Spherical finite rate of innovation with an application to diffusion magnetic resonance imaging

Deslauriers-Gauthier, Samuel

2014

Deslauriers-Gauthier, S. (2014). Spherical finite rate of innovation with an application to diffusion magnetic resonance imaging. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/61821>

<https://doi.org/10.32657/10356/61821>

Spherical Finite Rate of Innovation

with an Application to
Diffusion Magnetic Resonance Imaging

Samuel Deslauriers-Gauthier

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement of the degree of
Doctor of Philosophy

2013

À toi, ma Joanie.

Acknowledgement

I would like to thank my advisor Pina Marziliano. This obviously would not have been possible without you. While I may not always show it, I am grateful for your relentless efforts to put me on the right track and hope that you are happy with the result. If you are less than half as satisfied as I am that I decided to come to Singapore, it is more than twice what any student deserves.

Thanks to all my Singapore friends and colleagues, Pertami J. Hayuningtyas, Shripad Kulkarni, Zhang Kunlei, Nikhil S. Narayan, Gaetano Sciacovelli, Nini Soe, Leow-How Seok Lai, and Srinivasan Yadhunathan for making my stay a memorable one and for introducing me to new cultures and delicious food. A special thanks to Amrish Nair and Talal Bin Amin for the many discussions over kopi and the Friday outings. I hope that I have not pushed you in the wrong direction too often and look forward to your future discoveries. A very special thanks to Tian Jing, the first person to welcome me in Singapore and very likely the last one to wave goodbye. It is truly amazing, and it makes me very happy, that we can get along so well despite being so very different.

I would like to thank Maxime Descoteaux of the Université de Sherbrooke for kindly providing the data used in this work.

I would also like to thank the NTU cycling team, especially Amreet Singh, Yuvaraman Viswanathan, and Ang Weizhe for the brutal training sessions. I hope that we will have the opportunity to ride again some day.

Thanks to my friends from home, in particular François Brown Mastropaolo, François Gagnon-Desmeules and Étienne Lemay. If anything can reduce the stress of living far far away, it is knowing that when you return home, your friends will treat you as if you had never left at all.

Of course, thank you to my parents for their constant support and encouragements. Our weekly videos and chats made me feel close to home despite being almost 15000km away.

Finally, thank you, with all my heart, to Joanie for understanding that going on this Singaporean adventure was not an attempt to get away. I cannot express how important your support over the last 3 years, 5 months, and 18 days has been to me. I would not wish the difficulties of a long distance relationship on anyone, but if they all went as well as ours, they just might be worth it.

Abstract

Sampling theorems describe which types of signals can be reconstructed and under which conditions. In the 64 years since Shannon’s sampling theorem for bandlimited functions, we have learned to sample many other classes of signals, some of them non-bandlimited. However, for signals defined on the sphere, the vast majority of sampling theorems still deal with bandlimited signals. In many applications, the assumption of bandlimitedness is not optimal and hampers the accuracy of the reconstruction. A notable example is diffusion magnetic resonance imaging, the application considered in this thesis.

In this work, we produce new sampling theorems that prove that certain non-bandlimited signals on the sphere can be sampled using a finite number of samples. We use them to improve the reconstruction of fiber orientations in diffusion magnetic resonance imaging.

We proceed by first showing that certain non-bandlimited signals defined on the unit sphere in three dimensions can be sampled and reconstructed using a finite number of samples. When the sample locations are equiangular, we prove that the reconstruction is exact. We also argue and validate using numerical simulations that fewer samples can be used if they are uniformly distributed. In both cases, the number of samples required depends on the number of degrees of freedom, or rate of innovation, of the signals. Next, we consider different sampling kernels and signal models that can also be sampled at their rate of innovation. We develop an optimal kernel that allows us to sample and reconstruct signals at the critical rate, the number of degrees of freedom. We also consider antipodally symmetric kernels, a case which is not covered by the previous theory. Finally, we prove that orientations integrated along the azimuth and great circles can also be sampled at their rate of innovation. Because signals are often observed in noisy environments, we then investigate how to accurately recover the signals previously described from noisy measurements. We propose two major changes to the reconstruction algorithm: one that improves the results when we do not oversample and one that uses the additional information provided by oversampling. Finally, we apply our results to the recovery of fiber orientations in diffusion magnetic resonance imaging. The main advantage of our method over existing ones is that the angular resolution does not depend on the number of samples that is acquired. In theory, our reconstruction algorithm is able to recover fiber orientations regardless of their crossing angle, given an appropriate sampling scheme. We show, through numerical simulations, that our method is indeed able to distinguish fibers that cross at a very narrow angle.

Contents

Acknowledgement	iii
Abstract	iv
List of Figures	vii
List of Tables	x
1 Introduction	1
1.1 Contributions	3
1.2 Organisation of the thesis	3
2 Background	5
2.1 Sampling signals with a finite rate of innovation	5
2.2 Signals on the sphere	6
2.3 Fourier analysis on the sphere	7
2.3.1 Spherical harmonics	7
2.3.2 Signal approximation using spherical harmonics	9
2.3.3 Rotational harmonics	10
2.4 Spherical convolution	12
2.5 Sampling bandlimited signals on the sphere	13
2.5.1 Sampling schemes on the sphere	13
2.5.2 Sampling theorems for bandlimited signals	15
3 Sampling FRI signals on the sphere	18
3.1 Spherical FRI signals	19
3.2 Signal parameters from spherical harmonics	21
3.3 Bandlimited kernels	26
3.4 Numerical simulations	28
3.4.1 Metrics and signal generation	28
3.4.2 Results	29
3.5 Comparison with compressed sensing	32
3.6 Summary	36
4 Spherical FRI variations	37
4.1 Designing the optimal kernel	37
4.2 Antipodally symmetric kernels	39
4.3 Integrated orientations	44
4.4 Sampling Great Circles	48

4.5	Summary	52
5	Sampling noisy FRI signals on the sphere	56
5.1	Noisy spherical FRI without oversampling	57
5.1.1	Using all spherical harmonic coefficients	57
5.1.2	Selecting the rotation matrix	60
5.1.3	Numerical simulations	61
5.2	Noisy spherical FRI with oversampling	63
5.2.1	Cadzow's iterative denoising	63
5.2.2	The annihilating filter method as an optimization problem . .	65
5.2.3	Initialization matters	67
5.2.4	Adding bounds	68
5.2.5	Numerical simulations	70
5.2.6	Other solutions methods	71
5.3	Summary	71
6	Application to diffusion magnetic resonance imaging	73
6.1	Modelling the diffusion process	73
6.2	Spherical deconvolution	74
6.3	Spherical FRI for diffusion MRI (sFRIdMRI)	75
6.3.1	Estimating the spherical harmonic coefficients of orientations in dMRI	76
6.3.2	Kernel bandlimit and b-value	77
6.3.3	The advantage of sFRIdMRI	77
6.4	Numerical simulations	79
6.4.1	Data simulation and performance metrics	79
6.4.2	Parameter selection	80
6.4.3	Results	81
6.5	In vivo data	85
6.5.1	Data acquisition	86
6.5.2	Results	88
6.6	Summary	93
7	Conclusions and Future Work	94
A	Distributing points on the sphere	97
B	Publications of the author arising from this work	99

List of Figures

2.1	Illustration of the spherical function $f(\theta, \varphi) = \cos \theta$	8
2.2	Illustration of the real component of the spherical harmonics up to degree 2.	9
2.3	Approximation of a function on the sphere using spherical harmonics.	11
2.4	Example of the spherical convolution operation.	14
2.5	Comparison of equiangular and uniform locations on the sphere.	15
3.1	Comparison of the FRI signal model in Cartesian and spherical coordinates.	19
3.2	Illustration of a signal with 2 orientations whose parameters are $a_1 = a_2 = 1, \theta_1 = 2.67, \theta_2 = 2.93, \varphi_1 = 4.12, \varphi_2 = 0.22$	20
3.3	The sampling scheme used to sample and reconstruct spherical FRI signals.	21
3.4	Illustration of the ideal and exponential bandlimited kernels used in the simulations.	30
3.5	Illustration of the reconstruction process using equiangular sample locations and an ideal sampling kernel for a signal with $K = 3$	31
3.6	The angular error and mean squared error of spherical FRI using uniform locations, an ideal kernel, and no noise.	33
3.7	The angular error and mean squared error of spherical FRI using uniform locations, an exponential kernel, and no noise.	34
3.8	Illustration of the reconstruction process using uniform sample locations and an ideal sampling kernel for a signal with $K = 2$	35
4.1	Illustration of the ideal optimal sampling kernels as a function of the colatitude.	39
4.2	Illustration of the linear optimal sampling kernels as a function of the colatitude.	40
4.3	The mean angular error and the mean squared error as a function of the number of samples using uniform sample locations and the <i>ideal</i> optimal kernel.	41
4.4	The mean angular error and the mean squared error as a function of the number of samples using uniform sample locations and the <i>linear</i> optimal kernel.	42
4.5	Illustration of the ideal antipodal sampling kernels as a function of the colatitude.	45

4.6	The mean angular error and mean squared error as a function of the number of samples using uniform sample locations and an exponential antipodally symmetric sampling kernel.	46
4.7	Illustration of signals generated by integrating orientations.	47
4.8	The mean angular error and mean squared error as a function of the number of samples when reconstructing integrated orientation.	49
4.9	Example of a sum of 2 great circles.	50
4.10	Illustration of the reconstruction process for great circles.	53
4.11	The mean angular error and mean squared error as a function of the number of samples for great circles.	54
5.1	Magnitude of the spherical harmonic coefficients of $Gg_2(\mathbf{u})$ with $G = G_y(0)$, $G = G_y(\pi/4)$, and $G = G_y(\pi/2)$	59
5.2	The angular error as a function of G , described in terms of its y and z rotation.	62
5.3	The angular error as a function of the signal to noise ratio for a signal with $K = 2$	63
5.4	The angular error as a function of the signal to noise ratio for a signal with $K = 3$	64
5.5	The value of $\zeta(h)$ as a function of the SNR for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, and $a_k = 1$ sampled using $N = 9$	68
5.6	The value of $\zeta(h)$ as a function of the SNR for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, $a_k = 1$ and $a_k = 1$ sampled using $N = 16$	69
5.7	The value of $\zeta(h)$ as a function of the number of samples N for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, and $a_k = 1$ with signal to noise ratio of 15dB.	70
5.8	The angular error as a function of the signal to noise ratio for a signal with $K = 2$ orientations.	72
6.1	Magnitude of the rotational harmonic coefficients associated with a b -value of 1000 s/mm ² , 3000 s/mm ² , and 10000 s/mm ²	78
6.2	Example of synthetic dMRI signals generated using the multi-tensor model for $K = 1, 2$ and $b=1000, 3000$ s/mm ²	79
6.3	Detection rate as a function of the separation angle between 2 fibers.	82
6.4	Fiber reconstruction examples for separation angles of 90, 60, and 50°.	83
6.5	Fiber reconstruction examples for separation angles of 40, 30, and 30°.	84
6.6	Angular error as a function of the signal to noise ratio for a signal with two fibers.	85
6.7	Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for 2 fibers fanning or merging.	86
6.8	Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for 2 fibers kissing.	87

6.9	Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for a curving fiber.	88
6.10	Recovered fiber orientations for the first region of interest.	90
6.11	Recovered fiber orientations for the second region of interest.	91
6.12	Full volume fiber tracking results on real diffusion MRI data.	92
A.1	Illustration of the evolution of the position of the points $\mathbf{u}_m$ using $N = 64$ points. The black and red dots represent the initial and final position of the points, respectively. The blue dots are the evolution of the position at each iteration.	98

List of Tables

3.1	Mean angular error and mean square error observed using equiangular samples and the <i>ideal</i> bandlimited sampling kernel.	29
3.2	Mean angular error and mean square error observed using equiangular samples and the <i>exponential</i> bandlimited sampling kernel.	32
4.1	Mean angular error and mean square error observed using equiangular samples and the exponential antipodally symmetric bandlimited sampling kernel. For errors reported as 0, the observed values where below machine precision ($\approx 10^{-16}$).	44
4.2	Mean angular error and mean square error observed using equiangular samples and the ideal bandlimited sampling kernel when sampling integrated orientations.	48
4.3	Mean angular error and mean square error observed using equiangular samples and the ideal bandlimited sampling kernel when sampling great circles.	52

Notation and Symbols

The notation and symbols used in this thesis are detailed here. As Carl de Boor says in *A practical guide to splines* (de Boor, 2001): “Readers will have come across some of them, perhaps most of them. Still, better to bore them now than to mystify them later.”

Vectors are denoted by lowercase bold symbols such as $\mathbf{x}$ or $\boldsymbol{\omega}$ and matrices are denoted by uppercase bold symbols such as $\mathbf{X}$ or $\boldsymbol{\Omega}$. Sets are identified with blackboard bold letters such as $\mathbb{A}$ and $\mathbb{B}$.

$\equiv$ is the symbol that indicates equal by definition. It must always be read from left to right. For example, $a \equiv b$ means that a is defined as b .

$$j \equiv \sqrt{-1}.$$

$\{a, b, c, \dots\}$ is defined as the set that contains the elements $a, b, c, \dots$

$\{x \in \mathbb{X} : f(x)\}$ is defined as the set of elements in $\mathbb{X}$ satisfying the formula $f(x)$.

$$\mathbb{N} \equiv \{1, 2, 3, \dots\}.$$

$$\mathbb{N}_0 \equiv \{0, 1, 2, 3, \dots\}.$$

$$\mathbb{Z} \equiv \{\dots, -1, 0, 1, \dots\}.$$

$\mathbb{R}$ is defined as the set of all real numbers.

$\mathbb{C}$ is defined as the set of all complex numbers.

$|x|$ is defined as the absolute value of $x \in \mathbb{C}$.

$\lfloor x \rfloor$ is defined as the floor of $x \in \mathbb{R}$, that is, the largest integer not greater than x .

$$\|\mathbf{x}\|_p \equiv \left(\sum_{n=1}^N |x_n|^p \right)^{1/p}.$$

$$\mathbb{S}^2 \equiv \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\|_2 = 1\}.$$

$\mathbb{SO}(3)$ is defined as the special orthogonal group in 3 dimensions, that is, the set of all 3×3 rotation matrices with determinant 1.

$\Re x$ and $\Im x$ are defined as the real and imaginary part of x , respectively.

$[a, b]$ is defined as the interval between a and b , with limits included. In a similar manner, $]a, b[$ is defined as the interval between a and b , with limits excluded.

$\bar{x}$ is defined as the complex conjugate of x .

$\mathbf{x}^t$ and $\mathbf{X}^t$ are defined as the transpose of the vector $\mathbf{x}$ and matrix $\mathbf{X}$, respectively.

$\mathbf{x}^*$ and $\mathbf{X}^*$ are the conjugate transpose of the complex vector $\mathbf{x}$ and complex matrix $\mathbf{X}$, respectively.

$\mathbf{X}^{-1}$ is the inverse of $\mathbf{X}$.

$\mathbf{X}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{X}$. When $\mathbf{X}$ has full column rank, it is $\mathbf{X}^+ = (\mathbf{X}^* \mathbf{X})^{-1} \mathbf{X}^*$.

$\mathcal{N}(\mathbf{X})$ is a matrix whose columns form a basis for the null space of $\mathbf{X}$.

$\text{vec}(\mathbf{X})$ is the vectorization operation that transforms an $M \times N$ matrix into a $MN \times 1$ vector by stacking its columns.

$\mathbf{0}$ is defined as a matrix of zeros. Its dimensions are determined by the context.

$\delta(x)$ is the Dirac delta function.

δ_{mn} is the Kronecker delta function. Its value is 1 if $m = n$ and 0 otherwise.

$D_\ell^{mn}(\mathbf{G})$ with $\mathbf{G} \in \mathbb{SO}(3)$ is the rotational harmonic of degree ℓ and orders m and n .

$Y_\ell^m(\mathbf{u})$ with $\mathbf{u} \in \mathbb{S}^2$ is the spherical harmonic of degree ℓ and order m . $\mathbf{Y}$ is the spherical harmonic transform matrix that projects spherical harmonic coefficients onto points on the sphere.

$\mathcal{U}(a, b)$ is the continuous uniform distribution with minimum $a \in \mathbb{R}$ and maximum $b \in \mathbb{R}$.

Acronyms

CS	Compressed Sensing
CSD	Constrained Spherical Deconvolution
dMRI	Diffusion Magnetic Resonance Imaging
FRI	Finite Rate of Innovation
HARDI	High Angular Resolution Diffusion Imaging
MRI	Magnetic Resonance Imaging
ODF	Orientation Distribution Function
ROI	Region of Interest
SFRI	Spherical Finite Rate of Innovation
SFRIG	Spherical Finite Rate of Innovation with rotations
SNR	Signal to Noise Ratio
sFRIdMRI	Spherical Finite Rate of Innovation for Diffusion Magnetic Resonance Imaging
super-CSD	Super-resolved Constrained Spherical Deconvolution

I

Introduction

As defined by [Prince and Links \(2006\)](#), “Signals are mathematical functions of one or more independent variables, capable of modeling a variety of physical processes.” Typical examples of signals are the temperature value as a function of the time of day or the pixel intensity as a function of its position in an image. Generally speaking, it is not possible to continually observe a signal. Instead, we are able to measure its value at a few discrete locations. This process is called sampling a signal and the resulting measurements are named samples. By observing the samples, we have an incomplete view of the signal; we know its value only at a few points. This leads us to the following question: given the samples, can we reconstruct the original signal? Rephrasing, given the value of a signal at a few points, can we infer its value at any point? Without any supplemental information on the class of the signal at hand, the answer is no. There are simply too many different signals that will yield the same set of samples. On the other hand, we are often able to make assumption about characteristics of the signal. For example, if the signal is bandlimited, meaning it contains no frequencies higher than some bandlimit L , the Nyquist-Shannon sampling theorem ([Shannon, 1949, 1998](#)) states that it is completely determined by $2L$ samples. A lot of progress has since been made and we now have sampling theorems for many signal classes. An excellent review of the evolution of sampling 50 years after Shannon’s seminal paper was published by [Unser \(2000\)](#). Of special interest to us is the work of [Vetterli et al. \(2002\)](#) which shows that a certain class of signals which have a finite number of degrees of freedom, called finite rate of innovation (FRI) signals, can be sampled and reconstructed perfectly. In this case, it is the fact that FRI signals have a parametric representation that is exploited.

As the astute reader will have noticed from the title of this thesis, we are specifically interested in signals that are defined on the sphere. This class of signals appears naturally in a multitude of fields including astrophysics ([Spergel et al., 2007](#)), acoustics ([Jarrett et al., 2012](#); [Ahrens and Spors, 2012](#)), computer graphics ([Kronander et al., 2012](#); [Sloan et al., 2002](#)), and medical imaging ([Alexander et al., 2002](#)) to name a few. Similarly to one dimensional signals, we are able to sample bandlimited signals on the

sphere. A succinct review of available theorems can be found in the introduction of the recent work of [McEwen and Wiaux \(2011\)](#). For non-bandlimited signals on the sphere, there are very few guidelines to follow. The typical approach is to acquire as many samples as possible and reconstruct a lowpass approximation of the signal. This is obviously suboptimal.

An example of non-bandlimited functions we wish to sample and reconstruct, and the primary motivation for our work, is found in Magnetic Resonance Imaging (MRI). One of its variants, Diffusion Magnetic Resonance Imaging (dMRI), allows us to observe the diffusion of water molecules in biological tissues. Because the motion of water is hampered by structures, for example axonal fiber tracts in the white matter of the brain, observing the dMRI signal can reveal the microstructure of the brain. On a larger scale, if we recover the orientation of the fiber tracts over the complete volume of the brain, then we can map how cortical areas of the brain are connected ([Mori et al., 1999](#); [Conturo et al., 1999](#)).

Given a small volume of tissue, called a voxel, we can acquire a dMRI signal in a 3D space called the q -space. Given a voxel and a q -space vector, we have a single number which represents all of the diffusion information within the small volume. Diffusion MRI signals are therefore 6D objects, depending on the x , y , and z position of the voxel and the 3 dimensions of the q -space. Some methods process the 6D information all at once ([Wedeen et al., 2005, 2008](#)), but we are interested here in those that treat each voxel independently. While each voxel can be processed independently, a full volume is still acquired for each q -space sample. This gives us a strong incentive to limit the number of samples used when detecting fibers, since it has a direct impact on the amount of time a patient must stay in the scanner. A popular method to reduce the amount of information while maintaining angular resolution is to acquire all the samples on a spherical shell in q -space. This is what is called High Angular Resolution Diffusion Imaging, or HARDI ([Tuch et al., 1999](#)). Several reconstruction techniques that deal with HARDI data have been developed, for example: the multi-tensor model ([Alexander et al., 2001](#); [Tuch et al., 2002](#); [Hosey et al., 2005](#); [Kreher et al., 2005](#); [Peled et al., 2006](#); [Behrens et al., 2007](#); [Jian et al., 2007](#); [Melie-García et al., 2008](#)), generalized tensor ([Ozarslan and Mareci, 2003](#); [Liu et al., 2004](#)), spherical harmonic decomposition ([Alexander et al., 2002](#)), diffusion Kurtosis imaging ([Jensen et al., 2005](#); [Lu et al., 2006](#); [Minati et al., 2007](#)), and q -ball imaging [Tuch \(2004\)](#); [Hess et al. \(2006\)](#); [Descoteaux et al. \(2007\)](#); [Haldar and Leahy \(2013\)](#). A review of several of these techniques can be found in ([Assemlal et al., 2011](#)).

To our knowledge, the angular resolution of all HARDI techniques currently available depends on the number q -space samples that are acquired. This is a severe limitation because neural fibers often cross at narrow angles, requiring a large number of samples to differentiate them. New sampling theorems for non-bandlimited functions on the sphere have the potential to improve the situation by breaking the dependence between the angular resolution and the number of samples. Such a break is possible because sampling theorems state the sufficient conditions to sample and reconstruct a class of signal. In other words, if the sufficient conditions are met, *any* signal within

this class can be sampled and reconstructed, even if it contains fibers crossing at narrow angles. Indeed, we will show in Chapter 6 that the theorems and algorithms developed in this thesis can improve the reconstruction of crossing fibers in dMRI. While the application considered in this work is dMRI, we highlight that the theorems we prove have a much wider applicability. Indeed, the discussion in Chapters 3, 4, and 5 is kept generic so that our results may be used in other applications.

1.1 Contributions

Non-bandlimited signals on the sphere appear in many applications, including diffusion magnetic resonance imaging, and are impossible to sample and reconstruct using current techniques. We produce new sampling theorems that prove that certain non-bandlimited signals on the sphere can be sampled using a finite number of samples and use them to improve the reconstruction of fiber orientations in diffusion magnetic resonance imaging.

In more detail, the main contributions of this thesis are:

- Providing a new model for finite rate of innovation signals on the sphere and proving that they can be sampled and reconstructed using a finite number of samples.
- Generalizing the sampling scheme by including different kernels and extending the class of finite rate of innovation signals that can be reconstructed to include great circles and signals integrated along the azimuth.
- Developing two reconstruction algorithms for finite rate of innovation signals on the sphere that are robust to noise and validating them using several numerical simulations.
- Showing how finite rate of innovations signals on the sphere can be used to model diffusion magnetic resonance imaging signals and improve their reconstruction.

1.2 Organisation of the thesis

The remainder of this thesis is organized as follows.

Chapter 2 contains the background information that we will use throughout this thesis. We first review the main ideas behind sampling signals at their rate of innovation. Then, we define the terminology associated with signals on the sphere and highlight some of their properties and operations. This includes a formal definition of what we mean by signals on the sphere and their related operations. Finally, we review the state of the art in sampling bandlimited signals on the sphere.

Chapter 3 contains our base model, which we call a sum of K orientations, for FRI signals on the sphere and show that it can be reconstructed from a subset of its spectral coefficients. We then show that these coefficients can be computed from filtered

samples, leading to a new sampling theorem for non-bandlimited signals. Finally, we validate our results using numerous numerical simulations and examples.

In Chapter 4, we extend the type of non-bandlimited signals by modifying our sampling scheme and signal model. This allows our approach to be used with a much wider range of signals. The main results include the derivation of an optimal sampling kernel that allows the reconstruction of orientations using the critical number of samples. We also show that antipodally symmetric kernels, which do not meet the requirements of Chapter 3, can be used to sample and reconstruct orientations. To generalize our signal model, we prove that a sum of orientations integrated along the azimuth can also be recovered using our methods. Furthermore, we develop an exact reconstruction algorithm for great circles, which are also non-bandlimited.

In Chapter 5, we reluctantly admit that the world is not perfect and propose modifications to our reconstruction algorithm that makes it robust to noise. The modifications fall into two categories: those that require oversampling and those that do not. We then put these modifications to the test in numerical simulations.

Finally, in Chapter 6, we investigate a possible application of our reconstruction theorems in the form of the recovery of fiber orientations in diffusion magnetic resonance imaging. Using numerical simulations we show that our methods are able to distinguish fibers even when they cross at a very narrow angle. We also qualitatively assess the performance of our algorithm on real diffusion MRI data by identifying anatomical landmarks in our reconstructions.

II

Background

In this Chapter, we present a set of notions which are essential to understand the content of this thesis. We first review the main ideas behind sampling signals at their rate of innovation. Next, we define the terminology associated with signals on the sphere and highlight some of their properties and operations. The definition and properties of spherical harmonics are also reviewed as they are fundamental to spectral analysis on the sphere. Rotational harmonics are also reviewed, as they are a powerful tool to describe rotations of functions on the sphere. Finally, we present the state of the art in sampling bandlimited signals on the sphere.

2.1 Sampling signals with a finite rate of innovation

Consider a τ -periodic stream of K Diracs defined as

$$x(t) = \sum_{n \in \mathbb{Z}} \sum_{k=1}^K a_k \delta(t - t_k - n\tau) \quad (2.1)$$

where $a_k \in \mathbb{R}$ and $t_k \in [0, \tau[$ represent the amplitudes and locations of the Diracs, respectively. This signal has an infinitesimal support in the time domain and an infinite support in the spectral domain, making it particularly difficult to sample and reconstruct. Its saving grace, as recognized by [Vetterli et al. \(2002\)](#), is that it is completely determined by a finite number of parameters, or degrees of freedom, per unit of time. Hence the name *signal with a Finite Rate of Innovation*, or FRI signal. The Fourier series coefficients of the signal $x(t)$ are

$$\hat{x}_m = \frac{1}{\tau} \sum_{k=1}^K a_k u_k^m \quad (2.2)$$

with $m \in \mathbb{Z}$ and $u_k = \exp(-j2\pi t_k/\tau)$. Observing (2.1) and (2.2), we see that $x(t)$ and $\hat{x}_m$ are both completely defined by $2K$ parameters. As a result, it is intuitive that

sampling and reconstructing this signal will require at least $2K$ measurements. The main results of the FRI literature show that $2K + 1$ measurements are in fact sufficient to recover the parameters of $x(t)$ and hence $x(t)$ itself. We briefly outline the major steps of the sampling and reconstruction algorithm here.

