



**Louvain School of Engineering**

Institute of Information and Communication  
Technologies, Electronics and Applied Mathematics  
B - 1348 Louvain-la-Neuve, Belgium.

# **Solving Inverse Problems in Imaging using Robust and Regularized Optimization**

**Adriana Cristina González González**

Thesis presented in partial fulfillment of the requirements  
for the PhD degree in Engineering Sciences

## PhD Committee

Laurent Jacques, <i>Supervisor</i>	Université catholique de Louvain, Belgium
David Bol, <i>President</i>	Université catholique de Louvain, Belgium
Christophe De Vleeschouwer	Université catholique de Louvain, Belgium
Philippe Antoine	Lambda-X, Belgium
Christine De Mol	Université libre de Bruxelles, Belgium
Pierre Weiss	ITAV Toulouse, France

May 2016



# Abstract

Digital images play an important role in the human life since they allow observing, analyzing, studying and characterizing the world surrounding us. Their use is ubiquitous in many applications such as medicine, biology, astronomy, security, archeology, art conservation and industrial manufacture. Nonetheless, we usually do not have a direct access to the desired digital image, which needs to be recovered from corrupted, incomplete and/or indirect observations. The determination of the unknown image from the available observations is called an inverse problem. In this thesis, we study several inverse problems appearing in imaging applications and we investigate how to solve them using optimization techniques that are robust to the distortions in the acquisition and modeling processes.

In the first part of the thesis, we study an inverse problem that exists in a novel imaging modality called optical deflectometric tomography. Through indirect measurements, we can generate an image describing the spatial distribution of the refractive index of a transparent object, which allows the optical characterization of the object. For this imaging modality, we formulate a fast and accurate discrete forward model characterizing the physical acquisition process and we analyze the existing distortions. We propose a robust constrained optimization method based on the consistency of the model with the observations and on the physical properties of the refractive index map. The proposed approach is shown to outperform traditional tomographic reconstruction methods, providing a high quality image of the refractive index map from noisy and incomplete observations.

In the second part of the thesis, we analyze the 2-D phase unwrapping problem. This inverse problem appears in every optical application where the phase image is not directly observed but is extracted from the output of the

instrument and, hence, it is wrapped inside a certain interval (e.g.,  $[-\pi, \pi)$ ). By working with a forward model formulated in the phase derivative domain, the unwrapping inverse problem is relaxed and can be solved using convex optimization techniques. In contrast to existing unwrapping methods, we propose a robust approach that aims at simultaneously unwrap and denoise the phase image. The proposed approach is shown to compare favorably to those that perform the unwrapping and denoising in two separate steps.

In the last part of the thesis, we study the blind deconvolution problem. This inverse problem appears in every imaging application where the acquired image results from the convolution between the original image and the (unknown) response of the imaging device. Through blind deconvolution techniques, we aim at simultaneously retrieving the original enhanced image and the response of a solar telescope (also called point spread function or PSF) from corrupted solar observations. In contrast to most common approaches in the literature, which are based on a parametric model of the PSF, the proposed method relies on the information available on the image (sparsity and strong knowledge on some image pixels values and their exact location) while imposing soft constraints on the PSF. The blind deconvolution problem is solved through a proximal alternating algorithm that allows us to guarantee that a critical point of the problem is always reached. Through synthetic and experimental data, we show that the proposed non-parametric blind deconvolution method can estimate the core of the PSF with a similar quality to parametric methods proposed in the literature, and it presents the advantage of being general and applicable to any instrument.

# Acknowledgments

First of all, I would like to thank my supervisor, Prof. Laurent Jacques, for being my mentor since the beginning, for the insightful discussions, constructive criticism and encouragement. His guidance, his scientific rigor in both mathematical and physics fields and his enthusiasm have been crucial for the accomplishment of the thesis.

My initial supervisors, Dr. Philippe Antoine and Prof. Christophe De Vleeschouwer, have been of great influence too. I appreciate their help on my first project in optical deflectometric tomography, which led to the third chapter of the thesis. I would also like to thank them and the other members of the jury, Prof. Christine De Mol, Prof. Pierre Weiss and Prof. David Bol, for the time dedicated to review the thesis and for their valuable comments, which have greatly improved the final version of the manuscript. Thanks are extended to Prof. Raphaël Jungers for agreeing to be the president of the public defense of the thesis.

In addition, I am very grateful to Dr. Véronique Delouille for a great collaboration on the subject of blind deconvolution of astronomical images, for the expertise she provided in this field, for the fruitful discussions and for her valuable comments.

I would like to express my sincere thanks to Dr. Hong Jiang from Nokia Bell-Labs for giving me the opportunity to work in his team as a summer intern and also to all the team in Bell-Labs and my fellow interns. Although my work there has not been included in the thesis, I believe this summer internship was an important experience for my scientific training and for my personal and professional development.

I would also like to thank Université catholique de Louvain for the financial support and for giving me the opportunity of being a teaching assistant.

Thanks are extended to all the members of ICTEAM for all the nice activities and discussions and for the help whenever needed. I would like to specially mention my office colleagues for creating such a nice working environment, from the initial ones (Rémi, Farshad and Lyazid) until the most recent ones (Amir, Adrià and Augustin). Additionally, I would like to mention some of my colleagues from the corridor and thank them for the Bday cakes, the Friday's lunch and all the fun activities organized throughout these years: Amit, Arnab, Arnaud×2, Arsalan, Benoît, Cédric, François, Ivan, Jeevan, Kaori, Kévin, Mireia, Nafiseh, Pascaline, Pierre-Yves, Prasad, Sebastien, Shambhu, Simon, Stéphanie, Thomas, Máximo, Valerio. Finally, I would like to name some members of the technical staff that are always there to help us with any technical and administrative issue: Anthony, David, François, Isabelle, Jean, Nathalie and Patricia.

I am also very grateful to my friends that have been there through the years and with whom I have spent incredible moments and enjoyed unforgettable trips. Some of them since I was a kid (Anaka, Renata, Gaby, Marinell, Yeyi, Rosestela, Miren, Carmen, ...), some others since the university (Clau, Ele, Vane, Diego, Javi, ...), and some since my arrival in Belgium (Jaza, Hiam, Samer, Annabelle, Amin, Sandrine, Simon, Luce, Rony, ...).

I would like to dedicate special thanks to Mohieddine for his support in the good and the bad moments, for encouraging me at all times specially in those stressful moments that come when pursuing a PhD, for his help and valuable comments and finally for all the shared moments that have made these years so lovely and enjoyable.

Last but not least, I would like to thank all my family, specially my parents, my brother and my sister, because without their endless love, support and encouragement this journey would not have been possible.

# Notations

$x$	denotes a scalar
$\mathbf{x}$	denotes a vector
$\mathbf{X}$	denotes a matrix or a linear operator
$\mathbf{x}^T, \mathbf{X}^T$	denotes the transpose of a vector $\mathbf{x}$ or a matrix $\mathbf{X}$
$\mathbf{X}^*$	denotes the conjugate transpose (or adjoint) of a matrix (or operator) $\mathbf{X}$
$x_i, (\mathbf{x})_i$	denotes the $i^{\text{th}}$ component of a vector $\mathbf{x}$
$\mathbf{x}_i$	denotes the $i^{\text{th}}$ element of a vector set
$ x $	denotes the absolute value of a scalar $x$
$x^*$	denotes the complex conjugate of a scalar $x$
$\tilde{\mathbf{x}}, \tilde{\mathbf{X}}$	denotes the optimal value of a vector $\mathbf{x}$ or a matrix $\mathbf{X}$
$\hat{\mathbf{x}}, \hat{F}$	denotes the Fourier transform of a vector $\mathbf{x}$ or a function $F$
$F^*$	denotes the convex conjugation of a convex function $F$
$\partial F$	denotes the subgradient of a function $F$
$\mathcal{C}$	denotes a set
$ \mathcal{C} $	denotes the cardinality of a set $\mathcal{C}$ , measuring the number of elements of the set
$\mathcal{C} \setminus \mathcal{X}$	denotes the relative complement of $\mathcal{X}$ in $\mathcal{C}$ . This set contains all those elements of $\mathcal{C}$ that are not in $\mathcal{X}$ , i.e., $\mathcal{C} \setminus \mathcal{X} = \{x \in \mathcal{C}   x \notin \mathcal{X}\}$
$[N]$	denotes the set of indices in $\mathbb{R}^N$ , i.e., $[N] = \{1, \dots, N\}$
$\mathbb{N}$	denotes the set of all natural numbers
$\mathbb{R}$	denotes the set of all real numbers
$\mathbb{C}$	denotes the set of all complex numbers
$\mathbb{Z}$	denotes the set of all integers

$\mathbb{R}_+$	denotes the set of the non-negative real numbers
$\mathbf{I}_N$	denotes the $N$ -by- $N$ identity matrix
$\mathbf{0}_N$	denotes a vector in $\mathbb{R}^N$ with all elements equal to zero
$\mathbf{1}_N$	denotes a vector in $\mathbb{R}^N$ with all elements equal to one
$\mathbf{x}_C$ ( $\mathbf{X}_C$ )	denotes the vector (resp. matrix) obtained by retaining the components (resp. columns) of $\mathbf{x} \in \mathbb{R}^D$ (resp. $\mathbf{X} \in \mathbb{R}^{D' \times D}$ ) belonging to the subset $C \subset [D]$ . It is alternatively denoted as $\mathbf{x}_C = \mathbf{R}_C \mathbf{x}$ (resp. $\mathbf{X}_C = \mathbf{X} \mathbf{R}_C^T$ ), where $\mathbf{R}_C := ((\mathbf{I}_D)_C)^T \in \{0,1\}^{ C  \times D}$ is the <i>restriction operator</i>
$\text{supp } \mathbf{x}$	denotes the support of a vector $\mathbf{x} \in \mathbb{R}^N$ , which is defined as $\text{supp } \mathbf{x} = \{i \in [N] : x_i \neq 0\}$
$\text{Ker } \mathbf{A}$	denotes the kernel or <i>null space</i> of a matrix $\mathbf{A}$ , which is defined as $\text{Ker } \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^N : \mathbf{A}\mathbf{x} = \mathbf{0}\}$
$\lambda(\mathbf{X})$	denotes the vector containing the eigenvalues of matrix $\mathbf{X}$
$\lambda_M(\mathbf{X})$	denotes the largest (in absolute value) eigenvalue of a matrix $\mathbf{X}$
$\ \mathbf{x}\ _p$	denotes the $\ell_p$ norm of a vector $\mathbf{x}$ , which is defined as $\ \mathbf{x}\ _p = (\sum_i  x_i ^p)^{1/p}$ for $p \geq 1$
$\ \mathbf{x}\ _0$	denotes the $\ell_0$ counting function, which is defined as the cardinality of the support of $\mathbf{x}$ , <i>i.e.</i> , $\ \mathbf{x}\ _0 =  \text{supp } \mathbf{x} $
$\ \mathbf{X}\ _F$	denotes the Frobenius norm of a matrix $\mathbf{X}$ , which is defined as $\ \mathbf{X}\ _F = \sqrt{\sum_i \sum_j  x_{ij} ^2}$
$\ \mathbf{K}\ $	denotes the induced norm of an operator $\mathbf{K}$ , defined as $\ \mathbf{K}\  = \max_{\mathbf{x}: \ \mathbf{x}\ _2=1} \ \mathbf{K}\mathbf{x}\ _2 = \sqrt{\lambda_M(\mathbf{K}^* \mathbf{K})}$
$\mathbf{x}^T \mathbf{y} = \mathbf{x} \cdot \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$	denotes the scalar product between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ for some dimension $N \in \mathbb{N}$
$\mathbf{x} \otimes \mathbf{y} = \mathbf{y} \otimes \mathbf{x}$	denotes the convolution between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ for some dimension $N \in \mathbb{N}$
$C^k$	denotes the class of $k$ times continuously differentiable functions
$\Gamma_0(\mathcal{V})$	denotes the class of proper, convex and lower-semicontinuous functions from a finite dimensional vector space $\mathcal{V}$ ( <i>e.g.</i> , $\mathbb{R}^N$ ) to $(-\infty, +\infty]$

$\text{int } \mathcal{C}$	denotes the interior of a set $\mathcal{C}$
$\partial \mathcal{C}$	denotes the boundary of a set $\mathcal{C}$
$\mathbf{1}_{\mathcal{C}}(\mathbf{x})$	denotes the indicator function of the set $\mathcal{C}$ , which is equal to 1 if $\mathbf{x} \in \mathcal{C}$ and 0 otherwise
$i_{\mathcal{C}}(\mathbf{x})$	denotes the (convex) indicator or characteristic function of the set $\mathcal{C}$ , which is equal to 0 if $\mathbf{x} \in \mathcal{C}$ and $+\infty$ otherwise
$\text{dom } F$	denotes the domain of a function $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$ , which is defined as $\text{dom } F = \{\mathbf{x} \in \mathbb{R}^N : F(\mathbf{x}) < +\infty\}$
$\text{graph } F$	denotes the graph of a function $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$ , which is defined as $\text{graph } F = \{(\mathbf{x}, \lambda) \in \mathbb{R}^{N+1} : F(\mathbf{x}) = \lambda\}$
$\text{epi } F$	denotes the epigraph of a function $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$ , which is defined as $\text{epi } F = \{(\mathbf{x}, \lambda) \in \mathbb{R}^{N+1} : F(\mathbf{x}) \leq \lambda\}$
$(\mathbf{x})_+$	denotes the non-negativity thresholding function defined componentwise as $(x_i)_+ = (x_i +  x_i )/2$
$\nabla$	denotes the discrete gradient or finite difference operator
$\delta(t)$	denotes the continuous delta function or Dirac distribution
$\delta_0(m, n)$	denotes the 2-D discrete delta or Kronecker function, which is equal to 1 if $m = n = 0$ , and 0 otherwise
$\mathcal{N}(0, \sigma^2)$	denotes the standard Gaussian distribution with mean 0 and variance $\sigma^2$
$\mathcal{U}(0, 1)$	denotes the standard Uniform distribution on the interval $[0, 1]$
$\text{mod}$	denotes the modulo operation
$\log$	denotes the logarithm in base 2, also denoted as $\log_2$

x

**Notations**

# Acronyms

1-D	One dimensional
2-D	Two dimensional
3-D	Three dimensional
ADMM	Alternating direction method of multipliers
AIA	Atmospheric imaging assembly
AT	Absorption tomography
APG	Accelerated proximal gradient
AWGN	Additive white Gaussian noise
BSNR	Blurred signal-to-noise ratio
BV	Bounded-variation
CCD	Charge-coupled device
CP	Chambolle-Pock
CS	Compressed sensing
CT	Computed tomography
DFST	Deflectometric Fourier slice theorem
DFT	Discrete Fourier transform
DN	Digital number
DRS	Douglas-Rachford splitting
DWT	Discrete wavelet transform
EM	Electromagnetic
EUV	Extreme ultraviolet
EUVI	Extreme ultraviolet imager
ERS	European remote sensing
FB	Forward-backward
FBP	Filtered back projection
FDM	Frequency deflectometric measurements

FFT	Fast Fourier transform
FoV	Field-of-view
FST	Fourier slice theorem
i.i.d.	Independently and identically distributed
InSAR	Interferometric synthetic aperture radar
ISNR	Increase in signal-to-noise ratio
KL	Kurdyka-Łojasiewicz
l.s.c.	Lower semi-continuous
LS	Least-squares
MAP	Maximum <i>a posteriori</i>
ME	Minimum energy
ML	Maximum likelihood
MSNR	Measurement signal-to-noise ratio
NDFT	Non-equispaced discrete Fourier transform
NFFT	Non-equispaced fast Fourier transform
ODT	Optical deflectometric tomography
PSF	Point spread function
RIM	Refractive index map
RME	Robust median estimator
RSNR	Reconstruction signal-to-noise ratio
SECCHI	Sun Earth connection coronal and heliospheric investigation
SDO	Solar dynamics observatory
SNR	Signal-to-noise ratio
s.t.	Subject to
STEREO	Solar terrestrial relations observatory
TV	Total-variation
UDWT	Undecimated discrete wavelet transform
UT	Universal time
VST	Variance stabilization transform

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgments</b>	<b>v</b>
<b>Notations</b>	<b>vii</b>
<b>Acronyms</b>	<b>xi</b>
<b>Table of contents</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivations . . . . .	4
1.2 Dissertation Outline and Contributions . . . . .	8
<b>2 Inverse Problems and Optimization</b>	<b>13</b>
2.1 Discrete Inverse Problems in Imaging . . . . .	13
2.2 Solving Inverse Problems . . . . .	17
2.3 Regularization of Inverse Problems . . . . .	18
2.3.1 Different Regularized Formulations . . . . .	19
2.3.2 Sparse Regularization . . . . .	20
2.4 Optimization Tools . . . . .	30
2.4.1 Useful Definitions and Theorems . . . . .	30
2.4.2 Convex Optimization through Proximal Algorithms . . .	36
2.4.3 Non-Convex Optimization through Alternating Algorithms . . . . .	43
Appendix 2.A Robust Median Estimator . . . . .	46
Appendix 2.B Discrete Gradient Operator . . . . .	47
Appendix 2.C Wavelets Dictionaries . . . . .	48

Appendix 2.D	Details on the Expanded Version of the Chambolle-Pock Algorithm . . . . .	52
2.D.1	Useful Computations . . . . .	52
2.D.2	Iterations of the Extended Chambolle-Pock Algorithm . . . . .	53
<b>3</b>	<b>Compressive Optical Deflectometric Tomography</b>	<b>55</b>
3.1	Introduction . . . . .	55
3.1.1	Contributions . . . . .	57
3.1.2	Outline . . . . .	58
3.2	Optical Deflectometric Tomography . . . . .	59
3.2.1	Principles . . . . .	59
3.2.2	Deflection Measurements . . . . .	63
3.3	Discrete Forward Model . . . . .	64
3.3.1	Discrete Domains . . . . .	64
3.3.2	Discretized Functions . . . . .	65
3.3.3	Forward Model . . . . .	67
3.4	Related Works . . . . .	69
3.5	Refractive Index Map Reconstruction . . . . .	71
3.5.1	Filtered Back Projection . . . . .	71
3.5.2	Minimum Energy Reconstruction . . . . .	71
3.5.3	TV- $\ell_2$ Minimization . . . . .	72
3.6	Noise Identification, Estimation and Analysis . . . . .	76
3.6.1	Noise Sources . . . . .	76
3.6.2	Comparative Study of the Noise Impact on AT and ODT . . . . .	78
3.7	Numerical Methods . . . . .	79
3.7.1	ODT Numerical Reconstruction . . . . .	80
3.7.2	Adaptive Optimization Procedure . . . . .	82
3.8	Experiments . . . . .	82
3.8.1	Synthetic Data . . . . .	84
3.8.2	Experimental Data . . . . .	100
3.9	Conclusions . . . . .	105
Appendix 3.A	On the Deflectometric Model . . . . .	106
3.A.1	Derivation of the Deflection Equation . . . . .	106
3.A.2	Deflectometric Fourier Slice Theorem . . . . .	108
3.A.3	Ray Tracing . . . . .	108
Appendix 3.B	Non-Equispaced Fourier Transform . . . . .	111

Appendix 3.C Convex Conjugate Functions . . . . .	112
<b>4 Phase Unwrapping</b> . . . . .	<b>115</b>
4.1 Introduction . . . . .	115
4.1.1 Contributions . . . . .	118
4.1.2 Outline . . . . .	118
4.2 Discrete Forward Model . . . . .	118
4.3 Convex Optimization Approach . . . . .	120
4.4 Phase Unwrapping and Denoising Algorithm . . . . .	122
4.5 Experiments . . . . .	124
4.5.1 Synthetic Data . . . . .	125
4.5.2 Experimental Data . . . . .	130
4.6 Conclusion . . . . .	135
<b>5 Blind Deconvolution</b> . . . . .	<b>137</b>
5.1 Introduction . . . . .	138
5.1.1 Contribution . . . . .	141
5.1.2 Outline . . . . .	143
5.2 Problem Statement . . . . .	143
5.2.1 Discrete Model . . . . .	143
5.2.2 Noise Estimation . . . . .	145
5.3 Blind Deconvolution Problem . . . . .	146
5.3.1 Prior Information on the Image . . . . .	146
5.3.2 Constraints on the PSF . . . . .	147
5.3.3 Final Formulation . . . . .	149
5.4 Proximal Alternating Minimization . . . . .	150
5.4.1 Bounded sequence . . . . .	150
5.4.2 Cost-to-Move Parameters . . . . .	151
5.4.3 Stopping Criteria . . . . .	152
5.4.4 Numerical Reconstruction . . . . .	153
5.5 Filter Validation using Non-Blind Deconvolution . . . . .	156
5.6 Experiments . . . . .	157
5.6.1 Synthetic Data . . . . .	157
5.6.2 Experimental Data . . . . .	162
5.7 Conclusion . . . . .	175

Appendix 5.A On the Approximation of the Long-Range PSF Impact by a Constant . . . . .	178
Appendix 5.B On the Convergence of the Proximal Alternating Algorithm . . . . .	180
Appendix 5.C Algorithm Initialization . . . . .	182
5.C.1 Regularization Parameter . . . . .	182
5.C.2 Image and Filter . . . . .	183
Appendix 5.D Numerical Reconstruction Algorithms . . . . .	184
5.D.1 First Subproblem: Image Estimation . . . . .	184
5.D.2 Second Subproblem: Filter Estimation . . . . .	185
Appendix 5.E Convolutions with Unknown Boundary Conditions . . . . .	186
<b>6 Conclusions and Future Work</b>	<b>189</b>
6.1 Summary of the Thesis . . . . .	189
6.2 Future Work . . . . .	192
<b>List of Publications</b>	<b>197</b>
<b>Bibliography</b>	<b>199</b>

# Introduction

# 1

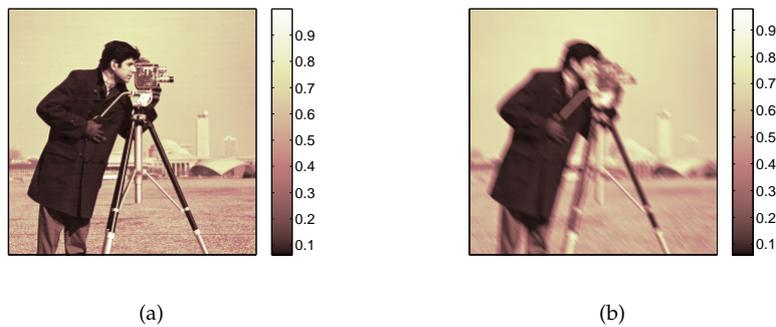
Images play an important role in the human life as they depict a visual perception of the world. An image can be either captured through the human eye and optical devices (*e.g.*, cameras, telescopes, microscopes and mirrors), or generated using a machine (*e.g.*, a computer or a medical imaging device). A wide variety of these images are digital since they are either generated by a digital computer or recorded by a digital device. A digital image is defined as a two-dimensional finite function, where a given object or scene is depicted by associating different intensity values to different locations in a two-dimensional pixel grid.

Imaging applications include all the fields where a digital image is either obtained from some measured data or processed to obtain some relevant information. The importance of digital images captured by modern devices or generated by a machine, is that they allow us to go beyond the human perception, which is restricted to the visible light range (wavelengths of 380-750 nm). The inside of objects and some physical phenomena cannot be imaged using only this limited range of the electromagnetic (EM) spectrum. Existing imaging devices are not only able to cover almost all the EM spectrum, from the gamma rays (wavelengths smaller than  $10^{-2}$  nm) to the radio waves (wavelengths of  $10^{-3}$  -  $10^5$  m), but they also use other types of sources (*e.g.*, electrons and sound waves). Additionally, the current devices are also able to image objects that are very small (in the nm scale) and objects that are located very far (at more than 100 light years from Earth).

In many applications such as medicine, biology, astronomy, security, archeology, art conservation, ecology, geology, oceanography and industrial manufacture; digital images are used for observation, analysis, characterization and detection. For instance, in medicine and biology, we can use X-rays, gamma-rays, magnetic field, ultra-sound waves, visible light, etc., to obtain images of the human body. These images allow the analysis of the anatomic, physiologic or metabolic behavior of the internal organs and structures and help in the diagnosis and treatment of diseases and in surgery guidance [1–5]. Since the T-rays, X-rays and gamma-rays have the ability to penetrate objects, they can also be used to image the interior of objects. We can find their application in airport security systems to detect illegal items or in industrial manufacture processes to analyze failures, detect flaws and perform reverse engineering [1, 6, 7]. In astronomy, extreme ultraviolet rays, visible light, radio waves, etc., are used to acquire images from the Sun, the Earth and many astronomical objects inside and outside the Solar System [8–11]. These images allow us to study physical phenomena in the universe, to observe the characteristics of the planets, stars and galaxies, among many other studies. Optical and electron microscopes and spectrometers are used, for example, to characterize a given biological sample, providing a deeper understanding of microorganisms and biological structures [5, 12, 13]. These are few of the many examples we can find where digital images are employed. Nevertheless, the process of acquiring or generating those images is not trivial. In the following, we describe some problems that appear and challenges to overcome when acquiring or generating digital images.

The image captured by a digital imaging device may be affected by distortions that are intrinsic to the measurement process. For example, let us consider that we capture the scene in Figure 1.1-(a) by a digital camera, but instead of obtaining the actual scene, we acquire the image in Figure 1.1-(b). We can notice that the acquired image is a blurred version of the actual one, it has a low visual quality and most details cannot be distinguished. Similar problems can appear when we acquire images using other digital imaging devices (*e.g.*, telescopes and microscopes). In all these cases, the acquired image is the result of combining the actual image with a function that corresponds to the response of the instrument, also called point-spread function (PSF) [12–18]. The response of the instrument depends on the acquisition pro-

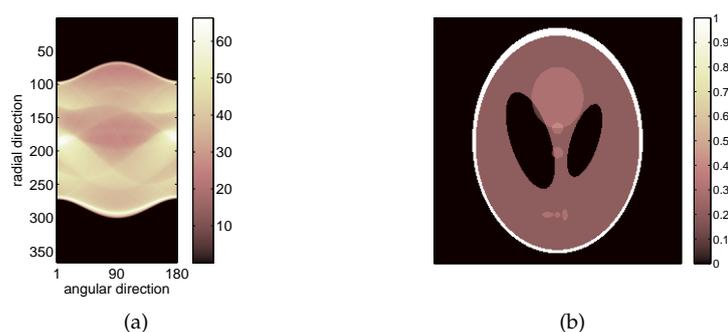
cess and the optical properties of the instrument. It can be a linear motion blur caused by a moving camera (as in Figure 1.1-(b)), an out-of-focus blur caused by objects that are outside the focal plane of the instrument, optical aberrations and light diffraction or scattering caused by instrumental effects, etc. In other applications, such as interferometry and radar imaging, we face other types of distortions where the acquired image is given by a function (*e.g.*, a modulo operation) of the actual image [19,20]. Through a careful study of the acquisition process and the imaging device, we can build an image formation model. Then, the issue is how to “invert” that model in order to remove the distortions and obtain the original “ideal” image.



**Figure 1.1** (a) The cameraman image as an example of a scene. (b) Illustration of a blurred version of the scene in (a).

When a digital image is generated by a machine, the actual acquisition does not correspond to an image. Instead, the acquisition corresponds to a set of reflected or transmitted EM rays, sound waves, magnetic pulses, etc. For illustration, let us consider as an example the absorption tomography (AT), also known as X-ray computed tomography (CT) or CT scan. In this imaging modality, a patient is placed inside a CT scanner, a machine that, for different orientations, emits and collects X-rays traversing a given area of the patient’s body. The X-rays emitted by the machine are absorbed in different proportions by the tissues inside the body, then the machine collects the rays going out of the body. As described by the Beer-Lambert law [4], the collected rays have different intensities depending on how much they have been absorbed by the tissues in the body. Using this information, we can generate an image that

contains the spatial distribution of the absorption coefficient. Since we know in advance the absorption coefficient of a given tissue, the generated image allows the characterization of the tissues inside the body. Figure 1.2-(a) illustrates a typical observation in CT using parallel X-rays, where the horizontal axis corresponds to the orientations of the scanner (angular direction) and the vertical axis corresponds to the different X-rays (radial direction). This observation corresponds to the Shepp-Logan phantom in Figure 1.2-(b). In this case, the forward model that describes how to obtain the observations in Figure 1.2-(a) from the object in Figure 1.2-(b), is well-known to be defined by the Beer-Lambert law. However, the issue is how to invert the forward model in order to generate, from the observations made by the CT scan, an image of the absorption map that can be used by a medical doctor.



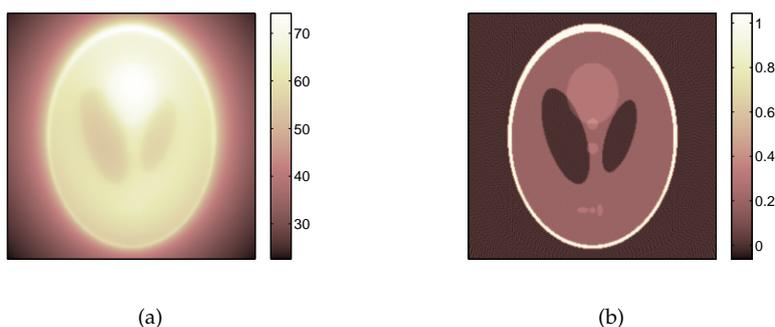
**Figure 1.2** (a) Illustration of a typical observation in absorption tomography. (b) Absorption map that generated the observations in (a).

## 1.1 Motivations

Inverse problems are ubiquitous in imaging applications and accurately solving them is crucial in order to acquire or generate good quality digital images. To solve an inverse problem is not an easy task since there are several challenges that need to be addressed [21]. A first challenge is in the formulation of an accurate forward model that characterizes the physical acquisition and the image formation process and that also has a fast and simple numerical implementation in a computer. Once the forward model has been built, the

next challenges consist of how to obtain the best representation of the original image in a short time (crucial in real-time applications), with low complexity (less consumption of computational resources) and with the lowest impact to the object being imaged (*e.g.*, no damage to the patient, object or sample being imaged) or to the application (*e.g.*, not affecting the manufacturing time).

At first though, one would like to estimate the original image by performing a “simple inversion” through fast analytic methods that are based on the characteristics of the forward model. In most inverse problems, however, due to the mathematical properties of the forward model, it is not possible to perform such inversion and the resulting image does not provide any useful information. Let us observe what happens in the example of the CT scan from Figure 1.2 if we perform a “simple inversion” of the forward model. The CT forward model for the observations in Figure 1.2-(a) consists of projections of the Shepp-Logan phantom for different orientations. These projections are given by the collection of integrals of the parallel X-rays [22, Section 3.1]. A naive inversion of such model, consists on the *back projection* of the measurements for all the orientations, which results in the image of Figure 1.3-(a). We notice that the estimated image cannot even be compared with the actual image from Figure 1.2-(b) and no useful information can be extracted from it.

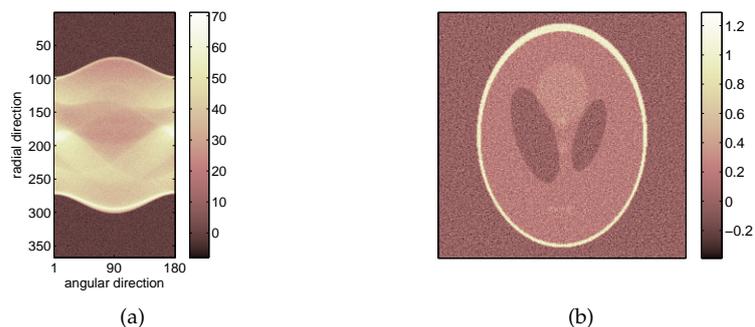


**Figure 1.3** Illustration of simple analytic inversions of the CT forward model. Reconstruction of the Shepp-Logan phantom from the observations in Figure 1.2-(a) using (a) the back projection method and (b) the filtered back projection method.

A more involved study of the forward model is therefore necessary in order to generate a useful digital image. In the CT reconstruction, there exists

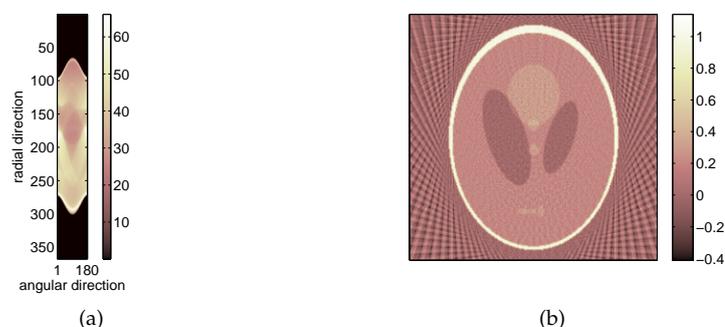
a common analytic approach that allows improving the quality of the estimated image by adding an adequate filter to the back projection method. This method, called filtered back projection (FBP) [22, Section 3.3], provides the image in Figure 1.3-(b). This image has a higher quality compared to the one in Figure 1.3-(a). Nevertheless, when comparing it with the original image in Figure 1.2-(b), we evidence artifacts, variations in the pixels intensities and negative values that are not plausible given the physical characteristics represented in the image. These results are a clear example of the limited application of non-regularized inversion methods to obtain a good representation of the image of interest.

In addition to the difficulty of inverting the forward model, there may also be distortions caused by the measurement device and by the acquisition and modeling processes. Figure 1.4-(a) depicts an example of a CT measurement presenting random fluctuations from their actual value that are due to the electronics in the imaging device. After applying the FBP method to such a noisy observation, we obtain the image in Figure 1.4-(b). Analytic inversion methods do not take into account the noise in the reconstruction process. Hence, the noise tends to be amplified, affecting the intensity of the pixels and preventing a clear distinction of small structures in the image. A good inversion method should take into account all the possible distortions that are affecting the acquisition and modeling processes.



**Figure 1.4** Illustration of analytic inversion methods applied to noisy observations. (a) Noisy CT observations. (b) Reconstruction of the Shepp-Logan phantom from the observations in (a) using the filtered back projection method.

The quality of the digital image can also be affected by the acquisition time, which is preferably short. For instance, in medicine and biology, a short acquisition time is required in order to avoid having a negative impact on the patient or biological sample. A short acquisition time is also desirable in industrial manufacturing, so that the imaging technique does not affect the object being imaged neither the manufacturing time. One of the consequences of forcing a short acquisition time is that less information is obtained, yielding to incomplete observations. Figure 1.5 depicts an example of how an analytic inversion method, such as FBP, deals with an incomplete set of CT observations. In Figures 1.2-(a) and 1.4-(a), we show the full set of observations “theoretically” required for a good CT reconstruction, which corresponds to 180 orientations of the scanner equally sampled in the interval  $[0, \pi)$  radians. The example in Figure 1.5-(a) contains only 25% of the full set of observations, corresponding to 45 angles equally sampled in the interval  $[0, \pi)$  radians. Figure 1.5-(b) depicts the resulting image from applying the FBP method to the incomplete measurements in Figure 1.5-(a). As in the noisy case, the FBP method is not able to provide a good representation of the original image when there is a lack of observations.



**Figure 1.5** Illustration of analytic inversion methods applied to incomplete observations. (a) Incomplete CT observations (25% of the full set of observations). (b) Reconstruction of the Shepp-Logan phantom from the observations in (a) using the filtered back projection method.

An analytic method is therefore not ideal when enhancing an acquired image or generating a digital image from incomplete or deteriorated observations. We believe that a good inversion can be achieved by adapting the fol-

lowing framework. First, a dedicated study of the underlying physical model must be performed, including the analysis of the existing distortions. Then, the inverse problem should be formulated as a minimization of a distance function based on the error model and on some additional functions and constraints that, based on the prior information available on the original image, decrease the amount of possible solutions to obtain the best approximation of the original image. Finally, the resulting minimization problem needs to be solved by means of efficient algorithms that guarantee the convergence to an approximate solution in a reasonable number of iterations.

This general framework has been successfully applied in existing applications, such as medicine and biology [2,4,13,23], photo enhancement [14,15,24], astronomy [11, 16, 25], industrial manufacturing [26, 27], etc. Nevertheless, each inverse problem is inherently different and needs to be studied and treated independently in order to guarantee the best results. Additionally, there are still many challenges to overcome in all those applications and in new ones yet to be discovered, which generate new inverse problems that need to be solved.

## 1.2 Dissertation Outline and Contributions

In this dissertation, we study several inverse problems existing in imaging applications and we investigate how to solve those problems such that the obtained solution is a good approximation of the original image and it is also robust with respect to the distortions present in the measurement and modeling processes. In each application, we show the advantage of the proposed approach against traditional non-regularized methods proposed in the literature. We outline the main contributions of this dissertation as follows.

In Chapter 2, we provide an overview of the literature related to solving inverse problems in imaging. We start by a formal description of discrete inverse problems and the process of building the image formation model, taking into account the possible distortions encountered in imaging applications. Next, we explain how inverse problems are generally solved for a given noise model and how to regularize them in order to obtain a meaningful solution.

We finally conclude this chapter with a description of the optimization tools used in the thesis to solve inverse problems.

In Chapter 3, we study an inverse problem existing in optical deflectometric tomography (ODT). ODT is a novel imaging modality that allows the optical characterization of a transparent material by using an image of the refractive index spatial distribution. We aim at obtaining a high quality digital image of the refractive index map from corrupted and incomplete observations that consist of light deflection measurements for multiple orientations. Based on the existing literature, we first perform a dedicated study of the physical acquisition process. Then, we formulate a simplified linear forward model and provide some insight about the impact of the adopted approximations. Since this modality is fairly new in the literature, little effort has been done in the discretization of the forward model and in the reconstruction of the refractive index map. Hence, we provide a detailed explanation about the discretization of the image grid and the formulation of an accurate discrete forward model that characterizes the acquisition process. We then perform a thoughtful analysis of the distortions present in the acquisition and modeling processes. Finally, we propose a constrained optimization method that enforces the proximity between the observations and the forward model while promoting a certain structure in the image based on the physical properties of the refractive index map. The proposed approach is shown to have a unique solution and it outperforms traditional non-regularized tomographic reconstruction methods for both synthetic and experimental data. The contents in this chapter are largely based on [28].

In Chapter 4, we analyze the unwrapping inverse problem appearing in imaging applications where the phase image is not directly observed but is computed from the output of the instrument. In several imaging applications such as interferometry, optical deflection tomography and radar imaging, the phase carries useful information and an accurate estimation of the phase image is crucial in order to avoid the propagation of errors in subsequent processing stages. For instance, in the imaging modality described in Chapter 3, the light deflection is not directly measured by the imaging device but it is obtained from the phase image. Hence, an accurate estimation of the refractive index is directly dependent on the correctness of the computed phase image. When a phase image is indirectly observed, its intensity can only be

estimated inside a certain interval (e.g.,  $[-\pi, \pi]$ ). If the actual intensity values of the phase image are located outside that interval, then the phase is said to be “wrapped” through a modulo operation. Existing unwrapping methods in the literature are based on analytic solutions, combinatorial optimization and graph-cut techniques, which are highly affected by the noise and by the presence of discontinuities in the phase image. We propose a robust phase unwrapping method that aims at simultaneously unwrap and denoise the phase image. The forward model is formulated in the phase derivative domain, which allows relaxing this highly non-linear and non-convex problem. This relaxation introduces some additional errors that are taken into consideration by the proposed approach. The problem is formulated as a constrained optimization that enforces the fidelity of the model to the observations, taking into account the characteristics of the noise in the phase image and the relaxation errors, while promoting a certain structure in the phase image to avoid reconstructing fake jumps. The performance of the proposed approach is illustrated through synthetically generated data and also through experimental data from interferometric synthetic aperture radar (InSAR) imaging. The results show that the proposed approach compares favorably with state-of-the-art methods that perform the unwrapping and denoising in two separate steps. The contents in this chapter are largely based on [29].

In Chapter 5, we study the blind deconvolution problem, an inverse problem that appears in every imaging application where the acquired image is the result of the convolution between the original image and the response of the imaging device. This issue is common in photography, astronomy and microscopy, among other applications. The response of the imaging device, also called point spread function (PSF), depends on instrumental effects and on problems during the acquisition process. The PSF can sometimes be estimated from the optical properties of the instrument, however, in general, it is unknown. In this chapter, we focus on the astronomical imaging application, where we aim at estimating the PSF of solar telescopes from noisy observations using blind deconvolution techniques. The estimated PSF is then used to enhance the images that have been acquired by the telescope. In contrast to most common approaches in the literature, which are based on building a parametric model of the PSF, the proposed method relies on the prior information available on the image while imposing only soft constraints on the PSF,

making the method general and applicable to any instrument. Crucial to the PSF estimation is the acquisition of images from a solar transit, where we have access to the specific values of well localized pixels in the image (such characteristics can also be found in images from other applications such as computer vision and medical imaging). The blind deconvolution is formulated as a non-convex optimization problem, where we enforce a small distance between the observations and the model and we promote certain conditions on both the image and the PSF. This problem is solved through a recently proposed alternating algorithm, which allows us to guarantee that a critical point of the problem is always reached. Additionally, we study an innovative stopping criterion that is based on the spectral characteristics of the residual image, given by the difference between the observation and the model at each iteration. We show, through synthetic and experimental data, that the non-parametric blind deconvolution method proposed in this chapter is able to estimate the core of the PSF with a similar quality to parametric methods proposed in the literature. The contents in this chapter are largely based on [30].

In Chapter 6, we summarize the main contributions of this dissertation and we provide some insight for future research works.



# Inverse Problems and Optimization

## 2

In this chapter, we provide a brief overview on the literature related to solving inverse problems in imaging. We first present a formal description of discrete inverse problems, emphasizing on the image formation model and the distortions typically encountered in the literature. Then we explain how inverse problems are generally solved for a given noise model and how to regularize them in order to obtain a meaningful solution. The chapter concludes with a description of the optimization tools used in the thesis to solve inverse problems.

## 2.1 Discrete Inverse Problems in Imaging

In many imaging applications, the desired digital image is not directly accessible for measurement. Instead, it needs to be recovered from corrupted, incomplete and/or indirect observations that are connected to the desired image by physical laws. In this context, the notion of inverse problems corresponds to the determination of the unknown image from the available observations by “inverting” these physical laws [21].

Mathematically, an imaging acquisition process can be described as a function  $F$  that is applied to a physical quantity  $x$  that varies continuously in time and/or space, producing an observation

$$z = F(x). \tag{2.1}$$

The problem of finding  $z$  given  $x$  described by (2.1) is known as the direct or forward problem. The function  $F(\cdot)$  may correspond to a differential operation, an integration or any other mathematical transformation [21, 31]. For instance, in the example of Figure 1.1,  $F(\cdot)$  represents an integration of portions of the actual scene weighted by the response of the imaging device (a linear motion blur); while in the example of Figure 1.2, it represents the integration of the absorption coefficient along the trajectory of the X-rays, as described by the Beer-Lambert law [4]. In both examples, thanks to the convolution theorem [12] and the Fourier slice theorem [22, Section 3.2], function  $F(\cdot)$  can be replaced by a linear map based on the Fourier transform. Since the Fourier transform has a fast computation, these are good examples of the desirable behavior of function  $F(\cdot)$ . Therefore, many efforts are dedicated to finding a fast linear map  $A$  between the object of interest  $x$  and the observations  $z$ , that can accurately characterize the acquisition process as follows:

$$z = Ax. \quad (2.2)$$

Nonetheless, in those examples and in general imaging applications, the physical model cannot be exactly represented by the continuous forward model in (2.2). First, due to the physical limitations of the imaging devices, a given object of interest can only be observed in a limited area of space called the field-of-view (FoV). This involves some boundary conditions that need to be considered in the modeling process. Second, since the imaging devices consist of a finite number of sensing elements, the observed data is discretized in space. In what follows, we describe in more detail the spatial discretization process, which will allow formulating a more accurate forward model.

The spatial discretization of the observations is due to the limited number of sensors in an actual imaging device, where the value recorded at each sensor is proportional to the intensity of the waves (*e.g.*, light rays, sound waves, etc.) traversing or reflecting off *objects*. In a digital camera, for example, each sensor corresponds to a pixel in the acquired image, with values proportional to the intensity of light striking at that point. In the CT example of Figure 1.2, the spatial discretization of the measurements corresponds to the number of sensors available to capture the intensity of a different X-ray and the number of scanner orientations where the acquisitions are made.

The acquisitions from an imaging device are, therefore, sampled in a  $M_1 \times M_2$  grid of  $M := M_1 M_2$  samples. This measurement grid can be, for instance, a regular Cartesian grid, as in the example in Figure 1.1, or a polar grid, as in the CT example in Figure 1.2. The spacing of the elements in the measurement grid depends on the physical characteristics of the imaging device, which will also determine the FoV of observation, the spatial sampling of the FoV and the resolution of the digital image that we aim at recovering.

We consider, hereafter, that the object of interest  $x$ , observed in the experimental FoV (denoted  $\Omega$ ), is sampled in a regular  $N_1 \times N_2$  2-D Cartesian grid of  $N := N_1 N_2$  samples, described as

$$\mathcal{C}_N = \{\mathbf{r}_{m,n} := (m \delta r, n \delta r) : -N_1/2 \leq m \leq N_1/2, -N_2/2 \leq n \leq N_2/2\}, \quad (2.3)$$

where the spatial spacing  $\delta r$  is adjusted to  $\Omega$  and to the resolution of the device. This 2-D description can always be arranged into a one-dimensional vector  $\mathbf{x} \in \mathbb{R}^N$ , given a convenient mapping between the components of indices  $\mathbf{x} = (x_1, \dots, x_N)^T$  and the pixel coordinates in  $\mathcal{C}_N$ . This “vectorization” allows having a simplified representation of any linear operator acting on the sampled data as a matrix multiplying  $\mathbf{x}$ . Similarly, the 2-D description of the measurements can be “vectorized” into the one-dimensional vector  $\mathbf{z} \in \mathbb{R}^M$ . The physical model, represented by the map  $A$  in (2.2), is also discretized as the linear operator  $\mathbf{A} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ . The discrete forward model results in

$$\mathbf{z} = \mathbf{A}\mathbf{x}. \quad (2.4)$$

We should note that the model in (2.4) is an idealization of the actual discrete model. In some applications [26, 32], the actual sensing model goes through slight modifications or approximations in order to obtain a linear, efficient and computationally tractable model represented by operator  $\mathbf{A}$ . In other applications [15, 33], the sensing operator  $\mathbf{A}$  is not exactly known *a priori*. This introduces distortions and the actual observations do not exactly match the model  $\mathbf{A}\mathbf{x}$ .

Moreover, there are external distortions intrinsic to the physical measurement process. One example is the noise caused by the electronics in the measurement devices (thermal noise, dark current and read-out noise). This

noise is typically characterized by an additive white Gaussian noise (AWGN) model [16]. Another example is the shot noise caused by the random process of counting the photons that arrive to the sensor. The probability of arrival of the photons is governed by a Poisson distribution, making it more apparent when the number of collected photons is relatively small. As the number of photons increases, the probability distribution of the shot noise approximates a standard Gaussian function and the AWGN model holds.

In imaging applications, we can expect to have a combination of both Poisson and Gaussian noise models [16,34,35]. In some cases, the noise is assumed to follow a pure Poisson distribution, either because the measurement devices are designed such that the electronic noise can be neglected [4] or because the Poisson process can include in its mean value the Gaussian part of the noise [25,34]. In most cases, however, the noise model is assumed to be pure Gaussian. The Poisson-Gaussian model can be approximated as pure Gaussian either by considering that the photon count is high enough such that the shot noise has a normal distribution [20,36,37], or by introducing a Variance Stabilization Transform (VST) in the model (2.5) [16,35,38]. Due to its simplicity, an additive (Gaussian) model is also typically adopted in scenarios where the true noise model is unknown (*e.g.*, [3,14,24,39,40]).

The noise models discussed above are the most common ones in the literature. Nevertheless, depending on the application, there exist other types of noise models that can be adopted. Some examples include the multiplicative noise and the quantization noise.

Taking into account the external noise and the errors caused by the discretization and approximations of the modeling process, the observation model (2.4) becomes

$$\mathbf{y} = \Theta(\mathbf{z}) = \Theta(\mathbf{Ax}), \quad (2.5)$$

where the function  $\Theta$  accounting for the distortions and the corrupting noise, is defined according to the considered noise model. For example, under an AWGN assumption, we have  $\Theta(\mathbf{u}) := \mathbf{u} + \boldsymbol{\eta}$ , with  $\boldsymbol{\eta} \in \mathbb{R}^M$  a random variable with normal distribution of mean zero and variance  $\sigma^2$ , *i.e.*,  $\eta_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$ . Under a Poisson noise assumption, the function is described as  $\Theta(\mathbf{u}) := \mathcal{P}(\mathbf{u})$ , which provides a random variable with Poisson distribution of mean  $\mathbf{u}$ .

In the following section, we explain how to formulate an optimization problem that, based on the adopted noise model, allows recovering the image of interest  $\mathbf{x}$  from the corrupted observations in (2.5).

## 2.2 Solving Inverse Problems

Inverse problems of the form (2.5) are often solved by minimizing the distance between the observations and the model

$$\tilde{\mathbf{x}} \in \underset{\mathbf{x}}{\operatorname{argmin}} D(\mathbf{y}, \mathbf{A}\mathbf{x}). \quad (2.6)$$

The distance function  $D(\mathbf{y}, \mathbf{A}\mathbf{x})$ , also referred as data fidelity function, is defined with respect to the noise model. In the thesis, we assume an AWGN model. Therefore, hereafter, the distance function is measured by the  $\ell_2$  norm<sup>1</sup> [39,41,42], and the optimization problem in (2.6) becomes

$$\tilde{\mathbf{x}} \in \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2. \quad (2.7)$$

This problem is commonly referred to as least-squares (LS) problem. Let us note that, if we consider matrices instead of vectors [33], the  $\ell_2$  norm can be replaced by the Frobenius norm, defined as  $\|\mathbf{X}\|_F = \sqrt{\sum_i \sum_j |x_{ij}|^2}$ .

