in \mathbb{R}^{m_s}$ a column of $\mathbf{D}$. We constrain $\mathbf{d}_k$ to be an instance of the 3D function $\Psi_k(\boldsymbol{\gamma_k}, \boldsymbol{\nu_k}, q\mathbf{u})$ in (7.2). Here, we do not try to directly estimate $\mathbf{d}_k$ but the vectors of parameters $\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$, that characterize the atom $\mathbf{d}_k$. For each observation $\mathbf{s}_i$, we define a coefficient vector $\mathbf{c}_i \in \mathbb{R}^K$, which forms the i^{th} column of the coefficient matrix $\mathbf{C} \in \mathbb{R}^{K \times M}$.

Problem statement: Given a training data set $\mathbf{S}$, we search for the dictionary $\mathbf{D}$ that gives the sparsest representation of this set (i.e. for each column $\mathbf{s}_i$ of $\mathbf{S}$). Mathematically, the problem is to find the dictionary $\mathbf{D}$ and the vectors $\mathbf{c}_i$ in $\mathbf{C}$ by solving :

$$\arg\min_{\mathbf{c}_i, \mathbf{D}} \{\|\mathbf{S} - \mathbf{DC}\|_2^2\} \text{ subject to } \forall_i \|\mathbf{c}_i\|_1 \leq \epsilon \tag{7.7}$$

with ϵ a small real defining the degree of sparsity of the dictionary. The minimization of the first term in (7.7) enables the signal estimation $\mathbf{DC}$ to remain close to the training data set $\mathbf{S}$ and the constraint imposes the sparsity of each signal representation $\mathbf{c_i}$ with respect to the dictionary $\mathbf{D}$. The method to solve (7.7) is described in the following and a summary of the algorithm is given in Alg. 1. This algorithm iteratively alternates between sparse signal estimations (i.e. $\{\mathbf{c}_i\}_{i=1}^{M}$) and updates of the dictionary (i.e. $\mathbf{D}$) so to better fit the training data set (i.e. $\mathbf{S}$).

First step (Sparse signal estimation): In the first step, the estimation of the column vector $\mathbf{c}_i$ is performed separately for each signal $\mathbf{s}_i$, i.e for each column of $\mathbf{S}$. Sparse estimation is achieved by solving the LASSO (Least Absolute Shrinkage and Selection Operator) problem [119]. It consists in minimizing the following objective function

$$\min_{\mathbf{c}_i} \|\mathbf{s}_i - \mathbf{D}\mathbf{c}_i\|_2^2 + \lambda \|\mathbf{c}_i\|_1. \tag{7.8}$$

λ is a constant that controls the degree of sparsity of the coefficients estimated. Note that we relax the constraint in (7.7) by using a Lagrangian multiplier (λ). There exist numerous iterative algorithms for efficiently solving such kind of constrained problems. These include coordinate descent, least-angle regression (LARS) [50], fast iterative thresholding shrinkage algorithm (FISTA) [12], etc. A number of these methods are available under the Python$^{\text{TM}}$ library Scikit-learn [103]. We use a Python$^{\text{TM}}$ implementation [103] of coordinate descent to solve (7.8).

Second step (Dictionary update): In the second step, we update the dictionary $\mathbf{D}$. For this purpose, we compute an absolute averaged coefficient vector $\hat{\mathbf{c}} \in \mathbb{R}^{n_c}$, such that $\hat{\mathbf{c}} = 1/M \sum_i |\mathbf{c}_i|$ ($|.|$ denotes the absolute value of each vector component), and find the atoms associated with the non zero values of $\hat{\mathbf{c}}$. It gives a rough idea of which atoms are used for modeling the signal and enables one to discard some unnecessary atoms and, thus, to enforce sparsity. Then, in this set of atoms, we update one atom at a time, while fixing all the others. This process is repeated for all the atoms associated with the non-zero

coefficients of $\hat{\mathbf{c}}$.

The in-update atom is denoted $\mathbf{d}_{k_0}$. To update this atom, we begin by decomposing the error term in (7.7) as in [2], i.e.

$$\begin{aligned}\|\mathbf{S}-\mathbf{DC}\|_{\ell 2}^2 &= \left\|\mathbf{S}-\sum_{k=1}^{K}\mathbf{d}_k\mathbf{c}_k^r\right\|_{\ell 2}^2 \\ &= \left\|\left(S-\sum_{k\neq k_0}\mathbf{d}_k\mathbf{c}_k^r\right)-\mathbf{d}_{k_0}\mathbf{c}_{k_0}^r\right\|_{\ell 2}^2 \\ &= \left\|\mathbf{E}_{k_0}-\mathbf{d}_{k_0}\mathbf{c}_{k_0}^r\right\|_{\ell 2}^2, \end{aligned} \tag{7.9}$$

where $\mathbf{c}_k^r$ is the k^{th} row of $\mathbf{C}$. The error matrix, denoted $\mathbf{E}_{k_0}$, contains the error between each observation $\mathbf{s}_i$ (the i^{th} column of $\mathbf{S}$) and its respective estimation with the dictionary where the k_0^{th} atom is removed. We could directly use the LM algorithm in order to fit the atom $\mathbf{d}_{k_0}$ to the error matrix $\mathbf{E}_{k_0}$. However, because it takes into account all the observations $\mathbf{s}_i$, this dictionary update would not impose sparsity. Instead, we enforce the sparsity by constraining the atom $\mathbf{d}_{k_0}$ to fit only a subset of observations and not the entire data set. For this purpose, we define the group of observations that use the atom $\mathbf{d}_{k_0}$, i.e $w_{k_0} = \{i, 1 \leq i \leq M, c_i(k_0) \neq 0\}$. In other words, they are the observations whose coefficients, associated with the atom $\mathbf{d}_{k_0}$, are non zeros. Then, we compute the error matrix $\mathbf{E}_{w_{k_0}} \in \mathbb{R}^{m_s \times \mathrm{card}(w_{k_0})}$. It corresponds to the estimation error between the observation vector $\{\mathbf{s}_i\}_{i\in w_{k_0}}$ that forms the columns of $\mathbf{S}_{w_{k_0}} \in \mathbb{R}^{m_s\times \mathrm{card}(w_{k_0})}$ and the signal estimated for the group of observation w_{k_0} (The $k_0{}^{\text{th}}$ atom is still removed from the dictionary), i.e $\tilde{\mathbf{S}}_{w_{k_0}} = \sum_{k\neq k_0}\mathbf{d}_k\mathbf{c}_k^r(i), i \in w_{k_0}$. Mathematically speaking, we have $\mathbf{E}_{w_{k_0}} = \mathbf{S}_{w_{k_0}} - \sum_{k\neq k_0}\mathbf{d}_k\mathbf{c}_k^r(i), i \in w_{k_0}$. Finally, we estimate the vector of parameters $\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$ by constraining $\mathbf{d}_{k_0}$ to fit the error matrix $\mathbf{E}_{w_{k_0}}$. This part is performed via the non linear Levenberg-Marquardt algorithm (LMA) . The atom update procedure is repeated for every atom $\mathbf{d}_{k_0}$ associated with the non zeros coefficients of $\hat{\mathbf{c}}$.

The method is given in Algorithm 1 as a whole.

Convergence : The sparse coding step (7.8) is well known to be convex and the coordinate descent algorithm allows one to converge to the unique solution specific to the current dictionary $\mathbf{D}$. The dictionary update step, where $\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$ are estimated using the LMA may converges to a local minima, depending on the initial solutions. Then, (7.7) is convex for $\mathbf{c}$ and converge for $\mathbf{D}$, which do not ensure a convergence to a global minimum. Nevertheless, in our experiments, a stationary point has been reached after few iterations and the resulting dictionaries were proved very good experimentally.

Initialization : The problem in (7.7) admits local minima, and the solution may vary depending on the initial parameters. We tried several ways to initialize the algorithm among which random initialization and initialization from signals selected at random in the training data set.. However, these kinds of initialization were not satisfactory since the corresponding solutions were too different between several attempts to build the dictionary. After many experiments, we finally selected each initial atom as a random combination of several training signals. This gave us the best satisfactory results, with the smallest sensitivity to the initialization.

Algorithm 1 Semi-parametric dictionary learning

1. Initialize the dictionary by fixing its dimension K and the vectors of parameters $\boldsymbol{\gamma_k}$ and $\boldsymbol{\nu_k}$ for $k = 1...K$ as random.
2. Sparse estimation of the observations $\{\mathbf{s}_i\}_{i=1}^{M}$. We use the coordinate descent algorithm to solve for $\mathbf{c_i}$ associated to each observation :

$$\min_{\mathbf{c}_i} \|\mathbf{s}_i - \mathbf{D}\mathbf{c}_i\|_2^2 + \lambda\|\mathbf{c}_i\|_1.$$

3. Updating the dictionary. Compute the absolute averaged coefficients vector $\hat{\mathbf{c}} = 1/M \sum_i |\mathbf{c}_i|$. Repeat until all the atoms of the dictionary, with non zeros value in $\hat{\mathbf{c}}$, have been scanned :
   - Let the current atom be the k_0^{th}.
   - Define the group of observation that use this atom : $w_{k_0} = \{i, 1 \leq i \leq M, c_{k_0}(i) \neq 0\}$.
   - Compute the error matrix $\mathbf{E}_{w_{k_0}} = \mathbf{S}_{w_{k_0}} - \sum_{k \neq k_0} \mathbf{d}_k \mathbf{c}_k^r(i), i \in w_{k_0}$. $\mathbf{S}_{w_{k_0}}$ contains the observation $s_i, i \in w_{k_0}$.
   - Apply the Levenberg - Marquardt algorithm to estimate the vectors of parameters $\boldsymbol{\gamma_{k_0}}$ and $\boldsymbol{\nu_{k_0}}$, which constrain $\mathbf{d}_{k_0}$ to best fit $\mathbf{E}_{w_{k_0}}$
   - Update $\mathbf{d}_{k_0}$ according to $\boldsymbol{\gamma_{k_0}}$ and $\boldsymbol{\nu_{k_0}}$.
4. Go back to the step 2 unless the overall error does not vary anymore

7.3.2 Signal estimation via the learned dictionary

The purpose of section 7.3.1 was to design a parametric dictionary $\mathbf{D}$. Now, we are able to recover any sparse representation $\mathbf{c}$ of the diffusion signal $\mathbf{s}$ using the dictionary $\mathbf{D}$ by solving the LASSO problem :

$$\min_{\mathbf{c}} \|\mathbf{s} - \mathbf{D}\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1. \tag{7.10}$$

We use the same algorithm as in the learning step to solve (7.8), i.e. the coordinate descent algorithm.

7.4 Experiments on synthetic data

We first train and validate our parametric dictionary on synthetic data. We assume the normalized diffusion signal $E(\mathbf{q})$ is generated from the multi-tensor described in Sec. 10.7

In the remainder of this section, we describe the steps required to correctly design our parametric dictionary, i.e. the choice of the dictionary radial and angular orders and the learning phase, then we validate the learned dictionary on synthetic data using the analytical formulae we have just described.

7.4.1 Which radial and angular order for the dictionary ?

We need to fix the angular and radial order, respectively denoted J and I in (7.2), for the dictionary generation.

We begin by defining the dictionary angular order J, which is related to the SH order L as $J = (L+1)(L+2)/2$. For this purpose, we generate synthetic ODFs corresponding to two fibers crossing at different degrees : 0°, 40°, 60°, 90°. Then, we fit each ODF with the Spherical Harmonic basis of order $L = 0, 2, 4, 6, 8, 10, 12$ using a least squares fitting technique, and we compute the Normalized Mean Square Error (NMSE) between the original synthetic ODFs (ψ) and the estimated ODFs in terms of SH (ψ_e), i.e. $NMSE = \frac{\|\psi - \psi_e\|_2^2}{\|\psi\|_2^2}$. The resulting NMSEs are shown in Fig. 7.1. Regarding this figure, we find that a SH order of $L = 8$ is sufficient to correctly estimate an ODF (In particular for fibers crossing at 40° or more). This leads to set the dictionary angular order to $J = 45$.

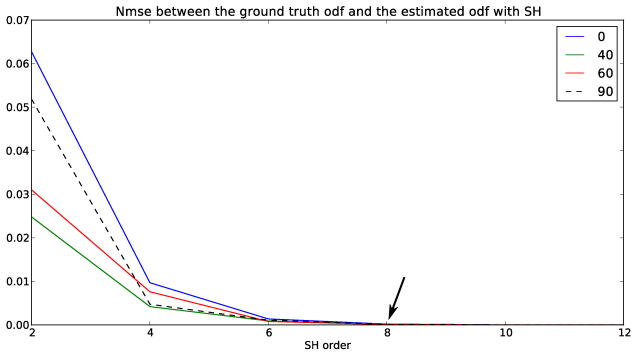


Figure 7.1: NMSE between the ground truth ODFs and the estimated ODFs in terms of SH. The arrow indicates the SH order corresponding to a NMSE considered as sufficiently close to zero.

The radial order is related to the number of fiber crossing in the voxel of interest. We assume a maximum of three crossing fibers in each voxel. This is the case in the region where the corticospinal, corpus callosum and superior longitudinal fasciculus fiber bundles are crossing. Voxel with more than three fibers are considered as part of the noisy background. Consequently, setting the radial order at three appears to be the best choice and was found experimentally satisfactory. We show in Fig. 7.2 some examples of radial attenuation in arbitrarily chosen directions simulated using a multi-Gaussian model mimicking three crossing fibers (plain line). We fit each of these radial attenuation with the radial part of the model in (7.2) and radial orders $I = 1, 2, 3$, using the Levenberg-Marquardt algorithm (The angular order J is set to zero). Fig. 7.2 indicates that we need a radial order $I = 3$ to accurately estimate radial attenuation characterized by three crossing fibers.

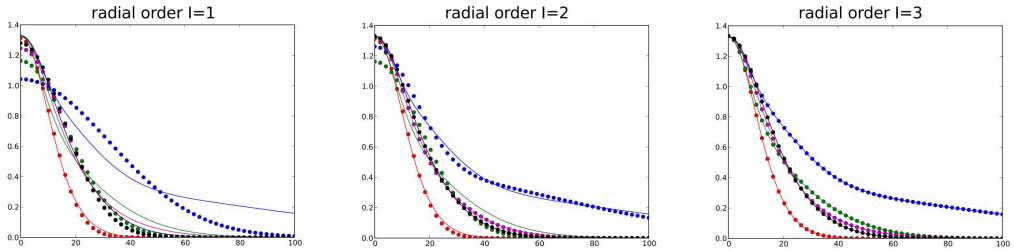


Figure 7.2: Examples of radial attenuation in arbitrarily chosen directions simulated using a multi-Gaussian model mimicking three crossing fibers (plain line). We fit each of these radial attenuation with the radial part of the model in (7.2) and radial orders $I = 1, 2, 3$, using the Levenberg-Marquardt algorithm (The angular order J is set to zero).