Due to its sparse nature, the signal $x(t)$ cannot be sampled directly and it is more convenient to consider its filtered version. The measurements are obtained by filtering the signal $x(t)$ with a periodic sinc kernel

$$\phi(t) = \frac{\sin(\pi Bt)}{B\tau \sin(\pi t/\tau)}$$

and sampling at a period $T_s = \tau/N$, namely,

$$\begin{aligned} y_n &= \langle x(t), \phi(nT_s - t) \rangle \\ &= \sum_{|m| \leq \lfloor B\tau/2 \rfloor} T_s \hat{x}_m \exp(j2\pi mn/N) \end{aligned} \quad (2.3)$$

with $n = 1, \dots, N$ and B is the bandwidth of the sinc kernel. Equation (2.3) highlights the Fourier relationship between $\hat{x}_m$ and y_n . It follows that the spectral components of the signal, $\hat{x}_m$ in (2.3), can be obtained by computing the discrete Fourier transform of y_n .

Using Prony's method (Kay and Marple, 1981; Kay, 1988; Marple, 1989), which is called the annihilating filter method in the context of FRI, Vetterli et al. (2002) showed that the locations t_k can be computed from the coefficients $\hat{x}_m$. The details are presented in Section 3.2 so we will simply state here that $2K$ contiguous values of $\hat{x}_m$ are sufficient to recover t_k for $k = 1, \dots, K$. Using these locations, the weights can be computed with from the linear system

$$\mathbf{a} = \mathbf{U}^{-1} \hat{\mathbf{x}} \quad (2.4)$$

with $\hat{\mathbf{x}} = [\hat{x}_1 \ \cdots \ \hat{x}_K]^t$, $\mathbf{a} = [a_1 \ \cdots \ a_K]^t$, and

$$\mathbf{U} = \begin{bmatrix} u_1^1 & \cdots & u_K^1 \\ \vdots & \ddots & \vdots \\ u_1^K & \cdots & u_K^K \end{bmatrix}.$$

The dimensions of $\hat{\mathbf{x}}$, $\mathbf{a}$, and $\mathbf{U}$ are $K \times 1$, $K \times 1$, and $K \times K$, respectively.

The sampling and reconstruction algorithm described above, which we will refer to simply as FRI, is optimal in a noiseless setting. It allows us to recover a signal $x(t)$ with $2K$ parameters using $2K + 1$ samples and has a low complexity.

2.2 Signals on the sphere

We say a signal lives on the sphere if it is defined on the set of points that lie on the unit sphere in three dimensional Euclidean space, which we denoted by $\mathbb{S}^2$. To be precise, we have

$$\mathbb{S}^2 \equiv \{\mathbf{u} \in \mathbb{R}^3 : \|\mathbf{u}\|_2 = 1\}.$$

Any point in $\mathbb{S}^2$ can be represented with a unit norm vector $\mathbf{u} \in \mathbb{R}^3$ or with the pair of angles (θ, φ) which we use to denote the colatitude and the azimuth, respectively. The two representations are related by

$$\mathbf{u} = \begin{bmatrix} \sin \theta \cos \varphi & \sin \theta \sin \varphi & \cos \theta \end{bmatrix}^t \quad (2.5)$$

which is simply a change from the Cartesian to the spherical coordinate system. We consider a function that maps a point on the sphere to a point on the complex plane, that is, $f : \mathbb{S}^2 \rightarrow \mathbb{C}$. To simplify the notation, we will interchangeably write $f(\mathbf{u})$ and $f(\theta, \varphi)$ where we assume both functions are equal if θ , φ , and $\mathbf{u}$ satisfy the relation of (2.5). The inner product of two functions on the sphere $f(\mathbf{u})$ and $g(\mathbf{u})$ is defined as

$$\begin{aligned} \langle f, g \rangle &\equiv \int_{\mathbb{S}^2} f(\mathbf{u}) \bar{g}(\mathbf{u}) d\mathbf{u} \\ &= \int_0^{2\pi} \int_0^\pi f(\theta, \varphi) \bar{g}(\theta, \varphi) \sin \theta d\theta d\varphi. \end{aligned} \quad (2.6)$$

The definition of the inner product leads to the norm $\|f\|_2 = \langle f, f \rangle^{1/2}$ which we use to define the space of square-integrable functions on the sphere $L^2(\mathbb{S}^2)$ that satisfy $\|f\|_2 < \infty$.

Several functions on the sphere are displayed throughout this thesis and we take a moment here to explain our representation. A function $f(\theta, \varphi)$ is illustrated by plotting a deformed sphere whose radius in one direction represents the magnitude of the real or imaginary component in that direction. The sign of the component is coded using color. For example, consider the function

$$f(\theta, \varphi) = \cos \theta$$

illustrated in Figure 2.1. In the direction $\theta = 0$ (up) the radius is 1, at $\theta = \pi/2$ (horizontal) the radius is 0, and at $\theta = \pi$ (down) the radius is -1 .

2.3 Fourier analysis on the sphere

As in the classical one dimensional case, the spectrum of a signal on the sphere is a powerful tool that can be used to investigate its properties. In this case, the spectrum is computed using the spherical harmonic transform, the natural extension of the Fourier transform. In this Section, we briefly review the theory of spherical harmonics and rotational harmonics that will be needed in the subsequent chapters. For a more complete review, we suggest the texts of [Varshalovich et al. \(1988\)](#), [Brink and Satchler \(1993\)](#), or [Edmunds \(1996\)](#).

2.3.1 Spherical harmonics

The spherical harmonic of degree ℓ and order m is defined as

$$Y_\ell^m(\theta, \varphi) = C_\ell^m P_\ell^m(\cos \theta) \exp(jm\varphi) \quad (2.7)$$

Figure 2.1: Illustration of the spherical function $f(\theta, \varphi) = \cos \theta$. The radius in a direction is the magnitude of the real component in that direction and the color indicates the sign.

where C_ℓ^m is a constant and $P_\ell^m(\cos \theta)$ is the associated Legendre polynomial of degree ℓ and order m . The associated Legendre polynomials are given by (Olver et al., 2010)

$$P_\ell^m(x) = (-1)^m (1-x)^{m/2} \frac{d^m}{dx^m} P_\ell(x) \quad (2.8)$$

with $\ell \in \mathbb{N}_0$, $0 \leq m \leq \ell$ and where

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell} (x^2 - 1)^\ell \quad (2.9)$$

is the Legendre polynomial of degree ℓ . Inserting (2.9) into (2.8), we find that for negative values of m

$$P_\ell^{-m}(x) = (-1)^m \frac{(\ell - m)!}{(\ell + m)!} P_\ell^m(x).$$

Finally, selecting the normalization constant

$$C_\ell^m = (-1)^m \left[\frac{2(\ell + 1)}{4\pi} \frac{(\ell - m)!}{(\ell + m)!} \right]^{1/2},$$

we find that

$$\bar{Y}_\ell^m = (-1)^m Y_\ell^{-m}$$

which follows the Condon and Shortley (1935) phase convention. Moreover, this normalization makes the spherical harmonics orthonormal with respect to both ℓ and m ,

$$\langle Y_\ell^m, Y_{\ell'}^{m'} \rangle = \delta_{\ell\ell'} \delta_{mm'}.$$

The real component of the spherical harmonics up to degree 2 are illustrated in Figure 2.2.

Figure 2.2: Illustration of the real component of the spherical harmonics up to degree 2.

The set of all spherical harmonics forms a basis for the space of square integrable functions on the sphere. As a direct consequence, any function $f(\theta, \varphi) \in L^2(\mathbb{S}^2)$ can be expanded as a sum of spherical harmonics

$$f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m Y_{\ell}^m(\theta, \varphi)$$

where the coefficients $\hat{f}_{\ell}^m$ are the spherical harmonic coefficients of $f(\theta, \varphi)$. They can be computed from

$$\hat{f}_{\ell}^m = \langle f, Y_{\ell}^m \rangle$$

which we will refer to as the *spherical Fourier transform* of $f(\theta, \varphi)$, as in the work of [Varshalovich et al. \(1988\)](#).

2.3.2 Signal approximation using spherical harmonics

Consider a function $f(\mathbf{u}) \in L^2(\mathbb{S}^2)$ that we wish to approximate using its value at the discrete locations $\mathbf{u}_n$ with $n = 1, \dots, N$. The samples f_n are given by

$$f_n = f(\mathbf{u}_n) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m Y_{\ell}^m(\mathbf{u}_n). \quad (2.10)$$

Let

$$\mathbf{Y}_{\ell}(\mathbf{u}) = \begin{bmatrix} Y_{\ell}^{-\ell}(\mathbf{u}) & Y_{\ell}^{-\ell+1}(\mathbf{u}) & \dots & Y_{\ell}^{\ell}(\mathbf{u}) \end{bmatrix} \quad (2.11)$$

and define

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_0(\mathbf{u}_1) & \mathbf{Y}_1(\mathbf{u}_1) & \dots & \mathbf{Y}_{L-1}(\mathbf{u}_1) \\ \mathbf{Y}_0(\mathbf{u}_2) & \mathbf{Y}_1(\mathbf{u}_2) & \dots & \mathbf{Y}_{L-1}(\mathbf{u}_2) \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{Y}_0(\mathbf{u}_N) & \mathbf{Y}_1(\mathbf{u}_N) & \dots & \mathbf{Y}_{L-1}(\mathbf{u}_N) \end{bmatrix} \quad (2.12)$$

which we refer to as the spherical harmonic transform matrix. To approximate $f(\mathbf{u})$, we can truncate the infinite sum at $L - 1$ and rewrite (2.10) in matrix form as

$$\mathbf{f} = \mathbf{Y} \hat{\mathbf{f}}$$

where

$$\mathbf{f} = \begin{bmatrix} f_1 & f_2 & \cdots & f_N \end{bmatrix}^t$$

and

$$\hat{\mathbf{f}} = \begin{bmatrix} \hat{f}_0 & \hat{f}_1 & \cdots & \hat{f}_{L-1} \end{bmatrix}^t$$

with $\hat{f}_\ell = \begin{bmatrix} \hat{f}_\ell^{-\ell} & \hat{f}_\ell^{-\ell+1} & \cdots & \hat{f}_\ell^\ell \end{bmatrix}$. To clarify, the dimensions of $\mathbf{f}$, $\hat{\mathbf{f}}$, and $\mathbf{Y}$ are respectively $N \times 1$, $L^2 \times 1$, and $N \times L^2$. We can then estimate the coefficients with

$$\hat{\mathbf{f}} = \mathbf{Y}^+ \mathbf{f}$$

where $\mathbf{Y}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{Y}$. The accuracy of this estimation will greatly depend on L , the number of samples N , and the sample locations $\mathbf{u}_n$. The approximation of a function using various values of L is illustrated in Figure 2.3.

While all square integrable functions on the sphere can be approximated by a finite sum of spherical harmonics, a subset of these can be exactly represented by a truncated sum. We call those bandlimited functions on the sphere.

Definition 2.1 (Bandlimited function on the sphere). *A function $f(\mathbf{u}) \in L^2(\mathbb{S}^2)$ on the sphere is bandlimited at or has a bandlimit L if $\hat{f}_\ell^m = 0$ for $\ell \geq L$.*

Bandlimited functions are of special interest because they can be completely described by a few coefficients. Indeed, if we know the spherical harmonic coefficients of the function up to the bandlimit, then we can compute the value of the function at any point using

$$f(\mathbf{u}) = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m Y_\ell^m(\mathbf{u}).$$

2.3.3 Rotational harmonics

A rotation in 3D space can be identified by three angles α , β , and γ with $\alpha, \gamma \in [0, 2\pi[$ and $\beta \in [0, \pi]$. Each of these angles represents a rotation about an axis in a fixed coordinate system. We opt for the zyz Euler convention where α , β , and γ are rotations about the z , y , and z axis, respectively. The three angles can also be combined into a 3×3 real orthogonal rotation matrix $\mathbf{G}$ with a determinant of one. The group of all rotation matrices $\mathbf{G}$ is denoted by $\mathbb{SO}(3)$. Analogously to functions on the sphere, functions on the rotation group can be written as $r(\alpha, \beta, \gamma)$ or $r(\mathbf{G})$. We will use the two notations interchangeably and assume that both functions are equal if (α, β, γ) and $\mathbf{G}$ are related by

$$\mathbf{G} = \mathbf{R}_z(\alpha) \mathbf{R}_y(\beta) \mathbf{R}_z(\gamma)$$

Figure 2.3: Approximation of a function on the sphere using spherical harmonics. The original function is illustrated in (a) and (b), (c), (d), and (e) show the approximations with $L = 2, 4, 6$, and 8 , respectively.

with

$$\mathbf{R}_z(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{R}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}$$

and $\alpha \in [0, 2\pi[$, $\beta \in [0, \pi[$, and $\gamma \in [0, 2\pi[$ are Euler angles.

The Wigner functions (Wigner, 1988), also called rotational harmonics, are the most well-known basis for square integrable functions on the rotation group. The rotational harmonic $D_\ell^{mn}(\alpha, \beta, \gamma)$ of degree ℓ and orders m and n is given by

$$D_\ell^{mn}(\alpha, \beta, \gamma) = \exp(-jm\alpha) d_\ell^{mn}(\cos \beta) \exp(-jn\gamma) \quad (2.13)$$

with $\ell \in \mathbb{N}_0$ and $|m|, |n| \leq \ell$. The real d -functions d_ℓ^{mn} can be expressed as (Healy et al., 1998)

$$d_\ell^{mn}(\cos \beta) = j^{m-n} \frac{\sin^{n-m} \beta (1 + \cos \beta)^m}{2^\ell ((\ell + m)!) (\ell - m)!^{1/2}} \left(\frac{(\ell - n)!}{(\ell + n)!} \right)^{1/2} \\ \times \frac{d^{\ell+n}}{d(\cos \beta)^{\ell+n}} (\cos \beta - 1)^{\ell+m} (\cos \beta + 1)^{\ell-m}.$$

The d -function can also be written recursively as (Nikiforov et al., 1991; Risbo, 1996)

$$d_\ell^{mn}(\theta) = j^{n-m} \sum_{k=-\ell}^{\ell} d_\ell^{km}(\pi/2) d_\ell^{kn}(\pi/2) \exp(jk\theta), \quad (2.14)$$

a property that we will use shortly. Because the Wigner functions form a basis for the space of square integrable functions on the rotation group, any function $r(\mathbf{G}) \in$

$L^2(\mathbb{SO}(3))$ can be expanded as

$$r(\mathbf{G}) = \sum_{\ell=0}^{\infty} \frac{(2\ell+1)}{8\pi^2} \sum_{m=-\ell}^{\ell} \sum_{n=-\ell}^{\ell} \hat{r}_{\ell}^{mn} \bar{D}_{\ell}^{mn}(\mathbf{G}) \quad (2.15)$$

where $\hat{r}_{\ell}^{mn}$ is the rotational harmonic coefficient of degree ℓ and orders m and n . They can be computed from

$$\hat{r}_{\ell}^{mn} = \int_{\mathbb{SO}(3)} r(\mathbf{G}) D_{\ell}^{mn}(\mathbf{G}) d\mathbf{G}. \quad (2.16)$$

Some functions on the rotation group can be expanded using a finite number of terms in (2.15).

Definition 2.2 (Bandlimited function on the rotation group). *A function $r(\mathbf{G}) \in L^2(\mathbb{SO}(3))$ on the rotation group is bandlimited at or has a bandlimit L if $\hat{r}_{\ell}^{mn} = 0$ for $\ell \geq L$.*

The rotational harmonics and the spherical harmonics are related by (Brink and Satchler, 1993)

$$Y_{\ell}^m(\mathbf{G}^{-1}\mathbf{u}) = \sum_{i=-\ell}^{\ell} Y_{\ell}^i(\mathbf{u}) D_{\ell}^{im}(\mathbf{G}). \quad (2.17)$$

and also

$$Y_{\ell}^m(\theta, \varphi) = \sqrt{\frac{2\ell+1}{4\pi}} \bar{D}_{\ell}^{m0}(\varphi, \theta, 0). \quad (2.18)$$

2.4 Spherical convolution

The convolution of a signal $f(\mathbf{u}) \in L^2(\mathbb{S}^2)$ on the sphere with a function $r(\mathbf{G}) \in L^2(\mathbb{SO}(3))$ on the rotation group is defined as

$$(f * r)(\mathbf{u}) = \int_{\mathbb{SO}(3)} f(\mathbf{G}^{-1}\mathbf{u}) r(\mathbf{G}) d\mathbf{G}. \quad (2.19)$$

This definition is analogous to the conventional Cartesian convolution, but instead of integrating over all translations, we integrate over all rotations. This operation can also be performed in harmonic space, as detailed by Healy et al. (1998). We present similar results here, using a slightly different approach. First, we can expand (2.19) using the spherical harmonic expansion of $f(\mathbf{u})$ and the identity in (2.17) to get

$$\begin{aligned} s(\mathbf{u}) &= \int_{\mathbb{SO}(3)} \left(\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m Y_{\ell}^m(\mathbf{G}^{-1}\mathbf{u}) \right) r(\mathbf{G}) d\mathbf{G} \\ &= \int_{\mathbb{SO}(3)} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m \left(\sum_{i=-\ell}^{\ell} Y_{\ell}^i(\mathbf{u}) D_{\ell}^{im}(\mathbf{G}) \right) r(\mathbf{G}) d\mathbf{G}. \end{aligned}$$

By rearranging the order of the sums and the integral, and using the definition of the rotational harmonic coefficients in (2.16) we have

$$\begin{aligned} s(\mathbf{u}) &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m \sum_{i=-\ell}^{\ell} Y_{\ell}^i(\mathbf{u}) \int_{\mathbb{SO}(3)} r(\mathbf{G}) D_{\ell}^{im}(\mathbf{G}) d\mathbf{G} \\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m \sum_{i=-\ell}^{\ell} Y_{\ell}^i(\mathbf{u}) r_{\ell}^{im} \\ &= \sum_{\ell=0}^{\infty} \sum_{i=-\ell}^{\ell} Y_{\ell}^i(\mathbf{u}) \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m r_{\ell}^{im}. \end{aligned}$$

Comparing this result to the spherical harmonic expansion of $s(\mathbf{u})$, we find that

$$s(\mathbf{u}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{s}_{\ell}^m Y_{\ell}^m(\mathbf{u}) \quad (2.20)$$

with

$$\hat{s}_{\ell}^m = \sum_{i=-\ell}^{\ell} \hat{f}_{\ell}^i \hat{r}_{\ell}^{mi} \quad (2.21)$$

which proves the analog of the convolution theorem of the Fourier transform (Oppenheim and Willisky, 1997) for the spherical harmonics. The relation of (2.21) can be written in matrix form as

$$\hat{\mathbf{s}}_{\ell} = \mathbf{R}_{\ell} \hat{\mathbf{f}}_{\ell}$$

for $\ell = 0, \dots, \infty$ where

$$\hat{\mathbf{s}}_{\ell} = \begin{bmatrix} \hat{s}_{\ell}^{-\ell} & \hat{s}_{\ell}^{-\ell+1} & \dots & \hat{s}_{\ell}^{\ell} \end{bmatrix}^t, \quad \hat{\mathbf{f}}_{\ell} = \begin{bmatrix} \hat{f}_{\ell}^{-\ell} & \hat{f}_{\ell}^{-\ell+1} & \dots & \hat{f}_{\ell}^{\ell} \end{bmatrix}^t,$$

and

$$\mathbf{R}_{\ell} = \begin{bmatrix} \hat{r}_{\ell}^{-\ell, -\ell} & \hat{r}_{\ell}^{-\ell, -\ell+1} & \dots & \hat{r}_{\ell}^{-\ell, \ell} \\ \hat{r}_{\ell}^{-\ell+1, -\ell} & \hat{r}_{\ell}^{-\ell+1, -\ell+1} & \dots & \hat{r}_{\ell}^{-\ell+1, \ell} \\ \vdots & \ddots & \ddots & \vdots \\ \hat{r}_{\ell}^{\ell, -\ell} & \hat{r}_{\ell}^{\ell, -\ell+1} & \dots & \hat{r}_{\ell}^{\ell, \ell} \end{bmatrix}.$$

If we assume the infinite sum in (2.20) is truncated at $L - 1$, then we can further simplify the notation and write

$$\hat{\mathbf{s}} = \mathbf{R} \hat{\mathbf{f}}$$

where

$$\hat{\mathbf{s}} = \begin{bmatrix} \hat{s}_0 \\ \hat{s}_1 \\ \vdots \\ \hat{s}_{L-1} \end{bmatrix}, \quad \hat{\mathbf{f}} = \begin{bmatrix} \hat{f}_0 \\ \hat{f}_1 \\ \vdots \\ \hat{f}_{L-1} \end{bmatrix}, \quad \text{and } \mathbf{R} = \begin{bmatrix} \mathbf{R}_0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_{L-1} \end{bmatrix}.$$

Note that in computations, it is more efficient to use $\hat{\mathbf{s}}_{\ell} = \mathbf{R}_{\ell} \hat{\mathbf{f}}_{\ell}$ with $\ell = 0, \dots, L - 1$ because $\mathbf{R}$ is block diagonal. An example of the convolution operation on the sphere

Figure 2.4: Example of the spherical convolution operation. The function on the sphere is illustrated in (a), the function on the rotation group in (b) and the result of the convolution in (c).

is illustrated in Figure 2.4. We can see that $s(\mathbf{u})$ is obtained by “blurring” $f(\mathbf{u})$ using $r(\mathbf{G})$. Note that functions on the rotation group are difficult to display on a 2D medium, so we instead display $(\delta * r)(\mathbf{u})$ where δ is a Dirac delta function along the positive z axis.

2.5 Sampling bandlimited signals on the sphere

2.5.1 Sampling schemes on the sphere

Before we discuss sampling functions on the sphere, we must first consider sampling schemes. More specifically, we must choose the location of our samples. In the 1D case, the most common approach is to choose a fixed sampling period T and take a sample every T seconds. To extend this idea to the sphere, we can distribute sample locations evenly along the colatitude and azimuth. The exact position of the samples are then easy to calculate as they lie on a discrete grid. This approach allows the use of fast transforms to compute the spherical harmonic transform, as we will discuss shortly. The drawback is that the grid is dense near the poles, but sparse at the equator. In other words, the samples do not all represent the same area of the sphere. For the remainder of this work, we will refer to measurements obtained using this sampling scheme as equiangular samples.

Definition 2.3 (Equiangular samples). *Equiangular samples are samples on the sphere whose locations are given by*

$$\theta_t = \frac{\pi(2t+1)}{2L-1}$$

where $t = 0, \dots, L-1$ and

$$\varphi_p = \frac{2\pi p}{2L-1}$$

where $p = 0, \dots, 2L-2$.

Note that the number of samples is restricted to $(L-1)(2L-1) + 1$ for some $L \in \mathbb{N}$.

Figure 2.5: Comparison of (a) equiangular and (b) uniform locations on the sphere for 106 locations.

Another approach is to distribute the samples uniformly on the surface of the sphere. For an arbitrary number of samples, computing the exact position of points on the sphere that minimise a cost function is still a research problem in mathematics (Chazelle, 2000; Schwartz, to appear 2013). However, if we do not require provably optimal positions, then numerical methods are available that will compute almost uniform locations, for example Crittenden and Turok (1998); Jones et al. (1999); Górski et al. (2005); Doroshkevich et al. (2005). We can also get near uniform samples using the tessellation of the isocahedron (White et al., 1992). In this work, when we refer to uniform samples, we will use sample locations generated using a specific implementation of the electrostatic repulsion model, which is described in detail in Appendix A.

For comparison, equiangular sample locations and uniform sample locations are illustrated in Figure 2.5 for a sampling scheme with $N = 106$ samples.

2.5.2 Sampling theorems for bandlimited signals

As we discussed in the previous sections, bandlimited signals can be expanded as a finite sum of spherical harmonics. The task of sampling a bandlimited function $s(\mathbf{u})$ therefore revolves around evaluating the integral

$$\hat{s}_\ell^m = \langle s, Y_\ell^m \rangle = \int_{\mathbb{S}^2} s(\mathbf{u}) \bar{Y}_\ell^m(\mathbf{u}) d\mathbf{u} \quad (2.22)$$

for $\ell = 0, \dots, L-1$ and $-\ell \leq m \leq \ell$. The use of quadrature formula was proposed by Shukowsky (1986) as an improvement to the discrete least squares approach (Fougere, 1966). Another exact algorithm was proposed by Driscoll and Healy (1994) and was later revisited by Healy et al. (2003) to improve its stability (Kostelec and Rockmore, 2008). Fast algorithms were proposed by Potts et al. (1998), Kunis and Potts (2003). Some authors also focus on multi-resolution reconstruction of bandlimited signal such as Wiaux et al. (2008). Recently, McEwen (2011) developed a new sampling theorem

on the sphere which reduces the number of samples from $\approx 4L^2$ to $\approx 2L^2$. In addition, the algorithm is numerically stable and does not require pre-computation. Since this method has clear advantages over its predecessors, it will be our reference algorithm when computing spherical harmonic coefficients of bandlimited functions. Because we will make heavy use of these results, we reproduce the proof of [McEwen and Wiaux \(2011\)](#) here.

