Wavelet-Based Reconstruction for Magnetic Resonance Imaging

Matthieu Guerquin-Kern

Thèse N° 5361 (juillet 2012)

Thèse présentée à la faculté des sciences et techniques de l'ingénieur
pour l'obtention du grade de docteur ès sciences
et acceptée sur proposition du jury

Prof. Pierre Vandergheynst, *président*
Prof. Michael Unser, *directeur de thèse*
Prof. Jalal Fadili, *rapporteur*
Prof. Klaas Pruessmann, *rapporteur*
Prof. Dimitri Van de Ville, *rapporteur*

École polytechnique fédérale de Lausanne—2012

Cover design by Annette Unser
Printing and binding by Repro-EPFL
Typeset with LaTeX

Available at **http://bigwww.epfl.ch/**

À mes parents, Agnès et Jean-Luc

Abstract

Magnetic resonance imaging (MRI) scanners produce raw measurements that are unfit to direct interpretation, unless an algorithmic step, called reconstruction, is introduced. Up to the last decade, this reconstruction was performed by algorithms of moderate complexity. This worked because substantial efforts were devoted to adjust the MRI hardware to suit the algorithmic component. More recently, new techniques have reversed this trend by putting more emphasis on the algorithms and alleviating the constraints on the hardware. Whereas many new methods suffer from a marked increase in computational complexity, this thesis focuses on the development of reconstruction algorithms that are faster and simpler than state-of-the-art solutions, while preserving their quality.

First, we present the physical principles that underlie the acquisition of MRI data and motivate the classical linear model. Based on this continuous equation, we derive efficient implementations of a discrete model. Standard and state-of-the-art reconstruction algorithms are reviewed and presented in a general framework where reconstruction is regarded as an optimization problem that can naturally integrate regularization.

Next, we propose novel simulation tools for the validation of reconstruction methods. Those tools model the sensitivity of the receiving coil, which allows for the simulation of parallel MRI experiments. To honor the continuous nature of the underlying physics, we suggest the use of analytical phantoms. Unlike rasterized simulations, our phantoms do not introduce aliasing artifacts. Instead, they allow us to study how rasterization itself impacts the quality of reconstruction. To achieve this goal, we were able to work out closed-form solutions for the Fourier transform of parametric regions that can realistically reproduce anatomical features. Our results show that the inverse-crime situation impairs significantly the assessment of the performance of reconstruction methods,

particularly, the nonlinear ones.

Finally, we investigate the design of algorithms that achieve reconstruction with a sparsity constraint expressed in a wavelet domain. Based on the latest developments in large-scale convex optimization, we derive an acceleration strategy that can be tailored to the MRI setup and provide theoretical evidence of its benefit. We develop it into a practical method that combines the advantages of speed and quality. Applied on challenging reconstruction problems, with simulated and *in-vivo* data, we significantly reduce the reconstruction time over state-of-the-art techniques without compromising quality.

Keywords: MRI, inverse problem, wavelets, sparsity, nonlinear reconstruction, undersampling, spiral, non-Cartesian, total variation, compressed sensing, iterative-shrinkage thresholding algorithm, ISTA, FISTA, FWISTA, analytical simulation, Shepp-Logan phantom, inverse crime, parallel MRI

Résumé

Les appareils d'imagerie par résonance magnétique (IRM) fournissent des données brutes dont l'interprétation nécessite un processus algorithmique nommé reconstruction. Jusqu'à la dernière décennie, cette reconstruction se faisait au moyen d'algorithmes de complexité modérée, notamment grâce à d'importants efforts dans l'optimisation de l'instrumentation. Récemment, de nouvelles techniques ont inversé cette tendance en permettant, avec un renfort de la partie algorithmique, de réduire les contraintes instrumentales. Alors que ces nouvelles méthodes souffrent d'une complexité de calcul de plus en plus élevée, cette thèse se concentre sur le développement d'algorithmes de reconstruction algorithmiquement plus efficaces.

Dans un premier temps, nous présentons les principes sous-jacents à l'acquisition des données IRM et nous justifions le modèle linéaire classique. Sur la base de cette équation continue, nous présentons des implémentations efficaces du modèle discret équivalent. Une revue des algorithmes standards et des solutions actuellement proposées est présentée dans un cadre général où la reconstruction est vue comme un problème d'optimisation pouvant naturellement inclure une régularisation.

Nous présentons ensuite de nouveaux outils de simulation destinés à la validation des méthodes de reconstruction. Ces outils, modélisant la sensibilité de l'antenne de réception, permettent de simuler des expériences d'IRM parallèle. Pour tenir compte de la nature continue de la physique, nous proposons l'utilisation de fantômes analytiques. Contrairement aux simulations discrètes, ces fantômes ne génèrent pas de repliement spectral, nous offrant ainsi la possibilité d'étudier comment ce repliement spectral affecte la qualité des reconstructions. Dans ce but, nous résolvons analytiquement la transformée de Fourier de régions paramétriques permettant de reproduire de manière réaliste des caractéristiques ana-

tomiques. Nos résultats montrent que la situation dite de *crime inverse* affecte de manière significative l'évaluation de la performance des méthodes de reconstruction ; plus particulièrement celles non linéaires.

Finalement, nous examinons la conception d'algorithmes réalisant la reconstruction avec une contrainte de parcimonie dans un domaine d'ondelettes. En nous appuyant sur les dernier développements en optimisation convexe à large échelle, nous proposons une stratégie d'accélération qui tient compte du problème d'IRM et nous apportons des preuves théoriques de son avantage. Nous développons ceci en une méthode pratique combinant les bénéfices de la rapidité et de la qualité. Appliquée sur des problèmes ambitieux, avec des données simulées et *in vivo*, notre méthode réduit significativement le temps de reconstruction par rapport à d'autres méthodes de pointe, sans compromettre la qualité de reconstruction.

Mots-clés : IRM, problème inverse, ondelettes, parcimonie, reconstruction non-linéaire, sous-échantillonnage, spirale, non-Cartésien, variation totale, acquisition comprimée, algorithme de seuillage itératif, ISTA, FISTA, FWISTA, simulation analytique, fantôme de Shepp et Logan, crime inverse, IRM parallèle

Acknowledgements

BEFORE thanking the people who I met, collaborated with and supported me throughout my PhD, there are three things I would like to apologize for.

First, I apologize for the abusive use of the pronoun "I" in these acknowledgements. A lot of people contributed to the work presented in this manuscript and I would have felt very self-centered to continue this way. So other than in the acknowledgements I used the pronoun "we". I sincerely apologize if you do not find it appropriate, but for once, at least, I get the final word.

Second, the rather unconventional way of presenting these acknowledgements. When I tried to sum up all the people I wanted to thank, I realized that I wanted to thank a lot of people for multiple reasons. After careful consideration, I ended up with the messy Table 1, which is not the appropriate format for acknowledging/thanking people. Please note that Table 1, interpreted as a punched card, is not supposed to contain any meaningful information; the contrary would be a scientific evidence of the existence of fate. I might dedicate the next years to this life altering question.

Third, I would like to express my sincere sorrow if I have forgotten to acknowledge you in Table 1. I know, it is unforgivable so I don't expect you to accept my apologies. But, don't be offended. I am very glad that you show interest in this manuscript and I thank you for that. At the end of the table there are some blank lines left, feel free to fill in one of them.

Table 1: Acknowledgements.

			Mountaineering companion[1]	Dedicated lab member[2]	Supportive family[3]	Jury member[4]	Officemate[5]	Thirsty fellow[6]	ETHZ people[7]	EPFL people[8]	Iranian dude[9]	Serbian dude[10]	Swiss dude[11]	Flatmate[12]	Student[13]
	ABRANTES	Gil						•							
Dr	AGUET	François						•		•		•			
Dr	AMINI	Arash				•					•				
	AUGIER	Nicolas								•					•
	BALAĆ	Katarina								•					
	BANDARI	Ayush								•					
Dr	BARITAUX	Jean-Charles	•					•		•				•	
Dr	BARMET	Christoph							•						
Prof	BAYRAM	Ilker				•		•		•					
Prof	BERNARD	Olivier								•					
Prof	BLU	Thierry		•						•					
	BOSTAN	Emrah								•					
	BOURQUARD	Aurélien								•	•				
	BUTENKO	Vitalijs						•		•					
	CARMO	Leticia						•							
Dr	CHAUDHURY	Kunal													
Dr	CHENOUARD	Nicolas						•		•					
	ČKONJOVIĆ	Marija						•				•			
	DALBAN	Pierre	•					•	•						
	DE FILIPPO	Rosa			•										
	DE MASI	Davide	•					•		•				•	
	DEAK	Sylvie	•					•							
Dr	DEGHANI TAFTI	Pouya								•	•				
	DELGADO-GONZALO	Ricard								•					
	DESERT SNAKE	Tarik												•	
	DIEBOLD	Marc	•							•			•		
	DIEBOLD	Yannick	•										•		
Dr	DRAGOJEVIĆ	Aleksandar						•				•			
Dr	DUCROS	Nicolas	•					•		•					
Prof	FADILI	Jalal				•									
Dr	GAJIĆ	Vojislav						•							
Dr	GUERQUIN-KERN	Jean-Luc			•										
	HÄBERLIN	Maximillian					•[14]								
Dr	HASSLER	Kai	•							•					
	KAMILOV	Ulugbek								•					
Dr	KANDASWAMY	Djano	•				•	•		•				•	
	KARAHANOGLU	Isik								•					•
Dr	KHALIDOV	Ildar								•					
Dr	KIRSHNER	Hagai								•					
Dr	KNEŽEVIĆ	Nikola						•				•			
Dr	KOLUNDŽIJA	Mikailo						•							
	KOROVNIKOV	Sergey						•							
	LAQBAYLI	Yassine						•		•				•	•
Dr	LAURENCON	Johnatan	•											•	
Dr	LEFKIMMIATIS	Stamatis								•					
	LEJEUNE	Laurent								•					•
	LENOBLE-LENOIR	Patrick			•										
	LENOIR-LENOBLE	Agnès			•										
Dr	LEVY	Sarah								•				•	
	LOPEZ	Antonio	•							•					
Dr	LUISIER	Florian				•		•					•		
Dr	MAGGIO	Simona			•[15]					•					

continued

Table 1: Acknowledgements.

			Mountaineering companion[1]	Dedicated lab member[2]	Supportive family[3]	Jury member[4]	Officemate[5]	Thirsty fellow[6]	ETHZ people[7]	EPFL people[8]	Iranian dude[9]	Serbian dude[10]	Swiss dude[11]	Flatmate[12]	Student[13]
	MAGGIO	Luigi		•											
	MARY	Manuelle	•						•			•			
Dr	MATNAZAROVA	Ludmila					•								
Dr	OLIVIERI	Chiara				•			•						
Prof	PRÜSSMANN	Klaas			•			•[16]							
Dr	PULFER	Iris										•			
Dr	RAMANI	Sathish				•			•						
Dr	SAGE	Daniel	•						•			•			
	SANYAL	Sebanti							•						
Prof	SEKHAR SEELA-MANTULA	Chandra							•						
Dr	SHILONOSOV	Vladimir					•		•				•		
	SINGH	Amardeep					•								
Dr	STANISAVLJEVIĆ	Miloš					•								
	STOJANOVIC	Tanja					•				•	•			
	STREIT	Daniela					•					•	•		
	TAIRUM	Alexandre							•						•
	TERRÈS CRISTOFANI	Raquel							•						
Dr	THÉVENAZ	Phlippe	•[17]						•			•			
Prof	UNSER	Michael		•[18]					•						
Dr	VANDENBERGHE	Charlotte							•						
Prof	VAN DE VILLE	Dimitri	•[19]	•		•			•						
Prof	VANDERGHEYNST	Pierre		•					•						
Dr	VASIĆ	Nedeljko					•				•				
Dr	VONESCH	Cédric							•						
Dr	WARD	John Paul							•						
Dr	WILM	Bertram						•				•			

[1] We went hiking/climbing/skiing together. It was fantastic.
[2] I owe her/him a lot. I could not thank him/her enough.
[3] Merci pour vos encouragements. Grazie per l'incoraggiamento.
[4] He kindly accepted to be part of my Ph.D. exam committee.
[5] It has been a pleasure to share an office with her/him.
[6] We shared a couple (or more) of beverages in a relaxed mood.
[7] During my stays in Zürich, I had the pleasure to interact with him or he kindly hosted me.
[8] I had the pleasure to know her/him during my four years stay at EPFL.
[9] سلام، خوبی؟ برو به جهنم!
[10] Ја те волим Србијо. Живели сто година.
[11] She/he is one of the few Swiss persons I happened to know during my four years stay in Switzerland.
[12] We shared a flat together and it was awesome (I forgive the coward person that stole Barnabée).
[13] She/he kindly accepted to work on a student project I proposed and I am very happy of what has been realized.
[14] I have had very fruitful discussions with Max. In particular, he helped me in the design of the MRI experiments. I owe him a lot for his expertise and the time he spent to get the scanner data.
[15] Simona now shares my life, supports and stands me on a daily basis. She is a gift of every moment. This thesis could have been dedicated to her.
[16] I would like to particularly thank Klaas who kindly hosted me in his lab at ETHZ. I admire his approach to science and research. His advices and support have been received as an honor.
[17] For his rigor, generosity, kindness, and dedication, Philippe deserves a very specific place here. If this thesis reads nicely, this is Philippe's signature. I learned a lot from his advices.
[18] He was much more than a jury member, because he is the Director of this thesis. He proposed me this project and got the funding with Klaas Prüssmann. In short, without him nothing would have been possible.
[19] Many thanks to Dimitri for his friendliness, his wise advices, his original ideas and the fruitful discussions we had.

Contents

List of notations

Acronyms

MRI	Magnetic resonance imaging
(p)MR(I)	(Parallel) magnetic resonance (imaging)
FOV	Field of view
ROI	Region of interest
S(E)(N)R	Signal-to-(error) (noise) ratio
(N)(R)MSE	(Normalized) (root-)mean-squared error
SL	Shepp-Logan
TV	Total variation
EPI	Echo planar imaging
ISTA	Iterative shrinkage/thresholding algorithm
(F)(W)(S)ISTA	(fast) (weighted) (subband adaptive) ISTA
CS	Compressed sensing
D(W)(C)(F)T	Discrete (wavelet) (cosine) (Fourier) transform
CG	Conjugate gradient
IRLS	Iteratively reweighted least squares
MSE	Mean-squared error
RS	Random shifting

Continuous Domain and Functions

$\mathbf{r}$	$\in \mathbb{R}^2$	spatial coordinates (XY plane)
$\mathbf{k}$	$\in \mathbb{R}^2$	k-space coordinates (XY plane)
$\rho(\mathbf{r})$	$\in \mathbb{R}^+$	object (proton density) in space
$m(\mathbf{k})$	$\in \mathbb{C}$	observation of the object in k-space
$\varphi(\mathbf{r})$	$\in \mathbb{R}$	generating function
$\hat{f}(\mathbf{k})$	$\in \mathbb{C}$	function f in the k-space domain
$\mathscr{C}$	$\in \mathbb{C}^{M^{\mathbb{R}}}$	cost function of a vector representing an image
$\mathrm{T}_{\boldsymbol{\tau}}$	$\mathbb{C}^M \mapsto \mathbb{C}^M$	shrinkage operator with thresholds $\boldsymbol{\tau}$
$\boldsymbol{\omega}$	$\in \mathbb{R}^2$	Fourier angular frequency

$m_S(\mathbf{k})$	$\in \mathbb{C}$	k-space observation from receiving coil S
$S(\mathbf{r})$	$\in \mathbb{C}$	spatial sensitivity of the receiving coil
$\hat{f}(\boldsymbol{\omega})$	$=$	$\int_{\mathbb{R}^d} f(\mathbf{x})\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{x}}\mathrm{d}\mathbf{x} \in \mathbb{C}$ (Fourier transform)
$\chi_{\mathscr{R}}(\mathbf{r})$	$\in \{0,1\}$	characteristic function of a region $\mathscr{R}$
$\partial\mathscr{R}$	$\subset \mathbb{R}^d$	closed contour of a region $\mathscr{R}$
J_n	$\in \mathbb{R}^{\mathbb{R}}$	n-th order Bessel function of the first kind
erf	$\in \mathbb{C}^{\mathbb{C}}$	error function of a complex argument
$\gamma(s,z)$	$\in \mathbb{C}^{\mathbb{R}\times\mathbb{C}}$	lower incomplete gamma function
Discrete Data and Linear Algebra		
j	$\in \mathbb{C}$	imaginary unit such that $\mathrm{j}^2 = -1$
$\mathbf{p}$	$\in \mathbb{Z}^2$	discrete spatial coordinates
M	$\in \mathbb{N}$	number of pixels in the ROI
N	$\in \mathbb{N}$	number of k-space samples
R	$\in \mathbb{N}$	number of receiving channels
$\mathbf{k}_n$	$\in \mathbb{R}^2$	nth k-space sampling position
m_n	$\in \mathbb{C}$	nth k-space observation
$\mathbf{m}$	$\in \mathbb{C}^N$	measurement vector
$\mathbf{b}$	$\in \mathbb{C}^N$	noise vector
$\mathbf{E}_0$	$\in \mathbb{M}_{\mathbb{C}}(N,M)$	Fourier encoding matrix (single homogeneous coil)
$\mathbf{E}$	$\in \mathbb{M}_{\mathbb{C}}(RN,M)$	SENSE encoding matrix
$\mathbf{M}$	$\in \mathbb{M}_{\mathbb{C}}(RN,M)$	system matrix (wavelet domain to k-space)
$\mathbf{W}$	$\in \mathbb{M}_{\mathbb{C}}(M,M)$	DWT matrix
$c[\mathbf{p}]$	$\in \mathbb{C}$	reconstructed spatial coefficient
$\mathbf{c}$	$\in \mathbb{C}^M$	vector of spatial coefficients
$\mathbf{w}$	$\in \mathbb{C}^M$	vector of wavelet coefficients
$\mathbf{X}^{\mathrm{H}}$	$\in \mathbb{M}_{\mathbb{C}}(N,M)$	Hermitian transpose of the matrix $\mathbf{X} \in \mathbb{M}_{\mathbb{C}}(M,N)$
$\lambda_{\max}(\mathbf{X})$	$\in \mathbb{R}^+$	largest eigenvalue of a symmetric matrix $\mathbf{X}$
$\kappa(\mathbf{X})$	$\in [1,+\infty[$	ℓ_2-condition number of a matrix $\mathbf{X}$
$\langle \mathbf{x}, \mathbf{y} \rangle$	$\in \mathbb{C}$	regular inner product
$\mathbf{x}\cdot\mathbf{y}$	$\in \mathbb{R}$	regular inner product
$\langle \mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{\Lambda}}$	$\in \mathbb{C}$	weighted inner product ($\mathbf{x}^{\mathrm{H}}\boldsymbol{\Lambda}\mathbf{y}$)
$\|\mathbf{x}\|_{\ell_2}$	$\in \mathbb{R}^+$	regular quadratic norm
$\|\mathbf{x}\|_{\boldsymbol{\Lambda}}$	$\in \mathbb{R}^+$	weighted quadratic norm ($\sqrt{\mathbf{x}^{\mathrm{H}}\boldsymbol{\Lambda}\mathbf{x}}$)
$\|\mathbf{x}\|_{\ell_1}$	$\in \mathbb{R}^+$	ℓ_1 norm ($\sum \lvert x_i \rvert$)
$\mathbf{e}_i$	$\in \mathbb{R}^d$	the canonical vectors such that $\mathbf{x}\cdot\mathbf{e}_i = x_i$
$\delta_{k,l}$	$\mathbb{N}^2 \mapsto \{0,1\}$	Kronecker's delta (1 if $k = l$ and 0 otherwise)
Operators on Vectors		
$\mathbf{a}\times\mathbf{b}$	$\in \mathbb{R}^3$	vector product for $\mathbf{a},\mathbf{b} \in \mathbb{R}^3$

$\boldsymbol{\nabla} f$	$\mathbb{R}^d \mapsto \mathbb{R}^d$	gradient operator for $f : \mathbb{R}^d \mapsto \mathbb{R}$
$\boldsymbol{\nabla} \cdot \mathbf{f}$	$\mathbb{R}^d \mapsto \mathbb{R}$	divergence operator for $\mathbf{f} : \mathbb{R}^d \mapsto \mathbb{R}^d$
$\boldsymbol{\nabla} \times \mathbf{f}$	$\mathbb{R}^3 \mapsto \mathbb{R}^3$	curl operator for $\mathbf{f} : \mathbb{R}^f \mapsto \mathbb{R}^3$

Multi-Index Notations

$\mathbf{z}^{\boldsymbol{\alpha}}$	$=$	$\prod z_i^{\alpha_i} \quad \in \mathbb{R} \quad$ for $\mathbf{z} \in \mathbb{R}^d$ and $\boldsymbol{\alpha} \in \mathbb{N}^d$
$\lvert\boldsymbol{\alpha}\rvert$	$=$	$\sum \alpha_i \quad \in \mathbb{N} \quad$ for $\boldsymbol{\alpha} \in \mathbb{N}^d$
$\mathbf{p}!$	$=$	$\prod p_i ! \quad \in \mathbb{N} \quad$ for $\mathbf{p} \in \mathbb{N}^d$
$\binom{\mathbf{p}}{\mathbf{q}}$	$=$	$\prod \binom{p_i}{q_i} = \frac{\mathbf{p}!}{(\mathbf{p}-\mathbf{q})!\mathbf{q}!} \quad \in \mathbb{N} \quad$ for $\mathbf{p}, \mathbf{q} \in \mathbb{N}^d$
$\sum_{\mathbf{p}=\mathbf{a}}^{\mathbf{b}}$	$=$	$\sum_{p_1=a_1}^{b_1} \sum_{p_2=a_2}^{b_2} \cdots \sum_{p_d=a_d}^{b_d}$ for $\mathbf{p}, \mathbf{a}, \mathbf{b} \in \mathbb{N}^d$

Chapter 1

Introduction

MAGNETIC resonance imaging (MRI) is a non-invasive imaging technique that dates back to the 70s [1]. It is now commonly used in medicine and biology because of its specific advantages over other imaging modalities: versatile contrasts, speed, good spatial resolution, multidimensional imaging capabilities, and structural as well as functional information. Still, MRI is subject of active research. While, in the past, efforts were mainly concentrated towards the improvement of the acquisition hardware, the focus is now shifting towards the use of more sophisticated post-acquisition processes. The idea is that, by relying on proper signal processing, one can relax the acquisition requirements. The prospects are of different kinds:

– Fast acquisition is a key point to increase resolution in MRI. It can be achieved by reducing the number of acquired samples. This complicates the reconstruction task.

– Some scanner defects can be compensated provided that the distortions are accurately modeled. An improved reconstruction process would then allow the use of lower-quality, and potentially cheaper hardware.

– In a clinical perspective, one would like to reduce acquisition time and use smoother acquisition schemes—to lower the acoustic noise—in order improve the comfort of the patients.

Some recent techniques are indicative of this trend: off-resonance correction [2], parallel MRI [3], motion correction [4, 5], compressed-sensing [6], and higher-order field imaging [7].

This thesis originates from a collaboration between the Biomedical Imaging Group of EPFL and the Institute for Biomedical Engineering of the University and ETH in Zürich. The aim is to develop reconstruction methods specific to magnetic resonance imaging relying upon the hypothesis that nonlinear reconstruction—for instance, wavelet regularization—are capable of enhancing image quality in challenging reconstruction tasks. Following the same research trend, several recent works [6, 8–16] confirm the potential benefits of nonlinear reconstruction for MRI. Another idea of this project is that significant acceleration of the reconstruction process can be achieved if the structure of the physical model is exploited. Surprisingly, this aspect is not yet subject to much investigation in the MRI community.

In signal-processing terms, the reconstruction problem can be stated as follows. We must estimate a discrete image that represents an existing continuous function. The MRI scanner samples the Fourier domain of this function. Medical and instrumental constraints influence the design of the trajectory in the Fourier domain. The resulting sampling locations are not uniformly distributed and this situation yields undesirable artifacts in conventional image reconstructions. We consider this kind of challenging reconstruction task and regard it as a linear inverse problem. Two characteristics must be considered:

– The problem is *large-scale* as it involves a considerable amounts of data and unknowns. Efficient algorithms are required to apply the data-formation model and the reconstruction process is performed iteratively.

– Due to the presence of noise and a potential lack of data, the problem is *ill-conditioned*. This is to say that many, possibly very different, images can explain the measured data. This is where nonlinear reconstruction schemes come into play for their ability to favor certain classes of solutions.

The downside is that iterative algorithms are much slower than traditional FFT-based reconstruction. While this is already true in the linear case, where convergence speed directly depends on the conditioning, current nonlinear reconstruction algorithms are even slower. In this context, our goal is to develop nonlinear reconstruction algorithms that enhance quality over standard reconstructions while being practically fast enough to compete with linear reconstruction.

1.1 Main Contributions

During our in-depth investigation of the MRI reconstruction pipeline, the problem of the proper assessment of the quality of reconstructions was pervasive. Simulations facilitate the evaluation and comparison of reconstruction software because the experimental conditions are perfectly controlled and ground-truth data are well-known. As first contribution, we introduce a new theoretical framework that allows for the analytical description of realistic phantoms in both the spatial and Fourier domains. This solution shall be preferred to standard discrete simulations that introduce aliasing artifacts impairing the evaluation of the reconstruction quality. Some analytical phantoms are already available for MRI simulation but, contrarily to the brain phantom that we propose, they suffer from a lack of realism in describing anatomical regions and are limited to the simulation of MRI with a homogeneous receiving antenna. The original framework we present is general enough to simulate parallel MRI experiments and the parametric shapes that describe our phantom are very flexible. We propose experiments that validate our method and illustrate the impact of using discrete simulations in the assessment of the performance of several reconstruction techniques. It turns out that the performance bias can be very pronounced, particularly with state-of-the-art reconstruction techniques. This emphasizes the importance of our contribution.

A second innovative aspect of this thesis consists in a new reconstruction method that imposes a wavelet-domain regularization. The starting point for our work is ISTA [17–19], a simple and robust iterative algorithm that achieves sparsity-promoting reconstruction. Recent developments have shown that this algorithm can be substantially accelerated using a multistep method [20] (fast ISTA, aka FISTA). The convergence rate of both ISTA and FISTA have been derived. An alternative acceleration strategy, namely SISTA [21], was proposed that adjusts the operations depending on the wavelet subband in order to accelerate reconstruction. We revisit all these algorithms using weighted norms. This perspective gives birth to a new algorithm that combines the FISTA and SISTA acceleration strategies synergistically. We derive the convergence bounds for the proposed algorithm, showing its theoretical advantage in terms of speed. Random shifting [17] (RS) is a modification in wavelet-based algorithms that brings sensible image quality enhancements but lacked a theoretical

foundation. We provide a new interpretation of RS as a greedy and practically efficient method to perform shift-invariant wavelet regularization. We come up with an algorithm for MRI that benefits from the advantages of all these techniques: convergence efficiency and reconstruction quality. Based on realistic MRI and parallel MRI simulations, we study the influence of the wavelet basis and the use of RS on reconstruction quality. We also verify the convergence superiority of our new algorithm. Finally, compared with other state-of-the-art reconstruction techniques on both simulated and *in vivo* data, our practical method proves to converge rapidly while achieving competitive reconstruction quality.

1.2 Organization of the Thesis

This thesis is organized as follows: In Chapter 2, we expose the principles of magnetic resonance imaging. We follow a conceptual path from the spin magnetic moment to the production of an image. The origin of scanner data is presented and the mathematical data-formation model is established. In Chapter 3, we describe the discrete model together with efficient implementations. Classical and state-of-the-art reconstruction techniques for MRI are presented within a general framework for reconstruction. Then, in Chapter 4, we present a new analytical simulation tool that is adapted to parallel MRI and addresses the issues occurring with standard rasterized simulations. The theory is exposed and experiments show the interest of the tool. Finally, in Chapter 5, we discuss an algorithmic strategy to perform competitive nonlinear reconstructions using wavelet regularization. The performance of the method is studied and its potential is demonstrated in several challenging MRI and parallel MRI experiments.

Chapter 2

MRI Principles

In this chapter, we present the basics of magnetic resonance imaging. Advanced concepts are covered by Haacke *et al.* [22] as well as Liang and Lauterbur [23].

This chapter is structured as follows: In Section 2.1, we present the physical principles of magnetic resonance. Then, in Section 2.2, we show how the signal in MRI is observed using receiving antennas and exploiting the magnetic resonance effect. In Section 2.3, we explain the special approach to imaging that is adopted in MRI, which relies on gradients of magnetic field. Finally, in Section 2.4, we present the most common MR imaging modalities with a particular focus on parallel MRI.

2.1 Nuclear Magnetic Resonance

2.1.1 Spins and Magnetization

Quantum physics [24] postulate the existence of a spin quantity associated to nucleon particles. Depending on the spin value, the particle may have a magnetic moment. It turns out that the proton presents such a magnetic moment. The proton is also a hydrogen nucleus, a compound which is present in water molecules. Its density in biological tissues makes it of major interest for biomedical imaging.

In the absence of an external magnetic field, a set of nuclei has no bulk (that is to say, macroscopic) magnetization because the spin magnetic moments have independent and randomly distributed directions. In pres-

ence of a static magnetic field $\mathbf{B}_0$, the magnetic moments tend to align in the direction of the field, which decreases the energy of the system. This effect is counteracted by the thermal energy, which ensures that some nucleons retain anti-parallel spins. Boltzmann statistics rules the distribution of these populations at thermal equilibrium as

$$N_\uparrow = N_\downarrow \exp\left(\gamma\hbar B_0/(k_B T)\right), \tag{2.1}$$

where $N_\uparrow$ and $N_\downarrow$ are the numbers of parallel and anti-parallel spins in a given volume, γ is their gyromagnetic ratio, $\hbar$ is Planck's constant, k_B is Boltzmann's constant, and T is the temperature.

In practice, the ratio of the populations is accurately described by a first-order approximation. [1] Consequently, the bulk magnetization $\mathbf{M}$, defined as the volume density of the spin magnetic moments, is aligned and proportional to the magnetic field. Its intensity characterizes the imbalance between the two spin populations.

2.1.2 Motion Law for the Magnetization

The kinetic moment resulting from the magnetization is $\mathbf{M}/\gamma$. In presence of an external field $\mathbf{B}$, the torque generated on the magnetization writes $\mathbf{M} \times \mathbf{B}$. This leads to the motion law

$$\mathrm{d}\mathbf{M}/\mathrm{d}t = \gamma \mathbf{M} \times \mathbf{B}. \tag{2.2}$$

This relation implies that $\mathrm{d}\mathbf{M}/\mathrm{d}t$ is perpendicular to $\mathbf{M}$. In turn, the magnetization intensity $\|\mathbf{M}\|$ is conserved.

Let us denote by $\mathbf{B}_0$ a static magnetic field. We define its direction to be the longitudinal axis. The plane that is perpendicular is referred to as the transverse plane. In presence of such a static field, the angle formed between $\mathbf{M}$ and the longitudinal axis is conserved. Specifically, $\mathbf{M}$ precesses clockwise around $\mathbf{B}_0$ at the angular frequency $\omega_0 = \gamma \left\|\mathbf{B}_0\right\|$, known as the Larmor frequency. We illustrate in Figure 2.1 the precession of the magnetization vector in a static field.

At thermal equilibrium, the magnetization is perfectly aligned with the static magnetic field, because the phases of the spin magnetic moments

1. At temperature $T = 300\,\mathrm{K}$, with a magnetic field $B_0 = 1\,\mathrm{T}$ and for protons of gyromagnetic ratio $\gamma = 2.678 \times 10^8\,\mathrm{rad\,T^{-1}\,s^{-1}}$, the ratio $\gamma\hbar B_0/(k_B T)$ is equal to 6.8×10^{-6}.

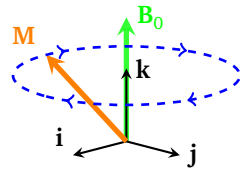

Figure 2.1: In presence of a static magnetic field, the spin magnetic moments and the resulting bulk magnetization precesses at the Larmor frequency.

are uncorrelated. [2] The detection of the magnetization is difficult in this situation.

2.1.3 Excitation Pulse and Resonance

The nuclear magnetic resonance (NMR) is a phenomenon that makes the magnetization measurable. It was first observed by Rabi [25] for a beam of atoms. The absorption of radio frequency waves by matter was characterized few years later by Purcell and Bloch [26, 27].

To provoke NMR, a pulse of oscillating electromagnetic field is applied transversally. Depending on its angular frequency ω, it excites the spin magnetic moments. Let us consider the magnetic field component $\mathbf{B}_1$ of this pulse. To facilitate the analysis, we define a frame of reference $(O, \mathbf{I}, \mathbf{J}, \mathbf{K})$ that rotates at the same angular frequency ω around the longitudinal axis $\mathbf{B}_0$. In this rotating frame, the magnetic fields $\mathbf{B}_0$ and $\mathbf{B}_1$ appear static at point O. The frames of reference are depicted in Figure 2.2.

In the rotating frame of reference, the motion law for the magnetization (2.2) rewrites

$$\left.\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t}\right|_{(O,\mathbf{I},\mathbf{J},\mathbf{K})} = \gamma \mathbf{M} \times \mathbf{B}_{\text{eff}}, \tag{2.3}$$

with $\gamma \mathbf{B}_{\text{eff}} = \boldsymbol{\Delta\omega} - \boldsymbol{\omega}_1$, $\boldsymbol{\Delta\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_0$, $\boldsymbol{\omega} = \omega \mathbf{k}$, $\boldsymbol{\omega}_0 = -\gamma \mathbf{B}_0$, and $\boldsymbol{\omega}_1 = -\gamma \mathbf{B}_1$. The magnetization vector precesses around the effective magnetic field $\mathbf{B}_{\text{eff}}$ at the angular frequency $\gamma \mathbf{B}_{\text{eff}}$. This movement takes

2. A consequence of local variations in the magnetic field and spin-spin interactions that we further discuss in Section 2.1.3.

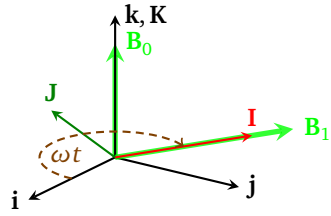

Figure 2.2: The rotating reference frame shares the longitudinal axis with the laboratory reference frame and rotates at the frequency of the excitation pulse.

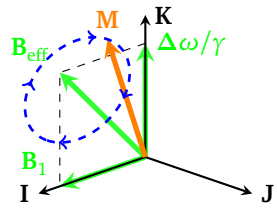

Figure 2.3: In the rotating frame, the magnetization precesses around the effective magnetic field.

place in the rotating frame and is illustrated in Figure 2.3. In the laboratory reference frame, this movement is composed with a rotation around the longitudinal axis at frequency ω.