The formulation in (2.7) derives from the maximum likelihood (ML) estimation under an AWGN assumption [15,16,43]. Nevertheless, the LS formulation can also be adopted in situations when there is a lack of noise model but the distortions in the observations are known to be bounded by the  $\ell_2$  norm [3,39].

The LS problem in (2.7) has a unique minimizer if operator  $\mathbf{A}$  has a trivial *null space*, *i.e.*, if  $\operatorname{Ker} \mathbf{A} = \{\mathbf{0}_N\}$ . However, in many practical situations, operator  $\mathbf{A}$  has a non-trivial *null space*, *i.e.*,  $\operatorname{Ker} \mathbf{A} \neq \{\mathbf{0}_N\}$ , either due to the properties of the operator (*e.g.*, the gradient operator contains the constant vectors in its *null space*) or because the problem is under-determined, *i.e.*, we have more unknowns than equations ( $M < N$ ). In those cases, the LS problem in (2.7) has infinitely many solutions and is thus ill-posed, *i.e.*, it does not satisfy the well-posedness conditions of Hadamard [21,31]: (i) a solution exists for any  $\mathbf{y}$ ; (ii)

<sup>1</sup>The  $\ell_p$  norm of a vector  $\mathbf{x}$  is defined as  $\|\mathbf{x}\|_p = (\sum_i |x_i|^p)^{1/p}$  for  $p \geq 1$ .

the solution is unique; and *(iii)* the solution has a continuous dependence on the observed data. Moreover, if operator  $\mathbf{A}$  has a large condition number, the problem is said to be ill-conditioned. These problems are unstable since small errors in the observed data (*e.g.*, due to some noise) cause large variations in the solution  $\tilde{\mathbf{x}}$  of (2.7).

In the thesis, we focus on solving ill-posed inverse problems of the form (2.7) that do not have a unique solution due to under-determinedness or due to the properties of the sensing operator or model. In order to be able to find a meaningful solution of  $\mathbf{x}$  among the infinitely many possible solutions, the problem in (2.7) needs to be regularized as explained in the following section. The regularization is also suitable for handling the problem of ill-conditioning [39].

## 2.3 Regularization of Inverse Problems

The regularization of inverse problems refers to the introduction of additional information to the ill-posed problem in (2.7). This helps to reduce the amount of possible solutions to those that are meaningful [39, 44, 45]. The regularization of the problem in (2.7) is usually done by adding constraints and/or penalty (regularization) functions to the data fidelity function. This additional information is based on prior knowledge we have on the image  $\mathbf{x}$ .

One of the most common forms of regularization, called the Tikhonov regularization, introduces an  $\ell_2$  norm as penalty function in order to retrieve either smooth or minimum energy solutions [45–47].

Another common form of regularization aims at recovering solutions with a certain structure. This is achieved by promoting a low-complexity model on the solution. In this context, we have the sparsity-based regularization, which is at the core of this work and is discussed in detail in Section 2.3.2.

Other forms of regularization include: accumulating several observations that follow the same sensing model (*e.g.*, multi-frame scenarios in [15, 16, 24]), constraining the solution to follow a specific model (*e.g.*, [16, 48, 49]), forcing some known values in the solution (*e.g.*, [16]), normalizing the solution (*e.g.*,

[33, 34]) and bounding the solution (*e.g.*, non-negativity constraint or other bounds [11, 24, 27, 33, 34, 50]).

In the following subsections, we describe different regularized versions of the problem in (2.7) and then we focus on the regularization based on promoting sparsity.

### 2.3.1 Different Regularized Formulations

A convex regularization function  $F(\mathbf{x})$  can be added to problem (2.7) in different ways. We can either have a constrained formulation

$$\min_{\mathbf{x}} F(\mathbf{x}) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{Ax}\|_2 \leq \varepsilon, \quad (2.8)$$

or a Lagrangian formulation

$$\min_{\mathbf{x}} \rho F(\mathbf{x}) + \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_2^2. \quad (2.9)$$

The parameter  $\varepsilon$  in the constrained formulation directly depends on the noise statistics (*e.g.*, the noise variance) while the parameter  $\rho$  in the Lagrangian formulation (called the Lagrangian multiplier) is a regularization parameter that controls the trade-off between the fidelity to the observations and the regularization of the solution. For each value of  $\varepsilon$ , there exists a value of  $\rho$  such that both problems (2.8) and (2.9) have a common solution (see [51, 52] and references therein).

From a Bayesian point of view, the Lagrangian formulation in (2.9) is equivalent to the maximum *a posteriori* (MAP) estimate for an AWGN model, which corresponds to the stabilization of the ML estimate in (2.7) by adding the prior distribution of the image  $\mathbf{x}$  [15, 43]. Due to its unconstrained form, the Lagrangian formulation in (2.9) is often preferred over the constrained one [15, 26, 32, 53, 54]. Also, the Lagrangian formulation is needed when the numerical algorithms used to solve the problem require the data fidelity term to be differentiable with a Lipschitz continuous gradient (see, *e.g.*, the algorithms in Sections 2.4.2.1 and 2.4.3). However, the regularization parameter  $\rho$  in the Lagrangian formulation can be sometimes difficult to estimate. Although several works have studied the choice of this parameter [40–42, 53, 55], there is no general strategy to set its value even when the noise level is known. This presents a disadvantage in the use of this formulation.

The constrained formulation in (2.8) is sometimes advantageous since the parameter  $\varepsilon$  can be directly estimated from the noise statistics (if they are known) and the data fidelity term is stronger. Under an AWGN assumption, where the noise  $\boldsymbol{\eta} \in \mathbb{R}^M$  has a normal distribution of mean zero and variance  $\sigma^2$ , the value of  $\varepsilon$  can be estimated using the Chernoff-Hoeffding bound [56]:

$$\varepsilon := \|\boldsymbol{\eta}\|_2 = \sigma\sqrt{M + c\sqrt{M}}, \quad (2.10)$$

which holds with high probability for  $c = \mathcal{O}(1)$ . The noise variance  $\sigma^2$  can be either known in advance or estimated using, for instance, a robust median estimator (see Appendix 2.A). Some algorithms, such as the proximal algorithms described in Section 2.4.2, are able to handle the constraint in (2.8). This has allowed the constrained formulation to be more frequently used in the literature [11, 27]. The constrained problem can be equivalently formulated by minimizing the fidelity function and expressing the regularization as a constraint, *e.g.*, we can constrain the solution to have a certain (known) level of sparsity [52, 57].

Note that additional constraints or penalty functions can be added to both formulations (2.8) and (2.9), resulting in expanded constrained or Lagrangian formulations or mixed ones.

In the explanations that follow we use, for simplicity, the constrained formulation in (2.8) but they can be extended to other formulations.

### 2.3.2 Sparse Regularization

An image  $\mathbf{x} \in \mathbb{R}^N$  is considered to be sparse if most of its entries are equal to zero and it is  $K$ -sparse if it has exactly  $K$  non-zero elements, with  $K \ll N$ . This definition can be relaxed by considering the notion of weakly sparse or compressible images, consisting of images where most of the entries are close to zero and only a few of them have a large magnitude.

In practice, most images are not sparse but can be “sparsified” using a system of elementary functions, *i.e.*, a “dictionary” [44, 45, 58]. This can be done in two ways. An image  $\mathbf{x}$  is sparse in a dictionary  $\boldsymbol{\Psi} : \mathbb{R}^W \rightarrow \mathbb{R}^N$  if it can be represented by few elements of  $\boldsymbol{\Psi}$ , *i.e.*, if there exists a sparse vector  $\boldsymbol{\alpha} \in \mathbb{R}^W$

such that

$$\mathbf{x} = \mathbf{\Psi}\boldsymbol{\alpha}. \quad (2.11)$$

The expression above corresponds to the *synthesis* relation because it is based on the *synthesis* of the signal  $\mathbf{x}$  with few atoms from a dictionary  $\mathbf{\Psi}$ .

We can also have an alternative expression, the *analysis* relation, where a sparse vector of coefficients  $\bar{\boldsymbol{\alpha}}$  is associated to an image  $\mathbf{x}$ , *i.e.*,

$$\bar{\boldsymbol{\alpha}} = \mathbf{\Psi}^* \mathbf{x}, \quad (2.12)$$

with  $\mathbf{\Psi}^*$  the adjoint of operator  $\mathbf{\Psi}$ , also called the *analysis* operator. This model can be interpreted as the image  $\mathbf{x}$  having sparse *projections* over a dictionary  $\mathbf{\Psi}$ . The *analysis* model has been recently investigated in terms of *co-sparsity* [59], which characterizes the coefficients vector  $\bar{\boldsymbol{\alpha}} \in \mathbb{R}^W$  in terms of the number of zero entries ( $W - K$ ) rather than the number of non-zeros ( $K$ ). We refer the reader to [59] for a more complete discussion on this subject.

We must be careful not to confuse the *synthesis* and the *analysis* relations since they are complementary but intrinsically different. In general, we have  $\bar{\boldsymbol{\alpha}} \neq \boldsymbol{\alpha}$ , and the equality is only reached when the dictionary  $\mathbf{\Psi}$  is an orthonormal basis, *i.e.*, if  $\mathbf{\Psi}^{-1} = \mathbf{\Psi}^*$ .

In the following, we provide more insight on the *synthesis* and *analysis* formulations and their impact in the regularization of inverse problems. We also describe the choice of the dictionary  $\mathbf{\Psi}$  and provide an overview of the functions used in the literature to promote sparse solutions.

### 2.3.2.1 Synthesis and Analysis Formulations

Let us consider a function  $F(\cdot)$  whose minimization promotes sparsity. The problem of recovering a sparse signal  $\mathbf{x}$  from the observations in (2.5) under an AWGN model is commonly formulated using the *synthesis*-based sparsity as [40, 44, 60]:

$$\tilde{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} F(\boldsymbol{\alpha}) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{\Psi}\boldsymbol{\alpha}\|_2 \leq \varepsilon, \quad (2.13)$$

where the solution is given by  $\tilde{\mathbf{x}} = \mathbf{\Psi}\tilde{\boldsymbol{\alpha}}$ .

If we consider the *analysis*-based sparsity model in (2.12), the minimization problem is formulated as [11, 27, 43]:

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} F(\mathbf{\Psi}^* \mathbf{x}) \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \varepsilon. \quad (2.14)$$

Note that, in opposition to the *synthesis* formulation, the *analysis* prior allows working directly on the domain of  $\mathbf{x}$  without increasing the dimension of the optimization problem. This is of particular importance when the sparsity dictionary  $\mathbf{\Psi}$  is redundant (see Section 2.3.2.2), *i.e.*, when  $W > N$ .

Since the *synthesis* and *analysis* models in (2.11) and (2.12) are only equivalent when  $\mathbf{\Psi}$  is orthonormal, then it is intuitive to assume that the same can be said about the *synthesis* and *analysis* formulations in (2.13) and (2.14). Elad *et al.* [43] have provided the theoretical and geometrical proofs of the differences between the two formulations.

It is not clear which formulation is better for which application [43, 45], however, better inversion results have been obtained using the *analysis*-based prior with redundant dictionaries for different inverse problems [11, 27, 43, 61]. This may be explained by the compactness promoted by the *synthesis* formulation. Since the *synthesis* formulation aims at retrieving the image that has the sparsest representation over the dictionary, this means that every element of the dictionary significantly contributes to the correct representation of the image and, hence, to its recovery. In this case, every small error in the choice of a certain dictionary element, which is highly probable in noisy scenarios, can cause a serious deviation from the desired image. This issue may become more important when using over-complete dictionaries, since they allow representing a wider variety of features (see Section 2.3.2.2 for more details). Instead, the *analysis* formulation uses all the dictionary elements simultaneously when describing the image. This minimizes the individual impact of each element of the dictionary in the image recovery process. See [43, 59] for more details on this comparison.

### 2.3.2.2 Dictionaries

The choice of dictionary is of great importance when promoting sparsity [44, 45, 58]. Intuitively, the dictionary  $\mathbf{\Psi}$  should be chosen as the one lead-

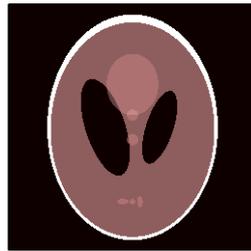
ing to the sparsest *synthesis* or *analysis* representation of the image  $\mathbf{x}$ . Additionally, the chosen dictionary must have a fast computation in both the forward and adjoint forms.

Consider a dictionary  $\Psi : \mathbb{R}^W \rightarrow \mathbb{R}^N$ . When  $W = N$ , the dictionary is considered to be complete, while when  $W > N$  the dictionary is said to be redundant or over-complete. A dictionary can be over-complete because of the way it is conceived (*e.g.*, the gradient operator or the redundant wavelet transform) or because it is obtained by merging two or several complete dictionaries (*e.g.*, a dictionary composed of the Fourier and wavelet dictionaries) [44,62]. Over-complete dictionaries have become popular in the literature of sparse regularization, specially in the *analysis*-based priors [11,27,61], due to their advantages over the traditional complete dictionaries. Since over-complete dictionaries contain a wider range of generating elements or *atoms*, they allow a better adaptation to the morphological content of the image, hence providing more flexibility to represent or project a given image [44,58].

In the literature, there exist two main approaches for choosing sparsity dictionaries. The first approach is based on learning an appropriate dictionary that adapts to a given data set [63,64]. The second approach consists of an analytic dictionary generated to represent the mathematical model of the data. In this work, we are interested in the second approach, where the dictionary does not require a previous learning process and it is highly structured, hence it has a fast numerical implementation (*e.g.*, Fourier, Wavelets).

There exist in the literature many analytic dictionaries that provide separate or joint representations in frequency, scale, translation and direction. Each dictionary allows distinguishing different types of singularities in an image. Some examples of such dictionaries include: Fourier, gradient, wavelets, curvelets, ridgelets and contourlets [44,65,66].

Our main focus in this work is on the following two dictionaries: the gradient operator suited for piecewise constant images (see Figure 2.1), and the wavelet dictionary suited for piecewise regular images (see Figure 2.2). The gradient operator is defined in Appendix 2.B and its description as a sparsity dictionary is presented in Section 2.3.2.3. The wavelet dictionary is described in Appendix 2.C.



(a)



(b)

**Figure 2.1** Illustration of the gradient operator applied to a piecewise constant image. (a) The Shepp-Logan phantom of size  $256 \times 256$ , an example of a piecewise constant image. (b) Absolute value of the finite differences of the image in (a) in both the horizontal (left) and vertical (right) directions. The image in (b) corresponds to  $\bar{\alpha}$  in (2.12), which is obtained by applying the gradient operator  $\Psi^* = \nabla$  (see Appendix 2.B) to image  $\mathbf{x}$  in (a), *i.e.*,  $\bar{\alpha} = \nabla \mathbf{x}$ . The intensity of the images belong to the interval  $[0, 1]$ , where the zero is represented as black and the one is represented as white.



**Figure 2.2** Illustration of the wavelet decomposition of a natural image using the Daubechies wavelet with 2 vanishing moments (db2). (a) The “cameraman” image, an example of a natural image. (b) Absolute value of the wavelet coefficients using DWT with 2 decomposition levels. (c) Absolute value of the wavelet coefficients using a redundant DWT (the UDWT) with 1 decomposition level. The images in (b) and (c) correspond to  $\bar{\alpha}$  in (2.12). They are obtained by applying, respectively, the non-redundant and redundant *analysis* wavelet operators  $\Psi^* \in \{\Psi_{\text{DWT}}^*, \Psi_{\text{UDWT}}^*\}$  (see Appendix 2.C) to image  $x$  in (a), *i.e.*,  $\bar{\alpha} = \Psi^*x$ . The intensity of the images belong to the interval  $[0, 1]$ , where the zero is represented as black and the one is represented as white.

### 2.3.2.3 Functions that Promote Sparse Solutions

The  $\ell_0$  counting function, also known as the  $\ell_0$  norm<sup>2</sup> of a vector  $\mathbf{x} \in \mathbb{R}^N$ , defined as

$$\|\mathbf{x}\|_0 = \lim_{p \rightarrow 0} \sum_{i=1}^N |x_i|^p = |\{i \in [N] : x_i \neq 0\}|,$$

counts the number of non-zero entries of a vector. Intuitively, this implies that the  $\ell_0$  norm is the best measure of the sparsity of a vector [45, Section 1.7]. In Figure 2.3, we can see that, as  $p$  tends to zero, the value of  $|x_i|^p$  becomes an indicator function on the set  $\mathbb{R} - \{0\}$ , *i.e.*,  $\mathbf{1}_{\{\mathbb{R}-\{0\}\}}$ , which is equal to 0 if  $x_i = 0$  and 1 otherwise. Hence, the summation of  $|x_i|^p$  for all the entries of  $\mathbf{x}$ , defining the  $\ell_0$  norm, provides a count of the non-zero values of the vector  $\mathbf{x}$ . Equivalently, we can see in Figure 2.4-(a) that the  $\ell_0$  ball favors solutions that are in the axis.

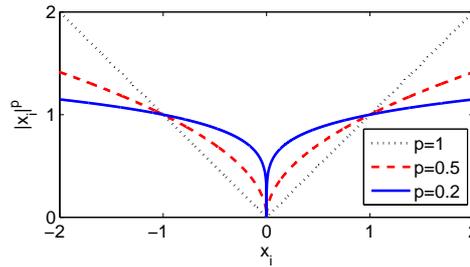


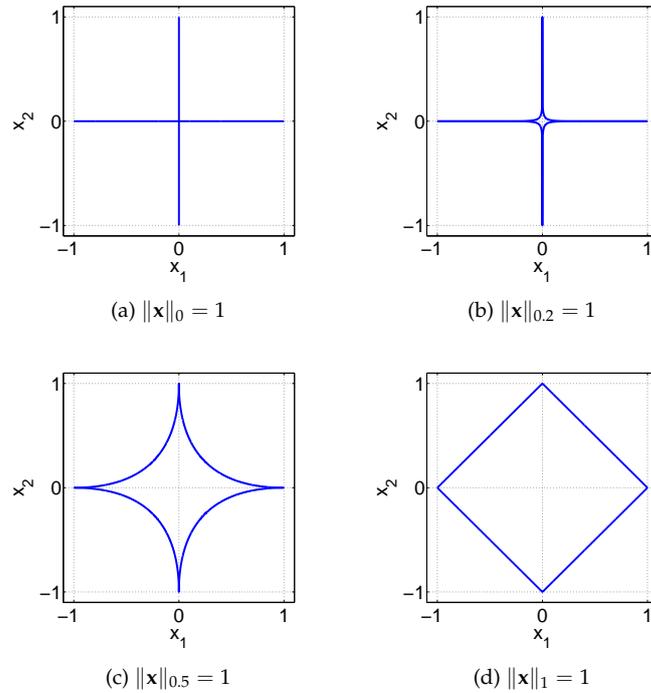
Figure 2.3 Behavior of  $|x_i|^p$  for several values of  $p$ .

If we assume that  $\mathbf{x}$  is sparse, by using the constrained formulation in (2.8) with an  $\ell_0$  norm prior, *i.e.*, with  $F(\mathbf{x}) = \|\mathbf{x}\|_0$ , we have

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \varepsilon. \quad (2.15)$$

The solution of this problem is given by the first intersection between the  $\ell_0$  ball (see Figure 2.4-(a)) and an  $\ell_2$  ball of radius  $\varepsilon$  about  $\mathbf{y}$ . Since the solution of the problem is located on the axis, it is sparse.

<sup>2</sup>With a slight abuse of notation, we use the word *norm* to refer to all the  $\ell_p$  norms even for  $p < 1$ . However, the  $\ell_p$  norms are not formally norms for  $p < 1$  because some of the properties of the norm are violated (*e.g.*, the triangle inequality  $\|\mathbf{u} + \mathbf{v}\|_p \leq \|\mathbf{u}\|_p + \|\mathbf{v}\|_p$  is violated for  $0 < p < 1$  and the absolute homogeneity  $\|\lambda\mathbf{u}\|_p = |\lambda|\|\mathbf{u}\|_p, \forall \lambda \neq 0$ , is violated for  $p = 0$ ).



**Figure 2.4** Illustration of the unit  $\ell_p$  balls in 2-D for (a)  $p = 0$ , (b)  $p = 0.2$ , (c)  $p = 0.5$  and (d)  $p = 1$ .

Due to its combinatorial nature, solving problem (2.15) is in general NP-hard (non-deterministic polynomial-time hard) [45,67]. Therefore, some alternative strategies have been investigated in order to find an approximate solution to (2.15). In [57], the authors have investigated a thresholding algorithm that allows working directly with the cost function in (2.15), however, due to the non-convexity of the  $\ell_0$  norm, the algorithm converges to local solutions that are highly dependent on the initial values. Some greedy techniques, such as matching pursuit type algorithms, have also been developed in order to approximately solve problem (2.15) [45,67,68]. These techniques proceed by a series of updates of the solution support, verifying at each step the satisfaction of the  $\ell_2$  fidelity constraint in order to find the non-zero values.

Another strategy that is commonly adopted for approximately solving (2.15) in a more tractable and efficient way, consists of a convex relaxation

of problem (2.15). This can be achieved, for instance, by replacing the  $\ell_0$  norm with the  $\ell_1$  norm:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{Ax}\|_2 \leq \varepsilon. \quad (2.16)$$

This problem is known in the literature as basis pursuit [45, 58, 59, 68, 69]. The solution to this problem, given by the first intersection of the  $\ell_1$  ball (see Figure 2.4-(d)) and the  $\ell_2$  ball of radius  $\varepsilon$  about  $\mathbf{y}$ , is located at the axis or close to it. Therefore, the  $\ell_1$  norm is used to promote sparse and weakly sparse solutions.

In the literature [39, 70, 71], there is a variant of the problem described above, which introduces the reweighted  $\ell_1$  norm, defined as  $\|\mathbf{x}\|_{1,w} = \|\mathbf{W}\mathbf{x}\|_1 = \sum_{i=1}^N w_i |x_i|$ , with  $\mathbf{W} = \text{diag}(w_1, \dots, w_N)$ . The solution of the new problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{1,w} \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{Ax}\|_2 \leq \varepsilon \quad (2.17)$$

has a higher concentration of values for the positions  $i$  with small weights  $w_i$  and, oppositely, its entries will tend to zero for the positions with large weights. The weights  $w_i$  are selected depending on the application. For instance, in [11, 70, 72] the authors use an iterative reweighting approach, where the weights are iteratively adapted based on the  $\ell_2$  norm of the solution. In this way, the resulting  $\ell_1/\ell_2$  norm is scale invariant and approximates better the  $\ell_0$  norm.

Almeida and Almeida [15] have proposed an alternative to the use of the  $\ell_1$  norm as sparsity norm. It consists of using  $\ell_p$  norms with  $0 < p < 1$ , which approximate better the  $\ell_0$  norm and promote sparser solutions (see Figures 2.3 and 2.4).

There exist other functions, different than the  $\ell_p$  norm, that allow promoting sparse solutions. We can also promote sparsity by using any function  $F(\mathbf{x}) = \sum_i \varphi(x_i)$ , with  $\varphi(x)$  a symmetric, monotonically non-decreasing function with a monotonic non-increasing derivative for  $x \geq 0$  [45, Section 1.6], e.g.,  $\varphi(x) = \log(1 + |x|)$ .

Note that the functions approximating better the  $\ell_0$  norm and hence promoting sparser solutions, are in general non-convex and the resulting regularized formulations tend to be harder to solve. As we will see in Section 2.4, the

$\ell_1$  norm represents a good compromise between promoting sparsity and having a closed-form solution for the proximal denoising. For this reason, it has been widely used in the literature for promoting sparse solutions [3,40,42,61]. Let us now introduce a special case of the  $\ell_1$  norm, called the Total-Variation norm, which is well-suited for edge preservation.

The Total-Variation (TV) norm is associated to the computation of the  $\ell_1$  norm in the gradient domain of the image and its minimization promotes the sparsity in this domain. The TV regularized problem was first introduced by Rudin, Osher and Fatemi [73] for denoising and, therefore, is also known as the ROF problem. It has been widely used in the literature [4,49,74,75] to regularize inverse problems where the desired solution presents a Bounded-Variation (BV) or “cartoon-shape” model (see Figure 2.1). This model includes images containing piece-wise constant areas, *e.g.*, images describing human-made objects or anatomic/metabolic functions of organs in the body.

There exist two main discrete formulations of the TV norm. The most commonly used formulation, the *isotropic* TV norm, takes the  $\ell_1$  norm of the magnitude of the gradient [76,77]

$$\|\mathbf{u}\|_{\text{TV}} := \|\nabla \mathbf{u}\|_{2,1}, \quad (2.18)$$

where  $\|\cdot\|_{2,1}$  is the mixed  $\ell_2$ - $\ell_1$  norm defined on any  $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2) \in \mathbb{R}^{N \times 2}$  as  $\|\mathbf{v}\|_{2,1} = \sum_{k=1}^N [(\mathbf{v}_1)_k^2 + (\mathbf{v}_2)_k^2]^{1/2}$ , and  $\nabla : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times 2}$  is the (finite differences) gradient operator defined in Appendix 2.B.

The second formulation, the *non-isotropic* TV norm, applies the  $\ell_1$  norm separately to each coordinate of the gradient [75, 77]:  $\|\mathbf{u}\|_{\text{TV}}^{\text{ni}} = \|(\nabla \mathbf{u})_1 + (\nabla \mathbf{u})_2\|_1$ , for  $\nabla \mathbf{u} = ((\nabla \mathbf{u})_1, (\nabla \mathbf{u})_2) \in \mathbb{R}^{N \times 2}$ .

The *isotropic* TV norm is more commonly used in imaging applications because it jointly penalizes the horizontal and the vertical directions of the gradient. This makes it more suitable to treat the arbitrary orientations of the edges that are typically present in natural images [72]. Note that both forms of TV regularizations are an example of an *analysis*-based regularization function with the gradient as the redundant sparsity dictionary, *i.e.*,  $\Psi^* = \nabla$ .

There exist also some variants of the TV norm. One variant consists of a regularized version of the TV norm, where a relaxation parameter is added to

avoid the non-smoothness of the TV function [4]. Another variant consists of a higher-order version of the TV norm (*e.g.*, the second-order Total Generalized Variation [78, 79] or the Hessian-Schatten norm [72]). By considering higher-order derivatives, this variant of the TV norm relaxes the prior for images that are not piece-wise constant everywhere, avoiding the staircase effects we might see in some TV reconstructions.

## 2.4 Optimization Tools

In this section, we present some optimization tools that are used throughout the thesis. We start by providing some useful definitions and theorems. Then, we present some proximal algorithms. These algorithms will be used in the thesis for solving convex optimization problems and a specific kind of non-convex problems.

### 2.4.1 Useful Definitions and Theorems

We start this section by providing some definitions that allow characterizing the sets and functions used in the thesis, specifically in Sections 2.4.2 and 2.4.3.

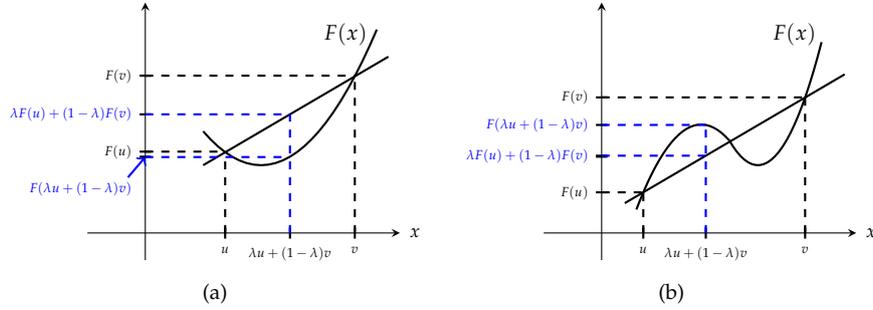
*Definition 2.1* (Convex set). A set  $\mathcal{C}$  is convex if the line segment between any two points in  $\mathcal{C}$  lies in  $\mathcal{C}$ . Formally, this means that, for any  $\mathbf{u}, \mathbf{v} \in \mathcal{C}$  and any  $\lambda \in [0, 1]$ , we have  $\lambda\mathbf{u} + (1 - \lambda)\mathbf{v} \in \mathcal{C}$ .

*Definition 2.2* (Convex function). A function  $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  is convex if  $F$  is below any line segment between two points of  $F$ . Formally, this means that  $F$  satisfies Jensen's inequality

$$F(\lambda\mathbf{u} + (1 - \lambda)\mathbf{v}) \leq \lambda F(\mathbf{u}) + (1 - \lambda)F(\mathbf{v}), \quad (2.19)$$

for any two points  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$  and for any  $\lambda \in [0, 1]$ . When the strict inequality holds in (2.19), the function  $F$  is said to be strictly convex.

Definition 2.2 is illustrated in Figure 2.5.



**Figure 2.5** Illustration of a convex and a non-convex function in 1-D. In (a), we have a convex function and we observe that there is no line segment that lies below  $F$  at any point. In (b), we have a non-convex function and we observe that the shown line segment lies above  $F$  at some points and below it at other points.

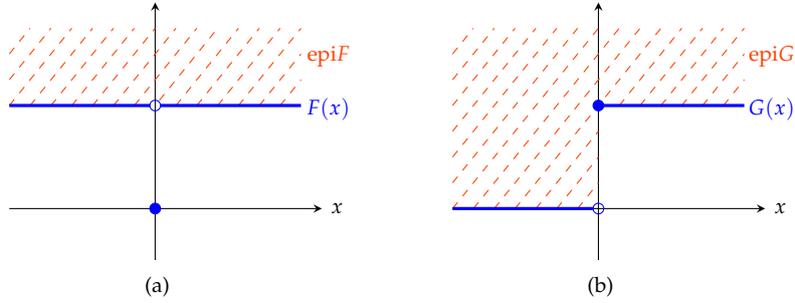
*Definition 2.3* (Proper function). A function  $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  is proper if its domain is non-empty, *i.e.*,  $\text{dom } F := \{\mathbf{x} \in \mathbb{R}^N : F(\mathbf{x}) < +\infty\} \neq \emptyset$ .

*Definition 2.4* (Lower semi-continuous function). A function  $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  is lower semi-continuous (l.s.c.) at a point  $\mathbf{x}_0 \in \mathbb{R}^N$  if  $F(\mathbf{x}_0) \leq \liminf_{\mathbf{x} \rightarrow \mathbf{x}_0} F(\mathbf{x})$  or, equivalently, if its epigraph  $\text{epi } F = \{(\mathbf{x}, \lambda) \in \mathbb{R}^{N+1} : F(\mathbf{x}) \leq \lambda\}$  is a closed set.

Let us illustrate Definition 2.4 with some examples. Consider functions  $F(x) = \{1 \text{ if } x \neq 0; 0 \text{ otherwise}\}$  and  $G(x) = \{1 \text{ if } x \geq 0; 0 \text{ otherwise}\}$  depicted in Figure 2.6. On one hand, the epigraph of  $F(\cdot)$  is a closed set and hence,  $F$  is a l.s.c. function. On the other hand, since the epigraph of  $G(\cdot)$  is not closed,  $G$  is not a l.s.c. function.

Using Definitions 2.2, 2.3 and 2.4, we can build the function class  $\Gamma_0(\mathcal{V})$ , which corresponds to the class of proper, convex and l.s.c. functions from a finite dimensional vector space  $\mathcal{V}$  (*e.g.*,  $\mathbb{R}^N$ ) to  $(-\infty, +\infty]$  [80]. This class of functions is very important in convex analysis since most results obtained in this field are for functions that belong to this class. Some interesting examples of functions belonging to this class include the  $\ell_p$  norm for  $p \geq 1$  and the convex indicator function<sup>3</sup> of a non-empty closed convex set.

<sup>3</sup>The convex indicator or characteristic function of the set  $\mathcal{C}$ , denoted as  $\iota_{\mathcal{C}}(\mathbf{x})$ , is equal to 0 if  $\mathbf{x} \in \mathcal{C}$  and  $+\infty$  otherwise.



**Figure 2.6** Illustration of Definition 2.4. (a) Function  $F(x) = \{1 \text{ if } x \neq 0; 0 \text{ otherwise}\}$  and its epigraph, an example of a l.s.c. function. (b) Function  $G(x) = \{1 \text{ if } x \geq 0; 0 \text{ otherwise}\}$  and its epigraph, an example of a function that is not l.s.c. The function is indicated with a solid line and the epigraph corresponds to the region with dashed lines. The discontinuity point is denoted by a solid circle if it belongs to the function or by an open circle if it does not belong.

*Definition 2.5* (Lipschitz continuous function). A function  $F: \mathbb{R}^N \rightarrow (-\infty, +\infty]$  is Lipschitz continuous with constant  $\lambda$ , if

$$|F(\mathbf{u}) - F(\mathbf{x})| \leq \lambda \|\mathbf{u} - \mathbf{x}\|_2,$$

for all  $\mathbf{x}, \mathbf{u} \in \mathbb{R}^N$  and  $\lambda \in \mathbb{R}$ .

*Definition 2.6* (Semi-algebraic set). A subset  $\mathcal{S}$  of  $\mathbb{R}^N$  is a real semi-algebraic set if there exists a finite number of real polynomial functions  $P_{ij}, Q_{ij}: \mathbb{R}^N \rightarrow \mathbb{R}$  such that

$$\mathcal{S} = \bigcup_{i=1}^m \bigcap_{j=1}^n \{\mathbf{x} \in \mathbb{R}^N : P_{ij}(\mathbf{x}) = 0, Q_{ij}(\mathbf{x}) < 0\}.$$

The class of semi-algebraic sets is stable under the operations of finite union, finite intersection, Cartesian product or complementation [81, 82].

*Definition 2.7* (Semi-algebraic function). A function  $F \in \Gamma_0(\mathbb{R}^N)$  is semi-algebraic if its graph, defined as  $\text{graph } F = \{(\mathbf{x}, \lambda) \in \mathbb{R}^{N+1} : F(\mathbf{x}) = \lambda\}$ , is a semi-algebraic subset of  $\mathbb{R}^{N+1}$ . The class of semi-algebraic functions is stable under the operations of finite sum, product and composition [81, 82].

Definitions 2.6 and 2.7 are of interest in this work since the algorithm presented in Section 2.4.3 applies to the semi-algebraic functions and sets. Some examples of semi-algebraic functions include the real polynomial functions,

the  $\ell_p$  norm with  $p$  rational, and the convex indicator function of a semi-algebraic set [82].

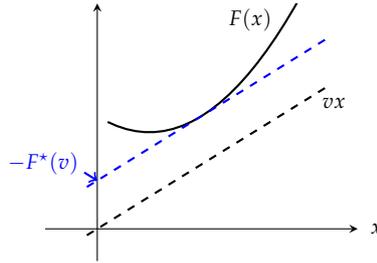
Let us now introduce the conjugacy of convex functions, a fundamental notion that will allow a better understanding of duality and, hence, of the primal-dual algorithm presented in Section 2.4.2.2.

*Definition 2.8* (Convex conjugate function). The convex conjugate of a function  $F : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  is denoted by  $F^* : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  and is provided by the Legendre transform as [83]:

$$F^*(\mathbf{v}) = \sup_{\mathbf{x} \in \mathbb{R}^N} \langle \mathbf{x}, \mathbf{v} \rangle - F(\mathbf{x}), \quad (2.20)$$

for any  $\mathbf{v} \in \mathbb{R}^N$ . Remark that when  $F \in \Gamma_0(\mathbb{R}^N)$ , then  $F^* \in \Gamma_0(\mathbb{R}^N)$  [84].

Definition 2.8 is illustrated in Figure 2.7.



**Figure 2.7** Illustration in 1-D of a convex function and its convex conjugate. The conjugate function  $F^*(v)$  is given by the maximum distance between the straight line of slope  $v$ , i.e.,  $vx$ , and the function  $F(x)$ . The primal description corresponds to the function  $F(x)$  while the dual description corresponds to the crossing points  $F^*(v)$ .

After understanding the conjugacy of convex functions from Definition 2.8, we can go deeper in the notion of duality and present the connection between primal and dual convex optimization problems. Consider the following minimization problem

$$\tilde{p} = \min_{\mathbf{x} \in \mathbb{R}^N} F(\mathbf{x}) + G(\mathbf{x}), \quad (2.21)$$

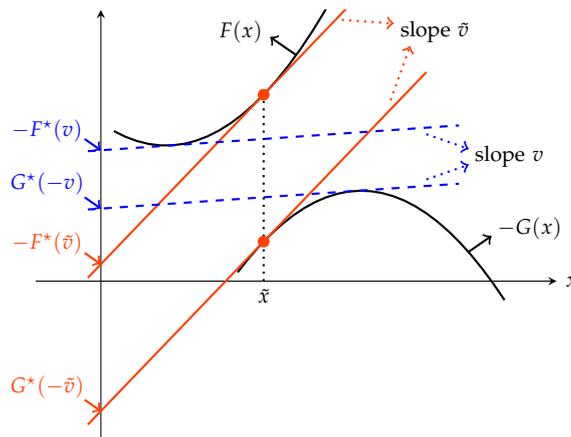
with  $F, G : \mathbb{R}^N \rightarrow (-\infty, +\infty]$  closed convex functions. If (2.21) is denoted as the primal problem, then its dual version corresponds to

$$\tilde{d} = \max_{\mathbf{v} \in \mathbb{R}^N} -F^*(\mathbf{v}) - G^*(-\mathbf{v}), \quad (2.22)$$

for  $F^*, G^*$  the convex conjugate functions of their unstarred versions computed using (2.20). The relation between these two problems is described by Fenchel's duality theorem [83–85].

**Theorem 2.1** (Fenchel's duality [83]). Let  $\tilde{p}$  and  $\tilde{d}$  be the solutions of the primal problem (2.21) and its dual version (2.22), respectively. If the functions  $F$  and  $G$  belong to  $\Gamma_0(\mathbb{R}^N)$  and the intersection of the interior domains of  $F$  and  $G$  is non-empty, *i.e.*,  $\{\text{int}(\text{dom } F) \cap \text{int}(\text{dom } G)\} \neq \emptyset$ , then  $\tilde{p} = \tilde{q}$ .

Theorem 2.1 is illustrated in Figure 2.8. Fenchel's duality is particularly important in the cases when the dual problem is easier to solve than the primal one [84]. Also, it allows formulating primal-dual algorithms that have shown to present certain advantages over the primal ones [86].



**Figure 2.8** Fenchel's duality in 1-D. The figure depicts, in solid red, the optimal pair  $(\tilde{x}, \tilde{v})$ , *i.e.*, the solution of the primal and dual problems. Remark that, for this optimal pair, we reach the minimal vertical distance between  $F(x)$  and  $-G(x)$  (solving the primal problem (2.21)) and we also reach the maximal difference between the crossing points  $-F^*(v)$  and  $G^*(-v)$  (solving the dual problem (2.22)). For comparison purposes, the result for an arbitrary value of the slope  $v$  is depicted in dashed blue.

In the rest of this section, we explain some important optimization notions that are crucial to understand the rest of the chapter. We first present the sub-gradient or subdifferential, which will allow the definition of a critical point for functions that are not differentiable. Then, we introduce the concept of

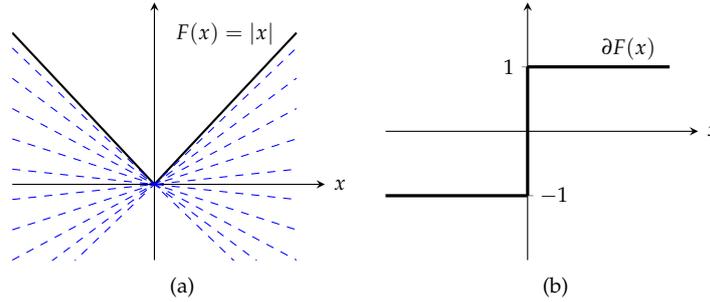
proximal operators, which are the basis of the proximal methods presented in Sections 2.4.2 and 2.4.3.

*Definition 2.9* (Subgradient or subdifferential). The subgradient or subdifferential of a function  $F \in \Gamma_0(\mathbb{R}^N)$  is the following set-valued operator [80]:

$$\partial F(\zeta) = \{\mathbf{s} \in \mathbb{R}^N : (\forall \mathbf{x} \in \mathbb{R}^N) F(\mathbf{x}) \geq F(\zeta) + (\mathbf{x} - \zeta)^T \mathbf{s}\}.$$

If  $F$  is differentiable at  $\zeta \in \mathbb{R}^N$ , its subgradient at  $\zeta$  is given by the gradient, *i.e.*,  $\partial F(\zeta) = \{\nabla F(\zeta)\}$ .

The subgradient is a generalization of the concept of gradient to non-differentiable functions. For instance, the  $\ell_1$  norm is a non-differentiable function at the origin, but we can compute its subgradient at the origin, which is given by the interval  $[-1, 1]$  (see Figure 2.9).



**Figure 2.9** Subgradient of function  $F(x) = |x|$ . (a) Illustration of the subgradient of  $F(x)$  at the origin ( $\zeta = 0$ ), which is given by  $\{s \in \mathbb{R} : \forall x \in \mathbb{R} |x| \geq sx\}$ , *i.e.*, the slopes of all the lines that are below the function  $|x|$  (shown as dashed lines in the figure). (b) Shape of  $\partial F$ . The subgradient of the  $\ell_1$  norm can be obtained by extending this concept to higher dimensions.

*Definition 2.10* (Critical point). A vector  $\tilde{\mathbf{x}} \in \mathbb{R}^N$  is a critical point of a function  $F \in \Gamma_0(\mathbb{R}^N)$  if  $0 \in \partial F(\tilde{\mathbf{x}})$ . Let us note that this is a necessary condition for  $\tilde{\mathbf{x}}$  to be a minimizer of  $F$ .

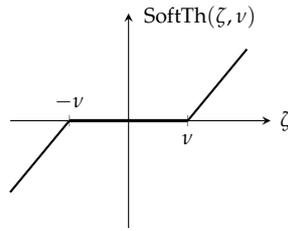
*Definition 2.11* (Proximal operator). The proximal or resolvent operator of a function  $F \in \Gamma_0(\mathbb{R}^N)$ , is the function  $\text{prox}_{\nu F} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  defined as [80, 87]:

$$\text{prox}_{\nu F} \zeta = (\mathbf{I}_N + \nu \partial F)^{-1}(\zeta) := \underset{\mathbf{u} \in \mathbb{R}^N}{\text{argmin}} F(\mathbf{u}) + \frac{1}{2\nu} \|\mathbf{u} - \zeta\|_2^2. \quad (2.23)$$

The proximal operator in (2.23) is uniquely defined and, for most interesting functions in imaging applications, it has a *simple* representation or can be computed fast and efficiently [51]. For instance, the proximal operator of the convex indicator function onto the convex set  $\mathcal{C}$  (commonly used to handle constraints), is given by the projection onto the set  $\mathcal{C}$ . Another useful example is the proximal operator of the  $\ell_1$  norm, which is given by the soft-thresholding operator [40, 42, 80]:

$$\text{prox}_{v\|\cdot\|_1} \zeta = \text{SoftTh}(\zeta, v) := \text{sign}(\zeta) (|\zeta| - v\mathbf{1}_N)_+. \quad (2.24)$$

This operator is illustrated in Figure 2.10. For a complete read on proximal operators we refer the reader to [51, 80].



**Figure 2.10** Illustration of the soft-thresholding operator, the proximal operator of the  $\ell_1$  norm.

## 2.4.2 Convex Optimization through Proximal Algorithms

Many inverse problems in imaging that aim at retrieving  $\mathbf{x}$  from the corrupted observations in (2.5), using the knowledge of the noise distribution and the regularization functions, can be formulated as a convex optimization problem of the form

$$\min_{\mathbf{x} \in \mathbb{R}^N} \sum_{j=1}^L F_j(\mathbf{K}_j \mathbf{x}) + G(\mathbf{x}), \quad (2.25)$$

where we have the sum of  $L$  functions belonging to  $\Gamma_0(\mathbb{R}^D)$  and a collection of  $L$  linear operators  $\mathbf{K}_j : \mathbb{R}^N \rightarrow \mathbb{R}^{W_j}$ , which can include the sensing matrix  $\mathbf{A}$ , the sparsity dictionary  $\Psi$  or simply the identity matrix  $\mathbf{I}_N$ .

When problem (2.25) contains a sum of differentiable functions, we can use Newton's method or the gradient descent method and its accelerated variants (see, e.g., [15]). However, in most inverse problems in imaging, we have

one or several non-smooth functions (*e.g.*, constraints and the  $\ell_1$  norm), which prevents us from using those methods. Non-smooth convex functions can be efficiently handled with the proximal operator in (2.23), and the resulting algorithms making use of the proximal operator are referred to as *proximal algorithms*.

In the following subsections, we review two main proximal algorithms widely used in the literature and that will be used throughout the thesis: the forward-backward algorithm [51,80] and the primal-dual algorithm proposed by Chambolle and Pock [74]. We also compare these algorithms with similar ones in the literature.

#### 2.4.2.1 Forward-Backward Algorithm

Let us first introduce the forward-backward (FB) splitting algorithm [80], also known as proximal gradient algorithm [51]. This algorithm allows the minimization of a smooth and a non-smooth convex functions and, hence, treats the following problem:

$$\min_{\mathbf{x} \in \mathbb{R}^N} F_1(\mathbf{x}) + G(\mathbf{x}), \quad (2.26)$$

which is a specific case of (2.25) for  $L = 1$ ,  $\mathbf{K}_1 = \mathbf{I}_N$ ,  $F_1 \in \Gamma_0(\mathbb{R}^N)$  and  $G : \mathbb{R}^N \rightarrow \mathbb{R}$  convex and differentiable with gradient  $\nabla G$ .

The FB algorithm proceeds by splitting the objective into two terms, using a forward Euler step for the differentiable function  $G$  (a gradient step) and a backward Euler step for the non-differentiable function  $F_1$  (a proximal step) [51,80]:

$$\mathbf{x}^{(k+1)} = \text{prox}_{\nu^{(k)} F_1} \left( \mathbf{x}^{(k)} - \nu^{(k)} \nabla G(\mathbf{x}^{(k)}) \right), \quad (2.27)$$

where  $\nu^{(k)} > 0$  is the step-size. When  $\nabla G$  is Lipschitz continuous with constant  $\lambda$  (see Definition 2.5), the step-size can be fixed to  $\nu^{(k)} = \nu \in (0, 2/\lambda)$  and the algorithm converges with a rate  $\mathcal{O}(1/k)$  [51,80,88]. If  $\lambda$  is not known, the step-sizes can be adaptively updated using, for instance, line search algorithms such as the one of Beck and Teboulle [51,89]. Algorithm (2.27) can also be used in general situations where function  $G$  is a composite function, *i.e.*,  $G(\mathbf{D} \cdot)$  with  $\mathbf{D} : \mathbb{R}^N \rightarrow \mathbb{R}^D$  a linear operator. In such case, the step-size  $\nu$  must be set accordingly [80] and the convergence rate might be affected by the behavior of operator  $\mathbf{D}$ .

Algorithm (2.27) is a general form of several known methods in the literature. For instance, it can be reduced to the gradient method for  $F_1(\cdot) = 0$  and to the proximal point algorithm for  $G(\cdot) = 0$ . Also, if  $F_1(\cdot) = \|\cdot\|_1$ , the algorithm in (2.27) is called Iterative Soft-Thresholding (IST) [39, 80, 90, 91], while it is referred as Iterative Hard-Thresholding (IHT) [57] if  $F_1(\cdot) = \|\cdot\|_0$ . Moreover, the projected Landweber [80] or (scaled) gradient projection method [4, 25, 34] can be found if  $F_1(\cdot)$  is the convex indicator function.

In the literature, there exist several accelerated versions of algorithm (2.27). For instance, Bioucas-Dias and Figueiredo [77] proposed the two-step IST (TwIST) algorithm; while Nesterov [92] and Beck and Teboulle [93] proposed accelerated versions based on extrapolation parameters. In the thesis, specifically in Chapter 5, we use the latter with the extrapolation parameter  $\beta$  suggested by Parikh and Boyd [51]. This algorithm is called accelerated proximal gradient (APG) algorithm and it iterates as follows:

$$\begin{cases} \mathbf{s}^{(k+1)} &= \mathbf{x}^{(k)} + \beta^{(k)} (\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) \\ \mathbf{x}^{(k+1)} &= \text{prox}_{\nu^{(k)} F_1} (\mathbf{s}^{(k+1)} - \nu^{(k)} \nabla G(\mathbf{s}^{(k+1)})) \\ \beta^{(k)} &= \frac{k}{k+3} \end{cases} \quad (2.28)$$

The step-size  $\nu^{(k)}$  can be defined in the same way as for the non-accelerated version in (2.27), either using the Lipschitz constant of  $\nabla G$ , in which case the algorithm converges with a rate  $\mathcal{O}(1/k^2)$  [51], or using an adaptive updating rule such as a line search algorithm [51, 89].