7.4.2 Training phase

We train the dictionary on multi-Gaussian signals, used for the HARDI contest at ISBI 2012 [2]. The contest was organized with the aim of providing a way for different groups to propose their own reconstruction algorithms and to fairly compare their methods against the others on a common set of ground-truth data.

Our training data set S_{train} is composed of the first $M = 5000$ instances of diffusion signal contained in the file *TestingIV*, in which the multi-Gaussian synthetic signals are generated with parameters taken at random (number of fibers, fractional anisotropy related to a fiber and crossing angle between these fibers). The dictionary angular and radial orders are respectively set to $J = 45$ and $I = 3$ (see Sec. 7.4.1). We take $m_s = 1000$ q-space samples for each instance of signal spread between $b_{min} = 0$ and $b_{max} = 10000\text{s/mm}^2$.

One difficulty in dictionary learning is the choice of the regularization parameter λ in (7.7). In order to assess λ, we use a cross validation (CV) procedure. For this purpose, we consider another set of signals S_{val}, called the validation data set, and composed of 1000 signals, which have not been used for training, and we repeat the following procedure for a range of λ,

1. Design a dictionary $\mathbf{D}_\lambda$ using Algorithm 1 with regularization parameter λ and the training data set S_{train}.
2. Using $\mathbf{D}_\lambda$, solve the LASSO problem for $\mathbf{c}$ (see (7.10) with S_{val} as entry, and compute the validation error $\epsilon_{\text{val}} = \|S_{\text{val}} - \mathbf{D}_\lambda \mathbf{c}\|_2^2 / \|S_{\text{val}}\|_2^2$

We keep λ that minimizes ϵ_{val}. This procedure is repeated after adding Rician noise, with SNR=10, 20 and 30, to the validation set. Rician noise is added in the following way : $S_{\text{val}_{\text{noisy}}} = \sqrt{(S_{\text{val}} + \epsilon_1)^2 + \epsilon_2^2}$, where $\epsilon_1, \epsilon_2 \sim \mathcal{N}(0, \sigma)$ with $\sigma = 1/SNR$. Validating λ on noisy data enables one to prevent from overfitting.

We show in Fig. 7.3, ϵ_{val} for λ in the range $[1 \cdot 10^{-5}, 1 \cdot 10^{-4}]$, in case of noisy and noiseless validation data. We also show the training error (blue curve in Fig. 7.3), i.e. $\epsilon_{\text{train}} = \|S_{\text{train}} - \mathbf{D}_\lambda \mathbf{c}\|_2^2 / \|S_{\text{train}}\|_2^2$ where $\mathbf{c}$ is the solution of the LASSO problem with S_{train} as entry. We observe, in Fig. 7.3, four different λ $(2.9 \cdot 10^{-5}, 4.4 \cdot 10^{-5}, 4.8 \cdot 10^{-5}, 5.8 \cdot 10^{-5})$ that minimize the validation error depending on the amount of noise we add to the validation data set. Because the noiseless case is not observed in practice, we discard the corresponding

[2] http://hardi.epfl.ch/

value of λ and consider the average value of $(4.4 \cdot 10^{-5}, 4.8 \cdot 10^{-5}, 5.8 \cdot 10^{-5})$, which gives $\lambda = 5 \cdot 10^{-5}$. This value is used to generate a new dictionary.

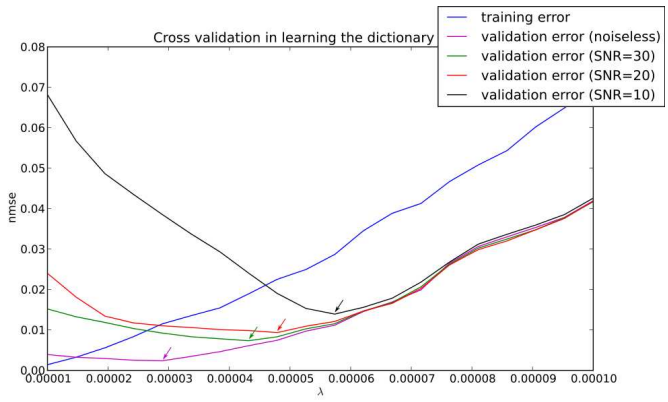


Figure 7.3: Training and validation error (ϵ_{train} and ϵ_{val}) computed for λ in the range $[1 \cdot 10^{-5}, 1 \cdot 10^{-4}]$. The blue curve represents the training error and the other curves represent the validation error in a noiseless case (purple curve), after adding Rician noise to the validation data set with SNR=30 (green curve), SNR=20 (red curve), SNR=10 (black curve). The arrows indicate the minimum of each curve corresponding to the validation error.

We obtain a dictionary containing 659 atoms. Fig. 7.4 shows the first 200 ODFs of these atoms. The atom ODFs are sorted in decreasing energy order from left to right and top to bottom. We observe various shape ranging from single fiber structures to more complex fiber configurations.

7.4.3 Validation

We validate the dictionary on the reconstruction of noisy multi-Gaussian signals, used for the HARDI contest at ISBI 2012 [3].

We consider 1000 signals, which have not been used for training the dictionary. Our preliminary and promising results on parametric dictionary learning (PDL) were published in the proceedings of this event [87]. At this stage, we obtained the best results in our category. Here, we enrich these previous results with a comparison of three different sampling schemes (displayed in Fig. 7.5) :

- A single shell sampling scheme with 64 measurements uniformly spread on one shell at a b-value $b = 3000 \text{ s} \cdot \text{mm}^{-2}$ (Fig. 7.5a).
- A multiple shells sampling scheme with 15 measurements spread on 2 shells at b-values $b = 1500, 2500 \text{ s} \cdot \text{mm}^{-2}$ (Fig. 7.5b).
- A multiple shells sampling scheme with 64 measurements spread on 2 shells at b-values $b = 1500, 2500 \text{ s} \cdot \text{mm}^{-2}$ (Fig. 7.5c).

[3] http://hardi.epfl.ch/

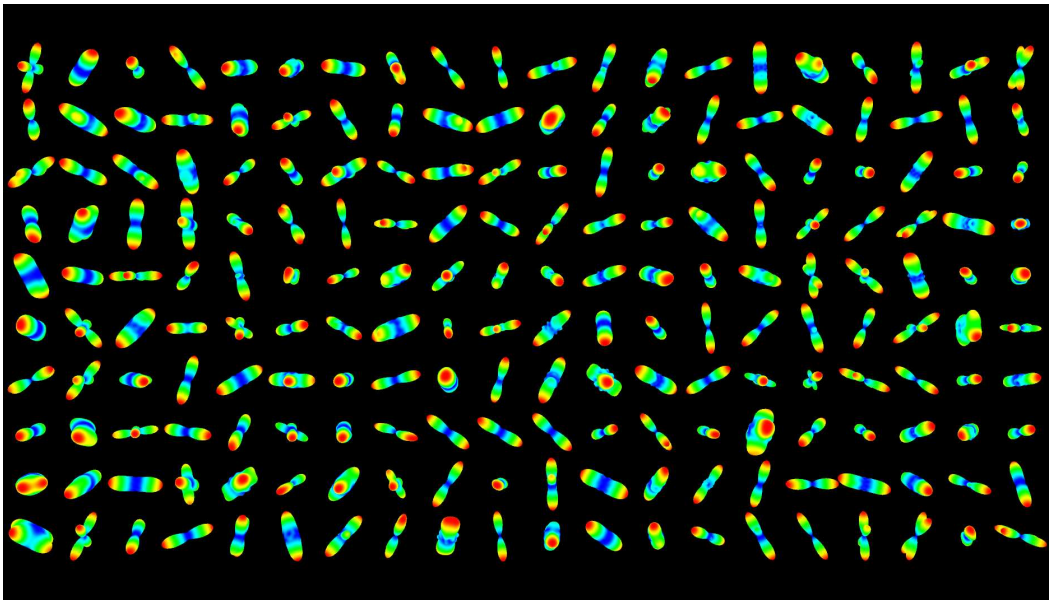


Figure 7.4: First 200 ODFs of the dictionary atoms. The atoms are sorted in decreasing energy order from left to right and top to bottom.

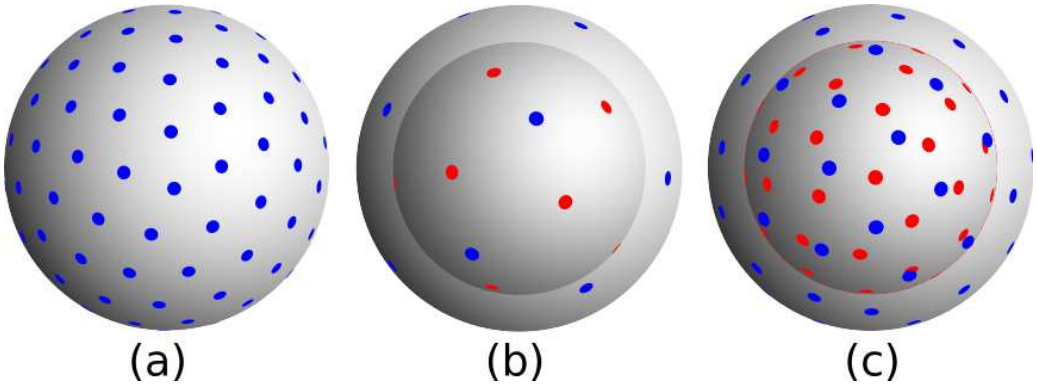


Figure 7.5: Sampling schemes used for validation. (a) A single shell sampling scheme with 64 measurements uniformly spread on one shell at a b-value $b = 3000\ \mathrm{s \cdot mm^{-2}}$. (b) A multiple shells sampling scheme with 15 measurements spread on 2 shells at b-values $b = 1500, 2500\ \mathrm{s \cdot mm^{-2}}$. (c) A multiple shells sampling scheme with 64 measurements spread on 2 shells at b-values $b = 1500, 2500\ \mathrm{s \cdot mm^{-2}}$.

To obtain the single shell (SS) sampling scheme, we use the algorithm given in [68, 41] to uniformly distribute points on a sphere. For the multiple shells (MS) sampling schemes, we use the algorithm given in [26] by setting the parameters in such a way that the the number of points on each shell is proportional to q^1. These particular parameters have proven efficient in [26, 85]. An important advantage of this algorithm is that the points from each shell have staggered directions and follow a near-optimal uniform distribution. You can generate and download sampling schemes for multiple Q-shell diffusion MRI with this web application : *http://www-sop.inria.fr/members/Emmanuel.Caruyer/q-space-sampling.php*.

We perform the experiments with two metrics used in the contest, i.e. the weighted difference in the number of fiber compartments (DNC) and the mean angular error (AE) at each voxel. For these two metrics we extract the maxima on the estimated ODFs and compare them to the ground truth maxima. Then, the DNC becomes the difference between the number of maxima extracted on the estimated ODFs M_e and the true number of maxima M_{gt}, weighted by the true number of maxima at each voxel, i.e., DNC $= \frac{|M_e - M_{gt}|}{M_{gt}}$. The AE is the mean angular error between the maxima extracted on the estimated ODFs and the respective maxima within the ground truth. We also enrich the results with a comparison of two other features, which have not been used in the contest validation, i.e. the diffusion signal and the Ensemble Average Propagator (EAP). To compare these features, we compute the Normalized Mean Square Error (NMSE) between the ground truth feature x and its estimation x_e given by $NMSE = \frac{\|x - x_e\|_2^2}{\|x\|_2^2}$. The DNC, AE and NMSE are, then, averaged on all the voxels. We add Rician noise to the normalized diffusion signal in the following way : $E_{\text{noisy}} = \sqrt{(E + \epsilon_1)^2 + \epsilon_2^2}$, where $\epsilon_1, \epsilon_2 \sim \mathcal{N}(0, \sigma)$ with $\sigma = 1/SNR$.

The three following sections present the results for the three proposed sampling schemes, i.e. the SS sampling scheme with 64 measurements, the MS sampling scheme with 15 measurements, and the MS sampling scheme with 64 measurements. We also compare with state of the art techniques such that QBI, using the solid angle ODF [1], and the SHORE reconstruction using a sparse prior (ℓ_1-SHORE) [99, 35, 90, 92, 86]. We show quantitative results in Tab. 7.1, 7.2, 7.3 and 7.4, and qualitative results in Fig. 7.6, 7.7 and 7.8. In each table, we write in blue letters the best score for a given SNR and metric.

Single shell sampling scheme with 64 measurements

Single shell sampling scheme with 64 measurements	Angular error		DNC	
	D-ODF	SA-ODF	D-ODF	SA-ODF
SNR 30	4.9398	6.6509	0.2068	0.2329
SNR 20	5.6386	7.1436	0.2102	0.2524
SNR 10	8.2530	12.419	0.2500	0.5993

Table 7.1: Dictionary based ODF estimation (D-ODF) versus solid angle ODF via QBI (SA-ODF) using a single shell sampling scheme with 64 measurements. We added rician noise from SNR=10 to 30. Two metrics are shown : the angular error and the difference in the number of compartments (DNC).

In these experiments, we first use the SS sampling scheme, i.e. 64 measurements uniformly spread on a shell at a b-value $b = 3000$ s $\cdot$ mm^{-2}, and we compare our dictionary

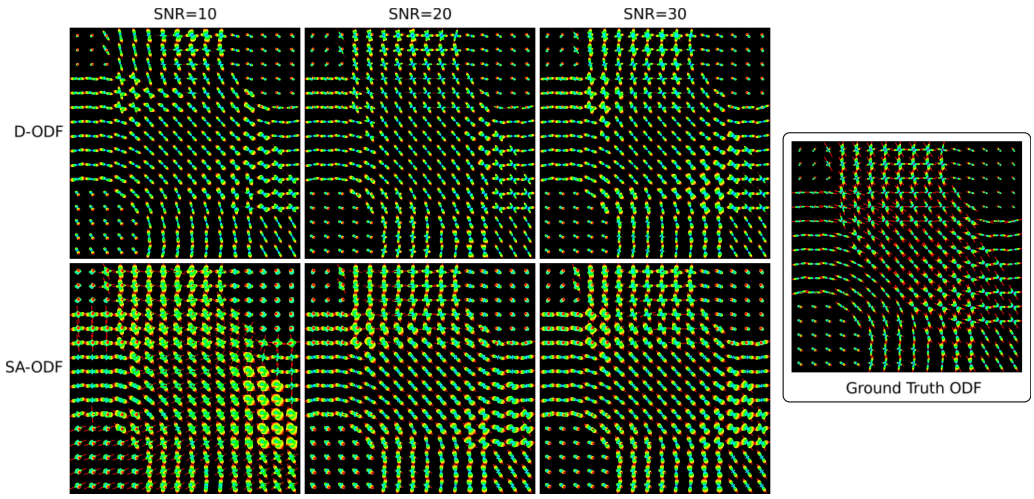


Figure 7.6: Dictionary based ODF estimation (D-ODF,top) versus solid angle ODF via QBI (SA-ODF,bottom). We added rician noise from SNR=10 to 30 (left to right). We also show the maxima extracted from the ODFs. On the right, we show the ground ODFs and their respective maxima.

based ODF estimation (D-ODF) to the solid angle ODF via QBI (SA-ODF) developed in [1]. For the SA-ODF, we set a SH order equal to the one used for the dictionary construction, i.e. $L = 8$. We adjust the Laplace-Beltrami regularization parameter [44, 1] using the generalized cross validation algorithm [40]. For our PDL approach, because we deal with ℓ_1 norm and not ℓ_2 norm, we use a simple cross validation procedure [129] to find the regularization parameter λ in (7.10).