Depending on the frequency of the excitation pulse, three situations can occur:

- $\|\boldsymbol{\Delta\omega}\| \gg \left\|\boldsymbol{\omega}_1\right\|$, which means that $\mathbf{B}_{\text{eff}}$ is almost aligned with $\mathbf{B}_0$. The bulk magnetization is precessing at angular frequency $\gamma \left\|\mathbf{B}_0\right\|$ around $\mathbf{k}$, as it does at thermal equilibrium. The impact of the excitation pulse on magnetization is insignificant.
- $\|\boldsymbol{\Delta\omega}\| \approx \left\|\boldsymbol{\omega}_1\right\|$. The effective field affecting the magnetization is in the plane $(O, \mathbf{I}, \mathbf{k})$ between vectors $\mathbf{k}$ and $\mathbf{I}$ (see Figure 2.3). The magnetization is affected by the excitation pulse but the resonance is not total. The magnetization never flips completely.
- $\|\boldsymbol{\Delta\omega}\| \ll \left\|\boldsymbol{\omega}_1\right\|$, which corresponds to nuclear magnetic resonance.

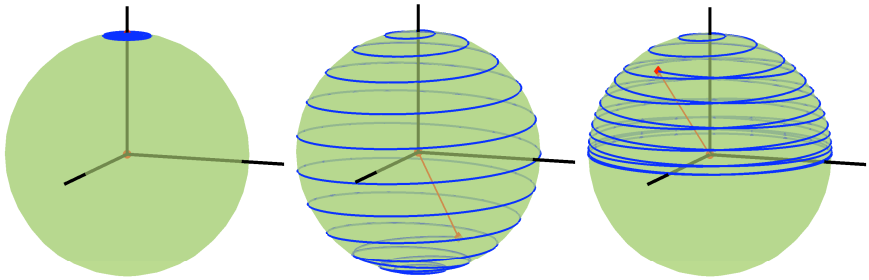

Figure 2.4: Motion of the magnetization as observed in the laboratory frame of reference in response to excitation pulses: From left to right: $\|\boldsymbol{\Delta\omega}\| = 10\left\|\boldsymbol{\omega}_1\right\|$, $\|\boldsymbol{\Delta\omega}\| = \left\|\boldsymbol{\omega}_1\right\|/10$, and $\|\boldsymbol{\Delta\omega}\| = \left\|\boldsymbol{\omega}_1\right\|$.

In this case, $\mathbf{B}_{\text{eff}} \approx \mathbf{B}_1$ and $\omega \approx \omega_0$. In the rotating frame, the bulk magnetization is precessing around $\mathbf{B}_1$ at angular frequency $\gamma\left\|\mathbf{B}_1\right\|$.

These three cases are illustrated in Figure 2.4.

2.1.4 Relaxation

The Bloch equation (2.2) imposes:

- A constant projection of the magnetization onto the longitudinal axis. In other words, the energy coming from the excitation pulse that resulted in a modified balance in the spins populations must be perpetually conserved by the system of spins.

- A constant intensity of magnetization in the transverse plane. This implies that the precessing magnetic moments of the spin remain synchronized in phase so as to maintain the transverse magnetization.

Contrarily to these predictions, the magnetization vector is observed to return to the thermal equilibrium state $\mathbf{M}_0 = M_0\mathbf{k}$, where the energy of the system of spins is minimal. This relaxation is exponential with two characteristic times: T_1 for the longitudinal magnetization and T_2 for the transverse component. In the reference frame rotating at angular frequency γB_0, the relaxation of the magnetization is modeled by the

phenomenological Bloch equations

$$\left.\frac{d\mathbf{M}_{\perp}}{dt}\right|_{(O,\mathbf{I},\mathbf{J})} = -\frac{\mathbf{M}_{\perp}}{T_2}, \tag{2.4}$$

$$\frac{d\mathbf{M}_{//}}{dt} = \frac{\mathbf{M}_0 - \mathbf{M}_{//}}{T_1}, \tag{2.5}$$

where $\mathbf{X}_{\perp}$ denotes the projection of $\mathbf{X}$ on the longitudinal axis and $\mathbf{X}_{//}$ denotes its projection on the transverse plane.

The relaxation times T_1 and T_2 have different origins.

– The longitudinal relaxation time T_1 describes the rate of the energy transfer from the spins to the environment. This transfer is mostly explained by the emission of the electromagnetic signal that might be captured by MRI receiving coils. As far as spins are concerned, this energy loss corresponds to a transfer of population from anti-parallel spins (higher energy) to parallel spins. The relaxation time T_1 can be interpreted as the average lifetime of anti-parallel spins.

– The transverse relaxation time T_2 characterizes the rate of desynchronization of the magnetic moments that takes place at the microscopic scale where the magnetic moments randomly influence each other. A second phenomenon causes dephasing. At the macroscopic scale, the B_0 field inhomogeneities provoke slight local changes in the precession frequency, leading to a transverse magnetization loss as the moments get out of phase. This trend can be reverted using a 180° pulse (see Section 2.1.5 and Figure 2.9 for further details). The combination of the two dephasing factors is characterized by the T_2^* transverse relaxation time directly observed in the magnetization decay.

In Table 2.1 we report the order of magnitude of T_1 and T_2 for various types of biological tissues. They are of particular importance in medicine; Damanian showed early on that their value can discriminate malignant from benign tumors [28].

During the excitation, relaxation times can be neglected as they are many times longer than the duration of the excitation pulse. [3]

3. For a 90° flip of proton magnetic moments with excitation field intensity $B_1 = 10^{-5}$ T at temperature 300 K, the required duration is about 590 µs, which is at least two orders of magnitude shorter than the characteristic times reported in Table 2.1.

Table 2.1: Typical characteristics of human brain tissues for $B_0 = 1.5\,\mathrm{T}$: Relative spin density, T_1, and T_2 relaxation times (according to the values gathered in [29]). CSF, GM, and WM stand for cerebrospinal fluid, gray matter, and white matter, respectively.

Tissue	CSF	GM	WM	Scalp	Marrow
ρ/ρ_{water}	0.98	0.745	0.617	0.8	0.12
T_1 (in ms)	4200	1000	680	340	550
T_2 (in ms)	2000	100	80	70	50

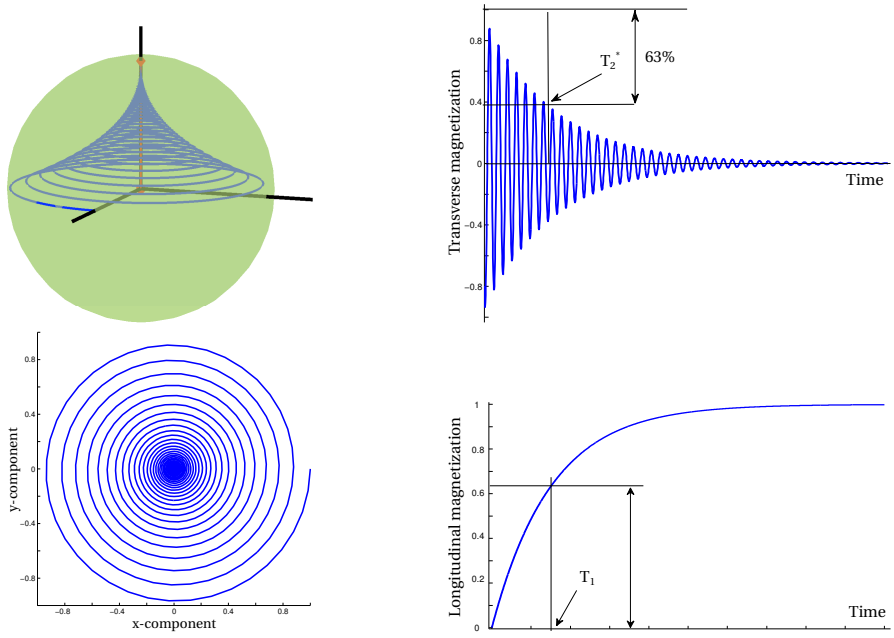

Figure 2.5: The motion of magnetization during a Free Induction Decay.

The transverse signal, which is measured, is called Free Induction Decay (FID). It carries the information on the tissue in the intensity of its envelope and the relaxation times. We show in Figure 2.5 an example of FID.

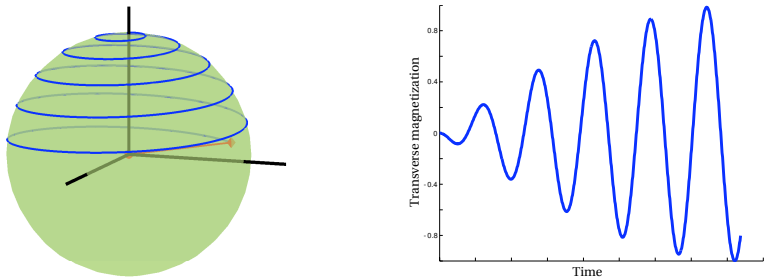

Figure 2.6: The motion of magnetization in response to a 90° excitation pulse.

2.1.5 Pulses

If the excitation pulse oscillates at the Larmor frequency γB_0 for a duration T_{pulse}, the magnetization gets flipped by an angle $\theta = \omega_1 T_{\text{pulse}}$. An excitation pulse of duration $T_{90^\circ} = \pi/(2\gamma B_1)$ moves the magnetization from the longitudinal axis to the transverse plane where it can be detected with maximal intensity. The effect of this pulse is illustrated in Figure 2.6.

During the relaxation, pulses that provoke a 180° flip (twice longer than the 90° ones) can be applied. The interest is to recover the loss of transverse signal caused by the static field inhomogeneities. The duration $T_E/2$ (for "Time Echo") is the time interval between the 90° and the 180° pulse. An echo in the FID is observed at time $T_E/2$ after the 180° pulse, as the spin phases get resynchronized. The signal loss that is due to the spin-to-spin interactions is not recovered, due to its stochastic nature. If several 180° pulses are performed, the characteristic time T_2 characterizes the decay in the intensity of the echoes as a function of T_E. The effect of a 180° pulse is shown in the laboratory reference frame in Figure 2.7. The production of a echo by means of a 180° pulse is illustrated in Figure 2.9.

2.2 Detection

In this section, we adopt a signal-processing perspective to analyze the signals provided by MRI scanners. The physical relationship with the magnetization is shown. As a reference, we suggest the book by Haacke [22].

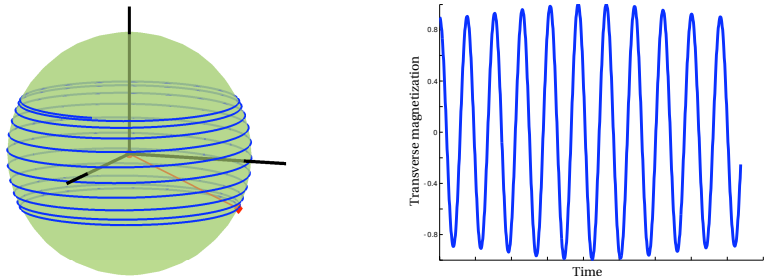

Figure 2.7: The motion of magnetization in response to a 180° excitation pulse.

2.2.1 Original Signal

The information of interest is the space distribution of the spin density. The magnetization vector $\mathbf{M}(\mathbf{r}, t)$ resulting from an excitation pulse is proportional to this quantity (see its definition in Section 2.1.1). We neglect the decay phenomena because their influence on signal detection is negligible. We model the transverse magnetization by a complex quantity $\underline{M}(\mathbf{r}, t) = M(\mathbf{r}) \exp\left(-\mathrm{j}\left(\omega_0 t + \theta(\mathbf{r}, t)\right)\right)$, where $M(\mathbf{r})$ is the initial magnetization (proportional to the spin density), $\omega_0 = \gamma B_0$, and $\theta(\mathbf{r}, t)$ accounts for a non-uniform and time-varying phase map of magnetization vectors.

2.2.2 Physics: Coil Sensitivity

The magnetization state is detected by a receiving antenna or coil. The principle is that the precessing magnetization generates a time-varying magnetic field that, in turn, induces an electromotive force in the coil.

First, the magnetization distribution admits an equivalent current distribution $\mathbf{J}_\mathbf{M}(\mathbf{r}, t) = \boldsymbol{\nabla} \times \mathbf{M}(\mathbf{r}, t)$. By choosing the Coulomb gauge [4] and neglecting propagation times [5], the magnetic vector potential is given by

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} \frac{\mathbf{J}_\mathbf{M}(\mathbf{r}', t)}{\|\mathbf{r} - \mathbf{r}'\|} \mathrm{d}\mathbf{r}'. \tag{2.6}$$

4. The Coulomb gauge is defined such that $\boldsymbol{\nabla} \cdot \mathbf{A} = 0$.
5. The propagation speed is close to the speed of light in vacuum.

The magnetic field is then expressed as $\mathbf{B}(\mathbf{r},t) = \boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r},t)$. Using Stokes' theorem, the flux of magnetic field through the coil surface $\mathscr{C}$ writes

$$\Phi(t) = \oint_{\partial\mathscr{C}} \mathbf{A}(\mathbf{r},t)\cdot \mathrm{d}\mathbf{r}. \tag{2.7}$$

Using the relation $\boldsymbol{\nabla}\times(a\mathbf{v}) = (\boldsymbol{\nabla}a)\times\mathbf{v} + a\boldsymbol{\nabla}\times\mathbf{v}$, and the triple product rules, we expand the flux expression into

$$\Phi(t) = \int_{\mathbb{R}^3} \mathbf{M}(\mathbf{r}',t)\cdot\mathbf{B}^u(\mathbf{r}')\mathrm{d}\mathbf{r}', \tag{2.8}$$

with

$$\mathbf{B}^u(\mathbf{r}') = \frac{\mu_0}{4\pi}\oint_{\partial\mathscr{C}} \frac{\mathrm{d}\mathbf{r}\times(\mathbf{r}-\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|^3}. \tag{2.9}$$

Note that (2.9) is the Biot-Savart law for a magnetic field generated at point $\mathbf{r}'$ by a unit-value steady current in the coil. This result is consistent with the principle of reciprocity.

The electromotive force induced in the coil is defined as $e(t) = -\mathrm{d}\Phi(t)/\mathrm{d}t$. Since the decay of the longitudinal magnetization is neglected, its expression simplifies to

$$e(t) = \mathrm{Re}\left(-\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3} \underline{M}(\mathbf{r},t)S(\mathbf{r})\mathrm{d}\mathbf{r}\right), \tag{2.10}$$

with the coil sensitivity map defined as $S(\mathbf{r}) = B_x^u(\mathbf{r}) - \mathrm{j}B_y^u(\mathbf{r})$.

By definition of $\underline{M}$, thanks to Fubini's theorem, and assuming that $|\mathrm{d}\theta/\mathrm{d}t| \ll \omega_0$, the electromotive force rewrites [6]

$$e(t) = -\omega_0\,\mathrm{Im}\left(\mathrm{e}^{-\mathrm{j}\omega_0 t}\int_{\mathbb{R}^3} S(\mathbf{r})M(\mathbf{r})\mathrm{e}^{-\mathrm{j}\theta(\mathbf{r},t)}\mathrm{d}\mathbf{r}\right). \tag{2.11}$$

2.2.3 Demodulation

The spectrum of the electromotive force induced in the coil is concentrated around the angular frequency ω_0. For further processing, it is

6. For protons in a field $B_0 = 1\,\mathrm{T}$, the Larmor frequency is $f_0 = 42.6\,\mathrm{MHz}$ and the bandwidth in MRI is typically in the range of 1 kHz.

convenient to shift this spectrum to the low frequencies. This operation is realized through phase-sensitive demodulation, as done in telecommunications with a Quadrature Amplitude Modulation (QAM) signal. First, the electromotive force is multiplied by two sines at frequency ω_0 which are in quadrature. This operation transposes the spectrum around the frequencies $\omega = 0$ and $\omega = 2\omega_0$. Second, a low-pass filter is used to attenuate the spectrum at $2\omega_0$. This type of demodulation returns two signals that are the real and imaginary parts of the MR scanner signal $m(t)$ which is complex-valued. Finally, the MR scanner signal is related to the magnetization through the linear integral

$$m(t) = \int_{\mathbb{R}^3} S(\mathbf{r}) M(\mathbf{r}) \mathrm{e}^{-\mathrm{j}\theta(\mathbf{r},t)} \mathrm{d}\mathbf{r}. \tag{2.12}$$

2.3 Gradients of Field for Imaging

The concept of phase encoding, introduced by Lauterbur [1], allows one to reconstruct spatial maps from the MR measurements. The idea is to exploit field gradients in order to express (2.12) as a Fourier transform. After sampling enough frequencies, the image is formed using an inverse discrete Fourier transform. In this section, the signal of interest is the spin density $\rho(\mathbf{r})$, rather than the magnetization. For simplicity, the receiving coil sensitivity is assumed to be homogeneous.

2.3.1 Spin Density Profile

We consider the case where the physical quantities only depend upon the transverse coordinate x, such that $\rho(\mathbf{r}) = \rho(x)$ and $\theta(\mathbf{r}, t) = \theta(x, t)$. According to (2.12), the MR scanner signal is

$$m(t) = \int \rho(x) \mathrm{e}^{-\mathrm{j}\theta(x,t)} \mathrm{d}x.$$

By superposing to B_0 a field with a gradient G that is constant in both space and time, the phase term becomes $\theta(x, t) = \gamma x G t$. In turn, the MR scanner signal writes

$$m(t) = \hat{\rho}(\gamma G t), \tag{2.13}$$

where $\hat{\rho}(\omega) = \int \rho(x) \mathrm{e}^{-\mathrm{j}\omega x} \mathrm{d}x$ stands for the Fourier transform of ρ. As a consequence, the spectrum of ρ is linearly scanned as time goes by.

It is reasonable to consider that the density of excited spins has a finite spatial support. Then, according to Shannon's sampling theorem applied in the k-space domain, the density of excited spins ρ can be exactly reconstructed from the samples of its spectrum if the k-space sampling frequency is higher than the width of the spatial support.

2.3.2 Phase Encoding for Imaging

The concept of phase encoding generalizes the idea of 1-D k-space scanning by the use of linear gradients along multiple directions.

A spatially-linear field component is added to the main static field B_0. The corresponding spatial gradient, which may vary over time, is noted $\mathbf{G}(t)$. The phase of the magnetization in the rotating frame of reference evolves accordingly as

$$\theta(\mathbf{r}, t) = \gamma \int_0^t \mathbf{G}(\tau) \cdot \mathbf{r} \mathrm{d}\tau. \tag{2.14}$$

It is customary in MRI to refer to the spatial frequency domain as the k-space. The gradient trajectory determines the k-space position through the relation

$$\mathbf{k}(t) = \frac{\gamma}{2\pi} \int_0^t \mathbf{G}(\tau) \mathrm{d}\tau, \tag{2.15}$$

and the corresponding data is given by

$$m(\mathbf{k}) = \int \rho(\mathbf{r}) \mathrm{e}^{-2\pi \mathrm{j} \mathbf{k} \cdot \mathbf{r}} \mathrm{d}\mathbf{r} = \hat{\rho}(2\pi \mathbf{k}), \tag{2.16}$$

where $\hat{\rho}$ stands for the multidimensional Fourier transform.

2.3.3 Slice Selection

When performing 2-D MRI, it is usual to let resonate only the spins located in a thin slice along the longitudinal axis. This selection is made possible by the superposition of a field gradient to the homogeneous longitudinal field. Thus, the static field intensity $B_0 + G_z(z - z_0)$ varies with the position. Consequently, an excitation pulse of bandwidth W centered around the frequency ω excites only the spins located in the slice of longitudinal coordinate $z = z_0 + (\omega/\gamma - B_0)/G_z$ with a thickness $\Delta z = W/(\gamma G_z)$. The relationships between the parameters involved in slice selection are illustrated in Figure 2.8.

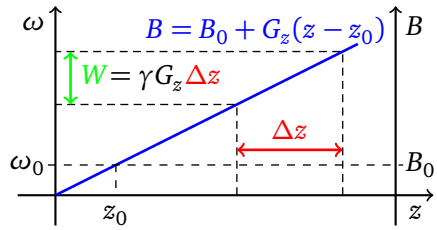

Figure 2.8: Illustration of the slice selection when applying an excitation pulse.

2.3.4 Sampling Strategies

According to (2.15), a displacement $\mathbf{\Delta}k$ in k-space can be obtained by maintaining a constant gradient of intensity G oriented like $\mathbf{\Delta}k$, during the time $\gamma G/(2\pi)$.

Several approaches to scan the k-space have been investigated. They have different advantages in terms of versatility, scan time, constraints on gradient switching, and robustness to reconstruction artifacts.

- **Cartesian trajectories** have been the first to be investigated. When the k-space is fully sampled that way, the reconstruction scheme comes trivially as an inverse discrete Fourier transform. This scheme leads to rapid acquisitions but puts hard constraints on the magnetic field gradients.
- **Radial trajectories** naturally favor a denser sampling in the center of the k-space where most of the information is concentrated. Thanks to the Fourier slice theorem, which relates a profile in k-space to a projection onto a line in space, this acquisition scheme benefits from the advanced reconstruction algorithms developed for tomography.
- **Spiral trajectories** can offer smooth gradient switching and are quite robust to off-resonance artifacts (inhomogeneities in the main field). Moreover, their design is versatile and they can be made to result in a reduced total scan time.

2.4 Imaging Modalities

In this section, we present the basic MRI contrasts that are used clinically. Next, we introduce parallel MRI, which is an advanced MRI technique

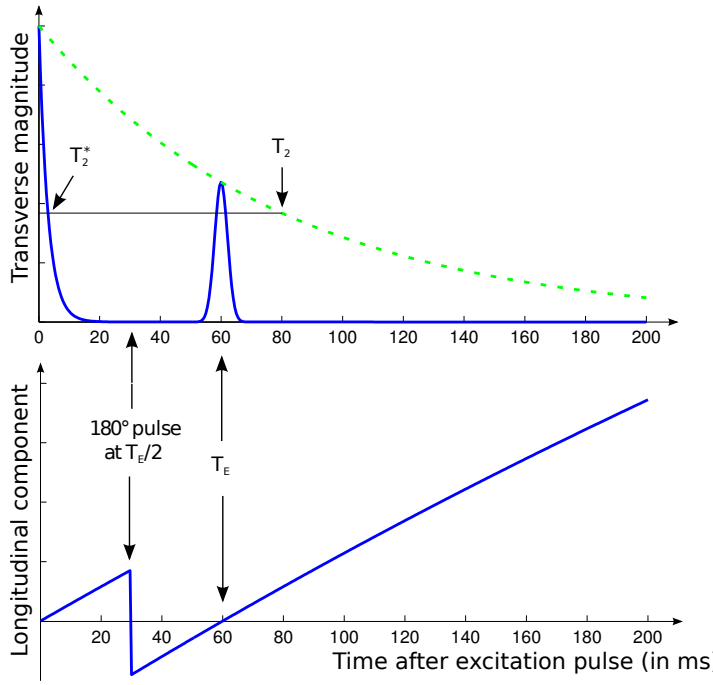

Figure 2.9: Time evolution of the magnetization after excitation and during the spin-echo provoked by a 180° pulse.

that we discuss in more details in the following chapters. It is beyond the scope of this thesis to present other advanced techniques such as diffusion MRI, functional MRI, real-time MRI, magnetization-transfer MRI, and compensation methods—*e.g.*, flow, motion, and field inhomogeneity.

2.4.1 Basic Contrasts

Classical MRI sequences offer several degrees of freedom one can play with to influence the contrast. The two main families of sequences, Gradient-Echo and Spin-Echo, aim at producing echoes in the FID signal. The echo is provoked in a way that is specific to each sequence. At time T_E after the excitation pulse, which can be controlled, the echo occurs. We illustrate the behavior of magnetization during a spin echo in Figure 2.9.

Other degrees of freedom include the repetition time T_R, which is the

Table 2.2: Basic MR contrasts and corresponding choices of T_E and T_R for Spin-Echo sequences.

Contrast	Spin density	T_1	T_2
T_R	$\gg T_1$	$\approx T_1$	$\gg T_1$
T_E	$\ll T_2$	$\ll T_2$	$\approx T_2$

period between two excitation pulses, and the flip angle provoked by the excitation.

For a tissue with spin density ρ and characteristic times T_1, T_2, and T_2^*, it is well-known (see for instance [23]) that the amplitude of the echo and, consequently, the signal being imaged, are proportional to the quantity

$$\rho\left(1-\exp(-T_R/T_1)\right). \tag{2.17}$$

In addition, the signal in a Spin-Echo sequence is weighted by

$$\exp(-T_E/T_2) \tag{2.18}$$

whereas, in a Gradient-Echo sequence using excitations pulses of flip angle α, it is weighted by

$$\frac{\exp(-T_E/T_2^*)\sin(\alpha)}{1-\cos(\alpha)\exp(-T_R/T_1)}. \tag{2.19}$$

It can be noticed that the contrast of the two sequences is functionally equivalent if the excitation flip angle is close to 90°. In that case, the only difference is that the Spin-Echo sequence gets rid of the dephasings inherent to field inhomogeneities, thus exhibiting a T_2 dependency, while Gradient-Echo sequences depend upon the shorter T_2^* time. For both sequences, the impact of T_1 on the contrast is weighted by T_R, while the impact of T_2 is weighted by T_E. The conditions favoring the different contrasts are listed in Tables 2.2 and 2.3.

T_1-weighted MRI images present a good contrast between fat, which appears dark, and water, which appears brighter. This type of contrast is used, for instance, in brain imaging to distinguish gray matter from white matter. Pathologies are often revealed by T_2-weighted MRI. Edemas (abnormal accumulation of fluids) appear bright, while tumors often appear darker than normal tissues.

Examples of T_1 and spin-density weighted images are shown in Figure 2.10.

Table 2.3: Basic MR contrasts and corresponding choices of T_E and T_R for Gradient-Echo sequences.

Contrast	Spin density	T_1	T_2^*
T_R	$\gg T_1$	$\approx T_1$	$\gg T_1$
T_E	$\ll T_2^*$	$\ll T_2^*$	$\approx T_2^*$
α	$\ll 90°$	$\gg 0°$	$\ll 90°$

Figure 2.10: Sagittal MRI scans of the author's knee with different contrasts. Left: T_1 ($T_E = 15\,\text{ms}$ and $T_R = 517\,\text{ms}$). Right: spin density ($T_E = 11\,\text{ms}$ and $T_R = 2630\,\text{ms}$).

2.4.2 Parallel MRI

In this section, we briefly describe the principle of parallel imaging and we present the main approaches to reconstruction. For further details, we suggest that the reader consults the detailed topical reviews by Blaimer *et al.* [30] and by Larkman and Numes [31].

2.4.2.1 Motivations

Parallel MRI (pMRI) is a method developed in the past decades to reduce the scan duration, breaking the limits imposed to the fastest gradient sequences. The scan time constraint is of particular importance in medicine as it conditions the discomfort underwent by the patients. It is also the limiting factor in crucial applications such as cardiac imaging. pMRI exploits the complementary spatial information of several receiving coil. A

well-designed array of receiving coils combined with an adequate reconstruction algorithm allows for the reduction of scan time, while preserving image details and contrast. The speedup achieved with pMRI is due to the multiple measurements being recorded by several coils in parallel, whereas classical MRI, relying on a single measurement coil, needs more time to get the data necessary for imaging.

2.4.2.2 Effect of Undersampling in Classical MRI

In early applications of pMRI, only Cartesian k-space sampling schemes were considered. We focus on Cartesian sampling strategies for the sake of simplicity. Specifically, in this section, we assume that the k-space trajectory is formed by scanning lines along the so-called frequency encoding direction. At the end of each line, a field gradient in the phase-encoding direction creates a perpendicular shift. That way, the next frequency-encoding step explores a new adjacent line in k-space. The lines acquired are generally equidistant. With such Cartesian sampling scheme, the scan-time is in direct in proportion to the number of lines scanned.

The physical speed limit for the acquisition of each line mainly resides in the performance of gradient switching. Once this hardware has been optimized, the only approach left to accelerate the scanning time is to reduce the number of lines. Sacrificing the highest frequency lines is a possible choice, but it directly impacts on the resolution of the reconstructed images. A second approach, which is dealt with in pMRI, is to increase the distance between lines. This frequency spacing is inversely proportional to the reconstruction FOV in the corresponding phase-encoding direction. According to Shannon's sampling theory, a loosely sampled k-space leads to an image suffering from aliasing artifacts. More precisely, when the object imaged is larger than the reconstruction FOV, its extremities appear folded in the reconstructed image. The effect of both approaches to reduce the number of lines is illustrated in Figure 2.11.

pMRI techniques aim at utilizing the localization of receiving coils in order to unfold the aliasing artifacts associated to the increased distance between k-space lines. From the k-space domain point of view, pMRI exploits the convolution properties of the receiving coil sensitivities in order to interpolate the missing k-space lines.

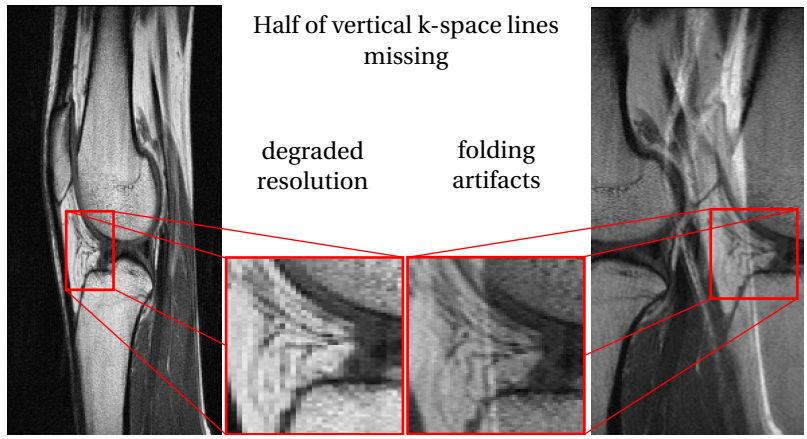

Figure 2.11: Reconstructed images from incomplete k-space measurements. Left: High frequency lines missing. Right: Each second line is missing.

2.4.2.3 Frequency-Domain Approaches

Frequency-domain approaches rely on linear combinations of the coil sensitivities. If they can reproduce sufficiently many spatial harmonics, the coefficients can be utilized to combine the acquired k-space lines such as to fill the entire k-space of the target image. The first technique of this kind proposed is SMASH [32]. It relies on prescanned coil sensitivities. Later, autocalibrated variants were proposed, with names like auto-SMASH [33] and VD-Auto-SMASH [34]. They save the time of the prescans. The coil calibration, which is realized by means of the acquisition of extra lines at the center of the k-space, robustifies reconstructions with respect to motion. GRAPPA [35] improves fitting of the extra calibration lines and differs from its predecessors in the sense that it recovers full k-space data for every coil channel. The image is reconstructed by inverse Fourier transform and a root sum of squares combination of the images from each channel.

2.4.2.4 Spatial-Domain Approaches

Alternative pMRI techniques such as PILS [36] and SENSE [3] undertake the reconstruction problem in the spatial domain. Both of them require prescans to estimate coil sensitivities. PILS is limited to particular coil-

array designs because it assumes very localized sensitivities, while SENSE deals with much more general situations. It is interesting to note that SENSE allows for a theoretical analysis of the physical limits of pMRI in terms of SNR. The so-called g-factor quantifies the noise sensitivity for each pixel of the final image. This factor can vary abruptly in space depending on the k-space trajectory and the coil-array configuration. A second major advantage of SENSE is that it can be mapped to an inverse problem approach which allows for very versatile k-space sampling strategies [37]. This inverse problem approach is the subject of the following chapter.

Chapter 3

Reconstruction Methods

WHILE in a majority of biomedical modalities, images are produced directly, in MRI they are only obtained after a reconstruction process. In this chapter, we present the current approaches to MRI reconstruction. We recall in Section 3.1 the continuous model of the data-collection process in MRI and derive its discretized version. In Section 3.2, we review the different approaches that lead to linear reconstructions and we introduce the key concepts of the inverse problem formalism. This approach maps image reconstruction into an optimization problem with the possibility to impose *a priori* constraints to distinguish the solution from other possible candidates and improve reconstruction quality. We legitimate in Section 3.3 the use of sparsity-promoting priors in MRI and explain how they can be imposed via a proper regularization term. Finally, we review in Section 3.4 the algorithmic procedures that are theoretically capable of achieving the desired reconstructions while being suited to the practical constraints encountered in MRI.

3.1 Model for MRI

3.1.1 Physics

A radio-frequency pulse is emitted to initiate nuclear magnetic resonance (NMR). It excites the spins in a 2-D plane or a 3-D volume, depending of the type of acquisition format. After excitation, the excited spins behave as radio-frequency emitters and have their precessing frequency and phase modified depending on their positions. This is achieved thanks to

the time-varying magnetic gradient fields that are applied during the relaxation, defining a trajectory $\mathbf{k}$ in the k-space domain. The modulated part of the signal received by a coil of sensitivity $S_i(\mathbf{r})$ is given by

$$m_i(\mathbf{k}) = \int_{\mathbb{R}^2} S_i(\mathbf{r})\rho(\mathbf{r})\mathrm{e}^{-2\mathrm{j}\pi\mathbf{k}\cdot\mathbf{r}}\mathrm{d}\mathbf{r}. \tag{3.1}$$

The signal ρ is referred to as *object*. This signal is proportional to the spin density, but might also depend upon other local characteristics. More details on the derivation of relation (3.1) are provided in Chapter 2.