Note that iterations (2.27) and (2.28) are defined for the minimization problem (2.26), *i.e.*, for  $L = 1$  and  $\mathbf{K}_1 = \mathbf{I}_N$  in (2.25). For  $L > 1$ , the optimization space can be extended, resulting in the generalized forward-backward algorithm [94]. Due to the presence of the proximal operator, algorithms (2.27) and (2.28) require  $\mathbf{K}_j \in \mathbb{R}^{W_j \times N}$  to be a tight frame, *i.e.*,  $\mathbf{K}_j \mathbf{K}_j^* = \gamma \mathbf{I}_{W_j}$  for  $\gamma > 0$ . In such case, the proximal operator of the composite function  $F_j(\mathbf{K}_j \cdot)$  can be computed by means of the semi-orthogonal linear transform property of the proximal operators [80]. If  $\mathbf{K}_j$  is not a tight frame, algorithms (2.27) and (2.28) depend on a fast inversion of  $\mathbf{K}_j^* \mathbf{K}_j + \mathbf{I}_{W_j}$  (see [94] for more details).

The FB algorithm presented in this section is an efficient proximal algorithm that allows solving a wide variety of inverse problems in imaging. In

the thesis, we use this algorithm in the context of filter estimation in Chapter 5, where we deal with a constrained optimization of several differentiable functions.

Nevertheless, the FB algorithm is not well suited to solve problems where operators  $\mathbf{K}_j$  are not a tight frame or where function  $G(\cdot)$  is not differentiable. These problems are frequently encountered in imaging applications, where the sensing operator characterizing the actual imaging process is not a tight frame (see, *e.g.*, Chapter 3), or where the inverse problem is formulated as a constrained optimization of non-differentiable functions (see, *e.g.*, Chapters 3, 4 and 5).

In the next section, we present the primal-dual algorithm of Chambolle and Pock [74] and its extended version for the sum of several convex functions. This algorithm is able to handle the problems that FB could not, *i.e.*, dealing with untight operators and with the presence of several non-differentiable convex functions.

#### 2.4.2.2 Chambolle-Pock Primal-Dual Algorithm

We now focus on the Chambolle-Pock (CP) algorithm [74], an efficient, robust and flexible algorithm that allows handling the sum of non-differentiable convex functions. Due to its primal-dual formulation, the CP algorithm is able to deal with a wide variety of operators  $\mathbf{K}_j$ , including those that are not a tight frame. This is of great importance in the thesis, since many of the sensing operators we face in actual imaging applications are not a tight frame (see, *e.g.*, Chapter 3).

The original formulation of the CP algorithm [74, Algorithm 1] aims at solving minimization problems as the one in (2.25) for  $L = 1$ , *i.e.*, a minimization of the form:

$$\min_{\mathbf{x} \in \mathbb{R}^N} F_1(\mathbf{K}_1 \mathbf{x}) + G(\mathbf{x}). \quad (2.29)$$

Based on Fenchel's duality (see Theorem 2.1), the CP algorithm solves the primal problem described above simultaneously with its dual problem, until the difference between their objective functions – the primal-dual gap – is zero [74, 95].

For any variable  $\mathbf{s} \in \mathbb{R}^{W_1}$ , the primal-dual optimization can be formulated as the following saddle-point problem:

$$\min_{\mathbf{x} \in \mathbb{R}^N} \max_{\mathbf{s} \in \mathbb{R}^W} \langle \mathbf{K}_1 \mathbf{x}, \mathbf{s} \rangle + G(\mathbf{x}) - F_1^*(\mathbf{s}), \quad (2.30)$$

where  $F_1^*$  is the convex conjugate of function  $F_1$  provided by the Legendre transform in (2.20).

Using the Legendre transform, we obtain the primal version described in (2.29) and also the dual version as follows:

$$\max_{\mathbf{s} \in \mathbb{R}^W} -F_1^*(\mathbf{s}) - G^*(-\mathbf{K}_1^* \mathbf{s}), \quad (2.31)$$

where  $\mathbf{K}_1^*$  is the exact adjoint of the operator  $\mathbf{K}_1$ , such that  $\langle \mathbf{K}_1 \mathbf{x}, \mathbf{s} \rangle = \langle \mathbf{x}, \mathbf{K}_1^* \mathbf{s} \rangle$ .

The CP algorithm is defined by the following iterations:

$$\begin{cases} \mathbf{s}^{(k+1)} &= \text{prox}_{\nu F_1^*} \left( \mathbf{s}^{(k)} + \nu \mathbf{K}_1 \bar{\mathbf{x}}^{(k)} \right) \\ \mathbf{x}^{(k+1)} &= \text{prox}_{\mu G} \left( \mathbf{x}^{(k)} - \mu \mathbf{K}_1^* \mathbf{s}^{(k+1)} \right), \\ \bar{\mathbf{x}}^{(k+1)} &= \mathbf{x}^{(k+1)} + \vartheta \left( \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right) \end{cases}, \quad (2.32)$$

with  $\vartheta \in [0, 1]$  a parameter that controls the over-relaxation in  $\mathbf{x}$ . Chambolle and Pock [74] proved that the algorithm in (2.32) converges for fixed values of  $\mu$  and  $\nu$  such that  $\mu\nu \|\mathbf{K}_1\|^2 < 1$ , with  $\|\mathbf{K}_1\|$  the induced norm of the operator  $\mathbf{K}_1$ . This provides a primal-dual gap that vanishes at a rate  $\mathcal{O}(1/k)$  [74]. In [96], the authors have observed that, when the operator  $\mathbf{K}_1$  is ill-conditioned and its norm is very large, the algorithm in (2.32) may have a slow convergence. In order to tackle this problem, Pock and Chambolle [96] replaced the step-size parameters  $\mu$  and  $\nu$  by special diagonal matrices that precondition operator  $\mathbf{K}_1$ .

In practice, the convergence of the primal-dual algorithm in (2.32) to the vicinity of a minimizer of (2.29) can be analyzed through the primal-dual gap. While iterating, the primal objective function (2.29) is greater than the dual objective function (2.31) and, when approaching the optimal solution, the difference between these objective functions tends to zero (see Theorem 2.1). Remark that if the primal and dual objective functions contain convex indicator functions [23], they can only be partially evaluated.

Another way of analyzing the algorithm convergence in practice is through the primal and dual residuals. These are defined by the subgradient (see Definition 2.9) of the saddle-point problem in (2.30) with respect to the primal and dual variables, namely,  $P(\mathbf{x}, \mathbf{s}) := \partial G(\mathbf{x}) + \mathbf{K}_1^* \mathbf{s}$  and  $D(\mathbf{x}, \mathbf{s}) := \partial F_1^*(\mathbf{s}) - \mathbf{K}_1 \mathbf{x}$ , respectively [97].

By definition, for an optimal point  $(\tilde{\mathbf{x}}, \tilde{\mathbf{s}})$  of (2.30), zero must belong to these residuals and, therefore, by tracking the norm of these residuals we can perform an analysis of the algorithm convergence. Explicit formulas can be obtained for the primal and dual residuals by observing the optimality conditions of (2.32) at each iteration. This provides

$$\begin{aligned} 0 &\in \mu \partial G(\mathbf{x}^{(k+1)}) + \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} + \mu \mathbf{K}_1^* \mathbf{s}^{(k+1)}, \\ 0 &\in \nu \partial F_1^*(\mathbf{s}^{(k+1)}) + \mathbf{s}^{(k+1)} - \mathbf{s}^{(k)} - \nu \mathbf{K}_1 \left( \mathbf{x}^{(k)} (1 + \vartheta) - \vartheta \mathbf{x}^{(k-1)} \right). \end{aligned}$$

or, equivalently,  $\mathbf{p}^{(k+1)} \in P(\mathbf{x}^{(k+1)}, \mathbf{s}^{(k+1)})$  and  $\mathbf{d}^{(k+1)} \in D(\mathbf{x}^{(k+1)}, \mathbf{s}^{(k+1)})$  with the primal and dual residual vectors

$$\begin{aligned} \mathbf{p}^{(k+1)} &= \frac{1}{\mu} \left( \mathbf{x}^{(k)} - \mathbf{x}^{(k+1)} \right), \\ \mathbf{d}^{(k+1)} &= \frac{1}{\nu} \left( \mathbf{s}^{(k)} - \mathbf{s}^{(k+1)} \right) + \mathbf{K}_1 \left( (1 + \vartheta) \mathbf{x}^{(k)} - \vartheta \mathbf{x}^{(k-1)} - \mathbf{x}^{(k+1)} \right). \end{aligned} \quad (2.33)$$

Goldstein *et al.* [97] show experimentally that a converging algorithm respects

$$\lim_{k \rightarrow \infty} \|\mathbf{p}^{(k)}\|_* + \|\mathbf{d}^{(k)}\|_* = 0, \quad (2.34)$$

for some norm  $\|\cdot\|_*$ , e.g.,  $\ell_1$ .

The CP algorithm has been widely used in the literature for solving inverse problems in imaging due to its robustness and its flexibility to handle a wide variety of operators  $\mathbf{K}_1$  [4, 23, 86]. This algorithm is equivalent to the one referred as primal-dual hybrid gradient method [95, 97, 98]. It also presents some similarities with well-known methods in the literature. For instance, the Douglas-Rachford splitting (DRS) algorithm [80, 99] is equivalent to the CP iterations in (2.32) for  $\mathbf{K}_1 = \mathbf{I}_N$  [74]. Also, the CP algorithm can be seen as a preconditioned version of the alternating direction method of multipliers (ADMM) [51, 80, 99], an algorithm that is often used for solving inverse problems in imaging [54, 61, 100, 101]. ADMM requires, however, operator

$\mathbf{K}_1 \in \mathbb{R}^{W_1 \times N}$  to be a tight frame or to be defined such that  $\mathbf{K}_1^* \mathbf{K}_1 + \mathbf{I}_{W_1}$  is easily invertible (see discussion in Section 2.4.2.1).

In order to handle problem (2.25) for  $L > 1$ , ADMM has been extended for an arbitrary sum of convex functions as illustrated in [61]. In this case, its usage is limited to those problems where  $\sum_{j=1}^L \mathbf{K}_j^* \mathbf{K}_j$  is easily invertible [61]. The CP algorithm does not have such constraint and, as explained hereafter, can also be extended for an arbitrary sum of convex functions. This extension is crucial in the thesis, since the inverse problems investigated in Chapters 3, 4 and 5 are formulated as optimization problems containing several convex functions and constraints.

The CP algorithm in (2.32) is easily adaptable to solve (2.25) by considering an  $L$ -times expanded optimization space  $\mathbb{R}^{LN}$ . This space is composed of  $\mathbf{x}' = (\mathbf{x}_1^T, \dots, \mathbf{x}_L^T)^T \in \mathbb{R}^{LN}$ , with  $\mathbf{x}_j \in \mathbb{R}^N$ . These auxiliary variables are all equal to each other, as described by the  $L - 1$  bisector planes  $\Pi_{1,j} = \{\mathbf{x}' \in \mathbb{R}^{LN} : \mathbf{x}_1 = \mathbf{x}_j, 2 \leq j \leq L\}$ . This allows us to work with the following equivalent primal problem:

$$\min_{\mathbf{x}' \in \mathbb{R}^{LN}} \sum_{j=1}^L F_j(\mathbf{K}_j \mathbf{x}_j) + \sum_{j=2}^L t_{\Pi_{1,j}}(\mathbf{x}') + H(\mathbf{x}_1). \quad (2.35)$$

The CP-shape in (2.29) is thus recovered by working in a bigger space  $\mathbb{R}^{LN}$  and by setting  $F(\mathbf{s}) = \sum_{j=1}^L F_j(\mathbf{s}_j)$ , with  $\mathbf{s} = (\mathbf{s}_1^T, \dots, \mathbf{s}_L^T)^T \in \mathbb{R}^{W=\sum_j W_j}$  and  $\mathbf{s}_j \in \mathbb{R}^{W_j}$ ,  $\mathbf{K} = \text{diag}(\mathbf{K}_1, \dots, \mathbf{K}_L) \in \mathbb{R}^{W \times LN}$  and  $G(\mathbf{x}') = \sum_{j=2}^L t_{\Pi_{1,j}}(\mathbf{x}') + H(\mathbf{x}_1)$ .

In this expanded optimization space, the equivalent dual problem is written as follows (see Appendix 2.D.1):

$$\max_{\mathbf{s} \in \mathbb{R}^W} - \sum_{j=1}^L F_j^*(\mathbf{s}_j) - H^* \left( - \sum_{j=1}^L \mathbf{K}_j^* \mathbf{s}_j \right). \quad (2.36)$$

It is easy to see that for any  $\nu > 0$  and  $\zeta = (\zeta_1^T, \dots, \zeta_L^T)^T \in \mathbb{R}^W$  we have:

$$\text{prox}_{\nu F^*} \zeta = \begin{pmatrix} \text{prox}_{\nu F_1^*} \zeta_1 \\ \vdots \\ \text{prox}_{\nu F_L^*} \zeta_L \end{pmatrix},$$

and, for any  $\mu > 0$  we have (more details in Appendix 2.D.1):

$$\text{prox}_{\mu G} \zeta = (\mathbf{I}_N, \dots, \mathbf{I}_N)^T \text{prox}_{\frac{\mu}{L} H} \left( \frac{1}{L} \sum_j \zeta_j \right).$$

Finally, making use of the above computations and taking  $\vartheta = 1$ , the extended CP algorithm is defined by the following iterations (see Appendix 2.D.2):

$$\begin{cases} \mathbf{s}_j^{(k+1)} &= \text{prox}_{\nu F_j^*} \left( \mathbf{s}_j^{(k)} + \nu \mathbf{K}_j \bar{\mathbf{x}}^{(k)} \right), j \in \{1, \dots, L\} \\ \mathbf{x}^{(k+1)} &= \text{prox}_{\frac{\mu}{L} H} \left( \mathbf{x}^{(k)} - \frac{\mu}{L} \sum_{j=1}^L \mathbf{K}_j^* \mathbf{s}_j^{(k+1)} \right) \\ \bar{\mathbf{x}}^{(k+1)} &= 2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \end{cases}. \quad (2.37)$$

The primal and dual residuals in (2.33) can be also computed for the iterations in (2.37) as follows:

$$\begin{aligned} \mathbf{p}^{(k+1)} &= \frac{L}{\mu} \left( \mathbf{x}^{(k)} - \mathbf{x}^{(k+1)} \right), \\ \mathbf{d}^{(k+1)} &= \left( \left( \mathbf{d}_1^{(k+1)} \right)^T, \dots, \left( \mathbf{d}_L^{(k+1)} \right)^T \right)^T, \\ \text{with } \mathbf{d}_j^{(k+1)} &= \frac{1}{\nu} \left( \mathbf{s}_j^{(k)} - \mathbf{s}_j^{(k+1)} \right) + \mathbf{K}_j \left( 2\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} - \mathbf{x}^{(k+1)} \right). \end{aligned} \quad (2.38)$$

### 2.4.3 Non-Convex Optimization through Alternating Algorithms

There exists a variety of inverse problems in imaging that cannot be formulated as a convex optimization problem like the one in (2.25) (see, for instance, the works of Almeida and Almeida [15], of Puy and Vandergheynst [24] and of Ravishankar and Bresler [102]). Some of them yield a non-convex optimization problem, where the objective function and/or the constraints are non-convex with respect to the optimization variables.

A non-convex problem is hard to minimize because the functions may have several local minima and the convergence to a global minima cannot be guaranteed. In the literature, there exist many types of non-convex optimization problems and there is no general rule for solving all kinds of non-convex problems. Hence, each problem needs to be treated according to its structure

and characteristics. In this section, we are interested in the following structured problem:

$$\min_{\mathbf{x} \in \mathbb{R}^N, \mathbf{u} \in \mathbb{R}^W} \{L(\mathbf{x}, \mathbf{u}) := F(\mathbf{x}) + Q(\mathbf{x}, \mathbf{u}) + G(\mathbf{u})\}, \quad (2.39)$$

with two regularization functions  $F: \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  and  $G: \mathbb{R}^W \rightarrow \mathbb{R} \cup \{+\infty\}$  that promote the prior information on  $\mathbf{x}$  and  $\mathbf{u}$ , respectively, and a function  $Q: \mathbb{R}^N \times \mathbb{R}^W \rightarrow \mathbb{R}$  that couples the vectors  $\mathbf{x}$  and  $\mathbf{u}$ . Our interest in this type of optimization problem arises from the inverse problem investigated in Chapter 5.

Non-convex problems of the form (2.39) are usually solved in an alternated fashion, by estimating the variables  $\mathbf{x}$  and  $\mathbf{u}$  alternately in a two-block Gauss-Seidel fashion [15, 103]. The general behavior of this type of algorithm is as follows:

$$\begin{cases} \mathbf{x}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{x}} \in \mathbb{R}^N} L(\bar{\mathbf{x}}, \mathbf{u}^{(k)}) \\ \mathbf{u}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{u}} \in \mathbb{R}^W} L(\mathbf{x}^{(k+1)}, \bar{\mathbf{u}}) \end{cases}. \quad (2.40)$$

It has been shown in [15] that this type of algorithm is able to provide fairly good results that approach to the solution of (2.39). However, the convergence of the simple alternating algorithm in (2.40) is either not guaranteed or it requires very strict conditions such as attaining a unique minimizer at each step [82]. Recently, Attouch *et al.* [104] proposed a proximal alternating minimization algorithm where cost-to-move functions are added to the common alternating algorithm. These functions are quadratic costs that penalize variations between two consecutive iterations. The algorithm iterates as follows:

$$\begin{cases} \mathbf{x}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{x}} \in \mathbb{R}^N} L(\bar{\mathbf{x}}, \mathbf{u}^{(k)}) + \frac{\lambda_x^{(k)}}{2} \|\bar{\mathbf{x}} - \mathbf{x}^{(k)}\|_2^2 \\ \mathbf{u}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{u}} \in \mathbb{R}^W} L(\mathbf{x}^{(k+1)}, \bar{\mathbf{u}}) + \frac{\lambda_u^{(k)}}{2} \|\bar{\mathbf{u}} - \mathbf{u}^{(k)}\|_2^2 \end{cases}, \quad (2.41)$$

where  $\lambda_x$  and  $\lambda_u$  are cost-to-move penalization parameters that control the variations between two consecutive iterations.

The convergence results of the proximal alternating algorithm in (2.41) are found in Attouch *et al.* [104, Theorem 9]. They state that, if the problem in (2.39) and the algorithm in (2.41) satisfy the following key hypotheses, then

every bounded sequence  $(\mathbf{x}^{(k)}, \mathbf{u}^{(k)})$  generated by the algorithm in (2.41) converges to a critical point of  $L(\mathbf{x}, \mathbf{u})$ .

*Hypothesis 2.1.* The functions in (2.39) satisfy the following properties:

- (i) The functions  $F : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $G : \mathbb{R}^W \rightarrow \mathbb{R} \cup \{+\infty\}$  are proper l.s.c. functions.
- (ii) The function  $Q : \mathbb{R}^N \times \mathbb{R}^W \rightarrow \mathbb{R}$  is a smooth  $C^1$  function<sup>4</sup>, and its gradient  $\nabla Q$  is Lipschitz continuous (see Definition 2.5) on bounded subsets of  $\mathbb{R}^N \times \mathbb{R}^W$ .

*Hypothesis 2.2.* The functions in (2.39) and the algorithm in (2.41) satisfy the following conditions:

- (i) The function  $L(\mathbf{x}, \mathbf{u})$  is bounded below, *i.e.*,  $\inf_{\mathbb{R}^N \times \mathbb{R}^W} L > -\infty$ .
- (ii) The function  $L(\cdot, \mathbf{u}^{(1)})$  is proper.
- (iii) For some positive  $c_- < c_+$ , the sequences  $(\lambda_x^{(k)}, \lambda_u^{(k)})$  belong to  $(c_-, c_+)$  for all  $k \geq 0$ .

*Hypothesis 2.3.* The objective function  $L(\mathbf{x}, \mathbf{u})$  in (2.41) satisfies the Kurdyka-Łojasiewicz (KL) property [104, Definition 7] at each point of the domain of  $L$ , *i.e.*, the objective function  $L(\mathbf{x}, \mathbf{u})$  is a KL function.

As shown in [104], the three hypotheses above are important to demonstrate the algorithm convergence to a critical point of  $L(\mathbf{x}, \mathbf{u})$ . Hypotheses 2.1 and 2.2 are the basis of the convergence proof, since they assure a sufficient decrease of each step of the algorithm and they guarantee that the distance between iterations tends to zero when the number of iterations increases; and that the generated sequence approaches the set of critical points. Nevertheless, the main convergence result in [104] is established through the KL property in Hypothesis 2.3, which guarantees that every bounded sequence generated by the algorithm has a finite length and is convergent. The study of the KL property is out of the scope of the thesis, the reader can refer to [81] for a complete discussion. However, it is important to note that there exist a wide variety of functions that satisfy the KL property, including the semi-algebraic

<sup>4</sup> $C^1$  is the class of continuously differentiable functions.

functions [81, 82, 104] (see Definition 2.7), which will be the main focus in this work.

Although the analysis provided in [104] assumes that each subproblem in (2.41) is solved exactly, in [81] the authors prove that the convergence results also hold for an inexact version of (2.41). This means that the proximal alternating minimization algorithm in (2.41) allows a relative error tolerance and, hence, it converges to a critical point of  $L(\mathbf{x}, \mathbf{u})$  in (2.39) even for inexact minimizations of the internal subproblems.

In [82], the authors provide convergence results of the proximal alternating algorithm in (2.41) when using well-known algorithms, such as the forward-back algorithm described in Section 2.4.2.1, to solve the minimization problem at each step.

## Appendix 2.A Robust Median Estimator

In this appendix, we describe the Robust Median Estimator (RME), a technique proposed by Donoho and Johnstone [41, 105] to estimate the noise variance  $\sigma^2$  of an AWGN corrupted observation model, *i.e.*, a model described by  $\mathbf{y} = \mathbf{x} + \boldsymbol{\eta}$ , with  $\eta_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$ . The estimated variance helps to determine the noise bound in (2.10) that is used in the constrained formulation (2.8).

The RME estimation is based on the assumption that the noise is an additional high frequency component in the observed signal  $\mathbf{y}$ . Let  $\Psi \in \mathbb{R}^{N \times W}$  be a wavelet dictionary (see Appendix 2.C). The variance  $\sigma_{RME}^2$  is estimated as follows [41]:

$$\sigma_{RME}^2 = \frac{\text{med } |\boldsymbol{\alpha}_\Lambda|}{0.6745}, \quad (2.42)$$

where  $\boldsymbol{\alpha}_\Lambda$  is a vector containing the finest scale wavelet coefficients of the observed vector  $\mathbf{y}$ , *i.e.*,  $\boldsymbol{\alpha}_\Lambda = \mathbf{S}_\Lambda \Psi^* \mathbf{y}$ , with  $\mathbf{S}_\Lambda \in \mathbb{R}^{L \times W}$  the selection operator of the set  $\Lambda$  that contains the finest scale wavelet coefficients.

The estimator in (2.42) can be generalized by applying other types of high-pass filters to the noisy data [106, Section 4.3], instead of working with the finest scale wavelet coefficients  $\boldsymbol{\alpha}_\Lambda$ .

## Appendix 2.B Discrete Gradient Operator

This appendix describes the discrete formulation of the gradient operator, also called the finite differences operator. This operator is used as a sparsity dictionary (see Section 2.3.2.2) when working with the TV norm function described in Section 2.3.2.3.

Consider a regular  $N_1 \times N_2$  2-D Cartesian grid of  $N := N_1 N_2$  samples as in (2.3). Consider also a one-dimensional vector  $\mathbf{u} \in \mathbb{R}^N$  representing the 2-D description of a continuous variable  $u$  in the Cartesian pixel grid, given a convenient mapping between the components of indices  $\mathbf{u} = (u_1, \dots, u_N)^T$  and the pixel coordinates. Reusing the 2-D coordinates of  $\mathbf{u}$ , the gradient operator is defined along each direction as [76]

$$\nabla \mathbf{u} = (\nabla_1 \mathbf{u}, \nabla_2 \mathbf{u})$$

with

$$(\nabla_1 \mathbf{u})_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N_1 \\ 0 & \text{if } i = N_1 \end{cases}, \quad (2.43)$$

$$(\nabla_2 \mathbf{u})_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N_2 \\ 0 & \text{if } j = N_2 \end{cases}. \quad (2.44)$$

In Figure 2.1, we show an example of the application of the gradient operator to a piecewise constant function.

The adjoint of the gradient operator is the negative of the divergence operator and it is defined as [76]:

$$\nabla^* \mathbf{v} = - \begin{cases} (\mathbf{v}_1)_{i,j} - (\mathbf{v}_1)_{i-1,j} & \text{if } 1 < i < N_1 \\ (\mathbf{v}_1)_{i,j} & \text{if } i = 1 \\ -(\mathbf{v}_1)_{i-1,j} & \text{if } i = N_1 \end{cases} - \begin{cases} (\mathbf{v}_2)_{i,j} - (\mathbf{v}_2)_{i,j-1} & \text{if } 1 < j < N_2 \\ (\mathbf{v}_2)_{i,j} & \text{if } j = 1 \\ -(\mathbf{v}_2)_{i,j-1} & \text{if } j = N_2 \end{cases},$$

for any  $\mathbf{v} \in \mathbb{R}^{N \times 2}$ .

The induced norm of the gradient operator is defined as [107, 108]

$$\|\nabla\| = \max\{\|\nabla\mathbf{x}\|_2 : \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_2 = 1\} = \sqrt{\lambda_M(\nabla^*\nabla)},$$

where  $\lambda_M(\nabla^*\nabla)$  is the maximum eigenvalue of the matrix  $\nabla^*\nabla$ , *i.e.*, the largest eigenvalue in absolute value.

The gradient matrix is defined as  $\nabla = (\nabla_1, \nabla_2)^*$ , with  $\nabla_1$  and  $\nabla_2$  the matrices corresponding to the gradient in the horizontal and vertical directions, respectively, obtained from (2.43) and (2.44). From the properties of the matrix  $\nabla^*\nabla$  we have the following relations [109, 110]

$$\begin{aligned} \lambda_M(\nabla^*\nabla) &\leq \lambda_M(\nabla_1^*\nabla_1) + \lambda_M(\nabla_2^*\nabla_2) \\ &\leq \lambda_M(\nabla_1^*)\lambda_M(\nabla_1) + \lambda_M(\nabla_2^*)\lambda_M(\nabla_2). \end{aligned}$$

Neglecting the border conditions in (2.43) and (2.44), the matrix  $\nabla_1$  can be written as a circulant matrix [108], where the first row  $(\nabla_1)_1$  corresponds to  $(-1, 1, 0, \dots, 0)$ . From the properties of circulant matrices and from Parseval's theorem we have [108]

$$\|\lambda(\nabla_1)\|_2 = \|(\nabla_1)_1\|_2,$$

where  $\lambda(\nabla_1)$  denotes the vector containing the eigenvalues of the matrix  $\nabla_1$ . Then, from the equivalence of norms in  $\mathbb{R}^N$  we obtain [107, Section 2.2]

$$\lambda_M(\nabla_1) = \|\lambda(\nabla_1)\|_\infty \leq \|\lambda(\nabla_1)\|_2 = \|(\nabla_1)_1\|_2 \leq \|(\nabla_1)_1\|_1 = 2.$$

The same equivalences can be obtained for the matrices  $\nabla_1^*$ ,  $\nabla_2$  and  $\nabla_2^*$ . The induced norm of the gradient operator is then bounded as

$$\|\nabla\| \leq 2\sqrt{2}.$$

## Appendix 2.C Wavelets Dictionaries

In this appendix, we briefly describe the wavelets dictionaries, a common set of dictionaries whose elements are a collection of parametrized waveforms.

This set of dictionaries allows obtaining sparse representations of natural images in both *synthesis* and *analysis* modes [11,41,42,91].

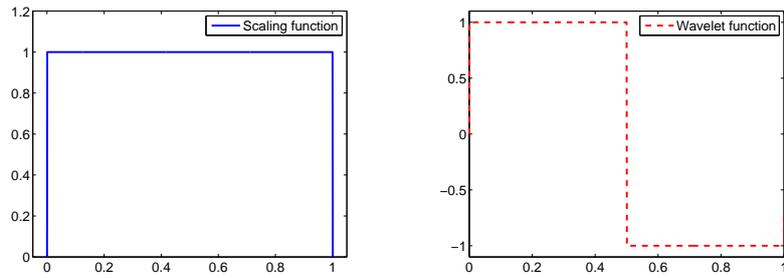
Compared to the frequency representation as obtained by the Fourier dictionary, 2-D wavelets dictionaries allow representing images with a *multi-resolution* scheme [111], *i.e.*, as a linear combination of localized *wave atoms* with variable sizes and locations. Wavelets dictionaries are based on a collection of translations and dilations of a basic “mother wavelet”, also called *wavelet function*, together with translations of a “father wavelet”, also called *scaling function* [58].

In the rest of this appendix, we focus on the family of Daubechies wavelets [112], the dictionary used in Chapters 4 and 5. First, the discussion is focused on the orthonormal basis and then the discussion is extended to the redundant case.

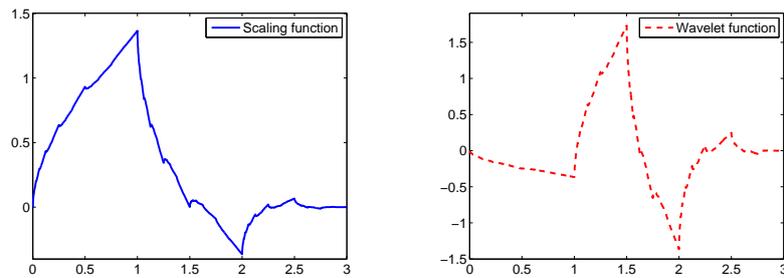
The set of Daubechies wavelets is characterized by having a compact support and a number of vanishing moments  $v$  that corresponds to the maximal allowed by the support width (given by  $2v - 1$ ). The different wavelets belonging to this family are denoted as  $dbv$ , *e.g.*,  $db1$  denoting the Daubechies wavelet with one vanishing moment. The vanishing moments determine the time-frequency localization of the waveform, a larger  $v$  implying a better localization. The vanishing moments also determine the smoothness of the wavelet, hence their ability to represent polynomial behaviors in an image, *e.g.*, one vanishing moment represents constant components in an image while three vanishing moments represent constant, linear and quadratic components in an image. Some *wavelet* and *scaling* functions of the Daubechies family are illustrated in Figure 2.11 for different vanishing moments.

The Daubechies wavelets family does not have an analytic expression for the waveforms, with the exception of the  $db1$ , also called *Haar* wavelet. The waveforms in Figure 2.11 are originated from the multiresolution analysis theory under some normalization and orthogonality constraints [111,112].

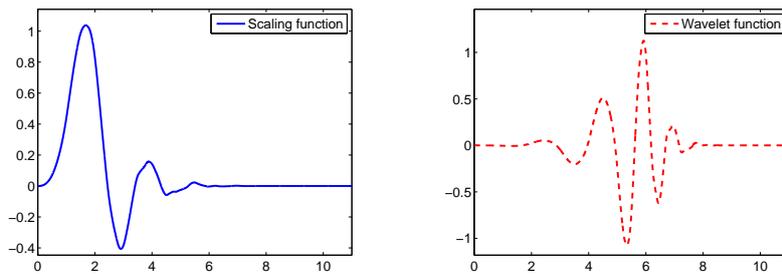
In practice, the wavelet decomposition is done through the discrete wavelet transform (DWT), a fast implementation consisting of a cascade algorithm applied to the image [111,112]. This algorithm is based on the application of two quadrature mirror filters: the *scaling* filter, a low pass filter



(a) db1 or Haar: 1 vanishing moment, filter with 2 taps



(b) db2: 2 vanishing moments, filter with 4 taps



(c) db6: 6 vanishing moments, filter with 12 taps

**Figure 2.11** Illustration of the Daubechies wavelet family for (a) 1 vanishing moment (db1 or Haar), (b) 2 vanishing moments (db2) and (c) 6 vanishing moments (db6). The left column contains the scaling functions and the right column contains the wavelet functions.

associated to the *scaling* function, and the *wavelet* filter, a high pass filter associated with the *wavelet* function. The length of these filters is defined by the number of vanishing moments of the wavelet function, *i.e.*, for  $v$  vanishing moments, the filters have a length of  $2v$  taps.

The wavelet decomposition of an image is done by scales or levels. At each level we apply the *scaling* and the *wavelet* filters to obtain, respectively, the *approximation* and *detail* coefficients (after a decimation process by a factor of 2). In the successive scales, the filters are again applied to the *approximation* coefficients. This process can be repeated for a maximum of  $L = \log N$  scales, with  $N$  the dimension of the image. Note that, for the 2-D wavelet decomposition, three sets of *detail* coefficients are obtained at each scale, each one representing a given direction (horizontal, vertical and diagonal).

An example of the wavelet decomposition of a natural image using DWT is given in Figure 2.2-(b). We can see that the wavelet coefficients are sparse in most scales. This shows their importance in the regularization of inverse problems by promoting sparsity [41, 42, 53, 60, 61]. However, we note that the coarsest scale (located in the upper left corner of the image), containing the scaling or *approximation* coefficients, is not sparse. In reconstruction algorithms, it is a common practice to remove the scaling coefficients in order to have a sparse vector of coefficients [41, 111, 113].

The wavelet transform described above, which is orthogonal and maximally-decimated, presents one main disadvantage: the transformation is translation variant, *i.e.*, the wavelet coefficients vary with shifts of position in the image [113]. One solution to this problem is to eliminate the decimation step in the DWT, which results in the Undecimated DWT (UDWT) [27, 65, 111]. The UDWT can also be seen as the union of all translations of an orthonormal DWT. Conversely to the DWT, this makes the UDWT translation invariant, *i.e.*, it enables an efficient characterization of all image features whatever is their location [114, 115], hence providing a better image reconstruction than the traditional DWT.

An example of a wavelet decomposition of a natural image using UDWT is given in Figure 2.2-(c). Remark that the decomposition is done using the *analysis* operator, *i.e.*,  $\bar{\mathbf{a}} = \mathbf{\Psi}^* \mathbf{x}$ . Since there is no decimation, all the scales in Figure 2.2-(c) are of the same size, equal to the size of the image  $\mathbf{x}$  in Figure 2.2-

(a). We can see that, as in the case of the DWT, the scaling coefficients, located in the upper left corner of the image, are not sparse and can be removed when used in reconstruction algorithms.

## Appendix 2.D Details on the Expanded Version of the Chambolle-Pock Algorithm

In this appendix, we provide some details on the Chambolle-Pock (CP) algorithm expanded for the sum of several convex functions.

### 2.D.1 Useful Computations

By comparing (2.25) and (2.35), we can define the function  $G(\mathbf{x}')$  as:  $G(\mathbf{x}') = \sum_{j=2}^L t_{\Pi_{1,j}}(\mathbf{x}') + H(\mathbf{x}_1)$ , for every vector  $\mathbf{x}' = (\mathbf{x}_1^T, \dots, \mathbf{x}_L^T)^T \in \mathbb{R}^{LN}$ . The set  $\Pi_{1,j}$  denotes the bisector plane  $\Pi_{1,j} = \{\mathbf{x}' \in \mathbb{R}^{LN} : \mathbf{x}_1 = \mathbf{x}_j\}$  for  $j \in [L]$ .

Consider a vector  $\mathbf{u} = (\mathbf{u}_1^T, \dots, \mathbf{u}_L^T)^T \in \mathbb{R}^{LN}$ . The dual function  $G^*(\mathbf{x}')$  is computed using the Legendre transform in (2.20) as follows:

$$\begin{aligned} G^*(\mathbf{x}') &= \max_{\mathbf{u}} \langle \mathbf{u}, \mathbf{x}' \rangle - G(\mathbf{u}) = \max_{\mathbf{u}: \mathbf{u}_1 = \dots = \mathbf{u}_L} \langle \mathbf{u}_1, \sum_{j=1}^L \mathbf{x}_j \rangle - H(\mathbf{u}_1) \\ &= \begin{pmatrix} H^* \left( \sum_{j=1}^L \mathbf{x}_j \right) \\ \vdots \\ H^* \left( \sum_{j=1}^L \mathbf{x}_j \right) \end{pmatrix} = \begin{pmatrix} \mathbf{I}_N \\ \vdots \\ \mathbf{I}_N \end{pmatrix} H^* \left( \sum_{j=1}^L \mathbf{x}_j \right). \end{aligned}$$

This result is used to build the extended dual problem in (2.36).

Let us now compute the proximal operator of the extended function  $G(\mathbf{x}')$  in order to formulate the CP iterations in the product space (2.37). For a vector  $\zeta = (\zeta_1, \dots, \zeta_L) \in \mathbb{R}^{LN}$ , the proximal operator of  $G(\mathbf{x}')$  reduces to

$$\tilde{\mathbf{x}}' = \text{prox}_{\mu G} \zeta = \underset{\mathbf{x}': \mathbf{x}_1 = \dots = \mathbf{x}_L}{\text{argmin}} \mu H(\mathbf{x}_1) + \frac{1}{2} \|\zeta - \mathbf{x}'\|_2^2.$$

As all the  $\mathbf{x}_j$  are equal, we necessarily have  $\tilde{\mathbf{x}}' = (\mathbf{I}_N, \dots, \mathbf{I}_N)^T \tilde{\mathbf{u}}$ , with  $\tilde{\mathbf{u}} \in \mathbb{R}^N$  defined by

$$\begin{aligned} \tilde{\mathbf{u}} &= \underset{\mathbf{u} \in \mathbb{R}^N}{\operatorname{argmin}} \mu H(\mathbf{u}) + \frac{1}{2} \sum_{j=1}^L \|\zeta_j - \mathbf{u}\|_2^2 \\ &= \underset{\mathbf{u} \in \mathbb{R}^N}{\operatorname{argmin}} \mu H(\mathbf{u}) + \frac{1}{2} \left( \|\zeta_1\|_2^2 + \dots + \|\zeta_L\|_2^2 - 2L \bar{\zeta}^T \mathbf{u} + L \|\mathbf{u}\|_2^2 \right) \\ &= \underset{\mathbf{u} \in \mathbb{R}^N}{\operatorname{argmin}} \mu H(\mathbf{u}) + \frac{1}{2} \left( L \|\bar{\zeta}\|_2^2 - 2L \bar{\zeta}^T \mathbf{u} + L \|\mathbf{u}\|_2^2 \right) \\ &= \operatorname{prox}_{(\mu/L)H} \bar{\zeta}, \end{aligned}$$

with  $\bar{\zeta} = \frac{1}{L} \sum_j \zeta_j \in \mathbb{R}^N$  and by subtracting and adding constants in the minimization (since this does not disturb its solution). Denoting by  $\mathbf{I}_N^L = (\mathbf{I}_N, \dots, \mathbf{I}_N) \in \mathbb{R}^{N \times LN}$  the operator such that  $\mathbf{I}_N^L \zeta = L \bar{\zeta}$ , this finally provides the following compact notation

$$\operatorname{prox}_{\mu G} \zeta = \left( \mathbf{I}_N^L \right)^T \operatorname{prox}_{\frac{\mu}{L} H} \frac{1}{L} \mathbf{I}_N^L \zeta.$$

### 2.D.2 Iterations of the Extended Chambolle-Pock Algorithm

We take the CP algorithm in (2.32) and we translate it into the expanded optimization space. For  $\vartheta = 1$ , we obtain the following:

$$\left\{ \begin{array}{l} \mathbf{s}^{(k+1)} = \begin{pmatrix} \mathbf{s}_1^{(k+1)} \\ \vdots \\ \mathbf{s}_L^{(k+1)} \end{pmatrix} = \begin{pmatrix} \operatorname{prox}_{\nu F_1^*} \left( \mathbf{s}_1^{(k)} + \nu \mathbf{K}_1 \bar{\mathbf{x}}_1^{(k)} \right) \\ \vdots \\ \operatorname{prox}_{\nu F_L^*} \left( \mathbf{s}_L^{(k)} + \nu \mathbf{K}_L \bar{\mathbf{x}}_L^{(k)} \right) \end{pmatrix} \\ \mathbf{x}'^{(k+1)} = \begin{pmatrix} \mathbf{x}_1^{(k+1)} \\ \vdots \\ \mathbf{x}_L^{(k+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_N \\ \vdots \\ \mathbf{I}_N \end{pmatrix} \operatorname{prox}_{\frac{\mu}{L} H} \left( \frac{1}{L} \sum_{j=1}^L \mathbf{x}_j^{(k)} - \mu \mathbf{K}_j^* \mathbf{s}_j^{(k+1)} \right) . \\ \bar{\mathbf{x}}'^{(k+1)} = \begin{pmatrix} \bar{\mathbf{x}}_1^{(k+1)} \\ \vdots \\ \bar{\mathbf{x}}_L^{(k+1)} \end{pmatrix} = 2 \begin{pmatrix} \mathbf{x}_1^{(k+1)} \\ \vdots \\ \mathbf{x}_L^{(k+1)} \end{pmatrix} - \begin{pmatrix} \mathbf{x}_1^{(k)} \\ \vdots \\ \mathbf{x}_L^{(k)} \end{pmatrix} \end{array} \right.$$

As  $\mathbf{x}_1 = \mathbf{x}_j$ , for  $2 \leq j \leq L$ , we can relabel the variable as  $\mathbf{x}^{(k)} = \mathbf{x}_1^{(k)} = \mathbf{x}_j^{(k)}$  and  $\mathbf{x}^{(k+1)} = \mathbf{x}_1^{(k+1)} = \mathbf{x}_j^{(k+1)}$ . In the same way,  $\bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_j$ , for  $2 \leq j \leq L$ , thus we can relabel the variable as  $\bar{\mathbf{x}}^{(k)} = \bar{\mathbf{x}}_1^{(k)} = \bar{\mathbf{x}}_j^{(k)}$  and  $\bar{\mathbf{x}}^{(k+1)} = \bar{\mathbf{x}}_1^{(k+1)} = \bar{\mathbf{x}}_j^{(k+1)}$ . We obtain the following algorithm:

$$\begin{cases} \mathbf{s}_j^{(k+1)} &= \text{prox}_{vF_j^*} \left( \mathbf{s}_j^{(k)} + v\mathbf{K}_j\bar{\mathbf{x}}^{(k)} \right), j \in \{1, \dots, L\} \\ \mathbf{x}^{(k+1)} &= \text{prox}_{\frac{\mu}{L}H} \left( \mathbf{x}^{(k)} - \frac{\mu}{L} \sum_{j=1}^L \mathbf{K}_j^* \mathbf{s}_j^{(k+1)} \right) \\ \bar{\mathbf{x}}^{(k+1)} &= 2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \end{cases} .$$

# Compressive Optical Deflectometric Tomography

## 3

In this chapter, we study an inverse problem existing in optical deflectometric tomography (ODT), where we aim at recovering the spatial distribution of the refractive index map (RIM) of a transparent object from light deflection measurements under multiple orientations. We show that the ODT imaging modality can be made *compressive, i.e.*, a correct RIM reconstruction is achievable with far less observations than required by traditional methods, *e.g.*, minimum energy (ME) methods or filtered back projection (FBP). Assuming a cartoon-shape RIM model, the reconstruction is driven by minimizing the map Total-Variation under a fidelity constraint with the available observations. Moreover, two other realistic assumptions are added to improve the stability of our approach: the map non-negativity and a frontier condition. Numerically, our method relies on an accurate ODT sensing model and on a primal-dual minimization scheme that can include the sensing operator and the proposed RIM constraints. We conclude this chapter by demonstrating the power of our method on synthetic and experimental data under various compressive scenarios. The contents in this chapter are largely based on [28].

## 3.1 Introduction

Optical deflectometric tomography (ODT) is an imaging modality that aims at reconstructing the spatial distribution of the refractive index of a transparent object from the deviation of the light passing through the object. By

reconstructing the refractive index map (RIM) we are able to optically characterize transparent materials like optical fibers or multifocal intra-ocular lenses (IOL), which is of great importance in manufacture and control processes.

ODT is attractive for its high sensitivity and effective detection of local details and object contours. The technique is insensitive to vibrations and it does not require coherent light. Compared to interferometry, ODT can be applied to objects with higher refractive index difference with respect to the surrounding solution.

The deflectometer used for the data acquisition is based on the phase-shifting Schlieren method [26]. For each orientation of the sample, two-dimensional (2-D) mappings of local light deviations are measured. As these light deviations are proportional to the transverse gradient of the RIM integrated along the light ray, ODT is able to reconstruct the RIM from the deflection measurements.

First works on deflectometric tomography [116, 117] have focused on the use of common reconstruction techniques like the filtered back projection (FBP). They have proved that FBP provides an accurate estimation of the RIM when a sufficient amount of object orientations is available. However, when we consider a scenario with a limited number of light incident angles, and in the presence of noise in the ODT observations, FBP induces the apparition of spurious artifacts in the estimated RIM, lowering the reconstruction quality.

Inspired by the compressed sensing (CS) paradigm [118, 119], which demonstrates that few measurements are enough for an accurate reconstruction of low complexity (*e.g.*, sparse) signals, recent works in ODT have started to exploit sparsity to reconstruct the RIM from few acquisitions. Fomouo *et al.* [26] and Antoine *et al.* [37] have used a sparse representation in a B-splines basis and regularized the problem using the  $\ell_1$  norm. The reconstruction was performed using iterative schemes and the results show that, although the method is capable of reproducing the shape of spatially localized objects, the image dynamics is not well preserved and the borders are smoothed.

Although sparsity based methods are new in ODT, they have been used in other applications, such as magnetic resonance imaging [2, 79, 102], absorption tomography [4, 23, 120], radio interferometry [11, 50], phase-contrast tomogra-

phy [75,121] and positron emission tomography [4], for image reconstruction under partial observation models. More details are given in Section 3.4.

An additional problem that rises in all physical applications is the modeling of the sensing operator that fits better the physical acquisition. Most operators present a non ideal behavior, which conditions the numerical methods to solve the inverse problem. In tomographic problems, this operator requires to map spatial data in a Cartesian grid to Fourier data in a Polar grid. Previous works have used *gridding* techniques to interpolate data from a polar to a Cartesian or pseudo polar grid before applying the Fourier Transform [122,123]. However, the error introduced when using these techniques is not bounded and introduces an uncontrolled distortion.

### 3.1.1 Contributions

In this work, we show that the ODT can be made both compressive and robust to Gaussian noise. In the context of a simplified 2-D sensing model, we propose a constrained method based on the minimization of the Total-Variation (TV) norm that provides high quality reconstruction results even when few acquisitions are available and in the presence of high levels of noise.

This is motivated by assuming the RIM to be composed of slowly varying areas separated by sharp transitions corresponding to material interfaces. Such a behavior is known to be represented by spatial functions having *bounded variations* (small TV norm), such as those following a cartoon-shape model. This also distinguishes our work from two previous studies focused on continuous RIM decomposition with a B-splines basis. Deflection integrals were there estimated in the spatial domain using complex numerical methods leading to a smoother RIM estimation [26,37].

To account for the noise and the raw data consistency, we add an  $\ell_2$  data fidelity term adapted to Gaussian and uniformly distributed noise. Moreover, the proposed method offers the flexibility to work with more than one constraint. In contrast with [124], we add here two more constraints based on general prior information on the RIM in order to converge to an optimal solution: (i) the RIM is positive everywhere and (ii) the object is completely con-

tained in the experimental field-of-view (FoV). These extra constraints probably guarantee the uniqueness of the ODT solution.

Since the ODT imaging modality is fairly new in the literature, little effort has been done in the appropriate characterization of the discrete forward model and of the noise corrupting the acquisition and modeling processes. Hence, in this work, we aim at performing an accurate formulation of the discrete forward model and a thoughtful analysis of the distortions corrupting this model. As for the operator, we use the NFFT algorithm<sup>1</sup>, a fast computation of the non-equispaced DFT. This algorithm provides an efficient estimation of the polar Fourier transform with a controlled distortion regarding the true polar transform.

The compressive ODT problem is solved by means of the primal-dual algorithm proposed by Chambolle and Pock [74], complemented by an adaptive choice of its parameters improving the global convergence speed, as proposed by Goldstein *et al.* [97]. The results are compared with a minimum energy (ME) method and a common FBP approach, showing clear gain compared to these traditional approaches in terms of compressiveness, noise robustness and reconstruction quality.

### 3.1.2 Outline

The rest of the chapter is organized as follows. In Section 3.2, we provide a brief background on optical deflectometric tomography, describing also its relation with other tomographic modalities and the experimental setup used for the data acquisition. Then, the ODT discrete model is presented in Section 3.3. Section 3.4 depicts related works on tomographic reconstruction, which provide a basis on the methods adopted to recover the RIM: the commonly used FBP method, a standard minimum energy (ME) approach (or penalized least square) and the proposed regularized method called TV- $\ell_2$ . These methods are described in Section 3.5. In Section 3.6, we present the identification and estimation of the noise in both synthetic and experimental data, and the analysis of the noise impact when comparing common absorption tomography and deflection tomography. Section 3.7 presents the numerical methods used to

---

<sup>1</sup>This toolbox is freely available here <http://www-user.tu-chemnitz.de/~potts/nfft/>.

recover the RIM from the noisy measurements by means of the proposed regularized formulation. In Section 3.8, some reconstruction results are shown, focusing first on the comparison between common tomographic and ODT reconstructions, and then on the comparison of the reconstruction methods on the basis of reconstruction quality and convergence for both synthetic and experimental data. Finally, Section 3.9 concludes the chapter.

## 3.2 Optical Deflectometric Tomography

In this section, the principles of optical deflectometric tomography are explained and completed with a brief description of the experimental setup used for actual deflectometric acquisition.

### 3.2.1 Principles

Optical deflectometric tomography aims at inferring the refractive index spatial distribution (or refractive index map – RIM) of a transparent object. This is achieved by measuring, under various incident angles, the deflection angles of multiple parallel light rays when passing through this transparent object (see Figure 3.1-(a)). The (indirect) observation of the RIM, allowing its further reconstruction, is guaranteed by the relation between the total light ray deflection and the integration of the RIM transverse gradient along the light ray path [116].