Overall, in Fig. 7.6 the D-ODFs are sharper than the SA-ODFs. Furthermore, the SA-ODFs appear more sensitive to noise than the ODFs based on our PDL estimation. Indeed, we observe that the D-ODFs are very robust to noise, even at SNR=10 where they are still correctly aligned with the underlying structure shown by the ground truth (On the right of Fig. 7.6), whereas the maxima extracted from the SA-ODFs mostly give corrupted or completely false fiber orientation estimation even for simple configuration as single fibers.

The quantitative results, in Tab. 7.1, confirms our previous remarks. In particular, at SNR=10, the SA-ODFs are not able to provide proper diffusion directions. The DNC mean value is higher than 0.5 meaning that, in average, more than half of the maxima in each voxel are not detected. The results regarding our PDL approach give more accurate diffusion directions at every SNR.

Besides the good directional information given by the ODFs estimated with our PDL approach, we can also estimate the EAP and interpolate/extrapolate the diffusion signal on the entire q-space, whereas QBI only estimates the ODFs. More results regarding the estimation of the latter features are given in Tab. 7.4, and we will discuss these results in Sec. 7.4.3.

MS sampling scheme with 15 measurements	Angular error		DNC	
	Dictionary	ℓ_1-SHORE	Dictionary	ℓ_1-SHORE
SNR 30	8.6066	14.670	0.2472	0.4010
SNR 20	9.7626	16.313	0.2540	0.4463
SNR 10	13.344	22.354	0.2734	0.4836
	Signal NMSE		EAP NMSE	
	Dictionary	ℓ_1-SHORE	Dictionary	ℓ_1-SHORE
SNR 30	0.0104	0.0433	0.0176	0.1040
SNR 20	0.0170	0.0578	0.0251	0.1122
SNR 10	0.0343	0.1027	0.0422	0.1350

Table 7.2: Dictionary based reconstruction versus ℓ_1-SHORE based reconstruction using a multiple shells sampling scheme with 15 measurements. We added Rician noise from SNR=10 to 30. Four metrics are shown : the angular error, the difference in the number of compartments (DNC), the signal NMSE and the EAP NMSE.

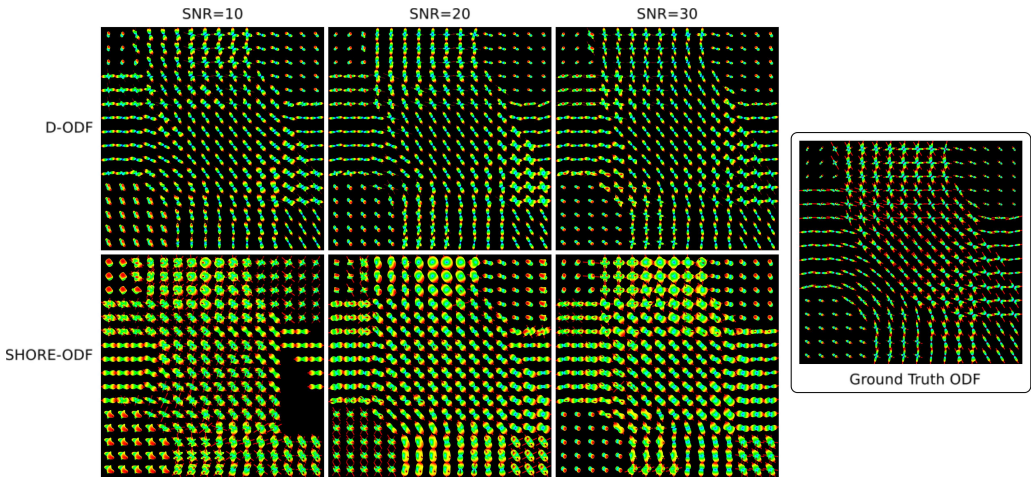


Figure 7.7: Dictionary based ODF estimation (D-ODF, top) versus ℓ_1-SHORE based ODF estimation (SHORE-ODF, bottom). We added Rician noise from SNR=10 to 30 (left to right). We also show the maxima extracted from the ODF. On the right, we show the ground ODFs and their respective maxima.

Multiple shells sampling scheme with 15 measurements

In these experiments, we compare a SHORE reconstruction using a sparse prior (ℓ_1-SHORE) to our dictionary reconstruction. The SHORE basis has been introduced by [99] and was used in [90, 92, 86] in the context of sparse recovery. The ℓ_1-SHORE method consists in solving the LASSO problem (see (7.10)) using coordinate descent while replacing the dictionary $\mathbf{D}$ by the SHORE basis [99]. In order to provide a fair comparison a SH order of $L = 8$ is used for the generation of the SHORE basis. In both methods (ℓ_1-SHORE and our PDL approach) we use cross validation [129] to assess the regularization parameter λ in (7.10). [99] and [35] respectively provide closed form formulae to estimate the ODF and the EAP when the diffusion signal is modeled in the SHORE basis.

We see in Tab. 7.2 that our PDL approach outperforms the ℓ_1-SHORE reconstruction in terms of angular error, difference in the number of compartments (DNC), signal NMSE and EAP NMSE.

Fig. 7.7 shows the ODFs estimated via our dictionary approach (D-ODF) and the ODFs estimated via the ℓ_1-SHORE method (SHORE-ODF). It qualitatively indicates an improvement of the angular information given by the D-ODFs over the SHORE-ODFs.

Regarding the PDL approach proposed at the ISBI contest, we obtained an angular error equal to 14.5° at SNR=10, equal to 11° at SNR=20, equal to 9.5° at SNR=30. The results based on our new proposed framework, shown in Tab. 7.2, indicates an improvement on the accuracy of the maxima estimation compared to the results obtained in the contest. Note that the PDL approach proposed in the contest was already the best in its category.

In the following we show the impact of an increase of the number of samples while keeping the two shells at b-values $b = 1500, 2500\ \mathrm{s \cdot mm^{-2}}$.

Multiple shells sampling scheme with 64 measurements

MS sampling scheme with 64 measurements	Angular error		DNC	
	Dictionary	ℓ_1-SHORE	Dictionary	ℓ_1-SHORE
SNR 30	5.6233	8.8950	0.2187	0.3106
SNR 20	6.3080	9.6641	0.2309	0.3401
SNR 10	8.3224	13.126	0.2511	0.3995
	Signal NMSE		EAP NMSE	
	Dictionary	ℓ_1-SHORE	Dictionary	ℓ_1-SHORE
SNR 30	0.0035	0.0368	0.0066	0.0732
SNR 20	0.0054	0.0386	0.0104	0.0746
SNR 10	0.0144	0.0752	0.0222	0.0825

Table 7.3: Dictionary based reconstruction versus ℓ_1-SHORE based reconstruction. We added Rician noise from SNR=10 to 30. Four metrics are shown : the angular error, the difference in the number of compartments (DNC), the signal NMSE and the EAP NMSE.

Now, we use the MS sampling scheme with 64 measurements and we compare once again the ℓ_1-SHORE method to our PDL approach. Our dictionary approach still outperforms the SHORE reconstruction in terms of angular error, DNC, signal NMSE and EAP

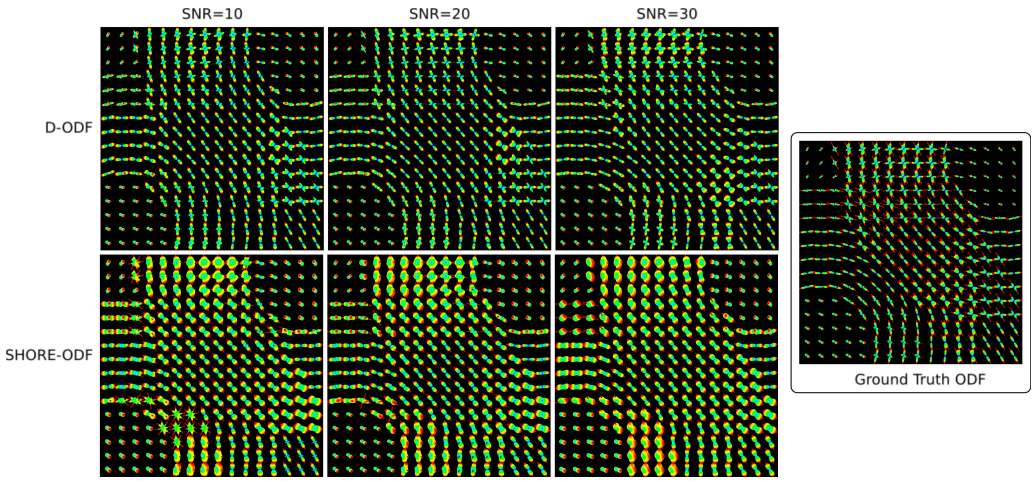


Figure 7.8: Dictionary based ODF estimation (D-ODF, top) versus ℓ_1-SHORE based ODF estimation (SHORE-ODF, bottom) using a multiple shells sampling scheme with 64 measurements. We added Rician noise from SNR=10 to 30 (left to right). We also show the maxima extracted from the ODFs. On the right, we show the ground ODFs and their respective maxima.

NMSE. Fig. 7.8 shows the ODFs estimated via our PDL approach (D-ODF) and the ODFs estimated via the ℓ_1-SHORE method (SHORE-ODF). We observe that the D-ODFs give a very accurate estimation of the underlying fiber structure where the SHORE-ODFs fail to provide coherent fiber direction estimation, especially at SNR=10. We confirm this by the quantitative results shown in Tab. 7.3.

We also compare the results of our PDL approach using this MS sampling scheme , i.e. 2 shells at b-values $b = 1500, 2500$ s $\cdot$ mm^{-2} with 64 measurements, to our PDL approach using the SS sampling scheme studied in Sec. 7.4.3, i.e. one shell at b-value $b = 3000$ s $\cdot$ mm^{-2} with 64 measurements. All the results are given in Tab. 7.4. Regarding the directional results (angular error and DNC), we observe a slight advantage with the SS sampling scheme. Moreover, if we want to reconstruct the full diffusion signal and the EAP , we notice that a MS sampling scheme is more adequate. This is because the radial information of the diffusion process is better considered when using 2 shells instead of only one shell.

7.4.4 Discussion on experiments with synthetic data

We studied the choice of the free parameters in the learning process, i.e. the dictionary angular and radial orders (in sec. 7.4.2), and λ in (7.7) (in Sec. 7.4.1). Our study led to a very good estimation of the diffusion direction (via the computation of the ODF), the diffusion signal and the EAP.

In particular, our PDL approach was shown to better estimate the diffusion directions than the solid angle ODF via QBI does and to compute the diffusion signal and the EAP in a more accurate way than a SHORE reconstruction (using a sparse prior) does.

Regarding the sampling scheme, we observe a slight advantage for the SS sampling

	Angular error		DNC		Signal NMSE		EAP NMSE	
	MS-SC	SS-SC	MS-SC	SS-SC	MS-SC	SS-SC	MS-SC	SS-SC
SNR 30	5.6233	4.9398	0.2187	0.2068	0.0035	0.0207	0.0066	0.0298
SNR 20	6.3080	5.6386	0.2309	0.2102	0.0054	0.0239	0.0104	0.0321
SNR 10	8.3224	8.2530	0.2511	0.2500	0.0144	0.0433	0.0222	0.0526

Table 7.4: Comparison between the multiple shells sampling scheme (MS-SC) with 64 measurements and the single shell sampling scheme (SS-SC) with 64 measurements, on reconstruction using the PDL approach. Four metrics are shown: the angular error, the difference in the number of compartments (DNC), the signal NMSE, the EAP NMSE. The simulations are performed with Rician noise at SNR=10, 20 and 30.

scheme considering the directional features, but we found that the MS sampling at b-values $b = 1500, 2500$ s $\cdot$ mm^{-2} is more appropriate, because it enables one to deal with the radial information in addition to the angular information. In Sec. 7.4.3 the dictionary reconstruction combined with a multiple shells sampling scheme was proved to efficiently approximate the diffusion signal, the EAP and the ODF.

Regarding the minimal number of measurements required before a large decrease of the reconstruction accuracy, we also perform a last experiment on synthetic data, in which we study the impact of the number of samples N on the error metrics presented in this experiment part with synthetic data. For this purpose, we still consider the same set of signals (different from the training data set) and add Rician noise with $SNR = 20$. We use the MS sampling scheme with 2 shells at b-values $b = 1500, 2500$ s $\cdot$ mm^{-2} and vary the number of samples between $N = 5$ and $N = 100$. Fig. 7.9 shows the resulting values. In this figure, we also plot a vertical line, which represents the number of samples N where the metric errors show a large increase of their values (which means that the estimations are not correct anymore). Overall, we define this bound to $N_{min} \approx 15$. In Fig. 7.9, we also represent the variance of each metric error. We observe from the variance, that the estimation of the diffusion signal and the EAP are robust to noise. However, the estimation of the diffusion direction is more sensitive to noise when we reduce the number of samples.

In conclusion, these synthetic experiments show that our PDL approach can sparsely model multi-fiber compartments signals with the assumption of mono-exponential signal decay with b-value. In addition, it overcomes the preliminary PDL approach presented at the HARDI contest at ISBI 2012, already ranked first in its category.

7.5 Experiments on real data

In this section, we propose to validate our parametric dictionary learning (PDL) method on real data from human brains. For this purpose, we acquired three distinct sets of data :

- A first set of measurements coming from a 7T scanner, used both to learn the dictionary and to validate it.
- A second set of measurements coming from a 3T scanner, used for the learning process.