Theorem 2.1 (Sampling a bandlimited function on the sphere). *If $s(\mathbf{u}) \in L^2(\mathbb{S}^2)$ has a bandlimit L then $(L-1)(2L-1)+1$ equiangular samples are sufficient to recover $s(\mathbf{u})$ exactly.*

Proof. We start with the definition of the spherical harmonic coefficients of $s(\theta, \varphi)$

$$\hat{s}_\ell^m = \int_0^{2\pi} \int_0^\pi s(\theta, \varphi) \bar{Y}_\ell^m(\theta, \varphi) \sin \theta d\theta d\varphi$$

and expand the spherical harmonics using the relation (2.18) to get

$$\hat{s}_\ell^m = \sqrt{\frac{2\ell+1}{4\pi}} \int_0^{2\pi} \int_0^\pi s(\theta, \varphi) D_\ell^{m0}(\varphi, \theta, 0) \sin \theta d\theta d\varphi.$$

Replacing the rotational harmonics by their definition given in (2.13) we get

$$\hat{s}_\ell^m = \sqrt{\frac{2\ell+1}{4\pi}} \int_0^{2\pi} \int_0^\pi s(\theta, \varphi) d_\ell^{m0}(\theta) \exp(-jm\varphi) \sin \theta d\theta d\varphi$$

and finally substituting the Fourier series decomposition of the small- d functions given in (2.14) to obtain

$$\hat{s}_\ell^m = j^m \sqrt{\frac{2\ell+1}{4\pi}} \sum_{k=-(L-1)}^{L-1} d_\ell^{km}(\pi/2) d_\ell^{k0}(\pi/2) g_{mk}$$

where

$$g_{mk} = \int_0^\pi g_m(\theta) \exp(-jk\theta) \sin \theta d\theta \quad (2.23)$$

and

$$g_m(\theta) = \int_0^{2\pi} s(\theta, \varphi) \exp(-jm\varphi) d\varphi. \quad (2.24)$$

Computing the spherical harmonic coefficients boils down to evaluating (2.23) and (2.24) exactly. Equation (2.24) can be expressed exactly by

$$g_m(\theta_t) = \frac{2\pi}{2L-1} \sum_{p=-(L-1)}^{L-1} s(\theta_t, \varphi_p) \exp(-jm\varphi_p) \quad (2.25)$$

since $g_m(\theta)$ is simply the Fourier transform of $s(\theta, \varphi)$ for a constant θ . To compute (2.23), we first use the same substitution as above to get

$$s(\theta, \varphi) = \sum_{m=-(L-1)}^{L-1} q_m(\theta) \exp(-jm\varphi) \quad (2.26)$$

with

$$q_m(\theta) = \sum_{m'=-L+1}^{L-1} q_{mm'} \exp(-jm'\theta) \quad (2.27)$$

and

$$q_{mm'} = j^{-m} \sum_{\ell=0}^{L-1} \sqrt{\frac{2\ell+1}{4\pi}} d_\ell^{m'm}(\pi/2) d_\ell^{m'0}(\pi/2) \hat{s}_\ell^m.$$

Substituting (2.26) into (2.25) gives

$$g_m(\theta) = 2\pi q_m(\theta) \quad (2.28)$$

and noting that

$$q_m(-\theta) = (-1)^m q_m(\theta)$$

we define a periodic extension of $g_m(\theta)$ whose discrete version is

$$\tilde{g}_m(\theta) = \begin{cases} g_m(\theta_t) & t \in \{0, \dots, L-1\} \\ (-1)^m g_m(\theta_{2L-2-t}) & t \in \{L, \dots, L-2\} \end{cases}. \quad (2.29)$$

Substituting (2.29) in (2.27) and inverting we get

$$q_{mm'} = \frac{1}{2\pi(L-1)} \sum_{t=-L+1}^{L-1} \tilde{g}_m(\theta_t) \exp(-jm'\theta_t) \quad (2.30)$$

and finally we substitute (2.30) into (2.23) to get

$$\begin{aligned} g_{mm'} &= \int_0^\pi \tilde{g}_m(\theta) \exp(-jm'\theta) \sin \theta d\theta \\ &= 2\pi \sum_{m''=-L+1}^{L-1} q_{mm''} w(m'' - m') \end{aligned}$$

where

$$\begin{aligned} w(m) &= \int_0^\pi \exp(jm'\theta) \sin \theta d\theta \\ &= \begin{cases} \frac{\pm j\pi}{2} & m = \pm 1 \\ 0 & m \text{ odd and } m \neq \pm 1 \\ \frac{2}{(1-m'^2)} & m \text{ even.} \end{cases} \end{aligned}$$

To compute the spherical harmonics, we must first evaluate $g_{mm'}$ for all m and m' which requires $\tilde{g}_m(\theta_t)$ with $t = 0, \dots, L-2$. Because $\tilde{g}_m(\theta_t)$ is a periodic extension of $g_m(\theta_t)$ we need $g_m(\theta_t)$ for $t = 0, \dots, L-1$. Finally each $g_m(\theta_t)$ required $s(\theta_t, \varphi_p)$ for $p = -(L-1), \dots, L-1$. We conclude that the minimum number of samples required to compute all the spherical harmonic coefficients is $(L-1)(2L-1) + 1$. $\square$

III

Sampling FRI signals on the sphere

As we discussed in the introduction, the work of [Marziliano \(2001\)](#); [Vetterli et al. \(2002\)](#); [Blu et al. \(2008\)](#) shows that certain non-bandlimited signals can be sampled and reconstructed using a finite number of samples. In this case, it is the fact that FRI signals have a parametric representation that is exploited. The typical approach is to compute the inner product of a signal of interest $f(t)$ with a known sampling kernel $h(t)$ to yield the observed signal $s(t)$. Given an appropriate sampling kernel, originally a sinc or Gaussian, the parameters of $f(t)$ can be recovered from a uniform set of samples of $s(t)$. These results were extended to other kernels including polynomial and exponential reproducing kernels ([Dragotti et al., 2007](#)) and to 2D signals ([Baboulaz and Dragotti, 2009](#); [Chen et al., 2012](#); [Maravić and Vetterli, 2004](#); [Shukla and Dragotti, 2007](#)). While signals on the sphere are essentially 2D, they rely on different definitions of inner product, convolution, and harmonic transform, making the current theory inapplicable. To illustrate this idea, consider the case where $f(t)$ is a stream of K Diracs. The observed signal $s(t)$ is then a sum of K scaled and *translated* versions of $h(t)$ as illustrated in Figure 3.1 (a). This is appropriate in Cartesian space but not on the sphere where the convolution involves a rotation. Indeed, we want s to be obtained by summing K scaled and *rotated* versions of h as illustrated in Figure 3.1 (b). In this chapter, we show that certain non-bandlimited signals on the sphere can be exactly recovered using a finite number of samples and an appropriate sampling scheme.

In Section 3.1, we define our signal of interest, namely orientations, and prove some of their properties. In Sections 3.2 and 3.3, we derive a sampling theorem for orientations that makes use of equiangular samples. In Section 3.4, we present the reconstruction algorithm and perform numerical simulations to validate our results. A comparison between the results of this Section and those of compressed sensing on the sphere is presented in Section 3.5.

Figure 3.1: Comparison of the FRI signal model in Cartesian and spherical coordinates. In (a), the signal $s(t)$ is obtained by translating and summing $h(t)$. In (b), the signal on the right is obtained by rotating and summing the sampling kernel on the left.

3.1 Spherical FRI signals

As presented in Section 2.1, the archetype finite rate of innovation signal is the τ -periodic stream of K Diracs given by

$$f(t) = \sum_{n \in \mathbb{Z}} \sum_{k=1}^K a_k \delta(t - t_k - n\tau)$$

where $\delta(\cdot)$ is the Dirac delta function. This signal is completely defined by $2K$ parameters and the work of [Vetterli et al. \(2002\)](#) indeed shows that $f(t)$ can be exactly recovered from $2K + 1$ samples. The extension of signals with a finite rate of innovation to the sphere has the form

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k) \quad (3.1)$$

where $a_k \in \mathbb{R}$, $\theta_k \in [0, \pi]$, and $\varphi_k \in [0, 2\pi[$. The variables θ and φ are respectively the colatitude and azimuth, as described in the preceding chapter. At this point, the use of $\cos \theta$ instead of θ may be perplexing, but it will be justified shortly. For the rest of this thesis, we will refer to signals of the form of (3.1) as *the sum of K orientations*. A sum of 2 orientations is illustrated in Figure 3.2.

Figure 3.2: Illustration of a signal with 2 orientations whose parameters are $a_1 = a_2 = 1$, $\theta_1 = 2.67$, $\theta_2 = 2.93$, $\varphi_1 = 4.12$, $\varphi_2 = 0.22$.

Because of their infinitesimal support, orientations cannot be sampled directly. As in the work of [Vetterli et al. \(2002\)](#), we propose a sampling scheme where the signal of interest is first convolved with a sampling kernel on the rotation group to yield the observed signal which is then sampled. This scheme is presented in Figure 3.3. The sum of K orientations $f(\mathbf{u})$ is convolved with a sampling kernel $r(\mathbf{G}) \in L^2(\mathbb{SO}(3))$, that is,

$$s(\mathbf{u}) = \int_{\mathbb{SO}(3)} f(\mathbf{G}^{-1}\mathbf{u})r(\mathbf{G})d\mathbf{G}. \quad (3.2)$$

The observed signal $s(\mathbf{u})$ is then sampled at the locations $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N\}$. Our objective is to recover $f(\mathbf{u})$ given the knowledge of $r(\mathbf{G})$ and K .

The convolution of the signal of interest $f(\mathbf{u})$ with the sampling kernel is difficult to implement as part of an acquisition system. The problem is that $f(\mathbf{u})$ is non-bandlimited and has a very small support in the space domain. Therefore, the convolution cannot be implemented efficiently and accurately in the space or harmonic domains. Nonetheless, the signal model of (3.2) is relevant because it occurs naturally in applications, notably in diffusion magnetic resonance imaging. In this context, $f(\mathbf{u})$ corresponds to the distribution of fibers in a voxel and $r(\mathbf{G})$ corresponds to the response of a single fiber. This particular application will be discussed further in Chapter 6. More generally, the model of (3.2) fits signals that can be modeled as a the sum of K rotations of a single function.

The reconstruction algorithm we propose proceeds in two steps. First, the spherical harmonic coefficients of $f(\mathbf{u})$ are recovered from the samples $s_n = s(\mathbf{u}_n)$ for $n =$

Figure 3.3: The sampling scheme considered in this work. The orientations $f(\theta, \varphi)$ are convolved with a sampling kernel $r(\mathbf{G})$ to yield the observed signal $s(\mathbf{u})$ which is then sampled. The samples s_n are used to recover the spherical harmonic (SH) coefficients $\hat{f}_\ell^m$ which are in turn used to find the signal parameters a_k, θ_k , and φ_k .

1, ..., N . Second, the parameters a_k, θ_k , and φ_k of the orientations are computed from the spherical harmonic coefficients. Because the first step depends on the number of coefficients needed in the second step, it is more intuitive to present the second step first. In Section 3.2, we show that a small set of spherical harmonic coefficients of $f(\mathbf{u})$ are sufficient to find its parameters. We then show in Section 3.3 that these coefficients can be obtained from the samples s_n with $n = 1, \dots, N$ where N depends on the sample locations and on the sampling kernel $r(\mathbf{G})$.

3.2 Signal parameters from spherical harmonics

In this section, we assume the spherical harmonic coefficients $\hat{f}_\ell^m$ of the orientations are available. We will derive an algorithm to recover the parameters of the orientation

using the coefficients. Specifically, we will prove the following.

Theorem 3.1. *Let $f(\theta, \varphi)$ be the sum of K orientations*

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k)$$

where $a_k \in \mathbb{R}$, $0 \leq \theta, \theta_k \leq \pi$, and $0 \leq \varphi, \varphi_k < 2\pi$. The spherical harmonic coefficients $\hat{f}_\ell^\ell$ with $\ell = 0, \dots, 2K - 1$ and $\hat{f}_\ell^{\ell-1}$ with $\ell = 1, \dots, K$ are sufficient to recover $f(\theta, \varphi)$.

Proof. Consider the spherical harmonics coefficient of the orientations

$$\begin{aligned} \hat{f}_\ell^m &= \langle f, Y_\ell^m \rangle \\ &= \int_0^{2\pi} \int_0^\pi f(\theta, \varphi) \bar{Y}_\ell^m(\theta, \varphi) \sin \theta d\theta d\varphi. \end{aligned}$$

We can replace $f(\theta, \varphi)$ and $Y_\ell^m(\theta, \varphi)$ by their respective definitions in (3.1) and (2.7) to get

$$\begin{aligned} \hat{f}_\ell^m &= C_\ell^m \sum_{k=1}^K a_k \left[\int_0^\pi \delta(\cos \theta - \cos \theta_k) P_\ell^m(\cos \theta) \sin \theta d\theta \right. \\ &\quad \left. \times \int_0^{2\pi} \delta(\varphi - \varphi_k) \exp(-jm\varphi) d\varphi \right] \\ &\stackrel{(a)}{=} C_\ell^m \sum_{k=1}^K a_k \left[\int_{-1}^1 \delta(\cos \theta - \cos \theta_k) P_\ell^m(\cos \theta) d(\cos \theta) \right. \\ &\quad \left. \times \int_0^{2\pi} \delta(\varphi - \varphi_k) \exp(-jm\varphi) d\varphi \right] \end{aligned}$$

where (a) is obtained by noting that $\sin \theta d\theta = -d(\cos \theta)$ (which justifies our use of $\cos \theta$ instead of θ) and adjusting the integral bounds. Using the sifting property of the Diracs to evaluate the integrals we find that the spherical harmonic coefficients of orientations are sums of spherical harmonics evaluated at the location of the orientations, that is,

$$\begin{aligned} \hat{f}_\ell^m &= C_\ell^m \sum_{k=1}^K a_k P_\ell^m(\cos \theta_k) \exp(-jm\varphi_k) \\ &= \sum_{k=1}^K a_k \bar{Y}_\ell^m(\theta_k, \varphi_k). \end{aligned} \tag{3.3}$$

To recover the parameters of $f(\mathbf{u})$, we exploit the form of specific spherical harmonics, namely those of equal degree and order, given by

$$\hat{f}_\ell^\ell = C_\ell^\ell \sum_{k=1}^K a_k P_\ell^\ell(\cos \theta_k) \exp(-j\ell\varphi_k). \tag{3.4}$$

The associated Legendre polynomial with equal degree and order in the preceeding equation is given by

$$P_\ell^\ell(\cos \theta) = -1^\ell \sin^\ell \theta \frac{d^\ell}{d(\cos \theta)} P_\ell(\cos \theta). \quad (3.5)$$

Noting from (2.9) that $P_\ell(\cos \theta)$ is a polynomial of degree ℓ and that its ℓ^{th} derivative is a constant, we can reduce (3.5) to

$$P_\ell^\ell(\cos \theta) = B_\ell \sin^\ell \theta$$

where B_ℓ is a constant. Inserting this result in (3.4) we get

$$\hat{f}_\ell^\ell = C_\ell^\ell B_\ell \sum_{k=1}^K a_k \sin^\ell \theta_k \exp(-j\ell\varphi_k).$$

To simplify the notation, we define $r_k = \sin \theta_k \exp(-j\varphi_k)$ and $z_\ell = \hat{f}_\ell^\ell / (C_\ell^\ell B_\ell)$ and get the normalized coefficients

$$z_\ell = \sum_{k=1}^K a_k r_k^\ell. \quad (3.6)$$

The normalized samples z_ℓ are of special interest because they are a linear combination of exponentials. We can thus use the annihilating filter method (Vetterli et al., 2002) to find the values of a_k and r_k . Because this step is essential to our recovery algorithm, we outline the annihilating filter method here.

Consider a filter whose coefficients are $\mathbf{h} = [h_0 \ \cdots \ h_K]^t$ and with a z -transform

$$H(z) = \prod_{k=1}^K (1 - r_k z^{-1}) = \sum_{k=0}^K h_k z^{-k}.$$

The roots of this filter correspond to the values of r_k and are uniquely determined by the coefficients in $\mathbf{h}$. Conversely, the coefficients are not uniquely defined by the roots. This minor annoyance can be rectified by adding a constraint, for example $\|\mathbf{h}\|_2 = 1$ or $h_0 = 1$, as proposed by Vetterli et al. (2002). To compute the coefficients, we observe the discrete convolution

$$\begin{aligned} (\mathbf{h} * \mathbf{z})_n &= \sum_{i=0}^K h_i z_{n-i} \\ &= \sum_{i=0}^K h_i \sum_{k=1}^K a_k r_k^{n-i} \\ &= \sum_{k=1}^K a_k r_k^n \sum_{i=0}^K h_i r_k^{-i} \\ &= \sum_{k=1}^K a_k r_k^n H(r_k) = 0 \end{aligned} \quad (3.7)$$

where the equality to zero is obtained because $H(r_k)$ is the filter being evaluated at one of its roots. For $n = K, \dots, 2K - 1$, we can rewrite the linear system of equations of (3.7) as

$$\begin{bmatrix} z_K & z_{K-1} & \cdots & z_0 \\ z_{K+1} & z_K & \cdots & z_1 \\ \vdots & \ddots & \ddots & \vdots \\ z_{2K-1} & z_{2K-2} & \cdots & z_{K-1} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_K \end{bmatrix} = \mathbf{Z}\mathbf{h} = 0. \quad (3.8)$$

The data matrix $\mathbf{Z}$ is Toeplitz so the linear system above yields a unique $\mathbf{h}$ up to a constant factor, provided that the values of r_k are distinct. If we assume that $h_0 = 1$, the linear system in (3.8) can be written as

$$\begin{bmatrix} z_{K-1} & \cdots & z_0 \\ z_K & \cdots & z_1 \\ \vdots & \ddots & \vdots \\ z_{2K-2} & \cdots & z_{K-1} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix} = - \begin{bmatrix} z_K \\ z_{K+1} \\ \vdots \\ z_{2K-1} \end{bmatrix}$$

and solved for the coefficients $h_1, \dots, h_K$. If we instead choose the constraint $\|\mathbf{h}\|_2 = 1$, the coefficients of the annihilating filter can be computed with a singular value decomposition. Indeed, $\mathbf{h}$ is the left eigenvector corresponding to the vanishing eigenvalue. The values of r_k can then be obtained by recalling that they correspond to the roots of the annihilating filter.

To recover the weights a_k , we rewrite the linear system of equations of (3.6) for $\ell = 0, \dots, K - 1$ which yields

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1^1 & r_2^1 & \cdots & r_K^1 \\ \vdots & \ddots & \ddots & \vdots \\ r_1^{K-1} & r_2^{K-1} & \cdots & r_K^{K-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{bmatrix} = \begin{bmatrix} z_0 \\ z_1 \\ \vdots \\ z_{K-1} \end{bmatrix} = \mathbf{R}\mathbf{a} = \mathbf{z}.$$

The matrix $\mathbf{R}$ is a square and full rank Vandermonde so the system above will yield a unique solution $\mathbf{a} = \mathbf{R}^{-1}\mathbf{z}$, provided that the values of r_k are unique. In a numerical implementation, the inversion of $\mathbf{R}$ must be done carefully as Vandermonde matrices can be ill-conditioned when the values of r_k are similar or the matrix is large. However, it should be noted that the inverse of a Vandermonde can be computed explicitly (Macon and Spitzbart, 1958) therefore mitigating the numerical numerical instability of the procedure. Furthermore, the matrices considered in this work are typically small, $K = 10$ being the largest matrix considered.

Now that the values of r_k and a_k are known, we can extract the parameters θ_k and φ_k . By recalling that $r_k = \sin \theta_k \exp(-j\varphi_k)$, we see that $\sin \theta_k$ and $-\varphi$ correspond to the magnitude and phase of r_k , respectively. However, because $\theta_k \in [0, \pi]$ the knowledge of $\sin \theta_k$ does not uniquely determine the value of θ_k . To compensate, we now turn to the spherical harmonics coefficients whose order is one less than its degree.

For $m = \ell - 1$ these coefficients have the form

$$\hat{f}_\ell^{\ell-1} = C_\ell^{\ell-1} \sum_{k=1}^K a_k P_\ell^{\ell-1}(\cos \theta_k) \exp(-j\varphi_k(\ell-1))$$

where

$$P_\ell^{\ell-1}(\cos \theta) = (-1)^{\ell-1} \sin^{\ell-1} \theta \frac{d^{\ell-1}}{d(\cos \theta)^{\ell-1}} P_\ell(\cos \theta). \quad (3.9)$$

The Legendre polynomial $P_\ell(\cos \theta)$ in (3.9) simplifies to

$$P_\ell^{\ell-1}(\cos \theta) = D_\ell \sin^{\ell-1} \theta \cos \theta,$$

for some constant D_ℓ , by observing that $P_\ell(\cos \theta)$ is a polynomial of degree ℓ so its $(\ell-1)^{\text{th}}$ derivative is proportional to $\cos \theta$. Making use of this simplification, we can rewrite the spherical harmonic coefficients whose order is one less than its degree as

$$\begin{aligned} \hat{f}_\ell^{\ell-1} &= C_\ell^{\ell-1} D_\ell \sum_{k=1}^K a_k \sin^{\ell-1} \theta_k \cos \theta_k \exp(-j\varphi_k(\ell-1)) \\ &= C_\ell^{\ell-1} D_\ell \sum_{k=1}^K a_k \cos \theta_k r_k^{\ell-1}. \end{aligned}$$

Defining the normalized coefficients as $w_\ell = \hat{f}_\ell^{\ell-1} / (C_\ell^{\ell-1} D_\ell)$ we have the linear system of equations

$$w_\ell = \sum_{k=1}^K a_k \cos \theta_k r_k^{\ell-1}$$

where the unknowns are the K values of $\cos \theta_k$. For $\ell = 1, \dots, K$ the system is written as

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1^1 & r_2^1 & \cdots & r_K^1 \\ \vdots & \ddots & \ddots & \vdots \\ r_1^{K-1} & r_2^{K-1} & \cdots & r_K^{K-1} \end{bmatrix} \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \vdots & a_K \end{bmatrix} \begin{bmatrix} \cos \theta_1 \\ \cos \theta_2 \\ \vdots \\ \cos \theta_K \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{K-1} \end{bmatrix} \quad (3.10)$$

or more compactly $\mathbf{R} \mathbf{A} \mathbf{c} = \mathbf{w}$. Because of the Vandermonde structure in $\mathbf{R}$, this system yields a unique solution $\mathbf{c} = (\mathbf{R} \mathbf{A})^{-1} \mathbf{w}$. Because $\theta_k \in [0, \pi]$, the values of $\cos \theta_k$ uniquely define the values of θ_k and we have now recovered all the parameters of $f(\mathbf{u})$.

In summary, we need the coefficients $\hat{f}_\ell^\ell$ for $\ell = 0, \dots, 2K-1$ to compute $2K$ values of z_ℓ . We then use the annihilating filter method to recover the values of a_k and φ_k . Finally, we need the coefficients $\hat{f}_\ell^{\ell-1}$ for $\ell = 1, \dots, K$ to compute K values of w_ℓ and solve a linear system to get the values of $\cos \theta_k$ and thus θ_k .

□

3.3 Bandlimited kernels

Theorem 3.1 states that we can perfectly recover a sum of K orientations, given the appropriate spherical harmonic coefficients. In this section, we investigate how these coefficients can be recovered from the observed samples. Recall our sampling scheme where the observed signal is given by the spherical convolution of the orientations with a sampling kernel $r(\mathbf{G})$ defined on the rotation group, that is,

$$s(\mathbf{u}) = \int_{\mathbb{SO}(3)} f(\mathbf{G}^{-1}\mathbf{u})r(\mathbf{G})d\mathbf{G}.$$

As was shown in Section 2.4, the convolution can be computed in the harmonic domain as

$$s(\mathbf{u}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{s}_{\ell}^m Y_{\ell}^m(\mathbf{u}) \quad (3.11)$$

with

$$\hat{s}_{\ell}^m = \sum_{i=-\ell}^{\ell} \hat{f}_{\ell}^i \hat{r}_{\ell}^{mi}. \quad (3.12)$$

Equation (3.12) gives us a straightforward relation between the spherical harmonic coefficients of the observed signal and the orientations, given the rotational harmonics of the sampling kernel. However, we still have to deal with the infinite sum in (3.11). To truncate the sum, we define a bandlimited rotational sampling kernel.

Definition 3.1 (Bandlimited rotational sampling kernel). *A bandlimited sampling kernel is a function $r(\mathbf{G}) \in \mathbb{SO}(3)$ whose rotational harmonic coefficients are given by*

$$\hat{r}_{\ell}^{mn} = \begin{cases} \hat{r}_{\ell}^m & \ell < L \text{ and } m = n \\ 0 & \text{otherwise} \end{cases} \quad (3.13)$$

with $\hat{r}_{\ell}^m \neq 0$ and $L > 0$.

Note that all bandlimited sampling kernels are bandlimited functions on the rotation group but the converse is not true because of the additional constraint $\hat{r}_{\ell}^m \neq 0$. A special case of the bandlimited rotational sampling kernel is when $\hat{r}_{\ell}^m = 1$ which we will define as the *ideal* bandlimited sampling kernel. Inserting a bandlimited sampling kernel in (2.21) leads to $\hat{s}_{\ell}^m = \hat{f}_{\ell}^m \hat{r}_{\ell}^m$ for $\ell < L$ and 0 otherwise. We can then use this result in (2.20) to get

$$\begin{aligned} s(\mathbf{u}) &= \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{s}_{\ell}^m Y_{\ell}^m(\mathbf{u}) \\ &= \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m \hat{r}_{\ell}^m Y_{\ell}^m(\mathbf{u}). \end{aligned}$$

Clearly the spherical harmonic coefficients of $s(\mathbf{u})$ below the bandlimit are those of $f(\mathbf{u})$ scaled by the kernel. We also notice that $s(\mathbf{u})$ is bandlimited at L meaning that

its spherical harmonic coefficients can be exactly recovered using a finite number of samples. An approach which guarantees the exact recovery of the spherical harmonic coefficients is to use equiangular samples as discussed in Section 2.5.2.

Recall that [McEwen and Wiaux \(2011\)](#) proved that to recover a function with a bandlimit L on the sphere we need $(L-1)(2L-1)+1$ equiangular samples. Now, according to Theorem 3.1, we need spherical harmonic coefficients up to degree $2K-1$ to recover the parameters of a sum of K orientations. Therefore, the minimum bandlimit of $r(\mathbf{G})$ is $L = 2K$. This proves the following corollary.

Corollary 3.1 (Bandlimited kernel and equiangular samples). *Let $f(\theta, \varphi)$ be the sum of K orientations*

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k)$$

where $a_k \in \mathbb{R}$, $0 \leq \theta, \theta_k \leq \pi$, and $0 \leq \varphi, \varphi_k < 2\pi$. Let $r(\mathbf{G})$ be a bandlimited sampling kernel with a bandlimit $L = 2K$. The equiangular samples $s_n = (f * r)(\mathbf{u}_n)$ with $n = 1, \dots, (L-1)(2L-1)+1$ are sufficient to recover $f(\theta, \varphi)$ exactly.