For an array of R receiving coils with sensitivities denoted by $S_1 \cdots S_R$ and a k-space trajectory sampled at N points $\mathbf{k}_n$, we represent the measurements concatenated in a global $RN \times 1$ vector

$$\mathbf{m} = \left(\left(m_{1,1}, \ldots, m_{N,1}\right), \ldots \left(m_{1,i}, \ldots, m_{N,i}\right), \ldots \left(m_{1,R}, \ldots, m_{N,R}\right)\right).$$

3.1.2 Model for the Original Data

3.1.2.1 Spatial Discretization of the Object

From here on, we consider that the Fourier domain and, in particular, the sampling points $\mathbf{k}_n$, are scaled to make the Nyquist sampling interval unity. This can be done without any loss of generality if the space domain is scaled accordingly. Therefore, we model the object as a linear combination of pixel-domain basis functions $\varphi_{\mathbf{p}}$ that are shifted replicates of some generating function φ, so that

$$\rho = \sum_{\mathbf{p}\in\mathbb{Z}^2} c[\mathbf{p}]\varphi_{\mathbf{p}}, \text{ with} \tag{3.2}$$

$$\varphi_{\mathbf{p}}(\mathbf{r}) = \varphi\left(\mathbf{r}-\mathbf{p}\right). \tag{3.3}$$

Given a sampled version of the coil sensitivity $s_i[\mathbf{p}]$, the sensitivity-weighted object is modeled by

$$S_i\rho = \sum_{\mathbf{p}\in\mathbb{Z}^2} s_i[\mathbf{p}]c[\mathbf{p}]\varphi_{\mathbf{p}}. \tag{3.4}$$

The standard implicit choice for φ is Dirac's delta even if it is hardly justified from an approximation theoretic point of view. Different discretizations have been proposed, for example by Sutton *et al.* [38] with φ as a

boxcar function or later by Delattre *et al.* [39] with B-splines. It is only recently [40] that the details have been worked out to get back the image for general φ that are non-interpolating, which is the case for instance for B-splines of degree greater than 1. The image to be reconstructed—*i.e.*, the sampled version of the object $\rho(\mathbf{p})$—is obtained by filtering the coefficients $c[\mathbf{p}]$ with the discrete filter

$$P\left(\mathrm{e}^{\mathrm{j}\boldsymbol{\omega}}\right) = \sum_{\mathbf{h}\in\mathbb{Z}^2} \hat{\varphi}\left(\boldsymbol{\omega} + 2\pi\mathbf{h}\right), \tag{3.5}$$

where $\hat{\varphi}$ denotes the Fourier transform of φ.

Since a finite field of view (FOV) determines sets of coefficients c and s_i with a finite number M of elements, we handle them as a vectors $\mathbf{c}$ and $\mathbf{s}_i$, keeping the discrete coordinates $\mathbf{p}$ as implicit indexing.

3.1.2.2 Wavelet Discretization

Due to sparsity properties that are discussed later in this chapter, it might be preferable to represent the object in terms of wavelet coefficients. In the wavelet formalism, some constraints apply to φ. It must be a scaling function that satisfies the properties for a multiresolution [41]. In that case, the wavelets can be defined as linear combinations of the $\varphi_{\mathbf{p}}$ and the object is equivalently characterized by its coefficients in the orthonormal wavelet basis. We refer to Mallat's reference book [42] for a full review on wavelets. There exists a discrete wavelet transform (DWT) that bijectively maps the coefficients $\mathbf{c}$ to the wavelet coefficients $\mathbf{w}$ that represent the same object ρ in a continuous wavelet basis. In the rest of the chapter, we represent this DWT by the synthesis matrix $\mathbf{W}$. Note that the matrix-vector multiplications $\mathbf{c} = \mathbf{W}\mathbf{w}$ and $\mathbf{w} = \mathbf{W}^{-1}\mathbf{c}$ have efficient filterbank implementations.

3.1.3 Matrix Representation of the Model

The data-formation model (3.1) and the object parameterization (3.4) are combined to model the measurement corresponding to every point $\mathbf{k}_n$ sampled in k-space. Accordingly, the measurement vector $\mathbf{m}$ is related to the coefficients $\mathbf{c}$ through the linear relation

$$\mathbf{m} = \mathbf{E}\mathbf{c}, \tag{3.6}$$

where the MRI encoding matrix $\mathbf{E}$ is formed as

$$\mathbf{E} = \left(\mathbf{I}_R \otimes \mathbf{E}_0\right) \left[\operatorname{diag}\left(\mathbf{s}_1\right), \ldots, \operatorname{diag}\left(\mathbf{s}_R\right)\right]^{\mathrm{T}}, \tag{3.7}$$

with the symbol $\otimes$ standing for the Kronecker product, $\mathbf{I}_R$ representing the $R \times R$ identity matrix, and $\mathbf{E}_0$ being the encoding matrix for the same MRI scan with a single homogeneous receiving coil

$$\mathbf{E}_0 = \operatorname{diag}\left(\left[\hat{\varphi}(2\pi\mathbf{k}_1), \ldots, \hat{\varphi}(2\pi\mathbf{k}_N)\right]\right) \left[\mathbf{v}_1, \ldots, \mathbf{v}_N\right]^{\mathrm{T}}. \tag{3.8}$$

There, $\mathbf{v}_n$ are vectors indexed like $\mathbf{c}$ such that $v_n[\mathbf{p}] = \exp\left(-2\mathrm{j}\pi\mathbf{k}_n \cdot \mathbf{p}\right)$.

Due to the presence of noise and other scanner inaccuracies, the introduction of a new term $\mathbf{b}$, accounting for an additive perturbation, makes the data-formation model

$$\mathbf{m} = \mathbf{E}\mathbf{c} + \mathbf{b} \tag{3.9}$$

more realistic. Equivalently, if the parameters of interest are the wavelet coefficients $\mathbf{w}$, the model writes

$$\mathbf{m} = \mathbf{M}\mathbf{w} + \mathbf{b} \tag{3.10}$$

with $\mathbf{M} = \mathbf{E}\mathbf{W}$.

In MRI, the major source of noise is a radio-frequency signal originating from the thermal motion in the object under investigation. When observed with a receiving array of coils, this noise presents non-negligible correlations across channels. In other terms, the $R \times R$ channel cross-correlation matrix $\boldsymbol{\Theta}$ has non-null off-diagonal entries. Accordingly, the additive perturbation is generally modeled as the realization of a centered multivariate Gaussian process $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi})$ with covariance matrix $\boldsymbol{\Psi} = \boldsymbol{\Theta} \otimes \mathbf{I}_N$.

3.2 Linear Solutions

The problem of imaging is to recover the M coefficients $\mathbf{c}$ (or equivalently $\mathbf{w}$) from the N corrupted measurements $\mathbf{m}$. In this section, we review the popular approaches that lead to reconstructions that depend linearly upon the observations. We show that they are functionally equivalent. Two of these approaches rely on a stochastic interpretation of the problem, where the matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Upsilon}$ are the known covariance matrices of the noise $\mathbf{b}$ and the object $\mathbf{c}$, respectively. The corresponding global

variances are given by $\nu_n = \mathrm{Tr}(\boldsymbol{\Psi})/N = \mathrm{Tr}(\boldsymbol{\Theta})$ and $\nu_s = \mathrm{Tr}(\boldsymbol{\Upsilon})/M$. We define the normalized covariance matrices as $\boldsymbol{\Psi}_0 = \boldsymbol{\Psi}/\nu_n$ and $\boldsymbol{\Upsilon}_0 = \boldsymbol{\Upsilon}/\nu_s$. Most linear solutions involve a balancing parameter λ which is necessarily positive and can be interpreted in terms of the signal-to-noise ratio $\lambda^{-1} = \mathrm{Tr}(\boldsymbol{\Upsilon})/\mathrm{Tr}(\boldsymbol{\Psi}) = M\nu_s/(N\nu_n)$.

3.2.1 Pseudoinverse

Depending on the scanner settings, the encoding matrix $\mathbf{E}$ is generally neither square nor invertible. In such cases, the Moore-Penrose pseudoinverse offers a solution to the reconstruction problem. The reconstruction matrix is then defined as

$$\mathbf{E}^{\dagger} = \lim_{\epsilon \to 0^{+}} \left(\mathbf{E}^{\mathrm{H}}\mathbf{E} + \epsilon \mathbf{I}_M\right)^{-1} \mathbf{E}^{\mathrm{H}}. \tag{3.11}$$

The Hermitian matrix transpose is used, denoted by the superscript $^{\mathrm{H}}$, because in MRI matrices have complex-valued entries. The problem of inverting a non-square matrix is tackled by considering the backprojected [1] problem

$$\mathbf{E}^{\mathrm{H}}\mathbf{m} = \mathbf{E}^{\mathrm{H}}\mathbf{E}\mathbf{c}, \tag{3.12}$$

because the matrix $\mathbf{E}^{\mathrm{H}}\mathbf{E}$ is square.

Considering the singular value decomposition $\mathbf{E} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}$, where $\boldsymbol{\Sigma}$ is an $RN \times M$ matrix whose diagonal entries are the singular values σ_n, one gets $\mathbf{E}^{\dagger} = \mathbf{V}\boldsymbol{\Sigma}^{\dagger}\mathbf{U}^{\mathrm{H}}$, with singular values

$$\sigma_n^{\dagger} = \begin{cases} 0 & \text{if } \sigma_n = 0, \\ 1/\sigma_n & \text{otherwise.} \end{cases}$$

The major concern with pseudoinverse reconstruction resides in the propagation of noise. Indeed, very small but non-null singular values lead to drastic amplification of the corresponding noise components. This effect is quantified by the condition number, defined as $\kappa(\mathbf{E}) = \max_n \sigma_n / \min_n \sigma_n$. This number, which is greater or equal to 1, is also representative of the numerical challenge faced when inverting $\mathbf{E}$. A linear inverse problem is termed "ill-conditioned" when the corresponding condition number is

1. Backprojection is a term used in tomography that refers to the multiplication by the transpose of the encoding matrix.

large. When the null space of $\mathbf{E}$ is not limited to $\{\mathbf{0}\}$, the problem is said to be ill-posed.

The aim of regularized reconstruction schemes is to improve reconstruction with respect to the pseudoinverse approach by limiting the propagation of noise in the images.

It is remarkable that (3.12) rewrites as $\boldsymbol{\nabla}_{\mathbf{c}}\left(\|\mathbf{m}-\mathbf{E}\mathbf{c}\|_2^2\right)=\mathbf{0}$, with $\boldsymbol{\nabla}_{\mathbf{c}}$ standing for the gradient operator with respect to $\mathbf{c}$. The More-Penrose pseudoinverse provides a least-squares solution $\mathbf{c}^\star=\mathbf{E}^\dagger\mathbf{m}$ to the reconstruction problem because it ensures $\mathbf{E}^{\mathrm{H}}\mathbf{m}=\mathbf{E}^{\mathrm{H}}\mathbf{E}\mathbf{c}^\star$. This least-squares solution makes sense when the noise term $\mathbf{b}$ is independent and identically distributed.

Instead, when the noise correlation matrix $\boldsymbol{\Psi}_0$ is available, this knowledge can be exploited using the weighted pseudoinverse

$$\mathbf{E}_{\mathbf{X}}^{\dagger}=\lim_{\epsilon\to 0^+}\left(\mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E}+\epsilon\mathbf{I}_M\right)^{-1}\mathbf{E}^{\mathrm{H}}\mathbf{X} \tag{3.13}$$

with the weighting matrix $\mathbf{X}=\boldsymbol{\Psi}_0^{\dagger}$. The interest of that type of solution is that it takes into account noise correlations and that it relies less on the noisier samples. Thanks to the relation $\mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E}\mathbf{E}_{\mathbf{X}}^{\dagger}\mathbf{m}=\mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, the weighted pseudoinverse provides a (weighted) least-squares solution.

3.2.2 Quadratic Regularization

The approach proposed by Phillips [43] and Twomey [44], for finite dimensional problems, and by Tikhonov [45], for infinite dimensional problems, defines the reconstruction as the minimization of the functional

$$\|\mathbf{m}-\mathbf{E}\mathbf{c}\|_{\mathbf{X}}^2+\lambda\|\mathbf{R}\mathbf{c}\|^2, \tag{3.14}$$

where the notation $\|\cdot\|_{\mathbf{X}}$ with $\mathbf{X}$ positive-definite stands for a weighted norm such that $\|\mathbf{v}\|_{\mathbf{X}}^2=\mathbf{v}^{\mathrm{H}}\mathbf{X}\mathbf{v}$. The functional is a trade-off between a fidelity term, which enforces consistency with the measurements, and a regularization term, which penalizes non-regular solutions with respect to the regularization matrix $\mathbf{R}$. The tuning parameter λ balances the influence of these two terms. The role of the regularization term is to limit the amplification of noise that can be dramatic for ill-conditioned problems (in MRI, see for instance [46]). In practice, it is often designed with a derivation operator to favor smooth solutions. Similar to the weighted

pseudoinverse solution, the weighting matrix can be chosen as $\mathbf{X} = \boldsymbol{\Psi}_0^{\dagger}$ yielding a reconstruction matrix that gives importance to the samples in inverse proportion to their level of noise. Another common choice is to take $\mathbf{X}$ diagonal such as to compensate for an inhomogeneous k-space sampling density [37]. This choice facilitates the reconstruction.

The minimization of a quadratic functional yields a linear solution. Indeed, by taking the gradient of the functional and setting it to zero, we find that the reconstruction matrix writes

$$\mathbf{F}_{\mathrm{QUAD}} = \left(\mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E} + \lambda\mathbf{R}^{\mathrm{H}}\mathbf{R}\right)^{-1}\mathbf{E}^{\mathrm{H}}\mathbf{X}. \tag{3.15}$$

3.2.3 Maximum *a posteriori*

Here, the reconstruction problem is tackled within a stochastic framework. The unknowns $\mathbf{c}$ and $\mathbf{b}$ are modeled as realizations of centered multivariate Gaussian distributions: $\mathbf{c} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Upsilon})$ and $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi})$.

According to the numerical model (3.9), the measurements also follow a multivariate Gaussian distribution $\mathbf{m} \sim \mathcal{N}(\mathbf{0}, \mathbf{E}\boldsymbol{\Upsilon}\mathbf{E}^{\mathrm{H}} + \boldsymbol{\Psi})$.

The maximum *a posteriori* solution (MAP) $\mathbf{c}$ is the vector that maximizes the posterior distribution given the measurements $\mathbf{m}$. Using Bayes' theorem, the probability density function of the posterior distribution of $\mathbf{c}$ writes

$$p(\mathbf{c} \mid \mathbf{m}) \propto p(\mathbf{m} \mid \mathbf{c})\, p(\mathbf{c}).$$

In the present stochastic setting, the probability density function can be expanded in

$$p(\mathbf{c} \mid \mathbf{m}) \propto \exp\left(-\|\mathbf{m} - \mathbf{E}\mathbf{c}\|^2_{\boldsymbol{\Psi}^{\dagger}}\right) \exp\left(-\|\mathbf{c}\|^2_{\boldsymbol{\Upsilon}^{\dagger}}\right). \tag{3.16}$$

Finally, the MAP solution is the vector $\mathbf{c}$ that minimizes the functional

$$\|\mathbf{m} - \mathbf{E}\mathbf{c}\|^2_{\boldsymbol{\Psi}_0^{\dagger}} + \lambda \|\mathbf{c}\|^2_{\boldsymbol{\Upsilon}_0^{\dagger}}. \tag{3.17}$$

We introduced the normalized covariance matrices in the later expression in order to have the parameter λ, which is the inverse of the signal to noise ratio, appear explicitly.

Similarly to the previous approaches, the functional to be minimized is composed of quadratic terms. As a consequence, the solution is linear, characterized by the reconstruction matrix

$$\mathbf{F}_{\mathrm{MAP}} = \left(\mathbf{E}^{\mathrm{H}}\boldsymbol{\Psi}_0^{\dagger}\mathbf{E} + \lambda\boldsymbol{\Upsilon}_0^{\dagger}\right)^{-1}\mathbf{E}^{\mathrm{H}}\boldsymbol{\Psi}_0^{\dagger}. \tag{3.18}$$

3.2.4 Linear Minimum Mean Squared Error estimator

The Gaussian model used in the MAP approach is hardly justified for true MRI images $\mathbf{c}$. This assumption can be substituted by the constraint that the reconstruction is affine with respect to the measurements. Accordingly, we write the reconstructed image $\mathbf{Fm}+\mathbf{g}$.

To determine adequate parameters $\mathbf{F}$ and $\mathbf{g}$, one can rely on the two first order statistics of the unknown data; that are, the expectation vectors $\bar{\mathbf{c}}$ and $\bar{\mathbf{b}}$, and covariance matrices $\boldsymbol{\Upsilon}$ and $\boldsymbol{\Psi}$. According to the data-formation model (3.9), the expectation and covariance of the reconstruction error $\mathbf{e}=\mathbf{Fm}+\mathbf{g}-\mathbf{c}$ are given by

$$\mathbb{E}\{\mathbf{e}\} = \mathbf{F}(\mathbf{E}\bar{\mathbf{c}}+\bar{\mathbf{b}})+\mathbf{g}-\bar{\mathbf{c}} \tag{3.19}$$

and

$$\mathbb{E}\left\{(\mathbf{e}-\mathbb{E}\{\mathbf{e}\})(\mathbf{e}-\mathbb{E}\{\mathbf{e}\})^{\mathrm{H}}\right\} = (\mathbf{FE}-\mathbf{I})\boldsymbol{\Upsilon}(\mathbf{FE}-\mathbf{I})^{\mathrm{H}}+\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^{\mathrm{H}}+\mathbb{E}\{\mathbf{e}\}\,\mathbb{E}\{\mathbf{e}\}^{\mathrm{H}}. \tag{3.20}$$

An unbiased reconstruction [2] is obtained when $\mathbf{g}=\bar{\mathbf{c}}-\mathbf{F}(\mathbf{E}\bar{\mathbf{c}}+\bar{\mathbf{b}})$. For the choice of $\mathbf{F}$, one would reasonably like to minimize the variance of the reconstruction error. Given that the estimator is unbiased, the variance also corresponds to the expectation of the mean-square error. It is is given by the trace of the covariance matrix

$$\mathrm{Var}\{\mathbf{e}\} = \mathrm{Tr}\big((\mathbf{FE}-\mathbf{I})\boldsymbol{\Upsilon}(\mathbf{FE}-\mathbf{I})^{\mathrm{H}}\big)+\mathrm{Tr}\big(\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^{\mathrm{H}}\big). \tag{3.21}$$

Interestingly, this relation reveals two distinct contributions to the error.

- The term $(\mathbf{FE}-\mathbf{I})\boldsymbol{\Upsilon}(\mathbf{FE}-\mathbf{I})^{\mathrm{H}}$ accounts for the reconstruction artifacts. They appear when the matrix $\mathbf{F}$ is not a left inverse of $\mathbf{E}$. Specifically, this situation occurs when $\mathbf{E}$ has a null space.
- The term $\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^{\mathrm{H}}$ corresponds to the noise propagated in the reconstructed image. It is significantly amplified if $\mathbf{F}$ presents large singular values.

The matrix $\mathbf{F}$ that minimizes the error variance, also referred to as mean-square error, can be computed using matrix calculus. Using the normalized covariance matrices, it writes

$$\mathbf{F}_{\mathrm{MMSE}} = \boldsymbol{\Upsilon}_0\mathbf{E}^{\mathrm{H}}(\mathbf{E}\boldsymbol{\Upsilon}_0\mathbf{E}^{\mathrm{H}}+\lambda\boldsymbol{\Psi}_0)^{-1}. \tag{3.22}$$

2. That is to say $\mathbb{E}\{\mathbf{e}\}=\mathbf{0}$.

3.2.5 Connections

First, Equations (3.15) and (3.18) show that quadratic regularization and MAP approaches are equivalent provided that $\mathbf{X} = \boldsymbol{\Psi}_0^\dagger$ and $\boldsymbol{\Upsilon}_0^\dagger = \mathbf{R}^H\mathbf{R}$. [3]

Second, the three following equalities reveal the connection between MAP and LMMSE solutions, (3.18) and (3.22), in the case where both matrices $\boldsymbol{\Upsilon}_0$ and $\boldsymbol{\Psi}_0$ are invertible:

$$\begin{aligned}
\mathbf{E}^H\boldsymbol{\Psi}_0^{-1}\mathbf{E}\boldsymbol{\Upsilon}_0\mathbf{E}^H + \lambda\boldsymbol{\Upsilon}_0^{-1}\boldsymbol{\Upsilon}_0\mathbf{E}^H &= \mathbf{E}^H\boldsymbol{\Psi}_0^{-1}\mathbf{E}\boldsymbol{\Upsilon}_0\mathbf{E}^H + \lambda\mathbf{E}^H\boldsymbol{\Psi}_0^{-1}\boldsymbol{\Psi}_0 \\
(\mathbf{E}^H\boldsymbol{\Psi}_0^{-1}\mathbf{E} + \lambda\boldsymbol{\Upsilon}_0^{-1})\boldsymbol{\Upsilon}_0\mathbf{E}^H &= \mathbf{E}^H\boldsymbol{\Psi}_0^{-1}(\mathbf{E}\boldsymbol{\Upsilon}_0\mathbf{E}^H + \lambda\boldsymbol{\Psi}_0) \\
\boldsymbol{\Upsilon}_0\mathbf{E}^H\big(\mathbf{E}\boldsymbol{\Upsilon}_0\mathbf{E}^H + \lambda\boldsymbol{\Psi}_0\big)^{-1} &= \big(\mathbf{E}^H\boldsymbol{\Psi}_0^{-1}\mathbf{E} + \lambda\boldsymbol{\Upsilon}_0^{-1}\big)^{-1}\mathbf{E}^H\boldsymbol{\Psi}_0^{-1}.
\end{aligned}$$

Last, the weighted pseudoinverse solution with $\mathbf{X} = \boldsymbol{\Psi}_0^\dagger$ corresponds to the other solutions in the limiting case where λ tends to 0. This is also the case for the regular Moore-Penrose pseudoinverse when the noise is independent and identically distributed; that is to say $\boldsymbol{\Psi}_0 = \mathbf{I}_{RN}/R$. As already mentioned, the pseudoinverse solutions are only valid when noise propagation is negligible. This situation occurs with well-conditioned ($\kappa(\mathbf{E}) \approx 1$) reconstruction problems that are largely overdetermined ($M \ll RN$) and/or subject to very little noise ($\mathrm{Tr}\big(\boldsymbol{\Upsilon}\big) \gg \mathrm{Tr}\big(\boldsymbol{\Psi}\big)$).

About the invertibility of $\boldsymbol{\Psi}_0$ and $\boldsymbol{\Upsilon}_0$: There is no particular reason for $\boldsymbol{\Psi}_0$ to be singular. Most of the time, the correlation between pixels in the image are not modeled; this translates in a matrix $\boldsymbol{\Upsilon}_0$ which is diagonal. When no signal is expected from some pixels of the image, (for instance outside a predetermined ROI), it could be tempting to set to 0 the corresponding entries in $\boldsymbol{\Upsilon}_0$, resulting in a singular matrix. However, a reasonable problem setting would exclude such entries in the unknown vector $\mathbf{c}$, restoring the invertibility of $\boldsymbol{\Upsilon}_0$.

3.3 Non-Quadratic Regularizations

We just saw that the linear approaches to reconstruction can be derived from the solution of some optimization problems. The corresponding

3. Let us mention that $\boldsymbol{\Upsilon}$ being a covariance matrix, it is necessarily Hermitian symmetric. Its pseudoinverse is also Hermitian symmetric and admits the same eigenvectors.

functionals were quadratic, yielding closed-form solutions. In this section, we consider other approaches that are popular in MRI and which involve non-quadratic regularization terms.

3.3.1 Inverse Problem Formalism

The solution $\mathbf{c}^\star$ is defined as the minimizer of a cost function that involves two terms: the data fidelity $\mathscr{F}(\mathbf{b})$ and the regularization $\mathscr{R}(\mathbf{c})$ that penalizes undesirable solutions. This is summarized as

$$\mathbf{c}^\star = \arg\min_{\mathbf{c}} \mathscr{F}(\mathbf{m} - \mathbf{E}\mathbf{c}) + \lambda \mathscr{R}(\mathbf{c}), \tag{3.23}$$

where the regularization parameter $\lambda \geq 0$ balances the two constraints. In MRI, the noise term $\mathbf{b} = \mathbf{m} - \mathbf{E}\mathbf{c}$ is usually assumed to be the realization of a Gaussian process with normalized covariance matrix $\mathbf{\Psi}_0$. From a Bayesian point of view, this justifies the choice $\mathscr{F}(\mathbf{b}) = \|\mathbf{b}\|^2_{\mathbf{\Psi}_0^\dagger} = \mathbf{b}^{\mathrm{H}}\mathbf{\Psi}_0^\dagger\mathbf{b}$ as a proper log-likelihood term. A more practical motivation for this choice is that a quadratic fidelity term yields a simple closed-form gradient that greatly facilitates the design and performance of reconstruction algorithms.

When the k-space sampling is dense enough and the signal-to-noise ratio is high, the quadratic regularization terms (presented in the previous section) yield satisfying reconstructions. But, the constraints to reduce the scan duration favor setups with reduced SNR and k-space trajectories that present regions of low sampling density. In these situations where the reconstruction problem is more challenging, the reconstructed image can often be enhanced by the use of a more suitable regularization term $\mathscr{R}(\mathbf{c})$.

3.3.2 Total Variation

Total Variation (TV) was introduced as an edge-preserving denoising method by Rudin *et al.* [47]. It is now a very popular approach to tackle image enhancement problems.

The TV regularization term corresponds to the sum of the Euclidean norms of the gradient of the object. In practice, it is defined as $\mathscr{R}(\mathbf{w}) = \|\boldsymbol{\nabla}\mathbf{c}\|_{\ell_1}$. In this context, the operator $\boldsymbol{\nabla}$ returns pixelwise the ℓ_2-norm of finite differences. The use of TV regularization is particularly appropriate for piecewise-constant objects such as the Shepp-Logan (SL) phantom

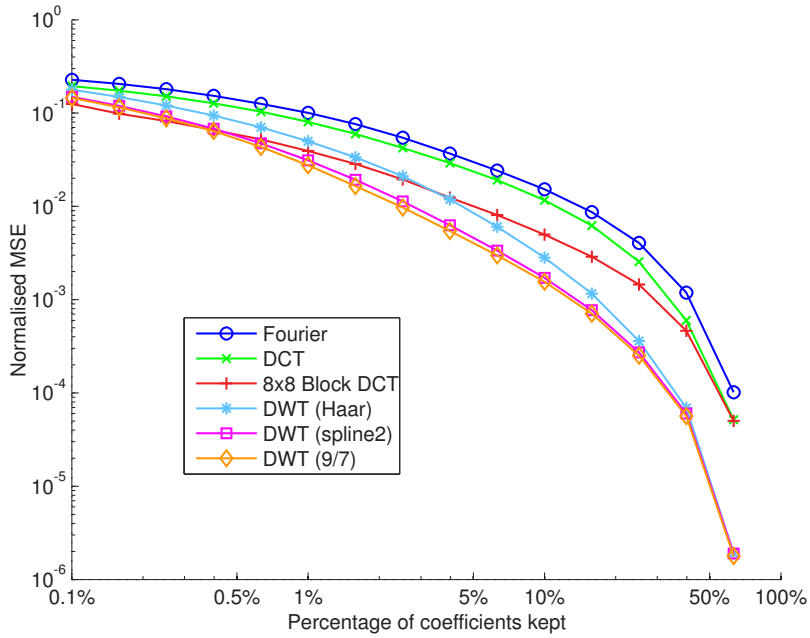

Figure 3.1: Sparse approximation errors of a typical MRI brain image using different orthonormal transforms. The Mean Squared Error is represented as a function of the percentage of coefficients kept. Wavelet transforms achieve the best sparse approximations.

used for simulations in tomography and MRI. Textured and noisy images exhibit a much larger total variation.

3.3.2.1 Sparsity-Promoting Regularization

Another popular idea is to exploit the fact that the object can be well represented by few non-zero coefficients (sparse representation) in an orthonormal basis of M functions $\phi_{\mathbf{p}}$. Formally, we write that

- $\exists S \subset \mathbb{Z}^2, \quad |S| \ll M$ *(sparse support)* and
- $\exists \mathbf{a}: \quad \left\| \rho - \sum_{\mathbf{p} \in S} a[\mathbf{p}] \phi_{\mathbf{p}} \right\| \ll \left\| \rho \right\|$ *(small error)*.

It is well-documented that typical MRI images admit sparse representation in bases such as wavelets or block DCT [6]. We illustrate this property in Figure 3.1.

The ℓ_1-norm is a good measure of sparsity with interesting mathematical properties (*e.g.*, convexity). Thus, among the candidates that are consistent with the measurements, we favor a solution whose wavelet coefficients have a small ℓ_1-norm. Specifically, the solution is formulated as

$$\mathbf{w}^\star = \arg\min_{\mathbf{w}} \mathscr{C}(\mathbf{w}), \tag{3.24}$$

with

$$\mathscr{C}(\mathbf{w}) = \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\ell_2}^2 + \lambda \|\mathbf{w}\|_{\ell_1}. \tag{3.25}$$

This is the general solution for wavelet-regularized inverse problems considered by [19] as well as by many other authors.

3.4 Algorithms

MRI gives rise to a *large-scale* inverse problem in the sense that the number of degrees of freedom—that is to say, the unknown pixel values—is large. Consequently, the matrices are generally too large to be stored in memory not to mention the fact that direct matrix multiplication involves too many operations. [4] We summarize in this section the strategies that make the reconstruction in MRI feasible with reasonable computer requirements and acceptable computation times.

3.4.1 Matrix-Vector Multiplications

The matrix-vector multiplications $\mathbf{y} = \mathbf{E}_0\mathbf{x}$ and $\mathbf{y} = \mathbf{E}_0^{\mathrm{H}}\mathbf{x}$ are two basic operations in MRI reconstruction. They can be implemented efficiently using the FFT algorithm. For non-Cartesian samples $\mathbf{k}_n$, the gridding method, based on FFT and interpolation, can provide accurate computations (see [48] for instance). Algorithms 1 and 2 describe the implementation of the operations $\mathbf{y} = \mathbf{E}_0\mathbf{x}$ and $\mathbf{y} = \mathbf{E}_0^{\mathrm{H}}\mathbf{x}$, respectively.

4. Take a single channel MRI problem with $M = 256 \times 256$ unknowns and $N = 256 \times 256$ measurements. The corresponding encoding matrix $\mathbf{E}$ is $65\,536 \times 65\,536$. With double precision floats that are required for accurate calculations, the storage of this complex-valued matrix would require 64 GiB of RAM memory which is a too large value for current personal computers. Performing direct matrix-vector multiplications involves an overwhelming amount of scalar multiplications and additions.

Algorithm 1: Matrix-Vector multiplication $\mathbf{y} = \mathbf{E}_0\mathbf{x}$, according to (3.8).

input : $\mathbf{x}$, $\mathbf{p}_1,\ldots,\mathbf{p}_M$, $\mathbf{k}_1,\ldots,\mathbf{k}_N$, and $\hat{\varphi}$;
$\mathbf{y} \leftarrow$ `Gridding`$(\mathbf{x}, \mathbf{p}_1,\ldots,\mathbf{p}_M, \mathbf{k}_1,\ldots,\mathbf{k}_N)$ (going to k-space domain);
for $n \leftarrow 1$ **to** N **do**
  $y_n \leftarrow \hat{\varphi}(2\pi\mathbf{k}_n)y_n$;
return $\mathbf{y}$;

Algorithm 2: Matrix-Vector multiplication $\mathbf{y} = \mathbf{E}_0^{\mathrm{H}}\mathbf{x}$, according to (3.8).

input : $\mathbf{x}$, $\mathbf{k}_1,\ldots,\mathbf{k}_N$, $\mathbf{p}_1,\ldots,\mathbf{p}_M$, and $\hat{\varphi}$;
for $n \leftarrow 1$ **to** N **do**
  $x_n \leftarrow \hat{\varphi}(2\pi\mathbf{k}_n)^* x_n$ (superscript * indicates the conjugate transpose);
$\mathbf{y} \leftarrow$ `Gridding`$(\mathbf{x}, -\mathbf{k}_1,\ldots,-\mathbf{k}_N, \mathbf{p}_1,\ldots,\mathbf{p}_M)$ (going to spatial domain);
return $\mathbf{y}$;

An interesting work by Wajer [49] identifies $\mathbf{E}_0^{\mathrm{H}}\mathbf{E}_0$ as a convolution matrix associated to the kernel

$$G[\mathbf{p}] = \sum_{n=1}^{N} \left|\hat{\varphi}(2\pi\mathbf{k}_n)\right|^2 \exp\left(2\mathrm{j}\pi\mathbf{k}_n \cdot \mathbf{p}\right). \tag{3.26}$$

When the kernel is precomputed for the lattice points belonging to the set $S = \{\mathbf{p}-\mathbf{q} \mid \mathbf{p} \in \mathrm{FOV}, \mathbf{q} \in \mathrm{FOV}\}$, one can avoid the use of Algorithms 1 and 2. An efficient implementation of the operation $\mathbf{y} = \mathbf{E}_0^{\mathrm{H}}\mathbf{E}_0\mathbf{x}$, which uses zero-padded multidimensional FFTs, is described in Algorithm 3.

Most of the time, in parallel MRI, the covariance matrices are block diagonal. In that case, they are sparse matrices and one can benefit from the related efficient memory storage and matrix operations. As already mentioned, $\mathbf{\Psi}_0$ is fully characterized by the channel cross-correlation matrix $\mathbf{\Theta}_0 = \mathbf{\Theta}/\nu_n$ such that $\mathbf{\Psi}_0 = \mathbf{\Theta}_0 \otimes \mathbf{I}_N$. Its pseudoinverse or inverse is then given by $\mathbf{\Psi}_0^{\dagger} = \mathbf{\Theta}_0^{\dagger} \otimes \mathbf{I}_N$. The matrix-vector multiplications with $\mathbf{E}^{\mathrm{H}}\mathbf{\Psi}_0^{\dagger}$ and $\mathbf{E}^{\mathrm{H}}\mathbf{\Psi}_0^{\dagger}\mathbf{E}$ are implemented as described in Algorithms 4 and 5, respectively.