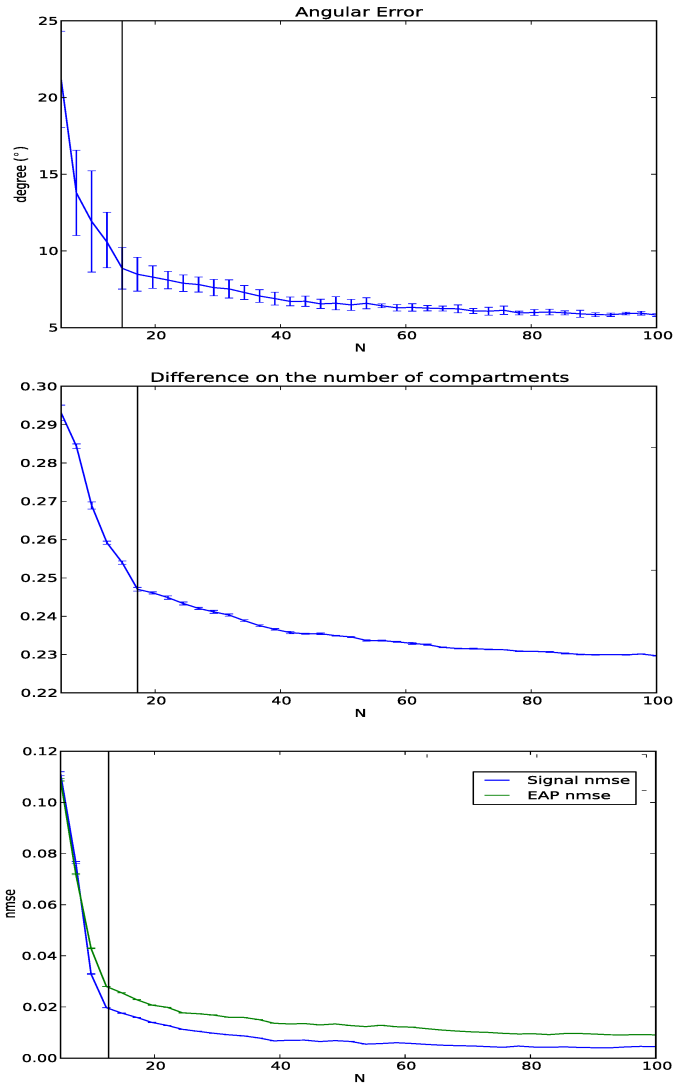


Figure 7.9: Evolution of the angular error, the DNC, the signal NMSE and the EAP NMSE in function of the number of samples N. We added Rician noise with SNR=20.

- A third set of measurements coming from a 3T scanner, used to validate the dictionary learned on the previous 3T scanner data.

7.5.1 Learning and reconstruction on a 7T scanner data

Training data was acquired on a 7T whole-body MR scanner (MAGNETOM, Siemens Healthcare) equipped with Siemens-AC072 whole body gradient coils, and an 24 channel phased array coil (Nova Medical). 12 axial slices were acquired with a 2D single shot DW-STEAM-EPI (TR/TE/Δ/δ = 3000/58/120/15 ms) sequence giving 2 mm isotropic resolution. The echo time (TE), diffusion time (Δ), and gradient duration (δ) were optimized to provide maximum SNR for a maximum b-factor. 8 different b-values b=500/1000/2000/3000/4000/5000/6000/7000 s/mm^2 are considered with 70 orientations at each b-value, and an imaging matrix of 96x96x12. The measurement locations are distributed using the algorithm given in [26] by setting the parameters in such way that there are a constant number of measurements per b-value. We use 11 slices as training data set, in which only the voxels corresponding to a Fractional Anisotropy (FA) superior to 0.20 are taken into account. The FA is computed from the diffusion tensor estimated with the whole set of measurements.

We validate the reconstruction, based on the previously learned dictionary, on the twelfth slice. We compare the ODFs estimated via our PDL approach (D-ODF) to the ODFs estimated with the ℓ_1-SHORE method (SHORE-ODF) on three different sampling schemes. The three sampling schemes consider measurements at b-values b=1000/2000 s/mm^2, and we change the number of measurements as N=15, 30 and 60. To distribute the samples between the two b-values, we use the algorithm given in [26] by setting the parameters in such a way that the the number of samples on each shell is proportional to q^1 [26, 85].

We choose a region of interest and show the estimated ODFs (see Fig. 7.10) along with the extracted maxima. This region contains several crossing configurations with different degree of crossing, and thus is appropriate for ODF validation. In Fig. 7.10, the middle corresponds to the D-ODFs and the bottom to the SHORE-ODFs. From the left to the right, we see the results for N=15, 30 and 60 samples. We consider the ground truth as the estimated signal using the ℓ_1-SHORE method when all the measurements are taken into account, i.e. 70x8 measurements. The corresponding ground truth ODF are shown at the top of Fig. 7.10.

Overall, the SHORE-ODFs lead to more false maxima than the D-ODFs. For instance, at $N = 60$ in the region A (in red), the SHORE-ODFs are not able to correctly resolve the crossing fiber configuration (erroneous number of detected maxima). This phenomenon is emphasized when the number of measurements decreases, whereas our dictionary estimation still provides a coherent map of ODFs.

7.5.2 Learning and reconstruction on a 3T scanner data

We also train our dictionary on data from a 3T Verio (MAGNETOM, Siemens Healthcare) scanner equipped with a 32-channel head coil. The data were acquired at a spatial resolution of 2 mm^3 isotropic, for 6 different b-values b=500/1000/2000/3000/5000/7000 s/mm^2, 70 orientations at each b-value, and an imaging matrix of 128x128x60. The dictionary is

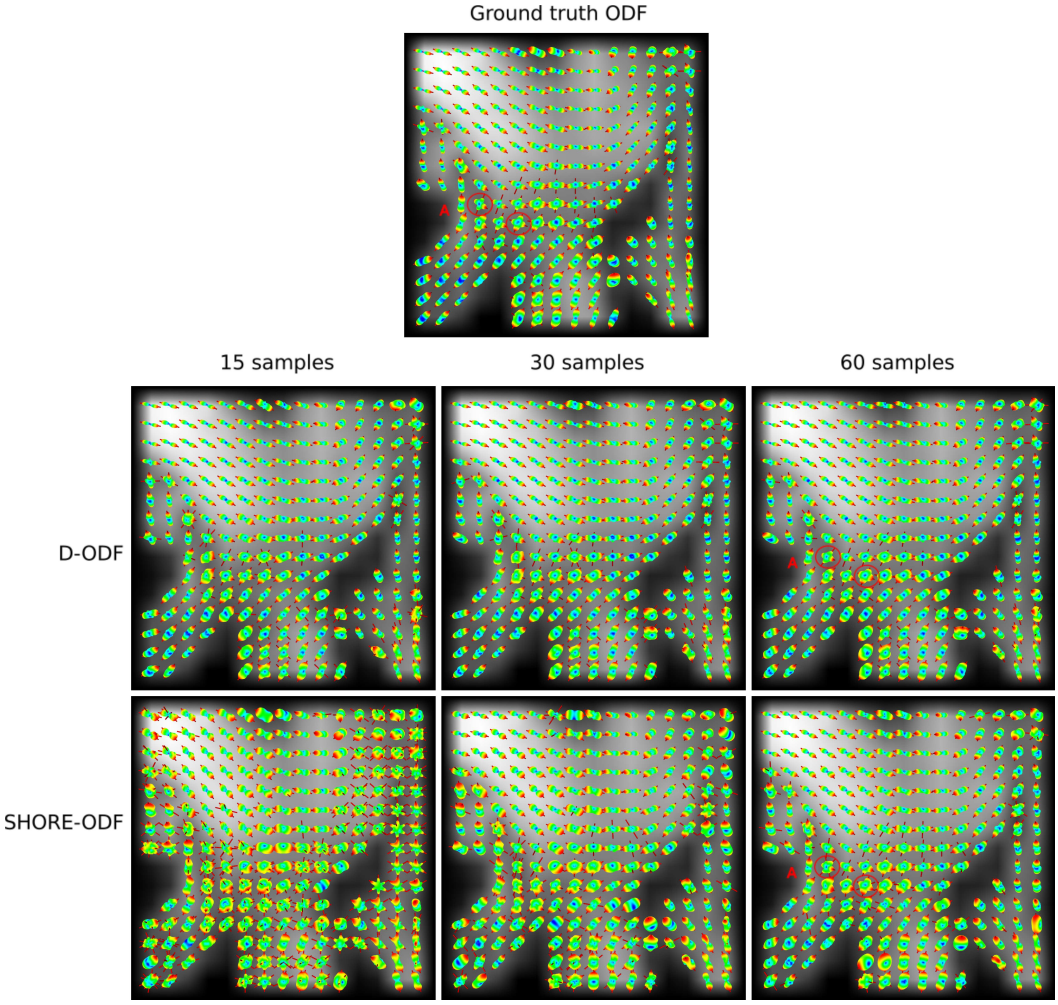


Figure 7.10: ODFs estimated from a 7T scanner data via our PDL approach (D-ODF, middle) and via the SHORE technique (SHORE-ODF, bottom). N=15, 30 and 60 samples are considered (left to right). We show the ground truth ODF at the top.

learned from all the measurements on the axial slices 25 to 35, for the voxels with $FA \geq 0.20$. The FA is computed from the diffusion tensor estimated with the whole set of measurements.

For the reconstruction, we use data from a 3T scanner (Philips Achieva) equipped with a 8-channel SENSE coil. The data were acquired at a spatial resolution of 2 mm^3 isotropic, for 6 different b-values b=500/1000/2000/4000/6000/8000 s/mm^2, 70 orientations at each b-value, and an imaging matrix of 128x128x60. We consider three sampling schemes with N=15, 30 and 60 samples and distribute them proportionally to q^1 on two b-values b=1000/2000 s/mm^2 [26, 85].

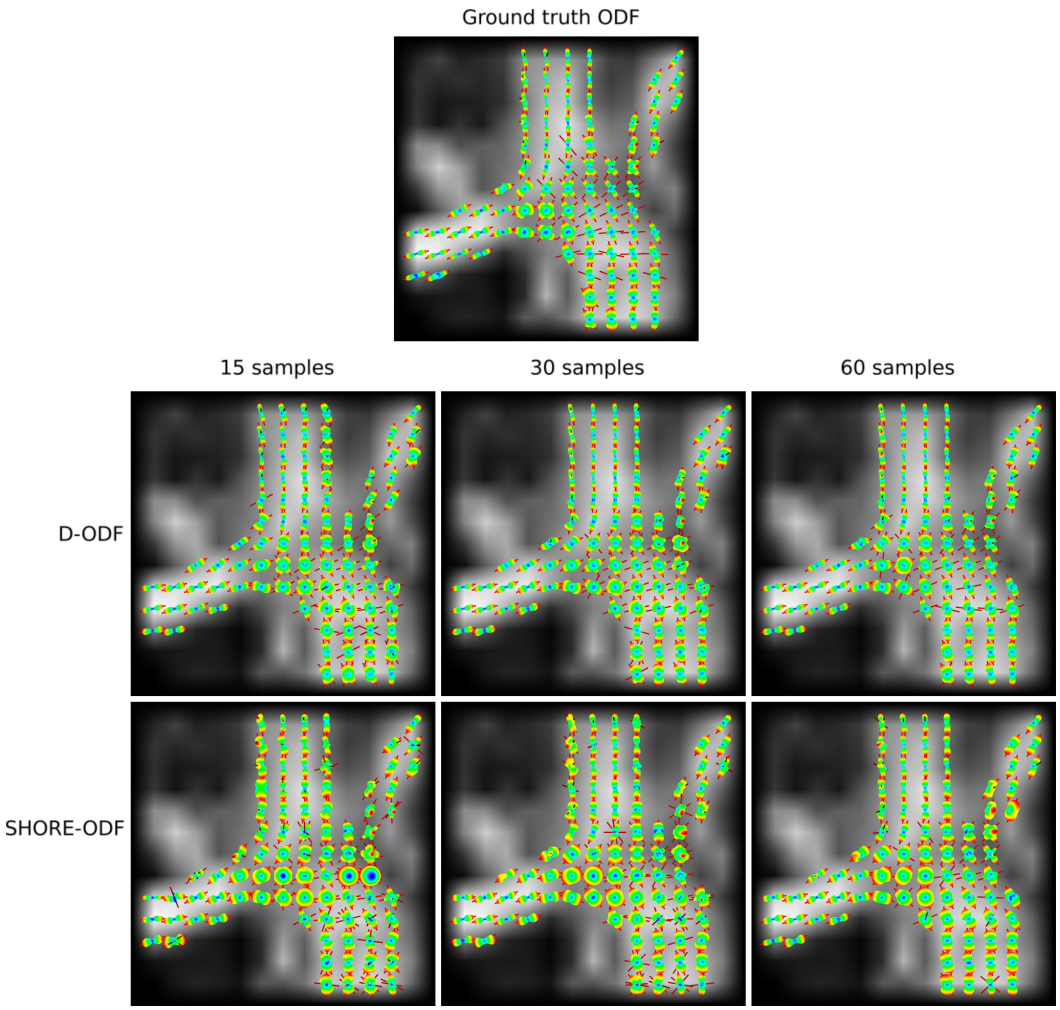


Figure 7.11: ODFs estimated from a 3T scanner data via our PDL approach (D-ODF, middle) and via the SHORE technique (SHORE-ODF, bottom). N=15, 30 and 60 samples are considered (left to right). We show the ground truth ODF at the top.

In Fig. 7.11, the top corresponds to the D-ODFs and the bottom to the SHORE-ODFs, on a selected region of interest. From the left to the right, we see the results for N=15, 30 and 60 samples. We also consider the ground truth as the estimated signal using the ℓ_1-SHORE method when all the measurements are taken into account, i.e. 70x6 measurements.

The corresponding ground truth ODFs are shown at the top of Fig. 7.11.

We obtain sharper ODFs estimation with our PDL approach on the selected region of interest. Again, we see that our dictionary is still able to model the fiber crossing configuration with only 15 samples, and provide less noisy ODFs than the ODFs estimated with SHORE.