The use of equiangular samples has two important advantages: it gives strong theoretical guarantees and it allows the use of [McEwen and Wiaux \(2011\)](#) algorithm which is fast and handles large bandlimits. The downside is that it is suboptimal in the number of samples. A signal with a bandlimit L has L^2 non-zero spherical harmonic coefficients but we need approximately $2L^2$ equiangular samples to recover it. To our knowledge, no sampling theorem reaches the lower bound of L^2 . However, in practice it is feasible to *accurately* recover the spherical harmonic coefficients of a bandlimited function using just L^2 samples. This has long been recognized by the diffusion magnetic resonance imaging community where the acquisition of superfluous samples is expensive. A common approach is to distribute the sample locations (almost) uniformly on the sphere using the electron repulsion model ([Jones et al., 1999](#)) or tessellation of the isocahedron. We then solve the linear system

$$\mathbf{s} = \mathbf{Y} \hat{\mathbf{s}}$$

which requires that $\mathbf{Y}$ has full column rank to yield a unique solution. Empirical results show that $N \geq L^2$ uniform samples will yield a well conditioned matrix $\mathbf{Y}$ and the spherical harmonic coefficients can be computed from

$$\hat{\mathbf{s}} = \mathbf{Y}^+ \mathbf{s} = (\mathbf{Y}^* \mathbf{Y})^{-1} \mathbf{Y}^* \mathbf{s}. \quad (3.14)$$

While this approach reduces the number of samples used for the accurate recovery of the spherical harmonic coefficients, it does not deal well with large bandlimits. For example, if $L = 1000$, then $\mathbf{Y}$ has dimensions $10^6 \times 10^6$ which common computers will not be able to store, much less invert. We emphasize again that there are no theoretical guarantees that $N = L^2$ uniform samples will yield a full rank column matrix $\mathbf{Y}$ even if

it is observed in practice. Nonetheless, the potential reduction in samples is significant and the performance of this approach will be investigated in the following section.

In (3.14), we propose to use the Moore-Penrose pseudoinverse to find the solution to the problem

$$\underset{\hat{s}}{\text{minimize}} \quad \|\mathbf{Y}\hat{s} - s\|_2^2$$

having the minimum norm $\|\hat{s}\|_2$. Because of the convex nature of this problem, many other algorithms are available (Boyd, 2004). These algorithms may have specific advantages, such as lower memory requirements or higher efficiency. They may also be used to regulate the problem by adding constraints. In this work, we have opted to use a pseudoinverse because of our target application. In dMRI, calculation are typically performed on a voxel per voxel basis, allowing us to reuse the matrix $\mathbf{Y}^+$ repeatedly. Because a dMRI volume may contain millions of voxels, the initial cost of computing the pseudoinverse becomes negligible. This may not be the case for other applications and specific implementations should be considered.

3.4 Numerical simulations

3.4.1 Metrics and signal generation

To validate the algorithms proposed in Section 3.3, we perform a number of numerical simulations in this section. The quality of a reconstruction can be evaluated using the mean angular error defined as

$$\alpha = \frac{180}{\pi K} \sum_{k=1}^K \cos \mathbf{v}_k^t \mathbf{v}_k' \quad (3.15)$$

where $\mathbf{v}_k = [\sin \theta_k \cos \varphi_k \quad \sin \theta_k \sin \varphi_k \quad \cos \theta_k]^t$ is the true k^{th} orientation and $\mathbf{v}_k'$ is the recovered k^{th} orientation. In addition, we consider the reconstruction mean square error as

$$\beta = \frac{1}{N} \sum_{n=1}^N [s(\mathbf{u}_n) - \hat{s}(\mathbf{u}_n)]^2. \quad (3.16)$$

To generate the samples, we will consider two bandlimited sampling kernels: the ideal bandlimited sampling kernel

$$\hat{r}_\ell^{mn} = \begin{cases} 1 & \ell < L \text{ and } m = n \\ 0 & \text{otherwise} \end{cases}$$

and an exponential kernel

$$\hat{r}_\ell^{mn} = \begin{cases} \exp(-6\ell/L) & \ell < L \text{ and } m = n \\ 0 & \text{otherwise} \end{cases}.$$

Note that the use of an exponential kernel is arbitrary and was added for illustrative purposes. Any other kernel whose rotational harmonic coefficients are non-zero below the

$K (N)$	1 (4)	2 (22)	5 (172)	10 (742)
α	1.61×10^{-7}	1.40×10^{-7}	2.27×10^{-7}	7.82×10^{-3}
β	0.000	0.000	0.000	0.078

Table 3.1: Mean angular error (α) and mean square error (β) observed using equiangular samples and the *ideal* bandlimited sampling kernel. For errors reported as 0, the observed values were below machine precision ($\approx 10^{-16}$).

bandlimit could be used in its place. The ideal bandlimited and exponential sampling kernels are illustrated in Figure 3.4 for various bandlimits. Given a sampling kernel, a number of orientations K , and a number of samples N we perform the following steps to generate sample values:

1. Choose either equiangular or uniform samples locations $\{\mathbf{u}_n\}$ for $n = 1, \dots, N$. For uniform sample locations, we use the algorithm described in Appendix A.
2. Generate the parameters a_k for $k = 1, \dots, K$ from a standard normal distribution.
3. Generate the parameters θ_k and φ_k for $k = 1, \dots, K$ from uniform distributions $\mathcal{U}(0, \pi)$ and $\mathcal{U}(0, 2\pi)$, respectively.
4. Set $L = 2K$ and compute the samples using

$$s_n = s(\mathbf{u}_n) = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m \hat{r}_{\ell}^m Y_{\ell}^m(\mathbf{u}_n)$$

where

$$\hat{f}_{\ell}^m = \sum_{k=1}^K a_k \bar{Y}_{\ell}^m(\theta_k, \varphi_k).$$

3.4.2 Results

The first series of tests we perform consider equiangular samples. A typical reconstruction using the ideal kernel is illustrated in Figure 3.5. The mean angular error and mean square error computed over 500 simulations are reported in Table 3.1 and 3.2 for the ideal and exponential kernel, respectively. As expected from Corollary 3.1, the angular errors are extremely small even for a relatively large value $K = 10$. The amplitude of the orientations are also very well recovered, as indicated by the low mean squared error. The large difference between the performance of the ideal and exponential kernels for the case $K = 10$ was caused by a single simulation. Because no constraints are placed on the proximity of the θ_k and φ_k parameters, two fibers were generated with near identical parameters for the exponential kernel simulations. This caused the annihilating filter to have near identical roots which were not precisely recovered.

The second series of tests evaluate the performance of using uniformly distributed samples. Figure 3.6 illustrates the angular error as a function of the number of samples

Figure 3.4: The two kernels used in the numerical simulations, illustrated for bandlimits of $L = 2$ in (a), $L = 6$ in (b) and $L = 10$ in (c).

Figure 3.5: Illustration of the reconstruction process using equiangular sample locations and an ideal sampling kernel for a signal with $K = 3$.

$K (N)$	1 (4)	2 (22)	5 (172)	10 (742)
α	1.66×10^{-7}	1.29×10^{-7}	2.05×10^{-5}	7.35×10^{-3}
β	0.000	0.000	1.23×10^{-14}	8.40×10^{-6}

Table 3.2: Mean angular error (α) and mean square error (β) observed using equiangular samples and the *exponential* bandlimited sampling kernel. For errors reported as 0, the observed values were below machine precision ($\approx 10^{-16}$).

for $K = 2, 3, 4$, and 5 using the ideal kernel. Figure 3.7 illustrates the same information, but using the exponential kernel. We observe that the angular error is very near 0 as soon as the number of measurements satisfies $N \geq L^2$ which is consistent with the observations in Section 3.3. The same is observed for the mean squared error, which falls below 10^{-3} when $N \geq L^2$. The spikes in the mean squared error observed when $N \geq L^2$ are caused by the orientations that are randomly generated very close to each other (that is, those that are separated by less than 1 degree). When this occurs, the annihilating filter has nearly identical roots which causes numerical problems when recovering the signal parameters. However, in all cases observed, the resulting angular error never exceeded 0.01 degree. The reconstruction algorithm using uniform samples is illustrated in Figure 3.8.

3.5 Comparison with compressed sensing

Advances in compressed sensing have shown that certain signals can be sampled and reconstructed using fewer samples than stated by standard sampling theorems. Because signal models considered in this Chapter are inherently sparse, one can assume that using the theory of compressed sensing will yield performance gains. However, the field of compressed sensing is not as mature on the sphere as it is in Cartesian space. Nonetheless, some ideas of compressed sensing have been extended to spherical signals. We compare these techniques to our own here.

To use the ideas of compressed sensing (CS), the signal of interest must be sparse or compressible. [Rauhut and Ward \(2011\)](#) and [Alem et al. \(2012, 2013\)](#) consider signals that are sparse in the harmonic domain. In both cases, this constraint is applied by assuming the signal of interest is bandlimited. In this work, we consider signals of the form of (3.1) which are sparse in the space domain and therefore non-bandlimited. However, the results of [Rauhut and Ward \(2011\)](#) and [Alem et al. \(2012, 2013\)](#) may be used to sample the bandlimited signal $s(\mathbf{u})$ in (3.2) more efficiently. Because the reconstruction with CS is possible with high probability and not guaranteed, these results do not improve on Corollary 3.1.

Another popular approach in CS is to assume the total variation of a signal is sparse. This idea was exploited for signals on the sphere by [McEwen et al. \(2011\)](#). They illustrate the potential of CS on an inpainting problem using a binary Earth map as a test signal. While they assume sparsity in the space domain, their assumption of a

(a)

(b)

Figure 3.6: The angular error (α) in (a) and mean squared error (β) in (b) of spherical FRI using uniform locations, an ideal kernel, and no noise.

(a)

(b)

Figure 3.7: The angular error (α) in (a) and mean squared error (β) in (b) of spherical FRI using uniform locations, an exponential kernel, and no noise.

Figure 3.8: Illustration of the reconstruction process using uniform sample locations and an ideal sampling kernel for a signal with $K = 2$.

sparse total variation does not fit with signals of the form of (3.2).

3.6 Summary

In this section, we have presented the following:

1. We defined a sum of K orientations as

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k)$$

and showed that their spherical harmonic coefficients are given by

$$\hat{f}_\ell^m = \sum_{k=1}^K a_k \bar{Y}_\ell^m(\theta_k, \varphi_k).$$

2. We proved that the parameters of orientations can be exactly recovered from a subset of their spherical harmonics. Specifically, we need the $3K$ coefficients $\hat{f}_\ell^\ell$ with $\ell = 0, \dots, 2K - 1$ and $\hat{f}_\ell^{\ell-1}$ with $\ell = 1, \dots, K$.
3. We proved that if K orientations are filtered using a bandlimited sampling kernel with a bandlimit $L = 2K$, then $(L - 1)(2L - 1) + 1$ equiangular samples are sufficient to recover the orientations.
4. We showed, through numerical simulations, that $N = 4K^2$ uniform samples are sufficient to recover K orientations very accurately using a bandlimited sampling kernel.

IV

Spherical FRI variations

The theorem and corollary presented in the previous chapter apply to orientations filtered using a bandlimited sampling kernel. Because the usefulness of a sampling theorem is proportional to the number of signal classes it applies to, we generalize our results here. Specifically, we consider different sampling kernels and signal models that can also be sampled at their rate of innovation.

The rest of this chapter is organised as follows. In Section 4.1, we derive an optimal kernel that allows the accurate recovery of orientations using the critical number of samples. In Section 4.2 we consider antipodally symmetric kernels, a case which is not covered by the theory of the previous chapter. In Section 4.3 we modify our signal model to consider orientation integrated along the azimuth. Finally, in Section 4.4 we show that great circles can also be sampled at their rate of innovation.

4.1 Designing the optimal kernel

Theorem 3.1 gives the minimum number of samples needed to recover orientation parameters from spherical harmonic coefficients. This theorem is optimal because it requires $3K$ coefficients to retrieve $3K$ parameters. However, using a bandlimited kernel, we need approximately $8K^2$ equiangular or $4K^2$ uniform samples to recover the spherical harmonic coefficients. Intuitively, we should be able to reduce the number of samples to $3K$, the number of degrees of freedom. In this section, we will derive an optimal sampling kernel that allows us to sample the sum of K orientations at the critical rate. These results are of course highly theoretical, because it is unlikely that we will encounter the situations outlined in this section in practical applications. Nonetheless, it is interesting to derive the optimal sampling conditions for a class of signals.

To derive the optimal filter, we first recall that to recover the parameters of K orientations we need the coefficients $\hat{f}_\ell^\ell$ for $\ell = 0, \dots, 2K - 1$ and $\hat{f}_\ell^{\ell-1}$ for $\ell = 1, \dots, K$ as stated by Theorem 3.1. Note that in Section 3.3, more coefficients were used because

we assumed the use of bandlimited sampling kernels. However, not all computed coefficients were required for the exact reconstruction of K orientations. To reduce the number of samples needed, we need to “eliminate” all superfluous coefficients. We therefore need a kernel whose rotational harmonic coefficients filter out any coefficients that are not needed. Recall the convolution property

$$\hat{s}_\ell^m = \sum_{i=-\ell}^{\ell} \hat{f}_\ell^i \hat{r}_\ell^{mi}.$$

An easy way to achieve our goal is to create a one-to-one correspondence between $\hat{s}_\ell^m$ and the needed coefficients $\hat{f}_\ell^m$. To do this we use the *optimal* sampling kernel, whose name comes from the fact that it allows us to sample and reconstruct orientations using the minimum number of samples $3K$.

Definition 4.1 (Optimal sampling kernel). *An optimal sampling kernel is a kernel whose rotational harmonics are given by*

$$\hat{r}_\ell^{mn} = \begin{cases} \hat{r}_\ell^\ell & \ell < L \text{ and } m = n = \ell \\ \hat{r}_\ell^{\ell-1} & 0 < \ell \leq L/2 \text{ and } m = n = \ell - 1 \\ 0 & \text{otherwise} \end{cases}$$

for $L > 0$ with L even.

Using this new definition, we find that

$$\hat{s}_\ell^{mn} = \begin{cases} \hat{f}_\ell^\ell \hat{r}_\ell^\ell & \ell < L \text{ and } m = n = \ell \\ \hat{f}_\ell^{\ell-1} \hat{r}_\ell^{\ell-1} & 0 < \ell \leq L/2 \text{ and } m = n = \ell - 1 \\ 0 & \text{otherwise} \end{cases}$$

and the spherical harmonic expansion of $s(\mathbf{u})$ is

$$s(\mathbf{u}) = \sum_{\ell=0}^{L-1} \hat{s}_\ell^\ell Y_\ell^\ell(\mathbf{u}) + \sum_{\ell=1}^{L/2} \hat{s}_\ell^{\ell-1} Y_\ell^{\ell-1}(\mathbf{u}). \quad (4.1)$$

In matrix form, this is $\mathbf{s} = \hat{\mathbf{Y}} \hat{\mathbf{s}}$ where $\hat{\mathbf{Y}}$ is a $N \times 3K$ matrix that contains only the columns of $\mathbf{Y}$ where $m = \ell$ or $m = \ell - 1$. For this system to yield a unique solution, $\hat{\mathbf{Y}}$ must have full column rank which means $N \geq 3K$. Again, there are no theoretical guarantees that it is possible to build a well conditioned matrix $\hat{\mathbf{Y}}$ from just $N = 3K$ sample locations. However, the numerical simulations of the next section show that $3K$ samples are indeed sufficient to recover $\hat{\mathbf{s}}$, and hence the signal parameters, accurately.

To validate our claims concerning the optimal sampling kernel, we generate signals using the same procedure as in Section 3.4. In this case, we use an ideal optimal kernel

$$\hat{r}_\ell^{mn} = \begin{cases} 1 & \ell < L \text{ and } m = n = \ell \\ 1 & 0 < \ell \leq L/2 \text{ and } m = n = \ell - 1 \\ 0 & \text{otherwise} \end{cases}$$

Figure 4.1: Illustration of the ideal optimal sampling kernels as a function of the colatitude. The bandlimits are $L = 4$ in (a) and $L = 10$ in (b).

and a linear optimal kernel

$$\hat{r}_\ell^{mn} = \begin{cases} 1 - \ell/L & \ell < L \text{ and } m = n = \ell \\ 1 - 2\ell/(L+1) & 0 < \ell \leq L/2 \text{ and } m = n = \ell - 1 \\ 0 & \text{otherwise} \end{cases}.$$

These two kernels are illustrated in Figures 4.1 and 4.2 for different bandlimits and rotations. Figures 4.3 and 4.4 illustrate the mean angular error (α) and mean squared error (β) function of the number of samples for the ideal and linear optimal kernels, respectively. The sharp drop in both angular error and squared error occurs when $N = 3K$ in all cases as predicted. It is at this point that the number of samples becomes sufficient to solve the linear system in (4.1) accurately. This transition is similar to the “phase transition” observed in compressed sensing (Donoho et al., 2009). Figures 4.3 and 4.4 illustrate that although exact recovery using $3K$ samples is not guaranteed theoretically, accurate reconstruction is observed in practice when using optimal kernels. In our experiments, the choice of the kernel, as long as it was optimal, did not affect the results.

4.2 Antipodally symmetric kernels

In Section 3.3, our definition of a bandlimited rotational sampling kernel required that the rotational harmonic coefficients be non-zero below the bandlimit. This ensured that all the spherical harmonic coefficients of the orientations could be computed from the spherical harmonics of the observed signal. It also meant that the sampling kernels could not exhibit antipodal symmetry because when a function is symmetric, its rota-

Figure 4.2: Illustration of the linear optimal sampling kernels as a function of the colatitude. The bandlimits are $L = 4$ in (a) and $L = 10$ in (b).

tional harmonics are 0 for odd degrees. There are many situations where the kernel is dictated by the process and is symmetric, a notable example being diffusion magnetic resonance imaging (Tournier et al., 2004). Although the orientations cannot be retrieved, we would still like to recover the directions¹ using a minimal number of samples.

Consider a kernel whose rotational harmonic coefficients with even degrees are zero above a bandlimit and whose rotational harmonic coefficients with odd degrees are all zero.

Definition 4.2 (Antipodally symmetric bandlimited sampling kernel). *An antipodally symmetric bandlimited sampling kernel is a kernel whose rotational harmonics are given by*

$$\hat{r}_{\ell}^{mn} = \begin{cases} \hat{r}_{\ell}^m & \ell \text{ even, } \ell < L, \text{ and } m = n \\ 0 & \text{otherwise} \end{cases}$$

where $\hat{r}_{\ell}^m \neq 0$, $L > 0$, and L odd.

Inserting the kernel in Definition 4.2 into the spherical harmonic expansion of $s(\mathbf{u})$ in

¹To clarify, we consider that two antipodal points on the sphere have the same direction, but different orientations.

Figure 4.3: The mean angular error and the mean squared error as a function of the number of samples using uniform sample locations and the *ideal* optimal kernel.

Figure 4.4: The mean angular error and the mean squared error as a function of the number of samples using uniform sample locations and the *linear* optimal kernel.

Equation (2.20), we have

$$\begin{aligned} s(\mathbf{u}) &= \sum_{\ell=0}^{(L-1)/2} \sum_{m=-2\ell}^{2\ell} \hat{s}_{2\ell}^m Y_{2\ell}^m(\mathbf{u}) \\ &= \sum_{\ell=0}^{(L-1)/2} \sum_{m=-2\ell}^{2\ell} \hat{f}_{2\ell}^m \hat{r}_{2\ell}^m Y_{2\ell}^m(\mathbf{u}) \end{aligned} \quad (4.2)$$

which is also symmetric. We observe that this signal could also have been obtained by assuming symmetric orientations, that is,

$$\hat{f}_{\ell}^m = \begin{cases} \sum_{k=1}^K a_k Y_{\ell}^m(\theta_k, \varphi_k) & \ell \text{ even} \\ 0 & \ell \text{ odd} \end{cases}.$$

Because the spherical harmonic coefficients of the orientations are only available for even degrees, we only have access to the even normalized coefficients

$$z_{\ell} = \begin{cases} \sum_{k=1}^K a_k r_k^{\ell} & \ell \text{ even} \\ 0 & \ell \text{ odd} \end{cases}.$$

We must therefore recover the parameters of $s(\mathbf{u})$ using z_{ℓ} for the even values of ℓ . To do so, consider the annihilating filter $\mathbf{h} = [h_0 \ h_1 \ \dots \ h_{2K}]$ whose z -transform is

$$H(z) = \prod_{k=1}^K (1 - r_k z^{-1})(1 + r_k z^{-1}).$$

Notice that this filter has the coefficients $h_i = 0$ for odd i . We can then simplify its z -transform to

$$H(z) = \sum_{i=0}^K h_{2i} z^{-2i}.$$

The discrete convolution of $\mathbf{h}$ with $\mathbf{z} = [z_0 \ z_2 \ \dots \ z_{4K-2}]$ is

$$\begin{aligned} (\mathbf{h} * \mathbf{z})_n &= \sum_{m=0}^K h_{2m} z_{n-2m} \\ &= \sum_{m=0}^K h_{2m} \left(\sum_{k=1}^K a_k r_k^{n-2m} \right) \\ &= \sum_{k=1}^K a_k z_k^n \sum_{m=0}^K h_{2m} r_k^{-2m} = 0 \end{aligned}$$

where the equality to zero is again obtained because the right hand sum is the filter being evaluated at one of its roots. Using $n = 2K, 2K+2, 2K+4, \dots, 4K-2$ which requires z_{ℓ} for $\ell = 0, 2, \dots, 4K-2$ we can build a linear system as in Section 3.2 to

recover the annihilating filter coefficients. Once again, the roots of $\mathbf{h}$ uniquely define the values of r_k and an additional set of roots which correspond to $-r_k$. Because we assume antipodal symmetry of $f(\theta, \varphi)$ we can discard all the values corresponding to $-r_k$. To recover the values of a_k and θ_k , we proceed as described in Section 3.2.

The procedure above requires the spherical harmonics $\hat{f}_\ell^\ell$ with $\ell = 0, 2, \dots, 4K - 2$ and $\hat{f}_\ell^{\ell-1}$ with $\ell = 2, 4, \dots, 2K$. This means that the bandwidth of the kernel must be adjusted to satisfy $L \geq 4K - 1$. We again appeal to the theorem of [McEwen and Wiaux \(2011\)](#) to recover all the spherical coefficients $\hat{s}_\ell^m$ using the equiangular samples s_n for $n = 1, \dots, (L - 1)(2L - 1) + 1$. We therefore have the following result.

Corollary 4.1. (*Antipodally symmetric sampling kernel*) *Let $f(\theta, \varphi)$ be the sum of K orientations*

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k)$$

where $a_k \in \mathbb{R}$ and $0 \leq \theta, \theta_k, \varphi, \varphi_k \leq \pi$. Let $r(\mathbf{G})$ be an antipodally symmetric sampling kernel with a bandlimit $L = 4K - 1$. The equiangular samples $s_n = (f * r)(\mathbf{u}_n)$ with $n = 1, \dots, (L - 1)(2L - 1) + 1$ are sufficient to recover $f(\theta, \varphi)$ exactly.

Corollary 4.1 guarantees the perfect recovery of $f(\theta, \varphi)$ given equiangular samples. As in the previous sections, in practice it is possible to reduce the number of samples and still recover the parameters of f accurately. Consider the linear system in (4.2) in matrix form $\mathbf{s} = \mathbf{Y} \hat{\mathbf{s}}$. The odd degree coefficients in $\hat{\mathbf{s}}$ are all zero and the corresponding columns of $\mathbf{Y}$ are not involved in the computations. We can therefore define a matrix $\hat{\mathbf{Y}}$ of dimensions $L(L + 1)/2 \times N$ which contains only the spherical harmonics with even degree. We can then estimate the spherical harmonic coefficients with

$$\hat{\mathbf{s}} = (\hat{\mathbf{Y}}^* \hat{\mathbf{Y}})^{-1} \hat{\mathbf{Y}}^* \mathbf{s}.$$

Again, our numerical experiments, presented in the next section, show that the system above is well conditioned for $N \geq L(L + 1)/2$ and that the parameters of $f(\theta, \varphi)$ are accurately recovered.

To validate our results, we use an exponential antipodally symmetric sampling kernel

$$\hat{r}_\ell^{mn} = \begin{cases} \exp(-(\ell + m)/L) & \ell \text{ even, } \ell < L, \text{ and } m = n \\ 0 & \text{otherwise} \end{cases}$$

to generate our signals. This kernel is illustrated in Figure 4.5. Table 4.1 reports the angular error and mean squared error using equiangular samples. As expected, the parameters are recovered to machine precision. Figure 4.6 illustrate the same performance measures, but using uniform samples. Again, we observe a sharp drop in both the mean angular error and mean squared error when $N \geq L(L + 1)/2$.

4.3 Integrated orientations

Similar to FRI results on streams of Diracs which can be extended to piecewise polynomials, our results on the sphere can also be extended to other non-bandlimited signals.

Figure 4.5: Illustration of the ideal antipodal sampling kernels as a function of the colatitude. The bandlimits are $L = 4$ in (a) and $L = 10$ in (b).

K (N)	1 (11)	2 (79)	5 (667)	10 (2927)
α	1.63×10^{-7}	1.92×10^{-7}	2.15×10^{-4}	0.5874
β	0.000	0.000	3.08×10^{-6}	0.9280

Table 4.1: Mean angular error and mean square error observed using equiangular samples and the exponential antipodally symmetric bandlimited sampling kernel. For errors reported as 0, the observed values were below machine precision ($\approx 10^{-16}$).

Consider signals of the form

$$\frac{\partial}{\partial \varphi} F(\theta, \varphi) = f(\theta, \varphi) \quad (4.3)$$

where $f(\theta, \varphi)$ is our usual sum of K orientations given by (3.1). Note that $F(\theta, \varphi)$ and $f(\theta, \varphi)$ have the same parameters and number of degrees of freedom. One such signal is illustrated in Figure 4.7.

By replacing $F(\theta, \varphi)$ in (4.3) by its spherical harmonics expansion and noting that

$$\frac{d}{d\varphi} Y_\ell^m(\theta, \varphi) = jm Y_\ell^m(\theta, \varphi)$$

(a)

(b)

Figure 4.6: The mean angular error and mean squared error as a function of the number of samples using uniform sample locations and an exponential antipodally symmetric sampling kernel.