Algorithm 3: Matrix-Vector multiplication $\mathbf{y} = \mathbf{E}_0^{\mathrm{H}}\mathbf{E}_0\mathbf{x}$, according to (3.8) and (3.26).

input : $\mathbf{x}$, S and $\mathbf{G}$;
Precompute $\mathbf{G} \leftarrow$ `FFT(G)` (DFT coefficients);
$\mathbf{x} \leftarrow$ `ZERO-PADDING(x,S)` (zero-padding $\mathbf{x}$ to the dimensions of $\mathbf{G}$);
$\mathbf{x} \leftarrow$ `FFT(x)` (computing DFT coefficients);
for $n \leftarrow 1$ **to** *number of elements of* $\mathbf{G}$ **do**
  $x_n \leftarrow G_n x_n$ (multiplication of DFTs);
$\mathbf{x} \leftarrow$ `IFFT(x)` (inverse DFT);
for $\mathbf{p} \in FOV$ **do**
  $y[\mathbf{p}] \leftarrow x[\mathbf{p}]$;
return $\mathbf{y}$;

Algorithm 4: Matrix-Vector multiplication $\mathbf{y} = \mathbf{E}^{\mathrm{H}}\boldsymbol{\Psi}_0^{\dagger}\mathbf{x}$, according to (3.7).

input : $\mathbf{x}$, $\mathbf{s}_1, \ldots, \mathbf{s}_R$, $\boldsymbol{\Theta}_0^{\dagger}$ and $\mathbf{E}_0^{\mathrm{H}}$;
$(\mathbf{x}_1, \ldots, \mathbf{x}_R) \leftarrow \mathbf{x}$;
for $r \leftarrow 1$ **to** R **do**
  $\mathbf{x}_r \leftarrow \mathbf{E}_0^{\mathrm{H}}\mathbf{x}_r$ (using Algorithm 2);
$\mathbf{y} \leftarrow \mathbf{0}$;
for $r \leftarrow 1$ **to** R **do**
  $\mathbf{y}_r \leftarrow \mathbf{0}$;
  for $r' \leftarrow 1$ **to** R **do**
    $\mathbf{y}_r \leftarrow \mathbf{y}_r + [\boldsymbol{\Theta}_0^{\dagger}]_{r,r'}\mathbf{x}_{r'}$;
  for $\mathbf{p} \in FOV$ **do**
    $(\mathbf{y}_r)[\mathbf{p}] \leftarrow (\mathbf{s}_r)^*[\mathbf{p}](\mathbf{y}_r)[\mathbf{p}]$;
  $\mathbf{y} \leftarrow \mathbf{y} + \mathbf{y}_r$;
return $\mathbf{y}$;

Algorithm 5: Matrix-Vector multiplication $\mathbf{y} = \mathbf{E}^{\mathrm{H}}\boldsymbol{\Psi}_0^{\dagger}\mathbf{E}\mathbf{x}$, according to (3.7).

input : $\mathbf{x}$, $\mathbf{s}_1, \ldots, \mathbf{s}_R$, $\boldsymbol{\Theta}_0^{\dagger}$ and $\mathbf{E}_0^{\mathrm{H}}\mathbf{E}_0$;
for $r \leftarrow 1$ **to** R **do**
  for $\mathbf{p} \in FOV$ **do**
    $(\mathbf{x}_r)[\mathbf{p}] \leftarrow (\mathbf{s}_r)[\mathbf{p}]x[\mathbf{p}]$;
  $\mathbf{x}_r \leftarrow \mathbf{E}_0^{\mathrm{H}}\mathbf{E}_0\mathbf{x}_r$ (using Algorithm 3);
$\mathbf{y} \leftarrow \mathbf{0}$;
for $r \leftarrow 1$ **to** R **do**
  $\mathbf{y}_r \leftarrow \mathbf{0}$;
  for $r' \leftarrow 1$ **to** R **do**
    $\mathbf{y}_r \leftarrow \mathbf{y}_r + [\boldsymbol{\Theta}_0^{\dagger}]_{r,r'}\mathbf{x}_{r'}$;
  for $\mathbf{p} \in FOV$ **do**
    $(\mathbf{y}_r)[\mathbf{p}] \leftarrow (\mathbf{s}_r)^*[\mathbf{p}](\mathbf{y}_r)[\mathbf{p}]$;
  $\mathbf{y} \leftarrow \mathbf{y} + \mathbf{y}_r$;
return $\mathbf{y}$;

3.4.2 Conjugate Gradient

The conjugate gradient method (CG) [50] is an iterative algorithm that is among the most efficient in solving large-scale linear problems $\mathbf{A}\mathbf{c} = \mathbf{b}$, characterized by symmetric and positive-definite matrices $\mathbf{A}$. The only operations involving the matrix $\mathbf{A}$ are matrix-vector multiplications $\mathbf{A}\mathbf{x}$. In parallel MRI, it is the method of reference [37] to perform linear reconstructions. The quadratic-regularized solution characterized by the reconstruction matrix in (3.15) is computed with CG solving the linear problem defined by the matrix $\mathbf{A} = \mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E} + \lambda\mathbf{R}^{\mathrm{H}}\mathbf{R}$ and vector $\mathbf{b} = \mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{m}$.

The idea of the method is to decompose the solution in a basis of mutually conjugate vectors; that is to say $\mathbf{c} = \sum_i \alpha_i \mathbf{p}_i$, with $\mathbf{p}_i^{\mathrm{H}}\mathbf{A}\mathbf{p}_j = 0$ for $i \neq j$. At iteration i, the estimate is $\mathbf{c}_i = \sum_{j \leq i} \alpha_j \mathbf{p}_j$ and the corresponding residue writes $\mathbf{r}_i = \mathbf{b} - \mathbf{A}\mathbf{c}_i$. For the next direction, the choice $\mathbf{p}_{i+1} = \mathbf{r}_i - \sum_{j \leq i} (\mathbf{p}_j^{\mathrm{H}}\mathbf{A}\mathbf{r}_i)\mathbf{p}_j / \left\|\mathbf{p}_j\right\|_{\mathbf{A}}$ ensures the conjugacy constraint. In this direction, the coefficient $\alpha_{i+1} = \mathrm{Re}\left(\mathbf{p}_{i+1}^{\mathrm{H}}\mathbf{A}\mathbf{r}_i\right) / \left\|\mathbf{p}_{i+1}\right\|_{\mathbf{A}}$ is optimal with respect to the cost $\mathscr{C}(\mathbf{c}) = \mathbf{c}^{\mathrm{H}}\mathbf{A}\mathbf{c} - \mathbf{c}^{\mathrm{H}}\mathbf{b} - \mathbf{b}^{\mathrm{H}}\mathbf{c}$. An efficient implementation of the method is described in Algorithm 6.

Algorithm 6: CG solving $\mathbf{Ac} = \mathbf{b}$ with $\mathbf{A}$ symmetric and positive-definite

input : $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}_0$ (optional, default: $\mathbf{c}_0 = \mathbf{0}$);
Initialization: $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{c}_0$, $\mathbf{p}_0 = \mathbf{r}_0$, and $i = 0$;
repeat
  $\mathbf{q}_i \leftarrow \mathbf{A}\mathbf{p}_i$;
  $\alpha_i \leftarrow \mathbf{r}_i^{\mathrm{H}}\mathbf{r}_i / (\mathbf{p}_i^{\mathrm{H}}\mathbf{q}_i)$;
  $\mathbf{c}_{i+1} \leftarrow \mathbf{c}_i + \alpha_i \mathbf{p}_i$;
  $\mathbf{r}_{i+1} \leftarrow \mathbf{r}_i - \alpha_i \mathbf{q}_i$;
  $\mathbf{p}_{i+1} \leftarrow \mathbf{r}_{i+1} + \mathbf{r}_{i+1}^{\mathrm{H}}\mathbf{r}_{i+1} / (\mathbf{r}_i^{\mathrm{H}}\mathbf{r}_i)\mathbf{p}_i$;
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{c}_i$;

The CG algorithm theoretically converges within a finite number of iterations. In practice, this result is compromised by the propagation of round-off errors. In the context of MRI, the property of practical interest is the linear convergence rate achieved by CG. Indeed, the distance to the desired solution decreases as a power of the iteration number, with the convergence rate

$$0 \leq r(\mathbf{A}) = \left(\sqrt{\kappa(\mathbf{A})} - 1\right) / \left(\sqrt{\kappa(\mathbf{A})} + 1\right) < 1.$$

When the condition number $\kappa(\mathbf{A})$ is large, the rate $r(\mathbf{A})$ gets close to the unity, characterizing a slower convergence. Using the weighted norm $\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\mathbf{x}^{\mathrm{H}}\mathbf{A}\mathbf{x}}$, the distance is upperbounded by

$$\left\|\mathbf{c}_i - \mathbf{c}^\star\right\|_{\mathbf{A}} \leq 2 \left\|\mathbf{c}_0 - \mathbf{c}^\star\right\|_{\mathbf{A}} r(\mathbf{A})^i. \tag{3.27}$$

With the regular Euclidean distance, the bound is looser

$$\left\|\mathbf{c}_i - \mathbf{c}^\star\right\|_2 \leq 2\kappa(\mathbf{A}) \left\|\mathbf{c}_0 - \mathbf{c}^\star\right\|_2 r(\mathbf{A})^i. \tag{3.28}$$

3.4.3 Iteratively Reweighted Least-Squares

The Iteratively Reweighted Least-Squares algorithm (IRLS), which is also known as the positive form of half-quadratic minimization[51], can be used to compute the solutions defined as

$$\mathbf{c}^\star = \arg\min_{\mathbf{c}} \|\mathbf{m} - \mathbf{E}\mathbf{c}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{R}\mathbf{c}\|_{\ell^p}^p. \tag{3.29}$$

In this context, the functional is strictly convex for $p > 1$. This condition ensures the unicity of the minimizer.

The principle of IRLS is to design an upperbounding quadratic proxy for the regularization term, tailored to the neighborhood of $\mathbf{c}_i$. In practice, one chooses the functional

$$\mathscr{Z}_i(\mathbf{c}) = p/2 \, \|\mathbf{Rc}\|^2_{\mathbf{D}_i} + (1 - p/2) \left\|\mathbf{Rc}_i\right\|^p_{\ell^p}, \tag{3.30}$$

where $\mathbf{D}_i$ is a diagonal matrix with entries $|\left(\mathbf{Rc}_i\right)_n|^{p-2}$. It has the following desirable properties

(i) $\mathscr{Z}_i(\mathbf{c}_i) = \left\|\mathbf{Rc}_i\right\|^p_{\ell^p}$,

(ii) $\boldsymbol{\nabla}_{\mathbf{c}} \mathscr{Z}_i(\mathbf{c}_i) = \boldsymbol{\nabla}_{\mathbf{c}} \|\mathbf{Rc}\|^p_{\ell^p} \, |_{(\mathbf{c}=\mathbf{c}_i)}$,

(iii) $\mathscr{Z}_i(\mathbf{c}) > \|\mathbf{Rc}\|^p_{\ell^p}$ for all $\mathbf{c} \neq \mathbf{c}_i$ and $p < 2$,

(iv) $\arg\min_{\mathbf{c}} \|\mathbf{m} - \mathbf{Ec}\|^2_{\mathbf{X}} + \lambda \mathscr{Z}_i(\mathbf{c}) = \left(\mathbf{E}^{\mathrm{H}}\mathbf{XE} + (\lambda p/2)\mathbf{R}^{\mathrm{H}}\mathbf{D}_i\mathbf{R}\right)^{-1}\mathbf{E}^{\mathrm{H}}\mathbf{Xm}$.

An implementation of the IRLS is described in Algorithm 7.

Algorithm 7: IRLS solving $\mathbf{c}^\star = \arg\min_{\mathbf{c}} \|\mathbf{m} - \mathbf{Ec}\|^2_{\mathbf{X}} + \lambda \|\mathbf{Rc}\|^p_{\ell^p}$.

input : $\mathbf{A} = \mathbf{E}^{\mathrm{H}}\mathbf{XE}$, $\mathbf{a} = \mathbf{E}^{\mathrm{H}}\mathbf{Xm}$, $\mathbf{R}$, p, λ, and $\mathbf{c}_0$;
$i \leftarrow 0$;
repeat
  $\mathbf{r}_i = \mathbf{Rc}_i$;
  for $n \leftarrow 1$ **to** *number of elements of* $\mathbf{r}_i$ **do**
    $\delta_n \leftarrow (\lambda p/2)|\left(\mathbf{r}_i\right)_n|^{p-2}$;
  $\mathbf{A}_i \leftarrow \mathbf{A} + \mathbf{R}^{\mathrm{H}}\mathrm{diag}([\delta_1, \ldots, \delta_n, \ldots])\mathbf{R}$;
  $\mathbf{c}_{i+1} \leftarrow \mathtt{CG}(\mathbf{A}_i, \mathbf{a}, \mathbf{c}_i)$ (using Algorithm 6);
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{c}_i$;

Let us remember that for $p \leq 1$ the minimization problem might not admit a unique solution. When the minimizer $\mathbf{c}^\star$ is unique, it is also the unique fixed-point of the algorithm. As long as $p < 2$, the sequence of functional values $\mathscr{C}(\mathbf{c}_i) = \left\|\mathbf{m} - \mathbf{Ec}_i\right\|^2_{\mathbf{X}} + \lambda \left\|\mathbf{Rc}_i\right\|^p_{\ell^p}$ generated by the IRLS is monotonically decreasing. This guarantees the convergence since the sequence is lower-bounded by the finite quantity $\mathscr{C}^\star = \min_{\mathbf{c}} \|\mathbf{m} - \mathbf{Ec}\|^2_{\mathbf{X}} + \lambda \|\mathbf{Rc}\|^p_{\ell^p}$.

The IRLS algorithm can be simply adapted in order to solve the minimization with mixed-norm regularization terms. A particular case is the total variation penalty which corresponds to the ℓ_1-norm of the pixelwise ℓ_2-norm of the spatial gradient [52]. The IRLS algorithm for TV regularization was first proposed by Wohlberg and Rodríguez [53]. It is described in Algorithm 8.

Algorithm 8: IRLS solving $\mathbf{c}^\star = \arg\min_{\mathbf{c}} \|\mathbf{m} - \mathbf{E}\mathbf{c}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{c}\|_{\mathrm{TV}}$.

input : $\mathbf{A} = \mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{E}$, $\mathbf{a} = \mathbf{E}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, λ, and $\mathbf{c}_0$;
Define the finite difference matrices $\mathbf{R}_d$ along every spatial dimension d;
$i \leftarrow 0$;
repeat
  for $d \leftarrow 1$ **to** *number of spatial dimensions* **do**
    $\mathbf{r}_d \leftarrow \mathbf{R}_d \mathbf{c}_i$;
  for $n \leftarrow 1$ **to** *number of elements of* $\mathbf{c}_i$ **do**
    $\delta_n \leftarrow \lambda / \left(2\sqrt{\sum_d |(\mathbf{r}_d)_n|^2}\right)$;
  $\mathbf{A}_i \leftarrow \mathbf{A} + \sum_d \mathbf{R}_d^{\mathrm{H}} \mathrm{diag}([\delta_1, \ldots, \delta_n, \ldots])\mathbf{R}_d$;
  $\mathbf{c}_{i+1} \leftarrow \mathtt{CG}(\mathbf{A}_i, \mathbf{a}, \mathbf{c}_i)$ (using Algorithm 6);
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{c}_i$;

Duality-based algorithms proved to be an efficient alternative to achieve TV regularization [54, 55].

3.4.4 Iterative Shrinkage/Thresholding Algorithm

The Iterative Shrinkage/Thresholding Algorithm (ISTA) [17–19], also known as thresholded Landweber (TL), aims at minimizing the functional

$$\mathscr{C}(\mathbf{w}) = \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{w}\|_{\ell_1} . \tag{3.31}$$

Here, we use the notation $\mathbf{w}$ because ISTA is often applied on wavelet coefficients.

An important observation to understand ISTA is to see that the nonlinear shrinkage operation, sometimes called soft-thresholding, solves a mini-

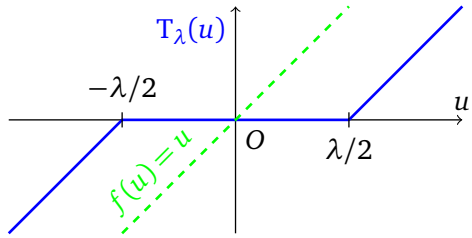

Figure 3.2: The shrinkage function T_λ.

mization problem[56], with

$$\begin{aligned} T_\lambda(u) &= (|u| - \min(\lambda/2, |u|)) \cdot \operatorname{sgn}(u) \\ &= \arg\min_{w\in\mathbb{C}} |u-w|^2 + \lambda|w|. \end{aligned} \tag{3.32}$$

By separability of norms, this applies component-wise to vectors of $\mathbb{C}^N$:

$$T_\lambda(\mathbf{u}) = \arg\min_{\mathbf{w}} \|\mathbf{u}-\mathbf{w}\|_{\ell_2}^2 + \lambda \|\mathbf{w}\|_{\ell_1}.$$

This means that the ℓ_1-regularized denoising problem (*i.e.*, when $\mathbf{M}$ and $\mathbf{X}$ are identity matrices) is precisely solved by a shrinkage operation.

The ISTA generates a sequence of estimates $\mathbf{w}_i$ that converges to the minimizer $\mathbf{w}^\star$ of (3.31) when it is unique. The idea is to define at each step a new functional $\mathscr{C}'(\mathbf{w}, \mathbf{w}_i)$ whose minimizer $\mathbf{w}_{i+1}$ will be the next estimate

$$\mathbf{w}_{i+1} = \arg\min_{\mathbf{w}} \mathscr{C}'(\mathbf{w}, \mathbf{w}_i). \tag{3.33}$$

Two constraints must be considered for the definition of $\mathscr{C}'$.

(i) It is sufficient for the convergence of the algorithm that $\mathscr{C}'(\mathbf{w}, \mathbf{w}_i)$ is an upper bound of $\mathscr{C}(\mathbf{w})$ with equality at $\mathbf{w} = \mathbf{w}_i$; this guarantees that the sequence $\{\mathscr{C}(\mathbf{w}_i)\}$ is monotonically decreasing.

(ii) The inner minimization (3.33) should be performed by a simple shrinkage operation to ensure the rapidity and accuracy of the algorithm.

In accordance with Constraint (i), $\mathscr{C}'$ can take the generic quadratically augmented form

$$\mathscr{C}'(\mathbf{w}, \mathbf{w}_i) = \mathscr{C}(\mathbf{w}) + \left\|\mathbf{w}-\mathbf{w}_i\right\|^2_{\mathbf{\Lambda}-\mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}}, \tag{3.34}$$

with the constraint that $(\mathbf{\Lambda}-\mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M})$ is positive definite, where the weighting matrix $\mathbf{\Lambda}$ plays the role of a tuning parameter.

Then, ISTA corresponds to the trivial choice $\mathbf{\Lambda} = \frac{L}{2}\mathbf{I}$, with the value of L chosen to be greater or equal to the Lipschitz constant of the gradient of $\|\mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2$, so that $L \geq 2\lambda_{\max}\left(\mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}\right)$.

Let us define $\mathbf{a} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, $\mathbf{A} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}$, and

$$\mathbf{z}_i = \mathbf{w}_i + 2(\mathbf{a} - \mathbf{A}\mathbf{w}_i)/L. \tag{3.35}$$

Then, using standard linear algebra, we can write

$$\begin{aligned} \mathbf{w}_{i+1} &= \arg\min_{\mathbf{w}} \left\|\mathbf{w} - \mathbf{z}_i\right\|_{\ell_2}^2 + (2\lambda/L)\,\|\mathbf{w}\|_{\ell 1} \\ &= \mathrm{T}_{2\lambda/L}\left(\mathbf{z}_i\right). \end{aligned}$$

This shows that Constraint (ii) is automatically satisfied.

Note that both the intermediate variable $\mathbf{z}_n$ in (3.35) and the threshold values will vary depending on L.

Algorithm 9: ISTA solving $\mathbf{w}^\star = \arg\min_{\mathbf{w}} \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2 + \lambda\,\|\mathbf{w}\|_{\ell_1}$

input : $\mathbf{A} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}$, $\mathbf{a} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, $\mathbf{w}_0$, and L;
$i \leftarrow 0$;
repeat
  $\mathbf{z}_{i+1} \leftarrow \mathbf{w}_i + 2(\mathbf{a} - \mathbf{A}\mathbf{w}_i)/L$;
  $\mathbf{w}_{i+1} \leftarrow \mathrm{T}_{2\lambda/L}\left(\mathbf{z}_{i+1}\right)$;
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{w}_i$;

Beck and Teboulle[20, Thm. 3.1] showed that this algorithm decreases the cost function in direct proportion to the number of iterations i.

Proposition 3.4.1. *Let $\{\mathbf{w}_i\}$ be the sequence generated by Algorithm 9 with $L \geq 2\lambda_{\max}(\mathbf{A})$. Then, for any $i > i_0 \in \mathbb{N}$,*

$$\mathscr{C}(\mathbf{w}_i) - \mathscr{C}(\mathbf{w}^\star) \leq \frac{L}{2(i - i_0)} \left\|\mathbf{w}_{i_0} - \mathbf{w}^\star\right\|_{\ell_2}^2. \tag{3.36}$$

Selecting L as small as possible will clearly favor the speed of convergence. It also raises the importance of a "warm" starting point.

Among the variants of ISTA, FISTA, proposed by Beck and Teboulle[20], ensures state-of-the-art convergence properties while preserving a comparable computational cost. Thanks to a controlled over-relaxation at each step, FISTA quadratically decreases the cost function, with

$$\mathscr{C}(\mathbf{w}_i) - \mathscr{C}(\mathbf{w}^\star) \leq \frac{2L}{(i+1)^2} \left\| \mathbf{w}_0 - \mathbf{w}^\star \right\|_{\ell_2}^2 . \tag{3.37}$$

More details on FISTA, as a particular case of FWISTA with the trivial choice $\mathbf{\Lambda} = \frac{L}{2}\mathbf{I}$, can be found in Section 5.2.3.

An implementation of FISTA is given Algorithm (10).

Algorithm 10: FISTA solving $\mathbf{w}^\star = \arg\min_{\mathbf{w}} \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{w}\|_{\ell_1}$

input : $\mathbf{A} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}$, $\mathbf{a} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, $\mathbf{w}_0$, and L;
Initialization: $i = 0$, $\mathbf{v}_0 = \mathbf{w}_0$, $t_0 = 1$;
repeat
  $\mathbf{w}_{i+1} \leftarrow \mathrm{T}_{2\lambda/L}\left(\mathbf{v}_i + 2(\mathbf{a} - \mathbf{A}\mathbf{v}_i)/L\right)$ (ISTA step);
  $t_{i+1} \leftarrow \left(1 + \sqrt{1 + 4t_i^2}\right)/2$;
  $\mathbf{v}_{i+1} \leftarrow \mathbf{w}_{i+1} + (t_i - 1)\left(\mathbf{w}_{i+1} - \mathbf{w}_i\right)/t_{i+1}$;
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{w}_i$;

Chapter 4

Simulation

In this chapter,[1] we present a new method for generating synthetic MRI data. The quantitative validation of reconstruction algorithms requires reliable data. Rasterized simulations are popular but they are tainted by an aliasing component that impacts the assessment of the performance of reconstruction. We introduce analytical simulation tools that are suited to parallel magnetic resonance imaging and allow one to build realistic phantoms. The proposed phantoms are composed of ellipses and regions with piecewise-polynomial boundaries, including spline contours, Bézier contours, and polygons. In addition, they take the channel sensitivity into account, for which we investigate two possible models. Our analytical formulations provide well-defined data in both the spatial and k-space domains. Our main contribution is the closed-form determination of the Fourier transforms that are involved. Experiments validate the proposed implementation. In a typical parallel MRI reconstruction experiment, we quantify the bias in the overly optimistic results obtained with rasterized simulations—the inverse-crime situation. We provide a package that implements the different simulations and contains tools to guide the design of realistic phantoms.

4.1 Motivations

An active area of research in magnetic resonance imaging (MRI) is the development of reconstruction algorithms. In particular, the inverse-

1. This chapter is partly Copyright ©, 2012, IEEE. Adapted, with permission, from [57].

problem approach is getting popular [58], where one relies on an accurate model of the measurement process and possibly on additional information about the object being imaged.

In general, the development of any reconstruction approach requires that it be evaluated and compared to other existing methods. There are several reasons to rely on simulations in a first step

– saving the costs inherent to getting real scanner data
– testing the suitability of the implemented discrete forward model
– quantitatively evaluating the performance of the reconstruction software
– providing reliable ground-truth data to compare with.

However, for the results to be meaningful, simulations must be accomplished carefully. For instance, the inverse-crime situation, where exactly the same discrete model is used for simulation and reconstruction, leads to artificially good results. In the context of MRI, many developers of algorithms base their simulations on rasterized images. One should just be aware that such testing does not account for the full continuous-domain reality, because it neglects the aliasing that is inherent to spatial discretization. More realistic simulations are required to remove this bias and to ensure that the methods will perform adequately in practice.

A method to obtain resolution-independent simulations is to formulate the simulation analytically in the continuous domain. This approach goes back to Shepp and Logan [59], who introduced an ellipse-based phantom (SL) for X-ray tomography. For MRI, several analytical phantoms have been proposed. The first works, based on the SL phantom, are by Smith *et al.* [60], followed by Van de Walle *et al.* [61]. More recently, Koay *et al.* [62] worked out the MR contribution of an ellipsoid for the 3-D extension of the SL phantom. Gach *et al.* [29] adapted these elliptical phantoms specifically for MRI, introducing realistic physical parameters as well as T_1 and T_2 relaxation times. The family of analytical phantoms is extended by two recent works by Greengard and Stucchio [63] that use Gaussian functions, and Ngo *et al.* [64] that introduce 3-D polyhedra.

The attractiveness of currently known analytical phantoms is limited for two reasons. First, the vast majority of currently available phantoms (except [64]) use ellipses as basic elements. While such simple shapes have the advantage of mathematical tractability, they do not lend themselves

well to the generation of images with realistic anatomical features. Secondly, to the best of our knowledge, no analytical phantom has been proposed that would take into account MRI receiving-coil sensitivities in the context of the simulation of parallel MRI experiments [3].

In this work, we extend the class of available analytical phantoms by introducing regions parameterized by spline contours which are general enough to reproduce polygons and Bézier contours. Our shapes are well suited to the description of realistic anatomical regions [65]. To accurately simulate image formation in parallel MRI, we also make use of analytical models for the coil sensitivity maps. Specifically, we investigate the use of two classes of basis functions—polynomials [3] and complex sinusoids—which both have the ability to generate maps that are physically realistic. These parametric forms are used to derive closed-form solutions for the MRI coil data. We have implemented and tested both models.

This chapter is organized as follows: in Section 4.2, we present the different models considered for the parallel MRI measurement process, the analytical phantom, and the coil sensitivities. We motivate and compare the polynomial and the proposed sinusoidal models. In Section 4.3, we propose the main theoretical elements that make the analytical MRI simulation possible, deferring the more technical considerations until Appendices A.1, A.2, and A.3. Finally, we present in Section 4.4 the experiments that validate our implementation of the theoretical tools and an application that quantifies the bias of rasterized simulations on linear and nonlinear reconstructions, in a typical parallel MRI setup.

4.2 Modeling

In this section, we present the MRI measurement model and building blocks that are used to define our phantom.

4.2.1 Parallel MRI

We use the well-established linear model for parallel MRI that relates the object ρ to the k-space signal m_{S_n} observed by each receiving coil, via the Fourier integral

$$m_{S_n}(\mathbf{k}) = \int S_n(\mathbf{r})\rho(\mathbf{r})\mathrm{e}^{-\mathrm{j}2\pi\mathbf{k}\cdot\mathbf{r}}\mathrm{d}\mathbf{r}, \tag{4.1}$$

where S_n accounts for the sensitivity map of the n-th receiving channel. We refer to Chapter 2 for more details, in particular, the relation with the Biot-Savart law.

4.2.2 Analytical Phantom

We mathematically define the phantom ρ as a simple function, involving R regions $\mathscr{R}_i$ of constant intensity ρ_i

$$\rho(\mathbf{r}) = \sum_{i=1}^{R} \rho_i \chi_{\mathscr{R}_i}(\mathbf{r}). \tag{4.2}$$

The term *region* refers to a connected and bounded set of $\mathbb{R}^d$. The symbol $\chi_{\mathscr{R}}$ denotes the characteristic function of a region $\mathscr{R}$. Such a phantom has a limited spatial support ($\bigcup_{i=1}^{R} \mathscr{R}_i$) that we call a *region of interest* (ROI).

This model allows us to render realistic phantoms of two kinds

- piecewise-constant phantoms that mimic segmented data with sharp contours (*e.g.*, the SL brain phantom)
- textured phantom via a triangular-mesh approach.

We investigate the first approach in this chapter. The contours that are considered are ellipses, polygons, and quadratic-spline curves. We show in Figure 4.1 three such phantoms that we use in our experiments.

4.2.3 Sensitivity Models

For computations, we need to parameterize the complex sensitivity maps. It is commonly admitted that they are smooth and slowly-varying spatially. It is therefore possible to generate physically-realistic sensitivity maps using a reasonably small number of low-pass basis functions. Here, we discuss two models that are well-suited for this task. They both relate linearly the parameters to the complex sensitivity values. Moreover, their corresponding MRI models involve the Fourier integrals of monomials over the regions of the phantom.

Definition 4.2.1. For $\boldsymbol{\alpha} \in \mathbb{N}^d$ and $\mathscr{R}$ a region of $\mathbb{R}^d$

$$f_{\mathscr{R}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \int_{\mathscr{R}} \mathbf{r}^{\boldsymbol{\alpha}} \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}} \mathrm{d}\mathbf{r}. \tag{4.3}$$

Figure 4.1: Phantoms parameterized by elliptical and Bézier-defined regions. From left to right: rectangle phantom used for validation, the Shepp-Logan phantom, and a proposed brain phantom. The PDF and SVG versions of these phantoms are available online at `http://bigwww.epfl.ch/algorithms/mriphantom/`.

Here, we adopted the multi-index notation $\mathbf{r}^{\boldsymbol{\alpha}}$ defined as $\mathbf{z}^{\boldsymbol{\alpha}} = \prod z_i^{\alpha_i} \quad \in \mathbb{R}$.

4.2.3.1 Polynomial Sensitivity

This model, first proposed in [3] to represent the local behavior of the sensitivity, assumes that the coil sensitivity S is represented by a polynomial of degree D inside the ROI as

$$S(\mathbf{r}) = \sum_{d=0}^{D} \sum_{|\boldsymbol{\alpha}|=d} s_{d,\boldsymbol{\alpha}} \mathbf{r}^{\boldsymbol{\alpha}}, \quad \forall \mathbf{r} \in \text{ROI}. \tag{4.4}$$

As the degree D increases, the model will reproduce sharper transitions. The number of polynomial coefficients in 2-D is $N_p = (D+1)(D+2)/2$.

The corresponding MR response is given by

$$m_S(\mathbf{k}) = \sum_{i=1}^{R} \rho_i \sum_{d=0}^{D} \sum_{|\boldsymbol{\alpha}|=d} s_{d,\boldsymbol{\alpha}} f_{\mathscr{R}_i}^{\boldsymbol{\alpha}}(2\pi\mathbf{k}). \tag{4.5}$$

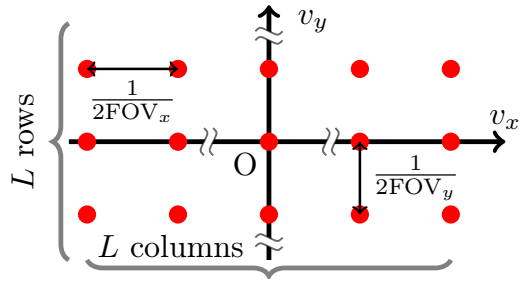

Figure 4.2: Grid of the angular frequencies involved in the sinusoidal model.

4.2.3.2 Sinusoidal Sensitivity

Alternatively, the coil sensitivity is defined by the linear combination of complex exponentials

$$S(\mathbf{r}) = \sum_{\mathbf{v}} s_{\mathbf{v}} e^{j\mathbf{r}\cdot\mathbf{v}}, \quad \forall \mathbf{r} \in \mathrm{ROI}. \tag{4.6}$$

We propose to constrain the problem to the angular frequencies $\mathbf{v}$ on a Cartesian grid with spacings that correspond to twice the considered field of view (FOV). The low-frequency properties are ensured by only considering the $L \times L$ angular frequencies around the origin (see Figure 4.2).

Similarly to the effect of the polynomial degree D, an increase in the parameter L allows one to reproduce sharper transitions. The number of coefficients in 2-D is given by $N_s = L^2$. The corresponding MR response is given by

$$m_S(\mathbf{k}) = \sum_{i=1}^{R} \rho_i \sum_{\mathbf{v}} s_{\mathbf{v}} f^{\mathbf{0}}_{\mathscr{R}_i}(2\pi\mathbf{k} - \mathbf{v}). \tag{4.7}$$

4.2.3.3 Comparison

In order to evaluate and compare the ability of the two models to describe realistic sensitivity maps, we considered a 256×256 rasterization of the SL phantom and the 27 648 pixels of its ROI. Using Biot-Savart's law (2.9), we simulated the complex sensitivity maps of a 24-channel circular head coil array (FOV: 28 cm, distance to center: 17 cm, radius: 5 cm) distributed around the phantom. In Figure 4.3, the average fitting

properties of the two models are presented as a function of the number of parameters.

We observe that the fitting accuracy of both models rapidly increases with the number of parameters, with a sensible advantage for the sinusoidal model. The downside is an increased condition number for the fitting operations. With respect to that criterion, the sinusoidal model behaves also better. The maximal spatial errors are comparable for both models.

4.3 Analytical MRI Measurements

4.3.1 Overview of Analytical Fourier Computations

In this section, we present the theoretical tools that are necessary to derive the analytical expression of the MRI measurements. Proofs and additional calculation details are provided in Appendices A.1, A.2, and A.3.

The models presented in the previous section allow us to decompose the analytical MRI measurements into Fourier integrals of the sensitivity over the regions that compose the phantom. Depending on the type of region or sensitivity model, we propose tailored methods to decompose the analytical response as a sum of special functions that can be computed accurately and rapidly. In Figure 4.4, we present the road map of these decompositions that are defined and worked out in the sequel.

4.3.2 Elliptical Regions

Let us consider an elliptical region $\mathscr{E}$ in 2-D parameterized by its center $\mathbf{r}_c$, the angle θ formed between its semimajor axis A and the abscissa, and its semiminor axis B. The linear transformation

$$\mathbf{r} \mapsto \mathbf{u} = \mathbf{D}^{-1}\mathbf{R}^{\mathrm{T}}\left(\mathbf{r} - \mathbf{r}_c\right), \tag{4.8}$$

with $\mathbf{D} = \mathrm{diag}(A, B)$ and $\mathbf{R}$ the rotation matrix of angle θ, maps $\mathscr{E}$ into a unit disk, that is to say, $\mathscr{E} = \{\mathbf{u} \mid \|\mathbf{u}\| \leq 1\}$. The Fourier transform of the unit disk involves the functions

$$G_n(\mathbf{x}) = J_n(\|\mathbf{x}\|)/\|\mathbf{x}\|^n, \tag{4.9}$$

where J_n denotes the n-th order Bessel function of the first kind [66].