7.6 Conclusion

We have proposed an original and efficient computational framework to model continuous diffusion MRI (dMRI) signal and to recover its important features such as the ODF and the EAP with a reduced number of measurements. The idea, we implemented, has been to use a parametric dictionary learning algorithm and to exploit the sparse property of a well designed dictionary to recover the diffusion signal and its features. Numerous experimental results have been carried out for validation on synthetic and human brain data acquired from 7-T and 3-T scanners. We have shown that we can clearly recover the diffusion signals and its features with a much better accuracy compared to state-of-the-art approaches and can accurately recover ODF in regions of multiple fiber crossing, even with a small and reduced number of measurements. This opens new perspectives for some dMRI applications, including for example tractography, where the improved characterization of the fiber orientations is likely to greatly and quickly help tracking through regions with and/or without crossing fibers

7.7 Appendix

7.7.1 Derivation of the normalization constant

We define a dictionary of M functions f_m, from $\mathbb{R}^3$ to $\mathbb{R}$. We want to normalize the atoms of the dictionary, for the classical ℓ_2 norm:

$$||f_m||_2^2 = \int_{\mathbb{R}^3} f_m(\mathbf{q})^2 \mathrm{d}\mathbf{q} = 1. \tag{7.11}$$

The functions f_m are constructed from elementary functions $g_{i,j}^{(m)}$ (in what follows, we drop the index m for the sake of clarity):

$$f_m = \sum_{i=1}^{I} \sum_{j=1}^{J} g_{i,j}(\mathbf{q}), \tag{7.12}$$

where

$$g_{i,j}(q \cdot \mathbf{u}) = \gamma_{i,j} e^{-\nu_i q^2} q^{\ell(j)} Y_j(\mathbf{u}). \tag{7.13}$$

The normalization in (7.11) rewrites as

$$\begin{aligned}1 &= \int_{\mathbb{R}^3} \left(\sum_{i,j} g_{i,j}(\mathbf{q}) \right)^2 \mathrm{d}\mathbf{q} \\ &= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{i'=1}^{I} \sum_{j'=1}^{J} \gamma_{i,j}\gamma_{i',j'} \int_0^\infty e^{-\nu_i q^2} e^{-\nu_{i'} q^2} q^{\ell(j)} q^{\ell(j')} q^2 \mathrm{d}q \\ &\qquad \int_{\mathcal{S}^2} Y_j(\mathbf{u}) Y_{j'}(\mathbf{u}) \mathrm{d}^2\mathbf{u} \end{aligned} \tag{7.14}$$

Provided that the spherical harmonics functions form an orthonormal basis, we have

$$\int_{\mathcal{S}^2} Y_j(\mathbf{u}) Y_{j'}(\mathbf{u}) \mathrm{d}^2\mathbf{u} = \delta_{j,j'}. \tag{7.15}$$

Therefore, the normality constraint rewrites as

$$1 = \sum_{i=1}^{N} \sum_{j=1}^{I} \sum_{i'=1}^{I} \gamma_{i,j}\gamma_{i',j} \int_0^\infty e^{-(\nu_i+\nu_{i'})q^2} q^{2\ell(j)+2} \mathrm{d}q. \tag{7.16}$$

Let's use the substitution $x = (u_i + u_{i'})q^2$ in the above integral, we have

$$1 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{i'=1}^{I} \frac{\gamma_{i,j}\gamma_{i',j}}{2(\nu_i+\nu_{i'})^{\ell(j)+3/2}} \int_0^\infty e^{-x} x^{\ell(j)+1/2} \mathrm{d}x \tag{7.17}$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{i'=1}^{I} \frac{\gamma_{i,j}\gamma_{i',j}}{2(\nu_i+\nu_{i'})^{\ell(j)+3/2}} \Gamma\left(\ell(j) + \frac{3}{2}\right) \tag{7.18}$$

7.7.2 Derivation of the Ensemble Average Propagator

The EAP $P(R\mathbf{r})$ is defined as

$$P(R\mathbf{r}) = \int_{q=0}^{\infty} \int_{\mathbf{u}\in S^2} E(q\mathbf{u}) \exp(+2\pi i q R \mathbf{u}.\mathbf{r}) d\mathbf{u} q^2 dq. \tag{7.19}$$

We use the spherical plane wave expansion

$$\exp(\pm 2\pi i q R \mathbf{u}.\mathbf{r}) = \frac{2\pi}{\sqrt{qR}} \sum_{j=0}^{\infty} (\pm i)^{\ell(j)} J_{\ell+1/2}(2\pi q R) Y_j(\mathbf{u}) Y_j(\mathbf{r}), \tag{7.20}$$

where $J_{\ell+1/2}$ is the standard bessel function of the first kind and order $\ell + 1/2$. Then we get

$$P(R\mathbf{r}) = \int_{q=0}^{\infty}\int_{\mathbf{u}\in S^2}\left[q^2\sum_{k=0}^{K}c_k\frac{1}{\sqrt{\chi_k}}\sum_{i=0}^{I}\sum_{j=0}^{J}\gamma_{kij}\exp\left(-\nu_{ki}q^2\right)q^{\ell(j)}Y_j(\mathbf{u})\right]\left[\frac{2\pi}{\sqrt{qR}}\sum_{j'=0}^{\infty}(\pm i)^{\ell(j')}j_l(2\pi qR)Y_{j'}(\mathbf{u})Y_{j'}(\mathbf{r})\right]d\mathbf{u}dq, \tag{7.21}$$

We shorten this expression using the orthonormal property of the spherical harmonic basis, i.e. $\int_{\mathbf{u}\in S^2}Y_j(\mathbf{u})Y_{j'}(\mathbf{u})d\mathbf{u} = \delta_{jj'}$. Then formula (7.21) becomes

$$P(R\mathbf{r}) = \sum_{k=0}^{K}c_k\frac{2\pi}{\sqrt{R\chi_k}}\sum_{i=0}^{I}\sum_{j=0}^{J}\gamma_{kij}(-1)^{\ell(j)/2}Y_j(\mathbf{r})\underbrace{\int_{q=0}^{\infty}\exp\left(-\nu_{ki}q^2\right)q^{\ell(j)+3/2}J_{\ell(j)+1/2}(2\pi qR)dq}_{I_{kij}(R)}, \tag{7.22}$$

We use the formula from [111], i.e. $\int_{q=0}^{\infty}x^{\nu+1}\exp(-\alpha x^2)J_\nu(\beta x) = \frac{\beta^\nu}{(2\alpha)^{\nu+1}}\exp(\frac{-\beta^2}{4\alpha})$

$$I_{kij}(R) = \int_{q=0}^{\infty}\exp\left(-\nu_{ki}q^2\right)q^{\ell(j)+3/2}J_{\ell(j)+1/2}(2\pi qR)dq \tag{7.23}$$

$$= \frac{(2\pi R)^{\ell+1/2}}{(2\nu_{ki})^{\ell+3/2}}\exp\left(\frac{-(2\pi R)^2}{4\nu_{ki}}\right) \tag{7.24}$$

Finally, we get a closed form for the propagator :

$$P(R\cdot\mathbf{r}) = \sum_{k=0}^{K}\frac{c_k}{\sqrt{\chi_k}}\sum_{i=0}^{I}\sum_{j=0}^{J}\gamma_{kij}(-1)^{\ell(j)/2}\left(\frac{\pi}{\nu_{ki}}\right)^{\ell(j)+3/2}R^{\ell(j)}\exp\left(\frac{-(\pi R)^2}{\nu_{ki}}\right)Y_j(\mathbf{r}) \tag{7.25}$$

7.7.3 Derivation of the Orientation Distribution Function

The ODF is given by

$$\Upsilon(\mathbf{r}) = \int_0^{\infty}P(R\cdot\mathbf{r})R^2dR. \tag{7.26}$$

We insert (7.5) in (7.26) and gather all the R-dependant terms within the integral to get

$$\Upsilon(\mathbf{r}) = \sum_{k=0}^{K} c_k \frac{(2\pi)^{3/2}}{\sqrt{\chi k}} \sum_{i=0}^{I} \sum_{j=0}^{J} \gamma_{kij} (-1)^{\ell(j)/2} \left(\frac{\pi}{\nu_{ki}}\right)^{\ell(j)+3/2} Y_j(\mathbf{r}) \underbrace{\int_0^{\infty} R^{\ell(j)+2} \exp\left(\frac{-(\pi R)^2}{\nu_{ki}}\right) dR}_{I(\nu_{ki},\ell(j))} \tag{7.27}$$

We use the formula from [111], i.e. $\int_0^{\infty} x^{2n} \exp(-\rho x^2) dx = \frac{(2n-1)!!}{2(2\rho)^n} \sqrt{\frac{\pi}{\rho}}$. Then,

$$I(\nu_{ki}, \ell(j)) = \frac{(\ell(j)+1)!!}{2\left(\frac{2\pi^2}{v_{ki}}\right)^{\ell(j)/2+1}} \sqrt{\frac{v_{ki}}{\pi}} \tag{7.28}$$

And,

$$\Upsilon(\mathbf{r}) = \sum_{k=0}^{K} \frac{c_k}{\sqrt{\chi k}} \sum_{i=0}^{I} \sum_{j=0}^{J} \gamma_{kij} (-1)^{\ell(j)/2} \left(\frac{\pi}{\nu_{ki}}\right)^{\ell(j)+1} \frac{(\ell(j)+1)!!}{2\left(\frac{2\pi^2}{v_{ki}}\right)^{\ell(j)/2+1}} Y_j(\mathbf{r}) \tag{7.29}$$

Part IV

Conclusion

Chapter 8

Conclusions and Perspectives

Contents

8.1 General conclusion 152
8.1.1 Local diffusion estimation 152
8.1.2 The multiple aspects of the Compressive Sensing theory 152
8.2 Applications and Collaborations 152
8.2.1 Fiber tractography via the Ensemble Average Propagator 153
8.2.2 Software 153

8.1 General conclusion

We have presented in this thesis several contributions related to the estimation of local diffusion information from the dMRI signal in PGSE experiments. In particular, this work is focused on the estimation of the Ensemble Average Propagator (EAP). The underlying theories and methods used in this thesis are related to the field of sparse coding techniques and in particular the Compressive Sensing (CS) theory. This work brings a new hope to estimate the EAP with a reduced number of samples. Other diffusion features are also studied to achieve the validation of the MR signal reconstruction. Finally, we have rendered possible the estimation of the EAP and its features in common clinical MRI scanners.

8.1.1 Local diffusion estimation: discrete vs analytical and continuous EAP representation

Regarding the estimation of local diffusion information, we have distinguished between two EAP representations. In **chapter 5**, we use the CS technique to reconstruction an EAP from DSI-like acquisition, i.e., the dMRI signal is acquired following a Cartesian sampling. After the application of the Fast Fourier Transform (FFT), the resulting EAP is discrete and is also defined on a Cartesian grid. A discrete representation causes expensive numerical computations if one wants to accurately derive other diffusion features. In **chapter 6** and **chapter 7**, the EAP is characterized as a linear combination of continuous and parametric functions. This leads to analytical formulae to perform useful computation of diffusion features for free. In particular, the continuous frameworks seen in these two chapters provide closed form formulae to estimate the Orientation Distribution Function (ODF) and allows one to extrapolate/interpolate the dMRI signal. However, these analytical formulae are not limited to the estimation of the ODF and other diffusion features can be analytically computed.

8.1.2 The multiple aspects of the Compressive Sensing theory

The Compressive Sensing theory is the central point of this thesis. Since the introduction of CS by Donoho and Candes [23, 46] a decade ago, many contributions have been presented on this topic. At first, the CS theory was described to reconstruct discrete and finite-length signals below the Shannon-Nyquist rate under the condition that the signal admits a sparse representation with respect to an orthonormal basis. This classical theory is used in **chapter 5** to accelerate the DSI technique. However, most of real world applications are outside the scope of this theory for two main reasons: 1-Real world signals are analog and 2-the signal to reconstruct may not be sparse in a orthonormal basis but in a overcomplete and redundant dictionary. For these two reasons, **chapter 6** and **chapter 7** handle the case of recovering analog signals. In particular, **chapter 7** introduces a dictionary learning algorithm based on a new continuous dMRI framework in order to design parametric and overcomplete dictionaries from a training data set of dMRI signals.

8.2 Applications and Collaborations

This work was partially done in collaboration with the Sherbrooke Connectivity Imaging Lab (SCIL) at the University of Sherbrooke (Canada). In particular we have been working

together in the application of CS to accelerate the DSI technique. The result of this collaboration is described in **chapter 5**. Another work have been done with the SCIL regarding a novel tractography algorithm based on the EAP [94]. During this PhD thesis, we have also implemented some of the Athena project team algorithms in the dipy software [56].

8.2.1 Fiber tractography via the Ensemble Average Propagator

It's well known that in diffusion MRI (dMRI), fibre crossing is an important problem for most existing diffusion tensor imaging (DTI) based tractography algorithms. To overcome these limitations, High Angular Resolution Diffusion Imaging (HARDI) based tractography has been proposed with a particular emphasis on the the Orientation Distribution Function (ODF). In our recent work [94], we advocate the use of the Ensemble Average Propagator (EAP) instead of the ODF for tractography in dMRI and propose an original and efficient EAP-based tractography algorithm that outperforms the classical ODF-based tractography, in particular, in the regions that contain complex fibre crossing configurations. Various experimental results including synthetic, phantom and real data illustrate the potential of the approach and clearly show that our method is especially efficient in handling regions where fiber bundles are crossing, and still handle well other fiber bundle configurations such as U-shape and kissing fibers.

8.2.2 Software

All the models and techniques presented in this thesis have been implemented in Python and are part of the project-team library called PyAthena. The dMRI signals representation with respect to some of the orthonormal bases presented in **chapter 6** were also implemented in dipy, a novel software library for diffusion MR and tractography [56].

Chapter 9

Conclusions (en français)

Contents

9.1 Conclusion générale 156
9.1.1 L'estimation de la diffusion locale 156
9.1.2 Les multiples aspects de la théorie du Compressive Sensing . . . 156
9.2 Applications et collaborations 157
9.2.1 Tractographie de fibres via l'EAP 157
9.2.2 Logiciel 157

9.1 Conclusion générale

Nous avons présenté dans cette thèse plusieurs contributions relatives à l'estimation de l'information locale de diffusion. En particulier, nous nous sommes concentrés sur l'estimation de l'EAP. Les théories sous-jacentes et les méthodes utilisées dans cette thèse sont liées aux techniques de reconstruction parcimonieuse et plus précisement la technique dite du Compressive Sensing (CS). Ce travail apporte un nouvel espoir pour estimer l'EAP avec un nombre réduit d'échantillons. D'autres caractéristiques de diffusion sont également étudiées pour réaliser la validation de la reconstruction du signal de diffusion. Enfin, nous avons rendu possible l'estimation de l'EAP et de ses caractéristiques dérivées dans la plupart des scanners IRM cliniques.