Figure 4.7: Illustration of signals generated by integrating orientations. The original orientations are presented in (a) and their lowpass version in (b). These orientations are integrated along the azimuth to produce the signal in (c).

we get

$$\begin{aligned}
 f(\theta, \varphi) &= \frac{d}{d\varphi} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{F}_{\ell}^m Y_{\ell}^m(\theta, \varphi) \\
 &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{F}_{\ell}^m \frac{d}{d\varphi} Y_{\ell}^m(\theta, \varphi) \\
 &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} jm \hat{F}_{\ell}^m Y_{\ell}^m(\theta, \varphi)
 \end{aligned} \tag{4.4}$$

By comparing (4.4) and the spherical harmonic expansion of $F(\theta, \varphi)$, we observe that

$$\hat{f}_{\ell}^m = jm \hat{F}_{\ell}^m$$

for $m \neq 0$. More generally, if we have

$$\frac{\partial^n}{\partial \varphi^n} F(\theta, \varphi) = f(\theta, \varphi)$$

then $\hat{f}_{\ell}^m = (jm)^n \hat{F}_{\ell}^m$ for $m \neq 0$. Now if we consider the filtered signal $s(\mathbf{u}) = (F * r)(\mathbf{u})$ where the kernel $r(\mathbf{G})$ has a bandlimit L we find that

$$s(\theta, \varphi) = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{F}_{\ell}^m \hat{r}_{\ell}^m Y_{\ell}^m(\theta, \varphi). \tag{4.5}$$

As before, the spherical harmonics coefficients $\hat{F}_{\ell}^m$ and thus $\hat{f}_{\ell}^m$ can be computed using Theorem 2.1. Because the computation of $\hat{f}_{\ell}^m$ for $m = 0$ involves a division by

$K (N)$	1 (11)	2 (37)	5 (211)	10 (821)
α	2.50×10^{-7}	1.56×10^{-7}	1.76×10^{-5}	0.381
β	0.000	0.000	0.000	8.10×10^{-3}

Table 4.2: Mean angular error and mean square error observed using equiangular samples and the ideal bandlimited sampling kernel when sampling integrated orientations. For errors reported as 0, the observed values were below machine precision ($\approx 10^{-16}$).

0, we use the spherical harmonics coefficients $\hat{f}_\ell^\ell$ with $\ell = 1, \dots, 2K$ and $\hat{f}_\ell^{\ell-1}$ with $\ell = 2, \dots, K+1$ to recover all the signal parameters a_k , θ_k , and φ_k . We therefore have the following results.

Corollary 4.2 (Reconstruction of integrated orientations). *Let $f(\theta, \varphi)$ be the sum of K orientations*

$$f(\theta, \varphi) = \sum_{k=1}^K a_k \delta(\cos \theta - \cos \theta_k) \delta(\varphi - \varphi_k)$$

and define $F(\theta, \varphi)$ such that

$$\frac{\partial^n}{\partial \varphi^n} F(\theta, \varphi) = f(\theta, \varphi).$$

Let $r(\mathbf{G}) \in L^2(\mathbb{S}\mathbb{O}^3)$ be a bandlimited sampling kernel with a bandlimit $L = 2K + 1$. Then the equiangular samples $s_n = (F * r)(\theta_n, \varphi_n)$ with $n = 1, \dots, (L-1)(2L-1)+1$ are sufficient to recover $F(\theta, \varphi)$ exactly.

While Corollary 4.2 considers equiangular samples, uniformly distributed samples could also be used as discussed in the previous sections. To verify our results, we generate signals using (4.5) and an ideal bandlimited sampling kernel. For equiangular samples, the results are reported in Table 4.2. As expected, both the angular error and mean squared error approach machine precision. For uniform samples, the performance of our reconstruction algorithm is presented in Figure 4.8. As usual, a sharp drop is observed when the number of samples reaches $N \geq (2K + 1)^2$.

4.4 Sampling Great Circles

In this section, we investigate the reconstruction of great circles. Great circles are formed by the intersection of a plane through the origin and a sphere centered at the origin. They are therefore completely defined by the orientation of the plane and the radius of the sphere. Of course, a plane can be described using its normal so it seems appropriate to describe great circles with orientations. The relation between the two notations, the great circles and the orientations, is the Funk transform (Funk, 1913, 1915) also called Funk-Radon transform because of its relation to Radon's work (Radon,

(a)

(b)

Figure 4.8: The mean angular error and mean squared error as a function of the number of samples when reconstructing integrated orientation. We used a uniform sample locations and ideal bandlimited sampling kernel.

Figure 4.9: Example of a sum of 2 great circles.

1917) or Funk-Minkowski transform given by

$$\begin{aligned} F(\mathbf{u}) &= \mathcal{R}f(\mathbf{u}) \\ &= \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) f(\mathbf{v}) d\mathbf{v}. \end{aligned} \quad (4.6)$$

Indeed, if we let $f(\mathbf{u})$ be K orientations then we have

$$\begin{aligned} F(\mathbf{u}) &= \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) \sum_{k=1}^K a_k \delta(\mathbf{v}^t \mathbf{v}_k - 1) d\mathbf{v} \\ &= \sum_{k=1}^K a_k \delta(\mathbf{u}^t \mathbf{v}_k) \end{aligned} \quad (4.7)$$

where we recall that $\mathbf{v}_k = [\cos \varphi_k \sin \theta_k \quad \sin \varphi_k \sin \theta_k \quad \cos \theta_k]^t$. The function $F(\mathbf{u})$ takes the value a_k when $\mathbf{u}$ is perpendicular to $\mathbf{v}_k$ and zero everywhere else. Note that for every great circle, we can identify two orientations: $\mathbf{v}_k$ and $-\mathbf{v}_k$. Because we are interested in the great circle itself and not the underlying orientation, we limit the azimuth to the interval $[0 .. \pi[$ (corresponding to all $\mathbf{v}_k$'s with a positive x component) which removes this ambiguity. A signal of the form of (4.7) is illustrated in Figure 4.9.

To obtain the relation between the spherical harmonics of $f(\mathbf{u})$ and those of $F(\mathbf{u})$, we use the Funk-Hecke Theorem (Andrews et al., 1999) which is presented here using our notation.

Theorem 4.1 (Funk-Hecke). *Let $f(t)$ be continuous on $[-1, 1]$ and let $Y_\ell^m(\mathbf{v})$ be any spherical harmonic of degree ℓ and order m . Then for a unit vector $\mathbf{u}$*

$$\int_{\mathbb{S}^2} f(\mathbf{u}^t \mathbf{v}) Y_\ell^m(\mathbf{v}) d\mathbf{v} = \lambda(\ell) Y_\ell^m(\mathbf{u})$$

where

$$\lambda(\ell) = 2\pi \int_{-1}^1 P_\ell(t) f(t) dt$$

and $P_\ell(t)$ is the Legendre polynomial of degree ℓ .

Using the Funk-Hecke Theorem, Descoteaux et al. (2007) proved a corollary whose proof we present here using the specific case of spherical harmonics and our notation.

Corollary 4.3 (Corollary of the Funk-Hecke Theorem). *Let $\delta(t)$ be the Dirac delta function and let $Y_\ell^m(\mathbf{v})$ be any spherical harmonic of degree ℓ and order m . Then given a unit vector $\mathbf{u}$*

$$\int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) Y_\ell^m(\mathbf{v}) \mathbf{v} = 2\pi P_\ell(0) Y_\ell^m(\mathbf{u}) \quad (4.8)$$

where $P_\ell(0)$ is the Legendre polynomial of degree ℓ evaluated at 0, given by

$$P_\ell(0) = \begin{cases} 0 & \ell \text{ odd} \\ (-1)^{\ell/2} \frac{1 \cdot 3 \cdot 5 \cdots (\ell-1)}{2 \cdot 4 \cdot 6 \cdots (\ell)} & \ell \text{ even.} \end{cases}$$

Proof. Using the Funk-Hecke Theorem on the left side of Equation (4.8) we have

$$\begin{aligned} \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) Y_\ell^m(\mathbf{v}) \mathbf{v} &= 2\pi \left(\int_{-1}^1 P_\ell(t) \delta(t) dt \right) Y_\ell^m(\mathbf{u}) \\ &\stackrel{(a)}{=} 2\pi P_\ell(0) Y_\ell^m(\mathbf{u}) \end{aligned}$$

where (a) is obtained by the sifting property of the Dirac. $\square$

If we expand $f(\mathbf{u})$ in (4.6) using spherical harmonics and then use Corollary 4.3, we find that

$$\begin{aligned} F(\mathbf{u}) &= \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) f(\mathbf{v}) d\mathbf{v} \\ &= \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m Y_\ell^m(\mathbf{u}) d\mathbf{v} \\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m \int_{\mathbb{S}^2} \delta(\mathbf{u}^t \mathbf{v}) Y_\ell^m(\mathbf{u}) d\mathbf{v} \\ &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m 2\pi P_\ell(0) Y_\ell^m(\mathbf{u}) d\mathbf{v}. \end{aligned} \quad (4.9)$$

Comparing Equation (4.9) with the spherical harmonic expansion of $F(\mathbf{u})$, we find that $\hat{F}_\ell^m = 2\pi P_\ell(0) \hat{f}_\ell^m$. The Funk transform acts similarly to an antipodally symmetric sampling kernel ($P_\ell(0) = 0$ for odd ℓ). A notable difference is that the relation between $\hat{f}_\ell^m$ and $\hat{F}_\ell^m$ holds for any ℓ , that is $F(\mathbf{u})$ is not bandlimited.

Given the coefficients $\hat{F}_\ell^\ell$ for $\ell = 0, 2, \dots, 4K - 2$ and $\hat{F}_\ell^{\ell-1}$ for $\ell = 2, \dots, 2K$ we can compute the corresponding $\hat{f}_\ell^\ell$ and $\hat{f}_\ell^{\ell-1}$. As described in Section 4.2, these are sufficient to recover all the parameters of $f(\mathbf{u})$ and therefore those of the great circles in $F(\mathbf{u})$. The final step in sampling and reconstructing great circles is to compute

the spherical harmonics coefficients from the samples. We again assume a sampling scheme

$$s_n = s(\theta_n, \varphi_n) = (F * r)(\theta_n, \varphi_n)$$

where (θ_n, φ_n) for $n = 1, \dots, N$ are equiangular samples. Assuming a sampling kernel $r(\mathbf{G})$ with a bandlimit L as described in Section 3.3, we find that

$$s_n = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{s}_{\ell}^m Y_{\ell}^m(\theta_n, \varphi_n) = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{F}_{\ell}^m \hat{r}_{\ell}^m Y_{\ell}^m(\theta_n, \varphi_n). \quad (4.10)$$

Again using the method of [McEwen and Wiaux \(2011\)](#) and assuming a bandlimit $L = 4K - 1$, we can recover all the coefficients $\hat{s}_{\ell}^m$. We then compute $\hat{F}_{\ell}^m = \hat{s}_{\ell}^m / \hat{r}_{\ell}^m$ and finally obtain $\hat{f}_{\ell}^m = \hat{F}_{\ell}^m / 2\pi P_{\ell}(0)$ for even ℓ . We therefore proved the following result.

Corollary 4.4 (Sampling Great Circles). *Let $F(\mathbf{u}) \in L^2(\mathbb{S}^2)$ be the sum of K great circles*

$$F(\mathbf{u}) = \sum_{k=1}^K a_k \delta(\mathbf{u}^{\dagger} \mathbf{v}_k)$$

and let $r(\mathbf{G}) \in L^2(\mathbb{SO}(3))$ be a rotational sampling kernel with a bandlimit $L = 4K - 1$. The equiangular samples $s_n = (F * r)(\theta_n, \varphi_n)$ for $n = 1, \dots, (L - 1)(2L - 1) + 1$ are sufficient to recover $F(\mathbf{u})$ exactly.

A great circle reconstruction is illustrated in Figure 4.10. The system in (4.10) can be written in matrix form as

$$\mathbf{s} = \mathbf{Y} \hat{\mathbf{s}} = \mathbf{Y} \mathbf{R} \hat{\mathbf{F}} = \mathbf{Y} \mathbf{R} \mathbf{P} \hat{\mathbf{f}}$$

where $\mathbf{Y}$ is the spherical harmonic transform matrix, $\mathbf{R}$ is the matrix associated with the kernel, and $\mathbf{P}$ is a diagonal matrix whose elements are the coefficients $2\pi P_{\ell}(0)$. Because $P_{\ell}(0) = 0$ for odd ℓ , the vector $\hat{\mathbf{f}}$ needs only consider even ℓ coefficients. The system therefore only has $(L/2 + 1)(L - 1)$ unknowns. As before, we expect to be able to recover the parameters of the great circles when $N \geq 8K^2 - 4K$.

Once again, we validate our results using numerical simulations. The signals are generated using (4.10) and an ideal bandlimited kernel. For equiangular samples, the results are reported in Table 4.3. As expected, both the angular error and mean squared error approach machine precision. For uniform samples, the performance of our reconstruction algorithm is presented in Figure 4.11. As usual, a sharp drop is observed when the number of samples reaches $N \geq 8K^2 - 4K$.

4.5 Summary

In this section, we have presented the following results:

- When using an optimal kernel, the parameters of the orientations can be accurately recovered using the critical number of samples $3K$.

Figure 4.10: Illustration of the reconstruction process using equiangular sample locations and an antipodally symmetric kernel. (a) original orientations (b) sample locations (c) sampling kernel (d) filtered signal (d) samples (f) recovered orientations.

(a)

(b)

Figure 4.11: The mean angular error and mean squared error as a function of the number of samples for great circles. We used uniform sample locations and the ideal optimal kernel.

$K(N)$	1 (11)	2 (79)	5 (667)	10 (2927)
α	1.15×10^{-7}	1.31×10^{-7}	1.65×10^{-5}	0.4188
β	0.000	0.000	2.84×10^{-12}	1.75×10^{-5}

Table 4.3: Mean angular error and mean square error observed using equiangular samples and the ideal bandlimited sampling kernel when sampling great circles. For errors reported as 0, the observed values were below machine precision ($\approx 10^{-16}$).

- When using an antipodally symmetric sampling kernel with a bandlimit $L = 4K - 1$, the parameters of the orientation can be exactly recovered using $N = (L - 1)(2L - 1) + 1$ equiangular samples. Our numerical results show that the parameters can be accurately recovered even when N is reduced to $L(L + 1)/2$ when using uniformly distributed sample locations.
- A signal that can be differentiated to a sum of K orientation can be sampled and reconstructed using an ideal bandlimited sampling kernel with a bandlimit $L = 2K + 1$. The number of required equiangular samples is $N = (L - 1)(2L - 1) + 1$.
- Great circles can be sampled and reconstructed using an ideal bandlimited sampling kernel with a bandlimit $L = 4K - 1$ and $N = (L - 1)(2L - 1) + 1$ equiangular samples. Our numerical results show that accurate recovery still occurs in $N = (L/2 + 1)(L - 1)$ uniform samples are used instead.

V

Sampling noisy FRI signals on the sphere

The reconstruction methods for finite rate of innovation signals on the sphere proposed in Chapters 3 and 4 offer excellent results when dealing with noiseless measurements. However, in any practical situation, the measurements will not be exact. For the simple case of orientations filtered with an ideal bandlimited kernel, the noisy samples can be written as

$$\mathbf{s} = \mathbf{Y}\hat{\mathbf{f}} + \boldsymbol{\epsilon}.$$

Here, $\boldsymbol{\epsilon} = [\epsilon_1 \ \cdots \ \epsilon_N]^t$ models the difference between expected samples and the observed samples. This discrepancy has several sources, including observation noise caused by the acquisition system and model mismatch. The simplest way to estimate $\hat{\mathbf{f}}$ is to ignore the noise and compute a least-squares solution $\hat{\mathbf{f}} \approx \mathbf{Y}^{-1}\mathbf{s}$. There are, of course, many ways to improve this estimate. Selecting the best approach usually depends on the specifics of the problem at hand, for example, the type of noise. For this reason, we will deal with this problem in the context of our application, in Chapter 6. Here, we assume that given $\mathbf{s}$ we can, at best, recover an approximation of $\hat{\mathbf{f}}$. The question is now: Can we accurately recover the parameters of $f(\mathbf{u})$ given the estimated coefficients $\hat{\mathbf{f}}$? The ubiquitous nature of this problem is reflected in the literature, where [Maravić and Vetterli \(2005\)](#), [Dragotti et al. \(2007\)](#), [Blu et al. \(2008\)](#), [Dragotti and Homann \(2009\)](#), [Ben-Haim et al. \(2010\)](#), [Tur et al. \(2011\)](#), and [Chen et al. \(2012\)](#) have considered sampling finite rate of innovation signals in the presence of noise. Note that all the previous references deal with 1D and 2D Cartesian signals.

In this Chapter, we investigate how to improve the recovery of orientations from noisy measurements. In Section 5.1, we propose a method to improve the reconstruction when the bandlimit is at the critical rate $L = 2K$. In this situation, we can only compute the minimum number of coefficients z_ℓ and cannot use superfluous information to reduce the impact of noise. In Section 5.2, we present a new algorithm to reduce the

impact of noise when the bandlimit is greater than the critical rate $L \geq 2K$, that is when we have redundant information (Deslauriers-Gauthier and Marziliano, 2013).

5.1 Noisy spherical FRI without oversampling

In this Section, for the sake of simplicity, we consider only orientations convolved with a bandlimited sampling kernel. However, we point out that the ideas presented here can be adapted to the signal models and sampling kernels of Chapters 3 and 4 with only slight modifications.

5.1.1 Using all spherical harmonic coefficients

Theorem 3.1 states that we can recover the parameters of a sum of K orientations using a subset of its spherical harmonic coefficients. Specifically, we need those with equal degree and order and those whose order is one less than their degree. For a signal with K orientations, this means we need only $3K$ coefficients to recover the parameters. However, in practice we usually compute more unless we are dealing with an optimal kernel. For the case of a bandlimited kernel we compute $4K^2$ which is much larger than $3K$. An obvious way to reduce the impact of noise is to use all of the available coefficients, which is what we propose here.

Consider the product of the spherical harmonic coefficients $\hat{\mathbf{f}}$ of a sum of K orientations with those of a second function $g(\theta, \varphi) \in L^2(\mathbb{S}^2)$ bandlimited at L :

$$\hat{\mathbf{f}}^* \hat{\mathbf{g}} = \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{g}_{\ell}^m \bar{\hat{f}}_{\ell}^m.$$

We can replace the coefficients $\hat{f}_{\ell}^m$ by their spherical harmonics representation

$$\hat{f}_{\ell}^m = \sum_{k=1}^K a_k \bar{Y}_{\ell}^m(\theta_k, \varphi_k)$$

which we proved in (3.3) to get

$$\begin{aligned} \hat{\mathbf{f}}^* \hat{\mathbf{g}} &= \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{g}_{\ell}^m \sum_{k=1}^K a_k Y_{\ell}^m(\theta_k, \varphi_k) \\ &= \sum_{k=1}^K a_k \sum_{\ell=0}^{L-1} \sum_{m=-\ell}^{\ell} \hat{g}_{\ell}^m Y_{\ell}^m(\theta_k, \varphi_k) \\ &= \sum_{k=1}^K a_k g(\theta_k, \varphi_k). \end{aligned} \tag{5.1}$$

In other words, the product $\hat{\mathbf{f}}^* \hat{\mathbf{g}}$ is a linear combination of $g(\theta, \varphi)$ evaluated at the location of the orientations. We can use this observation to get results similar to those presented in Section 3.2. Indeed, if we define

$$g_{\ell}(\theta, \varphi) = \sin^{\ell}(\theta) \exp(j\ell\varphi) \tag{5.2}$$

we find that

$$\hat{\mathbf{f}}^* \hat{\mathbf{g}}_\ell = \sum_{k=1}^K a_k \sin^\ell(\theta_k) \exp(j\ell\varphi_k) = \sum_{k=1}^K a_k r_k^\ell = z_\ell$$

where z_ℓ are the normalized samples defined in Section 3.2. This result is not surprising because the spherical harmonic coefficients $\hat{\mathbf{g}}_\ell$ are all zero except when the degree equals the order¹. When we compute the value of $\hat{\mathbf{f}}^* \hat{\mathbf{g}}_\ell$, we are simply selecting and scaling the coefficient $\hat{f}_\ell^\ell$. Furthermore, when we compute the products for $\ell = 0, \dots, 2K - 1$, we are automatically selecting the coefficients with equal degree and order used in Theorem 3.1.

As we discussed previously, using a single coefficient per degree is not ideal when the measurements are noisy. Instead, we would prefer to use all coefficients to improve the robustness to noise. To do this, we use a rotated version of $g_\ell(\theta, \varphi)$. Let

$$\mathbf{G}g_\ell(\mathbf{u}) \equiv g_\ell(\mathbf{G}\mathbf{u})$$

for some rotation matrix $\mathbf{G} \in \mathbb{SO}(3)$ and let $\mathbf{G}\hat{\mathbf{g}}_\ell$ be the vector of spherical harmonics coefficients of $\mathbf{G}g_\ell(\mathbf{u})$. In comparison to $g_\ell(\mathbf{u})$ whose harmonic representation is a single coefficient, $\mathbf{G}g_\ell(\mathbf{u})$ may have its power distributed across all coefficients of degree ℓ , depending on the matrix $\mathbf{G}$. This is illustrated in Figure 5.1 for two different rotations about the y axis.

The product $\hat{\mathbf{f}}^* \mathbf{G}\hat{\mathbf{g}}_\ell$ evaluates a rotated version of Equation (5.2) at the location of the orientation. If we write

$$g_\ell(\mathbf{u}_k) = \sin^\ell(\theta_k) \exp(j\ell\varphi_k)$$

then we have

$$\hat{\mathbf{f}}^* \mathbf{G}\hat{\mathbf{g}}_\ell = \sum_{k=1}^K g_\ell(\mathbf{G}\mathbf{u}_k) = \mathbf{G}z_\ell$$

where $\mathbf{G}z_\ell$ is a “rotated” version of z_ℓ . Using the values of $\mathbf{G}z_\ell$ for $\ell = 0, \dots, 2K + 1$ as an input to the annihilating filter method, we recover $\sin(\tilde{\theta}_k)$, a_k , and $\tilde{\varphi}_k$ where we have added the tilde to indicate that the parameters are in a rotated coordinate system. As in the noiseless case, we still have an ambiguity on the value of $\tilde{\theta}_k$. To recover the value of the colatitudes, we can additionally evaluate $\hat{\mathbf{f}}^* \mathbf{G}\hat{\mathbf{x}}_\ell$ for $\ell = 1, \dots, K$ where $\mathbf{G}\hat{\mathbf{x}}_\ell$ is the vector of spherical harmonic coefficients of a rotated version of x_ℓ and

$$x_\ell(\theta, \varphi) = \cos \theta \sin^\ell(\theta) \exp(j\ell\varphi).$$

Using the definition

$$x_\ell(\mathbf{u}_k) = \cos \theta_k \sin^\ell(\theta_k) \exp(j\ell\varphi_k)$$

¹This is trivial to prove since we already showed that $Y_\ell^\ell(\theta, \varphi) = C_\ell^\ell B_\ell \sin^\ell(\theta) \exp(-j\ell\varphi)$ for some constants C_ℓ^ℓ and B_ℓ .

Figure 5.1: Magnitude of the spherical harmonic coefficients of $\mathbf{G}g_2(\mathbf{u})$ with $\mathbf{G} = \mathbf{G}_y(0)$ in (a), $\mathbf{G} = \mathbf{G}_y(\pi/4)$ in (b), and $\mathbf{G} = \mathbf{G}_y(\pi/2)$ in (c).

and (5.1) we get

$$\hat{\mathbf{f}}^* \mathbf{G} \hat{\mathbf{x}}_\ell = \sum_{k=1}^K x_\ell(\mathbf{G} \mathbf{u}_k) = \mathbf{G} \mathbf{w}_\ell$$

where $\mathbf{G} \mathbf{w}_\ell$ is a rotated version of $\mathbf{w}_\ell$. Given the values of $\mathbf{G} \mathbf{w}_\ell$, we can follow the same procedure as in Section 3.2 and recover the value of $\tilde{\theta}_k$. Once all the rotated parameters are known, we can revert back to the original coordinate system using

$$\mathbf{v}_k = \mathbf{G}^{-1} \tilde{\mathbf{v}}_k$$

where $\mathbf{v}_k = \begin{bmatrix} \sin \theta_k \cos \varphi_k & \sin \theta_k \sin \varphi_k & \cos \theta_k \end{bmatrix}$.

While the procedure we just described improves the robustness to noise, the values of θ_k near zero are still troublesome to recover when we have imperfect measurements.

Recall the linear system in (3.10)

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1^1 & r_2^1 & \cdots & r_K^1 \\ \vdots & \ddots & \ddots & \vdots \\ r_1^{K-1} & r_2^{K-1} & \cdots & r_K^{K-1} \end{bmatrix} \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \vdots & a_K \end{bmatrix} \begin{bmatrix} \cos \theta_1 \\ \cos \theta_2 \\ \vdots \\ \cos \theta_K \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{K-1} \end{bmatrix} \quad (5.3)$$

which can be written compactly as $\mathbf{R}\mathbf{A}\mathbf{c} = \mathbf{w}$. When $\theta_k \approx 0$, then $|r_k| \approx 0$ and the corresponding column of $\mathbf{R}$ has little contribution to $\mathbf{w}$, with the exception of w_0 . This makes the estimation of $\cos \theta_k$ from the preceeding linear system unreliable. Only a single element of $\mathbf{w}$ can be used to determine its value.

Note that when we try to compute the values of θ_k from (5.3), the values of φ_k have already been computed using the annihilating filter method. Assuming a bandlimited kernel, we also have access to all the spherical harmonic coefficients $\hat{f}_\ell^m$ up to the bandlimit L . Recall that these coefficients are given by

$$\hat{f}_\ell^m = \sum_{k=1}^K a_k \bar{Y}_\ell^m(\theta_k, \varphi_k)$$

and in matrix form we have

$$\hat{\mathbf{f}} = \begin{bmatrix} \mathbf{Y}_{\theta_1 \varphi_1} & \cdots & \mathbf{Y}_{\theta_K \varphi_K} \end{bmatrix}^* \mathbf{a}$$

where

$$\mathbf{Y}_{\theta_k \varphi_k} = \begin{bmatrix} Y_0^0(\theta_k, \varphi_k) & Y_1^{-1}(\theta_k, \varphi_k) & Y_1^0(\theta_k, \varphi_k) & \cdots & Y_L^L(\theta_k, \varphi_k) \end{bmatrix}^t$$

and

$$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_K \end{bmatrix}^t.$$

The exact value of θ_k is not known, but we can approximate it using discrete values $\theta_1 = 0 < \theta_2 < \cdots < \theta_N = \pi$. For a sufficiently large N , one of these values will be close to the true θ_k . To identify the correct values of θ_k we can attempt to find the vector $\mathbf{a}$ that minimizes

$$\left\| \hat{\mathbf{f}} - \begin{bmatrix} \mathbf{Y}_{\theta_1 \varphi_1} & \cdots & \mathbf{Y}_{\theta_N \varphi_1} & \cdots & \mathbf{Y}_{\theta_1 \varphi_K} & \cdots & \mathbf{Y}_{\theta_N \varphi_K} \end{bmatrix}^* \mathbf{a} \right\|_2$$

where $\mathbf{a}$ is K sparse meaning it has K non-zero entries. To retrieve the correct columns, we use orthogonal matching pursuit (OMP) (Tropp and Gilbert, 2007; Needell and Vershynin, 2010).