In this work, we restrict ourselves to two-dimensional (2-D) ODT by assuming that the refractive index  $n$  of the observed object is constant along the  $\mathbf{e}_3$ -axis for a convenient coordinate system  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , *i.e.*,  $\partial n(\mathbf{r})/\partial r_3 = 0$  (with  $\mathbf{r} = (r_1, r_2, r_3)^T \in \mathbb{R}^3$ ) and deflections occur in the  $\mathbf{e}_1$ - $\mathbf{e}_2$  plane. This assumption is validated in the experimental setup described later in this section, where the light probes a very thin 2-D slice of the three dimensional (3-D) sample and where the particular geometry of the test objects makes the refractive index variation along the  $\mathbf{e}_3$  direction negligible, *i.e.*,  $\left| \frac{\partial}{\partial r_3} n \right| \ll \max \left\{ \left| \frac{\partial}{\partial r_1} n \right|, \left| \frac{\partial}{\partial r_2} n \right| \right\}$ .

Given the refractive index  $n : \mathbf{r} = (r_1, r_2)^T \in \mathbb{R}^2 \rightarrow n(\mathbf{r})$ , for a particular light ray trajectory  $\mathcal{R} = \{\mathbf{r}(s) : s \in \mathbb{R}\} \subset \mathbb{R}^2$  parametrized by

$\mathbf{r}(s) = (r_1(s), r_2(s))^T \in \mathbb{R}^2$  with  $s$  describing its *curvilinear* parameter<sup>2</sup>, the relation between light deflections and the refractive index  $n$  is provided by the *light ray equation*

$$\frac{d}{ds} \left( n \frac{d}{ds} \mathbf{r}(s) \right) = \nabla n, \quad (3.1)$$

established from the eikonal equation [125, Section 3.2].

For small deflection angles, we can adopt a (first order) approximation and assume the trajectory  $\mathcal{R}$  is a straight line. The error committed by removing the trajectory dependence in the deflection angle can be estimated by a *ray tracing* method relying on Snell's law (see Appendix 3.A.3). In our tests, for simple continuous object models, the absolute error between the two deflection models is lower than  $0.7^\circ$  for deflection angles smaller than  $7^\circ$  (the angular acceptance of the optical deflectometer). This is verified by the work of Antoine *et al.* [37], who have qualitatively shown that, for the same range of deflection angles and for limited refracted index variations, the model mismatch stays limited. Furthermore, this model mismatch is expected to be sparse since it only occurs for a few number of light rays (see Appendix 3.A.3).

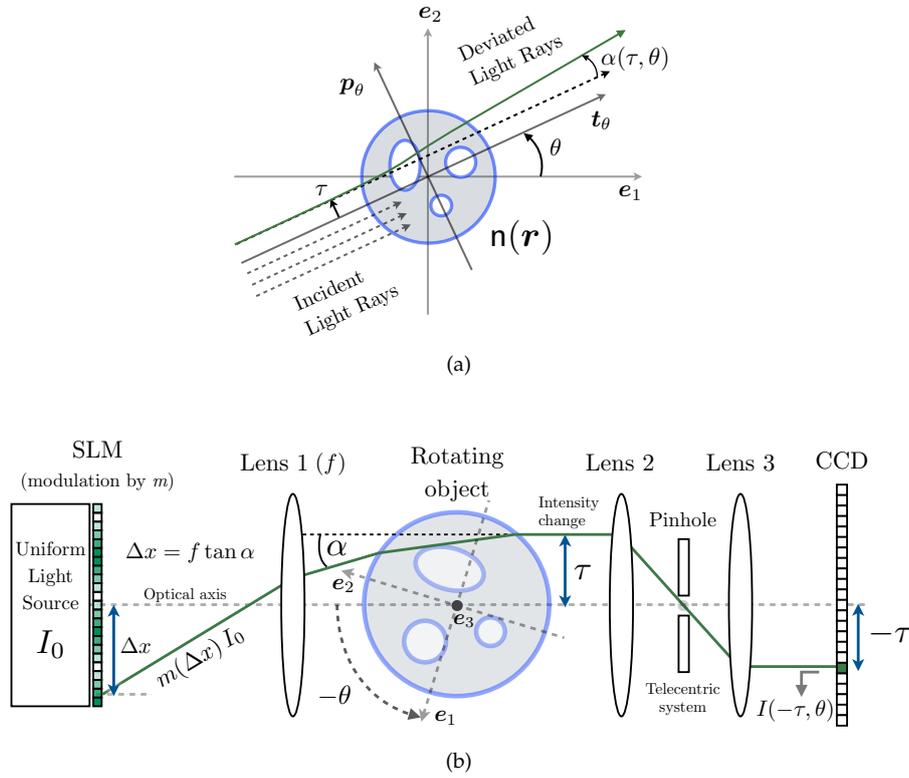
In this simplified model, the 2-D RIM is measured by  $\Delta(\tau, \theta; n)$ , the sinus of the deflection angle  $\alpha(\tau, \theta; n)$  of a light ray  $\mathcal{R}(\tau, \theta) = \{\mathbf{r} \in \mathbb{R}^2 : \mathbf{r} \cdot \mathbf{p}_\theta = \tau\}$ , where  $\tau \in \mathbb{R}$  is the *affine* parameter determining the distance between  $\mathcal{R}$  and the origin,  $\theta \in [0, 2\pi)$  is the incident angle of  $\mathcal{R}$  with the  $\mathbf{e}_1$  axis, and  $\mathbf{p}_\theta = (-\sin \theta, \cos \theta)^T$  is the (constant) perpendicular vector to the light ray direction  $\mathbf{t}_\theta = (\cos \theta, \sin \theta)^T$ .

The simplified ODT model depicted in Figure 3.1-(a) is then obtained from the light ray equation (3.1) as (see details in Appendix 3.A.1)

$$\Delta(\tau, \theta) := \frac{1}{n_r} \int_{\mathbb{R}} \nabla n(\mathbf{r}_{\tau, \theta}(s)) \cdot \mathbf{p}_\theta ds = \frac{1}{n_r} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) \delta(\tau - \mathbf{r} \cdot \mathbf{p}_\theta) d^2 \mathbf{r}, \quad (3.2)$$

where  $n_r$  is the (constant) reference refractive index of the surrounding medium,  $\mathbf{r}_{\tau, \theta}(s) = s\mathbf{t}_\theta + \tau\mathbf{p}_\theta \in \mathcal{R}$  and the Dirac distribution turns the line integral into an integration over  $\mathbb{R}^2$ . In short, the above equation relates the deflection angle  $\Delta(\tau, \theta)$  to the Radon transform of the transverse gradient of  $n$  within the straight line trajectory approximation.

<sup>2</sup>By a slight abuse of notation,  $\mathbf{r}$  denotes any point in  $\mathbb{R}^2$  while  $\mathbf{r}(s)$  represents a particular curve in  $\mathbb{R}^2$  parametrized by  $s \in \mathbb{R}$ .



**Figure 3.1** (a) The deflectometric model. (b) A scheme of a Phase-Shifting Schlieren Deflectometer (in refraction mode) measuring light ray deflection angles by encoding light deviation into light intensity variations.

*Remark 3.1.* We should note that the simplified ODT model in (3.2) is similar to the sensing model in X-ray differential phase contrast tomography [1, 75, 121], where the measured phase-shifts are related to the transverse gradient of the decrement of the real part of an object's refractive index (quantity of interest). However, due to the physics of the acquisition process and the properties of the X-ray beams, no first order approximation (as the one done in ODT) is necessary in order to obtain a linear model as the one in (3.2).

As for traditional parallel absorption tomography [22, Section 3.2], there exists a deflectometric Fourier slice theorem (DFST) that relates the 1-D (ra-

dial) Fourier transform of the deflection angle along the *affine parameter*  $\tau$ , i.e.,

$$y(\omega, \theta) := \int_{\mathbb{R}} \Delta(\tau, \theta) e^{-2\pi i \tau \omega} d\tau, \quad (3.3)$$

to the 2-D Fourier transform of the RIM. Mathematically, the DFST establishes the following equivalence [122] (proved in Appendix 3.A.2):

$$y(\omega, \theta) = \frac{2\pi i \omega}{n_r} \hat{n}(\omega \mathbf{p}_\theta), \quad (3.4)$$

where  $\hat{n}(\mathbf{k}) = \int_{\mathbb{R}^2} n(\mathbf{r}) e^{-2\pi i \mathbf{k} \cdot \mathbf{r}} d^2 \mathbf{r}$  stands for the 2-D Fourier transform of  $n$ . Hereafter, the value  $\omega$  is called *affine frequency*.

*Remark 3.2.* We should note that there are different ways to cover the 2-D frequency plane with deflectometric measurements. This can be observed from the following symmetry relations for  $n \in \mathbb{R}$ ,  $\tau, \omega \in \mathbb{R}$ ,  $\theta \in [0, 2\pi)$ :

$$\Delta(\tau, (\theta + \pi) \bmod 2\pi) = -\Delta(-\tau, \theta), \quad (3.5a)$$

$$y(\omega, (\theta + \pi) \bmod 2\pi) = -y(-\omega, \theta), \quad (3.5b)$$

$$y(\omega, \theta) = y^*(-\omega, \theta), \quad (3.5c)$$

where the two first relations come from the change  $\mathbf{p}_{\theta+\pi} = -\mathbf{p}_\theta$  in (3.2) and (3.4), and the last equation is due to the Fourier conjugate symmetry of real spatial functions. In particular, from the symmetry (3.5c), we can restrict  $\omega$  to positive values by taking  $\theta$  in the whole circle  $[0, 2\pi)$ . Or alternatively, from (3.5b) and (3.5c), we can deduce  $y(\omega, \theta) = -y^*(\omega, (\theta + \pi) \bmod 2\pi)$  and restrict  $\theta$  to the half circle  $[0, \pi)$  with  $\omega \in \mathbb{R}$ . We insist on the fact that this symmetry is only preserved when the first order approximation is validated. When this approximation is not valid, e.g., when imaging objects with large refractive index variations, the symmetry is broken [37].

*Remark 3.3.* When comparing the relation (3.4) with the standard tomographic Fourier slice theorem (FST) [22, Section 3.2], the main difference is provided by the transverse gradient in the deflectometric relation (3.2), which results in multiplying by  $2\pi i \omega / n_r$  the RIM Fourier transform.

*Remark 3.4.* Since  $\omega$  vanishes on the frequency origin, we can see from (3.2) or from (3.4) that the ODT sensing is blind to a constant RIM. As we will see in Section 3.5, this implies that the RIM can only be estimated relatively to the known reference RIM  $n_r$ .

Let us conclude this section by insisting on the impact of (3.4). Similarly to the use of the Fourier slice theorem in common (absorption) tomography, the relation in (3.4) is of great importance for defining a discrete ODT sensing model which can be computed efficiently in the Fourier domain given a discretized refractive index map  $n$ . Note that such fast and efficient computation in the Fourier domain can only be achieved because the first order approximation was adopted (an assumption that is not necessary in common X-ray absorption tomography).

### 3.2.2 Deflection Measurements

Experimentally, the deflection angles can be measured by phase-shifting Schlieren deflection tomography (schematically represented in Figure 3.1-(b)). We briefly explain this system for the sake of completeness in order to set the experimental background surrounding the actual deflection measurement process. More details can be found in [6, 26, 116, 122].

This system proceeds by encoding light ray deflection  $\alpha$  in intensity variations. A transparent object is illuminated with an incoherent uniform light source  $I_0$  modulated by a sinusoidal pattern  $m$  using a spatial light modulator (SLM). From classical optics, the light deviation angle  $\alpha$  is related to a phase shift  $\Delta x = f \tan \alpha$ , where  $f$  is the focal length of Lens 1. This phase shift is associated with the intensity variation thanks to the modulation  $m$  as  $I(-\tau, \theta) = m(\Delta x)I_0$ . These intensity variations are processed by phase-shifting methods [126] for recovering the deflection measurements  $\alpha(\tau, \theta)$  for each couple of parameters  $(\tau, \theta) \in \mathbb{R} \times [0, 2\pi]$ . Up to some linear coordinate rescaling, the *affine* parameter  $\tau$  corresponds to the horizontal pixel coordinate in the 2-D detector (consisting of a charge-coupled device – CCD) collecting light (assuming the object refractive index constant along the CCD vertical direction). This correspondence is implicitly allowed by the telecentric system formed by the combination of Lens 2 and Lens 3. The pinhole guarantees that only parallel light rays outgoing from the object are collected. Rather than rotating the whole incident light beam around the object, it is this one which is rotated by an angle  $-\theta$  along an axis parallel to the CCD pixel vertical direction [116]. Finally, since the system is invariant under time inversion, *i.e.*,

under light progression inversion, measuring the deflection angle  $\alpha$  in Figure 3.1-(b) is equivalent to measuring the same angle in Figure 3.1-(a).

It is important to note that, in order to obtain the deflection angles from the measured intensity variations, the phase-shifting methods require several acquisitions for different patterns generated by the SLM [37]. The number of acquisitions increases if the object is analyzed in three dimensions since, in this case, the deflection do not correspond to a single point (as it is assumed here) but to a deflection map instead [27]. Sudhakar *et al.* [27] demonstrated that the ODT imaging modality can be made compressive in terms of acquisition, *i.e.*, the (sparse) deflection map can be obtained from few acquisitions (few SLM patterns) made by the Schlieren deflectometer depicted in Figure 3.1-(b).

### 3.3 Discrete Forward Model

In order to reconstruct efficiently the RIM from ODT measurements, the recorded data must be treated appropriately considering, jointly, the data discretization, the polar geometry of the ODT sensing and the unavoidable measurement noise. We present here the discrete formulation of the ODT sensing and the construction of the forward model from the recorded data.

#### 3.3.1 Discrete Domains

Let us first assume that the object of interest is fully contained in a square field-of-view (FoV)  $\Omega \subset \mathbb{R}^2$  centered on the spatial origin. The physical dimensions of this FoV can be provided by the deflectometric device itself. In other words, the RIM is constant and equal to the reference index  $n_r$  outside of  $\Omega$ . This involves also that the deflection measurement vanishes, *i.e.*,  $\Delta(\tau, \omega) = 0$ , if  $|\tau|$  is bigger than the typical width of  $\Omega$  in a section of direction  $\theta + \pi/2$ .

We can consider a spatial sampling of  $\Omega$  as follows. We define a  $N_0 \times N_0$  2-D Cartesian grid of  $N := N_0^2$  pixels as

$$\mathcal{C}_N = \{\mathbf{r}_{m,n} := (m \delta r, n \delta r) : -N_0/2 \leq m, n < N_0/2\},$$

where the spatial spacing  $\delta r$  is adjusted to  $\Omega$  and to the resolution by imposing  $\Omega = [-\frac{1}{2}N_0 \delta r, \frac{1}{2}N_0 \delta r] \times [-\frac{1}{2}N_0 \delta r, \frac{1}{2}N_0 \delta r]$ .

Second, as the deflectometric experiments provide evenly sampled variables  $\tau$  and  $\theta$ ,  $\Delta$  is measured on a (signed) regular polar coordinate grid

$$\mathcal{P}_M := \{(\tau_s, \theta_t) : -(N_\tau/2) \leq s < (N_\tau/2), 0 \leq t < N_\theta\}, \quad \tau_s := s \delta\tau, \quad \theta_t := t \delta\theta,$$

of size  $M := N_\tau N_\theta$ , with  $N_\tau$  the number of parallel light rays passing through the object (assumed even),  $N_\theta$  the number of incident angles in ODT sensing,  $\delta\tau$  and  $\delta\theta = \pi/N_\theta$  the distance between two consecutive affine parameters and angles, respectively<sup>3</sup>.

The value  $\delta\tau$  can be known experimentally from the pixel size of the CCD detector in a Schlieren Deflectometer (see Section 3.2.2). Moreover, the value  $\delta\tau$  and the resolution  $N_\tau$  are also related to the FoV  $\Omega$  so that  $\delta\tau N_\tau \approx \delta r N_0$ . We consider hereafter that the resolution of the sampling  $\mathcal{C}_N$  is equal to that of the affine variations of the ODT measurements, *i.e.*,  $\delta\tau \approx \delta r$  and  $N_\tau \approx N_0$ .

Third, in this discretized setting, the affine frequency  $\omega$  in (3.4) must also be sampled with  $N_\tau$  values. As described in the next section, this comes from the replacement of the (radial) Fourier transform in (3.3) by its discrete counterpart. This leads to a (signed) frequency polar grid of same size

$$\begin{aligned} \widehat{\mathcal{P}}_M := \{(\omega_{s'}, \theta_t) : -(N_\tau/2) \leq s' < (N_\tau/2), 0 \leq t < N_\theta\}, \\ \omega_{s'} := s' \delta\omega, \quad \theta_t := t \delta\theta, \end{aligned}$$

with  $\delta\omega = 1/(N_\tau \delta\tau)$ . As it will become clearer in the following, only half of this polar grid will be necessary to bring independent observations of the RIM, *i.e.*, we will often work on

$$\widehat{\mathcal{P}}_M^+ := \{(\omega_{s'}, \theta_t) : 0 \leq s' < (N_\tau/2), 0 \leq t < N_\theta\},$$

with  $|\widehat{\mathcal{P}}_M^+| = M/2$ .

### 3.3.2 Discretized Functions

From the discrete domains defined above, the continuous RIM observed in the experimental FoV is discretized into a set of  $N$  values  $n(\mathbf{r}_{m,n})$  from the coordinates  $\mathbf{r}_{m,n} \in \mathcal{C}_N$ . This description is arranged into a one dimensional

<sup>3</sup>Notice that  $\delta\theta$  is not set to  $2\pi/N_\theta$  since  $\tau$  is allowed to be negative.

vector  $\mathbf{n} \in \mathbb{R}^N$  using a convenient mapping between the indices of the components of  $\mathbf{n} = (n_1, \dots, n_N)^T$  and the pixel coordinates in  $\mathcal{C}_N$ .

For the different functions discretized on  $\mathcal{P}_M$  or on  $\widehat{\mathcal{P}}_M$ , we use the same vectorization trick, namely, a function  $u$  defined on  $\mathcal{C}_N$  and sampled on  $\mathcal{P}_M$  is associated to a vector  $\mathbf{u} \in \mathbb{R}^M$  with the right correspondence between the components of  $\mathbf{u}$  and the polar coordinates in  $\mathcal{P}_M$  (and similarly for a function defined in the Fourier domain and sampled on  $\widehat{\mathcal{P}}_M$ ).

Therefore, the ODT observations  $\{\Delta(\tau, \theta) : (\tau, \theta) \in \mathcal{P}_M\}$  are gathered in a vector  $\mathbf{z} \in \mathbb{R}^M$  with  $z_j = \Delta(\tau_s, \theta_t)$  for a certain index mapping  $j = j(s, t) \in [M]$ . In this discrete representation, the equivalent transformation of (3.3) reads

$$\mathbf{y}_{\text{comp}} = (\sqrt{N_\tau} \delta\tau) \mathbf{F}^{\text{rad}} \mathbf{z}, \quad (3.6)$$

where  $\mathbf{y}_{\text{comp}} \in \mathbb{C}^M$  is associated to a (vectorized) sampling of  $y$  on  $\widehat{\mathcal{P}}_M$ , and  $\mathbf{F}^{\text{rad}} : \mathbb{R}^M \rightarrow \mathbb{C}^M$  performs a 1-D DFT on the radial  $\tau$ -variations of  $\mathbf{z}$ , *i.e.*,

$$(\mathbf{F}^{\text{rad}} \mathbf{z})_{k(s', t)} = \frac{1}{\sqrt{N_\tau}} \sum_s z_{j(s, t)} e^{-2\pi i s s' / N_\tau},$$

for a vectorized index  $k = k(s', t)$ . In other words, if  $\delta\tau$  is sufficiently small, *e.g.*, if  $\Delta(\tau, \theta)$  is band-limited with a cut-off frequency smaller than  $1/\delta\tau = N_\tau \delta\omega$ , then

$$\begin{aligned} y_{k(s', t)} &= \frac{\sqrt{N_\tau} \delta\tau}{\sqrt{N_\tau}} \sum_s \Delta(\tau_s, \theta_t) e^{-2\pi i (s \delta\tau)(s' / \delta\tau N_\tau)} \\ &= \sum_s \Delta(\tau_s, \theta_t) e^{-2\pi i \tau_s \omega_{s'}} \delta\tau \approx y(\omega_{s'}, \theta_t), \end{aligned}$$

using a Riemann sum approximation of (3.3) and knowing that  $\Delta$  vanishes on  $\Omega^c$ .

Despite the fact that  $\mathbf{y}_{\text{comp}}$  belongs to  $\mathbb{C}^M \simeq \mathbb{R}^{2M}$ , this vector brings only  $M$  independent real observations of  $\mathbf{n} \in \mathbb{R}^N$ . This is due to the central symmetry (3.5c) induced by the realness of  $\mathbf{z}$ , which allows considering  $\mathbf{y}_{\text{comp}}$  only on  $\widehat{\mathcal{P}}_M^+$ .

This is clarified by the definition of the useful operator  $\Theta : \mathbb{C}^M \rightarrow \mathbb{R}^M$ , which performs the two following linear operations. First, it restricts any vector  $\boldsymbol{\zeta} \in \mathbb{C}^M$  to the indices associated to the half grid  $\widehat{\mathcal{P}}_M^+$ . Second, it appends

the  $M/2$  imaginary values of the restricted vector to its  $M/2$  real values in order to form a vector in  $\mathbb{R}^M$ . The adjoint operation  $\Theta^*$ , which is also the inverse of  $\Theta$  for vectors in  $\mathbb{C}^M$  respecting the Hermitian symmetry, is obtained easily by first reforming a complex vector of length  $M/2$  from the separated real and imaginary parts, and by inserting the results in  $\mathbb{C}^M$  according to the indices of  $\widehat{\mathcal{P}}_M^+$  and by completing the information in  $\widehat{\mathcal{P}}_M \setminus \widehat{\mathcal{P}}_M^+$  with the central symmetry (3.5c).

Consequently, thanks to  $\Theta$ , we can form the real vector

$$\mathbf{y} = \Theta \mathbf{y}_{\text{comp}} = (\sqrt{N_\tau} \delta\tau) (\Theta \mathbf{F}^{\text{rad}}) \mathbf{z} \in \mathbb{R}^M, \quad (3.7)$$

with  $(\Theta \mathbf{F}^{\text{rad}}) : \mathbb{R}^M \rightarrow \mathbb{R}^M$ . We call  $\mathbf{y}$  the frequency deflectometric measurements (FDM) vector. This will be our direct source of ODT observations instead of  $\mathbf{z}$ .

### 3.3.3 Forward Model

We can now explain how we use the DFST relation (3.4) for defining the forward model that links any discrete 2-D RIM representation to its FDM vector.

In the work of Jacques *et al.* [122], the data available in the frequency polar grid  $\widehat{\mathcal{P}}_M$  were first interpolated to a Cartesian frequency grid in order to reconstruct the 2-D RIM using the Fast Fourier Transform (FFT). However, the polar to Cartesian frequency interpolation introduced a hardly controlled correlated distortion.

We use here another operator  $\mathbf{F} : \mathbb{R}^N \rightarrow \mathbb{C}^M$  performing a non-equispaced Discrete Fourier transform (NDFT) for directly relating functions sampled on the Cartesian spatial grid  $\mathcal{C}_N$  to those sampled on the polar frequency grid  $\widehat{\mathcal{P}}_M$ .

More precisely, given a function  $f : \Omega \rightarrow \mathbb{C}$  sampled on  $\mathcal{C}_N$ , the NDFT<sup>4</sup> computes

$$\widehat{f}(\mathbf{k}) = \sum_{m=0}^{N_0-1} \sum_{n=0}^{N_0-1} f(\mathbf{r}_{m,n}) e^{-2\pi i \mathbf{k} \cdot \mathbf{r}_{m,n}}, \quad (3.8)$$

<sup>4</sup>This NDFT formulation is strictly equivalent to the one given in [127] where  $\delta\tau = \delta r = 1$ .

on the  $M$  nodes  $\mathbf{k}$  of  $\widehat{\mathcal{P}}_M$ . Gathering all the values of  $\widehat{f}$  and  $f$  into vectors  $\widehat{\mathbf{f}} \in \mathbb{C}^M$  and  $\mathbf{f} \in \mathbb{R}^M$ , respectively, this relation is conveniently summarized as

$$\widehat{\mathbf{f}} = \mathbf{F}\mathbf{f}, \quad \text{with } F_{ij} = e^{-2\pi i \mathbf{k}_i \cdot \mathbf{r}_j}, \quad (3.9)$$

where the matrix  $\mathbf{F} \in \mathbb{C}^{M \times N}$  stands for the linear NDFT operation. Its internal entry indexing follows the one of the components of  $\mathbf{f}$  and  $\widehat{\mathbf{f}}$ . We explain in Appendix 3.B how the Non-equispaced Fast Fourier Transform (NFFT) algorithm allows us to compute efficiently in  $\mathcal{O}(N \log(N/\epsilon))$  the multiplications  $\mathbf{F}\mathbf{u}$  and  $\mathbf{F}^*\mathbf{v}$  for any  $\mathbf{u} \in \mathbb{R}^N$  and  $\mathbf{v} \in \mathbb{C}^M$ , with a controlled distortion  $\epsilon$  with respect to the true NDFT.

The action of  $\mathbf{F}$  on a discretized RIM  $\mathbf{n}$  is related to the continuous Fourier transform of  $n$  as follows. Let  $j = j(s', t) \in [M]$  be the  $j^{\text{th}}$  point  $\mathbf{k}_j$  of the grid  $\widehat{\mathcal{P}}_M$  associated to the polar coordinates  $(\omega_{s'}, \theta_t)$ . Then, for a sufficiently small  $\delta r$ , we have

$$(\mathbf{F}\mathbf{n})_{j(s', t)} = \sum_{m=0}^{N_0-1} \sum_{n=0}^{N_0-1} n(\mathbf{r}_{m,n}) e^{-2\pi i \mathbf{k}_j \cdot \mathbf{r}_{m,n}} \approx \frac{1}{(\delta r)^2} \widehat{n}(\omega_{s'} \mathbf{p}_{\theta_t}). \quad (3.10)$$

To take into account the multiplication by  $2\pi i \omega/n_r$  in (3.4) and the existence of a factor  $1/(\delta r)^2$  in the equivalence (3.10), we introduce the diagonal operator  $\mathbf{D} \in \mathbb{R}^{M \times M}$  defined as

$$\mathbf{D} = \frac{2\pi i (\delta r)^2}{n_r} \text{diag}(\omega_{(1)}, \dots, \omega_{(M)}),$$

where  $\omega_{(j)}$  refers to the  $\omega$ -coordinate of the  $j^{\text{th}}$  point of  $\widehat{\mathcal{P}}_M$ , *i.e.*, if as before  $j = j(s', t)$  is associated to  $(\omega_{s'}, \theta_t) \in \widehat{\mathcal{P}}_M$ , then  $\omega_{(j)} = \omega_{s'}$ . The operator  $\mathbf{D}$  models the effect of the transverse gradient in the Fourier domain.

In parallel to the discussion in Section 3.3.2, we also restrict the action of  $\mathbf{D}\mathbf{F}$  to the domain  $\widehat{\mathcal{P}}_M^+$ . Consequently, using the operator  $\Theta$  (see Section 3.3.2), the final linear forward model linking the real FDM to the 2-D NDFT of the discrete RIM  $\mathbf{n}$  reads

$$\mathbf{y} = \Phi \mathbf{n} + \boldsymbol{\eta} \in \mathbb{R}^M, \quad (3.11)$$

with  $\Phi = \Theta \mathbf{D}\mathbf{F} \in \mathbb{R}^{M \times N}$  the ODT sensing operator. The additional noise  $\boldsymbol{\eta} \in \mathbb{R}^M$  integrates the different distortions explained and estimated in Section 3.6, *i.e.*, the numerical computations, the model discretization, the discrepancy to the first order approximation (3.2), and the actual noise corrupting the observation of  $\mathbf{z}$ .

*Remark 3.5.* Notice that, in the absence of noise ( $\boldsymbol{\eta} = 0$ ), the model (3.11) could be easily turned into a classical tomographic model where the zero frequency is not observed. Indeed, forgetting the formal action of  $\Theta$ , if the frequency origin has been carefully removed from  $\widehat{\mathcal{P}}_M$ , then  $\mathbf{D}$  is invertible with  $\mathbf{D}^{-1} = -\frac{i n_r}{2\pi(\delta r)^2} \text{diag}(\omega_{(1)}^{-1}, \dots, \omega_{(M)}^{-1})$ , and we can solve the common tomographic problem

$$\bar{\mathbf{y}} := \mathbf{D}^{-1} \mathbf{y} = \mathbf{F} \mathbf{n}. \quad (3.12)$$

However, as we present in Section 3.8.1.1, this transformation is not suited to noisy ODT sensing. Even for a simple additive white Gaussian noise (AWGN)  $\boldsymbol{\eta}$ , the multiplication by  $\mathbf{D}^{-1}$  breaks the homoscedasticity of  $\boldsymbol{\eta}$ , *i.e.*, the variance of each  $(\mathbf{D}^{-1} \boldsymbol{\eta})_j$  varies with  $j \in [M]$ . This interferes with common reconstruction techniques used in classical tomography. Obviously, a noise whitening could be realized for stabilizing such methods but at least for AWGN, this strictly amounts to solve directly the model (3.11).

### 3.4 Related Works

This section describes some recently proposed methods for tomographic reconstruction in the domains of differential phase-contrast tomography and common absorption tomography.

In differential phase-contrast tomography, the refractive index distribution is recovered from phase-shifts measurements. These are composed by the derivative of the refractive index map, inducing the apparition of the *affine* frequency  $\omega$  when using the FST, as it happens in the ODT sensing model (see Remark 3.1). In this application, Pfeiffer *et al.* [1] have used the FBP algorithm to reconstruct the refractive index map from a fully covered set of projections. Cong *et al.* [75, 121] have used different iterative schemes based on the minimization of the TV norm to reconstruct the refractive index distribution over a region of interest. These methods are accurate and provide similar results, but the iterative scheme based on the TV norm has proved to be better than FBP when the amount of acquisitions decreases.

In common absorption tomography (AT), we deal with the reconstruction of the absorption index distribution from intensity measurements. As these measurements are directly related to the absorption index, the AT sensing

model does not include the *affine* frequency  $\omega$  (see Remarks 3.3 and 3.5). In this domain, several works have exploited sparsity based methods. Most recent works in AT have focused on promoting a small TV norm [4,23,120].

Sidky *et al.* [23] and Anthoine *et al.* [4] use a Lagrangian formulation for the tomographic reconstruction problem, promoting a small TV norm under a Kullback-Leiber data divergence (adapted to Poisson distributed noise) and a non-negativity constraint. In [23], the authors aim at reconstructing a breast phantom from 60 projections acquired using fan-beam scanning. For this, they use the Chambolle-Pock primal-dual algorithm described in Section 2.4.2.2. In [4], the authors are interested in the characterization of the abdominal region of a mouse, where soft tissues and spine bones are visible. The AT image is reconstructed from 90, 60 and 36 projections acquired using cone-beam scanning. The reconstruction was done with an accelerated version of the forward-backward algorithm described in Section 2.4.2.1. In both works [4,23], the proposed TV regularized methods result in a high quality reconstruction compared to FBP but with a convergence result that is highly dependent on the Lagrangian parameter chosen.

Ritschl *et al.* [120] use a constrained optimization formulation to reconstruct the absorption index from low amount of clinical data in the presence of metal implants and Gaussian noise. This problem is solved by means of an alternating method that allows optimizing separately the raw data consistency function and the sparsity cost function, without the need of prior information on the observations. The fast convergence of the method is based on the estimation of the optimization steps. The gradient descent method is used to minimize the TV norm and the consistency term is minimized via an algebraic reconstruction technique. The method demonstrated to provide better results than FBP.

These works have proved that some tomographic applications can be made compressive in the compressed sensing sense [118,119]. In short, accurate tomographic image reconstruction are reachable from a number of samples that is smaller than the desired image resolution, if the image is sparse in some basis, *i.e.*, the image expansion in that basis contains only a small number of non-zero coefficients. However, most of these works have considered Cartesian frequency grids while actual sensing often occurs in polar

or non-equispaced grids. Besides, they attack different problems than ODT, where the sensing model and type of measurement change from one to the other. One important difference between AT and ODT sensing models, relies on the presence of the *affine* frequency as materialized by a diagonal operator  $\mathbf{D}$  (see Remarks 3.3 and 3.5); whose impact is analyzed in detail in Sections 3.6.2 and 3.8.1.1 for perfect and noisy sensing scenarios.

## 3.5 Refractive Index Map Reconstruction

Three methods can be considered for recovering the discrete RIM  $\mathbf{n} \in \mathbb{R}^N$ : the common Filtered Back Projection (FBP); a penalized least square solution, which is related to a *minimal energy solution* (ME) [128]; and a regularized approach called TV- $\ell_2$  minimization. Since FBP and ME are widely used in tomographic problems, we will use them as a standard to compare with the quality of TV- $\ell_2$ .

### 3.5.1 Filtered Back Projection

The Filtered Back Projection (FBP) can be briefly defined as an analytical method that consists of first filtering the tomographic projections with an appropriate function, *i.e.*, a “ramp filter” for absorption tomography (AT) [22, Section 3.3] or a simple Hilbert transform for deflection tomography [1, 116]. The result is then *back projected* in the spatial domain by angular integration. Despite its simplicity, this technique suffers of severe artifacts when the number of angular observations decreases. Moreover, FBP does not integrate the noise distortion in its processing.

### 3.5.2 Minimum Energy Reconstruction

Consider the ODT sensing model in (3.11) for  $\eta = 0$ , *i.e.*,

$$\mathbf{y} = \Phi \mathbf{n} \in \mathbb{R}^M, \quad (3.13)$$

where  $\Phi = \Theta \mathbf{D} \mathbf{F} \in \mathbb{R}^{M \times N}$  is the ODT sensing operator with  $M \leq N$ . A common procedure to estimate  $\mathbf{n}$  consists of applying the (right) pseudo-inverse of

$\Phi$  to the observations, *i.e.*, computing  $\Phi^\dagger \mathbf{y} = \Phi^* (\Phi \Phi^*)^{-1} \mathbf{y}$  (for non-singular  $\Phi \Phi^*$ ).

This solution is actually equivalent to compute a *minimum energy* solution (or *penalized least square*)  $\tilde{\mathbf{n}}_{\text{ME}}$  by solving the convex problem

$$\tilde{\mathbf{n}}_{\text{ME}} = \underset{\mathbf{u} \in \mathbb{R}^N}{\operatorname{argmin}} \|\mathbf{u}\|_2 \text{ st. } \mathbf{y} = \Phi \mathbf{u}. \quad (3.14)$$

In other words, the inverse problem (3.13) is solved by a *regularization* promoting a small  $\ell_2$  norm (*i.e.*, a small *energy*). For large scale problems, solving this last convex formulation is preferred to the estimation of the pseudo-inverse that requires the costly inversion of an  $M \times M$  matrix.

Common reconstruction approaches follow this minimum energy principle [128]. They generally proceed by zeroing the Fourier coefficients of all unobserved frequencies, *i.e.*, a process that minimizes the energy of the solution for frequency sampling defined on regular grids. The ME solution is also close to a discretization of the FBP procedure for a densely covered frequency space.

As shown later, the ME method presents some strong limitations when the model (3.13) is severely ill-conditioned or when noise corrupts the observation of  $\mathbf{y}$ . For instance, in the case where the tomographic problem subsamples a regular sampling of the Fourier domain, artifacts and blurring appear from the convolution of the pure image with the filter associated to the subsampling pattern (or *dirty map* [50]).

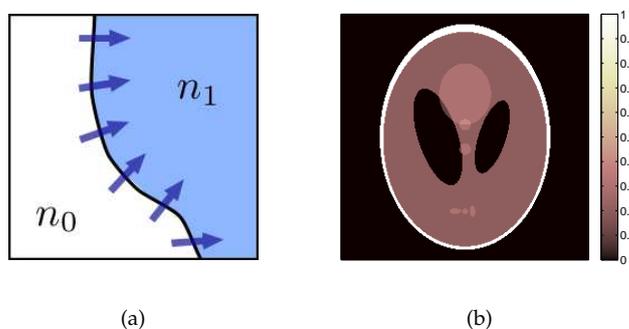
### 3.5.3 TV- $\ell_2$ Minimization

In order to overcome the limitations of both FBP and ME methods, we introduce a new reconstruction method which is both less sensitive to unwanted oscillations due to a low density frequency sampling  $\hat{\mathcal{P}}_M$  and to additional observational noise  $\boldsymbol{\eta}$  in (3.11).

In particular, since the spatial dimensions in  $\mathcal{P}_M$  and in  $\mathcal{C}_N$  are expected to be equal, *i.e.*,  $N_0 \approx N_\tau$  (see Section 3.3.1) we are interested in lowering the density of  $\hat{\mathcal{P}}_M$  in the Fourier plane by decreasing the number of angular observations  $N_\theta$ . In other words, with respect to this reduction, we aim at developing

a numerical reconstruction which makes optical deflectometric tomography “compressive” in a similar way other compressive imaging devices which, inspired by the compressed sensing paradigm [118, 119], reconstruct high resolution images from few (indirect) observations of its content [2, 23, 50, 129]. This ability would lead of course to a significant reduction of the ODT observation time with potential impact, for instance, in fast industrial object quality control relying on this technology.

This objective of compressiveness can only be reached by regularizing the ODT inverse problem by an appropriate prior model on the configuration of the expected RIM  $\mathbf{n}$ . Interestingly, the actual RIM of most human made transparent materials (*e.g.*, optical fiber bundles, lenses, ..) is composed by slowly varying areas separated by sharp boundaries (material interfaces) (see Figure 3.2-(a)). This can be interpreted with a bounded-variation (BV) or “cartoon-shape” model [73] as the typical Shepp-Logan phantom in Figure 3.2-(b). Therefore, the inverse problem in (3.11) can be regularized by promoting a small Total-Variation (TV) norm (see Section 2.3.2.3). In this work, we use the *isotropic* TV norm defined in (2.18).



**Figure 3.2** (a) RIM TV Model: the gradient of the RIM (represented by arrows) is non-zero only on the interface between the two RIM homogeneous areas  $n_0$  and  $n_1$ . This induces a small TV norm. (b) The Shepp-Logan phantom, an example of the “cartoon-shape” model.

In order to obtain a reconstruction method which is also robust to additive observation noise, we must lighten the strict fidelity constraint implicitly used by the ME method in (3.14). Therefore, assuming the data corrupted

by an additive white Gaussian noise (AWGN), we impose a data fidelity requirement using the  $\ell_2$  norm, *i.e.*, if  $\boldsymbol{\eta} = \mathbf{y} - \Phi \mathbf{n}$  is known to have a bounded norm (or *energy*)  $\|\boldsymbol{\eta}\|_2 \leq \varepsilon$ , we force any reconstruction candidate  $\mathbf{u}$  to satisfy  $\|\mathbf{y} - \Phi \mathbf{u}\|_2 \leq \varepsilon$ . The estimation of  $\varepsilon$  will be addressed in Section 3.6.

Additionally to a fidelity criterion with the observed data, other requirements can be imposed on the reconstruction. First, we can often assume that the reference refractive index  $n_r$  (*i.e.*, the one of some optical fluid surrounding the object) is lower than the object RIM. Second, if the object is completely contained in the field-of-view  $\Omega$  of the ODT experiment, we can force any candidate RIM  $\mathbf{u}$  to match  $n_r$  on the boundary of  $\Omega$ , *i.e.*, imposing  $u_k = n_r$  for all indices  $k$  belonging to the border of  $\mathcal{C}_N$ . These indices are associated to pixels  $\mathbf{r}_{m,n} \in \mathcal{C}_N$  for which at least one of the two 2-D coordinates is equal to either  $-N_0/2$  or  $N_0/2 - 1$ . The corresponding index set is denoted  $\partial\Omega$  for simplicity.

Gathering all these aspects, we could propose the following reconstruction program

$$\tilde{\mathbf{n}}_{\text{TV-}\ell_2} = \underset{\mathbf{u} \in \mathbb{R}^N}{\text{argmin}} \|\mathbf{u}\|_{\text{TV}} \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{u}\|_2 \leq \varepsilon, \mathbf{u} \succeq \mathbf{n}_r, \mathbf{u}_{\partial\Omega} = n_r(\mathbf{1}_N)_{\partial\Omega},$$

denoting by  $\mathbf{1}_N \in \mathbb{R}^N$  the vector of ones, *i.e.*, the unit RIM in  $\mathcal{C}_N$ , and recalling that  $\mathbf{v}_{\mathcal{A}} = \mathbf{R}_{\mathcal{A}} \mathbf{v}$  stands for the restriction of the components of  $\mathbf{v} \in \mathbb{R}^N$  to  $\mathcal{A} \subset [N]$ .

The reconstruction can be slightly simplified by observing that the kernel of the sensing operator  $\Phi = \Theta \mathbf{D} \mathbf{F}$  in ODT contains the set of constant vectors in  $\mathbb{R}^N$ . This is a consequence of the vanishing affine frequency  $\omega$  (which mainly defines the action of  $\mathbf{D}$ ) on the frequency origin, or more simply, this relies on the occurrence of the RIM gradient in the deflection model (3.2).

Therefore, a change of variable  $\mathbf{u} \rightarrow \mathbf{u} - n_r \mathbf{1}_N$  does not disturb the previous reconstruction, which can be recast as

$$\tilde{\mathbf{n}}_{\text{TV-}\ell_2} = \underset{\mathbf{u} \in \mathbb{R}^N}{\text{argmin}} \|\mathbf{u}\|_{\text{TV}} \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{u}\|_2 \leq \varepsilon, \mathbf{u} \succeq \mathbf{0}, \mathbf{u}_{\partial\Omega} = \mathbf{0}_{|\partial\Omega|}, \quad (3.15)$$

remembering that the true RIM estimation is actually  $\tilde{\mathbf{n}}_{\text{TV-}\ell_2} + n_r \mathbf{1}_N$ .

Without the *frontier constraint* ( $\mathbf{u}_{\partial\Omega} = \mathbf{0}_{|\partial\Omega|}$ ), the uniqueness of the solution is not guaranteed. In that case, for one minimizer  $\tilde{\mathbf{u}}$ , all the vectors of

$\{\tilde{\mathbf{u}} + \lambda \mathbf{1}_N : \lambda \in \mathbb{R}, \tilde{\mathbf{u}} + \lambda \mathbf{1}_N \succeq 0\}$  also minimize the problem since the kernels of both the sensing operator  $\Phi$  and the gradient operator  $\nabla$  contain constant vectors (i.e.,  $\text{Ker } \Phi \cap \text{Ker } \nabla = \{\lambda \mathbf{1}_N\}$ ). Considering the frontier constraint is thus essential for enforcing the uniqueness of the solution as expressed in the following lemma.

*Lemma 3.1.* If there is at least one feasible point for the constraints of (3.15), then the solution of this problem is unique.

**Proof.** Using the TV norm definition in (2.18) and squaring it, the TV- $\ell_2$  optimization (3.15) is equivalent to solve

$$\operatorname{argmin}_{\mathbf{u} \in \mathbb{R}^N} \|\nabla \mathbf{u}\|_{2,1}^2 \text{ s.t. } \|\mathbf{y} - \Phi \mathbf{u}\|_2 \leq \varepsilon, \mathbf{u} \succeq 0, \mathbf{u}_{\partial\Omega} = \mathbf{0}_{|\partial\Omega|}.$$

The kernel of  $\nabla$  is the set of constant vectors, while the one of  $\mathbf{R}_{\partial\Omega}$  (defining the frontier constraint) is the set of vectors equal to 0 on  $\partial\Omega$ . Therefore,  $\text{Ker } \nabla \cap \text{Ker } \mathbf{R}_{\partial\Omega} = \{\mathbf{0}_N\}$ . Moreover, since the domain of  $\|\nabla \cdot\|_{2,1}^2$  is  $\mathbb{R}^N$ , and since we assume at least one feasible point, (3.15) has at least one solution.

Let  $\mathbf{x}_1$  and  $\mathbf{x}_2$  be two distinct minimizers. Then,  $\mathbf{R}_{\partial\Omega} \mathbf{x}_1 = \mathbf{R}_{\partial\Omega} \mathbf{x}_2 = 0$  and  $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Ker } \mathbf{R}_{\partial\Omega}$  while  $\mathbf{x}_1 - \mathbf{x}_2 \neq \mathbf{0}_N$ . Therefore,  $\mathbf{x}_1 - \mathbf{x}_2 \notin \text{Ker } \nabla$  and  $\nabla \mathbf{x}_1 \neq \nabla \mathbf{x}_2$ . By the strict convexity of the function  $\varphi(\cdot) = \|\cdot\|_{2,1}^2$ , writing  $\mathbf{x}_\lambda = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$  for  $\lambda \in (0, 1)$ ,  $\varphi(\nabla \mathbf{x}_1) = \varphi(\nabla \mathbf{x}_2)$  involves

$$\varphi(\nabla \mathbf{x}_1) = \lambda \varphi(\nabla \mathbf{x}_1) + (1 - \lambda) \varphi(\nabla \mathbf{x}_2) > \varphi(\nabla \mathbf{x}_\lambda),$$

showing that  $\mathbf{x}_\lambda$ , which also satisfies the convex constraints, is a better minimizer, which is a contradiction.  $\square$

Therefore, we see that the uniqueness is actually reached by the stabilizing condition  $\mathbf{u}_{\partial\Omega} = \mathbf{R}_{\partial\Omega} \mathbf{u} = \mathbf{0}_{|\partial\Omega|}$ , which makes the optimization running outside of  $\text{Ker } \nabla \setminus \{\mathbf{0}_N\}$ .

As explained in Section 3.7, the program (3.15) can be efficiently solved using proximal methods [80] and operator splitting techniques, like the Chambolle-Pock primal-dual algorithm described in Section 2.4.2.2.

## 3.6 Noise Identification, Estimation and Analysis

In this section, we first discuss about the different sources of noise and how to estimate the noise energy present in the experimental data. Then, we analyze the noise impact in both AT and ODT measurements.

### 3.6.1 Noise Sources

When a real sensing scenario is being studied, such as the ODT, different sources of noise are present and they have to be considered when determining the global noise energy bound  $\varepsilon$  in (3.15).

First, we have the *observation noise*. Under high light source intensity, the images collected by a Schlieren deflectometer (Figure 3.1-(b)), and used for computing the ODT deflections  $\mathbf{z}$ , are mainly affected by electronic noise such as the CCD thermal noise. This induces a noise in the measured deflection angles that can reasonably be assumed Gaussian and homoscedastic, *i.e.*, with an homogeneous variance through all the measurements. By computing the 1-D Fourier transform of the ODT measurements using  $\Theta \mathbf{F}_{\text{rad}}$  in (3.7) the corresponding noise in the FDM  $\mathbf{y}$ , denoted  $\boldsymbol{\eta}_{\text{obs}}$ , remains Gaussian [130], *i.e.*,  $(\boldsymbol{\eta}_{\text{obs}})_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma_{\text{obs}}^2)$ . Actually, from the orthonormality of the Fourier basis, (3.7) provides  $\sigma_{\text{obs}}^2 = (\delta\tau)^2 N_{\tau} \sigma_z^2$ , where  $\sigma_z^2$  is the variance of the noise present in the ODT measurements. This one can be estimated from the noisy observations using the Robust Median Estimator described in Section 2.A.

Finally, this noise, defined as the difference between the noisy FDM  $\mathbf{y}$  and the noiseless FDM  $\mathbf{y}_{\text{true}}$ , has an energy that can be bounded using the Chernoff-Hoeffding bound in (2.10):

$$\|\mathbf{y} - \mathbf{y}_{\text{true}}\|_2^2 = \|\boldsymbol{\eta}_{\text{obs}}\|_2^2 < \varepsilon_{\text{obs}}^2 := \sigma_{\text{obs}}^2 \left( M + c\sqrt{M} \right),$$

with high probability for  $c = \mathcal{O}(1)$ .

Second, there is the *modeling error* that comes with every mathematical discrete representation of a physical continuous system. In the ODT system, this error is due to (i) the straight line approximation used to formulate (3.2), (ii) the sampling of the continuous RIM, and (iii) the discrete model itself. The *modeling* noise is related to the difference between the noiseless FDM and

the sensing model  $\Phi_{\text{true}}\mathbf{n} = \Theta\mathbf{D}\mathbf{F}_{\text{true}}\mathbf{n}$ , where  $\mathbf{F}_{\text{true}}$  performs the exact polar Fourier transform. As explained in Section 3.2.1, the *modeling* noise can be estimated by means of a *ray tracing* method based on Snell's law (see Appendix 3.A.3). This shows that, for simple objects, an absolute error of  $0.7^\circ$  is expected on light deflections smaller than  $7^\circ$ . A Gaussian noise model provides a rough estimation of 10 dB for the corresponding measurement SNR. This is equivalent to:

$$\|\mathbf{y}_{\text{true}} - \Phi_{\text{true}}\mathbf{n}\|_2 = \|\boldsymbol{\eta}_{\text{model}}\|_2 < \varepsilon_{\text{model}} = \frac{\|\mathbf{y}_{\text{true}}\|_2}{\sqrt{10}} \simeq \frac{\|\mathbf{y}\|_2}{\sqrt{10}}. \quad (3.16)$$

As it will be described in Section 3.8.1.4, the *modeling* error is sparse in the  $\tau$ -domain, which suggests that a better bound could be found using the  $\ell_1$  norm. Also, by means of the *ray tracing* method we could identify the few values of  $\tau$  where this error is significant so they are removed from the model. This would decrease the error in (3.16), yielding a stronger fidelity term. Nonetheless, remark that this information is only available in the  $\tau$ -domain and the measurements are in the Fourier domain. This is a matter of future study.