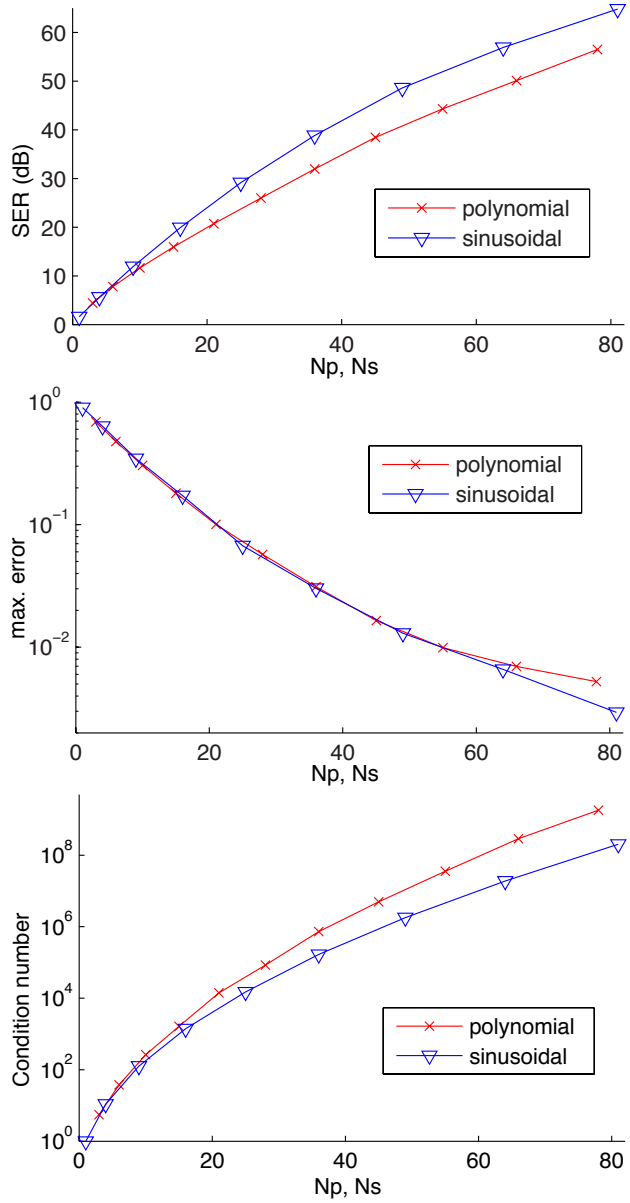

Figure 4.3: Fitting properties of the two sensitivity models as a function of the number of parameters. From left to right: approximation signal to error ratio (SER) in dB, maximal absolute error, and condition number of the fitting matrix.

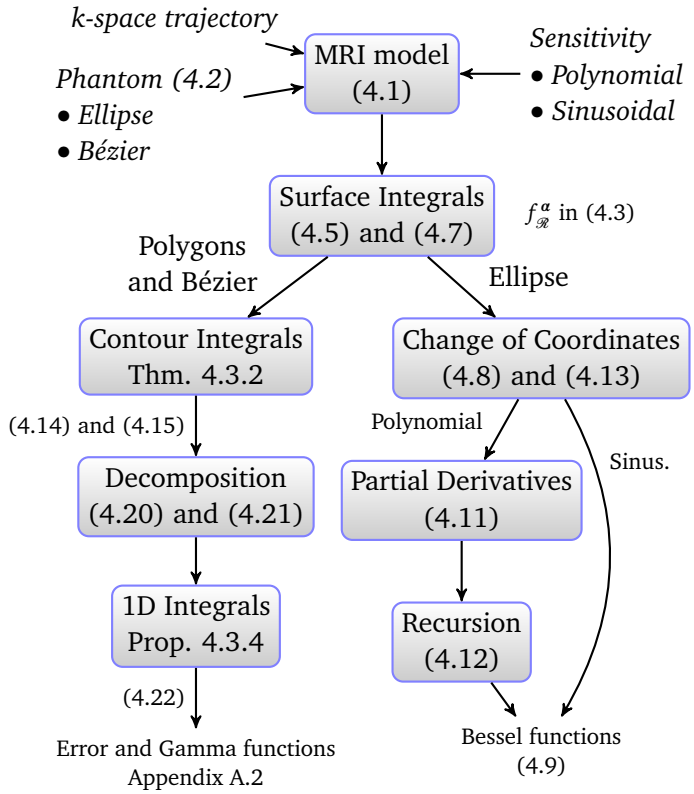

Figure 4.4: Road map of analytical Fourier computations.

Using the sinusoidal sensitivity model, the integral $f_{\mathscr{E}}^{\mathbf{0}}$ can be worked out [61] as

$$f_{\mathscr{E}}^{\mathbf{0}}(\boldsymbol{\omega}) = 2\pi|\mathbf{D}|\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}_c} G_1\left(\mathbf{D}\mathbf{R}^{\mathrm{T}}\boldsymbol{\omega}\right), \tag{4.10}$$

where $|\mathbf{D}|$ represents the absolute value of the determinant of matrix $\mathbf{D}$.

When considering the polynomial sensitivity model, we suggest to first consider the change of variables (4.8), rather than computing $f_{\mathscr{E}}^{\boldsymbol{\alpha}}$ directly. We write that

$$\int_{\mathscr{E}} \mathbf{u}^{\boldsymbol{\alpha}} \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}} \mathrm{d}\mathbf{r} = 2\pi|\mathbf{D}|\mathrm{j}^{|\boldsymbol{\alpha}|}\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}_c}\left(\frac{\partial^{|\boldsymbol{\alpha}|}G_1}{\partial\mathbf{x}^{\boldsymbol{\alpha}}}\right)\left(\mathbf{D}\mathbf{R}^{\mathrm{T}}\boldsymbol{\omega}\right). \tag{4.11}$$

The interesting point is that the partial derivatives $\partial^{|\boldsymbol{\alpha}|}G_1/\partial\mathbf{x}^{\boldsymbol{\alpha}}$ can be

Table 4.1: Closed-form MR contribution of elliptical and quadratic Bézier regions for the proposed sensitivity models.

	Model	Contribution of the region to MR measurements
Ellipse	Sinusoidal	$2\pi\lvert\mathbf{D}\rvert e^{-2\pi j\mathbf{k}\cdot\mathbf{r}_c}\sum_{\mathbf{v}} s_{\mathbf{v}} e^{j\mathbf{v}\cdot\mathbf{r}_c} G_1\left(\mathbf{DR}(-\theta)(2\pi\mathbf{k}vcv)\right)$
	Polynomial	$2\pi\lvert\mathbf{D}\rvert e^{-2\pi j\mathbf{k}\cdot\mathbf{r}_c}\sum_{d=0}^{D}\sum_{\lvert\boldsymbol{\alpha}\rvert=d} j^{\lvert\boldsymbol{\alpha}\rvert} t_{d,\boldsymbol{\alpha}} \frac{\partial^{\lvert\boldsymbol{\alpha}\rvert} G_1}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(2\pi\mathbf{DR}(-\theta)\mathbf{k})$
Bézier	Sinusoidal	$\sum_{\mathbf{v}} s_{\mathbf{v}} f_{\mathscr{B}}^{\mathbf{0}}(2\pi\mathbf{k}-\mathbf{v})$
	Polynomial	$\sum_{d=0}^{D}\sum_{\lvert\boldsymbol{\alpha}\rvert=d} s_{d,\boldsymbol{\alpha}} f_{\mathscr{B}}^{\boldsymbol{\alpha}}(2\pi\mathbf{k})$

decomposed recursively as a sum of G_n thanks to the property

$$\nabla G_n(\mathbf{x}) = -\mathbf{x} G_{n+1}(\mathbf{x}). \tag{4.12}$$

The coefficients of the polynomial in terms of the new coordinates (4.8) are required to satisfy

$$S(\mathbf{r}) = \sum_{d=0}^{D}\sum_{\lvert\boldsymbol{\alpha}\rvert=d} s_{d,\boldsymbol{\alpha}}\mathbf{r}^{\boldsymbol{\alpha}} = \sum_{d=0}^{D}\sum_{\lvert\boldsymbol{\alpha}\rvert=d} t_{d,\boldsymbol{\alpha}}\mathbf{u}^{\boldsymbol{\alpha}}. \tag{4.13}$$

They can be computed by inverting the matrix that relates the N_p coefficients to the sensitivity values at $N \geq N_p$ randomly chosen points in terms of the new coordinates.

The MR contribution of such an elliptical contour is presented in the upper part of Table 4.1.

4.3.3 Piecewise-Quadratic Contours

In this section, we first provide relations for the computation of the d-dimensional Fourier transform of a monomial delimited by a connected subset $\mathscr{B}$ of $\mathbb{R}^d$. With methods that are similar to the ones employed in [67], we show how to decompose the d-dimensional Fourier integral into a sum of integrals over the contour $\partial\mathscr{B}$. These summed integrals are of reduced dimensionality. In a second step, we show how quadratic-spline curves involve a family of 1-D integrals.

4.3.3.1 Fourier Transform of Monomials over a Connected Set

We show that the surface integral $f_{\mathscr{B}}^{\boldsymbol{\alpha}}$ in (4.3) can be decomposed into a sum of contour integrals.

Definition 4.3.1. We define

$$g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}) = \int_{\partial\mathscr{B}} \frac{\mathbf{r}^{\boldsymbol{\alpha}+\mathbf{e}_k}}{1+\alpha_k}\mathbf{e}_k \cdot \mathbf{n}\mathrm{d}\sigma, \quad \forall k \in [1,d], \tag{4.14}$$

$$g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \int_{\partial\mathscr{B}} \mathbf{r}^{\boldsymbol{\alpha}} \frac{\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}}}{\|\boldsymbol{\omega}\|^2}\boldsymbol{\omega} \cdot \mathbf{n}\mathrm{d}\sigma, \quad \forall \boldsymbol{\omega} \in \mathbb{R}^d \setminus \{\mathbf{0}\}, \tag{4.15}$$

where $\mathbf{n}$ stands for the outward-pointing unit normal of boundary element $\mathrm{d}\sigma$. Note that $g_{\mathscr{B}}^{\boldsymbol{\alpha}}$ is not continuous at the origin $\boldsymbol{\omega} = \mathbf{0}$.

Theorem 4.3.2. *For $\boldsymbol{\omega} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\boldsymbol{\alpha} \in \mathbb{N}^d$, with Definitions 4.2.1 and 4.3.1*

$$f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \mathrm{j}\sum_{\mathbf{m}=\mathbf{0}}^{\boldsymbol{\alpha}} \left(\frac{-\mathrm{j}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|^2}\right)^{\boldsymbol{\alpha}-\mathbf{m}} |\boldsymbol{\alpha}-\mathbf{m}|! \binom{\boldsymbol{\alpha}}{\mathbf{m}} g_{\mathscr{B}}^{\mathbf{m}}(\boldsymbol{\omega}), \tag{4.16}$$

and

$$f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}) = g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}). \tag{4.17}$$

The consequence of Theorem 4.3.2 is that the d-dimensional integral $f_{\mathscr{B}}^{\boldsymbol{\alpha}}$ can be decomposed into a sum of $(d-1)$-dimensional integrals. The proof is provided in Appendix A.1.

Note that the case $\boldsymbol{\omega} = \mathbf{0}$, which corresponds to the calculation of the moments of the region, has been worked out first by Jacob *et al.* [68] for parametric 2-D spline contours.

4.3.3.2 Parameterization of a Contour in 2-D

The region $\mathscr{B}$ is defined by its boundary, the contour $\partial\mathscr{B}$. In 2-D , a convenient way to parameterize the contour is by the use of a B-spline generating function φ such that

$$\forall \mathbf{r} \in \partial\mathscr{B}, \exists t \in \mathbb{R}, \quad \mathbf{r}(t) = \sum_{p\in\mathbb{Z}} \mathbf{c}_p \varphi(t-p). \tag{4.18}$$

The considered contour is closed. Consequently, the vector-valued function $\mathbf{r}$ must be periodic. In addition, the number N of coefficients $\mathbf{c}_p$ that

characterize the curve must be finite. The simplest way to satisfy these constraints is to impose that the sequence of coefficients $\mathbf{c}_p$ be N-periodic. This enforces the N-periodicity of $\mathbf{r}$.

If we note φ_p the N-periodized version of φ, the contour is parameterized either globally as

$$\forall t \in [0, N[\,, \quad \mathbf{r}(t) = \sum_{q=0}^{N-1} \mathbf{c}_q \varphi_p(t-q) \tag{4.19}$$

or piecewise, with $0 \leq t = n + \lambda < M$, $n \in 0 \dots N-1$ and $\lambda \in [0, 1[$, as

$$\mathbf{r}(\lambda + n) = \sum_{q=0}^{N-1} \mathbf{c}_{n-q} \varphi_p(\lambda + q). \tag{4.20}$$

4.3.3.3 Decomposition of the Contour Integrals

We introduce the notation $\mathbf{z}^{\perp}$ for the vector perpendicular to $\mathbf{z}$ with same norm and pointing outwards the region $\mathscr{B}$ at the considered point (see Figure 4.5). We write $\mathbf{r}'(t) = \frac{\partial \mathbf{r}}{\partial t}(t)$. The piecewise representation of the contour (4.20) can be exploited to decompose the contour integral of interest, for instance (4.14) or (4.15), which leads to the form

$$\int_{\partial \mathscr{B}} \mathbf{F}(\mathbf{r}) \cdot \mathbf{n} \mathrm{d}\sigma = \sum_{q=0}^{N-1} \int_0^1 \mathbf{F}\left(\mathbf{r}(q+\lambda)\right) \cdot \mathbf{r}'^{\perp}(q+\lambda) \mathrm{d}\lambda. \tag{4.21}$$

4.3.3.4 Quadratic Bézier Curves

In the sequel, we focus on contours represented by linear and quadratic B-splines. The former describe polygons while the latter give a piecewise description of quadratic Bézier curves. Three equivalent piecewise representations can be useful and are given in Table 4.2 with their relationships.

Definition 4.3.3.

$$h^{(m)}(a, b) = \int_0^1 \lambda^m \mathrm{e}^{-\mathrm{j}\lambda(a+\lambda b)} \mathrm{d}\lambda. \tag{4.22}$$

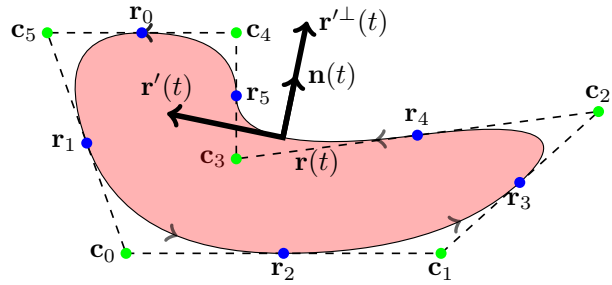

Figure 4.5: Example of a quadratic-spline-defined region with $N = 6$ control points. The boundary is counter-clockwise and the normal vector is outward-pointing. The contour is represented piecewise by quadratic Bézier curves.

Table 4.2: Piecewise representations of quadratic-B-spline contours.

Representation	$\mathbf{r}(t) = \mathbf{r}(\lambda + n)$ with $n \in \mathbb{N}$ and $\lambda \in [0,1]$
B-spline	$\frac{\lambda^2}{2}\mathbf{c}_n + \left(\frac{1}{2} + \lambda - \lambda^2\right)\mathbf{c}_{n-1} + \frac{1}{2}(1-\lambda)^2\,\mathbf{c}_{n-2}$
Polynomial	$\mathbf{r}_n + \lambda\boldsymbol{\beta}_n + \lambda^2\boldsymbol{\gamma}_n$
Bézier curve	$(1-\lambda)^2\mathbf{r}_n + 2\lambda(1-\lambda)\mathbf{c}_{n-1} + \lambda^2\mathbf{r}_{n+1}$
Relations	$\mathbf{r}_n = \frac{1}{2}\left(\mathbf{c}_{n-1} + \mathbf{c}_{n-2}\right)$ $\boldsymbol{\beta}_n = 2\left(\mathbf{c}_{n-1} - \mathbf{r}_n\right)$ $\boldsymbol{\gamma}_n = \mathbf{r}_{n+1} + \mathbf{r}_n - 2\mathbf{c}_{n-1}$

Proposition 4.3.4. *For $\boldsymbol{\omega} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and a contour $\partial\mathscr{B}$ parameterized piecewise by $\mathbf{r}(\lambda + n) = \mathbf{r}_n + \lambda\boldsymbol{\beta}_n + \lambda^2\boldsymbol{\gamma}_n$, with $n \in \mathbb{N}$ and $\lambda \in [0,1]$, using Definitions 4.3.1 and 4.3.3, we have that*

$$g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \sum_{n=0}^{N-1} \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}_n} \sum_{i=0}^{2|\boldsymbol{\alpha}|+1} d_{n,i}\, h^{(i)}(\boldsymbol{\omega}\cdot\boldsymbol{\beta}_n, \boldsymbol{\omega}\cdot\boldsymbol{\gamma}_n) \tag{4.23}$$

while

$$g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}) = \sum_{n=0}^{N-1} \sum_{i=0}^{2|\boldsymbol{\alpha}|+3} d'_{n,i}\, h^{(i)}(0,0), \tag{4.24}$$

where, with the notation $\sigma_{\mathbf{p}} = |\mathbf{p}_2| + 2|\mathbf{p}_3|$*, the symbol* $d_{n,i}$ *stands for*

$$\sum_{\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3=\boldsymbol{\alpha}} \frac{\mathbf{r}_n^{\mathbf{p}_1}\boldsymbol{\beta}_n^{\mathbf{p}_2}\boldsymbol{\gamma}_n^{\mathbf{p}_3}\boldsymbol{\alpha}!}{\mathbf{p}_1!\mathbf{p}_2!\mathbf{p}_3!}\boldsymbol{\omega}\cdot\left(\boldsymbol{\beta}_n^{\perp}\delta_{\sigma_{\mathbf{p}},i} + 2\boldsymbol{\gamma}_n^{\perp}\delta_{\sigma_{\mathbf{p}},i-1}\right), \tag{4.25}$$

and $d'_{n,i}$ *stands for*

$$\sum_{\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3=\boldsymbol{\alpha}+\mathbf{e}_k} \frac{\mathbf{r}_n^{\mathbf{p}_1}\boldsymbol{\beta}_n^{\mathbf{p}_2}\boldsymbol{\gamma}_n^{\mathbf{p}_3}\boldsymbol{\alpha}!}{\mathbf{p}_1!\mathbf{p}_2!\mathbf{p}_3!}\mathbf{e}_k\cdot\left(\boldsymbol{\beta}_n^{\perp}\delta_{\sigma_{\mathbf{p}},i} + 2\boldsymbol{\gamma}_n^{\perp}\delta_{\sigma_{\mathbf{p}},i-1}\right). \tag{4.26}$$

The values $h^{(m)}(a,b)$ follow a three-term recurrence relation [69]. More details on their numerical computation are given in Appendix A.2.

Note that the piecewise parameterization of the contour of a polygon corresponds to the particular case of a quadratic parameterization with $\boldsymbol{\beta}_n = \mathbf{r}_{n+1} - \mathbf{r}_n$ and $\boldsymbol{\gamma}_n = \mathbf{0}$. Such simpler polygonal models with homogeneous sensitivities have been considered in prior work [63, Prop. 3.2] using a similar formulation.

4.4 Experiments

4.4.1 Implementation Details

Our implementation uses Matlab 7.12 (Mathworks, Natick). The experiments run on a 64-bit 8-core computer, clock rate 2.8 GHz, 8 GiB RAM (DDR2 at 800 MHz), Mac OS X 10.6.7.

We implemented the analytical computations as described by the scheme in Figure 4.4, with double float precision. For efficient computations of the error function of a complex variable, we coded the critical parts of `erfz` in C++/MEX, with POSIX multithreading, following Marcel Leutenegger's recommendations.[2] The code implementing Theorem 4.3.2 utilizes Matt Fig's `npermutek`.[3] The rasterization of spline-defined regions, which is performed without approximation, partly relies on Bruno Luong's MEX implementation of `insidepoly`.[4] Our package also includes

2. Available online at `https://documents.epfl.ch/users/l/le/leuteneg/www/` .

3. Available at `http://www.mathworks.com/matlabcentral/fileexchange/11462-npermutek/` .

4. Available at `http://www.mathworks.com/matlabcentral/fileexchange/27840-2d-polygon-interior-detection/` .

graphical tools to design the analytical phantoms. For purposes of adequate visualization, export to the popular vector-graphics formats SVG 1.1 and PDF (via the PGF/Tikz LaTeX package) is supported. The package is distributed [5] in order to provide sensitivity fitting, phantom-design interface, analytical simulation tools, and to allow replication of the experiments of this section.

Unlike the sinusoidal model which is very robust to numerical errors, our current implementation of the three-term recurrence relation (see Appendix A.2) leads to instabilities when using the polynomial model. The theoretical relation $|h^{(m)}(a,b)| \leq 1/(m+1)$ is sometimes violated for orders $m \geq 2$ and large values of the first argument. This prevented us to present valid simulations of piecewise quadratic contours combined with a polynomial sensitivity. Given the comparison of the two models in Section 4.2.3.3, we considered the sinusoidal model with parameter $L = 7$, that is $N_s = 49$ in Figure 4.3, which lead to accurate representations of the physical sensitivities and numerically tractable inversions.

As an alternative to our analytical method, we consider the traditional simulation procedure that consists in *i)* sampling the phantom with a grid of a given size and *ii)* resampling the DFT of this discrete image according to the desired k-space trajectory. We call this procedure a rasterized simulation. It is expected to be consistent with our analytical method only when considering an infinitely dense sampling.

For reconstructions, we consider an optimization problem of the form

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}} \|\mathbf{m} - \mathbf{E}\mathbf{x}\|_2^2 + \lambda \mathscr{P}(\mathbf{x}), \tag{4.27}$$

where $\mathbf{x}$ represents an image, $\mathbf{x}^\star$ is the reconstructed one, $\mathbf{m}$ is the concatenated scanner data vector, and $\mathscr{P}$ is a regularization function. The MRI encoding matrix $\mathbf{E}$ is formed as defined in (3.7) using the usual implicit choice of Dirac's delta for the generating function.

We used two types of regularizations in our experiments

- $\mathscr{P}(\mathbf{x}) = \|\mathbf{x}\|_2^2$ corresponds to a Tikhonov regularization, which leads to linear reconstructions that we implemented with the conjugate-gradient method.
- $\mathscr{P}(\mathbf{x}) = \|\mathbf{x}\|_{\mathrm{TV}}$ is the isotropic total-variation pseudo-norm, which leads to a nonlinear reconstruction problem. This reconstruction scheme

5. Available online at `http://bigwww.epfl.ch/algorithms/mriphantom/` .

Table 4.3: Errors of our analytical simulations with the Rectangle phantom.

NRMSE	max. error in k-space	max. error inverse DFT
1.5e-15	2.8e-16	7.0e-15

is often used in compressed-sensing research and is particularly suited for dealing with piecewise-constant images such as our phantoms. We implemented it using the iteratively reweighted least-squares algorithm, also known as the additive form of the half-quadratic minimization [51, 70].

Please refer to Chapter 3 for more details on these regularization schemes.

4.4.2 Validation of the Implementation

4.4.2.1 Simple Example with Homogeneous Sensitivity

As first validation, we consider the simple phantom composed of a rectangular region that is represented in Figure 4.1. Under a proper change of variables, it yields a square and its Fourier transform is given by a product of sinc functions. This phantom is composed of a polygon and consequently falls in the category of the spline-defined contours. We test the accuracy of our proposed simulation method and of the rasterized approach against the closed-form solution. To do so, we consider the MR response associated with a homogeneous receiving coil sensitivity and a 256×256 Cartesian k-space sampling. The simulation errors are reported in Tables 4.3 and 4.4.

As expected, the error of rasterized simulations decreases when the sampling density increases. Meanwhile, the accuracy of our analytical implementation is as good as the machine `double float` precision would allow. Thus, we conclude that we can indistinctly use the closed-form ground truth or our proposed analytical model in the conditions of Section 4.4.2.1.

Table 4.4: Rasterized simulation errors with the Rectangle phantom.

Resolution	NRMSE	max. error in k-space	max. error inverse DFT
256	5.58e-02	5.5e-03	5.5e-01
352	2.51e-02	2.0e-03	1.9e-01
400	2.01e-02	1.5e-03	1.6e-01
512	1.25e-02	1.1e-03	1.0e-01
704	7.45e-03	5.5e-04	6.2e-02
800	6.04e-03	5.3e-04	5.4e-02
1024	3.85e-03	3.6e-04	3.9e-02
1408	1.59e-03	1.2e-04	1.5e-02
1600	1.27e-03	1.0e-04	1.2e-02
2048	8.32e-04	5.8e-05	6.7e-03

4.4.2.2 Validation with non-Homogeneous Sensitivity

We now use our analytical phantom as a gold standard to evaluate the accuracy the measurements obtained from rasterized simulations. We consider the SL and brain phantoms. The single sensitivity map is computed using Biot-Savart's law and is approximated on the support of each phantom with the sinusoidal model. The k-space is on a 128×128 Cartesian grid. Errors are reported in Tables 4.5 and 4.6.

We observe that the errors decrease with the same trend as in the rectangle case, which strongly suggests that our gold standard is accurate. Meanwhile, for a given sampling density, the errors occurring with the SL phantom are consistently larger than the ones corresponding to the brain phantom. This is explained by the fact that the SL phantom presents edge transitions of larger intensity.

4.4.3 Applications

4.4.3.1 Investigation of Aliasing Artifacts

Let us consider the function $f(\mathbf{u}) = S\rho(\mathbf{Mu})$ which depends on the spatial sampling step matrix $\mathbf{M}$. According to (4.1), the analytical MR data are given by $m_S(\mathbf{k}) = |\mathbf{M}|\hat{f}\,(2\pi\mathbf{Mk})$.

When the benefits of an analytical model are forsaken, the MRI data are

Table 4.5: Rasterized simulation errors versus our analytical simulations with the Brain phantom.

Resolution	NRMSE	max. error in k-space	max. error inverse DFT
128	1.45e-01	1.1e-02	2.3e-01
176	9.26e-02	5.7e-03	1.4e-01
256	5.45e-02	3.3e-03	1.1e-01
352	3.48e-02	2.6e-03	7.9e-02
512	2.13e-02	1.4e-03	5.1e-02
704	1.02e-02	6.6e-04	2.0e-02
1024	6.70e-03	4.1e-04	2.1e-02
1408	4.06e-03	2.4e-04	1.4e-02
2048	2.03e-03	1.5e-04	4.8e-03
2816	1.49e-03	9.5e-05	5.6e-03

Table 4.6: Rasterized simulation errors versus our analytical simulations with the SL phantom.

Resolution	NRMSE	max. error in k-space	max. error inverse DFT
128	2.76e-01	2.9e-02	4.7e-01
176	1.79e-01	1.6e-02	3.0e-01
256	9.74e-02	8.8e-03	1.6e-01
352	5.38e-02	4.9e-03	1.1e-01
512	2.85e-02	2.6e-03	6.5e-02
704	2.01e-02	1.7e-03	3.9e-02
1024	1.28e-02	1.0e-03	3.3e-02
1408	6.23e-03	6.1e-04	1.3e-02
2048	3.34e-03	3.0e-04	7.0e-03
2816	2.03e-03	1.7e-04	5.0e-03

generated from a rasterized version of the phantom and the sensitivity, using the (non-necessarily uniform) discrete Fourier transform (DFT)

$$m_{\mathbf{M}}(\mathbf{k}) = |\mathbf{M}| F\left(\mathrm{e}^{-2\pi \mathrm{j}\mathbf{M}\mathbf{k}}\right), \tag{4.28}$$

with $\|\mathbf{M}\mathbf{k}\|_\infty \leq 1/2$ and

$$F\left(\mathrm{e}^{2\mathrm{j}\pi\boldsymbol{\nu}}\right) = \sum_{\mathbf{p}\in\mathbb{Z}^2} f(\mathbf{p})\mathrm{e}^{-2\mathrm{j}\pi\mathbf{p}\cdot\boldsymbol{\nu}} = \sum_{\mathbf{q}\in\mathbb{Z}^2} \hat{f}\left(\boldsymbol{\nu}+\mathbf{q}\right). \tag{4.29}$$

The right-hand side of (4.29) can be worked out using Poisson's summation formula. The terms with $\mathbf{q} \neq \mathbf{0}$ represent the aliasing that occurs with rasterized simulations. Due to the intrinsically discontinuous nature of the phantom ρ, the Fourier transform $\hat{f}$ decreases slowly, leading to significant aliasing artifacts. However, as the sampling density increases ($\mathrm{Tr}(\mathbf{M}) \to 0$), the impact of aliasing is reduced, as we saw in Section 4.4.2.

Let us define an ideal anti-aliasing filter h in the Fourier domain as

$$\hat{h}(\boldsymbol{\nu}) = \begin{cases} 1 & \text{if} \quad \|\boldsymbol{\nu}\|_\infty \leq 1/2, \\ 0 & \text{otherwise.} \end{cases} \tag{4.30}$$

For normalized frequencies $\boldsymbol{\nu}$ such that $\|\boldsymbol{\nu}\|_\infty \leq 1/2$, the analytical simulation (unaliased) is characterized as the DFT of the samples of the lowpass-filtered continuous signal

$$\hat{f}\left(\boldsymbol{\nu}\right) = \sum_{\mathbf{p}\in\mathbb{Z}^2} \left(h * f\right)(\mathbf{p})\mathrm{e}^{-2\mathrm{j}\pi\mathbf{p}\cdot\boldsymbol{\nu}}, \tag{4.31}$$

where $\left(h * f\right)$ represents the spatial continuous convolution of h and f.

When using a full Cartesian k-space sampling, the classical approach to reconstruction is to perform an inverse DFT. In this case, the samples of the signal f will be perfectly recovered out of the rasterized simulation (4.29) which is not desired because it conceals the existence of the Gibbs phenomenon due to the antialiasing filter (see, for instance, [71]). By contrast, the data provided by our analytical model lead to a fairer reconstruction where the Gibbs phenomenon appears. This effect is illustrated in Figure 4.6.

Counterintuitively, the reconstructions out of rasterized simulations lead to aliasing effects that have a *positive* impact on visual quality. This

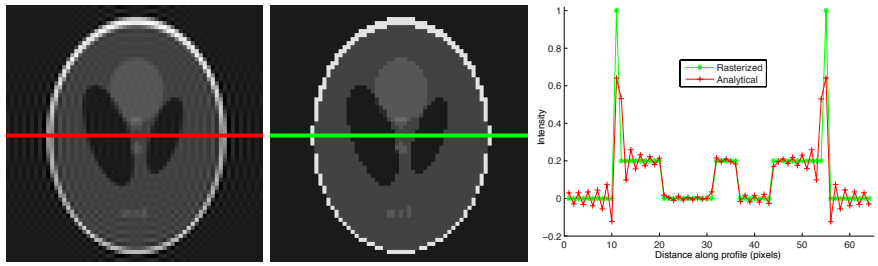

Figure 4.6: 64×64 SL full Cartesian sampling reconstructions. From left to right: analytical simulation, rasterized "inverse crime" simulation, and corresponding line profiles.

situation, which occurs when the same model is used for both simulation and reconstruction, is sometimes referred to as "inverse crime". It arises because of the artificially imposed consistency between the computational forward models used for simulation and reconstruction. In such an inverse-crime situation, the continuous nature of the underlying physical model is not taken into account.

4.4.3.2 Impact of Rasterized Simulations on Reconstruction

We consider a plausible pMRI setup. It involves an array of 8 receiver head-coils that are uniformly distributed around the phantom. The corresponding sensitivity maps are computed thanks to Biot-Savart's law. Spiral and EPI k-space trajectories are considered, both supporting a 256×256 reconstruction matrix with reduction factor $R = 4$. The channel data are generated using our analytical method as well as 256×256 and 512×512 rasterized simulations (see Section 4.4.3.1). The same realization of complex Gaussian noise is added to the simulated data with different intensities, according to three scenarios: very low noise (40 dB SNR), normal data (30 dB SNR), and very noisy data (20 dB SNR). Reconstructions are performed using quadratic (Tikhonov linear solution) and TV regularizations. The reconstruction algorithms exploit the same forward model, in the form of the same encoding matrix $\mathbf{E}$. The experiments only differ in terms of the input data. The regularization parameter is tuned to optimize the SER with respect to the ground-truth phantom (256×256 rasterization of the phantom). We report our results in Table 4.7 for the spiral trajectory and in Table 4.8 for the EPI experiments. Re-

Table 4.7: Reconstruction bias of rasterized simulations on Quadratic and TV regularization reconstructions of the brain phantom for the spiral SENSE experiment. Optimized SER and corresponding bias are shown in dB.

Channel data SNR		40dB		30dB		20dB	
Sampling density		256	512	256	512	256	512
Linear	SER	24.61	19.92	20.31	17.99	14.09	13.45
	Bias	5.07	0.37	2.56	0.24	0.75	0.11
TV	SER	33.75	20.80	27.60	20.26	19.61	17.72
	Bias	13.45	0.49	7.75	0.42	2.43	0.54

Table 4.8: Reconstruction bias of rasterized simulations on Quadratic and TV regularization reconstructions of the brain phantom for the EPI SENSE experiment. Optimized SER and corresponding bias are shown in dB.

Channel data SNR		40dB		30dB		20dB	
Sampling density		256	512	256	512	256	512
Linear	SER	36.25	20.77	26.30	19.79	16.73	15.31
	Bias	16.02	0.54	6.95	0.44	1.61	0.19
TV	SER	42.25	20.98	32.75	20.73	23.92	19.29
	Bias	21.85	0.58	12.57	0.55	5.02	0.39

constructed images are shown in Figures 4.7 and 4.8, together with their error maps, in order to illustrate the impact of the inverse-crime situation (the 256×256 rasterized simulation) in the different scenarios.

The reconstructions in the spiral experiment are penalized compared to the EPI ones, in the sense that the high-frequency corners of the k-space are not sampled which leads to slightly inferior resolution. This explains that, all other parameters remaining constant, the EPI reconstructions outperform the spiral ones qualitatively and quantitatively.