5.1.2 Selecting the rotation matrix

The selection of $\mathbf{G}$ will have an important impact on the performance of the algorithm described in the previous section. Instinctively, the best matrix is the one that distributes the power of $\mathbf{G}g_\ell(\mathbf{u})$ evenly across all coefficients of degree ℓ . However, our empirical evidence shows that this is not the case. Instead, the ideal matrix $\mathbf{G}$ varied

for each simulation. This phenomenon can be explained by our previous observation that values of θ_k near zero are challenging to recover. Suppose we choose a matrix $\mathbf{G}$ that represents a rotation of $\pi/2$ about the y -axis. This matrix will improve the recovery of the orientations with $\theta_k \approx 0$ by increasing their contribution to the data. On the other hand, orientations with a $\theta_k \approx \pi/2$ will be harder to see because they are rotated away. To illustrate this idea, consider the angular error as a function of $\mathbf{G}$ presented in Figure 5.2. In (a), we see the angular error for a single simulation with $K = 2$. The graph has regions where the angular error goes above 15° and regions where the angular error decreases to almost 0° . The angular error for an average of 100 simulations is illustrated in (b). The averaging has removed any local minimum and the error is almost constant at 10° .

Because we have no way of a priori determining the best rotation matrix, we propose to compute the parameters for several matrices and then selecting the solution that best fits the measurements. That is, we choose the parameters according to

$$\arg \min_{\theta_k, \varphi_k, a_k} \|\hat{\mathbf{s}} - \mathbf{Y} \mathbf{R} \hat{\mathbf{f}}\|_2$$

where the spherical harmonic coefficients of the vector $\hat{\mathbf{f}}$ are given by

$$\hat{f}_\ell^m = \sum_{k=1}^K a_k \bar{Y}_\ell^m(\theta_k, \varphi_k).$$

Obviously, the complexity of the reconstruction algorithm increases linearly with the number of rotation matrices to test. However, this increase in computation complexity also yields significant performance gains, as we will illustrate in the following section.

5.1.3 Numerical simulations

To verify our results, we simulated signals using

$$s_n = s(\theta_n, \varphi_n) = \sum_{\ell=1}^{L-1} \sum_{m=-\ell}^{\ell} \hat{f}_\ell^m Y_\ell^m(\theta_n, \varphi_n) + \epsilon_n$$

where the sample locations (θ_n, φ_n) with $n = 1, \dots, N$ were generated using the algorithm in Appendix A. The random variable ϵ_n is taken from a zero mean Gaussian distribution whose standard deviation is selected to achieve the desired signal to noise ratio. We use the usual definition of the signal to noise ratio which is

$$\text{SNR} = 20 \log_{10}(\sigma_s / \sigma_\epsilon) \quad (5.4)$$

where σ_s and σ_ϵ are the standard deviation of the signal and noise, respectively. The matrix used in the OMP step was generated using 50 values of θ in the interval $[0, \pi]$ for each φ_k , meaning that its dimensions are $L^2 \times 50K$. We used a total of 16 rotation matrices which correspond to every combination of rotations of $0, \pi/4, \pi/2$, and $3\pi/4$ about the y and z axes.

Figure 5.2: The angular error as a function of $\mathbf{G}$, described in terms of its y and z rotation. The angular error for a single simulation is reported in (a) and the average over 100 simulations in (b).

The angular error as a function of the signal to noise ratio is illustrated in Figure 5.3 for a signal with $K = 2$ using $N = 16$ samples. The blue line represents the average angular error for the reconstruction using several rotation matrices described in Section 5.1.2 and the red line is the original version for comparison purposes.

When the number of orientations is increased, the angular error also increases. This is illustrated in Figure 5.4 (a) for a signal with $K = 3$ and sampled with $N = 36$ samples and a bandlimit of $L = 6$. In this case, the poor performance is caused by the difficulty of estimating the spherical harmonic coefficients with a high degree. A simple way to improve the situation is to increase the number of samples N without increasing the bandlimit. The performance for an increase to $N = 49$ is illustrated in Figure 5.4 (b). The gain in performance is clearly visible, particularly in the low SNR range 0 to 10dB. Note that this still agrees with our idea of not oversampling, because keeping the bandlimit at the minimum value means we have minimum number of normalized coefficients z_ℓ to work with.

As was previously discussed, the improvement in performance comes with an increased complexity. For the signals with $K = 3$ orientations, the average computation time for the noiseless algorithm was 10ms whereas the improved version took 150ms in our MATLAB implementation. As expected, the increase in complexity is proportional to the number of rotation matrices used, in this case 16.

Figure 5.3: The angular error as a function of the signal to noise ratio for a signal with $K = 2$ orientations computed over 500 simulations. An ideal bandlimited kernel and $N = 16$ samples were used in the reconstruction. The blue line is the average and the dots are individual simulation results.

5.2 Noisy spherical FRI with oversampling

In the past Sections, the parameters of a sum of K orientations a_k , θ_k , and φ_k are recovered in two distinct steps. First, we used the spherical harmonics coefficients with equal degree and order $\hat{f}_\ell^\ell$ to compute the normalized coefficients z_ℓ with $\ell = 0, \dots, 2K - 1$. We then used these coefficients to build a data matrix $\mathbf{Z}$ and computed the vector in its null space $\mathbf{h}$. The elements of this vector were the coefficients of a polynomial whose roots are $\sin \theta_k \exp(-j\varphi_k)$. This gives us the value of φ_k . Second, we solved a linear set of equations with a sparse solution to compute the value of a_k and θ_k .

Computing $\mathbf{h}$ from $\mathbf{Z}$ as defined in (3.8) assumes that the rank of $\mathbf{Z}$ is K and that there will be exactly one vector in its null space (excluding the zero vector). When noise or model mismatch is added, this condition will not be satisfied and the roots of $\mathbf{h}$ will no longer correctly identify $\sin \theta_k$ or φ_k . This is a common problem faced when using the annihilating filter method and the commonly used solution was proposed by [Blu et al. \(2008\)](#) in the form of Cadzow's iterative denoising. We present this algorithm in Section 5.2.1 and an improved version in Sections 5.2.2 to 5.2.4.

5.2.1 Cadzow's iterative denoising

Cadzow's iterative denoising ([Cadzow, 1988](#); [Tufts and Kumaresan, 1982](#)) is a pre-processing step that forces the measured data to fit the model of FRI: a Toeplitz matrix of rank K . When applied to spherical FRI, it proceeds as follows:

1. Using the normalized coefficients z_ℓ with $\ell = 0, \dots, N$, build a wider Toeplitz

(a)

(b)

Figure 5.4: The angular error as a function of the signal to noise ratio for a signal with $K = 3$ orientations computed over 500 simulations. We used an ideal bandlimited kernel, $N = 36$ samples in (a), and $N = 49$ samples in (b). The blue line is the average and the dots are individual simulation results and the red line is the average performance of the noiseless algorithm.

data matrix

$$\mathbf{Z} = \begin{bmatrix} z_{[N/2]} & \cdots & z_0 \\ \vdots & \ddots & \vdots \\ z_N & \cdots & z_{[N/2]} \end{bmatrix}.$$

and compute its SVD.

2. Recover a total least-squares approximation of $\mathbf{Z}$, denoted by $\mathbf{Z}'$, by keeping only the K largest eigenvalues and their associated singular vectors.
3. Build a denoised version by averaging the diagonals of $\mathbf{Z}'$.
4. Repeat the Steps 2 and 3 until the $(K + 1)^{\text{th}}$ eigenvalue is below some threshold.

In summary, we attempt to find a Toeplitz data matrix that is as close as possible to $\mathbf{Z}$ but that has a rank of K . The flaw is that the least-squares estimate $\mathbf{Z}'$ obtained in Step 2 has a rank of K but is not Toeplitz. When the approximation is converted to a Toeplitz in Step 3, its rank is no longer K . By repeating the procedure, we do end up with a Toeplitz matrix of rank K , but it is most likely not the one that is closest to $\mathbf{Z}$. Nonetheless, Cadzow's iterative denoising method significantly reduces the impact of additive noise, but requires a minimum of redundancy to be effective. In the worst case scenario, when $N = 2K$, there are no superfluous eigenvalues that are set to zero and the denoising procedure has no effect.

5.2.2 The annihilating filter method as an optimization problem

Consider again the noiseless data matrix $\mathbf{Z}$. Computing the annihilating filter corresponds to finding $\mathbf{h}$ such that

$$\mathbf{Z}\mathbf{h} = 0 \tag{5.5}$$

which implies

$$\mathbf{Z} - \mathcal{N}(\mathbf{h}^*)(\mathcal{N}(\mathbf{h}^*)^\dagger \mathbf{Z}) = 0 \tag{5.6}$$

where $\mathcal{N}(\mathbf{h}^*)$ is a matrix whose columns form a basis for the null space of $\mathbf{h}^*$. Equation (5.6) clearly shows that finding the annihilating filter corresponds to maximizing the projection of $\mathbf{Z}$ into the space spanned by $\mathcal{N}(\mathbf{h}^*)$. In fact, the null space of $\mathbf{h}^*$ is the only subspace where $\mathbf{Z}$ projects onto itself which also makes $\mathbf{h}$ unique. When noise is added and $N > 2K + 1$, we make the data matrix taller

$$\mathbf{Z} = \begin{bmatrix} z_K & \cdots & z_0 \\ \vdots & \ddots & \vdots \\ z_N & \cdots & z_{N-K} \end{bmatrix}$$

and attempt to solve (5.5). In this case, no vector $\mathbf{h}$ will satisfy $\mathbf{Z}\mathbf{h} = 0$ and the total least-squares, usually implemented using an SVD, finds the vector $\mathbf{h}$ that minimizes

$$\|\mathbf{Z} - \mathcal{N}(\mathbf{h}^*)(\mathcal{N}(\mathbf{h}^*)^\dagger \mathbf{Z})\|_2 \tag{5.7}$$

where $\|\cdot\|_2$ is the Frobenius norm. The problem here is that the projection

$$\mathcal{N}(\mathbf{h}^*)(\mathcal{N}(\mathbf{h}^*)^+ \mathbf{Z})$$

is not necessarily a Toeplitz matrix and as such does not correspond to the model predicted by FRI. Essentially, we run into the same problem we observed during Cadzow's denoising.

An ideal approach would be to find the Toeplitz data matrix $\mathbf{Z}'$ that is the solution to the problem

$$\begin{aligned} & \underset{\mathbf{Z}'}{\text{minimize}} && \|\mathbf{Z} - \mathbf{Z}'\|_2 \\ & \text{subject to} && \mathbf{Z}'\mathbf{h} = 0 \end{aligned} \quad (5.8)$$

This optimization problem could be solved for all $\mathbf{h}$ and we could select the annihilating filter that yields the smallest optimal value. Needless to say, this is not feasible in practice. However, if the problem of (5.8) has a closed form solution we could attempt to optimize over $\mathbf{h}$.

Because Toeplitz matrices form a subspace under matrix addition and scalar multiplication, $\mathbf{Z}'$ can be represented as

$$\mathbf{Z}' = \sum_{k=1}^M c_k \mathbf{T}_k$$

with $M = (K + 1) + (N - K) - 1 = N$ and where $\mathbf{T}_k$ are orthonormal Toeplitz matrices. Let $\text{vec}(\cdot)$ be the vectorization operator that transforms a matrix into a vector by stacking its columns and let $\mathbf{w} = \text{vec}(\mathbf{X})$, $\mathbf{c} = \begin{bmatrix} c_1 & \cdots & c_N \end{bmatrix}$ and $\mathbf{T} = \begin{bmatrix} \text{vec}(\mathbf{T}_1) & \cdots & \text{vec}(\mathbf{T}_N) \end{bmatrix}$. We can rewrite the problem of (5.8) using $\mathbf{c}$ as an alternate variable and $\mathbf{Z}' = \mathbf{T}\mathbf{c}$,

$$\begin{aligned} & \underset{\mathbf{c}}{\text{minimize}} && \|\mathbf{w} - \mathbf{T}\mathbf{c}\|_2 \\ & \text{subject to} && \mathbf{U}\mathbf{c} = 0 \end{aligned} \quad (5.9)$$

where

$$\mathbf{U} = \begin{bmatrix} \mathbf{T}_1\mathbf{h} & \cdots & \mathbf{T}_N\mathbf{h} \end{bmatrix}^t.$$

For the constraint $\mathbf{U}\mathbf{c} = 0$ to hold, $\mathbf{c}$ must be in the null space of $\mathbf{U}$ and can therefore be expressed as $\mathbf{c} = \mathcal{N}(\mathbf{U})\mathbf{b}$ where $\mathbf{b}$ is an $(N - K + 1) \times 1$ column vector. This leads to

$$\|\mathbf{w} - \mathbf{T}\mathbf{c}\|_2 = \|\mathbf{w} - \mathbf{T}\mathcal{N}(\mathbf{U})\mathbf{b}\|_2$$

which has the least-squares solution

$$\mathbf{b} = (\mathbf{T}\mathcal{N}(\mathbf{U}))^+ \mathbf{w}.$$

We conclude that the optimum value of the problem of (5.8) is

$$\zeta(\mathbf{h}) = \|\mathbf{w} - \mathbf{A}\mathbf{A}^+ \mathbf{w}\|_2$$

with $\mathbf{A} = \mathbf{T}\mathcal{N}(\mathbf{U})$. Equivalently, the optimal solution can be written as

$$\zeta(\mathbf{h}) = \|\mathbf{A}'\mathbf{A}'^+\mathbf{w}\|_2$$

with $\mathbf{A}' = \mathbf{T}\mathbf{U}^*$, which slightly reduces the number of computations. In words, the function $\zeta(\mathbf{h})$ measures how well the subspace $\mathcal{N}(\mathbf{h}^*)$ fits the data while ensuring that $\mathbf{Z}'$ is a Toeplitz matrix. Now that we have a closed form solution as a function of the annihilating filter, the optimum filter can be identified by minimizing $\zeta(\mathbf{h})$ using standard optimization methods. Once the annihilating filter is known, we can proceed as described in the previous section to recover a_k , θ_k , and φ_k .

The projection $\mathbf{A}\mathbf{A}^+\mathbf{w}$ represents the data that is explained by a specific annihilating filter $\mathbf{h}$. Each filter is associated with exactly one set of roots $r_k = \sin \theta_k \exp(-j\varphi_k)$ so when minimizing $\zeta(\mathbf{h})$ we look for the values of r_k that best agree with our data. The underlying assumption is that if we could enumerate all sets of r_k , the set that generates the data closest to our measurements are the true or very near the true locations. In practice there are two situations where this assumption will not hold:

1. When two or more orientations have nearly identical r_k 's, i.e. they have near identical θ_k and φ_k .
2. When an orientation has a very low magnitude $|a_k|$.

When these two situations are encountered, minimizing $\zeta(\mathbf{h})$ will lead to parameters that explain the noise rather than the signal.

5.2.3 Initialization matters

To get a better understanding of how $\zeta(h)$ behaves, its value is illustrated in Figure 5.5. The parameters of the signal are $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, $a_k = 1$. Without loss of generality, we set $h_0 = 1$ and display the objective function with respect to the real and imaginary parts of h_1 . For the parameters enumerated above, the expected value of h_1 is $0 + \sqrt{2}j/2$. At a SNR of 100 and 20 dB, the objective function has a very clear minimum at the expected position. When the signal to noise ratio decreases, a minimum is still observed, but it no longer corresponds to the correct value of h_1 . At a very low SNR, for example 0 dB, the shape of $\zeta(h)$ varies greatly and it starts to exhibit local minima. This behavior is expected since the number of samples used is extremely low (we used $N = 9$). Figure 5.6 illustrates the objective function for the same signal, but sampled at $N = 16$. Oversampling tends to sharpen the location of the minimum and reduces “movement” of the minimum. However, lowering the SNR works in the opposite direction and widens the minimum. An important observation is that local minima are still present at SNR = 0 dB. To evaluate the impact of oversampling, Figure 5.7 illustrates $\zeta(h)$ for the same signal, but at a SNR = 5 dB and various number of samples. The sharpening effect around the optimal point as the number of sample increases is visible. We can also notice that the local minima are suppressed as the number of samples increases, but are not completely removed. Although Figures 5.5, 5.6, and 5.7 only illustrate the simplest signal with $K = 1$, they still provide

Figure 5.5: The value of $\zeta(h)$ as a function of the SNR for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, and $a_k = 1$ sampled using $N = 9$. The signal to noise ratio for the subplots are: (a) SNR = 100 dB (b) SNR = 20 dB (c) SNR = 10 dB (d) SNR = 0 dB

meaningful information about $\zeta(h)$. We can see that $\zeta(h)$ is well behaved and does not have a large number of local minima. Unfortunately, the possibility that $\zeta(h)$ has local minima means the initial value of h used to solve the optimization problem will have an impact on the performance.

5.2.4 Adding bounds

To improve the robustness of the optimization algorithm, we can derive bounds for each of the coefficients of the annihilating filter. We assume without loss of generality that $h_0 = 1$ and that the real and imaginary part of the remaining coefficients will be treated as two distinct variables. Consider the second degree polynomial $h(x)$ whose coefficients are

$$h(x) = (x - r_1)(x - r_2) = x^2 - (r_1 + r_2)x + r_1r_2$$

which can be used to recover the parameters of a signal with 2 orientations. Recalling that $r_k = \sin \theta_k \exp(-j\varphi_k)$ we observe that $\Re(h_1) = \Re(-(r_1 + r_2))$ is upper

Figure 5.6: The value of $\zeta(h)$ as a function of the SNR for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, $a_k = 1$ and $a_k = 1$ sampled using $N = 16$. The signal to noise ratio for the subplots are: (a) SNR = 100 dB (b) SNR = 20 dB (c) SNR = 10 dB (d) SNR = 0 dB

and lower bounded by 2 and -2 obtained when $(\theta_1, \varphi_1) = (\theta_2, \varphi_2) = (\pi/2, 0)$ and $(\theta_1, \varphi_1) = (\theta_2, \varphi_2) = (\pi/2, \pi)$, respectively. The imaginary part of the coefficient $\Im(h_1)$ has the same bounds, but obtained at $(\theta_1, \varphi_1) = (\theta_2, \varphi_2) = (\pi/2, 3\pi/2)$ and $(\theta_1, \varphi_1) = (\theta_2, \varphi_2) = (\pi/2, \pi/2)$. The real and imaginary part of the last coefficient h_2 are both bounded by -1 and 1 because it is composed only of a product of complex exponentials.

A more general method to obtain the upper and lower bounds for a coefficient can be devised by observing that the bounds depend only on the number of summed terms in the expansion of a coefficient. For a polynomial of degree K , the number of terms in the expansion of the coefficient h_k is given by the binomial coefficient which leads to the bounds

$$-b_k \leq \Re(h_k), \Im(h_k) \leq b_k \text{ where } b_k = \binom{K}{k}.$$

Figure 5.7: The value of $\zeta(h)$ as a function of the number of samples N for a signal with $K = 1$, $\theta_k = \pi/4$, $\varphi_k = 3\pi/2$, and $a_k = 1$ with signal to noise ratio of 15dB. The number of samples used are: (a) $N = 9$ dB (b) $N = 16$ dB (c) $N = 25$ dB (d) $N = 36$ dB

The constrained optimization problem to solve is

$$\begin{aligned} & \underset{\mathbf{h}}{\text{minimize}} && \zeta(\mathbf{h}) \\ & \text{subject to} && h_0 = 1 \\ & && -b_k \leq \Re(h_k), \Im(h_k) \leq b_k, \quad k = 1, \dots, K \end{aligned} \tag{5.10}$$

which is solved in the following section using the interior-point method.

5.2.5 Numerical simulations

As in Section 5.1.3, we verify our results by simulating signals using

$$s_n = s(\theta_n, \varphi_n) = \sum_{\ell=1}^{L-1} \sum_{m=-\ell}^{\ell} \hat{f}_{\ell}^m Y_{\ell}^m(\theta_n, \varphi_n) + \epsilon_n.$$

The random variable ϵ_n is taken from a zero mean Gaussian distribution whose standard deviation is selected to achieve the desired signal to noise ratio defined in Equa-

tion (5.4). In this instance, we set the bandlimit to $L = 3K$ instead of the critical value of $2K$.

The mean angular error defined in (3.15) as a function of the signal to noise ratio is illustrated in Figure 5.8 for three algorithms: spherical FRI, optimized spherical FRI, and optimized spherical FRI with multi-rotation. The signals were sampled with an ideal kernel bandlimited at $L = 6$ and using $N = 49$ samples. By comparing Figure 5.8 (a) and (b), we see that the optimization offers a significant reduction in angular error. For a signal to noise ratio of 15dB, the mean angular error is reduced by a factor of 3, going from 12 to 4°. At a signal to noise ratio near 5dB, the mean angular error is reduced from 20 to 8°, a reduction of more than 50%. Combining the optimization procedure with the multi-rotation approach of Section 5.1 yields further performance gains, as illustrated in Figure 5.8 (c). The gain is more significant on the low SNR range: at 0dB the mean angular error is reduced from 14 to 8°.

5.2.6 Other solutions methods

The method we propose to solve the problem in (5.8) is merely one of many. Specifically, the difficulty lies in the subproblem of minimizing $\zeta(\mathbf{h})$ which is nonconvex. This means that any minima that is identified may be local and not be the global minima. This is a common, and still current, problem in optimization. One possible solution is the use of heuristics, such as particle swarm optimization (Kennedy and Eberhart, 1995; Zhan et al., 2011) and simulated annealing (Kirkpatrick et al., 1983). Another solution is to use a local solver, such as interior-point methods, conjugate-gradient methods, or Newton's method for several starting point seeded randomly or using a heuristic. However, global solvers are generally slower than local solvers which is why we have opted for the latter.

5.3 Summary

In this section, we have presented the following:

- We proposed a new algorithm that improves the reconstruction of orientations when the bandlimit of the sampling kernel is at the critical value $L = 2K$. The gains are obtained by using all spherical harmonics coefficients instead of only the ones with equal degree and order.
- We proposed a new algorithm that improves the reconstruction of a sum of K orientations when the bandlimit of the sampling kernel is above the critical value, that is $L \geq 2K$. Using an optimization procedure, we found the parameters that agree with our model and fit the observed data. This approach also has the advantage of allowing further constraints, such as optimization bounds.

Figure 5.8: The angular error as a function of the signal to noise ratio for a signal with $K = 2$ orientations. The signals were sampled with an ideal kernel bandlimited at $L = 6$ and using $N = 49$ samples. The reconstruction algorithms are spherical FRI in (a), optimized spherical FRI in (b), and optimized spherical FRI multi-rotation in (c). The gray are individual simulation results and the line is the mean angular error.

VI

Application to diffusion magnetic resonance imaging

The theory and numerical simulations presented in Chapters 3 through 5 detailed how certain non-bandlimited signals defined on the sphere can be sampled and reconstructed. In this Chapter, we apply those findings to the recovery of fibers in diffusion magnetic resonance imaging.

In Section 6.1, we present the most common model for dMRI signals, the multi-tensor model. In Section 6.2, we review the concepts of spherical deconvolution, a popular method to recover fiber orientations. Our reconstruction algorithm, which we call sFRIdMRI, is presented in Section 6.3. Section 6.4 compares the performance of both algorithms through numerical simulations. Finally, Section 6.5 presents results obtained on real diffusion MRI data.

6.1 Modelling the diffusion process

The first technique, named diffusion tensor imaging, used to recover fiber tracts in dMRI was proposed by [Basser et al. \(1994\)](#). In diffusion tensor imaging, we assume the diffusion signal can be modelled as

$$s(\mathbf{q}) = s_0 \exp(-b\mathbf{q}^t \mathbf{D} \mathbf{q})$$

where $\mathbf{q} \in \mathbb{S}^2$ is the gradient direction, b is a parameter of the acquisition referred to as the b -value, and s_0 is a normalization parameter. The matrix $\mathbf{D}$ is a second order tensor that describes the diffusion signal. Specifically, the eigenvector of $\mathbf{D}$ that corresponds to the largest eigenvalue is assumed to match the fiber orientation.

The diffusion tensor model is sufficient to model the signal of several areas of the brain, however it is inadequate when multiple fibers intersect in a single voxel. To

model additional fibers, [Tuch et al. \(2002\)](#) proposed the multi-tensor model given by

$$s(\mathbf{q}) = s_0 \sum_{k=1}^K a_k \exp(-b\mathbf{q}^t \mathbf{D}_k \mathbf{q}) \quad (6.1)$$

where the K fiber populations have their own tensor $\mathbf{D}_k$. This model is used by several groups for simulation and validation purposes ([Alexander et al., 2002](#); [Tuch, 2004](#); [Descoteaux et al., 2006](#); [Michailovich et al., 2010](#)). If we assume that different fibers have the same diffusion characteristics, the multi-tensor model can be reduced to

$$s(\mathbf{q}) = (f * r)(\mathbf{q}) \quad (6.2)$$

where $f(\mathbf{q}) \in L^2(\mathbb{S}^2)$ is the fiber orientation distribution function (ODF) and $r(\mathbf{G}) \in L^2(\mathbb{SO}(3))$ is the response function of a single fiber ([Tournier et al., 2004](#)).

6.2 Spherical deconvolution

Spherical deconvolution ([Tournier et al., 2004](#); [Alexander, 2005](#); [Tournier et al., 2007, 2008](#); [Yeh et al., 2011](#)) is a parametric HARDI method to recover fiber orientations. It is based on the model of (6.2) where $r(\mathbf{G})$ is assumed to be known, either from a model or estimated from the data. To recover $f(\theta, \varphi)$, [Tournier et al. \(2004\)](#) rewrites (6.2) using spherical harmonics and the spherical deconvolution results of [Healy et al. \(1998\)](#). In this formalism, the spherical harmonics of $s(\theta, \varphi)$ are computed from the samples s_n using a linear least squares method ([Alexander et al., 2002](#)) and $\hat{\mathbf{f}}$ is then computed from

$$\hat{\mathbf{f}} = \mathbf{R}^{-1} \hat{\mathbf{s}} \quad (6.3)$$

where $\mathbf{R}$ is computed from $r(\mathbf{G})$ and assumed to be known. From $\hat{\mathbf{f}}$, we can evaluate $f(\theta, \varphi)$ at any points using a forward transform. The orientation of fibers is then estimated by finding local maximums on $f(\theta, \varphi)$ using a Newton-Raphson gradient descent method ([Press et al., 2007](#)).