9.1.1 L'estimation de la diffusion locale: représentation discrète vs représentation analytique et continue de l'EAP

En ce qui concerne l'estimation de l'information locale de diffusion , nous avons distingué deux types de représentation de l'EAP. Dans le **chapitre 5**, nous utilisons la technique du CS pour reconstruire l'EAP à partir d'acquisitions DSI, c'est-à-dire lorsque le signal de diffusion est acquis sur une grille cartésienne. Après l'application de la transformée de Fourier rapide, l'EAP qui en résulte est discret et est également défini sur une grille cartésienne. Une représentation discrète engendre des calculs numériques coûteux si l'on veut estimer avec précision d'autres caractéristiques de diffusion. Dans le **chapitre 6** et le **chapter 7**, l'EAP se caractérise par une combinaison linéaire de fonctions paramétriques et continues. Cela conduit à des formules analytiques afin d'estimer d'importantes caractéristiques de diffusion comme la fonction de distribution des orientations et permettent d'extrapoler/interpoler le signal de diffusion.

9.1.2 Les multiples aspects de la théorie du Compressive Sensing

La théorie du Compressive Sensing est le point central de cette thèse. Depuis l'introduction du CS par Donoho et Candes [23, 46] il y a une dizaine d'années, de nombreuses contributions ont été présentées à ce sujet. Dans un premier temps, la théorie du CS a été décrite pour reconstruire des signaux discrets et de longueur finie échantillonnés à un taux en dessous de celui de Shannon-Nyquist à la condition que le signal admette une représentation parcimonieuse dans une base orthonormale. Cette théorie classique est utilisée dans le **chapitre 5** afin d'accélérer la technique de DSI. Cependant, la plupart des applications du monde réel sont en dehors du champ d'application de cette théorie pour deux raisons principales: les signaux du monde réel sont analogiques et le signal à reconstruire est rarement parcimonieux dans une base orthonormale, mais plutôt dans un dictionnaire sur-complets et redondant. Pour ces deux raisons, le **chapitre 6** et le **chapitre 7** étendent la théorie classique du CS dans le cas des signaux analogiques. Le **chapitre 7** introduit un algorithme d'apprentissage de dictionnaire afin de concevoir des dictionnaires paramétriques et surcomplets à partir d'un ensemble de données d'apprentissage de signaux de diffusion.

9.2 Applications et collaborations

Ce travail a été partiellement réalisé en collaboration avec le laboratoire SCIL de l'Université de Sherbrooke (Canada). Nous avons travaillé ensemble dans l'application du CS afin d'accélérer la technique de DSI. Le résultat de cette collaboration est décrit dans le **chapter 5**. Un autre travail a été fait avec le CSIL sur un algorithme de tractographie basé sur l'EAP. Au cours de cette thèse, nous avons également mis en œuvre une partie des algorithmes de l'équipe-projet Athena dans le logiciel dipy [56].

9.2.1 Tractographie de fibres via l'EAP

Il est bien connu qu'en IRM de diffusion, le croisement de fibres est un problème important que le tenseur de diffusion ne peut pas gérer. Pour surmonter ces limitations, l'imagerie à haute résolution angulaire a été proposé pour effectuer des tractographies en se basant sur la fonction de distribution des orientations. Dans notre récent travail [94], nous préconisons l'utilisation de l'EAP au lieu de la fonction de distribution des orientations pour la tractographie. Ainsi, nous proposons un algorithme de tractographie originale et efficace basé sur l'EAP qui surpasse les algorithmes de tractography basés sur la fonction de distribution des orientations, en particulier dans les régions qui contiennent des configurations de fibres complexes. Divers résultats expérimentaux illustrent le potentiel de cette méthode et démontrent clairement que notre méthode est efficace pour gérer les régions où les faisceaux de fibres se croisent, et encore bien d'autres configurations de faisceaux de fibres tels que le U et les "baisers" de fibres.

9.2.2 Logiciel

Tous les modèles et techniques présentés dans cette thèse ont été implémentées en Python et font partie de la bibliothèque de l'équipe-projet appelée PyAthena. Certaines bases orthonormales qui ont été présentées dans le **chapitre 6** ont été implémentées dans dipy, une bibliothèque de logiciels d'IRM de diffusion et de tractographie [56].

Chapter 10

Appendix

Contents

10.1 Fourier transform 160
10.2 A world of wavelets 160
10.3 Inner product space 161
10.4 ℓ_p-norm in an n-dimensional Euclidean space 162
10.5 Orthonormal basis 162
10.6 Overcomplete and redundant dictionary 163
10.7 Analytical diffusion signal, EAP and ODF from the multiple tensor model 163

10.1 Fourier transform

The Fourier transform is a mathematical tool, which has found many applications in mathematical analysis, physics and engineering. It allows one to completely describe a signal in terms of its frequencies.

The Fourier transform, also called spectrum, of a signal $x(t)$ is

$$X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-2j\pi ft)dt, \tag{10.1}$$

where j is the imaginary unit, which is defined by its property $j^2 = -1$.

The inversion formula is

$$x(t) = \int_{-\infty}^{+\infty} X(f) \exp(2j\pi ft)df. \tag{10.2}$$

The Fourier transform has many properties. We give a non exhaustive list of them in Tab. 10.1 and add Fourier transform of common functions.

	$x(t)$	$X(f)$	Comments
scaling	$x(at)$	$\frac{1}{\|a\|}X(\frac{f}{a})$	a is a non-zero real number
shift	$x(t-a)$	$X(f)\exp(-2j\pi fa)$	a is a real number
conjugate	$x^*(t)$	$X^*(-f)$	
convolution	$x(t) * h(t)$	$X(f)H(f)$	
linearity	$ax(t) + bh(t)$	$aX(f) + bH(f)$	a, b are complex numbers
Dirac comb	$\sum_{k=-\infty}^{+\infty} \delta(t - kT)$	$F\sum_{k=-\infty}^{+\infty} \delta(f - kF)$	$F = \frac{1}{T}$
sinusoid	$\sin(2\pi f_0 t)$	$\frac{1}{2}i(\delta(f - f_0) + \delta(f + f_0))$	

Table 10.1: Properties of Fourier Transform and Fourier transform of common functions

10.2 A world of wavelets

The Fourier transform efficiently describes periodic or stationary signals in terms of frequency, but is not able to represent transient phenomenon. Wavelets are well localized and can represent local transient structures and singularities. In the following, we describe the main idea behind the wavelet transform. More information can be found in [80].

Let us consider the mother wavelet $\mathcal{W} \in \mathbf{L^2}(\mathbb{R})$. One can construct the dilated and translated family [80]

$$\left\{ \mathcal{W}_{j,n}(t) = \frac{1}{\sqrt{2^j}} \mathcal{W}\left(\frac{t - 2^j n}{2^j} \right) \right\}_{(j,n)\in\mathbb{Z}^2}, \tag{10.3}$$

which forms an orthonormal basis of $\mathbf{L^2}$. (10.3) is called the wavelet basis.

The idea is to represent any function f as a combination of wavelet functions $\mathcal{W}_{j,n}$,

$$f(t) = \sum_{n,j\in\mathbb{Z}^2} \langle f, \mathcal{W}_{j,n} \rangle \mathcal{W}_{j,n}(t) \tag{10.4}$$

From here, comes the concept of multiresolution analysis. Multiresolution analysis was introduced to represent functions as successive approximations at resolution 2^{-j}. To define the multiresolution analysis, Mallat [79] introduced the dilated and translated family of the scaling function $\mathcal{S} \in \mathbf{L^2}(\mathbb{R})$,

$$\left\{ \mathcal{S}_{j,n}(t) = \frac{1}{\sqrt{2^j}} \mathcal{S}\left(\frac{t-n}{2^j} \right) \right\}_{(j,n)\in\mathbb{Z}^2}, \tag{10.5}$$

which forms an orthonormal basis of $\mathbf{L^2}$. This basis is the orthogonal complement of the wavelet basis in (10.3).

Then from the wavelet and scaling bases, Mallat [79] characterizes the difference between two approximations of f at a different resolution j as

$$\sum_{n\in\mathbb{Z}^2} \langle f, \mathcal{S}_{j-1,n} \rangle \, \mathcal{S}_{j-1,n} = \sum_{n\in\mathbb{Z}^2} \langle f, \mathcal{S}_{j,n} \rangle \, \mathcal{S}_{j,n} + \sum_{n\in\mathbb{Z}^2} \langle f, \mathcal{W}_{j,n} \rangle \, \mathcal{W}_{j,n} \tag{10.6}$$

The wavelet coefficients $\langle f, \mathcal{W}_{j,n} \rangle$ characterizes the information between two successive approximations. Roughly speaking, $\mathcal{W}$ models the details and $\mathcal{S}$ models the coarse components.

For discrete signals $f[n], n \in \mathbb{N}$, the multiresolution analysis is performed by a filtering and downsampling procedure with a particular class of filters, called the conjugate mirror filters. These filters characterize the wavelet functions (via a high pass filter h) and the scaling function (via a low pass filter g). We obtain the wavelet coefficients of $f[n]$ after applying a fast Discrete Wavelet Transform (DWT), which is implemented by cascading the filters h and g. Then, an inverse DWT (IDWT) ensures a perfect reconstruction, i.e., an error free reconstruction of $f[n]$ from its wavelet coefficients.

The DWT is an extremely well established tool in the image processing community and is used, for instance, in image compression because it provides highly sparse representation of natural images.

10.3 Inner product space

An inner product space is a vector space $\mathbf{V}$ together with an inner product on it. The inner product, noted $\langle \cdot, \cdot \rangle$, of two vectors $(\mathbf{u}, \mathbf{v}) \in \mathbf{V}$ produces a scalar $\alpha \in \mathbb{F}$ with $\mathbb{F}$ is either the set of complex number $\mathbb{C}$ or the set of real number $\mathbb{R}$, i.e.,

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F} \tag{10.7}$$

The inner product is required to satisfy the following three axioms for all $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in \mathbf{V}$ and $(\alpha, \beta) \in \mathbb{F}$:

Linearity	$\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{v}, \mathbf{w} \rangle$
Conjugate symmetry	$\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$
Positivity	$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ for $\mathbf{u} \neq 0$

Examples of inner product space:

Real number space $\mathbb{R}$	$\langle u, v\rangle = uv$
Euclidean space $\mathbb{R}^n$	$\langle (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n)\rangle$ $= x_1y_1 + x_2y_2 + ...x_ny_n$
Vector space of real functions defined on $[a, b]$	$\langle f, g\rangle = \int_a^b fgdx.$

10.4 ℓ_p-norm in an n-dimensional Euclidean space

In the manuscript, we will often deal with discrete and finite length signals $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ in an n-dimensional Euclidean space, denoted by $\mathbb{R}^n$. In this case, we define the ℓ_p-norm as

$$\|x\|_{\ell_p} = \begin{cases} (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}, & p \in [1, \infty) \\ \max\limits_{i=1,2,\ldots,n} |x_i|, & p = \infty \\ |\text{supp}(x)|, & p = 0 \end{cases} \tag{10.8}$$

Strictly speaking, the ℓ_0-norm is not a "norm". It is actually a quasi-norm.

The ℓ_p-norm (quasi-)norms have different properties. We show in Fig. 10.1, the unit sphere induced by each of these norms in $\mathbb{R}^2$.

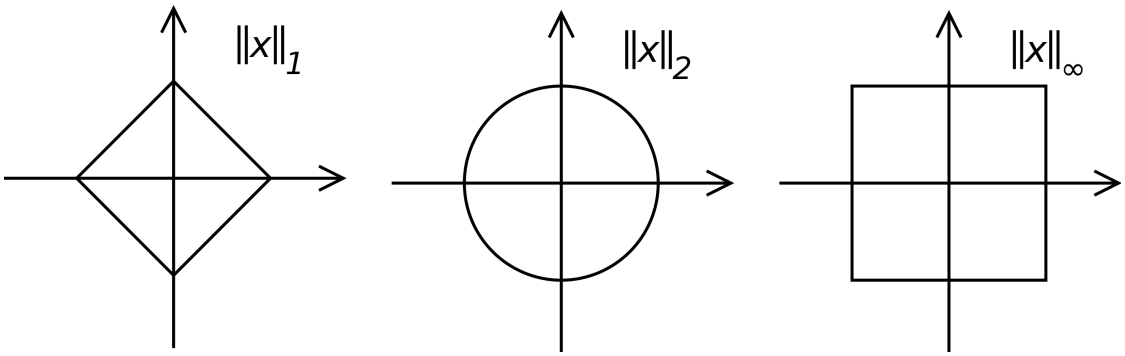


Figure 10.1: Unit sphere in $\mathbb{R}^2$ for the ℓ_p-norm with $p = 1, 2, \infty$.

10.5 Orthonormal basis

A subset of vectors $(\mathbf{u_1}, \cdots, \mathbf{u_n})$ of a inner product space $\mathbf{V}$ forms an orthonormal basis if

- $\langle \mathbf{u_i}, \mathbf{u_j}\rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$
- The set of vectors are linearly independent.

10.6 Overcomplete and redundant dictionary

In contrast to an orthonormal basis, an overcomplete and redundant dictionary is a set of vectors, which is by definition not linearly independent, i.e., we can write one vector as a linear conbination of many vectors in the set.

For instance, let us consider these 3 vectors in the 3-dimensional Euclidean space,

$$\underbrace{\begin{bmatrix}0\\0\\1\end{bmatrix}}_{\mathbf{v_1}}, \quad \underbrace{\begin{bmatrix}0\\2\\-2\end{bmatrix}}_{\mathbf{v_2}}, \quad \underbrace{\begin{bmatrix}1\\-2\\1\end{bmatrix}}_{\mathbf{v_3}}, \quad \underbrace{\begin{bmatrix}4\\2\\3\end{bmatrix}}_{\mathbf{v_4}}. \tag{10.9}$$

The four vectors together are linearly dependent because $\mathbf{v_1}$ can be written as a linear combination of $\mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}$ such that

$$\mathbf{v_1} = \left(-\frac{5}{9}\right)\mathbf{v_2} + \left(-\frac{4}{9}\right)\mathbf{v_3} + \frac{1}{9}\mathbf{v_4}. \tag{10.10}$$

10.7 Analytical diffusion signal, EAP and ODF from the multiple tensor model

We assume the normalized diffusion signal $E(\mathbf{q})$ is generated from the multi-tensor model for F fibers,

$$E(q\mathbf{u}) = \sum_{f=1}^{F} p_f \exp(-4\pi^2\tau q^2 \mathbf{u}^T D_f \mathbf{u}), \tag{10.11}$$

where a fibre f is defined by a tensor matrix D_f and weight p_f. q denotes the norm of the effective gradient and $\mathbf{u}$ is a unitary vector in Cartesian coordinate.

The analytical ground truth of the EAP for any radius R is then given by

$$P(R\mathbf{r}) = \sum_{f=1}^{F} p_f \frac{1}{\sqrt{(4\pi\tau)^3|D_f|}} \exp\left(\frac{-R^2\mathbf{r}^T D_f^{-1}\mathbf{r}}{4\tau}\right), \tag{10.12}$$

with $\mathbf{r}$ a unitary vector in Cartesian coordinate.