We observe that the reconstructions from rasterized simulations consistently outperform the ones obtained from analytical measurements. While large differences can occur between the inverse-crime scenario (the 256×256 rasterized simulations) and the analytical simulation data, the 512×512 simulations yield much closer performance, with at most a 0.6 dB

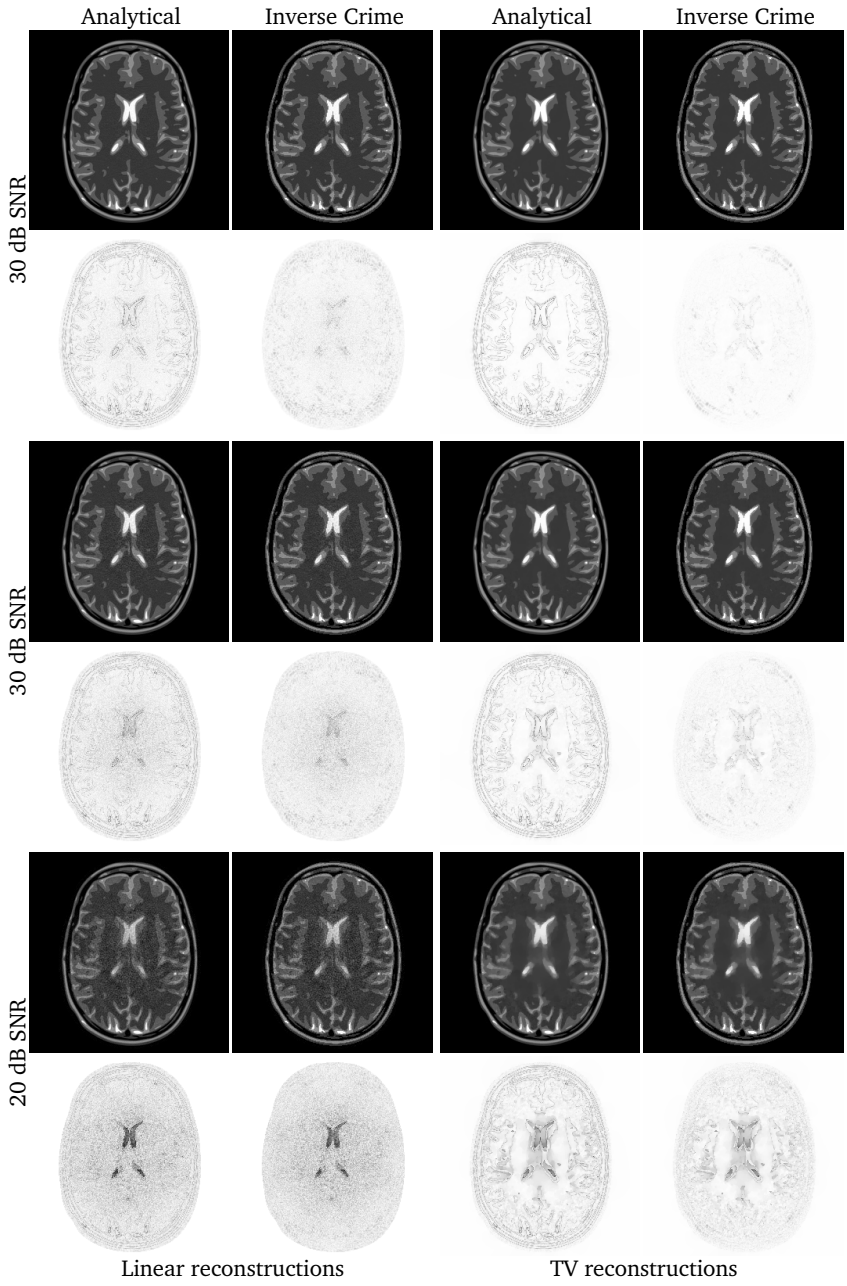

Figure 4.7: Reconstructed brain phantom and error maps for the spiral SENSE experiments.

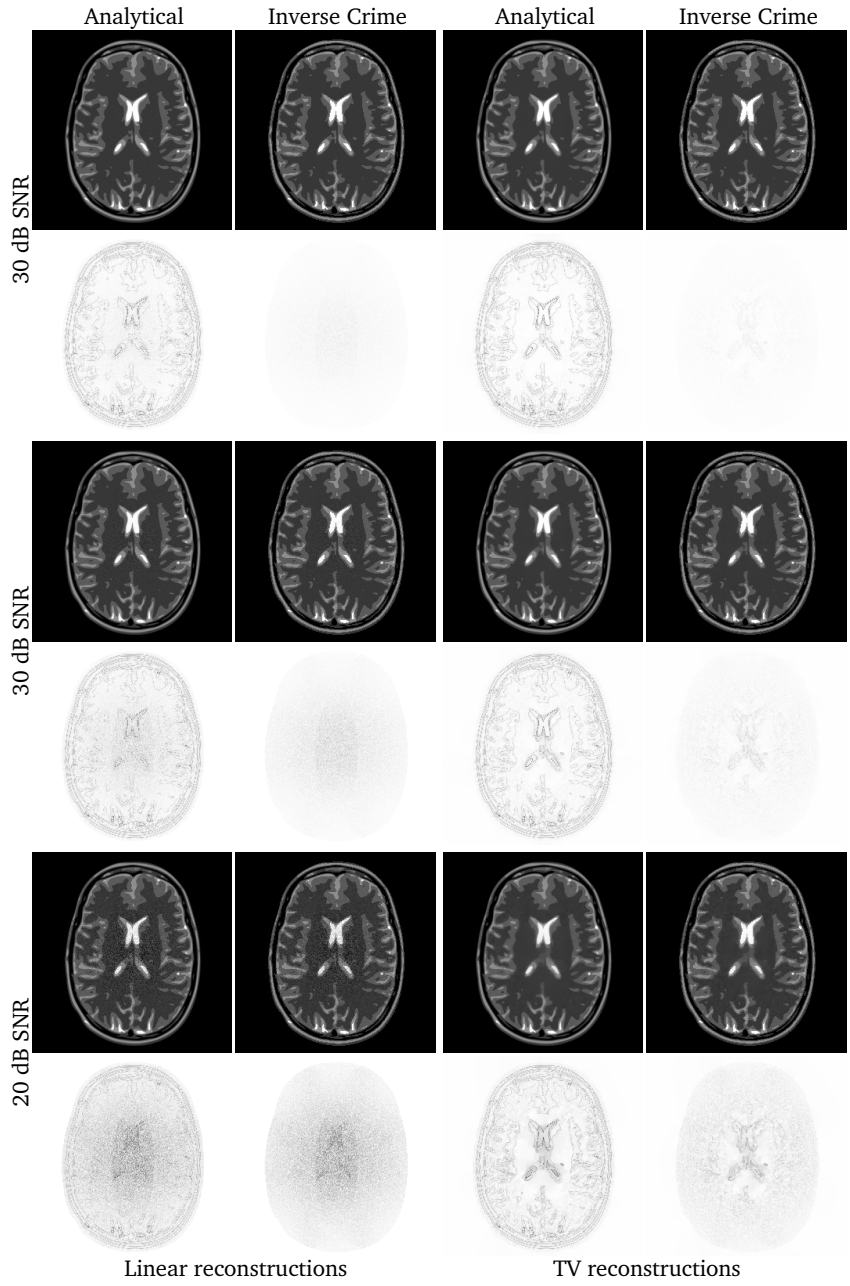

Figure 4.8: Reconstructed brain phantom and error maps for the EPI SENSE experiments.

SER difference. This is explained by the reduced aliasing artifacts when doubling the sampling density (see Section 4.4.3.1). As expected for this type of piecewise-constant phantom, the TV reconstructions consistently outperform the linear ones. Whatever the simulation method is, TV brings a significant improvement in the very noisy scenario. However, for the other scenarios (SNR 30dB and 40dB), the improvement over linear reconstruction is modest when using the analytic measurements, whereas it is artificially spectacular using the 256×256 rasterized simulations. We believe that our quality assessment, obtained analytically, offers fairer predictions of the practical worth of a reconstruction method than its overly optimistic rasterized version.

4.5 Summary

We proposed a method to develop realistic analytical phantoms for parallel MRI. Our analytical phantom approach offers strong advantages for the quantitative validation of MRI and pMRI reconstruction software: it is flexible enough to represent general imaging targets, it provides highly accurate representation of the physical continuous model and avoids overly optimistic reconstructions. This kind of framework is also applicable to the assessment of advanced MRI reconstruction methods such as autocalibrating parallel imaging, B0 correction [2], motion correction [4, 5], or higher order field imaging [7].

Implementations of the phantom are made available to the community.

Chapter 5

Efficient Wavelet-Based Reconstruction

In this chapter, [1] we exploit the fact that wavelets can represent magnetic resonance images well, with relatively few coefficients. We use this property to improve MRI reconstructions from undersampled data with arbitrary k-space trajectories. Reconstruction is posed as an optimization problem that could be solved with the iterative shrinkage/thresholding algorithm (ISTA) [17–19] which, unfortunately, converges slowly. To make the approach more practical, we propose a variant that combines recent improvements in convex optimization and that can be tuned to a given specific k-space trajectory. We present a mathematical analysis that explains the performance of the algorithms. Using simulated and *in vivo* data, we show that our nonlinear method is fast, as it accelerates ISTA by almost two orders of magnitude. We also show that it remains competitive with TV regularization in terms of image quality.

5.1 Motivations

MRI scanners provide data that are samples of the spatial Fourier transform (also known as *k-space*) of the object under investigation. The Shannon-Nyquist sampling theory in both spatial and k-space domains suggests that the sampling density should correspond to the field of view

1. This chapter is partly Copyright ©, 2011, IEEE. Adapted, with permission, from [40] and [72].

(FOV) and that the highest sampled frequency is related to the pixel width of the reconstructed images. However, constraints in the implementation of the k-space trajectory that controls the sampling pattern (*e.g.*, acquisition duration, scheme, smoothness of gradients) may impose locally reduced sampling densities. Insufficient sampling results in reconstructed images with increased noise and artifacts, particularly when applying gridding methods.

The common and generic approach to alleviate the reconstruction problem is to treat the task as an inverse problem [3]. In this framework, ill-posedness due to a reduced sampling density is overcome by introducing proper regularization constraints. They assume and exploit additional knowledge about the object under investigation to robustify the reconstruction.

Earlier techniques used a quadratic regularization term, leading to solutions that exhibit a linear dependence upon the measurements. Unfortunately, in the case of severe undersampling (*i.e.*, locally low sampling density) and depending on the strength of regularization, the reconstructed images still suffer from noise propagation, blurring, ringing, or aliasing errors. It is well known in signal processing that the blurring of edges can be reduced via the use of nonquadratic regularization. In particular, ℓ_1-wavelet regularization has been found to outperform classical linear algorithms such as Wiener filtering in the deconvolution task [17].

Indicative of this trend as well is the recent advent of *Compressed Sensing* (CS) techniques in MRI [6, 9]. These let us draw two important conclusions.

– The introduction of randomness in the design of trajectories favors the attenuation of residual aliasing artifacts because they are spread incoherently over the entire image.

– Nonlinear reconstructions—more precisely, ℓ_1-regularization— outperform linear ones because they impose constraints that are better matched to MRI images.

Many recent works in MRI have focused on nonlinear reconstruction via Total Variation (TV) regularization, choosing finite differences as a sparsifying transform [6, 8, 10, 12]. Nonquadratic wavelet regularization has also received some attention [6, 11, 13, 14, 16, 73, 74], but we are not aware of a study that compares the performance of TV against ℓ_1-wavelet regularization.

Various algorithms have been recently proposed for solving general linear inverse problems subject to ℓ_1-regularization. Some of them deal with an approximate reformulation of the ℓ_1 regularization term. This approximation facilitates reconstruction sacrificing some accuracy and introducing extra degrees of freedom that make the tuning task laborious. Instead, the iterative shrinkage/ thresholding algorithm [17–19] (ISTA) is an elegant and nonparametric method that is mathematically proven to converge. A potential difficulty that needs to be overcome is the slow convergence of the method when the forward model is poorly conditioned (*e.g.*, low sampling density in MRI). This has prompted research in large-scale convex optimization on ways to accelerate ISTA. The efforts so far have followed two main directions:

– generic multistep methods that exploit the result of past iterations to speed up convergence, among them: two-step iterative shrinkage/ thresholding [75] (TwIST), Nesterov schemes [76–78], fast ISTA [20] (FISTA), and monotonic FISTA [79] (MFISTA);

– methods that optimize wavelet-subband-dependent parameters with respect to the reconstruction problem: multilevel thresholded Landweber (MLTL) [80, 81] and subband adaptive ISTA (SISTA) [21].

In this work, we exploit the possibility of combining and tailoring the two generic types of accelerating strategies to come up with a new algorithm that can speedup the convergence of the reconstruction and that can accommodate for every given k-space trajectory. Here, we first consider single-coil reconstructions that do not use sensitivity knowledge. In a second time, we confirm the results with SENSE reconstructions [3].

We propose a practical reconstruction method that turns out to sensibly outperform linear reconstruction methods in terms of reconstruction quality, without incurring the protracted reconstruction times associated with nonlinear methods. This is a crucial step in the practical development of nonlinear algorithms for undersampled MRI, as the problem of fixing the regularization parameter is still open. We also provide a mathematical analysis that justifies our algorithm and facilitates the tuning of the underlying parameters.

This chapter is structured as follows: In Section 5.2, we propose a fast algorithm for solving the nonlinear reconstruction problem and present theoretical arguments to explain its superior speed of convergence. Finally, we present in Section 5.3 an experimental protocol to validate and

compare our practical method with existing ones. We focus mainly on reconstruction time and signal-to-error ratio (SER) with respect to the reference image.

5.2 Wavelet Regularization Algorithms

In this section, we present reconstruction algorithms that handle constraints expressed in the wavelet domain while solving the classical ℓ_1-regularized minimization problem

$$\mathbf{w}^\star = \arg\min_{\mathbf{w}} \mathscr{C}(\mathbf{w}), \tag{5.1}$$

with

$$\mathscr{C}(\mathbf{w}) = \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{w}\|_{\ell_1} \tag{5.2}$$

that is justified in Section 3.3.2.1.

By introducing weighted norms instead of simple Lipschitz constants, we revisit the principle of ISTA algorithm and simplify the derivation and analysis of this class of algorithms. We end up with a novel algorithm that combines different acceleration strategies and we provide a convergence analysis. Finally, we propose an adaptation of the fast algorithm to implement the random-shifting technique that is commonly used to improve results in image restoration.

5.2.1 ISTA with weighted norms

The standard algorithm ISTA is particularly well-suited for this minimization task. Let us recall the key properties of ISTA that are detailed in Section 3.4.4.

In ISTA, each iteration aims at minimizing in a simple shrinkage step a surrogate functional that is locally tailored to the objective. The functional

$$\mathscr{C}'(\mathbf{w}, \mathbf{w}_i) = \mathscr{C}(\mathbf{w}) + \left\|\mathbf{w} - \mathbf{w}_i\right\|_{\mathbf{\Lambda} - \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}}^2 \tag{5.3}$$

satisfies the constraints provided that:

- $\mathbf{\Lambda}$ is diagonal,
- $\mathbf{\Lambda} - \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}$ is positive-definite.

ISTA corresponds to the trivial choice $\mathbf{\Lambda} = \frac{L}{2}\mathbf{I}$ with $L \geq 2\lambda_{\max}\left(\mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}\right)$.

5.2.2 Subband Adaptive ISTA

SISTA is an extension of the ISTA that was introduced by Bayram and Selesnick [21]. Here, we propose an interpretation of SISTA as a particular case of (5.3) with a weighting matrix that replaces advantageously the step size $2/L$. The idea is to use a diagonal weighting matrix $\mathbf{\Lambda}^{-1} = \mathrm{diag}(\boldsymbol{\tau})$ with coefficients that are constant within a wavelet subband.

Accordingly, SISTA is described in Algorithm 11.

Algorithm 11: SISTA solving $\mathbf{w}^\star = \arg\min_{\mathbf{w}} \|\mathbf{m} - \mathbf{M}\mathbf{w}\|_{\mathbf{X}}^2 + \lambda \|\mathbf{w}\|_{\ell_1}$

input : $\mathbf{A} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{M}$, $\mathbf{a} = \mathbf{M}^{\mathrm{H}}\mathbf{X}\mathbf{m}$, $\mathbf{w}_0$, and $\boldsymbol{\tau}$;
$\mathbf{\Lambda}^{-1} \leftarrow \mathrm{diag}(\boldsymbol{\tau})$;
$i \leftarrow 0$;
repeat
  $\mathbf{z}_{i+1} \leftarrow \mathbf{w}_i + \mathbf{\Lambda}^{-1}(\mathbf{a} - \mathbf{A}\mathbf{w}_i)$;
  $\mathbf{w}_{i+1} \leftarrow \mathrm{T}_{\lambda\boldsymbol{\tau}}\left(\mathbf{z}_{i+1}\right)$;
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{w}_i$;

5.2.2.1 Convergence Analysis

By considering the weighted scalar product $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbf{\Lambda}}$ instead of $\frac{L}{2}\langle \mathbf{x}, \mathbf{y} \rangle$, we can adapt the convergence proof of ISTA by Beck and Teboulle (see Proposition 3.4.1). This result is new, to the best of our knowledge.

Proposition 5.2.1. *Let* $\{\mathbf{w}_i\}$ *be the sequence generated by Algorithm 11 with* $\mathbf{\Lambda} - \mathbf{A}$ *positive definite. Then, for any* $i > i_0 \in \mathbb{N}$,

$$\mathscr{C}(\mathbf{w}_i) - \mathscr{C}(\mathbf{w}^\star) \leq \frac{1}{i - i_0} \left\| \mathbf{w}_{i_0} - \mathbf{w}^\star \right\|_{\mathbf{\Lambda}}^2 . \tag{5.4}$$

Therefore, by comparing Propositions 3.4.1 and 5.2.1, an improved convergence is expected. The main point is that, for a "warm" starting point $\mathbf{w}_0$ or after few iterations (i_0), the weighted norm in (5.4) can yield significantly smaller values than the one weighted by $L/2$ in (3.36). The proof is provided in Appendix B.1.

5.2.2.2 Selection of Weights

Bayram and Selesnick [21] provide a method to select the values of $\boldsymbol{\tau}$ for SISTA. To present this result, let us introduce some notations. We denote by s an index that scans all the S wavelet subbands, coarse scale included, by τ_s the corresponding weight constant, and by $\mathbf{M}_s$ the corresponding block of $\mathbf{M}$. We also define $\gamma_{s_1,s_2} = \sqrt{\lambda_{\max}(\mathbf{M}_{s_2}^{\mathrm{H}}\mathbf{M}_{s_1}\mathbf{M}_{s_1}^{\mathrm{H}}\mathbf{M}_{s_2})}$. The authors of [21] show that, for each subband, the condition

$$\frac{1}{\tau_s} > \sum_{s'=1}^{S} \gamma_{s,s'} \tag{5.5}$$

is sufficient to impose the positive definiteness of $(\boldsymbol{\Lambda} - \mathbf{M}^{\mathrm{H}}\mathbf{M})$ that is required in Equation 5.3. In the present context, we propose to compute the values $\gamma_{s,s'}$ by using the power iteration method, once for a given wavelet family and k-space sampling strategy.

5.2.3 Best of Two Worlds: Fast Weighted ISTA

Taking advantage of the ideas developed previously, we derive an algorithm that corresponds to the subband adaptive version of FISTA. In the light of the minimization problem (5.2), FWISTA generalizes the FISTA algorithm using a parametric weighted norm. We give its detailed description in Algorithm 12, where the difference with respect to FISTA resides in using the SISTA step.

Algorithm 12: FWISTA

input : $\mathbf{A}$, $\mathbf{a}$, $\mathbf{w}_0$, and $\boldsymbol{\tau}$;
Initialization: $i = 0$, $\mathbf{v}_0 = \mathbf{w}_0$, $t_0 = 1$, $\boldsymbol{\Lambda}^{-1} = \mathrm{diag}(\boldsymbol{\tau})$;
repeat
  $\mathbf{w}_{i+1} \leftarrow \mathrm{T}_{\lambda\boldsymbol{\tau}}\left(\mathbf{v}_i + \boldsymbol{\Lambda}^{-1}(\mathbf{a} - \mathbf{A}\mathbf{v}_i)\right)$ (SISTA step);
  $t_{i+1} \leftarrow \left(1 + \sqrt{1 + 4t_i^2}\right)/2$;
  $\mathbf{v}_{i+1} \leftarrow \mathbf{w}_{i+1} + (t_i - 1)\left(\mathbf{w}_{i+1} - \mathbf{w}_i\right)/t_{i+1}$;
  $i \leftarrow i + 1$;
until *desired tolerance is reached*;
return $\mathbf{w}_i$;

In the same fashion as for SISTA, we revisit the convergence results of FISTA [20, Thm. 4.4] for FWISTA.

Proposition 5.2.2. *Let* $\{\mathbf{w}_i\}$ *be the sequence generated by Algorithm 12 with* $\mathbf{\Lambda} - \mathbf{A}$ *positive definite. Then, for any* $i \geq 1$*,*

$$\mathscr{C}(\mathbf{w}_i) - \mathscr{C}(\mathbf{w}^\star) \leq \left(\frac{2}{i+1}\right)^2 \left\|\mathbf{w}_0 - \mathbf{w}^\star\right\|_{\mathbf{\Lambda}}^2. \tag{5.6}$$

If $\mathscr{C}$ *is* ϵ*-strongly convex, i.e.* $\epsilon = 2\lambda_{min}(\mathbf{A}^{\mathrm{T}}\mathbf{A}) > 0$*, then*

$$\left\|\mathbf{w}_i - \mathbf{w}^\star\right\|_2 \leq 2\sqrt{\frac{2}{\epsilon}} \frac{\left\|\mathbf{w}_0 - \mathbf{w}^\star\right\|_{\mathbf{\Lambda}}}{i+1}. \tag{5.7}$$

The proof is provided in Appendix B.2.

This result shows the clear advantage of FWISTA compared to ISTA (Proposition 3.4.1) and SISTA (Proposition 5.2.1). Moreover, we note that FWISTA can be simply adapted in order to impose a monotonic decrease of the cost functional value, in the same fashion as MFISTA. The same convergence properties apply [79, Thm. 5.1].

5.2.4 Random Shifting

Wavelet bases perform well the compression of signals but can introduce artifacts that can be attributed to their relative lack of shift-invariance. In the case of regularization, this can be avoided by switching to a redundant dictionary. The downside, however, is a significant increase in computational cost. Alternatively, the practical technique referred to as random shifting (RS) [17] can be used. Applying random shifting is much simpler and computationally more efficient than considering redundant transforms and leads to sensibly improved reconstruction.

Here, we propose a variational interpretation that motivates our implementation of FWISTA with RS (see Algorithm 13). We consider the DWT $[\mathbf{W}_1 \cdots \mathbf{W}_{N_s}]^{\mathrm{H}}$, with $\mathbf{W}_s = \mathbf{S}_s\mathbf{W}$, where $\mathbf{S}_s$ represent the different shifting operations required to get a translation-invariant DWT. The desired reconstruction would be defined as the minimizer of

$$\mathscr{C}(\mathbf{c}) = \|\mathbf{m} - \mathbf{E}\mathbf{c}\|_{\ell_2}^2 + \frac{\lambda}{N_s} \left\|[\mathbf{W}_1 \cdots \mathbf{W}_{N_s}]^{\mathrm{H}}\mathbf{c}\right\|_{\ell_1}. \tag{5.8}$$

In 1-D , this formulation includes TV regularization, that is the ℓ_1-norm of a single-level undecimated Haar wavelet transform without coarse-scale.

Rewriting (5.8) in terms of wavelet coefficients, we get

$$N_s \, \mathscr{C}(\mathbf{c}) = \sum_s \mathscr{C}_s(\mathbf{W}_s^{-1}\mathbf{c}), \tag{5.9}$$

with

$$\mathscr{C}_s(\mathbf{w}_s) = \left\|\mathbf{m} - \mathbf{M}\mathbf{S}_s^{-1}\mathbf{w}_s\right\|_{\ell_2}^2 + \lambda \left\|\mathbf{w}_s\right\|_{\ell_1}. \tag{5.10}$$

For a current estimate, we select a transform $\mathbf{W}_i$ and perform a step towards the minimization of the cost with respect to $\mathbf{w}_s$ while keeping $\mathbf{w}_{s'}$, for $s' \neq s$ fixed. As the minimization subproblem (5.10) takes the form (5.2), a SISTA iteration is appropriate. It is expected to have the minimizers of the functionals $\mathscr{C}_s$ and the minimizer of (5.8) look very similar. In the first iterations of the algorithm, the minimization steps with respect to any $\mathbf{w}_s$ are functionally equivalent (*i.e.*, the modification is mostly explained by the gradient step). This is why FWISTA can be used to speedup the first iterations. When the solution gets close to the solution ISTA steps are used.

As the scheme is intrinsically greedy, we do not have a theoretical guarantee of convergence. Yet, in practice, we have observed that the SER stabilizes at a much higher value than it does when using ISTA schemes without RS (see Figure 5.4).

Our method is described in Algorithm 13. Note that it has no more matrix-vector multiplications per iteration than ISTA, SISTA and FISTA. Therefore, the computational cost of an iteration is expected to be equivalent.

5.3 Experiments

5.3.1 Implementation Details

Our implementation uses Matlab 7.9 (Mathworks, Natick). The reconstructions run on a 64-bit 8-core computer, clock rate 2.8 GHz, 8 GiB RAM (DDR2 at 800 MHz), Mac OS X 10.6.5. For all iterative algorithms, a key point is that matrices are not stored in memory. They only represent operations that are performed on vectors (images). In particular, $\mathbf{a} = \mathbf{M}^{\mathrm{H}}\mathbf{m}$ is computed once per dataset. Matrix-to-vector multiplication with $\mathbf{A} = \mathbf{M}^{\mathrm{H}}\mathbf{M}$, specifically, $\mathbf{E}^{\mathrm{H}}\mathbf{E}$, have an efficient implementation thanks to the convolution structure of the problem [49, 82]. For these Fourier

Algorithm 13: Our method (uses FWISTA+RS and moves to ISTA+RS in the neighborhood of the solution)

input : $\mathbf{A} = \mathbf{E}^{\mathrm{H}}\mathbf{E}$, $\mathbf{a} = \mathbf{E}^{\mathrm{H}}\mathbf{m}$, $\mathbf{c}_0$, $\mathbf{\Lambda}^{-1} = \mathrm{diag}(\boldsymbol{\tau})$, K;
Consider the sequence of DWT with RS: $\mathbf{W}_i$;
Initialization: $i = 0$, $k = 0$, $\gamma = 1$, $t_0 = 1$, $C_0 = \mathscr{C}(\mathbf{c}_0)$, $\mathbf{y}_0 = \mathbf{c}_0$;
repeat
  $\mathbf{w}_i \leftarrow {\mathbf{W}_i}^{-1}\mathbf{y}_i$;
  $\mathbf{c}_{i+1} \leftarrow \mathbf{W}_i \mathrm{T}_{\lambda\tau}\left(\mathbf{w}_i + \mathbf{\Lambda}^{-1}\mathbf{W}_i^{-1}(\mathbf{a} - \mathbf{A}\mathbf{y}_i)\right)$;
  $C_{i+1} \leftarrow \mathscr{C}(\mathbf{c}_i)$ (cost-function evaluation);
  if $C_{i+1} > C_i$ **then**
    $k \leftarrow k + 1$;
  if $k = K$ **then**
    $\gamma \leftarrow 0$;
    $\mathbf{\Lambda}^{-1} \leftarrow \max\{\boldsymbol{\tau}\}\mathbf{I}$;
    $\boldsymbol{\tau} \leftarrow \max\{\boldsymbol{\tau}\}\mathbf{1}$;
  $t_{i+1} \leftarrow \left(1 + \sqrt{1 + 4t_i^2}\right)/2$;
  $\mathbf{y}_{i+1} \leftarrow \mathbf{c}_{i+1} + \gamma\frac{t_i - 1}{t_{i+1}}(\mathbf{c}_{i+1} - \mathbf{c}_i)$;
  $i \leftarrow i + 1$;
until *stopping condition is met*;
return $\mathbf{c}_i$;

precomputations, we used the NUFFT algorithm [48] that is made available online. [2] For wavelet transforms, we used the code provided online [3] by the authors of [81]. This Fourier-domain implementation proved to be faster than Matlab's when considering reconstructed images smaller than 256×256 and the Haar wavelet. It must also be noted that the 2-D DFT was performed using the FFTW library which efficiently parallelizes computations.

For Tikhonov regularizations, we implemented the classical conjugate gradient (CG) algorithm, with the identity as the regularization matrix. For TV regularizations, we considered the iteratively reweighted least-squares algorithm (IRLS), which corresponds to the additive form of half-quadratic minimization [51, 70]. We used 15 iterations of CG to solve the

2. Available at `http://www.eecs.umich.edu/~fessler/code/` .
3. Available at `http://bigwww.epfl.ch/algorithms/mltldeconvolution` .

linear inner problems, always starting from the current estimate, which is crucial for efficiency. For the weights that permit the quadratic approximation of the TV term, we stabilized the inversion of very small values.

We implemented ISTA, SISTA, and FWISTA as described in Section 5.2, with the additional possibility to use random shifting (see Section 5.3.3.2). For the considered reconstructions using our method, described in Algorithm 13, $K = 30$ was a reasonable choice. The Haar wavelet transform was used, with 3 decomposition levels when no other values are mentioned. As is usual for wavelet-based reconstructions, the regularization was not applied to the coarse-level coefficients.

Reconstructions were limited to the pixels of the ROI for all algorithms. The regularization parameter λ was systematically adjusted such that the reconstruction mean-squared error (MSE) inside the ROI was minimal. For practical situations where the ground-truth reference is not available, it is possible to adjust λ by considering well-established techniques such as the discrepancy principle, generalized cross validation, or L-curve method [83].

5.3.2 Spiral MRI Reconstruction

In this section, we focused on the problem of reconstructing images of objects weighted by the receiving channel sensitivity, given undersampled measurements. This problem, which involves single-channel data and hence differs from SENSE, is challenging for classical linear reconstructions as it generates artifacts and propagates noise. We considered spiral trajectories with 50 interleaves, with an interleave sampling density reduced by a factor $R = 1.8$ compared to Nyquist for the highest frequencies and an oversampling factor 3.5 along the trajectory. Spiral acquisition schemes are attractive because of their versatility and the fact that they can be implemented with smooth gradient switching [84, 85].

We validate the results with the three sets of data that we present below. The corresponding reference images are shown in Figure 5.1.

5.3.2.1 MR Scanner Acquisitions

The data were collected on a 3T Achieva system (Philips Medical Systems, Best, The Netherlands). A field camera with 12 probes was used to monitor the actual k-space trajectory [86]. An array of 8 head coils pro-

vided the measurements. We acquired *in vivo* brain data from a healthy volunteer with parameters $T_R = 1000\,\text{ms}$ and $T_E = 30\,\text{ms}$. The excitation slice thickness was 3 mm with a flip angle of 30°. The trajectory was designed for a FOV of 25 cm with a pixel size 1.5 mm. It was composed of 100 spiral interleaves. The interleaf distance for the highest frequencies sampled defined a fraction of the Nyquist sampling density ($R = 0.9$).

The subset used for reconstruction corresponds to half of the 100 interleaves. The corresponding reduction factor, defined as the ratio of the distance of neighboring interleaves with the Nyquist distance, is $R = 1.8$.

5.3.2.2 Analytical Simulation

We used analytical simulations of the Shepp-Logan (SL) brain phantom with a similar coil sensitivity, following the method described in [87]. The values of these simulated data were scaled to have the same mean spatial value (*i.e.*, the same central k-space peak) as the brain reference image. A realization of complex Gaussian noise was added to this synthetic k-space data, with a variance corresponding to 40 dB SNR. The 176×176 rasterization of the analytical object provided a reliable reference for comparisons.

5.3.2.3 Simulation of a Textured Object

A second simulation was considered with an object that is more realistic than the SL phantom. We chose a 512×512 MR image of a wrist that showed little noise and interesting textures. We simulated acquisitions with the same coil profile and spiral trajectory (176×176 reconstruction matrix), in presence of a 40 dB SNR Gaussian complex noise. The height of the central peak was also adjusted to correspond to that of the brain data. The reference image was obtained by sinc-interpolation, by extracting the lowest frequencies in the DFT.

5.3.3 Results

In this section, we present the different experiments we conducted. The two main reconstruction performance measures that we considered are

– **Reconstruction duration**, which excludes all aforementioned precomputations and the superfluous monitoring operations.

Figure 5.1: Reference images from left to right: *in vivo* brain, SL reference, and wrist.

– **Signal to error ratio** with respect to a reference, defined as

$$\mathrm{SER}(\widetilde{\rho}, \rho_{\mathrm{ref}}) = \left\|\rho_{\mathrm{ref}}\right\|_{\ell_2} / \left\|\rho_{\mathrm{ref}} - \widetilde{\rho}\right\|_{\ell_2}$$

and $\mathrm{SERdB}(\widetilde{\rho}, \rho_{\mathrm{ref}}) = 20\log_{10}\big(\mathrm{SER}(\widetilde{\rho}, \rho_{\mathrm{ref}})\big)$. Practically, the references are either the ground-truth images or the minimizer of the cost functional. It is known that SER is not a foolproof measure of visual improvement but large SER values are encouraging and generally correlate with good image quality.

5.3.3.1 Convergence Performance of IST-Algorithms

In this first experiment, we compared the convergence properties of the different ISTA-type algorithms, as presented in Section 5.2, with the Haar wavelet transform. The data we considered are those of the MR wrist image. The regularization parameter was adjusted to maximize the reconstruction SER with respect to the ground-truth data. The actual minimizer of the cost functional, which is the common fixed-point of this family of algorithms, was estimated by iterating FWISTA 100 000 times.

The convergence results are shown in Figures 5.2 and 5.3 for the simulation of the MR wrist image. Similar graphs are obtained using the other sets of data.

For a fixed number of iterations, FISTA schemes (FISTA and FWISTA) require roughly 10% additional time compared to ISTA and SISTA. In spite of this fact, their asymptotic superiority appears clearly in both figures. The slope of the decrease of the cost functional in the log-log plot of Figure 5.3 reflects the convergence properties in Propositions 3.4.1, 5.2.1,

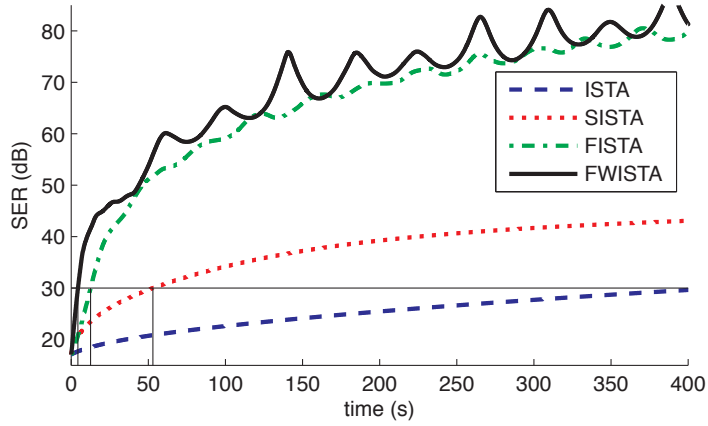

Figure 5.2: Time evolution of the SER with respect to the minimizer for several ISTA algorithms. Times to reach 30 dB are delineated.