The estimation of the fiber ODF $f(\theta, \varphi)$ from the measurements is greatly simplified by the fact that $r(\mathbf{G})$ can be assumed to be antipodally symmetric, real, and axisymmetric. These properties means the rotational harmonic coefficients $\hat{r}_\ell^{mn}$ are reduced to a single real value for each degree. In other words, we have $\hat{f}_\ell^m = \hat{s}_\ell^m / \hat{r}_\ell$ with $\hat{r}_\ell \in \mathbb{R}$ which makes $\mathbf{R}$ diagonal. Estimating the fiber ODF directly from (6.3) has two important disadvantages. First, if the fiber response attenuates high frequency components (high degrees), then r_ℓ will become small as ℓ increases. This leads to an amplification of the noise when we divide by $\hat{r}_\ell$ making high degree coefficients more sensitive to noise ([Tournier et al., 2004](#)). The second limitation is that the angular resolution of the technique is directly related to the number of degrees of $\hat{f}_\ell^m$ that can be estimated. To get a high angular resolution, $\hat{f}_\ell^m$ must be estimated for high values of ℓ , leading to an explosion in the number of samples required. Both of these limitations are improved by [Tournier et al. \(2007\)](#) and validated by [Tournier et al. \(2008\)](#) in a method named super-resolved constrained spherical deconvolution (super-CSD).

The two ideas behind super-CSD are to use a Tikhonov regularisation (Hansen, 1994) to improve the robustness to noise and a super-resolution technique (Starck et al., 2002) to improve the angular resolution. The problem to solve is

$$\underset{\hat{\mathbf{f}}}{\text{minimize}} \quad \|\mathbf{Y}\mathbf{R}\hat{\mathbf{f}} - \mathbf{s}\|_2^2 + \lambda^2 \|\mathbf{P}\hat{\mathbf{f}}\|_2^2 \quad (6.4)$$

where λ is a regularization parameter and $\mathbf{P}$ is a regularization matrix. The problem of (6.4) is solved several times, updating the matrix $\mathbf{P}$ as follows. Given an initial estimation of $\hat{\mathbf{f}}$, we compute its projection using

$$\mathbf{v} = \mathbf{Y}'\hat{\mathbf{f}}$$

where $\mathbf{Y}'$ is an $M \times L^2$ spherical harmonic transform matrix. The vector $\mathbf{v}$ represents the estimated fiber ODF along the M directions determined by the matrix $\mathbf{Y}'$. Because $f(\theta, \varphi)$ is expected to be sharp, $\mathbf{v}$ should be sparse. To enforce this assumption, $\mathbf{P}$ is defined as

$$\mathbf{P}_{mn} = \begin{cases} \mathbf{Y}'_{mn} & v_m < \tau \\ 0 & \text{otherwise} \end{cases}$$

for some threshold τ . The term $\|\mathbf{P}\hat{\mathbf{f}}\|_2^2$ therefore increases when the fiber ODF has non-zero values below τ . The rational behind this approach is that the fiber ODF represents a physical distribution of fibers and so cannot take on negative values.

To increase the number of parameters that can be estimated, Tournier et al. (2007) propose to rewrite the problem as

$$\underset{\hat{\mathbf{f}}}{\text{minimize}} \quad \|\mathbf{Q}\hat{\mathbf{f}} - \mathbf{s}'\|_2^2 \quad (6.5)$$

where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Y}\mathbf{R} \\ \lambda\mathbf{P} \end{bmatrix} \quad \text{and} \quad \mathbf{s}' = \begin{bmatrix} \mathbf{s} \\ \mathbf{0} \end{bmatrix}.$$

As discussed by the authors, the problem in (6.5) can be solved as long as $\mathbf{v}$ has enough values below τ to make $\mathbf{Q}$ square or overdetermined.

6.3 Spherical FRI for diffusion MRI (sFRIdMRI)

The model used in super-CSD obviously has several similarities to the ones we presented in Chapter 3, 4, and 5. In fact, if the fiber ODF can be modeled as a sum of orientations, the algorithms we described in this thesis can be used directly to recover their parameters. Specifically, the fiber response function modeled by the matrix $\mathbf{R}$ corresponds to the sampling kernel and each sample location corresponds to a diffusion weighted direction. However, the interpretation of $f(\theta, \varphi)$ changes significantly. In spherical deconvolution, this function represents the fiber distribution as a function of the orientation in a voxel. It is a physical quantity and, as a direct consequence, $f(\theta, \varphi)$ should be non-negative over its domain. In comparison, $f(\theta, \varphi)$ in the context of spherical FRI is a mathematical object that represents the rotations and scaling

applied to the sampling kernel. If all the weights a_k are positive, then $f(\theta, \varphi)$ is non-negative everywhere. However, what we recover in (6.3) is not $f(\theta, \varphi)$ itself but its lowpass version. Indeed, if $f(\theta, \varphi)$ is a sum of K orientations, its bandlimited version is expected to take on negative values. Consider the analogous situation of a Dirac and its lowpass approximation, a sinc. While the Dirac is non-negative everywhere, the sinc has negative sidelobes.

The recovery of the spherical harmonics coefficients $\hat{\mathbf{f}}$ is the first step of our algorithm. In the previous chapters, we recovered them using a least-squares approach as in (6.3). This yielded good results because we had the luxury of choosing our kernel and we selected kernels that led to well-conditioned matrices $\mathbf{R}$. The least-squares approach is not ideal here because, as we discussed, the kernel that models the diffusion process has many small values, especially for low b -values. A possible solution is to use the coefficients estimated using super-CSD in (6.4) as an input to spherical FRI. While this idea is attractive, our numerical experiment using this methodology did not yield good results. The cause of the problem is the difference in the interpretation of $f(\mathbf{u})$ of the techniques. The active suppression of negative values of $f(\mathbf{u})$ in super-CSD leads to a function with sharp peaks, but whose spherical harmonic coefficients do not correspond to those of orientations. They are therefore a poor choice as an input to spherical FRI. In the following subsection, we propose an alternative approach to estimating the vector $\hat{\mathbf{f}}$ which is aligned with our assumptions on $f(\mathbf{u})$.

6.3.1 Estimating the spherical harmonic coefficients of orientations in dMRI

As we showed in Section 3.2, the spherical harmonic coefficients of a sum of K orientations are given by

$$\hat{f}_\ell^m = \sum_{k=1}^K a_k Y_\ell^m(\theta_k, \varphi_k).$$

In matrix form, we have

$$\hat{\mathbf{f}} = \mathbf{\Lambda}^* \mathbf{a} \quad (6.6)$$

where $\hat{\mathbf{f}}$ is the usual vector of spherical harmonic coefficients, $\mathbf{a} = [a_1 \ \cdots \ a_K]^t$ and

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{Y}_0(\theta_1, \varphi_1) & \mathbf{Y}_1(\theta_1, \varphi_1) & \cdots & \mathbf{Y}_{L-1}(\theta_1, \varphi_1) \\ \mathbf{Y}_0(\theta_2, \varphi_2) & \mathbf{Y}_1(\theta_2, \varphi_2) & \cdots & \mathbf{Y}_{L-1}(\theta_2, \varphi_2) \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{Y}_0(\theta_K, \varphi_K) & \mathbf{Y}_1(\theta_K, \varphi_K) & \cdots & \mathbf{Y}_{L-1}(\theta_K, \varphi_K) \end{bmatrix}.$$

with $\mathbf{Y}_\ell(\theta, \varphi)$ defined in (2.11). We can also add the constraint that $a_k > 0$ because fibers can only cause a loss of signal, not a signal gain. Building the matrix $\mathbf{\Lambda}$ in (6.6) requires the knowledge of the parameters (θ_k, φ_k) which we do not have. However, if we discretize the angles (θ, φ) , we can build an overdetermined matrix $\mathbf{\Lambda}$ whose additional rows correspond to 0 entries in $\mathbf{a}$, making it K sparse. We can then approximate

$\hat{\mathbf{f}}$ by solving the problem

$$\begin{aligned} & \underset{\mathbf{a}}{\text{minimize}} \quad \|\mathbf{s} - \mathbf{Y}\mathbf{R}\mathbf{\Lambda}^*\mathbf{a}\|_2 + \lambda\|\mathbf{a}\|_1 \\ & \text{subject to} \quad \mathbf{a} \geq 0 \end{aligned} \quad (6.7)$$

where λ is a regularization parameter. The parameter λ quantifies the trade-off between sparsity and data agreement. Its value therefore depends on noise power, as will be discussed shortly. The constraint $\mathbf{a} \geq 0$ must hold element-wise, that is, all elements of $\mathbf{a}$ must be non-negative. Given the solution $\mathbf{a}$, the approximated coefficients are given by $\mathbf{\Lambda}^*\mathbf{a}$. In practice this approximation is accurate for low degree coefficients, but tends to underestimate high degree coefficients because of their non-linearity. Because the power spectrum of a sum of K orientations is flat, we add a second Tikhonov style regularization term to the problem:

$$\begin{aligned} & \underset{\mathbf{a}}{\text{minimize}} \quad \|\mathbf{s} - \mathbf{Y}\mathbf{R}\mathbf{\Lambda}^*\mathbf{a}\|_2 + \lambda_1\|\mathbf{a}\|_1 + \lambda_2\|\mathbf{\Lambda}^*\mathbf{a}\|_2 \\ & \text{subject to} \quad \mathbf{a} \geq 0 \end{aligned} \quad (6.8)$$

The problem of (6.8) is convex and can be solve using standard convex optimization methods. We used CVX, a package for specifying and solving convex programs (Grant and Boyd, 2008, 2012). Using the approximated coefficients $\hat{\mathbf{f}} = \mathbf{\Lambda}^*\mathbf{a}$, we can recover the parameters of $f(\mathbf{u})$, i.e. the fibers orientations and amplitudes, using the algorithm described in Section 5.2. This method is what we refer to as sFRIdMRI.

The quality of the solution found when solving the problem of (6.8) depends on the parameters λ_1 and λ_2 . Their optimal value will in turn depend on the signal to noise ratio of the samples $\mathbf{s}$. Because recovering the spherical harmonics $\hat{\mathbf{f}}$ is one step in the full fiber reconstruction scheme, which may include denoising procedures, near optimal values for λ_1 and λ_2 cannot be determined until the whole scheme is known. In this work, the parameters used have been selected using trial-and-error. This approach is obviously suboptimal in a real setting as it is time consuming and error prone. However, when the full reconstruction scheme is selected, several methods are available to choose near optimal parameters automatically (Belge et al., 2002; Bhave et al., 2014; Ramani et al., 2008; Candes et al., 2013; Ramani et al., 2012).

6.3.2 Kernel bandlimit and b-value

In the signal model described by (6.2), the kernel $r(\mathbf{G})$ is dictated by the diffusion process. While we cannot alter or choose the kernel directly, we can influence it using the parameters of the acquisition, specifically the b -value. The magnitude of the rotational harmonic coefficients, i.e. the coefficients of the sampling kernel, associated with different b -values is presented in Figure 6.1. As the b -value increases, the bandlimit of the kernel also increases. Examples of synthetic signals obtained for b -values of 1000 and 3000 s/mm² are illustrated in Figure 6.2.

Figure 6.1: Magnitude of the rotational harmonic coefficients associated with a b -value of 1000 s/mm^2 in (a), 3000 s/mm^2 in (b), and 10000 s/mm^2 in (c).

Figure 6.2: Example of synthetic dMRI signals generated using the multi-tensor model for $K = 1, b = 1000 \text{ s/mm}^2$ in (a), $K = 1, b = 3000 \text{ s/mm}^2$ in (b), $K = 2, b = 1000 \text{ s/mm}^2$ in (c), and $K = 2, b = 3000 \text{ s/mm}^2$ in (d).

6.3.3 The advantage of sFRIdMRI

Like all fiber reconstruction algorithms used in dMRI, the angular resolution of spherical deconvolution is limited by the number of measurements. This dependency is detrimental because to have a high angular resolution we must acquire more data, an expensive resource in dMRI. In contrast, the angular resolution of sFRIdMRI is independent of the number of measurements. In theory, its ability to recover fibers depends only on the bandlimit of the kernel. For example, suppose we wish to recover two fibers using an antipodal kernel. As we showed in Section 4.2, the bandlimit must satisfy $L \geq 4K - 1$. This means that given $N \geq 49$ encoding directions and a sufficiently high SNR, we should be able to accurately recover the two fibers *independently of their crossing angle*. This is a significant advantage that none of the current fiber reconstruction algorithms can claim.

6.4 Numerical simulations

6.4.1 Data simulation and performance metrics

To generate synthetic data, we do the following:

1. Choose the encoding directions $\mathbf{q} = 1, \dots, N$. We use the algorithm in Appendix A to distribute the N encoding directions on the half-sphere $\varphi > \pi$.
2. For each fiber, generate a random orientation $\mathbf{u}_k$ with $\|\mathbf{u}_k\|_2 = 1$ and a weight a_k satisfying

$$\sum_{k=1}^K a_k = 1.$$

As in the work of Descoteaux et al. (2007) we add the constraints $0.3 \leq a_k \leq 0.7$ if $K = 2$ and $0.2 \leq a_k \leq 0.4$ if $K = 3$ to make sure all fibers contribute significantly to the data.

3. For each fiber, compute the diffusion tensor $\mathbf{D}_k$ whose eigenvalues are 300×10^{-6} , 300×10^{-6} , and 1700×10^{-6} mm²/s.
4. For each encoding direction and a specific b -value, compute the noiseless signal $s(\mathbf{q})$ using (6.1).
5. Add complex Gaussian noise with a standard deviation selected according to the desired signal to noise ratio and compute the magnitude of the result to obtain the noisy signal, as described in Descoteaux et al. (2006) and Sijbers et al. (1998).

To evaluate the performance of our reconstruction algorithm, we use the same parameters as Michailovich et al. (2010) and Descoteaux et al. (2007): the angular error and the rate of correct detection. The angular error is defined as

$$\alpha_k = \frac{180}{\pi} \cos^{-1}(|v_k^* v'_k|)$$

and the mean angular error as

$$\alpha = \frac{1}{K} \sum_{k=1}^K \alpha_k$$

which is similar to the measure used in Chapter 3. The only difference is the addition of the absolute value, because recovering an orientation v'_k or its opposite $-v'_k$ is equivalent in the context of dMRI. Typically, the rate of correct detection is the number of times the algorithm detected the correct number of fibers over the total number of simulations. However, in sFRIdMRI, the number of fibers is provided as an input leading to 100% rate of correct detection. To make a fair comparison, we modify the definition to be the rate at which all fibers are recovered to 10 degrees, that is $\alpha_k < 10$ for all k .

6.4.2 Parameter selection

Both super-CSD and sFRIdMRI have parameters whose values must be selected *a priori*. For super-CSD, we used the values suggested by Tournier et al. (2008) which are $\lambda = 1$, $L = 13$, and $\tau = 10\%$ of the mean initial ODF. For sFRIdMRI, we used $\lambda_1 = \lambda_2 = 0.005$ for normalized measurements i.e. when the maximum value of the observed samples is 1. This value was obtained empirically and, in our experiments, offered a good trade-off between angular resolution and accuracy. The matrix Λ was computed using 100 orientations uniformly distributed over the half sphere $\varphi > \pi$. As in Section 5.1.3, the matrix used in the OMP step was generated using 50 values of θ in the interval $[0, \pi]$. In addition, sFRIdMRI requires the value of K as an input. While we could devise an algorithm to choose the value of K on a voxel by voxel basis, we propose to set it to $K = 2$ for all voxels. When the value of K is overestimated, our simulations show that sFRIdMRI tends to return the “extra” orientation with an amplitude very near zero. This means that these superfluous fibers will have little to no effect on the following fiber tracking algorithms. Because we choose $K = 2$ and the kernel used in dMRI is antipodal, we find that the minimum bandlimit required by sFRIdMRI is $L = 4K - 1 = 7$ which is what we use.

6.4.3 Results

We first evaluate the angular resolution that can be achieved using sFRIdMRI by computing its detection rate. We simulated signals from a voxel with two fiber populations crossing at a specific angle. The signal to noise ratio is set to 20 dB and the separation angle varied from 20 to 90° by steps of 10°. The experiment was repeated 512 times for each separation angle and for b -values of 1000 and 3000 s/mm². For comparison, we also included the results of super-CSD. The results are illustrated in Figure 6.3. Typical results obtained for different crossing angles are also illustrated in Figure 6.4 and 6.5. At a b -value of 3000, both algorithms are able to distinguish the two fiber populations with a rate near 1 as long as they are separated by 40° or more. For super-CSD, its ability to retrieve the two fibers then reduces rapidly, reaching 0 at 20-30°. In contrast, sFRIdMRI is able to recover fibers separated by 20-30° 60% of the time. As one would expect, using higher b -values leads to a better rate at lower separation angle. With a b -value of 3000, super-CSD begins to fail at 40-50° whereas sFRIdMRI still has a detection rate above 95% at 30-40°. The lower detection rate of super-CSD can be explained by observing Figures 6.4 and 6.5. At low separation angles, the function $f(u)$ recovered in super-CSD exhibits a single maximum which is returned by the Newton-Raphson gradient descent. In the case of sFRIdMRI, a single maximum also occurs, but because the algorithm does not rely on peak finding, it is still able to recover multiple fibers.

To evaluate the angular resolution, we simulated signals from a voxel with two fiber populations at a crossing angle uniformly distributed in the range 20 to 90°. We then computed the mean angular error as a function of the signal to noise ratio, which was varied from 0 to 30dB. The experiment was repeated 512 times for each SNR. The

Figure 6.3: Detection rate as a function of the separation angle between 2 fibers. The signals were generated using a b -value of 1000 s/mm^2 in (a) and 3000 s/mm^2 in (b), and a signal to noise ratio of 20dB.

results obtained are illustrated in Figure 6.6. As expected, the mean angular error decreases as the SNR increase. For a SNR below 5dB, super-CSD offers a slightly lower angular error than sFRIdMRI. Above 5dB, the situation is reversed and sFRIdMRI has a lower mean angular error.

We also simulated whole volumes to see if the improved angular resolution of sFRIdMRI has a significant impact on tractography results. All the fiber tracking results were computed using DSISStudio (<http://dsi-studio.labsolver.org>). A $10 \times 10 \times 10$ volume with two fibers fanning or braching is illustrated in Figure 6.7. The ground truth is the fiber tracking obtained when using the known fiber orientations for each voxel. By comparing the results of super-CSD and sFRIdMRI, it is clear that the latter recovers fibers that are closer to the ground truth. Specifically, the sFRIdMRI results are free from the overshooting fibers that miss the braching. This a direct consequence of the ability of sFRIdMRI to distinguish fibers crossing at a narrow angle at

Figure 6.4: Fiber reconstruction examples for separation angles of 90, 60, and 50°. Other angle are presented in Figure 6.5.

Figure 6.5: Fiber reconstruction examples for separation angles of 40, 30, and 20°.

Figure 6.6: Angular error as a function of the signal to noise ratio for a signal with two fibers. The signals were generated using a b -value of 1000 in (a) and 3000 s/mm² in (b), and a separation angle uniformly distributed in the range 20 to 90°.

the branching point. Another example where 2 fiber populations cross at a narrow angle is fiber kissing, illustrated in Figure 6.8. The fibers computed using the super-CSD result again overshoot near the kissing fibers. On the other hand, sFRIdMRI recovers the fibers more reliably. Finally, to emphasize that overestimating the value of K will not have a severe impact on the results, we simulated a volume with a single curving fiber, which is illustrated in Figure 6.9. In this case, super-CSD finds a solution that is almost identical to the ground truth. When sFRIdMRI is used with a value of $K = 2$, the fibers orientations are still recovered, but are not as smooth. More importantly, the tracking algorithm does not find an additional incorrect fiber population, therefore supporting our claim that overestimating K is acceptable. For comparison, if we set $K = 1$, the solution found resembles the one of super-CSD and the ground truth.

Figure 6.7: Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for 2 fibers fanning or merging. The signals were simulated using a b -value of 1000 s/mm^2 and a signal to noise ratio of 20dB.

6.5 In vivo data

In this section, we present reconstruction results obtained by applying sFRIdMRI to real diffusion MRI data. Because diffusion MRI is the only imaging technique that can probe the micro-structure of the brain in vivo, no ground truth can be provided for comparison. A typical approach, used in most diffusion MRI publications, is to compare the reconstruction results to known anatomical landmarks (see [Catani and Thiebaut de Schotten \(2012\)](#) as a reference of brain connections). While this does not quantitatively assess the results, it does provide a qualitative evaluation which reinforces our confidence in the results. We show results for two regions of interest where the corpus callosum intersects the corticospinal tract.

For completeness, we also provide reconstructions obtained using super-CSD on the same data. Reconstructions were performed using the MRtrix package (J-D Tournier, Brain Research Institute, Melbourne, Australia, <http://www.brain.org.au/software/>)([Tournier et al., 2012](#)). Finally, we add an image of fiber tracking on

Figure 6.8: Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for 2 fibers kissing. The signals were simulated using a b -value of 1000 s/mm^2 and a signal to noise ratio of 20dB.

the whole brain using the results of sFRIdMRI.

6.5.1 Data acquisition

The data we used was acquired by [Anwander et al. \(2007\)](#) and kindly provided by Maxime Descoteaux. The DWI data and high-resolution T1-weighted images were acquired in a healthy right-handed 25 year old volunteer on a whole-body 3T Magnetom Trio scanner (Siemens, Erlangen) equipped with an 8-channel head array coil. Written informed consent was obtained from all subjects in accordance with the ethical approval from the University of Leipzig. The spin-echo echo-planar-imaging sequence, $TE = 100 \text{ ms}$, $TR = 12 \text{ s}$, 128×128 image matrix, $FOV = 220 \times 220 \text{ mm}^2$, consists of 60 diffusion encoding gradients [Jones et al. \(1999\)](#) with a b -value of 1000 s/mm^2 . Seven images without any diffusion weightings are placed at the beginning of the sequence and after each block of 10 diffusion weighted images as anatomical reference for offline motion correction. The measurement of 72 slices with 1.7mm thickness (no gap) covered the whole brain. Random variations in the data were reduced by averaging

Figure 6.9: Comparison of fiber tracting results using fibers orientations computed with super-CSD and sFRIdMRI for a curving fiber. The signals were simulated using a b -value of 1000 s/mm^2 and a signal to noise ratio of 20dB.

3 acquisitions, resulting in an acquisition time of about 45 minutes. No cardiac gating was employed to limit the acquisition time. The issue of cardiac gating is discussed in Jones et al. (2002). The SNR in the white matter of this S_0 image was estimated to be approximately 37. Additionally, fat saturation was employed and we used 6/8 partial Fourier imaging, a Hanning window filtering and parallel acquisition (generalized auto-calibrating partially parallel acquisitions, reduction factor = 2) in the axial plane.

6.5.2 Results

Figure 6.10 illustrates the reconstruction results for the first region of interest. A highlight of the region is illustrated in Figure 6.10 (a) on a coronal slice of the fractional anisotropy map. Both sFRIdMRI and super-CSD recover the major landmarks which are the corpus callosum, the corticospinal tract and the cingulum. In comparison to super-CSD, sFRIdMRI finds more fiber crossings. This is particularly visible in the left side of the image where we can see the outward radiation of the corpus callosum pass through the corticospinal tract. It should be noted that although sFRIdMRI al-

ways find two fibers ($K = 2$ is a parameter), fibers which do not fit the data would have an amplitude of 0. These are not displayed in the figure or are barely visible. In other words, when two fibers are recovered with similar amplitudes, they fit the data similarly and are not spurious fibers added to satisfy $K = 2$.

Figure 6.11 illustrates the reconstruction results for the second region of interest. A highlight of the region is illustrated in Figure 6.11 (a) on a coronal slice of the fractional anisotropy map. Both algorithms recover the expected landmarks, sFRIdMRI again showing more regions of crossing fibers. Comparing the center of Figure 6.11 (b) and (c) shows a larger quantity of fibers crossing left to right. This represents where the corticospinal tract diverge into the motor cortex and is intersected by the lateral projections of the corpus callosum. Crossings are also more prominent for sFRIdMRI in the bottom left of the image where the superior longitudinal fasciculus crosses the projections of the corpus callosum.

Finally, Figure 6.12 illustrates the fiber tracking results on the full volume. These fiber tracking results were computed using DSISudio¹. While they are visually impressive, fiber tracking of complete brain volumes are difficult to validate and interpret. Some tools have been developed to quantify fiber tracking results (Côté et al., 2013), but validation of tractography results is out of the scope of this work. Figure 6.12 is included for aesthetic purposes.

¹<http://dsi-studio.labsolver.org>

Figure 6.10: Recovered fiber orientations for the first region of interest highlighted in a coronal slice of the fractional anisotropy map in (a). The anatomical landmarks are: the corpus callosum in green (bottom left), the corticospinal tract in blue (bottom right), and the cingulum in red (top left).

(a) ROI

(b) super-CSD $L = 9$ (maximum degree 8 in MRtrix) (c) sFRIdMRI $K = 2, L = 9$

Figure 6.11: Recovered fiber orientations for the second region of interest highlighted in a coronal slice of the fractional anisotropy map in (a). The anatomical landmarks are: the corpus callosum in green (middle right), the corticospinal tract in blue (bottom center), the cingulum in red (middle right), and the superior longitudinal fasciculus in red (bottom left).

Figure 6.12: Full volume fiber tracking results on real diffusion MRI data. The three views are sagittal in (a), frontal in (b), and perspective in (c). The color of a fiber represent its mean direction.

6.6 Summary

- The model diffusion magnetic resonance imaging signals used in spherical deconvolution is the same as the one we propose for spherical FRI. However, the interpretation of the function $f(\theta, \varphi)$ differs significantly. In spherical deconvolution, $f(\theta, \varphi)$ is a physical quantity that represent the distribution of fibers in a voxel. In sFRIdMRI, it is a mathematical tool that represents rotations applied to the kernel.
- We proposed a method to estimate the spherical harmonic coefficients of a sum of K orientations in the context of dMRI. This method uses both an ℓ_1 and ℓ_2 regularization term to improve the results. When this method is used with spherical FRI, we refer to the resulting algorithm as sFRIdMRI.
- In contrast to the fiber reconstruction algorithms used in dMRI, the angular resolution of sFRIdMRI does not depend on the number of encoding directions. This means that, if the kernel and SNR meet the requirement of spherical FRI, the fibers can be recovered independently of their separation angles.
- Our numerical simulations show that sFRIdMRI offers a better detection rate than super-CSD while having a comparable or lower mean angular error. As a result, sFRIdMRI offers better fiber tracking results when fibers cross at a narrow angle.
- When sFRIdMRI is used on real diffusion MRI data, the expected anatomical landmarks are recovered. Furthermore, our results correlate well with the results of super-CSD.