We can also derive an analytical and closed form expression for the ODF solid angle [1, 123],

$$\Upsilon(r) = \sum_{f=1}^{F} p_f \frac{1}{4\pi|D_f|^{\frac{1}{2}}(\mathbf{r}^T D_f^{-1}\mathbf{r})^{\frac{3}{2}}}, \tag{10.13}$$

Chapter 11

Publications of the author

Journal Publications

Sylvain Merlet, Rachid Deriche. Continuous diffusion signal, EAP and ODF estimation via Compressive Sensing in diffusion MRI. *Medical Image Analysis*. Elsevier, 2013, Online, doi-10.1016/j.media.2013.02.010.

Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, Rachid Deriche. A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features. *Medical Image Analysis*. Elsevier, 2013, Online, doi-10.1016/j.media.2013.04.011.

Sylvain Merlet, Michael Paquette, Rachid Deriche and Maxime Descoteaux. Accelerated DSI with efficient sampling and sparse discrete representation. *Magnetic Resonance in Medicine*, Wiley, 2013, submitted.

Conferences and Workshops with Proceedings

Sylvain Merlet, Anne-Charlotte Philippe, Rachid Deriche, Maxime Descoteaux. Tractography via the Ensemble Average Propagator in diffusion MRI. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012* , Sep 2013, Nice, France. Springer, Volume 7511, pp 339-346 , Lecture Notes in Computer Science.

Sylvain Merlet, Emmanuel Caruyer, Rachid Deriche. Parametric dictionary learning for modeling EAP and ODF in diffusion MRI. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012* , Sep 2013, Nice, France. Springer, Volume 7511, pp 339-346 , Lecture Notes in Computer Science.

Sylvain Merlet, Michael Paquette, Rachid Deriche, and Maxime Descoteaux. Ensemble Average Propagator Reconstruction via Compressed Sensing: Discrete or Continuous Bases, *20th ISMRM annual meeting*, May 2012, Melbourne, Australia.

Sylvain Merlet, Rachid Deriche, Kevin Whittingstall, and Maxime Descoteaux. Diffusion and multiple orientations from 1.5 MR systems with limited gradient tables. *20th ISMRM annual meeting*, May 2012, Melbourne, Australia.

Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, Rachid Deriche. Parametric Dictionary Learning in Diffusion MRI. *HARDI reconstruction workshop - ISBI - 9th IEEE International Symposium on Biomedical Imaging*, May 2012, Barcelona, Spain.

Sylvain Merlet, Emmanuel Caruyer and Rachid Deriche. Accelerating Diffusion MRI via Compressive Sensing. *ITWIST 2012*, May 2012, Marseille, France.

Sylvain Merlet, Emmanuel Caruyer and Rachid Deriche. Impact of radial and angular sampling on multiple shells acquisition in diffusion MRI. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011* , Sep 2011, Toronto, Canada. Springer, Volume 6892, pp 116-123, Lecture Notes in Computer Science.

Sylvain Merlet, Jian Cheng, Aurobrata Ghosh and Rachid Deriche. Spherical Polar Fourier EAP and ODF Reconstruction via Compressed Sensing in Diffusion MRI. *ISBI - 10th IEEE International Symposium on Biomedical Imaging*, May 2011, Chicago, United states.

Sylvain Merlet and Rachid Deriche. Compressed Sensing for Accelerated EAP Recovery in Diffusion MRI. *MICCAI Workshop on Computational Diffusion MRI - CDMRI'10*, Sep 2010, Beijing, China.

Bibliography

[1] I. Aganj, C. Lenglet, G. Sapiro, E. Yacoub, K. Ugurbil, and N. Harel. Reconstruction of the ODF in single and multiple shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2):554–566, 2010.

[2] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11):4311–4322, 2006.

[3] DanielC. Alexander. Maximum entropy spherical deconvolution for diffusion mri. In GaryE. Christensen and Milan Sonka, editors, *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 76–87. Springer Berlin Heidelberg, 2005.

[4] Vincent Arsigny. Processing data in lie groups: an algebraic approach. application to non-linear registration and diffusion tensor mri. *These de Science*, 2006.

[5] H.E. Assemlal, D. Tschumperlé, and L. Brun. Efficient and robust computation of pdf features from diffusion mr signal. *Medical Image Analysis*, 13(5):715–729, 2009.

[6] Richard Baraniuk and Philippe Steeghs. Compressive radar imaging. In *Radar Conference, 2007 IEEE*, pages 128–133. IEEE, 2007.

[7] Peter J Basser. New histological and physiological stains derived from diffusion-tensor mr images. *Annals of the New York Academy of Sciences*, 820(1):123–138, 1997.

[8] Peter J Basser and D Le Bihan. Fiber orientation mapping in an anisotropic medium with nmr diffusion spectroscopy. In *11th Annual Meeting of the SMRM, Berlin*, volume 1221, 1992.

[9] Peter J. Basser, Jim Mattiello, and Denis Le Bihan. Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance*, B(103):247–254, 1994.

[10] Peter J. Basser, Jim Mattiello, and Denis Le Bihan. Mr diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267, 1994.

[11] Peter J Basser, Carlo Pierpaoli, et al. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor mri. *Journal of magnetic resonance. Series B*, 111(3):209–219, 1996.

[12] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.

[13] Christian R Berger, Shengli Zhou, James C Preisig, and Peter Willett. Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing. *Signal Processing, IEEE Transactions on*, 58(3):1708–1721, 2010.

[14] J.C. Bezdek, R.J. Hathaway, R.E. Howard, C.A. Wilson, and M.P. Windham. Local convergence analysis of a grouped variable version of coordinate descent. *Journal of Optimization Theory and Applications*, 54(3):471–477, 1987.

[15] B. Bilgic, K. Setsompop, J. Cohen-Adad, A. Yendiki, L.L. Wald, and E. Adalsteinsson. Accelerated diffusion spectrum imaging with compressed sensing using adaptive dictionaries. *Magnetic Resonance in Medicine*, 2012.

[16] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, Oct 1946.

[17] Jérôme Bobin, J-L Starck, and Roland Ottensamer. Compressed sensing in astronomy. *Selected Topics in Signal Processing, IEEE Journal of*, 2(5):718–726, 2008.

[18] Paul T. Callaghan. *Principles of nuclear magnetic resonance microscopy*. Oxford University Press on Demand, 1993.

[19] E. Candes and Y. Plan. A probabilistic and ripless theory of compressed sensing. *Information Theory, IEEE Transactions on*, (99):1–1, 2010.

[20] E. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse problems*, 23:969, 2007.

[21] E.J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9-10):589–592, 2008.

[22] E.J. Candes, Y.C. Eldar, D. Needell, and P. Randall. Compressed sensing with coherent and redundant dictionaries. *Applied and Computational Harmonic Analysis*, 31(1):59–73, 2011.

[23] E.J. Candès and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[24] Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94(3):630, 1954.

[25] Melissa K Carroll, Guillermo A Cecchi, Irina Rish, Rahul Garg, and A Ravishankar Rao. Prediction and interpretation of distributed neural activity with sparse models. *NeuroImage*, 44(1):112–122, 2009.

[26] Emmanuel Caruyer, Jian Cheng, Christophe Lenglet, Guillermo Sapiro, Tianzi Jiang, and Rachid Deriche. Optimal design of multiple q-shells experiments for diffusion mri. In *MICCAI Workshop on Computational Diffusion MRI - CDMRI'11*, Toronto, Canada, September 2011.

[27] Emmanuel Caruyer and Rachid Deriche. Diffusion MRI signal reconstruction with continuity constraint and optimal regularization. *Medical Image Analysis*, 16(6):1113–1120, 2012.

[28] Emmanuel Caruyer, Aurobrata Ghosh, and Dr. Deriche, Rachid. A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features. *Medical Image Analysis*, May 2013.

[29] Emmanuel Caruyer, Christophe Lenglet, Guillermo Sapiro, and Rachid Deriche. Incremental gradient table for multiple q-shells diffusion mri. In *HBM 17th Annual Meeting*, Quebec, Canada, June 2011.

[30] R. Chartrand. Fast algorithms for nonconvex compressive sensing: Mri reconstruction from very few data. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2009.

[31] C ChefD'Hotel, Deriche Tschumperlé, R Deriche, and O Faugeras. Regularizing flows for constrained matrix-valued images. *Journal of Mathematical Imaging and Vision*, 20(1-2):147–162, 2004.

[32] C Chefd'Hotel, D Tschumperlé, Rachid Deriche, and O Faugeras. Constrained flows of matrix-valued functions: Application to diffusion tensor regularization. In *Computer Vision—ECCV 2002*, pages 251–265. Springer, 2002.

[33] Jian Cheng, Aurobrata Ghosh, Rachid Deriche, and Tianzi Jiang. Model-free, regularized, fast, and robust analytical orientation distribution function estimation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 6361 of *Lecture Notes in Computer Science*, pages 648–656. Springer, 2010.

[34] Jian Cheng, Aurobrata Ghosh, Tianzi Jiang, and Rachid Deriche. Model-free and analytical eap reconstruction via spherical polar fourier diffusion mri. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 6361, pages 590–597, 2010.

[35] Jian Cheng, Tianzi Jiang, and Rachid Deriche. Theoretical analysis and practical insights on eap estimation via a unified hardi framework. In *MICCAI workshop on Computational Diffusion MRI*, 2011.

[36] Jian Cheng, Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, Tianzi Jiang, and Rachid Deriche. Compressive sensing ensemble average propagator estimation via l1 spherical polar fourier imaging. In *MICCAI workshop on Computational Diffusion MRI*, 2011.

[37] R.R. Coifman, D.L. Donoho, et al. Translation-invariant de-noising. *Lecture notes in statistics-New York -Springer Verlag*, pages 125–125, 1995.

[38] 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. In *14th ISMRM*, Seattle, USA, 2006.

[39] DG Cory. Measurement of translational displacement probabilities by nmr: an indicator of compartmentation. *Magnetic resonance in medicine*, 14(3):435–444, 1990.

[40] P Craven and G Wahba. Smoothing noisy data with spline functions. *Numerische Mathematik*, 31(4):377–403, 1985.

[41] R. Deriche, J. Calder, and M. Descoteaux. Optimal real-time q-ball imaging using regularized kalman filtering with incremental orientation sets. *Medical image analysis*, 13(4):564–579, 2009.

[42] M. Descoteaux. *High angular resolution diffusion mri: from local estimation to segmentation and tractography*. PhD thesis, 2008.

[43] M. Descoteaux, R. Deriche, T.R. Knosche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *Medical Imaging, IEEE Transactions on*, 28(2):269–286, 2009.

[44] Maxime Descoteaux, Elaine Angelino, Shaun Fitzgibbons, and Rachid Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007.

[45] Maxime Descoteaux, Rachid Deriche, Denis Le Bihan, Jean-François Mangin, and Cyril Poupon. Multiple q-shell diffusion propagator imaging. *Medical Image Analysis*, 15:603–621, 2011.

[46] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[47] D.L. Donoho. Method and apparatus for compressed sensing, January 2010. US Patent 7,646,924.

[48] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, T. Sun, K.F. Kelly, and R.G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, 2008.

[49] M.F. Duarte and Y.C. Eldar. Structured compressed sensing: From theory to applications. *Signal Processing, IEEE Transactions on*, 59(9):4053–4085, 2011.

[50] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.

[51] Albert Einstein. On the movement of small particles suspended in stationary liquids required by the molecular-kinetic theory of heat. *Annalen der Physik*, 17(549-560):16, 1905.

[52] Yonina C. Eldar. Uncertainty relations for shift-invariant analog signals. *IEEE Trans. Inf. Theor.*, 55(12):5742–5757, December 2009.

[53] Pierre Fillard, Maxime Descoteaux, Alvina Goh, Sylvain Gouttard, Ben Jeurissen, James Malcolm, Alonso Ramirez-Manzanares, Marco Reisert, Ken Sakaie, Fatima Tensaouti, Ting Yo, Jean-François Mangin, and Cyril Poupon. Quantitative analysis of 10 tractography algorithms on a realistic diffusion MR phantom. *Neuroimage*, 56(1):220–234, 2011.

[54] Pierre Fillard, Xavier Pennec, Vincent Arsigny, and Nicholas Ayache. Clinical dt-mri estimation, smoothing, and fiber tracking with log-euclidean metrics. *Medical Imaging, IEEE Transactions on*, 26(11):1472–1482, 2007.

[55] A. Ganesh, VR Edward, et al. Reordering for improved constrained reconstruction from undersampled k-space data. *International Journal of Biomedical Imaging*, 2008, 2008.

[56] Eleftherios Garyfallidis, Matthew Brett, Bagrat Amirbekian, Christopher Nguyen, Fang-Cheng Yeh, Y Halchenko, and I Nimmo-Smith. Dipy–a novel software library for diffusion mr and tractography. In *17th Annual Meeting of the Organization for Human Brain Mapping*, 2011.

[57] A. Gramfort, P. Cyril, M. Descoteaux, et al. Sparse dsi: Learning dsi structure for denoising and fast imaging. In *MICCAI*, 2012.

[58] A. Gramfort, C. Poupon, and M. Descoteaux. Sparse dsi: Learning dsi structure for denoising and fast imaging. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 288–296, 2012.

[59] Erwin L Hahn. Spin echoes. *Physical Review*, 80(4):580, 1950.

[60] L. He, T.C. Chang, S. Osher, T. Fang, and P. Speier. Mr image reconstruction by using the iterative refinement method and nonlinear inverse scale space methods. *UCLA CAM Report*, 6:35, 2006.

[61] Matthew A Herman and Thomas Strohmer. High-resolution radar via compressed sensing. *Signal Processing, IEEE Transactions on*, 57(6):2275–2284, 2009.

[62] A. Pasha Hosseinbor, Moo K. Chung, Yu-Chien Wu, and Andrew L. Alexander. Bessel fourier orientation reconstruction (bfor): An analytical diffusion propagator reconstruction for hybrid diffusion imaging and computation of q-space indices. *NeuroImage*, 64(0):650 – 670, 2013.

[63] Ameer Hosseinbor, Moo Chung, Yu-Chien Wu, and Andrew Alexander. Bessel fourier orientation reconstruction: An analytical eap reconstruction using multiple shell acquisitions in diffusion mri. In Gabor Fichtinger, Anne Martel, and Terry Peters, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, volume 6892 of *Lecture Notes in Computer Science*, pages 217–225. Springer Berlin / Heidelberg, 2011.

[64] Jens H. Jensen, Joseph A. Helpern, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005.

[65] J.H. Jensen and J.A. Helpern. Mri quantification of non-gaussian water diffusion by kurtosis analysis. *NMR in Biomedicine*, 23(7):698–710, 2010.