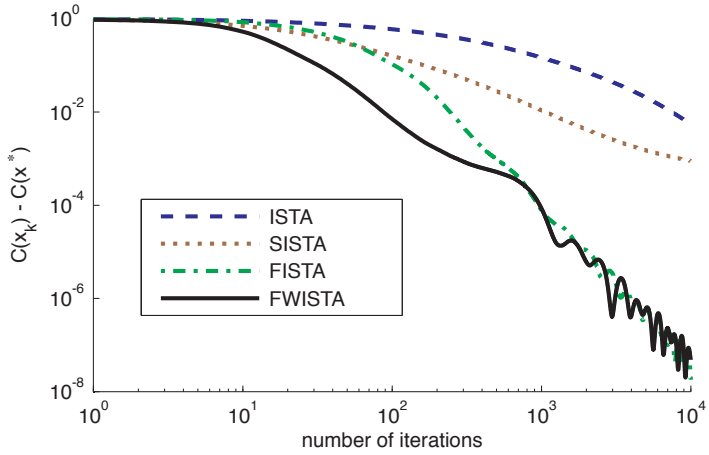

Figure 5.3: Time evolution of the difference in cost function value with respect to the minimizer for several ISTA algorithms.

Table 5.1: Values of the optimal SER and corresponding regularization parameters are shown for the different wavelet bases.

	Wavelet Basis	Without RS		With RS	
		λ opt.	SER (dB)	λ opt.	SER (dB)
SL	Haar	1 870	**12.65**	5 650	**13.38**
	Spline 2	2 510	12.16	3 900	12.53
	Spline 4	830	10.75	7 770	11.58
	Spline 6	1 460	9.70	1 370	10.38
Wrist	Haar	1 600	15.93	1 490	**18.70**
	Spline 2	946	**17.33**	850	18.24
	Spline 4	1 070	17.32	1 190	18.05
	Spline 6	1 350	17.07	1 260	17.87

and 5.2.2. When considering the first iterations, which are of greatest practical interest, the algorithms with optimized parameters (SISTA and FWISTA) perform better than ISTA and FISTA (see Figure 5.3). The times required by each algorithm to reach a 30 dB SER (considered as a threshold value to perceived changes) are 415 s (ISTA), 53 s (SISTA), 12.7 s (FISTA), and 4.4 s (FWISTA). With respect to this criterion, SISTA presents an 8-fold speedup over ISTA, while FWISTA presents a 12-fold speedup over SISTA and nearly a 3-fold speedup over FISTA. It follows that FWISTA is practically close to two orders of magnitude faster than ISTA.

5.3.3.2 Choice of the Wavelet Transform and Use of Random Shifting

The algorithms presented in Section 5.2 apply for any orthogonal wavelet basis. For the considered application, we want to study the influence of the basis on performance. In this experiment we considered the Battle-Lemarié spline wavelets [42] with increasing degrees, taking into account the necessary postfilter mentioned in (3.5).

We compared the best results for several bases. They were obtained with FWISTA after practical convergence and are reported in Table 5.1. Figure 5.4 illustrates the time evolution of the SER using ISTA and FWISTA in the case of the SL reconstruction. Similar graphs are obtained with the other sets of data.

It is known that the Haar wavelet basis efficiently approximates piecewise-

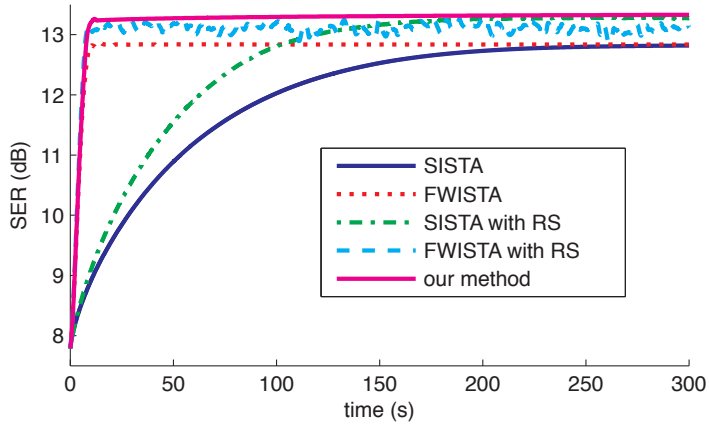

Figure 5.4: Time evolution of the SER for several algorithms for the SL simulation using Haar wavelets.

constant objects like the SL phantom, which is consistent with our results. On the other hand, splines of higher degree, which have additional vanishing moments, perform better on the textured images (upper part of Table 5.1).

We present in the lower part of Table 5.1 the performances observed when using ISTA with RS. We conclude that, in the case of realistic data, it is crucial to use RS as it improves results by at least 0.7 dB, whatever the wavelet basis is. The remarkable aspect there is that the Haar wavelet transform with RS consistently performs best. Two important things can be seen in Figure 5.4: FWISTA is particularly efficient during the very first iterations, while SISTA with RS yields the best asymptotic results in terms of SER and stability. Our method combines both advantages.

In Table 5.2, we present the results obtained using different depths of the wavelet decompositions. Our reconstruction method is used together with the Haar wavelet transform and RS. The performances are similar but there seems to be an advantage in using several decomposition levels both in terms of SER and reconstruction speed. The FWISTA scheme seems to recoup the cost of the wavelet transform operations associated to an increase in the depth of decomposition.

Table 5.2: Results of the proposed wavelet method for different wavelet decomposition depths. Values of the regularization parameter, the final SER, the relative maximal spatial domain error, and the time to reach $-0.5\,\mathrm{dB}$ of the final SER.

	Depth	λ opt.	SER (dB)	ℓ_∞ error (%)	$t_{-0.5\,\mathrm{dB}}$ (s)
SL	1	5 330	13.25	**50**	6.56
	2	5 110	13.34	51	**6.19**
	3	5 680	**13.39**	54	6.93
	4	5 200	13.35	53	7.59
Wrist	1	1 500	18.54	**15.5**	5.88
	2	1 520	18.66	16	**5.23**
	3	1 570	**18.71**	16	5.33
	4	1 700	18.70	16.5	4.25
Brain	1	9 650	18.78	17.3	8.99
	2	10 800	**19.00**	17.5	7.03
	3	11 400	18.99	17.49	**6.33**
	4	11 400	18.96	**17.0**	7.97

5.3.3.3 Practical Performance

We report in Table 5.3 the results obtained for different reconstruction experiments using state-of-the-art linear reconstruction, TV regularization, and our method. The images obtained when running the different algorithms after approximately 5 s, and after practical convergence as well, are shown in Figure 5.5. We display in Figure 5.6 the time evolution of the SER for the different experiments. In each case, we emphasize the time required to reach $-0.5\,\mathrm{dB}$ of the asymptotic value of SER. Finally, we present in Figure 5.7 the reconstructions and in Figure 5.8 the error maps of the different IST-algorithms at different moments of reconstruction. This was done with the wrist simulated experiment using the Haar wavelet basis and RS.

Firstly, we observe that TV and our method achieve similar SER (Table 5.3) and image quality (Figure 5.5). They both clearly outperform linear reconstruction, with a SER improvement from 1.5 to 3 dB, depending on the degree of texture in the original data. Moreover, the pointwise maximal reconstruction error appears to always be smaller with nonlinear reconstructions. Due to the challenging reconstruction task, which

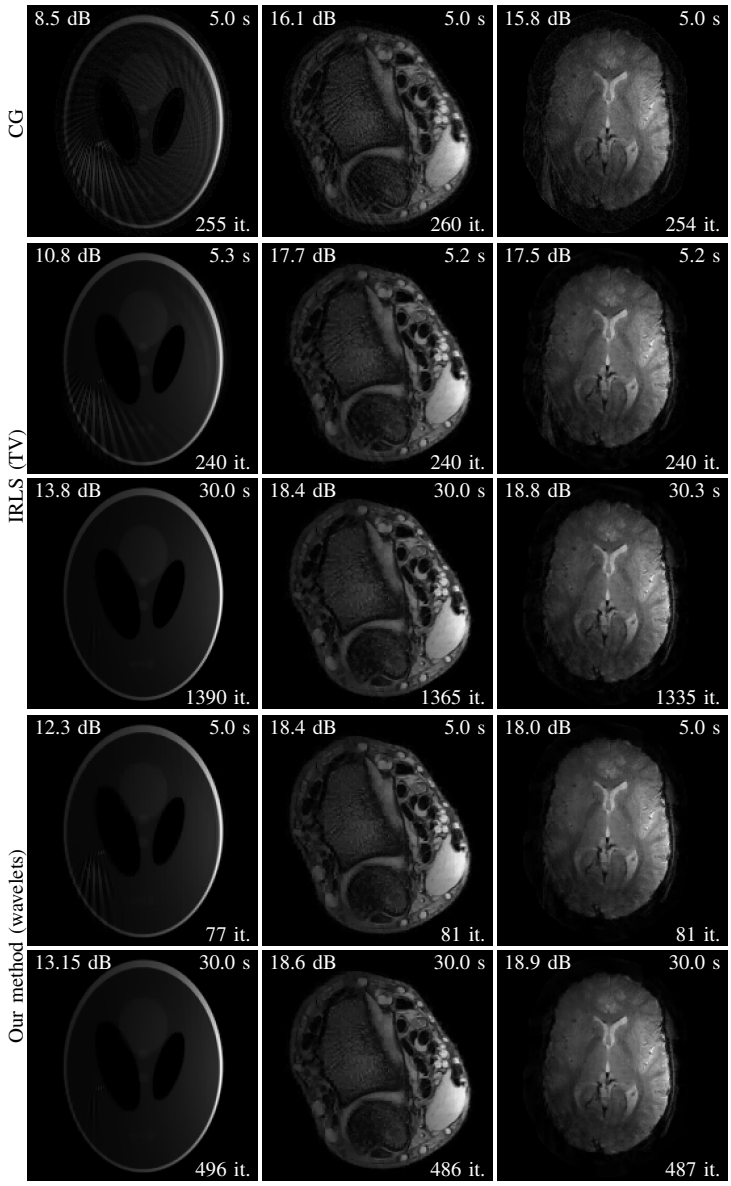

Figure 5.5: Result of different reconstruction algorithms for the three experiments. For each reconstruction, the performance in SER with respect to the reference (top-left), the reconstruction time (top-right) and the number of iterations (bottom-right) are shown.

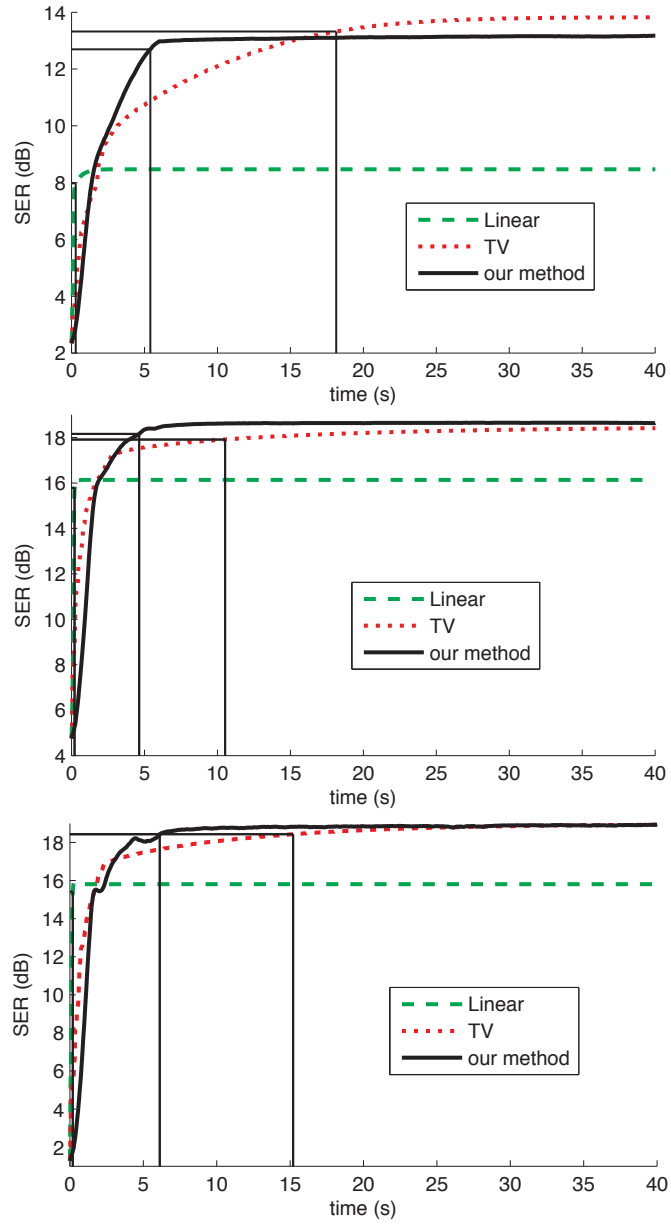

Figure 5.6: Evolution of the performance of the algorithms. From top to bottom: SL simulation, wrist simulation, and brain data. Times required to reach -0.5 dB of the asymptotic value are indicated.

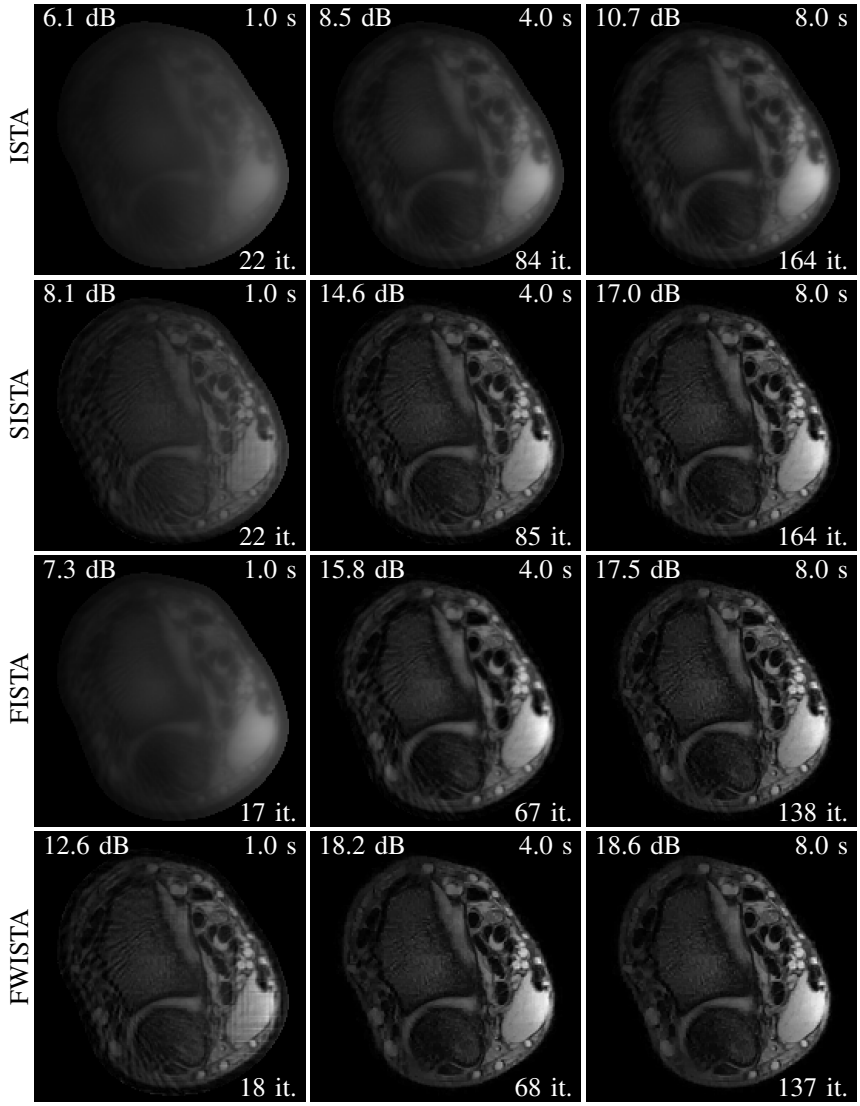

Figure 5.7: Reconstructions of different IST-algorithms with RS for the wrist experiment. For each reconstruction, the performance in SER with respect to the reference (top-left), the reconstruction time (top-right) and the number of iterations (bottom-right) are shown.

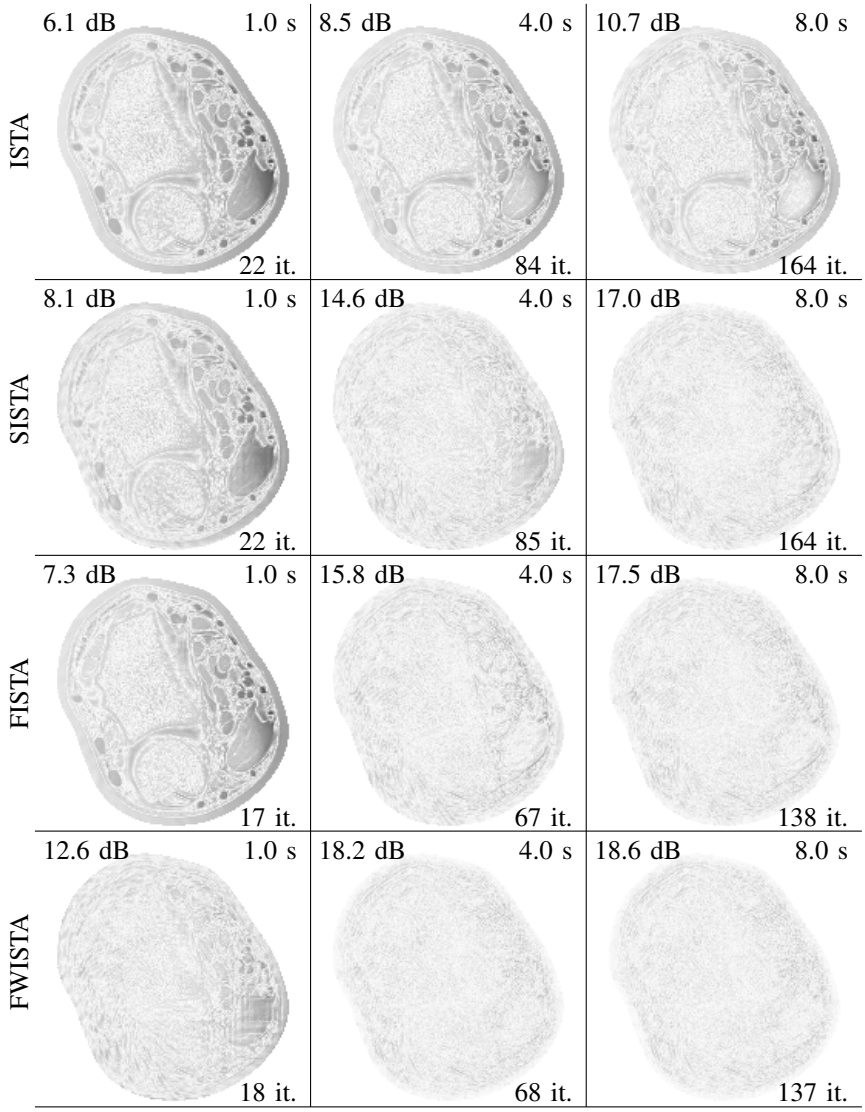

Figure 5.8: Error maps of different IST-algorithms with RS for the wrist experiment. The performance in SER with respect to the reference (top-left), the reconstruction time (top-right) and the number of iterations (bottom-right) are shown.

Table 5.3: Results of the algorithms CG (linear), IRLS (TV), and our method (wavelets). Values of the regularization parameter, the final SER, the relative maximal spatial domain error, and the time to reach $-0.5\,\mathrm{dB}$ of the final SER.

	Method	λ opt.	SER (dB)	ℓ_∞ error (%)	$t_{-0.5\,\mathrm{dB}}$ (s)
SL	linear	0.0247	8.46	**48**	**0.286**
	TV	4 090	**13.82**	49	18.1
	wavelets	6 380	13.17	51	5.40
Wrist	linear	0.436	16.14	21	**0.209**
	TV	760	18.41	**16**	10.5
	wavelets	1 620	**18.64**	**16**	4.64
Brain	linear	0.471	15.81	29	**0.205**
	TV	6 050	18.88	12	15.2
	wavelets	16 800	**18.93**	**11**	6.13

significantly undersamples of the k-space, residual artifacts remain in the linear reconstructions and at early stages of the nonlinear ones. Although the k-space trajectory is exactly the same in the three cases, artifacts are less perceived in the *in vivo* reconstructions, while they stand out for the synthetic experiments.

Secondly, it clearly appears that the linear reconstruction, implemented with CG, leads to the fastest convergence, unfortunately with suboptimal quality. For a reconstruction time one order of magnitude longer, our accelerated method provides better reconstructions. This is illustrated in Figure 5.5 for reconstruction times of the order of five seconds (rows 1, 2, and 4).

Finally, we observe in Figures 5.7 and 5.8 the superiority of the proposed FWISTA over the other IST-algorithms. For the given reconstruction times, it consistently exhibits better image quality as can be seen in both reconstructions and error maps.

5.3.4 SENSE MRI Reconstruction

Our reconstruction method is applicable to linear MR imaging modalities. In this section, we report results obtained on a SENSE reconstruction problem.

5.3.4.1 Sensitivity Weighted Reconstruction

For illustration, let us consider a simple MRI problem with one receiving channel and a full k-space Cartesian sampling. In such a case, the measurements are expressed as $\mathbf{m} = \mathbf{E}\mathbf{c}_{\text{orig}} + \mathbf{b}$. The system matrix is $\mathbf{E} = \mathbf{F}\mathbf{S}$ where $\mathbf{F}$ corresponds to a unitary Fourier matrix and $\mathbf{S}$ is the diagonal receiving sensitivity matrix.

Applying FISTA with $L = 2\tau^{-1}\|\mathbf{S}\|_\infty^2$ and $\tau \approx 1$ results in a gradient step of the form $\mathbf{c}_i + \tau\|\mathbf{s}\|_\infty^{-2}\mathbf{S}^{\text{H}}\mathbf{S}(\mathbf{c}_{\text{orig}} - \mathbf{c}_i) + \tau\|\mathbf{s}\|_\infty^{-2}\mathbf{S}^{\text{H}}\mathbf{F}^{\text{H}}\mathbf{b}$. The latter tends to be ineffective for the pixels whose locations correspond to a relatively small sensitivity, this fact being measured by the condition number of $\mathbf{S}$; i.e., the product $\|\mathbf{S}\|_\infty \left\|\mathbf{S}^{-1}\right\|_\infty$. In regard to the convergence result in Proposition 5.2.2, we see that the constant term of the upper bound for FISTA is proportional to $\|\mathbf{S}\|_\infty^2$ as it amounts to $\|\mathbf{S}\|_\infty^2 \left\|\mathbf{c}_0 - \mathbf{c}^\star\right\|_2^2$.

FWISTA, by contrast, can take advantage of step sizes (gradient and threshold) that are tailored to the individual pixel sensitivities. In our simple example, we apply FWISTA with $\mathbf{\Lambda} = \tau^{-1}\mathbf{S}^{\text{H}}\mathbf{S}$. This results into the simplified gradient step $\mathbf{c}_i + \tau(\mathbf{c}_{\text{orig}} - \mathbf{c}_i) + \tau\mathbf{S}\mathbf{F}^{\text{H}}\mathbf{b}$. For the convergence bound, we achieve a constant that is $\tau^{-1}\left\|\mathbf{c}_0 - \mathbf{c}^\star\right\|_{\mathbf{S}^{\text{H}}\mathbf{S}}^2$ and tends be significantly smaller than the one obtained with FISTA, especially if the range of sensitivity values is large.

That is why we propose, in this section, to adapt the threshold and step-sizes in FWISTA depending both on the spatial localization and the wavelet subband of the coefficients.

5.3.4.2 Synthetic Data

We first considered a 2-D brain imaging setup. The data are recorded by four head coils with known sensitivity maps distributed around the sample. Meanwhile, a radial k-space trajectory with 90 lines that supports a 176×176 reconstruction is imposed. Our simulation is achieved with a 704×704 rasterized version of the SL phantom. We are also using realistic coil sensitivities computed using Biot-Savart's law. The 2-levels Haar wavelet basis is chosen to impose the sparsity constraints. The sum-of-squares sensitivity is considered for the weights of FWISTA. We exploit the localization properties of the wavelets to impose the weights on the wavelet coefficients. The wavelet regularization parameter is the same across the wavelet subbands except for the coarse coefficients where the

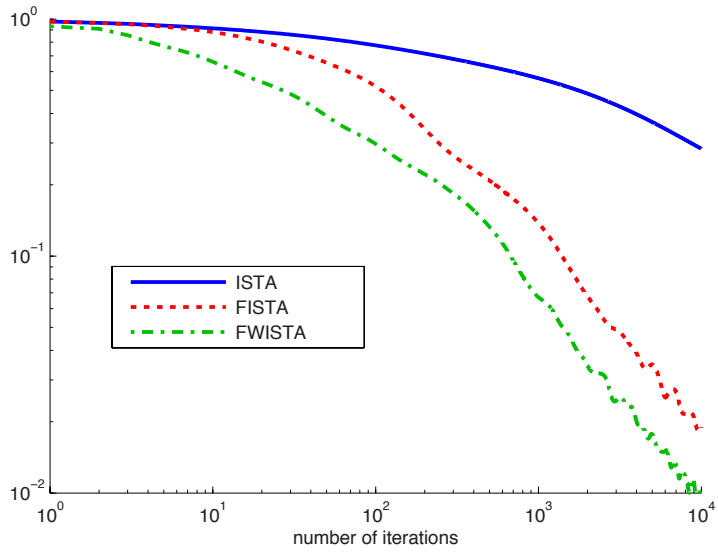

Figure 5.9: SENSE MRI simulation. $\left\|\mathbf{x}_k - \mathbf{x}^\star\right\|_2$.

value 0 is imposed. The value λ is optimized for MSE performance. The initial estimate is $\mathbf{c}_0 = \mathbf{0}$.

We compare the convergence speed of ISTA, FISTA, and FWISTA in the reconstruction task described above. The reference minimizer was determined by running FISTA for 100 000 iterations. Results are shown in Figures 5.9 and 5.10.

We observe that FWISTA yields nearly 3-fold acceleration over FISTA in terms of cost functional. The asymptotic rates of FISTA and FWISTA both on the cost functional value and the distance to the minimizer are similar and the speedup of FWISTA seems primarily due to a better constant, which is consistent with our theoretical prediction.

5.3.4.3 Scanner Data

The data were acquired with the same scanner setup as in Section 5.3.2.1. This time, the data from the 8 receiving channels were used for reconstruction, as well as an estimation of the sensitivity maps. An *in vivo* gradient echo EPI sequence of a brain was performed with T_2^* contrast. The data was acquired with the following parameters: excitation slice

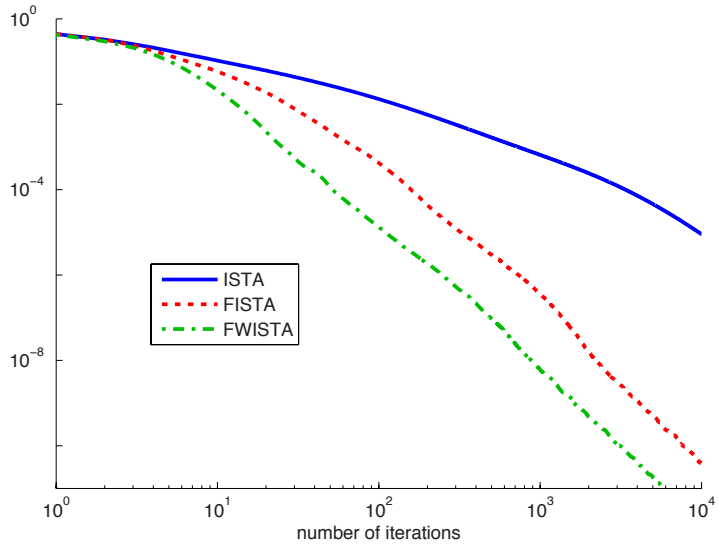

Figure 5.10: SENSE MRI simulation. $\mathscr{C}(\mathbf{x}_k) - \mathscr{C}(\mathbf{x}^\star)$.

thickness of 4 mm, $T_E = 35$ ms, $T_R = 900$ ms, flip angle of 80°, and trajectory composed of 13 interleaves, supporting a 200×200 reconstruction matrix with pixel resolution 1.18 mm × 1.18 mm. The oversampling ratio along the readout direction was 1.62.

The reference image was obtained using the complete set of data and performing an unregularized CG-SENSE reconstruction. The reconstruction involved 3 of the 13 interleaves, representing a significant undersampling ratio $R = 4.33$.

The images obtained using regularized linear reconstruction (CG), TV (IRLS), and our method are presented in Figure 5.11. In Figure 5.12, the SER evolution with respect to time is shown for the three methods. The times to reach -0.5 dB of the asymptotic SER value are 5.9 s (CG), 49.1 s (IRLS), and 22.8 s (our method).

With this high undersampling, the errors maps show that reconstructions suffer from noise propagation mostly in the center of the image. It appears that TV and our method improve qualitatively and quantitatively image quality over linear reconstruction (see Figure 5.11). As it was observed in Section 5.3.2 with the spiral MRI reconstructions, this *in vivo* SENSE experiment confirms that our method is competitive with TV. In

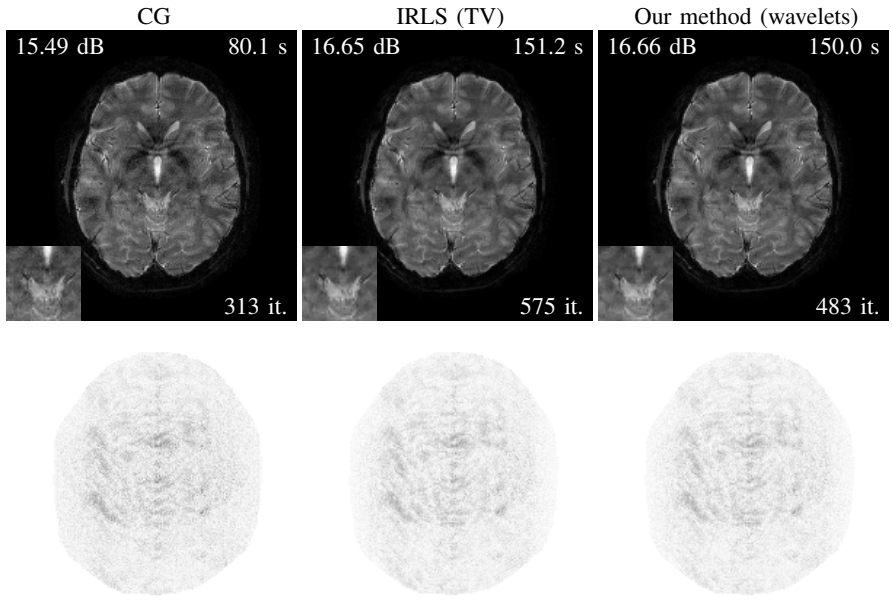

Figure 5.11: Reconstructions (top row) and error maps (bottom row) for the SENSE EPI experiment using CG (first column), IRLS-TV (second column), and our method (third column). For each reconstruction, the performance in SER with respect to the reference (top-left corner), the reconstruction time (top-right corner), the number of iterations (bottom-right corner) and a magnification of the central part (bottom-left) are shown.

terms of reconstruction duration, our method proves to converge in a time that is of the same order of magnitude as CG (see Figure 5.12).

5.4 Summary

We proposed an accelerated algorithm for nonlinear wavelet-regularized reconstruction that is based on two complementary acceleration strategies: use of adaptive subband thresholds plus multistep update rule. We provided theoretical evidence that this algorithm leads to faster convergence than when using the accelerating techniques independently. In the context of MRI, the proposed strategy can accelerate the reference algo-

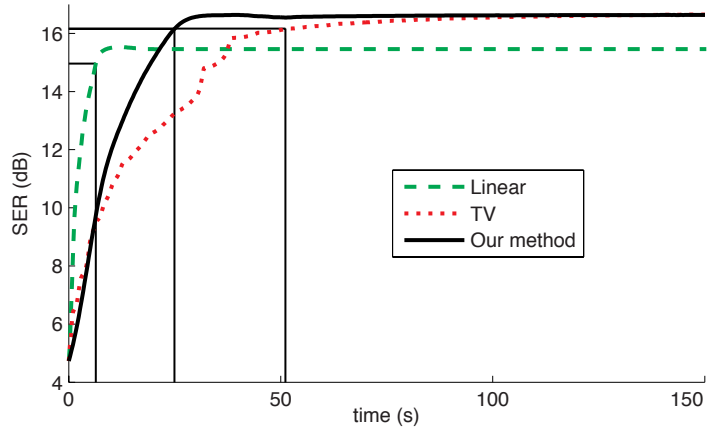

Figure 5.12: Evolution of the performance of the algorithms. Times required to reach -0.5 dB of the asymptotic value are indicated.

rithm up to two orders of magnitude. Moreover, we demonstrated that, by using the Haar wavelet transform with random shifting, we are able to boost the performance of wavelet methods to make them competitive with TV regularization. Using different simulations and *in vivo* data, we compared the practical performance of our reconstruction method with other linear and nonlinear ones.

The proposed method is proved to be competitive with TV regularization in terms of image quality. It typically converges within five seconds for the single channel problems considered. This brings nonlinear reconstruction forward to an order of magnitude of the time required by the state-of-the-art linear reconstructions, while providing much better quality.

Chapter 6

Conclusion

In this thesis, we presented a collection of new and competitive solutions for the reconstruction of magnetic-resonance images. We now summarize the main research directions and results in the first section of this chapter. The potential areas of interest for future research related to our work are listed in the second section.

6.1 Summary of Results

Modeling the MRI acquisition setup We described the principles of magnetic resonance imaging (MRI) from a signal-processing perspective. This allowed us to derive a clean linear discrete model which is consistent with the equations that govern the continuous physical world. For a given MRI setup, this numerical model relates the parameters that characterize the object under investigation to the corresponding scanner data. We detailed careful and efficient implementations of this model which is the cornerstone of any reconstruction method.

Realistic simulations for validation The use of simulations of the MRI data-acquisition process is very convenient for assessing the validity of reconstruction methods. To that end, we introduced a new theoretical framework. Among its important novel aspects are the facts that *(i)* it fully accounts for the continuous nature of the equations that govern the physics and that *(ii)* the parameterization is flexible enough to allow for the analytical description of realistic MRI setups and phantoms. We succeeded in designing such a realistic setup and did conduct validation

experiments with the help of this new simulation tool. In particular, we measured the image quality obtained from state-of-the-art reconstruction methods that were applied on data synthesized with both the conventional simulation approach—which is based on a discrete model—and our analytical method. The results differ significantly: the reconstruction performance is systematically overestimated in conventional simulations. This tendency is particularly pronounced with the nonlinear reconstruction schemes that are increasingly popular in MRI research. Our conclusion is that MRI reconstruction algorithms should not be evaluated using conventional simulations only. The novel simulation framework that we propose is a reliable alternative.