VII

Conclusions and Future Work

In this thesis, we have considered the reconstruction of non-bandlimited functions on the sphere with an application to diffusion magnetic resonance imaging. We give here a brief summary of each chapter and propose avenues of research for future work.

In Chapter 3, we developed a model for signals with a finite rate of innovation on the sphere, which we called a sum of K orientations. Just like FRI signals in a Cartesian space, a sum of K orientations is non-bandlimited and cannot be sampled and reconstructed using conventional approaches. Furthermore, because operations like convolution and Fourier transform have different definitions on the sphere, the current FRI methods are not adequate to handle sums of orientations. To sample and reconstruct them, we developed a new algorithm which proceeds in two steps. First, the sum of K orientations is convolved with a bandlimited kernel and it is the resulting function that is sampled. We proved, using the sampling theorem of [McEwen and Wiaux \(2011\)](#), that given an appropriate number of samples, it is possible to recover the spherical harmonic coefficients of the sum of orientations. Second, we proved that the form of these coefficients can be exploited to recover the parameters of the orientations. Specifically, we showed that certain coefficients can be reduced to a sum of exponentials and that the annihilating filter method ([Vetterli et al., 2002](#)) can be used to recover a subset of the parameters. The remaining parameters can then be obtained by solving two linear systems. One assumption we used throughout this work is that the observed signal is generated by convolving a single kernel with a single sum of orientations. An avenue of research is to consider a more general model that would allow signals generated by multiple kernels and multiple orientations. For example, let $f_m(\mathbf{u})$ and $r_m(\mathbf{G})$ be the m^{th} sum of orientations and kernel, respectively. Then the signal of interest could be defined as

$$s(\mathbf{u}) = \sum_{m=1}^M (f_m * r_m)(\mathbf{u}). \quad (7.1)$$

This would significantly broaden the range of signals that could be sampled and reconstructed. Multiresolution analysis on the sphere uses a model similar to 7.1) where each $r_m(\mathbf{G})$ corresponds to a different resolution. The typical approach, used in Michailovich et al. (2010), is to discretize the rotations that can be applied to $r_m(\mathbf{G})$ and to recover $s(\mathbf{u})$ using a sparse reconstruction algorithm such as OMP. Using an FRI approach to recover the function $f_m(\mathbf{u})$ could lead to better reconstructions by removing the need to discretize.

In Chapter 4, we focused on extending the ideas behind spherical FRI to other kernels and signal models. We defined an optimal kernel which allows the reconstruction of a sum of K orientations using only $3K$ measurements, the number of degrees of freedom of the signal. While it is unlikely that this kernel will be encountered in practice, it is interesting to know the optimal sampling conditions of a class of signals. In addition to the optimal kernel, we also considered antipodally symmetric kernels whose odd degree spherical harmonic coefficients are all zero. This characteristic is not compatible with the assumptions of Chapter 3 so we devised a modified approach specifically for antipodally symmetric kernels. To generalize the signal model, we investigated the reconstruction of integrated orientations. We showed that, just like orientations, a function whose derivative is a sum of orientations can also be sampled and reconstructed using a finite number samples. Finally, we presented how great circles, the intersection between a plane through the origin and a sphere centered at the origin, can be sampled at their rate of innovation. Obviously, many other parametric signal models can be considered. For example, a sum of cones which can be defined as

$$f(\mathbf{u}) = \sum_{k=1}^K a_k H(\mathbf{u}^t \mathbf{v}_k - \tau_k) \quad (7.2)$$

where $H(t)$ is the Heaviside function could also be considered. A potential for this model is dMRI, where requiring that all fibers be modeled as orientations is a bit restrictive. Allowing each fiber to have a “width” or fanning angle, controlled by τ_k in (7.2), would greatly enhance the model. However, it remains to be seen if sum of cones can be sampled and reconstructed efficiently.

In Chapter 5, we dealt with imperfect measurements and model mismatch. When dealing with noisy signals and a bandlimit set to the critical value, we showed that computing the parameters of orientations in a rotated coordinate system improved the performance of our algorithm. However, because we did not find a way to choose the best rotation beforehand, we proposed a brute force approach. While it does improve the reconstruction, this method is the source of an increase in computation time that is proportional to the number of rotations tested. A better solution could probably be devised by choosing the rotation as a function of the measurements. When the bandlimit is above the critical value, we can further improve the result by exploiting the redundancy. The most popular way to improve the performance of the annihilating filter method is a model fitting method called Cadzow’s denoising (Blu et al., 2008). While it does improve the recovery of the signals parameters, Cadzow’s denoising has some drawbacks discussed in Section 5.2.1. What we propose is to treat the recovery

of the annihilating filter as an optimization problem. We developed a cost function that measures how well a given model fits the observed data. We showed, using numerical simulations, that optimizing over this cost function greatly reduces the angular error when sampling and reconstructing a sum of K orientations. In addition to being non-iterative, this approach allows the use of all the facilities of an optimization problem, such as bounds, constraints, and regularization terms that are not available using Cadzow's denoising.

In Chapter 6, we applied our new reconstruction algorithm to the recovery of fiber orientations in diffusion magnetic resonance imaging. Because the kernel induced by the diffusion process has many low magnitude coefficients, we proposed a new method to recover the spherical harmonic coefficients from the measurements. Unlike spherical deconvolution and super-CSD, our method does not penalize the use of negative side-lobes. We justify this choice by observing that the signal created by lowpass filtering a sum of orientations is expected to take on negative values. When used in the context of dMRI, our algorithm is called sFRIdMRI. The main advantage of sFRIdMRI is that its angular resolution is not directly linked to the number of samples acquired. This means that, given an appropriate sampling scheme, it has the ability to differentiate crossing fibers irrespective of their crossing angle. This claim is substantiated by many numerical simulations. Indeed, our results show that sFRIdMRI has a better detection rate than super-CSD for fibers crossing at narrow angle. Even better, our results show that sFRIdMRI has an angular error similar to super-CSD, meaning we do not need to sacrifice precision for angular resolution. Of course, further testing and analysis on real data is required to determine the advantages and disadvantages of sFRIdMRI in a clinical setting. This is the subject of our current research.

Distributing points on the sphere

In this section, we describe an algorithm to distribute points almost uniformly on the sphere. Let $\mathbf{u}_i \in \mathbb{S}^2$ with $i = 1, \dots, N$ be the points we wish to distribute on the sphere. The points are randomly initialized and the algorithm proceeds as follows:

1. Compute the distance $d_{mn} = \arccos(\mathbf{u}_m^t \mathbf{u}_n)$ for $m, n = 1, \dots, N$.
2. For every point compute the force applied $f_{mn} = 1/d_{mn}^2$ and the direction of the force $\mathbf{v}_{mn} = \mathbf{u}_m - \mathbf{u}_n$. Note that the direction is not normalized.
3. Compute the resulting force at each point $\mathbf{u}_m$

$$\mathbf{f}_m = \sum_{n=1}^N f_{mn} \mathbf{v}_{mn}.$$

4. Project the resulting force on a plane orthogonal to the point where it is applied, that is

$$\boldsymbol{\mu}_m = \lambda \mathcal{N}(\mathbf{u}_m) \mathcal{N}(\mathbf{u}_m)^\dagger \mathbf{f}_m.$$

where $\lambda = 1/N$ and $\mathcal{N}(\mathbf{u}_m)$ is the 3×2 matrix whose columns are a basis for the null space of $\mathbf{u}_m$.

5. Update the position of the points $\mathbf{u}_m = \mathbf{u}_m + \boldsymbol{\mu}_m$.
6. Repeat all steps until the maximum displacement $\max \|\boldsymbol{\mu}_m\|_2$ is below some threshold (we use 10^{-5}).

The evolution of the position of the points $\mathbf{u}_m$ is illustrated in Figure A.1 using $N = 64$ points.

Figure A.1: Illustration of the evolution of the position of the points $\mathbf{u}_m$ using $N = 64$ points. The black and red dots represent the initial and final position of the points, respectively. The blue dots are the evolution of the position at each iteration.

B

Publications of the author arising from this work

Samuel Deslauriers-Gauthier and Pina Marziliano. Magnetic resonance image reconstruction using the annihilating filter method. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI 2011)*, March 30-April 2, 2011. Chicago, USA.

Samuel Deslauriers-Gauthier, Pina Marziliano, and Cher Heng Tan. Application of finite rate of innovation methods to the reconstruction of magnetic resonance images of the liver. *9th International Conference on Sampling Theory and Applications (SampTA)*, May 2-6, 2011. Singapore.

Samuel Deslauriers-Gauthier and Pina Marziliano. Spherical finite rate of innovation theory for the recovery of fiber orientations. *34th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC 2012)*, August 28-September 1, 2012. San Diego, USA.

Samuel Deslauriers-Gauthier and Pina Marziliano. Noisy finite rate of innovation beyond Cadzow. *Sampling Theory in Signal and Image Processing SampTA 2011 Special Issue*, 2013.

Samuel Deslauriers-Gauthier and Pina Marziliano. Sampling signals with a finite rate of innovation on the sphere. *IEEE Transactions in Signal Processing*, to appear 2013.

Samuel Deslauriers-Gauthier and Pina Marziliano. Sampling great circles at their rate of innovation. *SPIE: Wavelets and sparsity XV*, August 26-29, 2013. San Diego, USA.

Bibliography

- J. Ahrens and S. Spors. A modal analysis of spatial discretization of spherical loud-speaker distributions used for sound field synthesis. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(9):2564–2574, 2012.
- Y. Alem, D. Chae, and R. Kennedy. Sparse signal recovery on the sphere: Optimizing the sensing matrix through sampling. In *6th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pages 1–6, 2012.
- Y. Alem, S. Salehin, D. Chae, and R. Kennedy. Sparse recovery of spherical harmonic expansions from uniform distribution on sphere. In *7th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pages 1–5, 2013.
- A. L. Alexander, K. M. Hasan, M. Lazar, J. S. Tsuruda, and D. L. Parker. Analysis of partial volume effects in diffusion tensor MRI. *Magnetic Resonance in Medicine*, 45(5):770–780, 2001.
- D. C. Alexander. *Maximum entropy spherical deconvolution for diffusion MRI*. Information Processing in Medical Imaging, Springer Berlin Heidelberg, 2005.
- D. C. Alexander, G. J. Barker, and S. R. Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2):331–340, 2002.
- G. E. Andrews, R. Askey, and R. Roy. *Special functions*, volume 71 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1999.
- A. Anwender, M. Tittgemeyer, D. von Cramon, A. Friederici, and T. Knäuper. Connectivity-based parcellation of broca’s area. *Cerebral Cortex*, 17(4):816–825, 2007.
- H. E. Assemlal, D. Tschumperlé, L. Brun, and K. Siddiqi. Recent advances in diffusion MRI modeling: Angular and radial reconstruction. *Medical Image Analysis*, 15(4):369–396, 2011.
- L. Baboulaz and P. L. Dragotti. Exact feature extraction using finite rate of innovation principles with an application to image super-resolution. *IEEE Transactions on Image Processing*, pages 281–298, 2009.
- P. J. Basser, J. Mattiello, and D. Lebihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66(1):259–267, 1994.
- T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, and M. W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, 34(1):144–155, 2007.

- M. Belge, M. E. Kilmer, and E. L. Miller. Efficient determination of multiple regularization parameters in a generalized l-curve framework. *Inverse Problems*, 18(4): 1161, 2002.
- Z. Ben-Haim, T. Michaeli, and Y. C. Eldar. Performance bounds for the estimation of finite rate of innovation signals from noisy measurements. In *IEEE Sensor Array and Multichannel Signal Processing Workshop*, pages 97–100, 2010.
- S. Bhawe, R. Eslami, and M. Jacob. Sparse spectral deconvolution algorithm for noncartesian mr spectroscopic imaging. *Magnetic Resonance in Medicine*, 71:469–476, 2014.
- T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot. Sparse sampling of signal innovations. *IEEE Signal Processing Magazine*, 25(2):31–40, 2008.
- S. Boyd. *Convex optimization*. Cambridge University Press, 2004.
- D. M. Brink and G. R. Satchler. *Angular momentum*. Oxford University Press, 1993.
- J. Cadzow. Signal enhancement: A composite property mapping algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36:49–62, 1988.
- E. J. Candes, C. A. Sing-Long, and J. D. Trzasko. Unbiased risk estimates for singular value thresholding and spectral estimators. *IEEE Transactions on Signal Processing*, 61(19):4643–4657, 2013.
- M. Catani and M. Thiebaut de Schotten. *Atlas of human brain connections*. Oxford University Press, 2012.
- B. Chazelle. *The discrepancy method: Randomness and complexity*. Cambridge University Press, 2000.
- C. Chen, P. Marziliano, and A. Kot. 2D finite rate of innovation reconstruction method for step edge and polygon signals in the presence of noise. *IEEE Transactions on Signal Processing*, 60(6):2851–2859, 2012.
- E. U. Condon and G. H. Shortley. *The theory of atomic spectra*. Cambridge University Press, 1935.
- T. E. Conturo, N. F. Lori, T. S. Cull, E. Akbudak, A. Z. Snyder, J. S. Shimony, R. C. McKinstry, H. Burton, and M. E. Raichle. Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18):10422–10427, 1999.
- M. Côté, G. Girard, A. Boré, E. Garyfallidis, J. Houde, and M. Descoteaux. Tractometer: Towards validation of tractography pipelines. *Medical Image Analysis*, 17: 844–857, 2013.

- R. G. Crittenden and N. G. Turok. Exactly azimuthal pixelizations of the sky. *arXiv*, 1998.
- C. de Boor. *A practical guide to splines*. Springer-Verlag, 2001.
- M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. *Magnetic Resonance in Medicine*, 56:395–410, 2006.
- M. Descoteaux, E. Angelino, S. Fitzgibbons, and D. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58:497–510, 2007.
- S. Deslauriers-Gauthier and P. Marziliano. Noisy finite rate of innovation beyond Cadzow. *Sampling Theory in Signal and Image Processing SampTA 2011 Special Issue*, 2013.
- D. L. Donoho, A. Maleki, and A. Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.
- A. G. Doroshkevich, P. Naselsky, O. Verkhodanov, D. Novikov, V. Turchaninov, I. Novikov, P. Christensen, and L.-Y. Chiang. Gauss-legendre sky pixelization (GLESP) for CMB maps. *arXiv*, 2005.
- P. L. Dragotti and F. Homann. Sampling signals with finite rate of innovation in the presence of noise. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2941–2944, 2009.
- P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Transactions on Signal Processing*, 55(5):1741–1757, 2007.
- J. R. Driscoll and D. M. Healy. Computing Fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15(2):202–250, 1994.
- A. R. Edmunds. *Angular momentum in quantum mechanics*. Princeton University Press, 1996.
- P. F. Fougere. A defense of the Gram-Schmidt orthogonalization procedure applied to spherical harmonic analysis. *Journal of Geophysical Research*, 71(21):5171–5174, 1966.
- P. Funk. Über flächen mit lauter geschlossenen geodätischen linien. *Mathematische Annalen*, 74(2):278–300, 1913.
- P. Funk. Über eine geometrische anwendung der abelschen integralgleichung. *Mathematische Annalen*, 77(1):129–135, 1915.

- K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622:759–771, 2005.
- M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.
- M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. <http://cvxr.com/cvx>, September 2012.
- J. P. Haldar and R. M. Leahy. Linear transforms for Fourier data on the sphere: Application to high angular diffusion MRI of the brain. *NeuroImage*, 71:233–247, 2013.
- C. Hansen. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms*, 6(1):1–35, 1994.
- D. M. Healy, H. Hendriks, and P. T. Kim. Spherical deconvolution. *Journal of Multivariate Analysis*, 67(1):1–22, 1998.
- D. M. Healy, D. N. Rockmore, P. J. Kostelec, and S. Moore. FFTs for the 2-sphere: Improvements and variations. *The Journal of Fourier Analysis and Applications*, 9(4):341–385, 2003.
- C. P. Hess, P. Mukherjee, E. T. Han, D. Xu, and D. B. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine*, 56(1):104–117, 2006.
- T. Hosey, G. Williams, and R. Ansorge. Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54(6):1480–1489, 2005.
- D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor. Rigid sphere room impulse response simulation: Algorithm and applications. *The Journal of the Acoustical Society of America*, 132(3):1462–1472, 2012.
- J. H. Jensen, J. A. Helpert, A. Ramani, H. Lu, and K. 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.
- 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.
- D. Jones, M. A. Horsfield, and A. Simmons. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine*, 42(3):515–525, 1999.

- D. Jones, S. R. Williams, D. Gasston, M. Horsfield, A. Simmons, and R. Howard. Isotropic resolution diffusion tensor imaging with whole brain acquisition in a clinically acceptable time. *Human Brain Mapping*, 15:216–230, 2002.
- S. M. Kay. *Modern spectral estimation: Theory and application*. Prentice Hall, 1988.
- S. M. Kay and S. Marple. Spectrum analysis: A modern perspective. *Proceedings of the IEEE*, 69(11):1380–1419, 1981.
- J. Kennedy and R. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948, 1995.
- S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
- P. J. Kostelec and D. N. Rockmore. Ffts on the rotation group. *Journal of Fourier Analysis and Applications*, 14(2):145–179, 2008.
- B. W. Kreher, J. F. Schneider, I. Mader, E. Martin, J. Hennig, and K. A. Il'yasov. Multi-tensor approach for analysis and tracking of complex fiber configurations. *Magnetic Resonance in Medicine*, 54(5):1216–1225, 2005.
- J. Kronander, D. Jonsson, J. Löw, P. Ljung, A. Ynnerman, and J. Unger. Efficient visibility encoding for dynamic illumination in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):447–462, 2012.
- S. Kunis and D. Potts. Fast spherical Fourier algorithms. *Journal of Computational and Applied Mathematics*, 161(1):75 – 98, 2003.
- C. Liu, R. Bammer, B. Acar, and M. E. Moseley. Characterizing non-Gaussian diffusion by using generalized diffusion tensors. *Magnetic Resonance in Medicine*, 51(5):924–937, 2004.
- H. Lu, J. H. Jensen, A. Ramani, and J. A. Helpert. Three-dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, 19(2):236–247, 2006.
- N. Macon and A. Spitzbart. Inverses of vandermonde matrices. *The American Mathematical Monthly*, 65(2):pp. 95–100, 1958.
- I. Maravić and M. Vetterli. Exact sampling results for some classes of parametric nonbandlimited 2D signals. *IEEE Transactions on Signal Processing*, 52(1):175–189, 2004.
- I. Maravić and M. Vetterli. Sampling and reconstruction of signals with finite rate of innovation in the presence of noise. *IEEE Transactions on Signals Processing*, 53(8):2788–2805, 2005.

- S. Marple. A tutorial overview of modern spectral estimation. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 2152–2157, 1989.
- P. Marziliano. *Sampling innovations*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2001.
- J. D. McEwen. Fast, exact (but unstable) spin spherical harmonic transforms. *All Results Journals: Phys*, 1(1):4–18, 2011.
- J. D. McEwen and Y. Wiaux. A novel sampling theorem on the sphere. *IEEE Transactions on Signal Processing*, 59(12):5876–5887, 2011.
- J. D. McEwen, G. Puy, J. P. Thiran, P. Vandergheynst, D. Van De Ville, and Y. Wiaux. Sampling theorems and compressive sensing on the sphere. In *SPIE: Wavelets and Sparsity XIV*, 2011.
- L. Melie-García, E. J. Canales-Rodríguez, Y. Alemán-Gómez, C.-P. Lin, Y. Iturria-Medina, and P. A. Valdés-Hernández. A Bayesian framework to identify principal intravoxel diffusion profiles based on diffusion-weighted MR imaging. *NeuroImage*, 42(2):750–770, 2008.
- O. Michailovich, Y. Rathi, and S. Dolui. On approximation of orientation distributions by means of spherical ridgelets. *IEEE Transactions on Image Processing*, 19(2):461–477, 2010.
- L. Minati, D. Aquino, S. Rampoldi, S. Papa, M. Grisoli, M. G. Bruzzone, and E. Maccagnano. Biexponential and diffusional kurtosis imaging, and generalised diffusion-tensor imaging (GDTI) with rank-4 tensors: A study in a group of healthy subjects. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 20(5-6):241–253, 2007.
- S. Mori, B. J. Crain, V. P. Chacko, and P. C. M. V. Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology*, 45(2):265–269, 1999.
- D. Needell and R. Vershynin. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):310–316, 2010.
- A. F. Nikiforov, S. K. Suslov, and V. B. Uvarov. *Classical orthogonal polynomials of a discrete variable*. Springer-Verlag, 1991.
- F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, 2010.
- A. V. Oppenheim and A. S. Willisky. *Signals and systems (Second edition)*. Prentice Hall, 1997.

- E. Ozarslan and T. H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50(5):955–965, 2003.
- S. Peled, O. Friman, F. Jolesz, and C. F. Westin. Geometrically constrained two-tensor model for crossing tracts in DWI. *Magnetic Resonance Imaging*, 24(9):1263–1270, 2006.
- D. Potts, G. Steidl, and M. Tasche. Fast and stable algorithms for discrete spherical Fourier transforms. *Linear Algebra and its Applications*, 275-276:33 – 450, 1998.
- W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes: The art of scientific computing (Third edition)*. Cambridge University Press, 2007.
- J. L. Prince and J. M. Links. *Medical Imaging Signals and Systems*. Pearson Prentice Hall, 2006.
- J. Radon. Über die bestimmung von funktionen durch ihre integralwerte längs gewissermannigfaltigkeiten. *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math. Nat. Kl.*, 69: 262–277, 1917.
- S. Ramani, T. Blu, and M. Unser. Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms. *IEEE Transactions on Image Processing*, 17(9):1540–1554, 2008.
- S. Ramani, L. Zhihao, J. Rosen, J.-F. Nielsen, and J. A. Fessler. Regularization parameter selection for nonlinear iterative image restoration and mri reconstruction using gcv and sure-based methods. *IEEE Transactions on Image Processing*, 21(8): 3659–3672, 2012.
- H. Rauhut and R. Ward. Sparse recover for spherical harmonic expansions. In *9th International Conference on Sampling Theory and Applications (SampTA)*, 2011.
- T. Risbo. Fourier transform summation of Legendre series and D-functions. *Journal of Geodesy*, 70(7):383–396, 1996.
- R. E. Schwartz. The 5-electron case of Thomson’s problem. *Experimental Mathematics*, to appear 2013.
- C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37 (1):10–21, 1949.
- C. E. Shannon. Classic paper: Communication in the presence of noise. *Proceedings of the IEEE*, 86(2):447–457, 1998.
- P. Shukla and P. L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 55(7):3670–3686, 2007.

- W. Shukowsky. A quadrature formula over the sphere with application to high resolution spherical harmonic analysis. *Bulletin géodésique*, 60(1):1–14, 1986.
- J. Sijbers, A. D. Dekker, J. V. Audekerke, M. Verhoye, and D. V. Dyck. Estimation of the noise in magnitude MR images. *Magnetic Resonance Imaging*, 16(1):87–90, 1998.
- P. Sloan, J. Kautz, and J. Snyder. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. volume 21, pages 527–536, 2002.
- D. N. Spergel, R. Bean, O. Dore, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-year Wilkinson microwave anisotropy probe (WMAP) observations: Implications for cosmology. *The Astrophysical Journal Supplement Series*, 170(2):377–408, 2007.
- J. L. Starck, E. Pantin, and F. Murtagh. Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific*, 114(800):pp. 1051–1069, 2002.
- J.-D. Tournier, F. Calamante, D. G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23(3):1176–1185, 2004.
- J.-D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35(4):1459–1472, 2007.
- J. D. Tournier, C.-H. Yeh, F. Calamante, K.-H. Cho, A. Connelly, and C.-P. Lin. Resolving crossing fibres using constrained spherical deconvolution: Validation using diffusion-weighted imaging phantom data. *NeuroImage*, 42(2):617–625, 2008.
- J.-D. Tournier, F. Calamante, and A. Connelly. Mrtrix: Diffusion tractography in crossing fiber regions. *International Journal of Imaging Systems and Technology*, 22: 53–66, 2012.
- J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12):4655–4666, 2007.
- D. S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52:1358–1372, 2004.
- D. S. Tuch, R. M. Weisskoff, J. W. Belliveau, and V. J. Wedeen. High angular diffusion imaging of the human brain. *Magnetic Resonance in Medicine*, page 321, 1999.
- D. S. Tuch, T. G. Reese, M. R. Wiegell, N. Makris, J. W. Belliveau, and V. J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic resonance in medicine*, 48:577–582, 2002.

- D. W. Tufts and R. Kumaresan. Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood. volume 70, pages 975–989, 1982.
- R. Tur, Y. C. Eldar, and Z. Friedman. Innovation rate sampling of pulse streams with application to ultrasound imaging. *IEEE Transactions on Signal Processing*, pages 1827–1842, 2011.
- M. Unser. Sampling: 50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, 2000.
- D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum theory of angular momentum*. World Scientific Publishing, 1988.
- M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.
- V. J. Wedeen, P. Hagmann, W.-Y. I. Tseng, T. G. Reese, and R. M. Weisskoff. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, 54:1377–1386, 2005.
- V. J. Wedeen, R. P. Wang, J. D. Schmahmann, T. Benner, W. Y. I. Tseng, G. Dai, D. Pandya, P. Hagmann, H. D’Arceuil, and A. de Crespigny. Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage*, 41(4):1267 – 1277, 2008.
- D. White, A. J. Kimerling, and W. S. Overton. Cartographic and geometric components of a global sampling design for environmental monitoring. *Cartography and Geographic Information Systems*, 19(1):5–22, 1992.
- Y. Wiaux, J. McEwen, P. Vandergheynst, and O. Blanc. Exact reconstruction with directional wavelets on the sphere. *Monthly Notices of the Royal Astronomical Society*, 388(2):770–788, 2008.
- E. P. Wigner. *Group theory and its application to the quantum mechanics of atomic spectra*. Academic Press, 1988.
- F.-C. Yeh, V. J. Wedeen, and W.-Y. I. Tseng. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *NeuroImage*, 55(3):1054–1062, 2011.
- Z.-H. Zhan, J. Zhang, Y. Li, and Y.-H. Shi. Orthogonal learning particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, 15(6):832–847, 2011.