[66] B. Jian, B.C. Vemuri, E. Özarslan, P.R. Carney, and T.H. Mareci. A novel tensor distribution model for the diffusion-weighted mr signal. *NeuroImage*, 37(1):164–176, 2007.

[67] Heidi Johansen-Berg and Timothy E. J. Behrens, editors. *Diffusion MRI : from quantitative measurement to in-vivo neuroanatomy*. Elsevier/Academic Press, Amsterdam, Boston, 2009. nlmcopyc.

[68] D.K. Jones, M.A. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(39):515 – 525, August 1999.

[69] Daniela Kuhnt, Miriam HA Bauer, Jan Egger, Mirco Richter, Tina Kapur, Jens Sommer, Dorit Merhof, and Christopher Nimsky. Fiber tractography based on diffusion tensor imaging compared with high-angular-resolution diffusion imaging with compressed sensing: Initial experience. *Neurosurgery*, 72:A165–A175, 2013.

[70] Li-Wei Kuo, Wen-Yang Chiang, Fang-Cheng Yeh, Van Jay Wedeen, and Wen-Yih Isaac Tseng. Diffusion spectrum mri using body-centered-cubic and half-sphere sampling schemes. *Journal of Neuroscience Methods*, 212(1):143 – 155, 2013.

[71] Bennett A. Landman, John A. Bogovic, Hanlin Wan, Fatma El Zahraa ElShahaby, Pierre-Louis Bazin, and Jerry L. Prince. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor mri. *NeuroImage*, 59(3):2175 – 2186, 2012.

[72] Bennett A. Landman, John A. Bogovic, Hanlin Wan, Fatma El Zahraa ElShahaby, Pierre-Louis Bazin, and Jerry L. Prince. Resolution of crossing fibers with constrained compressed sensing using diffusion tensor mri. *NeuroImage*, 59(3):2175 – 2186, 2012.

[73] Denis Le Bihan, E Breton, et al. Imagerie de diffusion in-vivo par résonance magnétique nucléaire. *Comptes-Rendus de l'Académie des Sciences*, 93(5):27–34, 1985.

[74] Denis Le Bihan, Eric Breton, Denis Lallemand, Philippe Grenier, Emmanuel Cabanis, Maurice Laval-Jeantet, et al. Mr imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2):401–407, 1986.

[75] Christophe Lenglet. *Geometric and variational methods for diffusion tensor mri processing*. PhD thesis, University Nice Sophia-Antipolis, FRANCE, 2006.

[76] Chunlei Liu, Roland Bammer, Burak Acar, and Michael E Moseley. Characterizing non-gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5):924–937, 2004.

[77] M. Lustig, D. Donoho, and J.M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[78] Mark F Lythgoe, Albert L Busza, Fernando Calamante, Christopher H Sotak, Martin D King, Anna C Bingham, Stephen R Williams, and David G Gadian. Effects of diffusion anisotropy on lesion delineation in a rat model of cerebral ischemia. *Magnetic resonance in medicine*, 38(4):662–668, 1997.

[79] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[80] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[81] M Mania, M Jacobb, A Guidon, C Liu, A Songc, V Magnottab, and J Zhonga. Acceleration of high angular and spatial resolution diffusion imaging using compressive sensing.

[82] Marion I. Menzel, Ek T. Tan, Kedar Khare, Jonathan I. Sperl, Kevin F. King, Xiaodong Tao, Christopher J. Hardy, and Luca Marinelli. Accelerated diffusion spectrum imaging in the human brain using compressed sensing. *Magnetic Resonance in Medicine*, 66:226–1233, November 2011.

[83] Klaus-Dietmar Merboldt, Wolfgang Hanicke, and Jens Frahm. Self-diffusion nmr imaging using stimulated echoes. *Journal of Magnetic Resonance (1969)*, 64(3):479–486, 1985.

[84] S. Merlet, E. Caruyer, and R. Deriche. Parametric dictionary learning for modeling eap and odf in diffusion mri. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 10–17, 2012.

[85] Sylvain Merlet, Emmanuel Caruyer, and Rachid Deriche. Impact of radial and angular sampling on multiple shells acquisition in diffusion mri. In *MICCAI*, volume 6891, pages 113–121. Springer, 2011.

[86] Sylvain Merlet, Emmanuel Caruyer, and Rachid Deriche. Parametric dictionary learning for modeling eap and odf in diffusion mri. In *Lecture Notes in Computer Science series MICCAI 2012*, volume 7512, page 7512, Nice, France, October 2012.

[87] Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, and Rachid Deriche. Parametric dictionary learning in diffusion mri. In *HARDI Contest, ISBI - Internation Symposium on Biomedical Imaging*, Barcelona, Spain, May 2012. Alessandro Daducci and Jean-Philippe Thiran and Yves Wiaux.

[88] Sylvain Merlet, Emmanuel Caruyer, Aurobrata Ghosh, and Rachid Deriche. A computational diffusion mri and parametric dictionary learning framework for modeling the diffusion signal and its features. *Medical Image Analysis*, (0):–, 2013.

[89] Sylvain Merlet, Jian Cheng, Aurobrata Ghosh, and Rachid Deriche. Spherical polar fourier eap and odf reconstruction via compressed sensing in diffusion mri. In *Proceedings of ISBI*, April 2011.

[90] Sylvain Merlet, Dr. Deriche, Rachid, Kevin Whittingstall, and Maxime Descoteaux. Diffusion and multiple orientations from 1.5 mr systems with limited gradient tables. In *ISMRM 20th Annual Meeting (2012)*, page 1922, Melbourne, Australie, May 2012.

[91] Sylvain Merlet and Rachid Deriche. Compressed sensing for accelerated eap recovery in diffusion mri. In *Proceedings Computational Diffusion MRI - MICCAI Workshop*, September 2010.

[92] Sylvain Merlet, Michael Paquette, Rachid Deriche, and Maxime Descoteaux. Ensemble average propagator reconstruction via compressed sensing: Discrete or continuous bases ? In *ISMRM 20th Annual Meeting (2012)*, page 2277, Australie, May 2012.

[93] Sylvain Merlet, Michael Paquette, Rachid Deriche, and Maxime Descoteaux. A comparison of sampling strategies and sparsifying transforms for compressed sensing dsi. *Magnetic Resonance in Medicine*, submitted.

[94] Sylvain Merlet, Anne-Charlotte Philippe, Rachid Deriche, and Maxime Descoteaux. Tractography via the ensemble average propagator in diffusion mri. In *Lecture Notes in Computer Science series*, volume 7511, page 7511, Nice, France, October 2012.

[95] Sylvain. L Merlet and Rachid Deriche. Continuous diffusion signal, eap and odf estimation via compressive sensing in diffusion mri. *Medical Image Analysis*, 17(5):556 – 572, 2013.

[96] O. Michailovich and Y. Rathi. On approximation of orientation distributions by means of spherical ridgelets. *Image Processing, IEEE Transactions on*, 19(2):461–477, 2010.

[97] O. Michailovich, Y. Rathi, and S. Dolui. Spatially regularized compressed sensing for high angular resolution diffusion imaging. *Medical Imaging, IEEE Transactions on*, 30(5):1100–1115, 2011.

[98] Radhouène Neji, Noura Azzabou, Nikos Paragios, and Gilles Fleury. A convex semi-definite positive framework for dti estimation and regularization. In *Advances in Visual Computing*, pages 220–229. Springer, 2007.

[99] E. Ozarslan, C.G. Koay, T.M. Shepherd, S.J. Blackband, and P.J. Basser. Simple harmonic oscillator based reconstruction and estimation for three-dimensional q-space mri. In *ISMRM 17th Annual Meeting and Exhibition, Honolulu,*, page 1396, 2009.

[100] Evren Ozarslan, T.M. Shepherd, Baba C. Vemuri, S.J. Blackband, and Thomas H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (dot). *NeuroImage*, 31(3):1086–1103, 2006.

[101] Evren Özarslan, Baba C Vemuri, and Thomas H Mareci. Generalized scalar measures for diffusion mri using trace, variance, and entropy. *Magnetic Resonance in Medicine*, 53(4):866–876, 2005.

[102] Jeffrey L Paulsen, HyungJoon Cho, Gyunggoo Cho, and Yi-Qiao Song. Acceleration of multi-dimensional propagator measurements with compressed sensing. *Journal of Magnetic Resonance*, 213(1):166–170, 2011.

[103] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python . *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[104] Carlo Pierpaoli, Peter Jezzard, Peter J Basser, Alan Barnett, and Giovanni Di Chiro. Diffusion tensor mr imaging of the human brain. *Radiology*, 201(3):637–648, 1996.

[105] C. Poupon, B. Rieul, I. Kezele, M. Perrin, F. Poupon, and J-F. Mangin. New diffusion phantoms dedicated to the study and validation of hardi models. *Magnetic Resonance in Medicine*, 60:1276–1283, 2008.

[106] Lingling Pu, Theodore P Trouard, Lee Ryan, Chuan Huang, Maria I Altbach, and Ali Bilgin. Model-based compressive diffusion tensor imaging. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 254–257. IEEE, 2011.

[107] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69:37–38, Jan 1946.

[108] A. Ramirez-Manzanares, M. Rivera, B.C. Vemuri, P. Carney, and T. Mareci. Diffusion basis functions decomposition for estimating white matter intravoxel fiber geometry. *Medical Imaging, IEEE Transactions on*, 26(8):1091–1102, 2007.

[109] Y. Rathi, O. Michailovich, K. Setsompop, S. Bouix, M. Shenton, and C. Westin. Sparse multi-shell diffusion imaging. In *MICCAI*, pages 58–65. Springer, 2011.

[110] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[111] I.M. Ryzhik, A. Jeffrey, and D. Zwillinger. *Table of integrals, series and products*. Academic Press, 2007.

[112] Etienne Saint-Amant and Maxime Descoteaux. Sparsity characterization of the diffusion propagator. In *Proceedings of ISMRM*, 2011.

[113] I.W. Selesnick, R.G. Baraniuk, and N.C. Kingsbury. The dual-tree complex wavelet transform. *Signal Processing Magazine, IEEE*, 22(6):123–151, 2005.

[114] K. Setsompop, J. Cohen-Adad, B.A. Gagoski, T. Raij, A. Yendiki, B. Keil, V.J. Wedeen, and L.L. Wald. Improving diffusion mri using simultaneous multi-slice echo planar imaging. *NeuroImage*, 63(1):569 – 580, 2012.

[115] Kawin Setsompop, Lawrence L. Wald, Vijayanand Alagappan, Borjan Gagoski, Franz Hebrank, Ulrich Fontius, Franz Schmitt, and Elfar Adalsteinsson. Parallel rf transmission with eight channels at 3 tesla. *Magnetic Resonance in Medicine*, 56(5):1163–1171, 2006.

[116] Claude Elwood Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.

[117] E.O. Stejskal and J.E. Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *Journal of Chemical Physics*, 42:288–292, 1965.

[118] Mehmet Süzen, Alexia Giannoula, and Turgut Durduran. Compressed sensing in diffuse optical tomography. *Optics Express*, 18(23):23676–23690, 2010.

[119] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[120] J-Donald Tournier, Fernando Calamante, and Alan Connelly. Robust determination of the fibre orientation distribution in diffusion mri: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459 – 1472, 2007.

[121] J.-Donald Tournier, Chun-Hung Yeh, Fernando Calamante, Kuan-Hung Cho, Alan Connelly, and Ching-Po Lin. Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *NeuroImage*, 42(2):617 – 625, 2008.

[122] A. Tristán-Vega and C.F. Westin. Probabilistic odf estimation from reduced hardi data with sparse regularization. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011*, pages 182–190, 2011.

[123] A. Tristán-Vega, SC.F. Westin, and S. Aja-Fernandez. Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. *NeuroImage*, 47(2):638–650, August 2009.

[124] J. Trzasko and A. Manduca. Highly undersampled magnetic resonance image reconstruction via homotopic l0-minimization. *IEEE Transactions on Medical Imaging*, 28(1):106–121, 2009.

[125] David Tschumperlé and Rachid Deriche. Variational frameworks for dt-mri estimation, regularization and visualization. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 116–121. IEEE, 2003.

[126] D. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6):1358–1372, 2004.

[127] David Solomon Tuch. *Diffusion MRI of complex tissue structure*. PhD thesis, Citeseer, 2002.

[128] Yong Wang, Qing Wang, Justin P Haldar, Fang-Cheng Yeh, Mingqiang Xie, Peng Sun, Tsang-Wei Tu, Kathryn Trinkaus, Robyn S Klein, Anne H Cross, et al. Quantification of increased cellularity during inflammatory demyelination. *Brain*, 134(12):3590–3601, 2011.

[129] R. Ward. Compressed sensing with cross validation. *Information Theory, IEEE Transactions on*, 55(12):5773 –5782, 2009.

[130] V.J. Wedeen, P. Hagmann, W. Tseng, T.G. Reese, and R.M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54(6):1377–1386, 2005.

[131] VJ Wedeen, TG Reese, DS Tuch, MR Weigel, JG Dou, RM Weiskoff, and D Chessler. Mapping fiber orientation spectra in cerebral white matter with fourier transform diffusion mri. In *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, page 82, 2000.

[132] Carl-Fredrik Westin, Sariel Peled, Hakon Gudbjartsson, Ron Kikinis, and Ferenc A Jolesz. Geometrical diffusion measures for mri from tensor basis analysis. In *Proceedings of ISMRM*, volume 97, page 1742, 1997.

[133] Yves Wiaux, Laurent Jacques, Gilles Puy, AMM Scaife, and Pierre Vandergheynst. Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, 395(3):1733–1742, 2009.

[134] Rebecca M. Willett, Roummel F. Marcia, and Jonathan M. Nichols. Compressed sensing for practical optical imaging systems: a tutorial. *Optical Engineering*, 50(7):072601–072601–13, 2011.

[135] Y.C. Wu and A.L. Alexander. Hybrid diffusion imaging. *NeuroImage*, 36(3):617–629, 2007.

[136] W. Ye, B.C. Vemuri, and A. Entezari. An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 940–943. IEEE, 2012.

[137] W. Ye, B.C. Vemuri, and A. Entezari. An over-complete dictionary based regularized reconstruction of a field of ensemble average propagators. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 940–943. IEEE, 2012.

[138] Yanjie Zhu, Yin Wu, Yuanjie Zheng, EX Wu, Leslie Ying, and Dong Liang. A model-based method with joint sparsity constraint for direct diffusion tensor estimation. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 510–513. IEEE, 2012.

[139] M. Zibulevsky and M. Elad. L1-l2 optimization in signal and image processing. *IEEE Signal Processing Magazine*, 27(3), May 2010.