Competitive reconstruction of MRI images We presented MRI reconstruction as a general inverse problem which, in turn, is reformulated as a minimization problem. We detailed the known approaches leading to linear reconstruction and demonstrated their strong connections. Applied on challenging reconstruction tasks such as imaging when k-space samples are missing, these methods can be largely outperformed by some promising nonlinear approaches. Among those, we particularly focused on wavelet regularization. From our investigations about the influence of the choice of the wavelet transform on image quality, it appears that the wavelet basis has a limited impact and that it is vain to decompose the data beyond three levels. We provided a new variational interpretation that motivates the use of random shifting and deepens the understanding of its benefits on computational complexity and reconstruction quality. In our MRI experiments, random shifting led to a substantial gain in reconstruction quality, particularly when used together with the Haar wavelet transform. We also investigated acceleration strategies for wavelet-based iterative reconstruction. Based on theoretical grounds, we demonstrated that two recently proposed strategies can be combined synergistically. In practice, one can tailor the reconstruction scheme to the MRI setting—k-space trajectory and sensitivity of the receiving coil—to improve reconstruction speed. We ended up with a practical algorithm that is optimized for performance in terms of reconstruction quality and time. We conducted experiments to validate the method with challenging reconstruction tasks involving single and multiple channel, simulated, and real-world data. Linear reconstructions were observed to fail in providing acceptable image quality, contrarily to our reconstruction method

which is on a par with total-variation regularization for the quality of the reconstructed images. Typically, our nonlinear method requires few seconds on a personal computer to perform a reconstruction, which is comparable to the time required by conventional linear reconstruction methods but brings a substantial increase in quality.

6.2 Outlook

Extending the analytical phantom The analytical simulation framework we described allows the integration of temporal aspects, for instance moving phantoms (of interest for cardiac MRI) or region-dependent T_1 and T_2 parameters.

The phantom we proposed still lacks some texture to look fully realistic. Good candidates for the description of image texture would arise from the simulation of mesh-based structures, an aspect of our framework that we left without investigation.

It is not clear yet how to extend the Bézier-based contour parameterization to a third dimension while keeping the benefits of a closed-form description of the simulated data. Since very popular 3-D phantoms are defined using nonuniform rational B-splines (NURBS) [65], it would be a great achievement to work out the corresponding analytical solution.

Cross-fertilization The ideas behind our analytical phantom can potentially have an impact in other domains such as spectroscopy.

Adjustment of the regularization parameter The problem of setting a proper regularization parameter, which is nontrivial for nonlinear reconstruction in the context of ill-posed problems, makes the reconstruction time a crucial point for methods involving regularization. Current approaches to tune this parameter include generalized cross validation [88], the L-curve method [89], and Monte-Carlo SURE [90]. A recent work focused on MRI and Gaussian noise seems promising for non-linear reconstruction in MRI [91].

Further speedup in reconstruction time Substantial speedups in the reconstruction process have been recently reported using parallelization on dedicated GPU units [92, 93]. There is no apparent obstacle that

would prevent to employ the same parallelization strategy with our method to further speedup the reconstruction.

Another promising direction is the recently proposed i-LET approach [94] that performs well on deconvolution problems. This method provides a general framework that includes the different multi-step first-order approaches proposed in recent years to solve large-scale ℓ_1-regularized minimization problems.

Design of the MRI acquisition setup In this work, we focused on improving the quality of MRI images reconstructed out of data obtained through conventional acquisition setups. However, the recent trend of *Compressed sensing* or *Compressive sampling* in MRI has brought to light the potential benefits in revisiting the design of the acquisition setup. The first observation is that the reconstruction artifacts are less striking when the undersampling trajectory is randomized. Other designs of the k-space trajectory are currently under investigation [95]. It is likely that future research will increasingly involve signal processing in the design of the MRI setup.

Advanced MRI problems We considered in this thesis the estimation of the image out of MRI data. This is a linear problem. However, several challenging MRI problem settings do not fall in this category. This is the case of the estimation of the receiving coil sensitivity (possibly estimated jointly with the image), B0 estimation and correction [2], motion correction [4, 5], or higher-order field imaging [7]. Efficient solutions to these problems might be found in the future, based on the results obtained with linear problems.

Appendix A

Simulation

A.1 Proof of Theorem 4.3.2

First, we consider the case $\boldsymbol{\omega} = \mathbf{0}$.

Proof. Take $\psi(\mathbf{r}) = \mathbf{r}^{\boldsymbol{\alpha}}$ and $\varphi(\mathbf{r}) = (\mathbf{e}_k \cdot \mathbf{r})^2/2 = r_k^2/2$. Then, $\boldsymbol{\nabla}\psi(\mathbf{r}) \cdot \mathbf{e}_k = \alpha_k \mathbf{r}^{\boldsymbol{\alpha}-\mathbf{e}_k}$, $\boldsymbol{\nabla}\varphi(\mathbf{r}) = r_k \mathbf{e}_k$, and $\Delta\varphi(\mathbf{r}) = 1$. Using Green's first identity yields

$$(1+\alpha_k) f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}) = \int_{\partial\mathscr{B}} \mathbf{r}^{\boldsymbol{\alpha}+\mathbf{e}_k} \mathbf{e}_k \cdot \mathbf{n} \mathrm{d}\sigma = (1+\alpha_k) g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\mathbf{0}). \qquad \square$$

For the case $\boldsymbol{\omega} \neq \mathbf{0}$, we need an intermediate result.

Lemma A.1.1. *For $\boldsymbol{\omega} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\boldsymbol{\alpha} \in \mathbb{N}^d$,*

$$f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \mathrm{j} g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) + \sum_i \left(\frac{-\mathrm{j}\omega_i}{\|\boldsymbol{\omega}\|^2} \right) \alpha_i f_{\mathscr{B}}^{\boldsymbol{\alpha}-\mathbf{e}_i}(\boldsymbol{\omega}). \tag{A.1}$$

Proof. Use Green's first identity with $\psi(\mathbf{r}) = \mathbf{r}^{\boldsymbol{\alpha}}$, and $\varphi(\mathbf{r}) = -\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}}$. Then, $\mathbf{x}\cdot\boldsymbol{\nabla}\psi(\mathbf{r}) = \sum_i x_i \alpha_i \mathbf{r}^{\boldsymbol{\alpha}-\mathbf{e}_i}$, $\boldsymbol{\nabla}\varphi(\mathbf{r}) = \mathrm{j}\mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}}\boldsymbol{\omega}$, and $\Delta\varphi(\mathbf{r}) = \|\boldsymbol{\omega}\|^2 \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}}$. $\square$

Let us continue the proof of Theorem 4.3.2 by induction on $n = |\boldsymbol{\alpha}|$. For $n = 0$, $\boldsymbol{\alpha} = \mathbf{0}$ and the result holds true according to Lemma A.1.1. When considering $n = 1$, $\boldsymbol{\alpha} = \mathbf{e}_i$, and Lemma A.1.1, we obtain $f_{\mathscr{B}}^{\mathbf{e}_i}(\boldsymbol{\omega}) = \mathrm{j} g_{\mathscr{B}}^{\mathbf{e}_i}(\boldsymbol{\omega}) + \frac{\omega_i}{\|\boldsymbol{\omega}\|^2} g_{\mathscr{B}}^{\mathbf{0}}(\boldsymbol{\omega})$. This is true for all i, hence the result holds true for

$n = 1$. Now, we suppose the result holds true at order n and we consider $\boldsymbol{\alpha}$ such that $|\boldsymbol{\alpha}| = n+1$. From Lemma A.1.1, we have that $f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \mathrm{j} g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) + \sum_i \left(\frac{-\mathrm{j}\omega_i}{\|\boldsymbol{\omega}\|^2}\right) \alpha_i f_{\mathscr{B}}^{\boldsymbol{\alpha}-\mathbf{e}_i}(\boldsymbol{\omega})$. Since $|\boldsymbol{\alpha}-\mathbf{e}_i| = n$, we substitute $f_{\mathscr{B}}^{\boldsymbol{\alpha}-\mathbf{e}_i}$ using the induction hypothesis and, after simplifications, we obtain

$$\begin{aligned} f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) &= \mathrm{j} g_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) + \mathrm{j} \sum_i \sum_{\mathbf{m}=\mathbf{0}}^{\boldsymbol{\alpha}} \left(\frac{-\mathrm{j}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|^2}\right)^{\boldsymbol{\alpha}-\mathbf{m}} \\ &\times |\boldsymbol{\alpha}-\mathbf{m}|! d_i(\boldsymbol{\alpha}-\mathbf{m}) \binom{\boldsymbol{\alpha}}{\mathbf{m}} g_{\mathscr{B}}^{\mathbf{m}}(\boldsymbol{\omega}) \end{aligned}$$

with $d_i(\mathbf{x}) = x_i/|\mathbf{x}|$ for $\mathbf{x} \neq \mathbf{0}$ and $d_i(\mathbf{0}) = 0$. By permutation of the sums and noting that $\sum_i d_i(\mathbf{x}) = 1$ for $\mathbf{x} \neq \mathbf{0}$ and $\sum_i d_i(\mathbf{0}) = 0$, we get

$$f_{\mathscr{B}}^{\boldsymbol{\alpha}}(\boldsymbol{\omega}) = \mathrm{j} \sum_{\mathbf{m}=\mathbf{0}}^{\boldsymbol{\alpha}} \left(\frac{-\mathrm{j}\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|^2}\right)^{\boldsymbol{\alpha}-\mathbf{m}} |\boldsymbol{\alpha}-\mathbf{m}|! \binom{\boldsymbol{\alpha}}{\mathbf{m}} g_{\mathscr{B}}^{\mathbf{m}}(\boldsymbol{\omega}).$$

This is valid for all $\boldsymbol{\alpha}$ such that $|\boldsymbol{\alpha}| = n+1$. Hence, we just proved the result for $n+1$ assuming it holds true for n.

A.2 Characterization and Computations of a Family of 1-D Integrals

Proposition A.2.1. *For $m \in \mathbb{N}$, $h^{(m)}$ follows the recursion rule*

$$2\mathrm{j}b h^{(m+1)}(a,b) + \mathrm{j}a h^{(m)}(a,b) - m h^{(m-1)}(a,b) + \mathrm{e}^{-\mathrm{j}(a+b)} - \delta_m = 0. \quad \text{(A.2)}$$

Proof. Integrate $\int_0^1 -\mathrm{j}(a+2b\lambda)\lambda^m \mathrm{e}^{-\mathrm{j}\lambda(a+\lambda b)} \mathrm{d}\lambda$ by parts and identify $h^{(m+1)}$, $h^{(m)}$ and $h^{(m-1)}$ if $m > 0$. □

Corollary A.2.2. *For small values of a and b, one can rely on the backward iteration starting from a higher order $M > m$ to get accurate results*

- $\tilde{h}^{(M+1)}(a,b) = \tilde{h}^{(M)}(a,b) = 0$
- $\tilde{h}^{(m)}(a,b) = \left(2\mathrm{j}b\tilde{h}^{(m+2)}(a,b) + \mathrm{j}a\tilde{h}^{(m+1)}(a,b) + \mathrm{e}^{-\mathrm{j}(a+b)}\right)/(m+1)$.

Proposition A.2.3. *For b nonzero and $m \geq 1$, the forward iteration is used*

- $h^{(0)}(a,b) = \frac{\sqrt{\pi}\mathrm{e}^{\frac{\mathrm{j}a^2}{4b}}}{2\sqrt{\mathrm{j}b}} \left[\mathrm{erf}\left(\frac{(a+2b)\sqrt{\mathrm{j}}}{2\sqrt{b}}\right) - \mathrm{erf}\left(\frac{a\sqrt{\mathrm{j}}}{2\sqrt{b}}\right)\right]$

– $h^{(m+1)}(a,b) = \left(mh^{(m-1)}(a,b) - \mathrm{j}ah^{(m)}(a,b) + \mathrm{e}^{-\mathrm{j}(a+b)}\right)/(2\mathrm{j}b)$

with $\mathrm{erf}(z) = \frac{2z}{\sqrt{\pi}} \int_0^1 \mathrm{e}^{-z^2t^2}\mathrm{d}t$.

Proof. From Proposition A.2.1 with $m = 0$ and $b = 0$, we get $h^{(0)}(a,0) = \mathrm{e}^{-\mathrm{j}a/2}\mathrm{sinc}(a/(2\pi))$. In the case $b \neq 0$, we define $t = \lambda + \frac{a}{2b}$ such that $\lambda(a + b\lambda) = \frac{a^2}{4b} - bt^2$. By Definition (4.22), we get

$$\mathrm{e}^{-\mathrm{j}\frac{a^2}{4b}} \int_{\frac{a}{2b}}^{\frac{a}{2b}+1} \mathrm{e}^{-\mathrm{j}bt^2}\mathrm{d}t,$$

after the change of variable. Splitting this integral, we get

$$\mathrm{e}^{\mathrm{j}\frac{a^2}{4b}} \left(\int_0^{\frac{a}{2b}+1} \mathrm{e}^{-\mathrm{j}bt^2}\mathrm{d}t - \int_0^{\frac{a}{2b}} \mathrm{e}^{-\mathrm{j}bt^2}\mathrm{d}t \right).$$

The result follows from normalizing the integration intervals. □

Proposition A.2.4. *For b small, the truncated Taylor series in $b = 0$ provides accurate results*

$$h^{(m)}(a,b) = \sum_{n=0}^{\infty} \frac{(-\mathrm{j}b)^n \gamma(m+2n+1, \mathrm{j}a)}{n!(\mathrm{j}a)^{m+2n+1}}, \tag{A.3}$$

where the lower incomplete gamma function is defined as

$$\gamma(s,z) = z^s \int_0^1 \lambda^{s-1}\mathrm{e}^{-\lambda z}\mathrm{d}\lambda.$$

Proof. Note that $\mathrm{e}^{-\mathrm{j}\lambda(a+\lambda b)} = \mathrm{e}^{-\mathrm{j}\lambda a} \sum_{n=0}^{\infty} (-\mathrm{j}\lambda^2 b)^n/n!$. By virtue of Fubini's theorem, we get

$$h^{(m)}(a,b) = \sum_{n=0}^{\infty} (-\mathrm{j}b)^n h^{(m+2n)}(a,0)/n!.$$

Identify $h^{(m+2n)}(a,0)$ to $\gamma(m+2n+1, \mathrm{j}a)/(\mathrm{j}a)^{m+2n+1}$. □

A.3 Proof of Proposition 4.3.4

Proof. We rewrite $g_{\mathscr{B}}^{\boldsymbol{\alpha}}$ using Equation (4.21) with $F(\mathbf{r}) = \mathbf{r}^{\boldsymbol{\alpha}} \mathrm{e}^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{r}} \boldsymbol{\omega} / \|\boldsymbol{\omega}\|^2$ for $\boldsymbol{\omega} \neq \mathbf{0}$ and $F(\mathbf{r}) = \mathbf{r}^{\boldsymbol{\alpha}+\mathbf{e}_k} \mathbf{e}_k / (1+\alpha_k)$ for $\boldsymbol{\omega} = \mathbf{0}$. The piecewise parameterization of the contour (Table 4.2) is then used, and by virtue of the multinomial theorem, we expand the terms $\mathbf{r}^{\boldsymbol{\alpha}}$ and $\mathbf{r}^{\boldsymbol{\alpha}+\mathbf{e}_k}$. □

Appendix B

Efficient Wavelet-Based Reconstruction

B.1 Proof of Proposition 5.2.1

Proof. We rewrite the cost function (5.2) with the change of variable $\mathbf{w}' = \mathbf{\Lambda}^{1/2}\mathbf{w}$. We then apply ISTA to solve the problem in terms of $\mathbf{w}'$. The new parameters are $\mathbf{a}' = \mathbf{\Lambda}^{-1/2}\mathbf{a}$, $\mathbf{A}' = \mathbf{\Lambda}^{-1/2}\mathbf{A}\mathbf{\Lambda}^{-1/2}$, and thresholds $\lambda\sqrt{\tau_k}$ that are specific to each coefficient. Noting that $\mathbf{\Lambda} - \mathbf{A}$ is positive-definite if and only if $\mathbf{I} - \mathbf{A}'$ is positive-definite leads us to $L = 2$. The iteration $\mathbf{w}'_{i+1} = \mathrm{T}_{\lambda\sqrt{\tau}}\left(\mathbf{w}'_i + (\mathbf{a}' - \mathbf{A}'\mathbf{w}'_i)\right)$ can be rewritten, in terms of the original variable, as $\mathbf{w}_{i+1} = \mathrm{T}_{\lambda\tau}\left(\mathbf{w}_i + \mathbf{\Lambda}^{-1}(\mathbf{a} - \mathbf{A}\mathbf{w}_i)\right)$. The latter is nothing but an iteration of SISTA (see Algorithm 11). According to Proposition 3.4.1, we have $\mathscr{C}(\mathbf{\Lambda}^{-1/2}\mathbf{w}'_i) - \mathscr{C}(\mathbf{w}^\star) \leq \left\|\mathbf{w}'_{i_0} - \mathbf{\Lambda}^{1/2}\mathbf{w}^\star\right\|^2 / (i - i_0)$, which translates directly into the proposed result. □

B.2 Proof of Proposition 5.2.2

Proof. In the spirit of the proof of Proposition 5.2.1, we consider the change of variable $\mathbf{w}' = \mathbf{\Lambda}^{1/2}\mathbf{w}$ and apply FISTA to solve the new reconstruction problem. The ISTA step $\mathbf{w}'_{i+1} = \mathrm{T}_{\lambda\sqrt{\tau}}\left(\mathbf{v}'_i + (\mathbf{a}' - \mathbf{A}'\mathbf{v}'_i)\right)$ is equivalent to a SISTA step in terms of the original variable $\mathbf{w}_{i+1} = \mathrm{T}_{\lambda\tau}\left(\mathbf{v}_i + \mathbf{\Lambda}^{-1}(\mathbf{a} - \mathbf{A}\mathbf{v}_i)\right)$. The convergence results of FISTA [20, Thm. 4.4] applies on the sequence $\left\{\mathbf{w}'_i\right\}$, which leads to $\mathscr{C}(\mathbf{\Lambda}^{-1/2}\mathbf{w}'_i) - \mathscr{C}(\mathbf{w}^\star) \leq$

$\left(\frac{2}{i+1}\right)^2 \left\|\mathbf{w}_0' - \mathbf{\Lambda}^{1/2}\mathbf{w}^\star\right\|^2$. In the strongly convex case, we have

$$\frac{\epsilon}{2}\left\|\mathbf{w}_i - \mathbf{w}^\star\right\|_2^2 \leq \mathscr{C}(\mathbf{w}_i) - \mathscr{C}(\mathbf{w}^\star).$$

□

Bibliography

[1] P. Lauterbur, "Image formation by induced local interactions: Examples employing nuclear magnetic resonance," *Nature*, vol. 242, pp. 190–191, 1973.

[2] L. Man, J. Pauly, and M. A., "Multifrequency interpolation for fast off-resonance correction," *Magnetic Resonance in Medicine*, vol. 37, no. 5, pp. 785–792, April 1997.

[3] K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, October 1999.

[4] M. Zaitsev, C. Dold, G. Sakas, J. Hennig, and O. Speck, "Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system," *NeuroImage*, vol. 31, no. 3, pp. 1038–1050, 2006.

[5] M. B. Ooi, S. Krueger, W. J. Thomas, S. V. Swaminathan, and T. R. Brown, "Prospective real-time correction for arbitrary head motion using active markers," *Magnetic Resonance in Medicine*, vol. 62, no. 4, pp. 943–954, 2009.

[6] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, pp. 1182–1195, 2007.

[7] B. J. Wilm, C. Barmet, M. Pavan, and K. P. Pruessmann, "Higher order reconstruction for MRI in the presence of spatiotemporal field perturbations," *Magnetic Resonance in Medicine*, vol. 65, no. 6, pp. 1690–1701, 2011.

[8] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint," *Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1086–1098, 2007.

[9] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic MRI," *Magnetic Resonance in Medicine*, vol. 59, no. 2, pp. 365–373, 2008.

[10] L. Ying, L. Bo, M. C. Steckner, W. Gaohong, W. Min, and L. Shi-Jiang, "A statistical approach to SENSE regularization with arbitrary k-space trajectories," *Magnetic Resonance in Medicine*, vol. 60, pp. 414–421, 2008.

[11] B. Liu, E. Abdelsalam, J. Sheng, and L. Ying, "Improved spiral sense reconstruction using a multiscale wavelet model," in *Proceedings of ISBI*, 2008, pp. 1505–1508.

[12] B. Liu, K. King, M. Steckner, J. Xie, J. Sheng, and L. Ying, "Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations," *Magnetic Resonance in Medicine*, vol. 61, pp. 145–152, January 2009.

[13] L. Chaâri, J.-C. Pesquet, A. Bebazza-Benyahia, and P. Ciuciu, "Autocalibrated regularized parallel MRI reconstruction in the wavelet domain," in *Proceedings of ISBI*, Paris, France, May, 14-17 2008, pp. 756–759.

[14] L. Chaâri, J.-C. Pesquet, A. Benazza-Benyahia, and P. Ciuciu, "A wavelet-based regularized reconstruction algorithm for SENSE parallel MRI with applications to neuroimaging," *Medical Image Analysis Journal*, vol. 15, no. 2, pp. 185–201, April 2011.

[15] F. Knoll, K. Bredies, T. Pock, and S. R., "Second order total generalized variation (TGV) for MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 480–491, Feb 2011.

[16] G. Puy, J. Marques, R. Gruetter, J. Thiran, D. Van De Ville, P. Vandergheynst, and Y. Wiaux, "Spread spectrum magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 586–598, March 2012.

[17] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Transactions on Signal Processing*, vol. 12, no. 8, pp. 906–916, 2003.

[18] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle, "A ℓ^1-unified variational framework for image restoration," *Lecture Notes in Computer Science*, vol. 3024, pp. 1–13, 2004.

[19] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[20] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[21] İ. Bayram and I. W. Selesnick, "A subband adaptive iterative shrinkage/thresholding algorithm," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1131–1143, 2010.

[22] M. E. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, Wiley-Liss, June 1999.

[23] Z.-P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*, Wiley-IEEE Press, October 1999.

[24] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique*, vol. I and II, Hermann, 1973.

[25] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch, "A new method of measuring nuclear magnetic moment," *Physical Review*, vol. 53, no. 4, pp. 318, Feb 1938.

[26] E. M. Purcell, H. C. Torrey, and R. V. Pound, "Resonance absorption by nuclear magnetic moments in a solid," *Physical Review*, vol. 69, no. 1–2, pp. 37–38, Jan 1946.

[27] F. Bloch, "Nuclear induction," *Physical Review*, vol. 70, pp. 460–473, 1946.

[28] R. V. Damadian, "Tumor detection by nuclear magnetic resonance," *Science*, vol. 171, pp. 1151–1153, March 1971.

[29] H. Gach, C. Tanase, and F. Boada, "2D & 3D Shepp-Logan phantom standards for MRI," in *19th International Conference on Systems Engineering, Las Vegas, NV, USA*, Los Alamitos, CA, USA, August 2008, pp. 521–526, IEEE Computer Society.

[30] M. Blaimer, F. Breuer, M. Mueller, R. Heidemann, M. Griswold, and P. Jakob, "SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method.," *Top. Magn. Reson. Imaging*, vol. 15, no. 4, pp. 223–236, 2004.

[31] D. J. Larkman and R. G. Nunes, "Parallel magnetic resonance imaging," *Physics in Medicine and Biology*, vol. 52, no. 7, pp. R15–R55, Apr. 2007.

[32] D. Sodickson and W. Manning, "Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays.," *Magnetic Resonance in Medicine*, vol. 38, no. 4, pp. 591–603, 1997.

[33] P. Jakob, M. Griswold, R. Edelman, and D. Sodickson, "AUTO-SMASH: a self-calibrating technique for SMASH imaging. SiMultaneous Acquisition of Spatial Harmonics.," *MAGMA*, vol. 7, no. 1, pp. 42–54, 1998.

[34] R. Heidemann, M. Griswold, A. Haase, and P. Jakob, "VD-AUTO-SMASH imaging.," *Magnetic Resonance in Medicine*, vol. 45, no. 6, pp. 1066–1074, 2001.

[35] M. A. Griswold, P. M. Jackob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisition (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1002–1010, 2002.

[36] M. Griswold, P. Jakob, M. Nittka, J. Goldfarb, and A. Haase, "Partially parallel imaging with localized sensitivities (PILS).," *Magnetic Resonance in Medicine*, vol. 44, no. 4, pp. 602–609, 2000.

[37] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magnetic Resonance in Medicine*, vol. 46, no. 4, pp. 638–651, 2001.

[38] B. P. Sutton, D. C. Noll, and J. A. Fessler, "Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 178–188, 2003.

[39] B. Delattre, J.-N. Hyacinthe, J.-P. Vallée, and D. Van De Ville, "Spline-based variational reconstruction of variable density spiral k-space data with automatic parameter adjustment," in *Proceedings of ISMRM 17th Annual Meeting*, Hawai'i, USA, 2009, p. 2066.

[40] M. Guerquin-Kern, M. Häberlin, K. P. Pruessmann, and M. Unser, "A fast wavelet-based reconstruction method for magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 30, no. 9, pp. 1649–1660, September 2011.

[41] M. Unser and T. Blu, "Wavelet theory demystified," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 470–483, 2003.

[42] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1999.

[43] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *J. ACM*, vol. 9, no. 1, pp. 84–97, Jan. 1962.

[44] S. Twomey, "On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," *J. ACM*, vol. 10, no. 1, pp. 97–101, Jan. 1963.

[45] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, Winston, Washington, D.C., 1977.

[46] C. Barmet, M. Haeberlin, and K. P. Pruessmann, "A Robust Alternative to Regularization in Parallel Imaging Reconstruction," in *Proceedings of ISMRM*, 2007.

[47] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[48] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Transactions on Signal Processing*, vol. 51, pp. 560–574, 2003.

[49] F. T. W. A. Wajer and K. P. Pruessmann, "Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories," in *Proceedings of ISMRM 9th Annual Meeting*, Glasgow, United Kingdom, 2001, p. 625.

[50] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 409–436, Dec. 1952.

[51] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, July 1995.

[52] J.-C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1143–1153, May 2011.

[53] B. Wohlberg and P. Rodríguez, "An iteratively reweighted norm algorithm for minimization of total variation functionals," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 948–951, Dec. 2007.

[54] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120–145, May 2011.

[55] M. J. Fadili and G. Peyré, "Total variation projection with first order schemes," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 657–669, March 2011.

[56] A. Chambolle, R. A. DeVore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Transactions on Signal Processing*, vol. 7, no. 3, pp. 319–335, 1998.

[57] M. Guerquin-Kern, L. Lejeune, K. P. Pruessmann, and M. Unser, "Realistic analytical phantoms for parallel magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 626–636, March 2012.

[58] A. Ribés and F. Schmitt, "Linear inverse problems in imaging," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 84–99, July 2008.

[59] L. Shepp and B. Logan, "The Fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, pp. 21–43, June 1974.

[60] M. Smith, L. Chen, Y. Hui, T. Mathews, J. Yang, and X. Zeng, "Alternatives to the use of the DFT in MRI and spectroscopic reconstructions," *International Journal of Imaging Systems and Technology*, vol. 8, no. 6, pp. 558–564, December 1997.

[61] R. Van de Walle, H. Barrett, K. Myers, M. Aitbach, B. Desplanques, A. Gmitro, J. Cornelis, and I. Lemahieu, "Reconstruction of MR images from data acquired on a general nonregular grid by pseudoinverse calculation," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1160–1167, December 2000.

[62] C. Koay, J. Sarlls, and E. Ãzarslan, "Three-dimensional analytical magnetic resonance imaging phantom in the Fourier domain," *Magnetic Resonance in Medicine*, vol. 58, no. 2, pp. 430–436, August 2007.

[63] L. Greengard and C. Stucchio, "Spectral edge detection in two dimensions using wavefronts," *Applied and Computational Harmonic Analysis*, vol. 30, no. 1, pp. 69–95, 2011.

[64] T. M. Ngo, G. S. Fung, B. M. Tsui, E. McVeigh, and H. D. A., "Three dimensional digital polyhedral phantom framework with analytical fourier transform and application in cardiac imaging," in *Proceedings of ISMRM*, 2011, p. 1310.

[65] W. Segars, D. Lalush, and B. Tsui, "A realistic spline-based dynamic heart phantom," *IEEE Transactions on Nuclear Science*, vol. 46, no. 3, pp. 503–506, June 1999.

[66] S. Vembu, "Fourier transformation of the n-dimensional radial delta function," *The Quarterly Journal of Mathematics*, vol. 12, no. 1, pp. 165–168, 1961.

[67] E. Sorets, "Fast Fourier transforms of piecewise constant functions," *Journal of Computational Physics*, vol. 116, no. 2, pp. 369–379, 1995.

[68] M. Jacob, T. Blu, and M. Unser, "An exact method for computing the area moments of wavelet and spline curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 633–642, June 2001.

[69] W. Gautschi, "Computational aspects of three-term recurrence relations," *SIAM review*, vol. 9, pp. 24–82, January 1967.

[70] F. Champagnat and J. Idier, "A connection between half-quadratic criteria and EM algorithms," *IEEE Signal Processing Letters*, vol. 11, no. 9, pp. 709–712, September 2004.

[71] R. Archibald and A. Gelb, "A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 305–319, April 2002.

[72] M. Guerquin-Kern, J.-C. Baritaux, and M. Unser, "Efficient image reconstruction under sparsity constraints with application to MRI and bioluminescence tomography," in *Proceedings of the Thirty-Sixth IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'11)*, Prague, Czech Republic, May 22-27 2011, pp. 5760–5763.

[73] M. Lustig, J. H. Lee, D. L. Donoho, and J. M. Pauly, "Faster imaging with randomly perturbed, undersampled spirals and $|L|_1$ reconstruction," in *Proceedings of ISMRM*, 2005.

[74] M. Guerquin-Kern, D. Van De Ville, C. Vonesch, J.-C. Baritaux, K. P. Pruessmann, and M. Unser, "Wavelet-regularized reconstruction for rapid MRI," in *Proceedings of ISBI*, 2009, pp. 193–196.

[75] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.

[76] Y. E. Nestorov, "Gradient methods for minimizing composite objective function," Tech. Rep., CORE report, 2007.

[77] P. Weiss, *Algorithmes rapides d'optimisation convexe. Application à la restoration d'images et à la détection de changements*, Ph.D. thesis, Université de Nice, Dec. 2008.

[78] S. Becker, J. Bobin, and E. J. Candès, "NESTA: A fast and accurate first-order method for sparse recovery," Tech. Rep., California Institute of Technology, 2009.

[79] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, pp. 2419–2434, 2009.

[80] C. Vonesch and M. Unser, "A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution," *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 539–549, 2008.

[81] C. Vonesch and M. Unser, "A fast multilevel algorithm for wavelet-regularized image restoration," *IEEE Transactions on Image Processing*, vol. 18, no. 3, pp. 509–523, 2009.

[82] J. A. Fessler, S. Lee, V. T. Olafsson, H. R. Shi, and D. C. Noll, "Toeplitz-based iterative image reconstruction for MRI with correction for magnetic field inhomogeneity," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3393–3402, 2005.

[83] W. C. Karl, "Regularization in image restoration and reconstruction," in *Handbook of image and video processing*, A. C. Bovik, Ed., pp. 183–202. Elsevier Science & Technology, 2005.

[84] G. H. Glover, "Simple analytic spiral k-space algorithm," *Magnetic Resonance in Medicine*, vol. 42, no. 2, pp. 412–415, 1999.

[85] D.-H. Kim, E. Adalsteinsson, and D. M. Spielman, "Simple analytic variable density spiral design," *Magnetic Resonance in Medicine*, vol. 50, no. 1, pp. 214–219, 2003.

[86] C. Barmet, N. De Zanche, B. J. Wilm, and K. P. Pruessmann, "A transmit/receive system for magnetic field monitoring of *in vivo* MRI," *Magnetic Resonance in Medicine*, vol. 62, no. 1, pp. 269–276, July 2009.

[87] M. Guerquin-Kern, F. I. Karahanoğlu, D. Van De Ville, K. P. Pruessmann, and M. Unser, "Analytical form of Shepp-Logan phantom for parallel MRI," in *Proceedings of ISBI, Rotterdam, The Netherlands*, April 14-17 2010, pp. 261–264.

[88] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.

[89] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Review*, vol. 34, pp. 561–580, December 1992.

[90] 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*, vol. 17, no. 9, pp. 1540–1554, 2008.

[91] S. Ramani, Z. Liu, 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," *to appear in IEEE Transactions on Image Processing*, 2012.

[92] F. Knoll, M. Unger, C. Diwoky, C. Clason, T. Pock, and R. Stollberger, "Fast reduction of undersampling artifacts in radial MR angiography with 3D total variation on graphics hardware," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, pp. 103–114, 2010.

[93] S. S. Vasanawala, M. J. Murphy, M. T. Alley, P. Lai, K. Keutzer, J. M. Pauly, and M. Lustig, "Practical parallel imaging compressed sensing MRI: Summary of two years of experience in accelerating body MRI of pediatric patients," in *Proceedings of ISBI*, 2011.

[94] H. Pan and T. Blu, "Sparse image restoration using iterated linear expansion of thresholds," in *Proceedings of the 2011 IEEE International Conference on Image Processing (ICIP'11)*, 11–14 September 2011, pp. 1905–1908.

[95] M. Haeberlin, B. Wilm, C. Barmet, S. Kozerke, G. Katsikatsos, and K. P. Pruessmann, "Sinusoidal perturbations improve the noise behavior of parallel EPI," in *Proceedings of ISMRM*, 2010.

Biography

Matthieu Guerquin-Kern was born in France in 1983. He studied electrical engineering at University of Paris-Sud 11 and ÉNS Cachan. In 2007, he obtained a M.Sc. in applied mathematics from ÉNS Cachan. Meanwhile, he became laureate from the competitive examination of the French public education system, know as “agrégation”, in applied physics and electrical engineering. He carried out his Ph.D. at the Biomedical Imaging Group at EPFL (Switzerland) under the supervision of Prof. M. Unser. Since September 2011, he is lecturer at ÉNSEA, a French graduate school in electrical engineering and computer science. His teaching interests include electrical engineering, mathematics and physics. His research interests include inverse problems, large-scale optimization, nonlinear reconstruction, and applications to biomedical imaging.