Third, we must consider the *interpolation noise* given by the mathematical error committed by estimating the polar Fourier transform with the NFFT algorithm, *i.e.*, the noise  $\Phi_{\text{true}}\mathbf{n} - \Phi\mathbf{n}$ . To determine a bound  $\varepsilon_{\text{nfft}}$  on the energy of this error we first estimate the NFFT distortion (*i.e.*, without the action of  $\mathbf{D}$ ), defined as the difference between the NFFT polar Fourier transform  $\mathbf{F}_{\text{app}}$  and the true NDFT  $\mathbf{F}$ . Theoretically, for any vector  $\mathbf{f} \in \mathbb{R}^N$ , the  $\ell_\infty$  norm of this distortion is bounded as  $\|\mathbf{F}_{\text{app}}\mathbf{f} - \mathbf{F}\mathbf{f}\|_\infty \leq \varepsilon C(\mathbf{f}) = \mathcal{O}(\varepsilon \|\mathbf{f}\|_1)$ , where  $\varepsilon$  controls both the accuracy and the complexity  $\mathcal{O}(N \log N / \varepsilon)$  of the NFFT [127] (see Appendix 3.B). Assuming that each component of  $\mathbf{F}_{\text{app}}\mathbf{f} - \mathbf{F}\mathbf{f}$  is i.i.d. with a uniform distribution  $\mathcal{U}([-C(\mathbf{f}), C(\mathbf{f})])$ , and using the Chernoff-Hoeffding bound [56, 131], we can estimate  $\|\mathbf{F}_{\text{app}}\mathbf{f} - \mathbf{F}\mathbf{f}\|_2^2 < \frac{C(\mathbf{f})^2}{3} (M + c\sqrt{M})$ , with high probability for  $c = \mathcal{O}(1)$ . Finally,  $\varepsilon_{\text{nfft}}$  can be crudely computed as

$$\begin{aligned} \|\Phi_{\text{true}}\mathbf{n} - \Phi\mathbf{n}\|_2 &= \|\Theta\mathbf{D}(\mathbf{F}_{\text{app}}\mathbf{n} - \mathbf{F}\mathbf{n})\|_2 \leq \|\Theta\mathbf{D}\| \|\mathbf{F}_{\text{app}}\mathbf{n} - \mathbf{F}\mathbf{n}\|_2 \\ &= \frac{2\pi(\delta r)^2 \omega_{\text{max}}}{n_r} \frac{\varepsilon C(\mathbf{n})}{\sqrt{3}} (M + c\sqrt{M})^{1/2} \\ &\approx \frac{\pi\delta r}{\sqrt{3}n_r} \varepsilon C(\mathbf{n}) (M + c\sqrt{M})^{1/2} =: \varepsilon_{\text{nfft}}, \end{aligned}$$

with  $\omega_{\text{max}} \approx \frac{1}{2}N\tau\delta\omega = \frac{1}{2\delta\tau} \approx \frac{1}{2\delta r}$  representing the maximum frequency amplitude in  $\widehat{\mathcal{P}}_M$ . In practice, because of the RIM shift  $\mathbf{n} \rightarrow \mathbf{n} - n_r\mathbf{1}_N$  explained

in Section 3.5.3, we can bound  $\|\mathbf{n}\|_1$  and hence  $C(\mathbf{n})$  with the expected RIM dynamics  $\delta n$ , *i.e.*,  $\|\mathbf{n}\|_1 \leq N\delta n$ , and we adjust  $\epsilon$  in order to have  $\epsilon_{\text{nfft}}$  much lower than the other sources of noises.

Finally, we may also have an error introduced by the *instrument calibration*, when determining the exact  $\tau$  and  $\theta$  associated to the projections. We are going to neglect this error by assuming a pre-calibration process that provides an exact knowledge of these values (see Section 3.8.2).

In conclusion, gathering all the noises sources previously identified and assuming the different noise are of zero mean and independent of each other, we can bound the difference between the actual ODT measurement and the sensing model as follows:

$$\begin{aligned} \|\mathbf{y} - \Phi\mathbf{n}\|_2^2 &= \|\mathbf{y} - \mathbf{y}_{\text{true}} + \mathbf{y}_{\text{true}} - \Phi_{\text{true}}\mathbf{n} + \Phi_{\text{true}}\mathbf{n} - \Phi\mathbf{n}\|_2^2 \\ &= \|\eta_{\text{obs}} + \eta_{\text{model}} + \eta_{\text{nfft}}\|_2^2 < \epsilon_{\text{obs}}^2 + \epsilon_{\text{model}}^2 + \epsilon_{\text{nfft}}^2. \end{aligned}$$

Therefore, we have  $\epsilon \approx \sqrt{\epsilon_{\text{obs}}^2 + \epsilon_{\text{model}}^2 + \epsilon_{\text{nfft}}^2}$ .

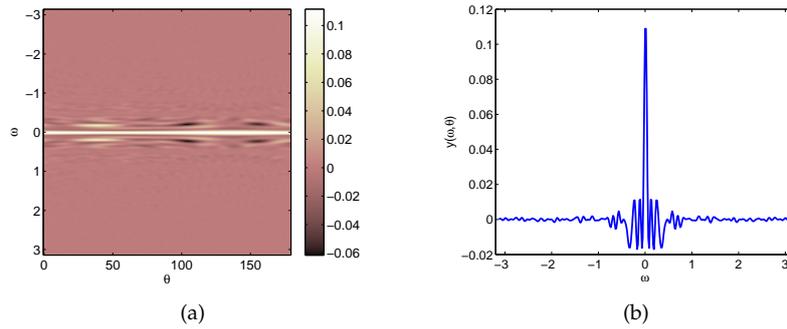
### 3.6.2 Comparative Study of the Noise Impact on AT and ODT

As discussed in Remarks 3.3 and 3.5, the main difference between the AT and ODT problems is the appearance of the diagonal operator  $\mathbf{D}$  in the last one. Therefore, the AT sensing operator is described as  $\Phi_{\text{AT}} = \Theta\mathbf{F}$  and the ODT sensing operator as  $\Phi_{\text{ODT}} = \Theta\mathbf{D}\mathbf{F}$ . We will now analyze the impact of an additive white Gaussian noise on the measurements regarding the presence of this operator.

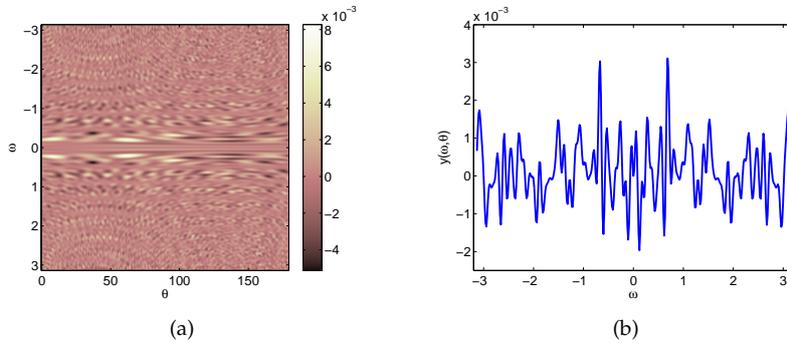
For this, we apply the sensing operators  $\Phi_{\text{AT}}$  and  $\Phi_{\text{ODT}}$  to a section of the fibers bundle (see Figure 3.5-(a)), in order to obtain the AT and ODT acquisitions, respectively. In Figures 3.3 and 3.4, we show the real part of the Fourier measurements in AT and ODT, respectively.

For the class of images we are interested in, we can notice that in AT the magnitude presents a peak around  $\omega = 0$  and then decreases significantly when the distance to the center increases, tending to zero in the borders (see Figure 3.3). Whereas in ODT, the presence of operator  $\mathbf{D}$  makes the image intensity to be quite spread through all the pixels (see Figure 3.4). This has

a direct impact on the reconstruction when the measurement is affected by additive Gaussian noise. As the noise spreads evenly through the image, the pixels that are not around  $\omega = 0$  will be more affected in the AT model because their intensity is significantly lower.



**Figure 3.3** AT Measurement (in Fourier) (a) on the whole grid  $\hat{\mathcal{P}}_M$  for  $N_\theta = 180$  and (b) on the slice  $\theta = 80^\circ$ .



**Figure 3.4** ODT Measurement (in Fourier) (a) on the whole grid  $\hat{\mathcal{P}}_M$  for  $N_\theta = 180$  and (b) on the slice  $\theta = 80^\circ$ .

### 3.7 Numerical Methods

This section describes the numerical methods used in this work to solve the ODT problem under multiple constraints and how to adaptively select the optimization parameters in order to obtain a faster convergence.

### 3.7.1 ODT Numerical Reconstruction

The TV- $\ell_2$  problem in (3.15) consists of the minimization of a non-smooth convex function and three constraints. The constraints can be handled by adding the convex indicator functions on the following sets:  $\mathcal{P}_0 = \{\mathbf{u} \in \mathbb{R}^N : u_i \geq 0 \text{ if } i \in \text{int } \Omega; u_i = 0 \text{ if } i \in \partial\Omega\}$  and  $\mathcal{C} = \{\mathbf{v} \in \mathbb{R}^M : \|\mathbf{y} - \mathbf{v}\|_2 \leq \varepsilon\}$ . The ODT problem can then be reformulated as

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\nabla \mathbf{x}\|_{2,1} + \iota_{\mathcal{C}}(\Phi \mathbf{x}) + \iota_{\mathcal{P}_0}(\mathbf{x}). \quad (3.17)$$

The resulting optimization, containing the sum of three functions belonging to  $\Gamma_0(\mathbb{R}^D)$ , for  $D \in \{2N, M, N\}$ , can be solved using the expanded Chambolle-Pock (CP) algorithm for  $L = 2$  in (2.25) (see Section 2.4.2.2).

We can easily show that (3.17) has the shape of (2.25) with  $F_1(\mathbf{s}_1) = \|\mathbf{s}_1\|_{2,1}$  for  $\mathbf{s}_1 \in \mathbb{R}^{N \times 2} \simeq \mathbb{R}^{2N}$ ,  $F_2(\mathbf{s}_2) = \iota_{\mathcal{C}}(\mathbf{s}_2)$  for  $\mathbf{s}_2 \in \mathbb{R}^M$ ,  $H(\mathbf{x}) = \iota_{\mathcal{P}_0}(\mathbf{x})$ ,  $\mathbf{K}_1 = \nabla \in \mathbb{R}^{2N \times N}$  and  $\mathbf{K}_2 = \Phi = \Theta \mathbf{D} \mathbf{F} \in \mathbb{R}^{M \times N}$ . We have then  $\mathbf{K} = \text{diag}(\mathbf{K}_1, \mathbf{K}_2) = \text{diag}(\nabla, \Phi) \in \mathbb{R}^{(2N+M) \times N}$ .

In order to build the dual problem as in (2.36), we need the conjugate functions of  $F_1, F_2$  and  $H$ , which are straightforwardly computed using the Legendre transform in (2.20). As a matter of fact,  $F_1^*$  is the indicator function onto the convex set  $\mathcal{Q} = \{\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2) \in \mathbb{R}^{N \times 2} : \|\mathbf{q}\|_{2,\infty} \leq 1\}$  with the mixed  $\ell_\infty \ell_2$  norm defined as  $\|\mathbf{q}\|_{2,\infty} = \max_k \sqrt{(\mathbf{q}_1)_k^2 + (\mathbf{q}_2)_k^2}$  [74]. The conjugate function  $F_2^*$  is computed as (see Appendix 3.C):

$$F_2^*(\mathbf{s}_2) = \iota_{\mathcal{C}}^*(\mathbf{s}_2) = (\mathbf{s}_2)^\top \mathbf{y} + \varepsilon \|\mathbf{s}_2\|_2,$$

while the convex conjugate of  $H$  is simply  $H^*(\mathbf{u}) = \iota_{\mathcal{D}}(-\mathbf{u})$ , where  $\mathcal{D} = \{\mathbf{u} \in \mathbb{R}^N : u_i \geq 0 \text{ if } i \in \text{int } \Omega\}$ .

The dual optimization problem is thus defined as:

$$\max_{\mathbf{s} \in \mathbb{R}^{2N+M}} -\iota_{\mathcal{Q}}(\mathbf{s}_1) - \langle \mathbf{s}_2, \mathbf{y} \rangle - \varepsilon \|\mathbf{s}_2\|_2 - \iota_{\mathcal{D}}(\nabla^* \mathbf{s}_1 + \Phi^* \mathbf{s}_2).$$

In order to apply the iterations in (2.37), we must compute the proximal operators of  $F_1^*, F_2^*$  and  $H$ . The one of  $F_1^*$  is given by [74]

$$\left( \text{prox}_{\nu F_1^*} \boldsymbol{\zeta} \right)_k = \frac{((\boldsymbol{\zeta}_1)_k, (\boldsymbol{\zeta}_2)_k)}{\max(1, \sqrt{(\boldsymbol{\zeta}_1)_k^2 + (\boldsymbol{\zeta}_2)_k^2})}, \quad \boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) \in \mathbb{R}^{N \times 2}.$$

The proximal operator of  $F_2^*$  is determined via the proximal operator of  $F_2$  by means of the conjugation property defined in [80]:

$$\text{prox}_{\nu F_2^*} \mathbf{s}_2 = \mathbf{s}_2 - \nu \text{prox}_{\frac{1}{\nu} F_2} \frac{1}{\nu} \mathbf{s}_2.$$

The proximal operator of  $F_2$  is given by the projection onto the convex set  $\mathcal{C}$ :

$$\text{prox}_{\frac{1}{\nu} F_2} \mathbf{s}_2 = \mathbf{y} + (\mathbf{s}_2 - \mathbf{y}) \min \left( 1, \frac{\varepsilon}{\|\mathbf{s}_2 - \mathbf{y}\|_2} \right).$$

The proximal operator of the function  $H$  is given by the projection onto the positive orthant with zero borders:

$$\text{prox}_{\frac{\mu}{2} H} \boldsymbol{\zeta} = \text{proj}_{\mathcal{P}_0} \boldsymbol{\zeta} = \begin{cases} (\zeta_i)_+ & \text{if } i \in \text{int } \Omega, \\ 0 & \text{if } i \in \partial \Omega, \end{cases}, \quad \boldsymbol{\zeta} \in \mathbb{R}^N.$$

Finally, making use of the above computations, the expanded CP algorithm in (2.37) applied to our TV- $\ell_2$  problem in ODT can be reduced to:

$$\begin{cases} \mathbf{s}_1^{(k+1)} &= \text{prox}_{\nu F_1^*} \left( \mathbf{s}_1^{(k)} + \nu \nabla \bar{\mathbf{x}}^{(k)} \right) \\ \mathbf{s}_2^{(k+1)} &= \text{prox}_{\nu F_2^*} \left( \mathbf{s}_2^{(k)} + \nu \Phi \bar{\mathbf{x}}^{(k)} \right) \\ \mathbf{x}^{(k+1)} &= \text{proj}_{\mathcal{P}_0} \left( \mathbf{x}^{(k)} - \frac{\mu}{2} \left( \nabla^* \mathbf{s}_1^{(k+1)} + \Phi^* \mathbf{s}_2^{(k+1)} \right) \right) \\ \bar{\mathbf{x}}^{(k+1)} &= 2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \end{cases}. \quad (3.18)$$

The estimated RIM is provided by the vector  $\mathbf{x}^{(k+1)}$ . The step-size parameters  $\mu$  and  $\nu$  in (3.18) are defined such that  $\mu\nu \|\mathbf{K}\|^2 < 1$  (see Section 3.7.2), where  $\|\mathbf{K}\|$  is the induced norm of the operator, estimated using the standard power iteration [23].

In our experiments, the variables  $\bar{\mathbf{x}}^{(1)}$ ,  $\mathbf{s}_1^{(1)}$  and  $\mathbf{s}_2^{(1)}$  are initialized to zero vectors and the variable  $\mathbf{x}^{(1)}$  is initialized with the FBP solution as  $\mathbf{x}^{(1)} = \tilde{\mathbf{n}}_{\text{FBP}}$ .

The algorithm presented in (3.18) stops when it achieves a stable behavior, *i.e.*, when  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{(k)}\|_2 \leq \text{Th}$ . The threshold  $\text{Th}$  is defined for ODT in the next section. In parallel, we analyze the convergence of the algorithm by means of the primal and dual residuals in (2.38) and using the energy condition in (2.34). For the iterates in (3.18), the primal and dual residuals are

defined as:

$$\begin{aligned} \mathbf{p}^{(k+1)} &= \frac{2}{\mu} \left( \mathbf{x}^{(k)} - \mathbf{x}^{(k+1)} \right), \\ \mathbf{d}^{(k+1)} &= \begin{pmatrix} \mathbf{d}_1^{(k+1)} \\ \mathbf{d}_2^{(k+1)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\nu} \left( \mathbf{s}_1^{(k)} - \mathbf{s}_1^{(k+1)} \right) + \nabla \left( 2\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} - \mathbf{x}^{(k+1)} \right) \\ \frac{1}{\nu} \left( \mathbf{s}_2^{(k)} - \mathbf{s}_2^{(k+1)} \right) + \Phi \left( 2\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} - \mathbf{x}^{(k+1)} \right) \end{pmatrix}. \end{aligned} \quad (3.19)$$

### 3.7.2 Adaptive Optimization Procedure

The algorithm in (3.18) is non-adaptive since it uses constant step-sizes  $\mu$  and  $\nu$ . However, such a procedure often presents slow convergence as presented in Section 3.8. Therefore, motivated by the work by Goldstein *et al.* [97], an adaptive version of the CP algorithm is used for the ODT reconstructions (see Algorithm 3.1). While this variant also depends on a couple of parameters that adjust the algorithm adaptivity, these are more naturally related to the characteristics of the inverse problem solved by this optimization.

In this adaptive approach, the step-size update are tuned to the size of the primal and dual residuals at each iteration. If the primal residual is large with respect to the dual residual, then the primal step-size  $\mu^{(k)}$  is increased by a factor  $\frac{1}{1-\rho^{(k)}}$ , and the dual step-size  $\nu^{(k)}$  is decreased by the same factor. If the dual residual is large with respect to the primal residual, then the primal step-size is decreased and the dual step-size is increased. The parameter  $\rho^{(k)}$  is a constant that controls the adaptivity level of the method, and it is updated as  $\rho^{(k+1)} = \beta\rho^{(k)}$ , for some  $\beta < 1$ . In Algorithm 3.1, the parameter  $\gamma > 1$  is used to compare the sizes of the primal and dual residuals and the scaling parameter  $c > 0$ , which depends on the image expected dynamics, is used to balance the residuals. The specific values selected for the parameters of Algorithm 3.1 are those recommended by Goldstein *et al.* [97].

## 3.8 Experiments

In this section, the ODT reconstruction is first compared to the common tomographic (AT) reconstruction using the FBP and ME methods. Then, the proposed regularized reconstruction (TV- $\ell_2$ ) is compared with the FBP and ME procedures on synthetic and experimental ODT data.

---

**Algorithm 3.1** Adaptive ODT reconstruction
 

---

**Initialization:**  $\bar{\mathbf{x}}^{(1)} = \mathbf{0}_N$ ,  $\mathbf{s}_1^{(1)} = \mathbf{0}_{2N}$ ,  $\mathbf{s}_2^{(1)} = \mathbf{0}_M$ ,  $\mathbf{x}^{(1)} = \tilde{\mathbf{n}}_{\text{FBP}}$ ,  $\nu^{(1)} = \mu^{(1)} = 0.9 / \|\mathbf{K}\|$ ,  
 $\gamma = 1.1$ ,  $\rho^{(1)} = 0.5$ ,  $\beta = 0.95$ ,  $c = 1000$ ,  $\text{MaxIter} = 50 \times 10^4$ .

1: **for**  $k = 1$  to  $\text{MaxIter}$  **do**

Compute the CP iterations (3.18):

- 2:  $\mathbf{s}_1^{(k+1)} = \text{prox}_{\nu^{(k)}F_1^*} \left( \mathbf{s}_1^{(k)} + \nu^{(k)} \nabla \bar{\mathbf{x}}^{(k)} \right)$
- 3:  $\mathbf{s}_2^{(k+1)} = \text{prox}_{\nu^{(k)}F_2^*} \left( \mathbf{s}_2^{(k)} + \nu^{(k)} \Phi \bar{\mathbf{x}}^{(k)} \right)$
- 4:  $\mathbf{x}^{(k+1)} = \text{proj}_{\mathcal{P}_0} \left( \mathbf{x}^{(k)} - \frac{\mu^{(k)}}{2} \left( \nabla^* \mathbf{s}_1^{(k+1)} + \Phi^* \mathbf{s}_2^{(k+1)} \right) \right)$
- 5:  $\bar{\mathbf{x}}^{(k+1)} = 2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$

Compute the primal and dual residual norms (3.19):

- 6:  $p^{(k+1)} = \left\| \frac{2}{\mu^{(k)}} \left( \mathbf{x}^{(k)} - \mathbf{x}^{(k+1)} \right) \right\|_1$
- 7:  $d^{(k+1)} = \left\| \begin{pmatrix} \frac{1}{\nu^{(k)}} \left( \mathbf{s}_1^{(k)} - \mathbf{s}_1^{(k+1)} \right) + \nabla \left( \bar{\mathbf{x}}^{(k)} - \mathbf{x}^{(k+1)} \right) \\ \frac{1}{\nu^{(k)}} \left( \mathbf{s}_2^{(k)} - \mathbf{s}_2^{(k+1)} \right) + \Phi \left( \bar{\mathbf{x}}^{(k)} - \mathbf{x}^{(k+1)} \right) \end{pmatrix} \right\|_1$

Parameters update:

- 8: **if**  $p^{(k+1)} > cd^{(k+1)}\gamma$  **then**  $\triangleright$  Primal residual is larger than dual
- 9:  $\mu^{(k+1)} = \mu^{(k)}(1 - \rho^{(k)})$ ;  $\nu^{(k+1)} = \nu^{(k)} / (1 - \rho^{(k)})$ ;  $\rho^{(k+1)} = \rho^{(k)}\beta$
- 10: **else if**  $p^{(k+1)} < cd^{(k+1)} / \gamma$  **then**  $\triangleright$  Dual residual is larger than primal
- 11:  $\mu^{(k+1)} = \mu^{(k)} / (1 - \rho^{(k)})$ ;  $\nu^{(k+1)} = \nu^{(k)}(1 - \rho^{(k)})$ ;  $\rho^{(k+1)} = \rho^{(k)}\beta$
- 12: **else**  $\triangleright$  Similar primal and dual residuals, i.e.,  $p^{(k+1)} \in [cd^{(k+1)} / \gamma, \gamma cd^{(k+1)}]$
- 13:  $\mu^{(k+1)} = \mu^{(k)}$ ;  $\nu^{(k+1)} = \nu^{(k)}$ ;  $\rho^{(k+1)} = \rho^{(k)}$
- 14: **end if**

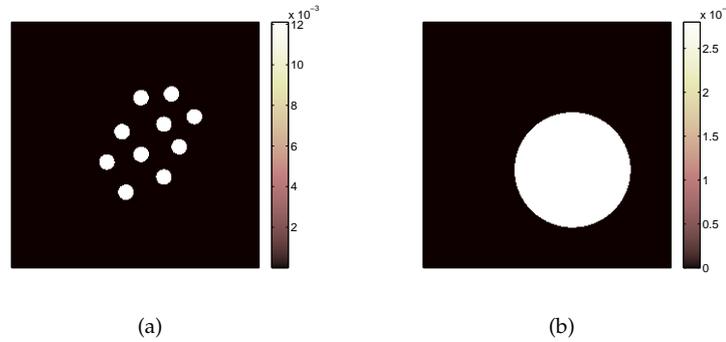
Stop if stable behavior:

- 15: **if**  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{(k)}\|_2 \leq \text{Th}$  **then** break.
  - 16: **end if**
  - 17: **end for**
  - 18: Return  $\tilde{\mathbf{n}}_{\text{TV-}\ell_2} = \mathbf{x}^{(k+1)}$
-

All algorithms were implemented in MATLAB and executed on a 3.2 GHz Intel i5-650 CPU with 3.7 GiB of RAM, running a 64 Bit Ubuntu 14.04 LTS operating system.

### 3.8.1 Synthetic Data

Three kinds of discrete synthetic 2-D RIM are selected to test the reconstruction. They are defined on a  $256 \times 256$  pixel grid ( $N = 256^2$ ). In the first object, the RIM ( $n$ ) simulates a 2-D section of a bundle of 10 fibers of radius 8 pixels each, immersed in an optical fluid (the background). The two media have a refractive index difference of  $\delta n = 12.1 \times 10^{-3}$  (see Figure 3.5-(a)). The second object consists in a homogeneous ball centered in the pixel (154, 154) with a radius of 60 pixels, immersed in a liquid with  $\delta n = 2.8 \times 10^{-3}$  (see Figure 3.5-(b)). These two objects were selected in correspondence to the available experimental data we use for reconstruction later in this section. The third object is the well-known Shepp-Logan phantom (see Figure 3.2-(b)), which is a more complex image in a “cartoon-shape” model.



**Figure 3.5** Realistic refractive index maps: (a) synthetic fibers bundle and (b) synthetic ball.

The measurements were simulated according to (3.11) by means of the operator<sup>5</sup>  $\Phi$ , and then, additive white Gaussian noise  $(\eta_{\text{obs}})_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma_{\text{obs}}^2)$  is added in order to simulate a realistic ODT scenario.

<sup>5</sup>The operators used in this work were numerically implemented using the SPARCO framework [132].

The operator  $\Phi$  is defined as  $\Phi_\epsilon = \Theta \mathbf{D} \mathbf{F}_\epsilon$ , with  $\epsilon$  representing the distortion of  $\mathbf{F}_\epsilon$  regarding a true operator  $\mathbf{F}_{\text{true}}$  that would provide the actual NDFT. As discussed in Section 3.3.3, the NDFT computational time is inversely proportional to this parameter  $\epsilon$  in  $\mathcal{O}(N \log(N/\epsilon))$ . Therefore, we need to do a compromise between an accurate and an efficient computation of the NDFT. For this reason, we use two different operators: (i) an accurate and high dimensional operator  $\Phi_{\epsilon_0} = \Theta \mathbf{D} \mathbf{F}_{\epsilon_0}$  for the acquisition, with a small  $\epsilon_0 = 10^{-14}$ ; and (ii) a less accurate but lower dimensional operator  $\Phi_{\epsilon_1} = \Theta \mathbf{D} \mathbf{F}_{\epsilon_1}$  for the reconstruction, with  $\epsilon_1 > \epsilon_0$ . The error caused for using a higher  $\epsilon$  for the reconstruction is taken into account in  $\epsilon_{\text{nfft}}$  (see Section 3.6).

For each object, the ODT measurements are obtained with  $N_\tau = 367$  according to a varying number of orientations  $N_\theta$ , which allows to analyze the compressiveness of the reconstruction method. In this synthetic experiment, the orientations  $\theta$  are taken in  $[0, \pi)$  so that  $N_\theta = 360$  corresponds to two orientations per degree. Hereafter, we consider this last situation, *i.e.*,  $\delta\theta = \pi/360$  as a “full observation” scenario since, given the considered RIM resolution, the discrete frequency plane is almost fully covered in this case. More generally, we say that a given orientation number  $N_\theta$  is associated to  $(100 \frac{N_\theta}{N_{\text{full}}})\%$  of the full coverage,  $N_{\text{full}}$  being the number of orientations when  $\delta\theta = \pi/360$ .

The reconstruction robustness with respect to the noise level has been considered for a Measurement signal-to-noise ratio

$$\text{MSNR} = 20 \log_{10} \frac{\|\Delta\|_2}{\|\eta\|_2}$$

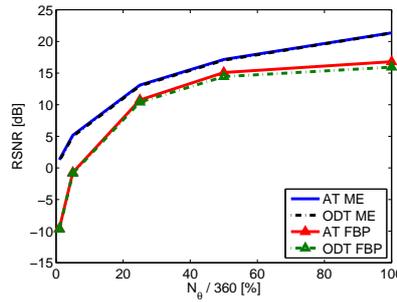
taken in  $\{10 \text{ dB}, 20 \text{ dB}, +\infty\}$ . This last case with MSNR close to  $+\infty$  corresponds to the noiseless scenario, where no Gaussian noise is added, only the NFFT interpolation error ( $\eta_{\text{nfft}}$ ) is taken into account. This actually provides a high MSNR value around 270 dB.

The reconstruction quality of  $\tilde{\mathbf{n}} \in \{\tilde{\mathbf{n}}_{\text{FBP}}, \tilde{\mathbf{n}}_{\text{ME}}, \tilde{\mathbf{n}}_{\text{TV-}\ell_2}\}$  is measured with respect to the true RIM  $\mathbf{n}$  using the Reconstruction SNR (RSNR), where

$$\text{RSNR} = 20 \log_{10} \frac{\|\mathbf{n}\|_2}{\|\mathbf{n} - \tilde{\mathbf{n}}\|_2}.$$

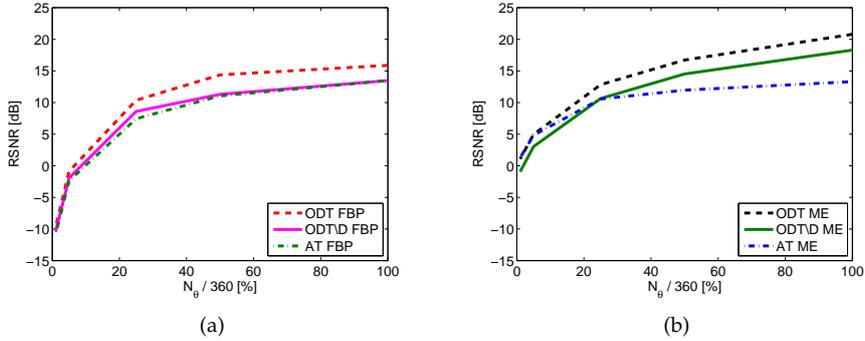
### 3.8.1.1 Robustness Comparison for AT and ODT

In order to assess numerically the impact of operator  $\mathbf{D}$  in ODT, we compare the RSNR between AT and ODT in similar noisy acquisition scenarios. The comparison is made using the FBP and ME procedures, commonly applied in tomographic reconstructions. In ODT, the mean of the image cannot be estimated but this is not taken into account by FBP and ME procedures, thus the mean of the reconstructed image was removed for the computation of the RSNR in order to correctly compare the two reconstructions. We analyze the impact of the affine frequency  $\omega$ , present in ODT, via the compressiveness and noise robustness. For this, we focus on the reconstruction of the bundle of fibers for different number of orientations  $N_\theta \in \{4, 18, 90, 180, 360\}$ . The results are depicted in Figure 3.6 and Figure 3.7.



**Figure 3.6** Absorption Tomography vs Deflectometric Tomography for different number of orientations  $N_\theta$  with  $\text{MSNR} = +\infty$ . Using FBP and ME procedures.

In Figure 3.6, we can see that, when no Gaussian noise is added, we obtain similar RSNR values for both AT and ODT. The same behavior is observed for both FBP and ME reconstructions, with FBP always providing a lower RSNR. The impact of the parameter  $\omega$  is evident only in the convergence time of ME, causing the ODT reconstruction to be 4 times slower than the AT one. However, when we add Gaussian noise in such a way that both data have  $\text{MSNR} = 20$  dB, the AT reconstruction presents a fast degradation while the ODT reconstruction remains almost unaffected by the noise (see Figure 3.7). These results, observed for both FBP and ME, corroborate the discussion in Section 3.6.2.



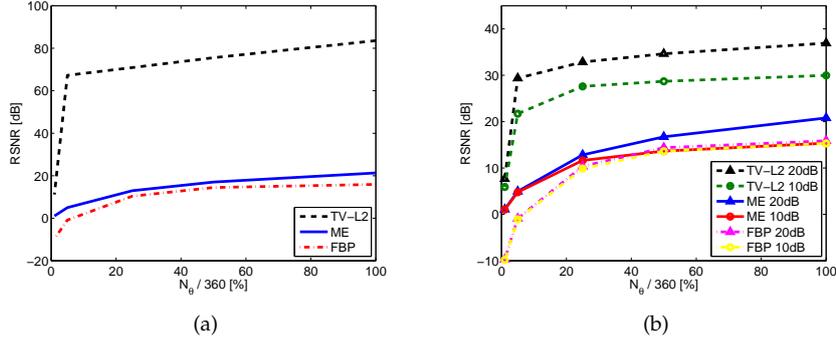
**Figure 3.7** Absorption Tomography vs Deflectometric Tomography for different number of orientations  $N_\theta$  with MSNR = 20 dB using (a) the FBP algorithm and (b) the ME procedure.

Following the discussion from Remark 3.5, we analyze now the reconstruction of the RIM using a simplified ODT sensing model that is close to the classical tomographic model in (3.12). In Figure 3.7, we show a third curve that corresponds to the RIM reconstruction from a noisy ODT sensing where the Fourier measurements are divided by the diagonal operator  $\mathbf{D}$  (denoted ODT \  $\mathbf{D}$ ) as in (3.12). The results were obtained using the FBP and ME procedures and for MSNR = 20 dB. As it was expected, when dividing the measurements by the operator  $\mathbf{D}$ , the reconstruction quality decreases significantly compared to the results obtained with the complete ODT sensing model in (3.11). Moreover, the regularized formulation TV- $\ell_2$  cannot be used for this ODT reconstruction because the noise is heteroscedastic.

### 3.8.1.2 Results for the TV- $\ell_2$ Reconstruction Method

The TV- $\ell_2$  reconstruction is compared with the common FBP and ME methods. The reconstruction quality is investigated with respect to compressiveness and noise robustness. On the contrary of FBP and ME, the TV- $\ell_2$  method takes into account the zero mean of the image by the *frontier constraint*. Therefore, the mean of the reconstructed image is only removed from the ME and FBP results. Figure 3.8 presents comparison graphs of FBP, ME and TV- $\ell_2$  showing the RSNR vs the number of orientations  $N_\theta \in \{4, 18, 90, 180, 360\}$  for

the three noise scenarios. These results correspond to the reconstruction of the bundle of fibers for  $\text{Th} = 10^{-5}$ .



**Figure 3.8** FBP, ME and TV- $\ell_2$  for different number of orientations  $N_\theta$  with (a)  $\text{MSNR} = +\infty$  and (b)  $\text{MSNR} = 20$  dB and  $\text{MSNR} = 10$  dB.

In Figure 3.8-(a), we present the scenario without added noise, *i.e.*,  $\text{MSNR} = +\infty$ . We can see that, for a full coverage, *i.e.*,  $N_\theta = 360$ , as the TV- $\ell_2$  method takes into account the small noise coming from the NFFT interpolation error, it provides a very good reconstruction that outperforms by 62 dB the ME reconstruction quality and by 68 dB the FBP reconstruction quality.

Both FBP and ME methods degrade rapidly when the problem is ill-posed, *i.e.*, when the projections space is not fully covered, whereas the TV- $\ell_2$  method maintains a high performance. By promoting a small TV-norm, the regularized method presents high compressiveness, as it can be observed in the graph where a high reconstruction quality is still achieved at only 5% of 360 incident angles, obtaining a gain of 62 dB over ME and of 68 dB over FBP. Although the performance of the algorithm decreases significantly for a coverage of 1%, it still provides a higher reconstruction quality than both ME and FBP.

The high compressiveness properties of the TV- $\ell_2$  method are preserved when we add Gaussian noise. We are able to obtain good quality images even for a compressive and highly noisy sensing. With TV- $\ell_2$ , at a  $\text{MSNR} = 10$  dB, we get a RSNR of 22 dB for a 5% radial coverage compared to 5 dB for ME and -1 dB for FBP. However, we can notice how the reconstruction quality of

$\text{TV-}\ell_2$  diminishes with respect to the noiseless scenario, whereas FBP and ME are less affected by the noise.

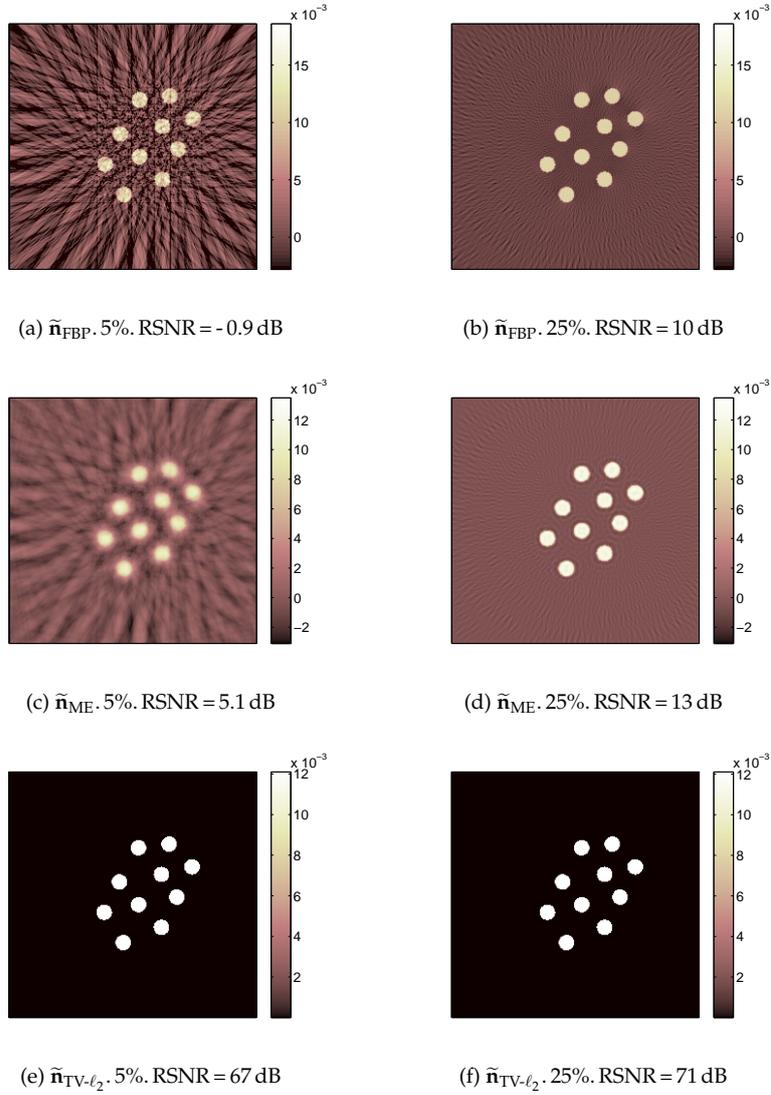
Figure 3.9 presents the resulting images when reconstructing the bundle of fibers in a noiseless scenario and for  $N_\theta = \{18, 90\}$ , which represents, respectively, a coverage of 5% and 25% of the frequency plane. The algorithm is set to stop when  $\text{Th} = 10^{-5}$  is reached. We notice how the  $\text{TV-}\ell_2$  method preserves the image dynamics even for 5% of coverage, while FBP and ME provide images with implausible negative values. Moreover, at low measurements regimes, some artifacts appear in both FBP and ME results. The image dynamics are less preserved in the FBP reconstruction, where the artifacts also affect the center of the fibers for a 5% of coverage.

About the computational time, since FBP is an analytic method, it takes less than one second for the reconstruction, while both ME and  $\text{TV-}\ell_2$  are iterative methods and require more time. Table 3.1 summarizes the time required by ME and  $\text{TV-}\ell_2$  to reach the same stopping threshold value of  $10^{-5}$ . The results are shown for a coverage of 25% and 5% of the frequency plane.

# Measurements [%]	Method	# iter	Time	RSNR [dB]
25	ME	6280	1h05'	13
	$\text{TV-}\ell_2$	1900	28'	71
5	ME	3450	51'	5.1
	$\text{TV-}\ell_2$	6620	1h49'	67

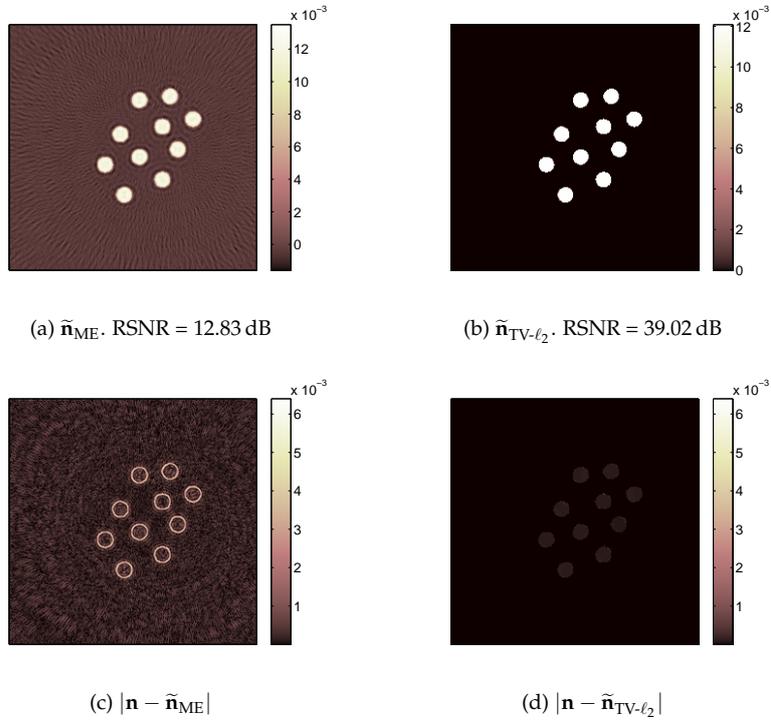
**Table 3.1** Computational time required by ME and  $\text{TV-}\ell_2$  methods to reach the stopping threshold value of  $10^{-5}$ .

We notice that, for 25% of coverage,  $\text{TV-}\ell_2$  outperforms both in quality and in computational time compared to ME. For a coverage of 5%, the  $\text{TV-}\ell_2$  method requires a higher time to converge, however, the reconstruction quality is clearly higher when using the proposed regularized method. In the case where the quality of the image reconstruction is sufficiently high, the threshold can be decreased to a less restrictive value. At the end of this section we perform a convergence analysis for different threshold values, which allows choosing the suitable threshold for a required quality or convergence time.



**Figure 3.9** Reconstructed images using FBP, ME and TV- $\ell_2$  methods for MSNR =  $+\infty$  and different number of orientations. In the upper row, we have the FBP reconstruction results for (a)  $N_\theta = 18$  and (b)  $N_\theta = 90$ . In the central row, we have the ME reconstruction results for (c)  $N_\theta = 18$  and (d)  $N_\theta = 90$ . In the bottom row, we have the TV- $\ell_2$  reconstruction results for (e)  $N_\theta = 18$  and (f)  $N_\theta = 90$ .

Although the FBP method requires less time than ME, we have noticed that the ME method outperforms FBP at low to medium reduction of the total number of angles. Henceforth, only the ME reconstruction method will be used for comparison with the  $\text{TV-}\ell_2$  method. In Figure 3.10, we present the reconstructed images for the bundle of fibers for a moderately noisy sensing ( $\text{MSNR} = 20 \text{ dB}$ ). We also show the error images in order to provide a better appreciation of the difference between both methods.

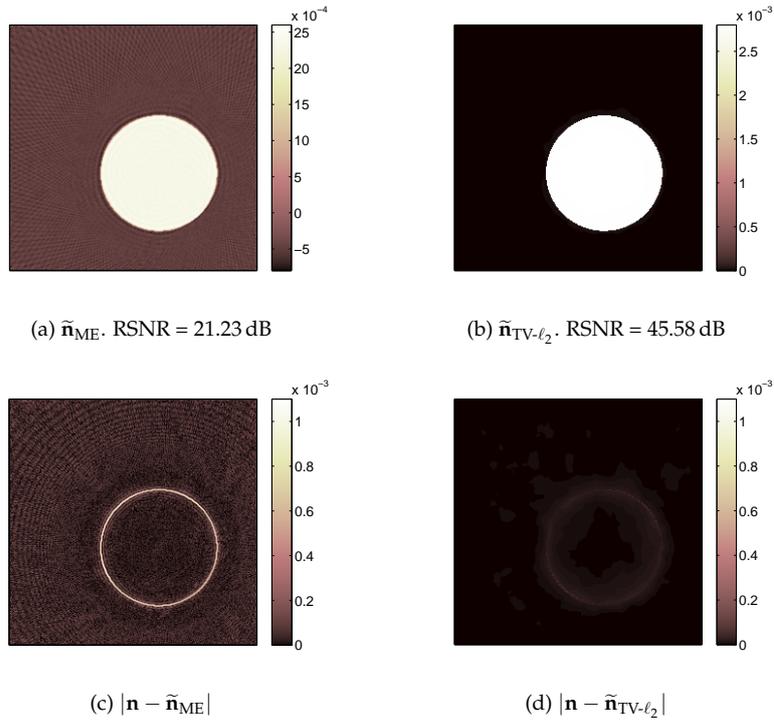


**Figure 3.10** Fiber reconstruction for  $\text{MSNR} = 20 \text{ dB}$  and  $N_\theta = 90$ . Reconstructed image  $\tilde{\mathbf{n}}$  using (a) ME and (b)  $\text{TV-}\ell_2$  reconstruction methods. Difference between the ground truth  $\mathbf{n}$  and the reconstructed image  $\tilde{\mathbf{n}}$  using (c) ME and (d)  $\text{TV-}\ell_2$  reconstruction methods.

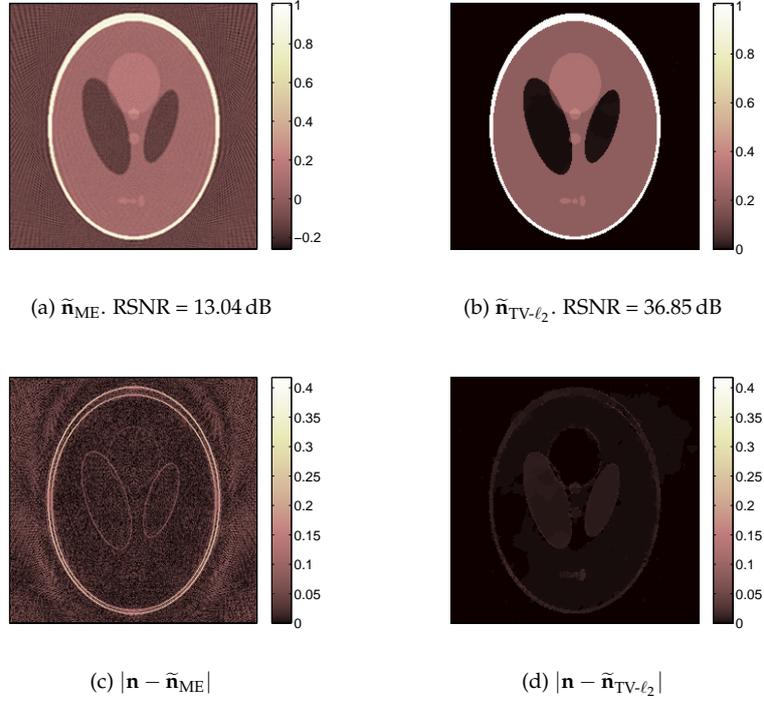
In this noisy scenario, the fiber contours are no longer well estimated using ME and, as we have coverage of only 25%, some oscillating (Gibbs) artifacts appear. On the contrary, the regularized method provides a good estimation

on the borders, with no visible artifacts. We notice certain loss in the dynamics of the image, which causes a lower RSNR compared to the noiseless scenario.

Let us show now some results obtained for the other two synthetic images. Figures 3.11 and 3.12 present the results obtained using ME and TV- $\ell_2$  on the sphere and the Shepp-Logan phantoms, respectively. This comparison was performed for MSNR = 20 dB and  $N_\theta = 90$ , *i.e.*, a coverage of 25% of the frequency plane. As the image expected dynamics change for the Shepp-Logan phantom, the parameter  $c$  for the residuals balancing was set to 250 instead of 1000.



**Figure 3.11** Sphere reconstruction for MSNR = 20 dB and  $N_\theta = 90$ . Reconstructed image  $\tilde{\mathbf{n}}$  using (a) ME and (b) TV- $\ell_2$  reconstruction methods. Difference between ground truth  $\mathbf{n}$  and Reconstructed image  $\tilde{\mathbf{n}}$  using (c) ME and (d) TV- $\ell_2$  reconstruction methods.



**Figure 3.12** Shepp-Logan phantom reconstruction for  $\text{MSNR} = 20$  dB and  $N_\theta = 90$ . Reconstructed image  $\tilde{\mathbf{n}}$  using (a) ME and (b)  $\text{TV-}\ell_2$  reconstruction methods. Difference between ground truth  $\mathbf{n}$  and Reconstructed image  $\tilde{\mathbf{n}}$  using (c) ME and (d)  $\text{TV-}\ell_2$  reconstruction methods.

The regularized method provides a better image dynamics for these phantoms than when reconstructing the fibers for the same noise level. For these phantoms, it can also be observed that ME reconstructions present a poor estimation on the borders and oscillating artifacts. Moreover, the error image shows a higher discordance with respect to the actual image.

Table 3.2 presents a more complete comparison of the RSNR obtained using ME and  $\text{TV-}\ell_2$  on the three synthetic images. The methods are analyzed for three scenarios, one noiseless with  $\text{MSNR} = +\infty$ , and the other two with noise such that we have  $\text{MSNR} = 20$  dB and  $\text{MSNR} = 10$  dB. Results are presented for  $\text{Th} = 10^{-5}$  and  $N_\theta = 90$ , *i.e.*, a coverage of 25% of the frequency plane. When comparing the behavior of the algorithm for the different synthetic im-

ages, we can notice that the TV- $\ell_2$  method outperforms the ME method for all cases.

	RSNR[dB]					
	MSNR = $+\infty$		MSNR = 20 dB		MSNR = 10 dB	
	TV- $\ell_2$	ME	TV- $\ell_2$	ME	TV- $\ell_2$	ME
Fibers	70.9	13.1	39.02	12.83	35.69	11.63
Sphere	53.59	21.54	45.58	21.23	37.70	18.79
Shepp-Logan	54.37	13.21	36.85	13.04	25.24	11.79

**Table 3.2** Comparison of the different RSNR obtained using ME and TV- $\ell_2$  on the three synthetic images for  $N_\theta = 90$ .

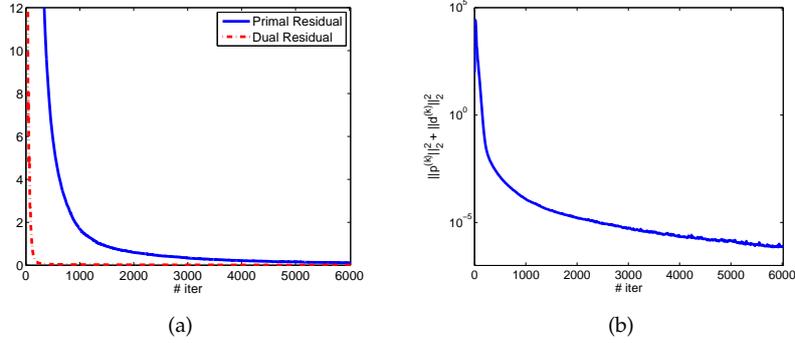
### 3.8.1.3 Algorithm Convergence

Finally, we analyze the convergence of the algorithm by studying the evolution of the primal and dual residual norms and of the RSNR for different threshold values. We also analyze the convergence difference between the adaptive and the non-adaptive method. For this, only the reconstruction of the bundle of fibers for  $N_\theta = 360$  and MSNR = 20 dB is used, however, the other cases present a similar behavior.

The evolution of the primal and dual residual norms in (3.19) and the convergence condition in (2.34) along the iterations is depicted in Figure 3.13. We notice that, when the number of iterations increases, the primal and dual residual norms tend to zero along with the energy ( $\|\mathbf{p}^{(k)}\|_2^2 + \|\mathbf{d}^{(k)}\|_2^2$ ).

In order to compare the adaptive and non-adaptive methods, we analyze the evolution of  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{(k)}\|_2$  and the RSNR along the iterations for two cases: (i) “Adapt” where the stepsizes are adaptively updated using Algorithm 3.1 and (ii) “Non-Adapt” where we use constant stepsizes equal to  $\mu = \nu = 0.9 / \|\mathbf{K}\|$ . Results are presented in Figure 3.14.

The evolution of  $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2 / \|\mathbf{x}^{(k)}\|_2$  helps us to analyze the stability of the algorithm. We can see the curves are not smooth specially for the non-adaptive method, which indicates a non stable behavior mainly due to a bad conditioning of the global operator  $\mathbf{K}$  in the product space optimization. This could be improved by a preconditioning procedure as described in [96, 133].



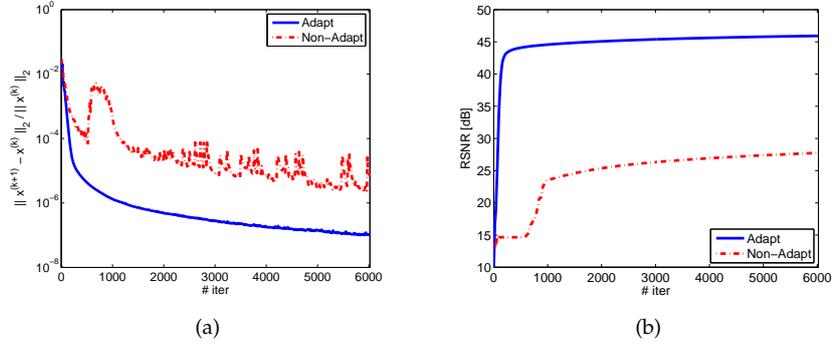
**Figure 3.13** Convergence results when reconstructing the bundle of fibers with  $N_\theta = 360$  and MSNR = 20 dB. (a) The primal and dual residual norms and (b) the convergence condition in (2.34).

We also notice that, for the same number of iterations, the adaptive method converges faster than the non-adaptive and provides a better result. Moreover, the adaptive method does not require to empirically set the parameters  $\mu$  and  $\nu$  as the algorithm will converge to the optimal parameters independently of the initialization.

Table 3.3 presents the values of the time and RSNR for different threshold values. These results show that a lower threshold provides higher reconstruction quality but significantly increases the number of CP iterations. In this specific reconstruction, setting the threshold to  $10^{-5}$  or running more than 500 CP iterations guarantees a RSNR higher than 42 dB.

Th	# iter	Time	RSNR [dB]
$10^{-4}$	190	5'	38.79
$10^{-5}$	420	11'	41.86
$10^{-6}$	1540	42'	43.86
$10^{-7}$	7150	3h15'	46.24

**Table 3.3** Convergence results when reconstructing the bundle of fibers with  $N_\theta = 360$  and MSNR = 20 dB.



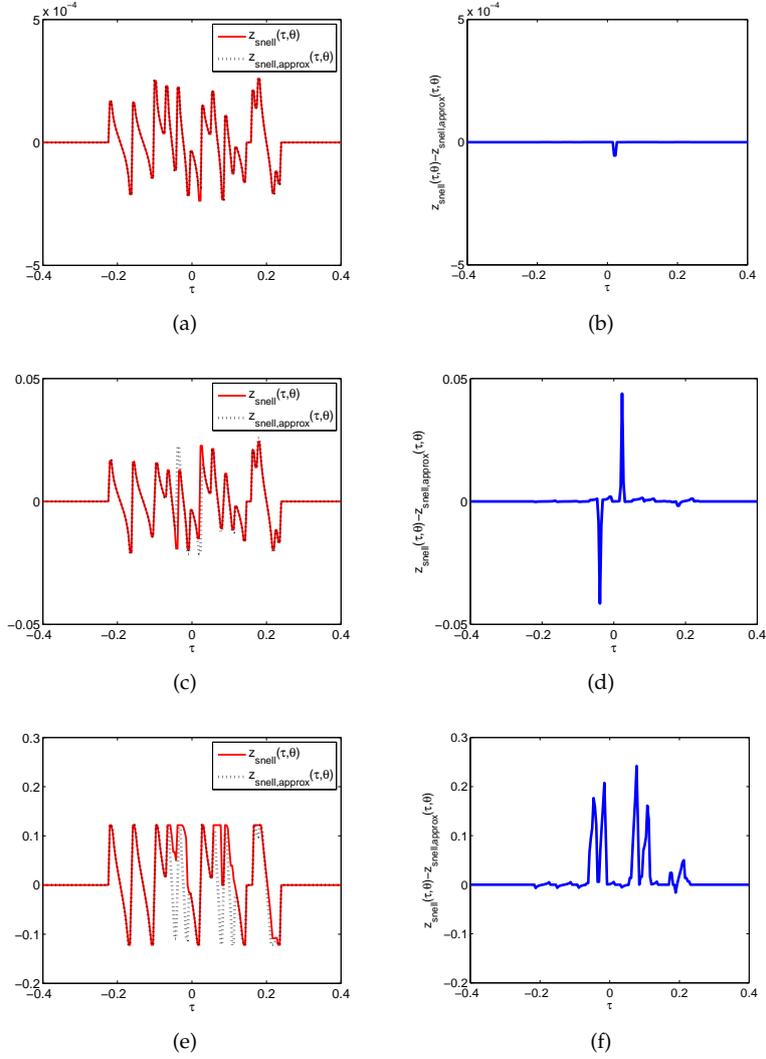
**Figure 3.14** Comparison between the adaptive and non-adaptive methods. Convergence results when reconstructing the bundle of fibers with  $N_\theta = 360$  and  $\text{MSNR} = 20$  dB. The progress of the values of (a)  $\|x^{(k+1)} - x^{(k)}\|_2 / \|x^{(k)}\|_2$  and (b) the RSNR, along the iterations.

#### 3.8.1.4 Impact of the Straight Line Trajectory Approximation

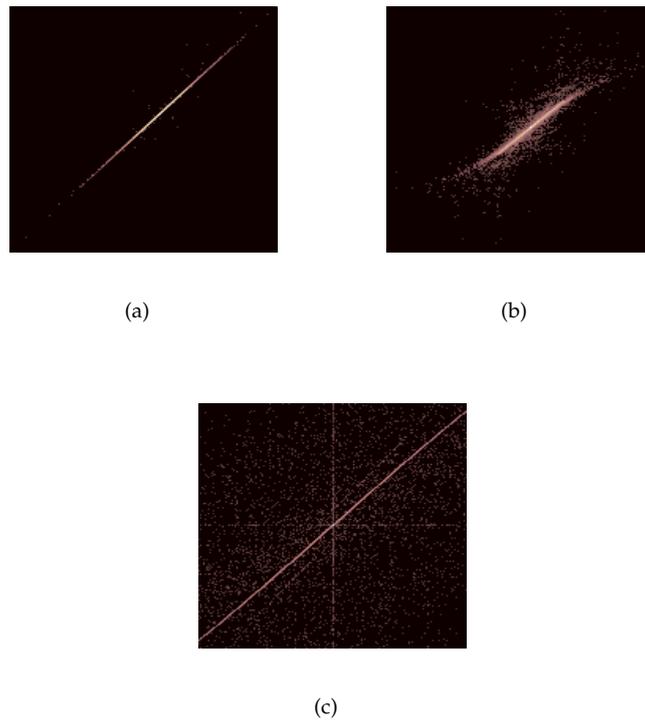
In this section, we aim at analyzing the impact of the straight line trajectory approximation that allowed formulating the simplified ODT model in (3.2). For this, we use the image of the synthetic bundle of fibers from Figure 3.5-(a) and simulate the measurements by means of the *ray tracing* method described in Appendix 3.A.3. The observations are obtained for  $N_\tau = 367$  and  $N_\theta = 180$  orientations taken in  $[0, \pi)$ .

The impact of this approximation is expected to change with respect to the deflection angles, which depend on the value of the refractive index difference  $\delta n$ . Therefore, in the simulations, we take several values of refractive index difference as  $\delta n \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ , covering a wide range of possible deflection angles.

In Figures 3.15 and 3.16, we show the difference between the measurements obtained using the *ray tracing* method with and without the straight line approximation. We observe that the error depends on the amplitude of the measured deflection angles, with an insignificant error for small deflection angles ( $\delta n = 10^{-4}$ ) and large errors for more important deflection angles ( $\delta n = 10^{-1}$ ). Interestingly, we notice that the error between the two models is sparse since it occurs at few values of  $\tau$ , with the sparsity level also changing with the magnitude of the deflection angles.

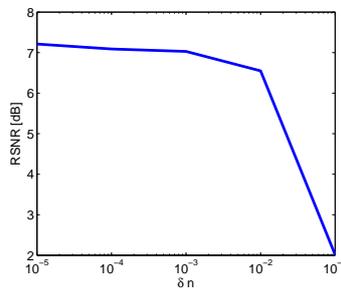


**Figure 3.15** Deflectometric measurements on the slice  $\theta = 0^\circ$  corresponding to the synthetic bundle of fibers in Figure 3.5-(a). In the left column, we present a comparison of the measurements obtained using the *ray tracing* method with and without approximation, *i.e.*,  $\mathbf{z}_{\text{snell,approx}}(\tau, \theta)$  and  $\mathbf{z}_{\text{snell}}(\tau, \theta)$ , respectively. In the right column, we present the difference between both measurements, *i.e.*,  $\mathbf{z}_{\text{snell}}(\tau, \theta) - \mathbf{z}_{\text{snell,approx}}(\tau, \theta)$ . The measurements in the top row correspond to  $\delta n = 10^{-4}$ , the ones in the middle row correspond to  $\delta n = 10^{-2}$ , and the ones in the bottom row correspond to  $\delta n = 10^{-1}$ . Remark that the scales are different in the three rows due to the high difference between the measured deflection angles in each case.



**Figure 3.16** Comparison between the deflectometric measurements obtained using the *ray tracing* method with (vertical axis) and without (horizontal axis) the straight line approximation. The figures show the logarithm of the 2-D histogram and the measurements correspond to the synthetic bundle of fibers in Figure 3.5-(a). The measurements in the top row correspond to  $\delta n = 10^{-4}$ , the ones in the middle row correspond to  $\delta n = 10^{-2}$ , and the ones in the bottom row correspond to  $\delta n = 10^{-1}$ . The intensity of the images belong to the interval  $[0, 1]$ , where the zero is represented as black and the one is represented as white.

We now use the TV- $\ell_2$  reconstruction method to recover the RIM from synthetic measurements of the bundle of fibers generated using the *ray tracing* method. No straight line approximation has been considered when generating the measurements. This allows evaluating the impact of the straight line assumption on the RIM reconstruction. Figure 3.17 presents the evolution of the reconstruction quality for different values of the refractive index difference.



**Figure 3.17** Reconstruction quality when using the TV- $\ell_2$  method to recover the synthetic bundle of fibers from measurements simulated using the *ray tracing* method. The reconstruction quality is analyzed for several values of the refractive index difference.

We observe that the reconstruction quality decreases when the refractive index difference increases. We notice that for the values of  $\delta n$  that are lower than  $10^{-2}$  the reconstruction quality does not significantly decrease. However, for  $\delta n > 10^{-2}$  the reconstruction quality drops considerably. This can be explained by a high modeling error caused by the straight line approximation being valid for less values of  $\tau$  (see Figure 3.15-(f)).

The impact of the modeling error could be decreased by replacing the additive noise model by a model that is more adapted to the actual behavior of the error (*e.g.*, a multiplicative model). Also, due to the sparsity of the error observed in Figure 3.15, we can also consider a fidelity term based on the  $\ell_1$  norm instead of the  $\ell_2$  norm. Additionally, the actual light ray trajectory could be estimated and inserted in an iterative process as done by Antoine *et al.* [37]. However, the forward model could not be represented in the frequency domain and we would lose its fast computation.

Remark that the values of the RSNR in Figure 3.16 are lower than the ones obtained in the previous sections, where the NFFT operator is used both for

the generation of the measurements and for the reconstruction. Instead, the measurements in this section are generated by means of the *ray tracing* method described in Appendix 3.A.3. This method tries to better simulate the experimental setup, taking into account the modeling and discretization errors. Let us now study the performance of the proposed reconstruction method in actual experimental data acquired by the Schlieren deflectometer.

### 3.8.2 Experimental Data

The reconstruction algorithm was tested with two particular transparent objects similar to the synthetic data studied in the previous section: a homogeneous sphere and a bundle of 10 fibers, both immersed in an optical fluid<sup>6</sup>. The reconstruction is based on  $N_\tau = 696$  parallel and equally spaced light rays. The experimental setup is based on the Schlieren Deflectometric Tomography described in Section 3.2.

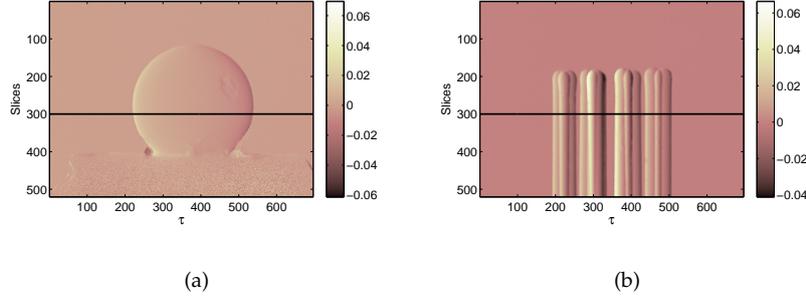
A  $696 \times 523$  pixels CCD camera was used for the acquisition, covering a field of view of  $3.25\text{mm} \times 2.43\text{mm}$ . This corresponds to  $N_\tau = 696$  parallel light rays and 523 2-D slices, which leads to  $\delta\tau = 4.7 \times 10^{-3}\text{mm}$ , and thus to  $\delta r = 4.7 \times 10^{-3}\text{mm}$ . Figure 3.18 presents one measurement of the deflection angles on the CCD camera grid for the two analyzed optical phantoms. This observation corresponds to  $\theta = 0^\circ$ .

The experimental configuration leads to a calibration problem. As the object is rotating, the rotation center is modified within a small range and the origin of the affine parameter  $\tau$  is altered. A post-acquisition calibration method was therefore implemented for correcting this effect. In short, for each angle, the method estimates the true centrum location by averaging the locations of the maximum and minimum deflection values along  $\tau$ .

The next two subsections present the characteristics of the two objects of interest and the reconstruction results obtained for the three tested methods (FBP, ME and TV- $\ell_2$ ) from the collected experimental observations. In all these experiments, and as discussed in Section 3.6.1, the TV- $\ell_2$  method was considered in the context of a 10 dB modeling noise. This choice seems somehow

---

<sup>6</sup>The data has been obtained from a collaboration between Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium and Lambda-X, Nivelles, Belgium.

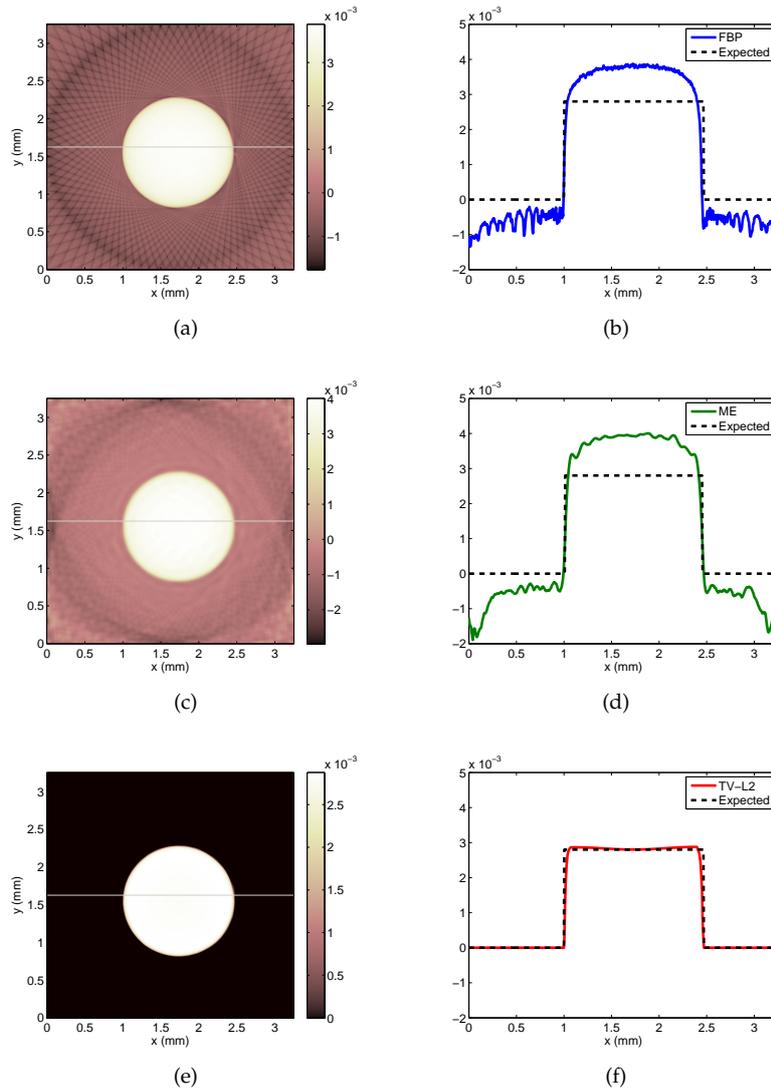


**Figure 3.18** Computed deflection angles (in radians) from CCD camera observations for  $\theta = 0^\circ$ , corresponding to (a) the homogeneous sphere and (b) the bundle of fibers. The 300<sup>th</sup> slice (over the 512 available slices) used for the reconstruction is indicated in the figures using a black line. In this specific slice, the refractive index maps are expected to have negligible variations along the vertical direction ( $\mathbf{e}_3$ ).

optimal for our experimental conditions. In several tests not reported here, higher or lower values of noise SNR lead either to severe artifacts in areas where no object is expected or to a significant loss in the expected RIM dynamics.

### 3.8.2.1 Homogeneous Sphere

The first observed object consists of a homogeneous sphere with a diameter of 1.5 mm. The difference of refractive index between the sphere and the optical fluid where it is immersed is  $\delta n = 2.8 \times 10^{-3}$ . The deviation map was measured for  $N_\theta = 45$  angular positions over 360 degrees (*i.e.*, 13% of measurements). The most important noise in the measurements is the modeling noise obtained by assuming  $\text{MSNR} = 10$  dB in (3.16), which provides  $\varepsilon_{\text{model}} = 0.035$ ; while the estimated observation noise provides  $\varepsilon_{\text{obs}} = 0.008$ . The NFFT interpolation noise is considerably smaller, with  $\varepsilon_{\text{nfft}} = 4.24 \times 10^{-16}$ . This noise estimation provides a  $\text{MSNR} = 9.79$  dB. Figure 3.19 shows the reconstruction results obtained when using FBP, ME and  $\text{TV-}\ell_2$  reconstruction methods, for the 300<sup>th</sup> 2-D slice of the observed 3-D object. The results are shown for a threshold  $\text{Th} = 10^{-5}$ , where ME converges in 7180 iterations and  $\text{TV-}\ell_2$  in 5180 iterations.



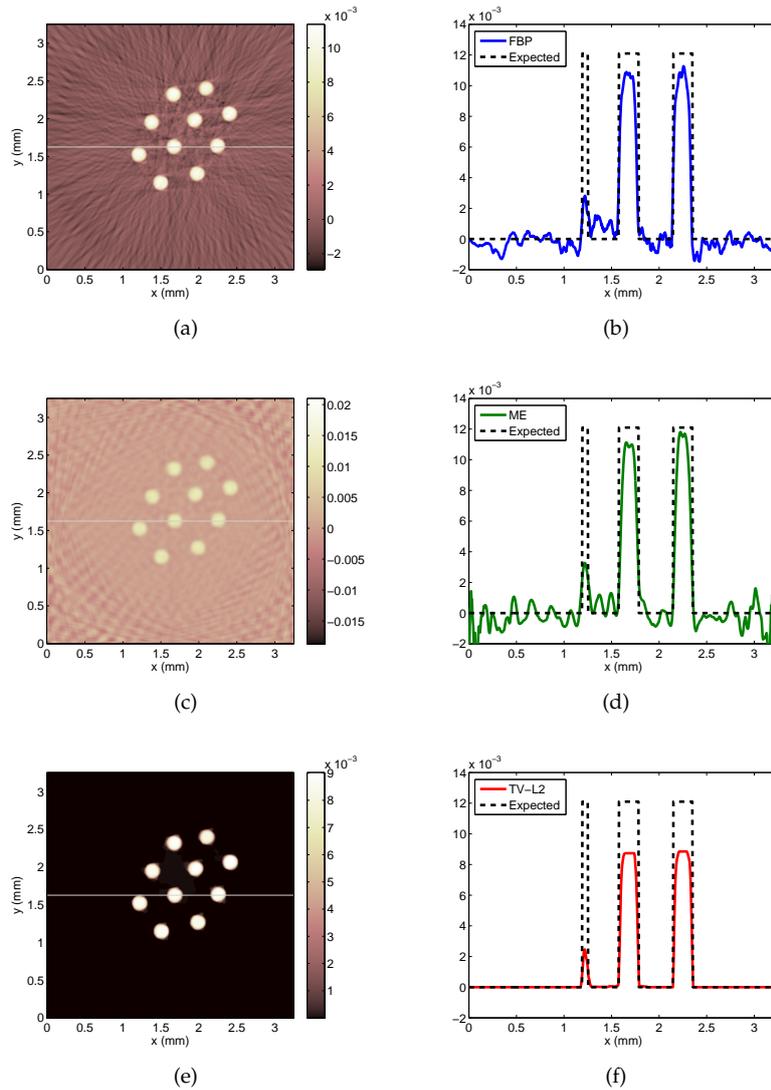
**Figure 3.19** Reconstructed sphere for 45 angular positions for the 300<sup>th</sup> 2-D slice. 2-D distribution using (a) FBP, (c) ME and (e) TV- $\ell_2$ . 1D profile along  $y = 1.625$  mm using (b) FBP (d) ME and (f) TV- $\ell_2$ .

We observe a similar behavior to the one found in the synthetic reconstruction. Compared to FBP and ME results, the sphere frontier is sharper with the TV- $\ell_2$  estimation and the RIM vanishes on the background. The image dynamics recovery is also more accurate using the proposed regularized method, whereas with FBP and ME the reconstructions present several artifacts with implausible negative values. It is important to notice that the preservation of the image dynamics depends mainly on the noise estimation and on the proper definition of the constants included in the operator  $\mathbf{D}$  (see Section 3.3). When considering the appropriate constants, we are able to make an equivalence between the physical problem and its discrete mathematical formulation.

### 3.8.2.2 Bundle of Fibers

The second measured object is a bundle of 10 fibers of 200  $\mu\text{m}$  diameter each. The refractive index difference with respect to the solution where the fibers are immersed is  $\delta n = 12.1 \times 10^{-3}$ . The experimental data was measured for 60 angular positions over 360 degrees (*i.e.*, 17% of measurements). As in the case of the sphere, the noise that has more influence in the measurements is the modeling noise obtained by assuming MSNR = 10 dB in (3.16), which provides  $\varepsilon_{\text{model}} = 0.093$ ; while the observation noise provides  $\varepsilon_{\text{obs}} = 0.005$ . The NFFT interpolation noise is considerably smaller ( $\varepsilon_{\text{nfft}} = 1 \times 10^{-15}$ ). This noise estimation provides MSNR = 9.98 dB.

Figure 3.20 shows the reconstruction results obtained using FBP, ME and TV- $\ell_2$  reconstruction methods, for the 300<sup>th</sup> 2-D slice. For a threshold of  $10^{-5}$ , ME converges in 33920 iterations and TV- $\ell_2$  in 4560 iterations. Compared to FBP and ME reconstructions, TV- $\ell_2$  provides a much sharper estimation of the true RIM and the background is correctly estimated to zero. However, the image dynamics is not properly recovered. Such reconstruction error is present for the fibers because the refractive index difference between the material and the optical fluid is higher than for the sphere. In this case, the deflection angle is higher and it causes the modeling error to increase (see Section 3.8.1.4). We can also notice that a section of the bundle of fibers represents a complex image to reconstruct, since the light enters and comes out of multiple fibers. This amplifies the modeling error.



**Figure 3.20** Reconstructed bundle of 10 fibers for 60 angular positions for the 300<sup>th</sup> 2-D slice. 2-D distribution using (a) FBP, (c) ME and (e) TV- $\ell_2$ . 1D profile along  $y = 1.625$  mm using (b) FBP, (d) ME and (f) TV- $\ell_2$ .

Furthermore, investigating the TV- $\ell_2$  program in (3.15) and assuming that the unknown true RIM is a feasible point of the constraints, *i.e.*, the noise power  $\varepsilon$  is correctly bounded, the TV norm of the solution is necessarily smaller than the one of the true map. The norm reduction actually increases with the noise power since the optimization has then more freedom to reduce the TV norm of the solution. A direct impact for cartoon-shape maps is thus a reduction of RIM dynamics in the reconstruction compared to the expected one. Investigating if this dynamics loss could be limited (*e.g.*, by constraining the mean) is a matter of future study.

### 3.9 Conclusions

We have demonstrated how regularized reconstruction methods, such as TV- $\ell_2$ , can be used in the framework of optical deflectometric tomography in order to tackle the lack of deflectometric observations and to make the ODT imaging process robust to noise. The proposed constrained optimization problem shows significant improvements in the reconstructions, compared to the well-known filtered back projection and minimum energy methods. The results confirm that, when dealing with a compressive setting, the Total-Variation regularization and the prior information constraints (non-negativity and FoV restriction) help in providing a unique and accurate estimation of the RIM, promoting sharp edges and preserving the image dynamics (when the modeling noise is limited). By working with the Chambolle-Pock algorithm we exploit the advantages of proximal operators and of primal-dual algorithms, and their flexibility to integrate multiple constraints. We have also shown that the use of the fast NFFT algorithm efficiently approximates (with a controlled error) the ODT sensing model involving the polar NDFT.

Noticeably, there still exist some artifacts in the experimental data reconstruction, coming from the modeling error. In order to handle this problem, it would be interesting to study how to avoid the loss in dynamics for a high modeling noise by introducing some additional constraints. Also, by taking into account the information provided by a ray tracing method, we could reformulate the problem such that the modeling error is decreased. This matter will be further discussed in Chapter 6.

It is important to note that the optical deflectometric framework treated in this chapter can also be applied to other imaging techniques, such as the X-ray phase contrast tomography. Interestingly, with the use of X-rays, the first order approximation involving a linear ray trajectory is no longer needed and the proposed methods are expected to provide better results.

In the imaging modality studied in this chapter, the light deflection  $\alpha$  is not directly measured but is related to the observed light intensities through the phase-shift  $\Delta x$ . In order to obtain the light deflection required to recover the RIM, the phase-shift needs to be first estimated from the observed light intensities using phase-shifting methods. However, the phase-shift can only be estimated inside a certain interval (e.g.,  $[-\pi, \pi)$ ) given by the SLM period [37]. If the actual phase-shift has values that are outside that interval, the estimated phase-shift is only a “wrapped” version of the actual one. Other applications such as magnetic resonance imaging and interferometry, are also concerned by this problem. In the next chapter, we address the issue of recovering the phase from its “wrapped” version such that it can be useful for further processing. The unwrapping inverse problem is formulated as a regularized convex optimization problem and we use similar reconstruction techniques as the ones used in this chapter.

## Appendix 3.A On the Deflectometric Model

In this appendix, we provide some details on the light deflection model, which allow understanding the origin of the ODT model presented in Section 3.2.1 and estimating the error committed when simplifying the actual ODT acquisition process.

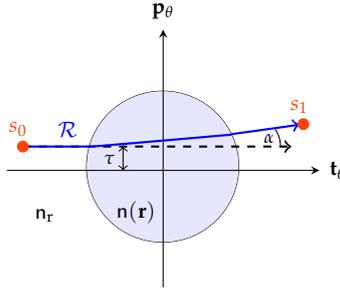
### 3.A.1 Derivation of the Deflection Equation

Using the notations of Section 3.2.1, we must prove that

$$\Delta(\tau, \theta) := \frac{1}{n_r} \int_{\mathbb{R}} \nabla n(\mathbf{r}_{\tau, \theta}(s)) \cdot \mathbf{p}_\theta \, ds = \frac{1}{n_r} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) \delta(\tau - \mathbf{r} \cdot \mathbf{p}_\theta) \, d^2 \mathbf{r}.$$

Let us consider the deflectometric model from Figure 3.1-(a) for one light ray at a distance  $\tau$  from the origin and one orientation  $\theta = 0^\circ$ . The light

ray departs from a point  $s_0$  with a velocity  $\frac{d}{ds}\mathbf{r}(s_0) = \dot{\mathbf{r}}(s_0)$  and arrives to a point  $s_1$  with a velocity  $\frac{d}{ds}\mathbf{r}(s_1) = \dot{\mathbf{r}}(s_1)$ . The departure and arrival points are located outside the object and inside the medium surrounding the object, *i.e.*,  $n(\mathbf{r}(s_0)) = n(\mathbf{r}(s_1)) = n_r$ . This is illustrated in Figure 3.21.



**Figure 3.21** The Deflectometric model for a given light ray at a distance  $\tau$  from the origin and for  $\theta = 0^\circ$ .

By integrating both sides of the light ray equation in (3.1) over the trajectory  $\mathcal{R}$  between  $s_0$  and  $s_1$  (see Figure 3.21), we have

$$\begin{aligned} \int_{s_0}^{s_1} \frac{d}{ds} \left( n(\mathbf{r}(s)) \frac{d}{ds} \mathbf{r}(s) \right) ds &= \int_{s_0}^{s_1} \nabla n(\mathbf{r}(s)) ds \\ n_r (\dot{\mathbf{r}}(s_1) - \dot{\mathbf{r}}(s_0)) &= \int_{s_0}^{s_1} \nabla n(\mathbf{r}(s)) ds. \end{aligned} \quad (3.20)$$

By projecting both sides of (3.20) on the direction  $\mathbf{p}_\theta$ , we obtain

$$\begin{aligned} (\dot{\mathbf{r}}(s_1) - \dot{\mathbf{r}}(s_0)) \cdot \mathbf{p}_\theta &= \frac{1}{n_r} \int_{\mathcal{R}(\mathbf{r}(s_0), \mathbf{r}(s_1))} \nabla n(\mathbf{r}(s)) \cdot \mathbf{p}_\theta ds \\ \sin \alpha &= \frac{1}{n_r} \int_{\mathcal{R}(\mathbf{r}(s_0), \mathbf{r}(s_1))} \nabla n(\mathbf{r}(s)) \cdot \mathbf{p}_\theta ds. \end{aligned} \quad (3.21)$$

Remark that the integral on the right hand side of (3.21) is done over a curved trajectory  $\mathcal{R}(\mathbf{r}(s_0), \mathbf{r}(s_1))$  that depends on the refractive index  $n$  [37]. In order to remove this dependency, we adopt in this work a first order approximation and assume that the trajectory  $\mathcal{R}$  is a straight line. In the general deflectometric model, a straight line trajectory depends on the distance to the origin  $\tau$  and the orientation  $\theta$ . Hence, the trajectory is described by  $\mathbf{r}$  that depends on  $(\tau, \theta) \in \mathbb{R} \times [0, 2\pi]$  and varies with the linear parameter  $s \in \mathbb{R}$  as

$\mathbf{r}_{\tau,\theta}(s) = s\mathbf{t}_\theta + \tau\mathbf{p}_\theta$ . This transforms (3.21) in the following line integral

$$\Delta(\tau, \theta) := \sin \alpha \approx \frac{1}{n_r} \int_{\mathbb{R}} \nabla n(\mathbf{r}_{\tau,\theta}(s)) \cdot \mathbf{p}_\theta \, ds.$$

The trajectory  $\mathcal{R}$  can also be written as  $\mathcal{R}(\tau, \theta) = \{\mathbf{r} \in \mathbb{R}^2 : \mathbf{r} \cdot \mathbf{p}_\theta = \tau\}$ . Using the Dirac distribution  $\delta(\cdot)$ , the line integral in the above equation transforms to the following integral over  $\mathbb{R}^2$

$$\Delta(\tau, \theta) = \frac{1}{n_r} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) \delta(\tau - \mathbf{r} \cdot \mathbf{p}_\theta) \, d^2\mathbf{r}.$$

### 3.A.2 Deflectometric Fourier Slice Theorem

Using the notations of Section 3.2.1, we must prove that

$$y(\omega, \theta) = \frac{2\pi i \omega}{n_r} \widehat{n}(\omega \mathbf{p}_\theta).$$

By the definition of  $y$  in (3.3), we have

$$\begin{aligned} y(\omega, \theta) &= \frac{1}{n_r} \int_{\mathbb{R}} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) \delta(\tau - \mathbf{r} \cdot \mathbf{p}_\theta) e^{-2\pi i \tau \omega} \, d^2\mathbf{r} \, d\tau \\ &= \frac{1}{n_r} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) e^{-2\pi i \mathbf{r} \cdot (\omega \mathbf{p}_\theta)} \, d^2\mathbf{r}. \end{aligned}$$

On the other hand, for any function  $F : \mathbb{R} \rightarrow \mathbb{C}$  integrable on  $\mathbb{R}$ , we have

$$\widehat{\frac{dF}{dt}}(\omega) = \int_{\mathbb{R}} \frac{d}{dt} F(t) e^{-2\pi i t \omega} \, dt = (2\pi i) \omega \widehat{F}(\omega).$$

Therefore, we compute straightforwardly for any  $\mathbf{a}, \boldsymbol{\zeta} \in \mathbb{R}^2$ , the following

$$\int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{a}) e^{-2\pi i \mathbf{r} \cdot \boldsymbol{\zeta}} \, d^2\mathbf{r} = (2\pi i) (\boldsymbol{\zeta} \cdot \mathbf{a}) \widehat{n}(\boldsymbol{\zeta}).$$

Setting  $\mathbf{a} = \mathbf{p}_\theta$  and  $\boldsymbol{\zeta} = \omega \mathbf{p}_\theta$ , we find finally

$$y(\omega, \theta) = \frac{1}{n_r} \int_{\mathbb{R}^2} (\nabla n(\mathbf{r}) \cdot \mathbf{p}_\theta) e^{-2\pi i \mathbf{r} \cdot (\omega \mathbf{p}_\theta)} \, d^2\mathbf{r} = \frac{2\pi i \omega}{n_r} \widehat{n}(\omega \mathbf{p}_\theta).$$

### 3.A.3 Ray Tracing

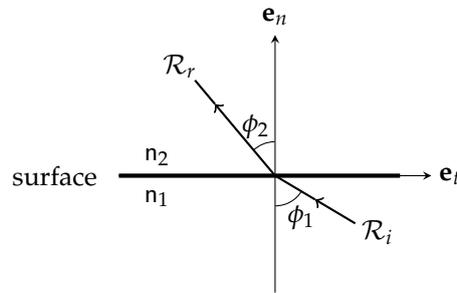
In this section, we describe the *ray tracing* method used to estimate the error committed by removing the trajectory dependence in the deflection angle (see Section 3.2.1). By the use of elementary geometry and the laws of

reflection and refraction, *ray tracing* methods [125, Section 4.9] allow an accurate determination of the path followed by the light from a source point to a detector. These methods have been used, for instance, in ultrasound tomography [134] and in terahertz tomography [7], to determine the actual trajectory of the rays and to introduce it in the reconstruction methods.

The behavior of the light when traversing a transparent object is described by the laws of reflection and refraction [125, Section 1.5]. When a light ray hits the surface of a transparent material, it is partly reflected and partly refracted into a second material. The direction of the incident and refracted rays,  $\mathcal{R}_i$  and  $\mathcal{R}_r$ , respectively, is described by the law of refraction or Snell's law [125, Section 1.5] as:

$$\sin \phi_2 = \frac{n_1}{n_2} \sin \phi_1, \quad (3.22)$$

where  $n_1$  and  $n_2$  are the refractive index of the two materials traversed by the light ray; and  $\phi_1$  and  $\phi_2$  represent the angles of the incident and refracted rays, respectively, with respect to the normal ( $\mathbf{e}_n$ ) of the surface separating the two materials. This phenomenon is depicted in Figure 3.22.

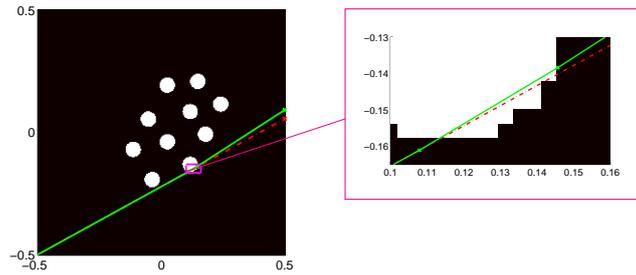


**Figure 3.22** Description of Snell's law for a light ray traversing a surface. In this figure, the ray goes from a material with refractive index  $n_1$  to another material with refractive index  $n_2$ , with  $n_1 < n_2$ .

Using the same notation as in Sections 3.2 and 3.3, we apply a *ray tracing* method based on Snell's law to the ODT problem. For this, we use the synthetic images depicted in Figure 3.5, which are defined in the following discrete grid of  $N$  pixels:  $\{\mathbf{r}_{m,n} : -0.5 \leq m, n \leq 0.5\}$ .

For each angle  $\theta$ , we generate a set of rays  $\mathcal{R}(\tau, \theta) = \{\mathbf{r} \in \mathbb{R}^2 : \mathbf{r} \cdot \mathbf{p}_\theta = \tau\}$  from a source point located outside the discrete grid. For each ray  $\mathcal{R}(\tau, \theta)$ , we estimate the first intersection point between the ray and each disk present in the image. If there is an intersection point, we apply Snell's law (3.22) in order to compute the angle of the refracted ray with respect to the normal of the surface, *i.e.*,  $\phi_2$  in Figure 3.22. The source of the refracted ray is given by the intersection point and its direction is computed using the angle  $\phi_2$ . We then estimate the next intersection point and the procedure is repeated until the end of the discrete grid is reached. The deflection angle  $\alpha_{\text{snell}}(\tau, \theta)$  is finally computed as the difference between the last refracted angle and the initial angle of incidence.

Figure 3.23 illustrates the *ray tracing* method for the bundle of fibers with a refractive index difference of  $\delta n = 5 \times 10^{-2}$ . The solid green line represents the deviated light ray according to Snell's law and the dashed red line represents the original trajectory of the light ray for  $\theta = 30^\circ$  and  $\tau = -0.1932$ . In this example, there are two intersection points and the initial ray deviates twice from its original direction  $\mathbf{t}_\theta$ . If a given ray traverses several objects, then the ray curvature is higher due to many intersection points. The refractive index difference in Figure 3.23 was chosen because it allows a better illustration of the light refraction (larger deviation angle). However, as it will be clear later in this section, the first order approximation considered in this work is not valid for this large refractive index difference.



**Figure 3.23** *Ray tracing* method applied to a synthetic image with a bundle of 10 fibers with  $\delta n = 5 \times 10^{-2}$ . The solid line represents the deviated light ray according to Snell's law and the dashed line represents the original trajectory of the light ray for  $\theta = 30^\circ$  and  $\tau = -0.1932$ .

Let us note that, in the actual ODT acquisition, the deflection angles are computed from the measured intensity values by means of phase-shifting techniques. Also, the angular aperture of the instrument allow measuring only the angles that are inside the interval  $[-7, 7]^\circ$ . To simulate this reality, the deflection angles obtained using Snell's law  $\alpha_{\text{snell}}(\tau, \theta)$  are first filtered by means of a median filter computed over three values and, then, are cropped inside the interval  $[-7, 7]^\circ$ . The ODT observations are computed as the sinus of the resulting angles and are gathered in the vector  $\mathbf{z}_{\text{snell}}(\tau, \theta)$ . The corresponding Fourier measurements  $\mathbf{y}_{\text{snell}}(\omega, \theta) \approx \mathbf{y}_{\text{true}}$  are computed using (3.7).

The straight line approximation adopted in this work, considers that the ray arriving at the detector has followed a single total deviation  $\alpha_{\text{snell,approx}}(\tau, \theta)$ , that is "equivalent" to the accumulation of all the actual deviations. Following the same procedure described above, the measured angles are filtered and cropped. Then their sinus is computed and gathered in the vector  $\mathbf{z}_{\text{snell,approx}}(\tau, \theta)$ . The corresponding Fourier measurements  $\mathbf{y}_{\text{snell,approx}}(\omega, \theta)$ , which represent the sensing model  $\Phi_{\text{true}}\mathbf{n}$  in (3.16), are computed using (3.7). The *modeling* error  $\eta_{\text{model}}$  in (3.16) is then estimated from the differences between these measurements and the "true" noiseless FDM described by Snell's law.

## Appendix 3.B Non-Equispaced Fourier Transform

The non-equispaced Fourier Transform (NFFT) allows a fast computation, *i.e.*, in  $\mathcal{O}(N \log N)$ , of the NDFT (defined in (3.8)) of a function defined on  $\mathcal{C}_N$ . This computation is performed with a controllable error which can be further reduced by increasing the computational time.

In a nutshell, the NFFT algorithm replaces the equivalent matrix multiplication  $\hat{\mathbf{f}} = \mathbf{F}\mathbf{f}$  of (3.9) by  $\hat{\mathbf{s}} = \bar{\mathbf{F}}\mathbf{f}$ , where  $\bar{\mathbf{F}}$  has a fast matrix-vector multiplication computation. In this scheme, the discrepancy between  $\hat{\mathbf{f}}$  and  $\hat{\mathbf{s}}$ , as measured by  $E_\infty(\mathbf{f}) := \|\hat{\mathbf{f}} - \hat{\mathbf{s}}\|_\infty$ , is controlled and kept small.

More precisely, the matrix  $\bar{\mathbf{F}}$  starts by embedding  $\mathbb{R}^N$  in a bigger regular space  $\mathbb{R}^n$ , with  $n > N$ . This is obtained from the multiplication of  $\mathbf{f}$  with a matrix  $\mathbf{W} \in \mathbb{R}^{n \times N}$  that performs a weighting of  $\mathbf{f}$  by a vector  $\mathbf{w} \in \mathbb{R}_+^N$  (component wise) followed by a symmetric zero-padding on each side of the

function domain  $\mathcal{C}_N$ . Once in  $\mathbb{R}^n$ , the common DFT matrix  $\mathbf{F}_n$  is applied. It can be computed with the FFT algorithm in order to obtain an oversampled Fourier transform of  $\mathbf{f} \in \mathbb{R}^N$ . Finally, a sparse matrix  $\mathbf{V} \in \mathbb{R}^{M \times n}$  multiplies the output of  $\mathbf{F}_n$  in order to end in the  $M$  dimensional space  $\widehat{\mathcal{P}}$ . Each row of  $\mathbf{V}$  corresponds to the translation in the 2-D Fourier domain of a compact and separable 2-D filter  $\psi$  on one specific point of the non-regular grid  $\widehat{\mathcal{P}}_M$ . As a final result, the matrix  $\bar{\mathbf{F}}$  is thus factorized as  $\bar{\mathbf{F}} = \mathbf{V}\mathbf{F}_n\mathbf{W}$  [127].

Without entering into unnecessary technicalities, the NFFT scheme is characterized by a precise connection between the function  $\psi$  defining  $\mathbf{V}$  and the weighting performed by  $\mathbf{W}$ . In particular, each component of  $\mathbf{w}$  is actually set to the inverse of the Fourier transform of a filter  $\varphi$ , while  $\psi$  is a periodization of (a truncation of) the same filter.

There exist several choices of windows  $\varphi/\psi$  associated to different numerical properties (*e.g.*, localized support in frequency and time, simple precomputations of the windows, ...). We select here the *translates of Gaussian bells* [135], involving a Gaussian behavior for  $\widehat{\varphi}(k)$ , which provides fast error decay for  $E_\infty(\mathbf{f})$ .

In particular, denoting by  $\kappa = n/N \geq 1$  the oversampling factor and using the FFT for matrix-vector multiplications involving  $\mathbf{F}_n$ , the total complexity  $T$  of the multiplications of  $\bar{\mathbf{F}}$  or  $\bar{\mathbf{F}}^*$  with vectors is

$$T(\bar{\mathbf{F}}) = \mathcal{O}\left(n \log n + N \left(\frac{2\kappa-1}{2\kappa-2}\right) \log \frac{1}{\epsilon}\right),$$

if we impose  $E_\infty(\mathbf{f}) \leq 4\|\mathbf{f}\|_1 \epsilon$ . For a fixed  $\kappa > 1$ , this reduces to  $T(\bar{\mathbf{F}}) = \mathcal{O}(N \log N/\epsilon)$ , which is far less than a direct computation of the DFT in  $\mathcal{O}(MN)$ , even for a small value of  $\epsilon$ .

### Appendix 3.C Convex Conjugate Functions

In this appendix, we describe how to compute the convex conjugate of the indicator function  $\iota_{\mathcal{C}}$  from Section 3.7.1.

Given the indicator function  $\iota_{\mathcal{C}}(\mathbf{u})$  of the convex set  $\mathcal{C} = \{\mathbf{v} \in \mathbb{R}^M : \|\mathbf{v} - \mathbf{y}\|_2 \leq \epsilon\}$ , its dual function can be computed via

the Legendre transform in (2.20) as follows:

$$\begin{aligned} i_C^*(\mathbf{v}) &= \max_{\mathbf{u}} \langle \mathbf{v}, \mathbf{u} \rangle - i_C(\mathbf{u}) = \langle \mathbf{v}, \mathbf{y} \rangle + \max_{\mathbf{u}: \|\mathbf{u}-\mathbf{y}\|_2 \leq \varepsilon} \langle \mathbf{v}, \mathbf{u} - \mathbf{y} \rangle \\ &= \langle \mathbf{v}, \mathbf{y} \rangle + \max_{\mathbf{b}: \|\mathbf{b}\|_2 \leq \varepsilon} \langle \mathbf{v}, \mathbf{b} \rangle. \end{aligned}$$

The value of  $\{\mathbf{b} : \|\mathbf{b}\|_2 \leq \varepsilon\}$  that maximizes the last expression is  $\mathbf{b} = \frac{\mathbf{v}}{\|\mathbf{v}\|_2} \varepsilon$ , and we obtain

$$i_C^*(\mathbf{v}) = \langle \mathbf{v}, \mathbf{y} \rangle + \varepsilon \|\mathbf{v}\|_2.$$



# Phase Unwrapping

# 4

In this chapter, we study the inverse problem of 2-D phase unwrapping, where we aim at retrieving a phase image from its modulo observations. Many applications, such as interferometry, optical deflection tomography and synthetic aperture radar imaging, are concerned by this problem since they proceed by recording complex or modulated data from which a “wrapped” phase is extracted. The 2-D phase unwrapping problem faces two main challenges that need to be addressed: noise and discontinuities in the phase image. In contrast to state-of-the-art techniques, this work aims at simultaneously unwrap and denoise the phase image. We propose a robust convex optimization approach that enforces data fidelity constraints expressed in the phase derivative domain while promoting a sparse phase prior. The resulting optimization problem is solved by the Chambolle-Pock primal-dual scheme. Using synthetic data, we show that under different observation noise levels, our approach compares favorably to those that perform the unwrapping and denoising in two separate steps. This is also illustrated for experimental data from interferometric synthetic aperture radar imaging (InSAR). The contents in this chapter are largely based on [29].

## 4.1 Introduction

The information contained in the “phase image” is essential in many applications such as optical deflection tomography (ODT – studied in Chapter 3) [37], magnetic resonance imaging [19, 136], interferometric synthetic

aperture radar (InSAR) [137, 138] and time-of-flight imaging [20]. In these techniques, the phase is not observed directly but computed from a complex or modulated data. Therefore, the measured values are wrapped in the interval  $[-\pi, \pi)$  and the observed signal presents  $2\pi$  jumps. Phase unwrapping is the procedure that allows removing these discontinuities to obtain the actual phase image.

Although the modulo operation is quite trivial, its inversion can be hard to solve. Phase unwrapping techniques need to be able to overcome, among other problems, discontinuities, noise and under-sampling of the phase.

In one dimension, the unwrapping process is straightforward since there is only one possible “path” and the phase can be recovered by a simple integration. However, this only works when the Itoh smoothness condition [139] is satisfied, *i.e.*, when the absolute value of the actual phase gradient is lower than or equal to  $\pi$ . The presence of noise or discontinuities could violate this condition, causing some unwrapping errors.

Most existing methods extend this integration principle to two dimensions (2-D) and are denominated path-following algorithms. The problem in 2-D is the error propagation when the smoothness condition is violated [138]. This occurs because the integration results depend on the chosen integration path and on the start and end points. The challenge remains in distinguishing jumps due to phase wrapping from those due to noise and discontinuities in the actual phase image. Several works have considered additional information such as pixel quality maps [140, 141] to appropriately update the integration path. However, such algorithms have some difficulties to deal with high levels of noise.

In addition to these algorithms, several efforts have been made in the development of path-independent methods. In these works, the phase unwrapping task is formulated as an inverse problem that is solved by the minimization of the error between the observed data and the acquisition model [136, 142]. In general, the acquisition model is formulated in the phase derivative domain and the error function is based on the  $\ell_p$  norm, with  $p < 2$  providing better results [138, 142]. However, this optimization problem has many possible solutions, making the unwrapping problem to be ill-posed. In order to restrict the amount of possible solutions to those that are significant

with respect to the prior information we have on the phase image, some regularization functions must be added to the optimization. Bioucas-Dias and Valadão [143] have regularized the problem by assuming that the phase image follows a Markov random field model and the resulting minimization problem is solved using graph-cut techniques. They have demonstrated that an exact phase recovery is possible if the objective function is convex and, if it is not the case, only an approximated solution can be obtained. The developed algorithm, denoted PUMA, is considered a state-of-the-art method in phase unwrapping. More recently, Kamilov *et al.* [72] have proposed an unwrapping method that relies on the minimization of a fidelity function based on the reweighted  $\ell_1$  norm and a regularization function based on the Hessian-Schatten norm, a higher-order version of the Total-Variation (TV) norm. The resulting optimization problem was solved by means of an iterative program that alternates between the phase estimation using an augmented Lagrangian scheme and the estimation of the weights of the  $\ell_1$  fidelity term using the  $\ell_2$  norm of the estimated phase.

It is important to note that the unwrapping methods described above do not deal with the presence of noise in the phase images. They tend to provide a noisy version of the phase that needs to be later denoised in order to be interpreted or further processed. In order to handle the noise, Huang *et al.* [144] proposed a method that starts with a TV-denoising step on the corrupted phase gradient and then it proceeds with the unwrapping of the denoised phase gradient using path integration techniques. Other works use unwrapping state-of-the-art techniques, such as PUMA, and they combine it with a denoising procedure that is performed either before [145] or after [146] the unwrapping step. Both approaches have proved to outperform the classical  $\ell_p$  norm minimization in the presence of noise. While the approach of denoising before unwrapping [145] deals better with high noise scenarios without discontinuities, it was reported that discontinuities are better preserved when the denoising is performed after the unwrapping [146]. In the context of time-of-flight imaging, Mei *et al.* [20] have proposed a simultaneous unwrapping and denoising technique that relies on a good characterization of the acquisition model and a wavelet sparsity prior. They show that, in such a scenario and for moderate levels of noise, the simultaneous unwrapping-denoising method has a better performance compared to those performing both steps separately.

### 4.1.1 Contributions

In this chapter, we aim at developing a general numerical reconstruction method that is able to simultaneously unwrap and denoise the observed phase. We propose a convex optimization approach based on the minimization of a sparsity prior on the phase image under an  $\ell_1$  norm fidelity constraint expressed in the phase derivative domain. The optimization also aims at retrieving the Gaussian noise component, which helps enforcing the consistency of the solution to the observations. Moreover, the algorithm is stabilized by an additional constraint forcing the first phase component to be zero. The problem is solved using the Chambolle-Pock primal-dual scheme [74]. The proposed algorithm is tested on synthetic data and the results are compared with the state-of-the-art technique PUMA whose output is denoised in a post-processing step. Although PUMA is faster and provides better results for the noiseless scenario, the proposed method provides better reconstruction quality for noisy scenarios. Finally, the proposed algorithm is shown to have a good performance on experimental InSAR data.

### 4.1.2 Outline

The rest of the chapter is organized as follows. In Section 4.2, we describe the discrete forward model related to the phase unwrapping problem. In Section 4.3, we relax the unwrapping problem and formulate it as a regularized convex optimization problem. Section 4.4 presents the numerical methods used to simultaneously unwrap and denoise the phase image by means of the proposed convex approach. In Section 4.5, some reconstruction results are shown and the proposed method is compared with a state-of-the-art algorithm in terms of noise robustness and of the amount of “wraps” in the measurements. Also, experimental InSAR data is used in order to illustrate the performance of the proposed method in a real imaging application. Finally, Section 4.6 concludes the chapter.

## 4.2 Discrete Forward Model

Our work is concerned by the reconstruction of a phase image  $\mathbf{x} \in \mathbb{R}^N$  discretized over  $N = n^2$  pixels on a  $n \times n$  regular Cartesian grid. The phase

measurement process can be defined via the centralized modulo operator, or *wrapping*,  $W$ :

$$\mathbf{y} = W(\mathbf{x} + \boldsymbol{\eta}), \quad (4.1)$$

where  $W(\lambda) := [((\lambda + \pi) \bmod 2\pi) - \pi] \in [-\pi, \pi)$  represents the component-wise wrapping,  $\mathbf{y} \in \mathbb{R}^N$  is the observed wrapped phase and  $\boldsymbol{\eta} \in \mathbb{R}^N$  stands for an additive phase sensing noise.

Because of the modulo operation, we know that the actual phase  $\mathbf{x}$  can be expressed as the sum of the observations  $\mathbf{y}$  with integer multiples of  $2\pi$ , *i.e.*,  $\mathbf{x} = \mathbf{y} + 2\pi\mathbf{k}$  for some  $\mathbf{k} \in \mathbb{Z}^N$ .

Interestingly, if there is no noise in the measurements (*i.e.*,  $\boldsymbol{\eta} = \mathbf{0}_N$ ), the image gradient can be indirectly observed through the following relation [139, 146]:

$$\mathbf{q} := W(\nabla\mathbf{y}) = \nabla\mathbf{x}, \quad (4.2)$$

where  $\nabla : \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$  is the (finite difference) gradient operator (see Section 2.B). Note that relation (4.2) is only valid when the Itoh smoothness condition is satisfied, *i.e.*, when  $|(\nabla\mathbf{x})_j| \leq \pi$  on every pixel  $j$  [139].

In the presence of low levels of noise such that the noisy phase satisfies the smoothness condition ( $|(\nabla(\mathbf{x} + \boldsymbol{\eta}))_j| \leq \pi$ ), *i.e.*, when  $|(\nabla\boldsymbol{\eta})_j| \leq \pi - |(\nabla\mathbf{x})_j|$  on every pixel  $j$ , the relation in (4.2) becomes:

$$\mathbf{q} := W(\nabla\mathbf{y}) = \nabla(\mathbf{x} + \boldsymbol{\eta}). \quad (4.3)$$

Assuming an additive white Gaussian noise model, *i.e.*,  $\eta_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$ , the Itoh condition is satisfied on all pixels if  $\|\nabla\boldsymbol{\eta}\|_\infty \leq \pi - \|\nabla\mathbf{x}\|_\infty$ . Since  $\|\boldsymbol{\eta}\|_\infty = \mathcal{O}(\sigma\sqrt{\log N})$  [147], we have  $\|\nabla\boldsymbol{\eta}\|_\infty = \mathcal{O}(\sigma\sqrt{\log N})$  and  $\|\nabla\mathbf{x}\|_\infty \leq \pi - C\sigma\sqrt{\log N}$ , for some  $C > 0$ . This shows that the Itoh condition is increasingly harder to satisfy for higher values of  $\sigma$  and  $N$ .

Note that for any  $\sigma$ , the  $\ell_2$  norm of  $\boldsymbol{\eta}$  can be bounded using the Chernoff-Hoeffding bound in (2.10):

$$\|\boldsymbol{\eta}\|_2 \leq \varepsilon_n := \sigma\sqrt{N + c\sqrt{N}},$$

which holds with high probability for  $c = \mathcal{O}(1)$ . Practically, the noise variance  $\sigma^2$  can be estimated by applying the robust median estimator (see Section 2.A) to the indirect observations  $\mathbf{q}$ .

When we encounter high levels of noise and phase discontinuities, the relation in (4.3) is not valid. The error between the indirect observations  $\mathbf{q}$  and the phase gradient occurs in the  $j$  pixels where the Itoh condition is not satisfied. Since  $j$  represents a small percentage of the total amount of pixels (increasing with  $\sigma$  and with the discontinuities), we can assume the error is sparse and it can be bounded using the  $\ell_1$  norm:

$$\|\mathbf{q} - \nabla(\mathbf{x} + \boldsymbol{\eta})\|_1 \leq \varepsilon_w, \quad (4.4)$$

where  $\varepsilon_w$  is the error bound. In this paper, we estimate this error by computing  $\|\mathbf{q} - \nabla\mathbf{x}_p\|_1$  using the reconstruction  $\mathbf{x}_p$  obtained by the PUMA method<sup>1</sup> [143]. This algorithm is known for estimating  $\mathbf{x} + \boldsymbol{\eta}$  without denoising.

### 4.3 Convex Optimization Approach

It is rather clear that the ensemble of images having the same wrapped observations  $\mathbf{y}$  forms a non-convex set. Even when  $\boldsymbol{\eta} = \mathbf{0}_N$ , if both  $\mathbf{x}_1$  and  $\mathbf{x}_2$  produce the same wrapped observation  $\mathbf{y}$ , there exist two integer vectors  $\mathbf{k}_1$  and  $\mathbf{k}_2$  such that  $\mathbf{y} = \mathbf{x}_1 + 2\pi\mathbf{k}_1 = \mathbf{x}_2 + 2\pi\mathbf{k}_2$  and obviously  $\mathbf{x}' = \lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$  does not satisfy such a relation for all  $\lambda \in [0, 1]$  since  $\lambda\mathbf{k}_1 + (1 - \lambda)\mathbf{k}_2 \notin \mathbb{Z}^M$  for most  $\lambda$ . In front of such a discrete formalism, several authors have proposed methods based on combinatorial optimization and graph-cut techniques [143, 146].

In this paper, we follow a different approach. We propose to relax the problem and, by leveraging the differential relation in (4.3), to solve it using a convex optimization formulation. Despite the wrapping operation, we expect that (4.4) holds for the phase signal  $\mathbf{x}$  given an appropriate value of  $\varepsilon_w$ .

Moreover, in order to circumvent the ill-posedness of the problem in the presence of noise and “non-Itoh” phase discontinuities, we also regularize our method by an appropriate wavelet *analysis* prior model based on the structure of the phase. More specifically, we assume that the unwrapped phase image has a sparse or compressible representation in an orthonormal basis

<sup>1</sup>MATLAB code: <http://www.lx.it.pt/~bioucas/code.htm>.

$\Psi \in \mathbb{R}^{N \times N}$ , *i.e.*, the coefficients vector  $\Psi^* \mathbf{x}$  has few important values and its  $\ell_1$  norm is expected to be small. The regularization process can then proceed by promoting a small  $\ell_1$  norm in the wavelet projection of the phase image. The rationale of this is also to prevent reconstructing fake phase jumps (since these increase locally the wavelet coefficient values) and to enforce the noise canceling. We also follow a common practice in the field which removes the (un)sparse scaling coefficients from the  $\ell_1$  norm computation.

Since both the differential fidelity and the wavelet prior are blind to the addition of a global constant, there is an ambiguity to estimate the phase up to such addition. In order to avoid this incertitude and to stabilize the convex optimization, we arbitrarily enforce the first phase component to be zero. We should note that the initial problem is itself ill-posed since, even if we solve it using (4.1) directly with a “perfect” data prior model, a “good” solution would be determined up to a global addition of a multiple of  $2\pi$ . This constraint also induces the uniqueness of the solution.

Finally, since the noise level is assumed to be known and the noise  $\ell_2$  norm to be bounded, we also propose to explicitly recover the noise part in an additive model where the unknown phase and noise are summed up to faithfully satisfy (4.4). We expect that, by explicitly recovering the noise component, the consistency of the solution is enforced [148], *i.e.*, that the estimated pair  $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\eta}})$  satisfies the following relation:

$$\mathbf{y} = W(\tilde{\mathbf{x}} + \tilde{\boldsymbol{\eta}}). \quad (4.5)$$

Gathering all these aspects, the proposed reconstruction program reads

$$\begin{aligned} & \underset{\mathbf{u}, \mathbf{v} \in \mathbb{R}^N}{\operatorname{argmin}} \quad \|\mathbf{S}_\Theta \Psi^* \mathbf{u}\|_1 \\ & \text{s.t.} \quad \|\mathbf{v}\|_2 \leq \varepsilon_n \\ & \quad \|\mathbf{q} - \nabla(\mathbf{u} + \mathbf{v})\|_1 \leq \varepsilon_w \\ & \quad u_1 = 0 \end{aligned} \quad (4.6)$$

where  $\mathbf{S}_\Theta \in \mathbb{R}^{S \times N}$  is the selection operator of the set  $\Theta$ , with  $|\Theta| = S$ , which contains the set of *detail* wavelet coefficients.

In the next section, we present the algorithm that allows solving (4.6) numerically.

## 4.4 Phase Unwrapping and Denoising Algorithm

We are interested in finding the phase candidate that minimizes (4.6), a problem that consists of the minimization of a non-smooth convex function and three constraints. The constraints can be handled by the convex indicator functions on the following convex sets:  $\mathcal{C}_1 = \{\mathbf{z} \in \mathbb{R}^N : \|\mathbf{z}\|_2 \leq \varepsilon_n\}$ ,  $\mathcal{C}_2 = \{\mathbf{z} \in \mathbb{R}^{2N} : \|\mathbf{q} - \mathbf{z}\|_1 \leq \varepsilon_w\}$  and  $\Omega = \{\mathbf{z} \in \mathbb{R}^{2N} : z_1 = 0\}$ . By forming the vector  $\mathbf{w} = (\mathbf{u}^T, \mathbf{v}^T)^T$ , the convex minimization in (4.6) can be recast as

$$\operatorname{argmin}_{\mathbf{w} \in \mathbb{R}^{2N}} \|\mathbf{S}_\Theta \mathbf{\Psi}^* \mathbf{S}_u \mathbf{w}\|_1 + \iota_{\mathcal{C}_1}(\mathbf{S}_v \mathbf{w}) + \iota_{\mathcal{C}_2}(\nabla(\mathbf{I}_N, \mathbf{I}_N) \mathbf{w}) + \iota_\Omega(\mathbf{w}), \quad (4.7)$$

where  $\mathbf{S}_u, \mathbf{S}_v \in \mathbb{R}^{N \times 2N}$  are the selection operators of the first and the last  $N$  elements of a vector in  $\mathbb{R}^{2N}$ , respectively.

The resulting phase unwrapping problem in (4.7) contains the sum of four functions belonging to  $\Gamma_0(\mathbb{R}^D)$ , for  $D \in \{N, 2N\}$ . This kind of problem can be efficiently solved using the expanded Chambolle-Pock (CP) algorithm for  $L = 3$  in (2.25) (see Section 2.4.2.2).

To match the formulation in (2.25) with the phase unwrapping problem in (4.7), we set  $F_1(\mathbf{s}_1) = \|\mathbf{s}_1\|_1$  for  $\mathbf{s}_1 \in \mathbb{R}^S$ ;  $F_2(\mathbf{s}_2) = \iota_{\mathcal{C}_1}(\mathbf{s}_2)$  for  $\mathbf{s}_2 \in \mathbb{R}^N$ ;  $F_3(\mathbf{s}_3) = \iota_{\mathcal{C}_2}(\mathbf{s}_3)$  for  $\mathbf{s}_3 \in \mathbb{R}^{2N}$ ;  $H(\mathbf{w}) = \iota_\Omega(\mathbf{w})$ ;  $\mathbf{K}_1 = \mathbf{S}_\Theta \mathbf{\Psi}^* \mathbf{S}_u \in \mathbb{R}^{S \times 2N}$ ;  $\mathbf{K}_2 = \mathbf{S}_v \in \mathbb{R}^{N \times 2N}$  and  $\mathbf{K}_3 = \nabla(\mathbf{I}_N, \mathbf{I}_N) \in \mathbb{R}^{2N \times 2N}$ . We have then  $\mathbf{K} = \operatorname{diag}(\mathbf{K}_1, \mathbf{K}_2, \mathbf{K}_3) \in \mathbb{R}^{(3N+S) \times 2N}$ .

In order to apply the algorithm in (2.37), we must compute the proximal operators of  $F_1^*$ ,  $F_2^*$ ,  $F_3^*$  and  $H$ , the first three functions being the Legendre-Fenchel conjugate of their unstarred version. The proximal operator of  $F_j^*$ , for  $j \in \{1, 2, 3\}$  is determined via the one of  $F_j$  thanks to the conjugation property [80]:

$$\operatorname{prox}_{\nu F_j^*} \zeta = \zeta - \nu \operatorname{prox}_{\frac{1}{\nu} F_j} \frac{1}{\nu} \zeta,$$

for  $\zeta \in \mathbb{R}^D$  and  $D \in \{S, N, 2N\}$ .

The proximal operator of  $F_1$  is given by the soft thresholding operator in (2.24):

$$\operatorname{prox}_{\frac{1}{\nu} F_1} \mathbf{s}_1 = \operatorname{SoftTh}\left(\mathbf{s}_1, \frac{1}{\nu}\right),$$

while the one of  $F_2$  is given by the projection onto the set  $\mathcal{C}_1$ :

$$\text{prox}_{\frac{1}{\nu}F_2} \mathbf{s}_2 = \mathbf{s}_2 \min \left( 1, \frac{\varepsilon_n}{\|\mathbf{s}_2\|_2} \right).$$

The proximal operator of  $F_3$  is given by the projection onto the set  $\mathcal{C}_2$  [51]:

$$\text{prox}_{\frac{1}{\nu}F_3} \boldsymbol{\zeta} = \mathbf{q} + \text{SoftTh}(\boldsymbol{\zeta} - \mathbf{q}, \lambda_w), \quad \boldsymbol{\zeta} \in \mathbb{R}^{2N}.$$

In the equation above, the threshold  $\lambda_w$  is equal to zero if  $\|\boldsymbol{\zeta} - \mathbf{q}\|_1 \leq \varepsilon_w$  and, otherwise, it is found by solving  $\sum_{i=1}^{2N} \max\{0, |\zeta_i - q_i| - \lambda_w\} = \varepsilon_w$ , using, for instance, the bisection method [149].

Finally, the proximal operator of  $H$  is given by the projection onto the convex set  $\Omega$ :

$$\text{prox}_{\mu H} \mathbf{w} = \text{diag}(0, 1, \dots, 1) \mathbf{w}.$$

Making use of the above computations, the expanded CP algorithm in (2.37) applied to the proposed phase unwrapping problem becomes:

$$\begin{cases} \mathbf{s}_1^{(k+1)} &= \text{prox}_{\nu F_1^*} \left( \mathbf{s}_1^{(k)} + \nu \mathbf{\Psi}^* \mathbf{S}_u \bar{\mathbf{w}}^{(k)} \right) \\ \mathbf{s}_2^{(k+1)} &= \text{prox}_{\nu F_2^*} \left( \mathbf{s}_2^{(k)} + \nu \mathbf{S}_v \bar{\mathbf{w}}^{(k)} \right) \\ \mathbf{s}_3^{(k+1)} &= \text{prox}_{\nu F_3^*} \left( \mathbf{s}_3^{(k)} + \nu \nabla (\mathbf{I}_N, \mathbf{I}_N) \bar{\mathbf{w}}^{(k)} \right) \\ \mathbf{w}^{(k+1)} &= \text{prox}_{\mu H} \left( \mathbf{w}^{(k)} - \frac{\mu}{3} \left( \mathbf{S}_u^* \mathbf{\Psi} \mathbf{s}_1^{(k+1)} + \mathbf{S}_v^* \mathbf{s}_2^{(k+1)} + (\mathbf{I}_N, \mathbf{I}_N) \nabla^* \mathbf{s}_3^{(k+1)} \right) \right) \\ \bar{\mathbf{w}}^{(k+1)} &= 2\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)} \end{cases} \quad (4.8)$$

The estimated phase, denoted  $\tilde{\mathbf{x}}_c$ , is provided by the first  $N$  components of the vector  $\mathbf{w}^{(k+1)}$ , *i.e.*,  $\tilde{\mathbf{x}}_c = \mathbf{S}_u \mathbf{w}^{(k+1)}$ . The estimated noise component, denoted  $\tilde{\boldsymbol{\eta}}$ , is provided by the last  $N$  components of the vector  $\mathbf{w}^{(k+1)}$ , *i.e.*,  $\tilde{\boldsymbol{\eta}} = \mathbf{S}_v \mathbf{w}^{(k+1)}$ .

The step-size parameters  $\mu$  and  $\nu$  in (4.8) are kept fixed through the iterations such that the condition  $\mu\nu \|\mathbf{K}\|^2 < 1$  is satisfied, *i.e.*,

$$\mu = \sqrt{\frac{0.9}{t}} \frac{1}{\|\mathbf{K}\|}, \quad \nu = t\mu,$$

with  $t = 10^6$  providing the best results. The induced norm of the operator  $(\|\mathbf{K}\|)$  is estimated using the standard power iteration algorithm [23].

In our experiments, the variables  $\bar{\mathbf{w}}^{(1)}$  and  $\mathbf{s}_j^{(1)}$ , for  $j \in \{1, 2, 3\}$ , are initialized to zero vectors. The phase component of the variable  $\mathbf{w}^{(1)}$  is initialized with the solution provided by the PUMA algorithm, while the noise component is initialized to a zero vector.

The algorithm in (4.8) stops when it achieves a stable behavior, *i.e.*, when  $\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|_2 / \|\mathbf{w}^{(k)}\|_2 \leq \text{Th}$ . Empirically, we set the value of the threshold Th to  $10^{-5}$  since it provided the best results.

The proposed phase unwrapping method is summarized in Algorithm 4.1.

---

**Algorithm 4.1** Convex phase unwrapping algorithm
 

---

**Initialization:**  $\mathbf{s}_1^{(1)} = \mathbf{0}_S, \mathbf{s}_2^{(1)} = \mathbf{S}_v \mathbf{w}^{(1)} = \mathbf{0}_N, \mathbf{s}_3^{(1)} = \bar{\mathbf{w}}^{(1)} = \mathbf{0}_{2N}, \mathbf{S}_u \mathbf{w}^{(1)} = \tilde{\mathbf{x}}_p, t = 10^6,$   
 $\mu = \sqrt{0.9t^{-1}} (\|\mathbf{K}\|)^{-1}, \nu = t\mu, \text{MaxIter} = 50 \times 10^4, \text{Th} = 10^{-5}.$

1: **for**  $k = 1$  to **MaxIter** **do**

*Compute the CP iterations (4.8):*

- 2:  $\mathbf{s}_1^{(k+1)} = \text{prox}_{\nu F_1^*} \left( \mathbf{s}_1^{(k)} + \nu \mathbf{\Psi}^* \mathbf{S}_u \bar{\mathbf{w}}^{(k)} \right)$
- 3:  $\mathbf{s}_2^{(k+1)} = \text{prox}_{\nu F_2^*} \left( \mathbf{s}_2^{(k)} + \nu \mathbf{S}_v \bar{\mathbf{w}}^{(k)} \right)$
- 4:  $\mathbf{s}_3^{(k+1)} = \text{prox}_{\nu F_3^*} \left( \mathbf{s}_3^{(k)} + \nu \nabla (\mathbf{I}_N, \mathbf{I}_N) \bar{\mathbf{w}}^{(k)} \right)$
- 5:  $\mathbf{w}^{(k+1)} = \text{prox}_{\mu H} \left( \mathbf{w}^{(k)} - \frac{\mu}{3} \left( \mathbf{S}_u^* \mathbf{\Psi} \mathbf{s}_1^{(k+1)} + \mathbf{S}_v^* \mathbf{s}_2^{(k+1)} + (\mathbf{I}_N, \mathbf{I}_N) \nabla^* \mathbf{s}_3^{(k+1)} \right) \right)$
- 6:  $\bar{\mathbf{w}}^{(k+1)} = 2\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}$

*Stop if stable behavior:*

- 7: **if**  $\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|_2 / \|\mathbf{w}^{(k)}\|_2 \leq \text{Th}$  **then** **break**.
  - 8: **end if**
  - 9: **end for**
  - 10: Return  $\tilde{\mathbf{x}}_c = \mathbf{S}_u \mathbf{w}^{(k+1)}, \tilde{\boldsymbol{\eta}} = \mathbf{S}_v \mathbf{w}^{(k+1)}$
- 

## 4.5 Experiments

In this section, we first validate the proposed convex phase unwrapping algorithm<sup>2</sup> on synthetic data by studying the quality of the unwrapped phase with respect to the amount of “wraps” in the measurements, the noise level

<sup>2</sup>MATLAB code: <http://sites.uclouvain.be/ispgroup/index.php/Softwares/HomePage>.

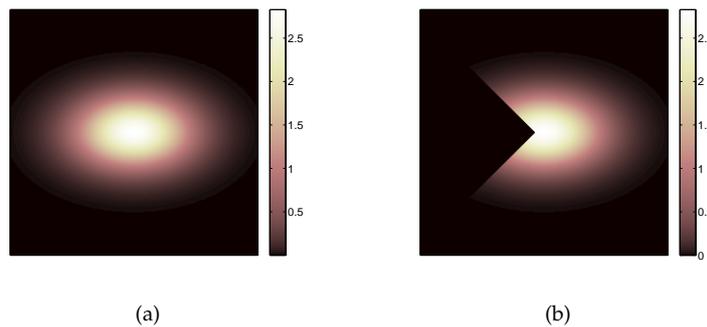
and the presence of non-Itoh discontinuities in the original phase image. After, we illustrate the behavior of the proposed algorithm on experimental data from interferometric synthetic aperture radar (InSAR) imaging.

During the experiments, the wavelet transformation used for the sparse image representation is the Daubechies wavelet basis<sup>3</sup> with seven vanishing moments and six levels of decomposition (see Section 2.C for more details).

All algorithms were implemented in MATLAB and executed on a 3.2 GHz Intel i5-650 CPU with 3.7 GiB of RAM, running a 64 Bit Ubuntu 14.04 LTS operating system.

#### 4.5.1 Synthetic Data

Two kinds of discrete synthetic phase images are selected in our experiments. They are defined on a  $256 \times 256$  pixel grid ( $N = 256^2$ ). In the first image, the phase is simulated by a 2-D Gaussian function of height  $0.9\pi$ , and standard deviations of 40 pixels horizontally and 25 pixels vertically (see Figure 4.1-(a)). In the second image, the phase is simulated by a truncated version of the 2-D Gaussian, where the image is masked by a side triangle (see Figure 4.1-(b)). By truncating the Gaussian image, we are able to simulate phase discontinuities.



**Figure 4.1** Synthetic phase images used in the experiments. (a) 2-D Gaussian phase image. (b) 2-D Truncated Gaussian phase image.

<sup>3</sup>The wavelet operator and the other operators used in this work were numerically implemented using the SPARCO framework [132].

The measurements are generated according to (4.1) with some additive white Gaussian noise  $\eta_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$  that simulates a realistic scenario. The robustness of our method is tested against three different noise levels, each characterized by a different Measurement SNR:

$$\text{MSNR} = 20 \log_{10} \frac{\|\mathbf{x}\|_2}{\|\boldsymbol{\eta}\|_2},$$

namely, 10 dB, 25 dB and  $+\infty$  dB (no noise).

Results are fairly compared with a post-denoised phase unwrapping obtained by the conjunction of the PUMA algorithm [143] with an optimal soft thresholding denoising [40] using the same wavelet basis as described above. Hereafter, the solutions of our convex approach and of the post-denoised PUMA are denoted as  $\tilde{\mathbf{x}}_c$  and  $\tilde{\mathbf{x}}_{dp}$ , respectively. Since there is no reliable estimation of the initial signal mean in any phase unwrapping method, all reconstruction qualities are measured with centralized reconstructions and ground truths, *i.e.*, by mean subtraction. After such procedure, the quality of a given reconstruction  $\tilde{\mathbf{x}} \in \{\tilde{\mathbf{x}}_c, \tilde{\mathbf{x}}_{dp}\}$  with respect to the true phase  $\mathbf{x}$ , is measured with the Reconstruction SNR:

$$\text{RSNR} = 20 \log_{10} \frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}.$$

For the behavior of the algorithm with respect to the amount of “wraps”, we analyze the Gaussian phase image in Figure 4.1-(a) and we vary its intensity by multiplying the image by a factor  $\rho \in [1, 20]$ , where  $\rho = 1$  does not cause phase wraps since  $\|\mathbf{x}\|_\infty < \pi$ . We noticed that for the noiseless scenario (MSNR =  $+\infty$  dB), the reconstruction quality is not affected by the value of  $\rho$ , since the Itoh condition is always satisfied. However, since there is no noise, PUMA outperforms the proposed method for all  $\rho$ . Table 4.1 presents a comparison for the two noisy scenarios, *i.e.*, MSNR = 25 dB and MSNR = 10 dB. The RSNR is presented for an average of 5 trials.

We can notice that the proposed convex approach (C) outperforms the denoised-PUMA (DP) for the scenarios where the Itoh condition is satisfied and, for the cases where this condition is affected by the noise corrupting the phase, C and DP provide similar results. Remark that, for  $\rho = 20$  and MSNR = 10 dB, both methods fail to provide an unwrapped phase that shares

$\rho$	RSNR[dB]			
	MSNR = 25 dB		MSNR = 10 dB	
	C	DP	C	DP
1	34.18	34.11	29.62	19.14
5	42.27	34.13	22.17	20.69
10	35.08	34.13	5.04	5.05
20	35.18	34.34	fail	fail

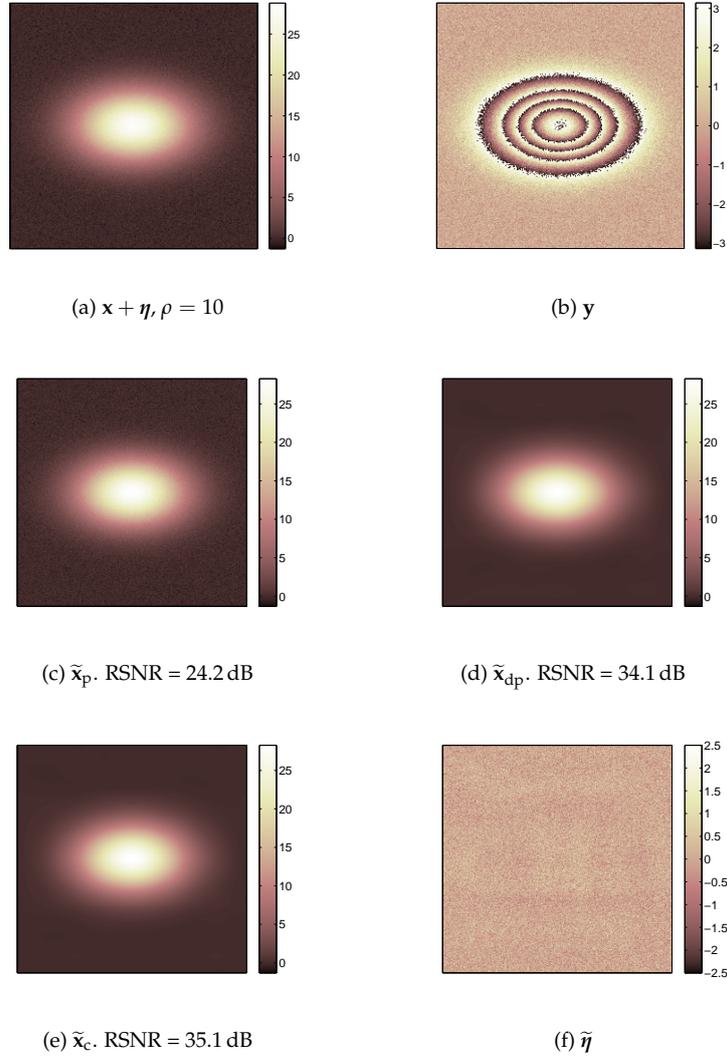
**Table 4.1** Comparison of the different RSNR obtained using denoised-PUMA (DP) and our convex approach (C) for different values of  $\rho$  on the Gaussian phase image.

any similarity with the actual phase. This happens because the high level of noise corrupting the phase image causes the wrapped phase to be pure noise, *i.e.*, the contributions due to the original phase image cannot be distinguished from the noise.

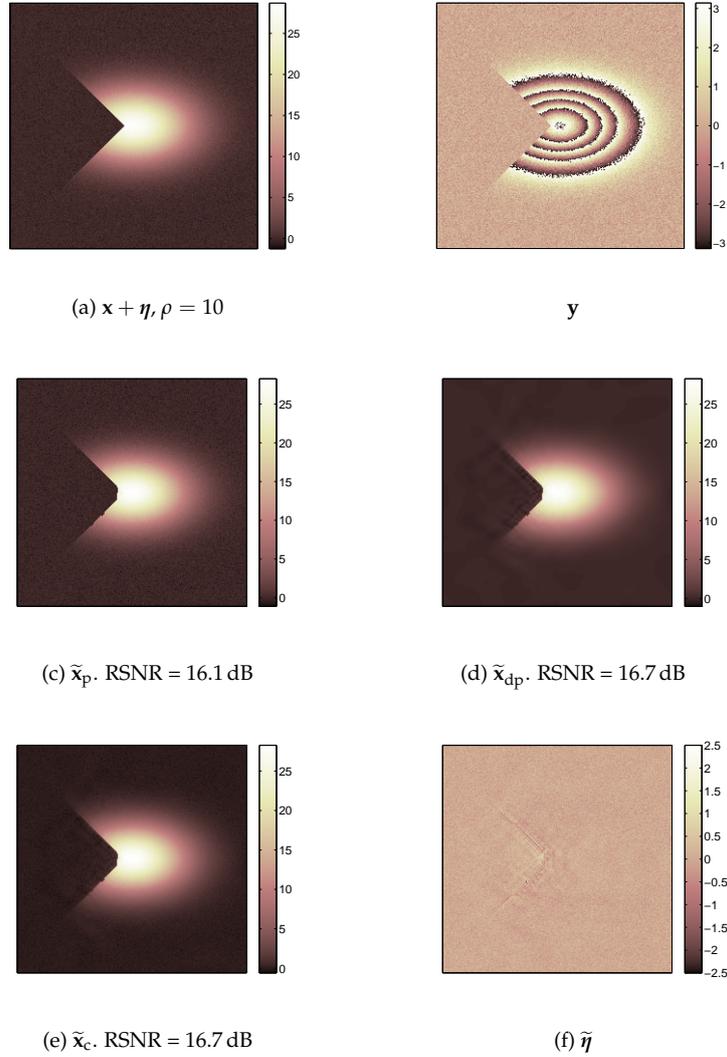
About the computational time, for the first noise scenario (MSNR = 25 dB), the convex algorithm reaches convergence for an average of 10000 iterations, which takes approximately 11 minutes. For the second noise scenario (MSNR = 10 dB), the convergence is reached for an average of 15000 iterations, which takes approximately 16 minutes. Note that the step-size parameters in (4.8) were set to a fixed value for all the iterations. To improve the convergence, these parameters can be adaptively updated using the procedure described in Section 3.7.2. Additionally, in order to reduce the computation time, we can make use of parallel computing techniques for the separable functions (*e.g.*, the proximal operators) and the algorithms or some functions can be implemented in C language instead of MATLAB.

Figures 4.2 and 4.3 depict the resulting images for the Gaussian and the Truncated Gaussian phases, respectively. Results are shown for  $\rho = 10$  and MSNR = 25dB.

For the Gaussian phase image in Figure 4.2, the convex approach compares favorably with respect to denoised-PUMA, with a difference of 1 dB in the image quality, while PUMA is able to exactly recover the noisy phase but not the original one. We can observe that the noise component recovered by the proposed convex approach is, as expected, composed of random values



**Figure 4.2** Reconstruction results corresponding to the 2-D Gaussian phase image in Figure 4.1-(a) for  $\rho = 10$ . In the upper row, we have (a) the noisy phase image for  $\text{MSNR} = 25\text{dB}$  and (b) the measurement obtained by wrapping the noisy phase in (a). In the central row, we have the phase recovered from the measurement in (b) using (c) PUMA and (d) denoised-PUMA. In the bottom row, we have (e) the phase recovered from the measurement in (b) using the proposed convex unwrapping algorithm and (f) the estimated noise.



**Figure 4.3** Reconstruction results corresponding to the 2-D truncated Gaussian phase image in Figure 4.1-(b) for  $\rho = 10$ . In the upper row, we have (a) the noisy phase image for MSNR = 25dB and (b) the measurement obtained by wrapping the noisy phase in (a). In the central row, we have the phase recovered from the measurement in (b) using (c) PUMA and (d) denoised-PUMA. In the bottom row, we have (e) the phase recovered from the measurement in (b) using the proposed convex unwrapping algorithm and (f) the estimated noise.

gathering most of the noise present in the observations. The estimated pair  $(\tilde{\mathbf{x}}_c, \tilde{\boldsymbol{\eta}})$  was verified to satisfy the consistency condition in (4.5).

The results for the truncated Gaussian phase in Figure 4.3 are less satisfactory. We can observe that both the proposed method and denoised-PUMA provide similar phase quality. However, they are not able to recover the original phase image due to the high discontinuity at the peak of the triangle. We can also note that the noise recovered by the proposed algorithm contains some structured artifacts. Hence, due to the problems in the recovery of both phase and noise components, the consistency condition in (4.5) is not satisfied in this case. One possible issue encountered when estimating the truncated Gaussian phase, is that the fidelity term in (4.4) is weakened by a poor estimation of  $\varepsilon_w$ , since PUMA is not able to perform an accurate phase unwrapping. In some applications, we can exploit the additional existing information (*e.g.*, quality and coherence maps in InSAR imaging) in order to identify possible phase discontinuities. Additionally, we believe that the wavelet sparsity prior may not be suited to handle high discontinuities in the phase image (such as the peak of the triangle), as it can be observed from the artifacts in the phase images recovered by denoised-PUMA and by the proposed convex approach. Different priors should be investigated in order to deal with the discontinuities in the phase image. This matter is part of a future study.

### 4.5.2 Experimental Data

The proposed convex unwrapping algorithm was tested on experimental data from interferometric synthetic aperture radar (InSAR) imaging. For the sake of completeness, the results presented in this section aim at illustrating the performance of the algorithm in more realistic scenarios. It does not aim, however, at a full comprehension of the InSAR acquisition process neither at a precise analysis of the recovered images.

InSAR is an imaging modality that, based on phase-difference images, aims at building accurate digital elevation models of the Earth's topography and measuring the deformations of the terrain [137, 138]. InSAR measurements are obtained by the appropriate combination of two SAR images. The latter are obtained by collecting, at a remote sensing system (*e.g.*, a radar at a satellite), the microwave field that is backscattered by the part of the Earth's

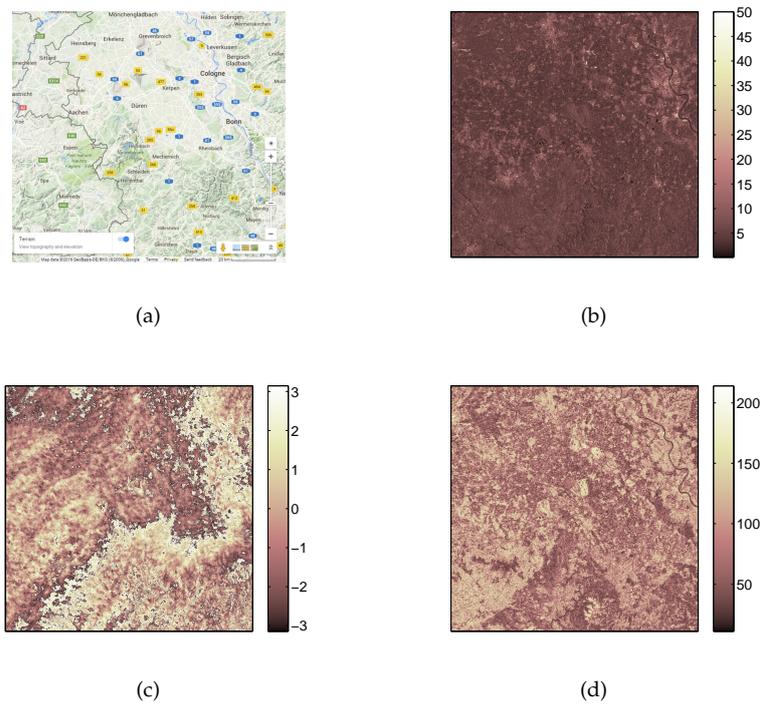
surface that is being illuminated. Each SAR image contains amplitude and phase information. The amplitude measures the radiation backscattered towards the radar, with the bright pixels representing areas of strong backscattered radiation (*e.g.*, urban areas) and the dark pixels representing areas of low backscattered radiation (*e.g.*, quiet water surfaces). The phase measures the two-way travel distance of the waves, with the bright pixels corresponding to a large distance between the scatterer and the radar (*e.g.*, flat surfaces) and the dark pixels corresponding to a small distance (*e.g.*, mountains). Since the transmitted signals are considered to be purely sinusoidal, they are periodic and the SAR phase can only measure the last fraction of the two-way travel distance that is smaller than the transmitted wavelength, *i.e.*, the phase is wrapped in the interval  $[-\pi, \pi)$ .

A SAR interferogram (also called InSAR image) is built by the combination of two SAR images acquired at slightly different view-points: by different radars mounted on the same platform or by the same radar but at different times. The InSAR image is generated using a pixel by pixel multiplication of the first SAR image and the complex conjugate of the second one. The resulting amplitude corresponds to the product of the amplitude of both images and the resulting phase corresponds to the phase difference between the images. The phase-difference information at each point of the image is proportional to the relative terrain altitude at each point and also to possible deformations of the terrain. For a more complete read on both SAR and InSAR imaging techniques, the reader may refer to [137, 138].

The data set used in this section for testing the proposed algorithm corresponds to an interferogram built by combining two SAR images captured on June 10th and 11th 1996 by the European remote sensing (ERS) satellite<sup>4</sup>. The available data, shown in Figure 4.4, is defined in a  $1024 \times 1024$  pixel grid ( $N = 1024^2$ ) and it corresponds to the Lower-Rhine-Embayment area in Germany (see Figure 4.4-(a)). All the images are oriented with the north pointing towards the top. For reference, we have the city of Cologne, Germany in the north-west (upper-right corner of the image) and the Rhine river is traversing the city. The InSAR amplitude is shown in Figure 4.4-(b) and the wrapped interferometric phase is depicted in Figure 4.4-(c).

---

<sup>4</sup>The data has been provided by the Institute of Digital Image Processing, Joanneum Research, Graz, Austria.



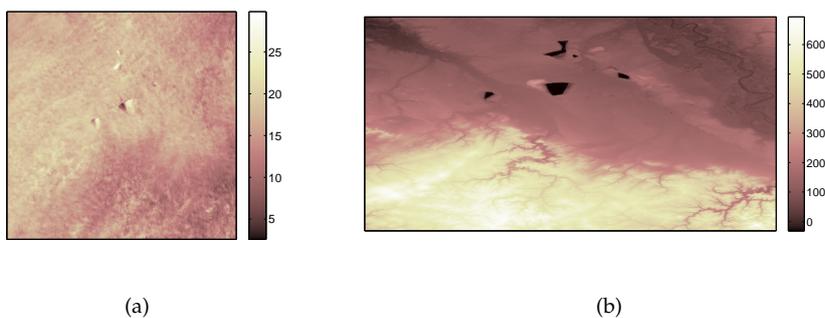
**Figure 4.4** Observation of the Lower-Rhine-Embayment area in Germany. The InSAR data was obtained by the combination of two SAR images acquired on June 10th and 11th 1996 by ERS. (a) Map of the area obtained using Google Maps. (b) Amplitude of the InSAR image (the colormap has been modified to improve the contrast in the image). (c) Phase-difference wrapped in the interval  $[-\pi, \pi)$  radians. (d) Coherence map.

In addition to the amplitude and phase, in InSAR imaging it is possible to obtain some extra information using the pair of SAR images. One example is the coherence map shown in Figure 4.4-(d). This map is estimated by means of the local coherence, which measures the cross-correlation of the SAR image pair over a small window that is moved around to cover the whole image [138]. This map provides information of the noise level in the interferometric phase image. Its values range from zero (the interferometric phase is pure noise) to one (there is no noise). For instance, in the coherence map of Figure 4.4-(d) the Rhine river is displayed in dark color, which means that the value corresponds to pure noise and it is more difficult to estimate the phase in that area. This occurs in general with water surfaces because, since they are in constant movement, they do not remain equal between the two SAR acquisitions (in this case, there is a difference of one day between the two SAR acquisitions).

The coherence map in Figure 4.4-(d) shows that the phase-difference in Figure 4.4 does not contain, in general, high levels of noise. Hence, accurate unwrapping results could be expected. Let us note that the coherence map is shown for illustration, however, it has not been used in the unwrapping procedure. The noise in the observed phase has been estimated using the robust median estimator described in Section 2.A, while the discontinuities are estimated from the unwrapped phase by PUMA. It would be interesting to include the information provided by the coherence map in the unwrapping procedure, nevertheless, this is a matter of future study.

Figure 4.5-(a) depicts the reconstruction results using the proposed convex optimization approach. The unwrapped phase is proportional to the altitude of the terrain, where bright pixels are associated with a low altitude while dark pixels are associated with high altitude terrains. The unwrapped phase is qualitatively compared to a topographic map of the area displayed in Figure 4.5-(b). Let us note that both images represent the same geographical area, nevertheless, since the exact coordinates and scale of the SAR images are not at our disposition, the images are not depicted using the same scale and orientation. Furthermore, there is a difference in the time the two images were acquired (the SAR images were acquired in 1996 while the topographic map is more recent) and in the resolution (the SAR image presents a higher resolution than the topographic map).

The two images, the unwrapped phase in Figure 4.5-(a) and the topographic map in Figure 4.5-(b), are analyzed qualitatively by observing different points and comparing the relative altitudes. For instance, the dark areas in the center of the topographic map, are man-made holes that correspond to mining areas in the district of Düren, Germany. Since the larger mining area is the deepest one, it corresponds to the region in the map with the lowest altitude. In the unwrapped phase, this area is shown at the center of the image and contains the brightest pixels. We can observe that, at the left of this bright region we can also find the darkest values of the unwrapped phase. Since this observation does not match the reality in the topographic map, we believe that it corresponds to an artifact either coming from the observations or resulting from the unwrapping algorithm. The presence of this artifact can be explained by the abrupt changes in altitude that are found in that part of the mining area. The other mining areas, which are not so deep, can also be appreciated in the unwrapped phase. In the lower-left corner of the unwrapped phase, we have a dark area of pixels that corresponds to the hilly area in the south of the topographic map. The rest of the unwrapped phase image has a relatively bright intensity, which corresponds to the flat areas in the topographic map where the cities are located.



**Figure 4.5** (a) Unwrapped phase using the proposed approach. The scale in the image corresponds to the phase and the colorbar is indicated in radians. The phase is proportional to the altitude, where the bright pixels correspond to low altitude terrains and the dark pixels correspond to high altitude terrains. (b) Topographic map of the Lower-Rhine-Embayment area obtained from <http://topographic-map.com/>. The scale in the image corresponds to the altitude of the terrain and the colorbar is indicated in meters.

## 4.6 Conclusion

We proposed a general convex optimization approach for robust phase unwrapping. In contrast to state-of-the-art techniques, the proposed approach aims at simultaneously unwrap and denoise the phase image. The proposed approach is shown to outperform the post-denoised PUMA for those scenarios where the noisy phase is smooth enough to satisfy the Itoh condition. When such condition is violated due to a high noise level or discontinuities in the phase image, the algorithm presents the same quality as the denoised-PUMA. As explained in more detail in Chapter 6, we believe that a careful attention should be paid to the prior model, since the wavelet sparsity prior may be preventing the reconstruction of discontinuities in the phase image. Additionally, future studies should include additional information on the phase image in order to form a better prior or to estimate the noise and the discontinuities.

By performing a joint estimation of the phase and the noise components, we aimed at enforcing the consistency between the solution and the wrapped observations. However, in the cases where the discontinuities in the phase image prevented a high quality reconstruction of both components, we found that the solution is not consistent with the observations. This issue needs to be further investigated in future research work.

In this chapter, we have studied the phase unwrapping problem, an inverse problem encountered in many imaging applications (*e.g.*, ODT, InSAR) where the observation corresponds to the modulo of the actual image. Besides the modulo operation, there exist other types of distortion that are also originated at the acquisition. In the following, we are interested in a type of distortion that is caused by the characteristics of the imaging device, where instead of measuring a discrete version of the actual scene, we obtain a function of the actual scene that is related to the response of the imaging device (*e.g.*, telescope, microscope, camera) through a convolution operation. This kind of distortion is frequently encountered in several imaging applications, such as astronomical imaging, microscopy and positron emission tomography. In the next chapter, we study this problem, more specifically, we investigate the estimation of the instrument response and the correction and enhancement of the observed images using blind deconvolution techniques.



# Blind Deconvolution

# 5

In this chapter, we present an inverse problem ubiquitous in several imaging applications (*e.g.*, microscopical and astronomical imaging, photo enhancement, ...), where the image measured by the real optical instrument is given by the convolution of the ideal image and the response of the instrument (also called point spread function – PSF). Additionally, the image acquisition process is also contaminated by other sources of noise (*e.g.*, read-out and photon-counting). Our focus in this chapter lays in the astronomical imaging application, where the optical characterization of instrumental effects is of high importance in order to extract accurate physical information from the observations. Nevertheless, the proposed techniques can be extended to other imaging applications.

The problem of estimating both the PSF and the true image, called blind deconvolution, is ill-posed. Hence, we propose a blind deconvolution scheme that relies on image regularization. Contrarily to most methods presented in the astronomical imaging literature, our method does not assume a parametric model of the PSF and can thus be applied to any telescope. Our scheme uses a wavelet *analysis* prior model on the image and weak assumptions on the PSF. We use observations from a celestial transit, where the occulting body can be assumed to be a black disk. These constraints allow us to retain meaningful solutions for the filter and the image, eliminating trivial, translated and interchanged solutions. Under an additive Gaussian noise assumption, they also enforce noise canceling and avoid reconstruction artifacts by promoting the whiteness of the residual between the blurred observations and the cleaned

data. Our method is applied to synthetic and experimental data. The PSF is estimated for the SECCHI/EUVI instrument using the 2007 Lunar transit, and for SDO/AIA using the 2012 Venus transit. Results show that the proposed non-parametric blind deconvolution method is able to estimate the core of the PSF with a similar quality to parametric methods proposed in the literature. We also show that, if these parametric estimations are incorporated in the acquisition model, the resulting PSF outperforms both the parametric and non-parametric methods. The contents in this chapter are largely based on [30].

## 5.1 Introduction

Deconvolution is an ubiquitous data processing method that arises in a variety of applications, *e.g.*, biomedical imaging [150], astronomy [25, 151], remote sensing [152] and video and photo enhancement [14]. Formally, this problem amounts to recovering a signal  $\mathbf{x}$  from *blurred* and *corrupted* observations  $\mathbf{y}$ :

$$\mathbf{y} = \Theta(\mathbf{h} \otimes \mathbf{x}), \quad (5.1)$$

where  $\otimes$  stands for the *convolution* operation between  $\mathbf{x}$  and some other signal  $\mathbf{h}$ , while  $\Theta$  is a general function accounting for some corrupting noise, *e.g.*,  $\Theta(\mathbf{u}) := \mathbf{u} + \boldsymbol{\eta}$  under an additive corruption model with some (Gaussian) noise  $\boldsymbol{\eta}$ .

For instance, when a scene is captured by an optical instrument, the observation is a blurred or degraded version of the original image, corrupted by noise and by some effects due to the instrument (motion, out-of-focus, light scattering, ...). In such a case, the imaging system is usually assumed to be well represented by the sensing model (5.1) where the ideal image  $\mathbf{x}$  is blurred by a point spread function (PSF)  $\mathbf{h}$ . Implicitly, the PSF describes the response of the imaging device to a Dirac point source, *i.e.*, the impulse response of the instrument, while  $\Theta$  can distinguish different sources of noise contaminating the image: read-out, photon counting, multiplicative and compression noise, among others.

If the true PSF can be determined in advance, then it is possible to recover the original, undistorted image by convolving the acquired image with the in-

verse PSF. This is the so called “known-PSF deconvolution.” In most practical applications, however, finding the true PSF is impossible and an approximation must be made. As the acquired image is corrupted by various sources of noise, both the PSF and a denoised version of the image should be recovered together. This process is known as “blind deconvolution.”

Since there are infinite combinations of images and filters that are compatible with the distorted observations, the blind deconvolution problem is severely ill-posed. One way to reduce the number of unknowns is to introduce a forward (parametric) model of the PSF. For instance, Oliveira *et al.* [48] have modeled the motion blur by a straight line, where the number of unknowns is reduced to two: the line length and angle. Babacan *et al.* [49] have modeled smoothly varying PSFs by using a simultaneous autoregressive prior, where the only parameter to estimate is the variance of a Gaussian function. However, such methods are limited to specific applications as they can only recover the expected model of the PSF, excluding the estimation of a generic non-parametric PSF. Furthermore, due to the ill-posedness of the blind deconvolution problem, a slight mismatch between the specified model and the true PSF can lead to poorly deconvolved images.

Blind-deconvolution is also very sensitive to the noise present in the observations. While some works [153, 154] have used fast and efficient techniques such as a simple inverse filtering to recover both the PSF and the undistorted image from the blurred observations, these methods present low tolerance to noise [155].

Robust blind deconvolution methods, on the other hand, optimize a data fidelity term, which is based on the acquisition model, stabilized by some additional regularization terms. If the observations are corrupted by an additive Gaussian noise, the deconvolution problem can be solved by a regularized least-squares (LS) minimization [15]. In the presence of Poisson noise, the problem can be formulated using the Kullback-Leibler divergence as the data fidelity term [25, 34, 155], which represents a more complex function to minimize. To avoid dealing with the Kullback-Leibler divergence, the Poisson corrupted data can also be handled through a Variance Stabilization Transform (VST) [16, 53, 151]. The VST provides an approximated Gaussian noise distributed data and thus allows working with a regularized LS formulation.

In this chapter, blind deconvolution is used to recover both the PSF and the undistorted image from blurred observations acquired by solar extreme ultraviolet (EUV) telescopes. We use the proximal alternating minimization method described in Section 2.4.3 in the context of additive Gaussian noise. This algorithm allows handling LS problems for a large class of regularization functionals. The advantage of this method over usual alternating minimization approaches, *e.g.*, [15, 103, 155], is that it provides theoretical convergence guarantees and it is general enough to include a wide variety of prior information. To our knowledge, it is the first time that blind deconvolution is solved using the proximal algorithm of Section 2.4.3.

In optical telescopes, the PSF originates from various instrumental effects, such as: optical aberrations (spherical, astigmatism, coma), diffraction (produced by, *e.g.*, an entrance filter mesh), scattering (from, *e.g.*, micro-roughness of the mirrors resulting in long-range diffuse illumination), charge spreading, etc. All these effects “spread” light, *i.e.*, incident photons which would otherwise focus to a single point of the focal plane, may get detected at a location that is slightly shifted or even far away. This does affect different types of measurements, such as the photometry of fine features (see, *e.g.*, the work of DeForest *et al.* [9]).

The PSF of solar telescopes is usually modeled using pre-flight instrument specifications [18, 156]. After building a specific model for a given instrument, few parameters need to be estimated in order to recover the PSF. For instance, Gburek *et al.* [154] fitted the central pixels of the TRACE PSF to a Moffat function, while DeForest *et al.* [9] modeled the long-range effect of the TRACE PSF as the sum of a measured diffraction pattern with a circularly symmetric scattering profile. The PSFs from the four EUVI instrument channels on STEREO-B were studied by Shearer *et al.* [16, 151], where the long-range scattering effect was assumed to follow a parametric piece-wise power-law model. A similar method was used to estimate the PSF of the SWAP instrument on board the PROBA2 satellite [157]. Finally, Poduval *et al.* [17] provided a semi-empirical model for the Atmospheric Imaging Assembly (AIA) onboard SDO, similar to the one used by DeForest *et al.* [9] for TRACE. In all these works, the complete set of parameters that were fitted are in the range of five to eleven values, just a few compared to the resolution of the PSF (millions of pixels).

In addition to building a parametric model of the PSF using the specifications of the instruments, some works [9, 16, 17] have used the information from a celestial transit to help inferring the PSF of a given instrument. A celestial transit, shown in Figure 5.1, is an astronomical event where the Moon or a planet move across the disk of the sun, hiding a part of it, as seen from the telescope. Since the EUV emission of transiting celestial bodies is zero, any apparent emission is known to be caused by the instrument. Therefore, this type of event provides strong prior information that allows the regularization of the blind deconvolution problem.



**Figure 5.1** Accumulation of observations of the Venus transit captured by SDO/AIA on June 5th - 6th 2012.

Let us note that all these parametric methods present the disadvantage of being specific for a given instrument and depending on a good characterization of the PSF, which is not always possible [158].

### 5.1.1 Contribution

Central to this work is the information provided by the transit of a celestial object whose apparent boundary on the recorded image can be predicted with high (sub-pixel) accuracy. This is typically the case for the observation of Moon or Venus transits for which ephemeris allows us to precisely know their apparent boundaries at the recording time, provided of course that (i) the small variations of the object topography around a perfect disk (*e.g.*, mountains, craters) are smaller than a pixel width, and (ii), that such a transiting object has no atmosphere that could blur its apparent boundary (*e.g.*, as for an Earth transit). As will be clear below, these hypotheses have been respected for at least two cases: for the instrument SECCHI/EUVI during the 2007 Lunar transit and for SDO/AIA during the 2012 Venus transit. For these two

situations, and for any other transit respecting the conditions above, we know a priori the values for a set of pixels in the image and also their exact location. Our work uses such information as constraint in the function to be optimized for blindly deconvolving images. Notice that our approach could also be applied to other situations where pixel values and exact location are known a priori (*e.g.*, dark calibration patterns in computer vision applications and anatomical landmarks in medical imaging applications).

Contrarily to previous works in solar physics, the blind deconvolution method demonstrated in this chapter is not based on a specific model of the PSF, but infers it from the observed data. This allows the method to be used for estimating the PSF of any instrument provided it has a prior information on a set of pixel values. Since no model is imposed on the PSF, the large number of unknowns makes the problem computationally intractable. We thus focus on the estimation of the PSF core, which is defined as the central pixels encompassing at least 99% of the total PSF energy. Note that this thresholding follows the radial energy characterization of the SDO/AIA PSF in the work of Poduval *et al.* [17] and a similar study of the available PSFs of SECCHI/EUVI (*e.g.*, the work of Shearer *et al.* [16]). The PSF core accounts for charge spreading, optical aberrations, and diffraction effects. Note that the same algorithm would be able to handle a parametric model of the PSF provided a precise instrument characterization is available. For simplicity, we consider an average additive Gaussian noise model that gathers all sources of noise present in the observation.

The proposed method is first tested through simulated realistic scenarios. Results show the ability to estimate a given PSF with high quality. We also show the importance of considering multiple transit observations in order to provide a better conditioning to the filter estimation problem.

We validate our method on AIA and EUVI observations using solar transits. The validation of the recovered filters is based on the reduction of the celestial body apparent emissions. The estimated PSFs were also tested for images containing active regions. Results show that, when recovering the PSF core, the proposed non-parametric scheme can achieve similar results to parametric methods. Also, we consider the case where the parametric PSF is incorporated in the imaging model as an additional filter convolving the actual

image [16, 151]. We show that the resulting parametric/non-parametric PSF provides better results than both parametric and non-parametric PSFs estimated independently. Let us note that, since parametric methods are based on the physics of the instrument, they tend to provide a more precise PSF containing both the core and the diffraction peaks. However, when the telescope presents unknown or unexpected behavior, an accurate model of the filter cannot be built. In such cases, we require non-parametric models as the one presented in this chapter, which due to their flexibility, can handle better some properties of the instrument.

### 5.1.2 Outline

The rest of the chapter is organized as follows. In Section 5.2, we describe the discrete forward model and the noise estimation. In Section 5.3, we present the formulation of the blind deconvolution problem based on the image and filter available prior information. Section 5.4 describes the alternating minimization method used for estimating the filter and the image. It briefly presents the method, including an analysis of its convergence, the initialization, the parameter estimation, the stopping criteria and the numerical reconstruction. In Section 5.5, we present a non-blind deconvolution method that is used for the experimental validation. Finally, Section 5.6 presents the results on both synthetic and experimental data for the AIA and EUVI telescopes, and Section 5.7 concludes the chapter.

## 5.2 Problem Statement

The telescope imaging process can be mathematically modeled as the instrument's PSF  $\mathbf{h}$  convolving the true image  $\mathbf{x}$ , *i.e.*,  $\mathbf{h} \otimes \mathbf{x}$ . The convolution operation, represented by  $\otimes$ , consists of integrating portions of the actual scene weighted by the PSF. In the following, we consider a discrete setting where the convolution integration is represented by a discrete sum.

### 5.2.1 Discrete Model

Let us consider that, during the transit, the EUV telescope acquires a set of  $P$  images containing a celestial body. In order to limit the computation time,

the effective field-of-view (FoV) of each observation is restricted to an image patch centered on the transiting celestial body with a size several times bigger than the body apparent diameter. The effects of this FoV truncation will be discussed in detail in Section 5.3 and Section 5.6. In this chapter, all  $n \times n$  images are represented as  $N = n^2$  vectors so that linear transformations of images are identifiable with matrices. Each observed patch is a collection of values gathered in a vector  $\mathbf{y}_j \in \mathbb{R}^N$ , with  $j = 1, \dots, P$ . The observed values are modeled as a “true” image  $\mathbf{x}_j \in \mathbb{R}^N$  convolved by the instrument’s PSF ( $\mathbf{h} \in \mathbb{R}^N$ ). The PSF is assumed to be spatially and temporally invariant, thus is the same for each observed patch. For simplicity, we consider that all sources of noise present in the observation can be gathered in an average noise model, which is assumed to be additive, white and Gaussian. The observations are then assumed to be affected by an additive white Gaussian noise (AWGN),  $\boldsymbol{\eta}_j \in \mathbb{R}^N$ , with  $(\boldsymbol{\eta}_j)_i \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$ . The acquisition process of the telescope is thus modeled as follows:

$$\mathbf{y}_j = \mathbf{h} \otimes \mathbf{x}_j + \boldsymbol{\eta}_j. \quad (5.2)$$

This model assumes that the observed images  $\mathbf{y}_j$  are uncompressed and have been corrected for CCD effects (dark current removal, flat fielding, despiking). In the case of SDO/AIA, this corresponds to level 1 data processing.

In 1-D, the discrete circular convolution  $\sum_{k=1}^n u_k g_{i-k+1}$  of a vector  $\mathbf{u} \in \mathbb{R}^n$  with a filter  $\mathbf{g} \in \mathbb{R}^n$  can always be described by the product of  $\mathbf{u}$  with a circulant matrix

$$\Phi(\mathbf{g}) \begin{pmatrix} g_1 & g_2 & \cdots & g_n \\ g_n & g_1 & \cdots & g_{n-1} \\ \vdots & & & \vdots \\ g_2 & g_3 & \cdots & g_1 \end{pmatrix} \in \mathbb{R}^{n \times n},$$

*i.e.*, a matrix where every row is a right cyclic shift of the *kernel*  $\mathbf{g}$  [108]. In 2-D and in particular for the model (5.2), the convolution of  $\mathbf{x}_j$  by  $\mathbf{h}$  can also be represented by the multiplication of  $\mathbf{x}_j$  by a matrix  $\Phi(\mathbf{h}) \in \mathbb{R}^{N \times N}$  that is now *block-circulant with circulant blocks*, *i.e.*, a matrix made of  $n \times n$  blocks, each block being a  $n \times n$  circulant matrix representing the action of one row of the 2-D filter  $\mathbf{h}$  in (5.2).

If we gather the  $P$  “true” images in a single matrix  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_P) \in \mathbb{R}^{N \times P}$ , the acquisition model can be transformed into the

following matrix form:

$$\mathbf{Y} = \Phi(\mathbf{h})\mathbf{X} + \mathbf{N}, \quad (5.3)$$

with  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_P) \in \mathbb{R}^{N \times P}$  and  $\mathbf{N} = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_P) \in \mathbb{R}^{N \times P}$  two matrices that similarly gather the  $P$  observed images and their noises, respectively.

### 5.2.2 Noise Estimation

For the sake of simplicity and to keep fast numerical methods, the model (5.3) implicitly considers an additive noise model. A way to generalize our method to Poisson noise would be to include a VST [16, 38, 53] in the acquisition model in (5.3), both on the observations  $\mathbf{Y}$  and on the convolution result  $\Phi(\mathbf{h})\mathbf{X}$ .

Under the AWGN assumption, the energy of the residual noise is known to be bounded using the Chernoff-Hoeffding bound in (2.10):

$$\|\mathbf{Y} - \Phi(\mathbf{h})\mathbf{X}\|_{\mathbb{F}}^2 = \|\mathbf{N}\|_{\mathbb{F}}^2 < \varepsilon^2 := \sigma^2 \left( NP + c\sqrt{NP} \right), \quad (5.4)$$

with high probability for  $c = \mathcal{O}(1)$ . A first guess of the noise variance  $\sigma^2$  can be estimated using the Robust Median Estimator ( $\sigma_{RME}^2$ ) described in Section 2.A.

Note that the variance could also be obtained from the physical noise characteristics using, for instance, dark areas in the images to estimate the dark noise. However, such computations may not consider all the noise sources present in the observations, providing an underestimation of the actual variance.

In this work, we prefer to adopt another strategy. Since the bound in (5.4) considers an average noise model and since the model (5.3) is an AWGN approximation of the actual data noise corruption, we aim to find an adaptive value of  $\sigma$  that optimizes the AWGN assumption, *i.e.*, which maximizes somehow the proximity of this simple model to the true corruptions. Section 5.6.2 describes in detail this adaptive strategy and demonstrates its advantage over the fixed choice  $\sigma = \sigma_{RME}$ .

### 5.3 Blind Deconvolution Problem

In this section, we aim at reconstructing both  $\mathbf{X}$  and  $\mathbf{h}$  from the noisy observations  $\mathbf{Y}$ . Since the data is corrupted by an AWGN, we can estimate the image and the filter using a least squares minimization, *i.e.*, by finding the values that minimize the energy of the noise:

$$\min_{\bar{\mathbf{X}}, \bar{\mathbf{h}}} \frac{1}{2} \|\mathbf{Y} - \Phi(\bar{\mathbf{h}})\bar{\mathbf{X}}\|_{\mathbb{F}}^2, \quad (5.5)$$

with  $\bar{\mathbf{X}} \in \mathbb{R}^{N \times P}$  and  $\bar{\mathbf{h}} \in \mathbb{R}^N$ .

The blind deconvolution problem in (5.5) is ill-posed and can have infinite possible solutions. By regularizing this problem we aim to reduce the amount of possible solutions to those that are meaningful. In the coming sections, we first describe the prior information and constraints on the image and the filter and then use this information to formulate a regularized blind deconvolution problem.

#### 5.3.1 Prior Information on the Image

In the following, we present in detail the available prior information on the image candidate.

**(a) Zero disk intensity:** Each image of the transit contains the observation of a celestial body, Venus or the Moon. For each observation  $\mathbf{y}_j$ , we know that the celestial bodies' EUV emission is zero and therefore can be represented as a black disk of constant radius  $R$  and center  $c_j$ , both being known from external astronomical observations: [10] for SDO/AIA and (J.-P. Wuelser *private communication*) for SECCHI/EUVI. The set of pixels of  $\mathbf{x}_j$  inside the disk is denoted by  $\Omega_j$ . In the minimization problem we can set to zero the pixels of the  $j^{\text{th}}$  image that are inside the set  $\{\Omega_j\}$ , *i.e.*,  $X_{ij} = (\mathbf{x}_j)_i = 0$  for all  $i \in \Omega_j$ . This prior information is crucial in the regularization of the blind deconvolution problem as it allows us to remove the issue of interchangeability between the filter and the image. Furthermore, it prevents “oppositely” translated solutions of the image and the filter, a problem occurring because the convolution process is blind to such translations, *i.e.*,  $\mathbf{h} \otimes \mathbf{x}_j = T_{\mathbf{a}}T_{-\mathbf{a}}(\mathbf{h} \otimes \mathbf{x}_j) = (T_{\mathbf{a}}\mathbf{h}) \otimes (T_{-\mathbf{a}}\mathbf{x}_j)$ , with  $T_{\mathbf{a}}$  a translation by  $\mathbf{a} \in \mathbb{Z}^2$ .

**(b) Analysis-based sparsity model:** As commonly done with ill-posed inverse problems associated to image restoration tasks [32, 40, 60], we regularize our method with a 2-D wavelet prior in *analysis* (see Section 2.3.2 for more details). As a dictionary, we use the system associated to the Undecimated Discrete Wavelet Transform (UDWT) described in Section 2.C. In this context, we assume that, if we represent the image candidate on a redundant wavelet dictionary  $\Psi \in \mathbb{R}^{N \times W}$  made of  $W$  vectors in  $\mathbb{R}^N$ , the wavelet coefficients are sparse, *i.e.*, the coefficients matrix  $\Psi^* \mathbf{X}$  has few important values and its  $\ell_1$  norm (computed over all its entries) is expected to be small. However, the wavelet coefficients with support touching the boundary of the Moon or Venus disk are not sparse. Hence, we do not consider in this prior model those coefficients that are affected by the occulting body. Also, we follow a common practice in the field which removes the (un)sparse scaling coefficients from the  $\ell_1$  norm computation (see Section 2.C). The set of *detail* coefficients that are not affected by the black disk is denoted by  $\Theta$ , with  $|\Theta| = Q$ . See Section 5.6.1 for further details on how to compute the set  $\Theta$ . The matrix  $\mathbf{S}_\Theta \in \mathbb{R}^{Q \times W}$  is the corresponding selection operator that extracts from any coefficients vector  $\mathbf{u} \in \mathbb{R}^W$  only its components indexed by  $\Theta$ . The global prior information can thus be exploited by promoting a small  $\ell_1$  norm on the wavelet coefficients belonging to  $\Theta$ . The rationale of this prior is to enforce noise canceling and avoid artifacts in the reconstructed image.

**(c) Image non-negativity:** Note that the image corresponds to a measure of a photon emission process, therefore, its pixel values must be non-negative everywhere.

### 5.3.2 Constraints on the PSF

Since the filter is not known a priori and it changes from one instrument to another, we are interested in adding soft constraints that are common to most solar EUV telescopes. We are not aiming at reconstructing all of the PSF's components but rather at estimating its core, which is responsible for most of the diffused light.

We first consider that, since the PSF corresponds to an observation of a point, it is non-negative everywhere. Additionally, by assuming that the amount of light entering in the instrument is preserved, the  $\ell_1$  norm

of the filter candidate must be equal to one. This is explained by taking the sum over the pixels of one observed patch in (5.3), which is equal to  $\mathbf{1}^T \mathbf{y}_j = \mathbf{1}^T \Phi(\mathbf{h}) \mathbf{x}_j$  in a noiseless process. Since the filter is non-negative, taking its  $\ell_1$  norm equal to one is equivalent to  $\mathbf{1}^T \Phi(\mathbf{h}) = \mathbf{1}^T$ , which results in  $\mathbf{1}^T \mathbf{y}_j = \mathbf{1}^T \mathbf{x}_j$ , *i.e.*, the light is preserved. Therefore, the filter candidate must belong to the Probability Simplex [51], defined as  $\mathcal{PS} = \{\mathbf{h} \in \mathbb{R}^N : h_i \geq 0, \|\mathbf{h}\|_1 = 1\}$ .

One important assumption in our proposed method is that the filter candidate is of size  $(2b + 1) \times (2b + 1)$ , for any  $b \in \mathbb{N}$ , and is centered in the spatial origin of the discrete grid. This means that the filter candidate has a limited support inside a set  $\Lambda$  of size  $(2b + 1)^2$ , *i.e.*, the filter has only important values in the central pixels and is negligible beyond the considered support. This allows us to work with a patch (only a part of the observation) and not with the complete FoV of the instrument in order to reduce the computation time.

EUV images have a high dynamic range. Hence, due to long-range effects, a given pixel value can be affected by the value of another high intensity pixel located as far as 100 arcsec away for SDO/AIA [17] and 1500 arcsec away for SECCHI/EUVI [16]. In such cases, despite a PSF that can rapidly decay far from its center, the high intensity pixel can induce a long-range effect that is not taken into account by a filter with truncated support. Let us define the complementary set of  $\Lambda$  as  $\Lambda^c = [N] \setminus \Lambda$ , with  $|\Lambda^c| = N - (2b + 1)^2$ . We assume that the actual PSF is composed by two different filters: the PSF core with support on  $\Lambda$ , denoted by  $\mathbf{h}_\Lambda$ , which accounts for short-range effects, and another one, denoted  $\mathbf{h}_{\Lambda^c}$  with support on  $\Lambda^c$ , accounting for long-range effects, *i.e.*,  $\mathbf{h} = \mathbf{h}_\Lambda + \mathbf{h}_{\Lambda^c}$ . An accurate estimation of the long-range PSF is out of the scope of this work. Its effect inside the disk of the celestial body is modeled as a constant  $\gamma$ , *i.e.*,  $(\mathbf{h}_{\Lambda^c} \otimes \mathbf{x}_j)_i \approx \gamma, \forall i \in \Omega_j$  (see Appendix 5.A for more details). For a given patch, the acquisition model in (5.2) can then be approximated as:  $\mathbf{y}_j = (\mathbf{h}_\Lambda + \mathbf{h}_{\Lambda^c}) \otimes \mathbf{x}_j = \mathbf{h}_\Lambda \otimes \mathbf{x}_j + \gamma \mathbf{1}_N$ . In this work, we aim at estimating only the PSF core, which is called  $\mathbf{h}$  hereafter. The effect of the long-range PSF, denoted by  $\gamma$ , is computed in a preprocessing stage by the average value inside the center of several observed transit disks (see Section 5.6.2). This simple estimation is motivated by the presence of a systematic intensity background at the center of each observed transit disk. This one is

quite independent of the Sun activity around such disk and its mean intensity (few tens of DNs) is also far above the estimated noise level (few DNs).

*Remark 5.1.* Note that our method could allow the addition of other convex constraints, e.g., if we know that the filter is sparse or that  $0 \leq h_i \leq g_i$ , for some upper bound  $g_i$  on  $h_i$  such as a specific power law decay. However, we will not consider this possibility here as we want to stay as agnostic as possible on the properties of the filter to reconstruct.

### 5.3.3 Final Formulation

From Sections 5.3.1 and 5.3.2, we can formulate the following regularized blind deconvolution problem:

$$\begin{aligned} \min_{\bar{\mathbf{X}}, \bar{\mathbf{h}}} \quad & \rho \|\mathbf{S}_\Theta \Psi^* \bar{\mathbf{X}}\|_1 + \frac{1}{2} \|\mathbf{Y} - \Phi(\bar{\mathbf{h}}) \bar{\mathbf{X}} - \gamma \mathbf{1}_N \mathbf{1}_P^T\|_{\mathbb{F}}^2 \\ \text{s.t.} \quad & (\bar{\mathbf{x}}_j)_i = 0 \text{ if } i \in \Omega_j; (\bar{\mathbf{x}}_j)_i \geq 0 \text{ otherwise} \\ & \bar{\mathbf{h}} \in \mathcal{PS}; \text{supp } \bar{\mathbf{h}} = \Lambda \end{aligned} \quad (5.6)$$

with  $\bar{\mathbf{X}} \in \mathbb{R}^{N \times P}$  and  $\bar{\mathbf{h}} \in \mathbb{R}^N$ . The term  $\gamma \mathbf{1}_N \mathbf{1}_P^T$  cancels the long-range part of the filter. The regularization parameter, denoted by  $\rho$ , controls the trade-off between the sparsity of the image projection in a wavelet dictionary and the fidelity to the observations. This essential parameter is estimated in Appendix 5.C.1.

*Remark 5.2.* It is important to note that the prior information related to the darkness of the occulting body ( $(\bar{\mathbf{x}}_j)_i = 0$  if  $i \in \Omega_j$ ), which is crucial in the regularization of the blind deconvolution problem, is taken into account in the first constraint of the problem (5.6).

The constraints in (5.6) can be handled by the convex indicator functions on the following convex sets:  $\mathcal{P}_0 = \{\mathbf{U} \in \mathbb{R}_+^{N \times P} : (\mathbf{u}_j)_i = 0 \text{ if } i \in \Omega_j\}$ ; and  $\mathcal{D} = \{\mathbf{v} \in \mathbb{R}_+^N : \|\mathbf{v}\|_1 = 1, \text{supp } \mathbf{v} = \Lambda\}$ . The problem of estimating the image and the filter in (5.6) can then be recast as

$$\{\tilde{\mathbf{X}}, \tilde{\mathbf{h}}\} = \underset{\bar{\mathbf{X}}, \bar{\mathbf{h}}}{\operatorname{argmin}} \{L(\bar{\mathbf{X}}, \bar{\mathbf{h}}) = \rho \|\mathbf{S}_\Theta \Psi^* \bar{\mathbf{X}}\|_1 + \frac{1}{2} \|\mathbf{Z} - \Phi(\bar{\mathbf{h}}) \bar{\mathbf{X}}\|_{\mathbb{F}}^2 + \iota_{\mathcal{P}_0}(\bar{\mathbf{X}}) + \iota_{\mathcal{D}}(\bar{\mathbf{h}})\}, \quad (5.7)$$

where  $L(\bar{\mathbf{X}}, \bar{\mathbf{h}})$  is the objective function and  $\mathbf{Z} = \mathbf{Y} - \gamma \mathbf{1}_N \mathbf{1}_P^T$  are the modified observations.

## 5.4 Proximal Alternating Minimization

In this section, we describe the numerical methods used to solve the problem (5.7) and hence estimating the image and the PSF from noisy observations.

Problem (5.7) is non-convex with respect to both  $\mathbf{X}$  and  $\mathbf{h}$ , but it is convex with respect to  $\mathbf{X}$  (resp.  $\mathbf{h}$ ) if  $\mathbf{h}$  (resp.  $\mathbf{X}$ ) is known. Such problem can be solved through the proximal alternating algorithm described in Section 2.4.3, which iterates as follows:

$$\begin{cases} \mathbf{X}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{X}} \in \mathbb{R}^N} L(\bar{\mathbf{X}}, \mathbf{h}^{(k)}) + \frac{\lambda_x^{(k)}}{2} \|\bar{\mathbf{X}} - \mathbf{X}^{(k)}\|_{\mathbb{F}}^2 \\ \mathbf{h}^{(k+1)} &= \operatorname{argmin}_{\bar{\mathbf{h}} \in \mathbb{R}^W} L(\mathbf{X}^{(k+1)}, \bar{\mathbf{h}}) + \frac{\lambda_h^{(k)}}{2} \|\bar{\mathbf{h}} - \mathbf{h}^{(k)}\|_2^2 \end{cases}. \quad (5.8)$$

The non-convexity of (5.7) prevents attaining a global minimizer. However, thanks to the works of Attouch *et al.* [81, 104], algorithm (5.8) is proven to converge to a critical point of  $L(\mathbf{X}, \mathbf{h})$  provided some conditions are met (see Section 2.4.3). As shown in Appendix 5.B and in the following sections, objective function  $L(\mathbf{X}, \mathbf{h})$  and algorithm (5.8) satisfy Hypotheses 2.1, 2.2 and 2.3. Therefore, if the sequence generated by algorithm (5.8) is bounded, then it converges to a critical point of problem (5.7).

Note that an exact solution of each subproblem in (5.8) is not required for the algorithm convergence (see discussion in Section 2.4.3): the work of Attouch *et al.* [81] suggests that an approximate solution of each subproblem is sufficient to guarantee the convergence of the algorithm to a critical point of (5.7). In practice, we do not check that the subproblems are solved with a sufficient precision, but the proposed implementation based on a threshold on the relative error (see Appendices 5.D.1 and 5.D.2) still works fine empirically.

In the following, we first demonstrate that the sequence generated by algorithm (5.8) is bounded. Then, we describe how parameters  $(\lambda_x^{(k)}, \lambda_h^{(k)})$  are tuned such that algorithm (5.8) satisfies Hypothesis 2.2. We present next how to stop the iterations so that a vicinity of a critical point of (5.7) is reached. Finally, we provide an overview of the numerical reconstruction.

### 5.4.1 Bounded sequence

In order to guarantee that sequence  $\{\mathbf{X}^{(k)}, \mathbf{h}^{(k)}\}_{k \in \mathbb{N}}$  generated by the proximal alternating algorithm (5.8) converges to a critical point of problem (5.7),

we must demonstrate that it is bounded. Remark that the filter must belong to the probability simplex, a constraint enforced in (5.7) by the indicator function onto the set  $\mathcal{D}$ . This constraint forces sequence  $\{\mathbf{h}^{(k)}\}_{k \in \mathbb{N}}$  to be bounded.

Let us now study if sequence  $\{\mathbf{X}^{(k)}\}_{k \in \mathbb{N}}$  is bounded. Due to the non-negativity constraint on the image, sequence  $\{\mathbf{X}^{(k)}\}_{k \in \mathbb{N}}$  is bounded below. We need to find also an upper-bound in order to guarantee that the sequence is bounded. From the work of Attouch *et al.* [104, Lemma 5], we know that algorithm (2.41) has the following property:

$$L(\mathbf{X}^{(k+1)}, \mathbf{h}^{(k+1)}) + \frac{1}{2} \left[ \lambda_h^{(k)} \|\mathbf{h}^{(k+1)} - \mathbf{h}^{(k)}\|_2^2 + \lambda_x^{(k)} \|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_{\mathbb{F}}^2 \right] \leq L(\mathbf{X}^{(k)}, \mathbf{h}^{(k)}), \quad (5.9)$$

which implies that  $L(\mathbf{X}^{(k)}, \mathbf{h}^{(k)})$  does not increase. Due to (i) the nature of the objective function, which is equal to  $L(\mathbf{X}, \mathbf{h}) = \rho \|\mathbf{S}_{\Theta} \mathbf{\Psi}^* \mathbf{X}\|_1 + \frac{1}{2} \|\mathbf{Z} - \mathbf{\Phi}(\mathbf{h}) \mathbf{X}\|_{\mathbb{F}}^2$  if the constraints are always satisfied during the iterations, (ii) the properties of the matrix  $\mathbf{\Phi}(\mathbf{h})$  (a convolution matrix of kernel  $\mathbf{h}$ ) for  $\mathbf{h} \in \mathcal{D}$ , (iii) the properties of the cost-to-move parameters  $\{\lambda_x^{(k)}, \lambda_h^{(k)}\}_{k \in \mathbb{N}}$  (they are bounded); a sequence  $\{\mathbf{X}^{(k)}\}_{k \in \mathbb{N}}$  with norm tending to infinity cannot satisfy the property in (5.9). Therefore, since algorithm (5.8) is non-increasing, sequence  $\{\mathbf{X}^{(k)}\}_{k \in \mathbb{N}}$  is bounded. Remark that an additional constraint can also be added to the problem such that we force the image to have a specific upper-bound based on the maximum intensity of the observations.

### 5.4.2 Cost-to-Move Parameters

The presence of cost-to-move parameters  $(\lambda_x, \lambda_h)$  different than zero ensures the convergence of the algorithm (2.41) to a critical point  $(\tilde{\mathbf{X}}, \tilde{\mathbf{h}})$  of (5.7). Although their values are not crucial in the reconstruction results, they must be bounded as required by Hypothesis 2.2, *i.e.*, for some positive  $c_- < c_+$ , the sequences  $(\lambda_x^{(k)}, \lambda_h^{(k)})$  belong to  $(c_-, c_+)$  for all  $k \geq 0$ .

The tuning criteria used in this work is based on the works of Puy and Vandergheynst [24]. The iterations start with high values of  $\lambda_x$  and  $\lambda_h$ , keeping the image and the filter estimations close to the initial values. When the number of iterations  $k$  increases, the filter and the image estimates become more and more accurate and the parameters  $\lambda_x$  and  $\lambda_h$  can be progressively decreased until some lower bound  $c_- \in \mathbb{R}_+$ , fixed here equal to  $10^{-15}$ .

This choice of the cost-to-move parameters is general and independent of the numerical algorithm used to solve each step of the alternating approach. If, for instance, we use the forward-backward algorithm to solve each step (see Section 2.4.2.1), the cost-to-move parameters can be modified based on the partial Lipschitz moduli of  $\nabla Q$  [82].

### 5.4.3 Stopping Criteria

It is important to find an automatic criterion for stopping the iterations in (2.41) when the solution reaches the vicinity of a critical point of (5.7). A usual stopping criterion is based on the quality of the reconstruction with respect to the ground truth image and filter, which are in general unavailable. As proposed by Almeida and Figueiredo [55], we use the spectral characteristics of the noise to analyze the quality of the estimation. This allows stopping the minimization when a high quality solution has been obtained. Let us define each residual image at iteration  $k$  as the difference between the observed image and the blurred estimate:

$$\mathbf{R}^{(k)} = \mathbf{Z} - \Phi(\mathbf{h}^{(k+1)}) \mathbf{X}^{(k+1)} = (\mathbf{r}_1^{(k)}, \dots, \mathbf{r}_p^{(k)}). \quad (5.10)$$

Since the observation model is degraded by an additive white noise, we know that the residual image is spectrally white if the estimation at iteration  $k$  has a good quality, otherwise the residual contains structured artifacts. Therefore, the iterations can be stopped when the residual is spectrally white, *i.e.*, when the 2-D autocorrelation  $C_{\mathbf{r}_j, \mathbf{r}_j}(m, n)$  of each residual image  $\mathbf{r}_j^{(k)}$  is approximately the 2-D Kronecker function  $\delta_0(m, n)$ , which is equal to 1 if  $m = n = 0$ , and 0 otherwise. Let us note that, for the computation of the 2-D autocorrelation function, each residual vector  $\mathbf{r}_j^{(k)}$  needs to be previously transformed into a matrix of  $n \times n$  pixels and normalized to zero mean and unit variance.

Considering a  $(2B + 1) \times (2B + 1)$  window, we compute the distance between the autocorrelation and the Kronecker function  $\delta_0$  as follows:

$$\begin{aligned} W(\mathbf{r}_j^{(k)}) &= - \sum_{(m,n)=(-B,-B)}^{(B,B)} \left( \delta_0(m, n) - C_{\mathbf{r}_j, \mathbf{r}_j}(m, n) \right)^2 \\ &= - \sum_{\substack{(m,n)=(-B,-B) \\ (m,n) \neq (0,0)}}^{(B,B)} \left( C_{\mathbf{r}_j, \mathbf{r}_j}(m, n) \right)^2, \end{aligned} \quad (5.11)$$

which is higher for whiter residuals. As suggested by Almeida and Figueiredo [55], in our experiments we have used  $B = 4$ . We then average over the measures obtained for each residual image:

$$W(\mathbf{R}^{(k)}) = \frac{1}{P} \sum_{j=1}^P W(\mathbf{r}_j^{(k)}). \quad (5.12)$$

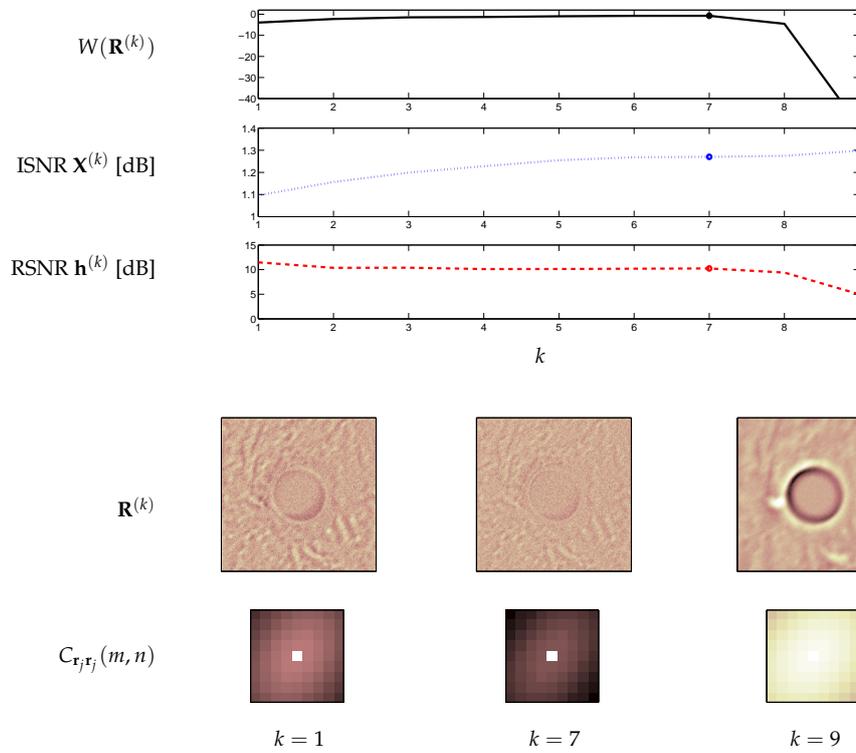
Figure 5.2 illustrates the behavior of whiteness measure  $W(\mathbf{R}^{(k)})$  when reconstructing the image of Figure 5.3-(a) and the PSF of Figure 5.4-(a). For a fixed value  $\rho = 0.81$ , the figure shows the evolution (along the iterations  $k$ ) of whiteness measure  $W(\mathbf{R}^{(k)})$ , the image quality measured by the ISNR in (5.14) and the PSF quality measured by the RSNR in (5.15). Additionally, for some iterations ( $k = 1, 7, 9$ ), the figure shows the residual image  $\mathbf{r}_1^{(k)}$  and the 2-D autocorrelation  $C_{\mathbf{r}_j, \mathbf{r}_j}(m, n)$  defined in a  $(2B + 1) \times (2B + 1)$  window with  $B = 4$ .

We observe that, in practice,  $W(\mathbf{R}^{(k)})$  has large negative values at the beginning of the iterations when the image and/or the PSF are not properly estimated (see  $k = 1$  in Figure 5.2). Then, the value of  $W(\mathbf{R}^{(k)})$  increases with  $k$  until it reaches a maximum close to zero (see  $k = 7$  in Figure 5.2). In this iteration, both the image and the PSF have a good quality and the residual image is almost spectrally white (evidenced by an auto-correlation close to  $\delta_0(m, n)$ ). Finally, the whiteness measure starts to decrease again after the algorithm has converged to the best estimations for a fixed value of  $\rho$  and starts overfitting the noise (see  $k = 9$  in Figure 5.2). This behavior implies that the iterations can be stopped when the maximum of the whiteness measure  $W(\mathbf{R}^{(k)})$  is reached. A similar behavior has been observed by Almeida and Figueiredo [55].

#### 5.4.4 Numerical Reconstruction

The proximal alternating minimization algorithm is summarized in Algorithm 5.1. This algorithm requires to set the initial values of the image, the filter and the regularization parameter. The initialization of the image and the filter is discussed in Appendix 5.C.2, and the regularization parameter is tuned iteratively as explained in Appendix 5.C.1.

The first step in Algorithm 5.1 estimates the image by minimizing a sum of two smooth and two non-smooth convex functions. Such problem



**Figure 5.2** Evolution of the whiteness criteria  $W(\mathbf{R}^{(k)})$  in (5.12), the image quality measured by the ISNR in (5.14) and the PSF quality measured by the RSNR in (5.15), along the iterations  $k$ . The example is shown for the image of Figure 5.3-(a), the PSF of Figure 5.4-(a) and  $\rho = 0.81$ . For  $k = 1$  (first iteration),  $k = 7$  (best value of  $W(\mathbf{R}^{(k)})$ ) and  $k = 9$  (last iteration), the figure depicts the residual image and the 2-D auto-correlation.

can be solved through the expanded Chambolle-Pock (CP) described in Section 2.4.2.2. In Appendix 5.D.1, we provide further details on the numerical reconstruction of the first step.

The second step in Algorithm 5.1 estimates the filter by minimizing a sum of two smooth and one non-smooth convex functions. The optimization problem can be solved using the accelerated proximal gradient (APG) algorithm described in Section 2.4.2.1. In Appendix 5.D.2 we provide more details on the numerical reconstruction of the second step.

---

**Algorithm 5.1** Proximal Alternating Minimization Algorithm

---

**Initialization:**  $\mathbf{X}^{(1)} = \mathbf{X}_0; \mathbf{h}^{(1)} = \mathbf{h}_0; \lambda_x^{(1)} = \lambda_h^{(1)} = \rho = \rho_0; \Delta = 0.75; \text{MaxIter} = 20$ 1: **for**  $k = 1$  to  $\text{MaxIter}$  **do**    *1st step, image estimation:*

2:  $\mathbf{X}^{(k+1)} = \underset{\bar{\mathbf{X}}}{\text{argmin}} L(\bar{\mathbf{X}}, \mathbf{h}^{(k)}, \rho) + \frac{\lambda_x^{(k)}}{2} \|\bar{\mathbf{X}} - \mathbf{X}^{(k)}\|_{\mathbb{F}}^2$

*2nd step, filter estimation:*

3:  $\mathbf{h}^{(k+1)} = \underset{\bar{\mathbf{h}}}{\text{argmin}} L(\mathbf{X}^{(k+1)}, \bar{\mathbf{h}}) + \frac{\lambda_h^{(k)}}{2} \|\bar{\mathbf{h}} - \mathbf{h}^{(k)}\|_2^2$

*Parameters update:*

4:  $\lambda_x^{(k+1)} = \min\{\lambda_x^{(k)} \Delta, 10^{-15}\}$

5:  $\lambda_h^{(k+1)} = \min\{\lambda_h^{(k)} \Delta, 10^{-15}\}$

*Compute residuals and whiteness measure:*

6:  $\mathbf{R}^{(k)} = \mathbf{Z} - \Phi(\mathbf{h}^{(k+1)}) \mathbf{X}^{(k+1)}$

7: Compute  $W(\mathbf{R}^{(k)})$  using (5.12)    *Stop when residual is spectrally whiter:*8: **if**  $W(\mathbf{R}^{(k+1)}) < W(\mathbf{R}^{(k)})$  **then** break.9: **end if**10: **end for**11: Return  $\mathbf{X}^{(k)}$  and  $\mathbf{h}^{(k)}$ .

---

*Remark 5.3.* The algorithms used to solve the minimization problems described above were chosen because they are well suited to the characteristics of each subproblem. Since the cost functions are different in each step, the same algorithm could not be used to optimally find a solution of both problems. The presence of two non-smooth functions in the first step, prevents the use of the APG algorithm. For the second step, the CP algorithm could be used to solve the optimization problem, however, this algorithm presented a slower convergence than APG, which is optimal when the optimized cost contains a differentiable function. This choice can be further investigated and other algorithms may be better suited for solving both problems. For instance, as done in [82], we could solve all steps using the forward-backward algorithm, but this would require to increase the optimization space by including auxiliary variables.

Let us note that, in the numerical reconstruction, the convolution operator is implemented using the Fast Fourier Transform (FFT). This operation assumes that the boundaries of the image are periodic, a condition that is far from reality. In actual imaging, no condition can be assumed on the boundary pixels, *i.e.*, they cannot be assumed to be zero, neither periodic, nor reflexive. Not taking care of the boundaries conditions would introduce ringing artifacts. Based on the work of Almeida and Figueiredo [61], in the numerical experiments we consider unknown boundary conditions, *i.e.*, the pixels belonging to the image border are not observed. To take this into account, the problem formulation needs to be appropriately modified as discussed in Appendix 5.E. A recent work by Simões *et al.* [159] have shown that, under these unknown boundary conditions, the unobserved pixels in the image border could also be estimated. However, including this estimation in (5.8) is beyond the scope of this work and should be investigated in the future.

## 5.5 Filter Validation using Non-Blind Deconvolution

This section is dedicated to the validation of the filter estimated using Algorithm 5.1. Since the true emissions of the Moon and Venus are zero, the deconvolution with a correct filter should remove all the celestial bodies' apparent emissions. Therefore, the validation process consists of taking an image from a transit observation that was not used for filter estimation (*i.e.*,  $\mathbf{y}_j \notin \{\mathbf{y}_1, \dots, \mathbf{y}_P\}$ ), deconvolving this image using the estimated filter  $\mathbf{h}$ , and verifying that, in the reconstructed image  $\tilde{\mathbf{x}}_j$ , the pixels inside the celestial body disk are zero.

To obtain this reconstructed image, we formulate a least squares minimization problem regularized using the prior information available on the image, that is, that the image is non-negative and has a sparse representation on a suitable wavelet basis  $\Psi \in \mathbb{R}^{W \times N}$ . The proposed non-blind deconvolution method reads as follows:

$$\tilde{\mathbf{x}}_j = \underset{\bar{\mathbf{x}}_j}{\operatorname{argmin}} \quad \rho \|\mathbf{S}_\Delta \Psi^* \bar{\mathbf{x}}_j\|_1 + \frac{1}{2} \|\mathbf{z}_j - \mathbf{h} \otimes \bar{\mathbf{x}}_j\|_2^2 \quad \text{s.t.} \quad \bar{\mathbf{x}}_j \in \mathcal{P} \quad (5.13)$$

where  $\mathbf{S}_\Delta \in \mathbb{R}^{S \times W}$  is the selection operator of the set  $\Delta$ , with  $|\Delta| = S$ , which contains the detail coefficients. The convex set  $\mathcal{P}$  is defined as

$\mathcal{P} = \{\mathbf{u} \in \mathbb{R}^N : u_i \geq 0\}$ . The non-blind image estimation is solved using the extended CP algorithm (see Section 2.4.2.2). Since this algorithm is able to minimize a sum of three non-smooth convex functions, the constraint can be handled as a convex indicator function. Notice that other algorithms could be used such as the extension in the product space of the proximal gradient algorithm presented in Section 2.4.2.1.

## 5.6 Experiments

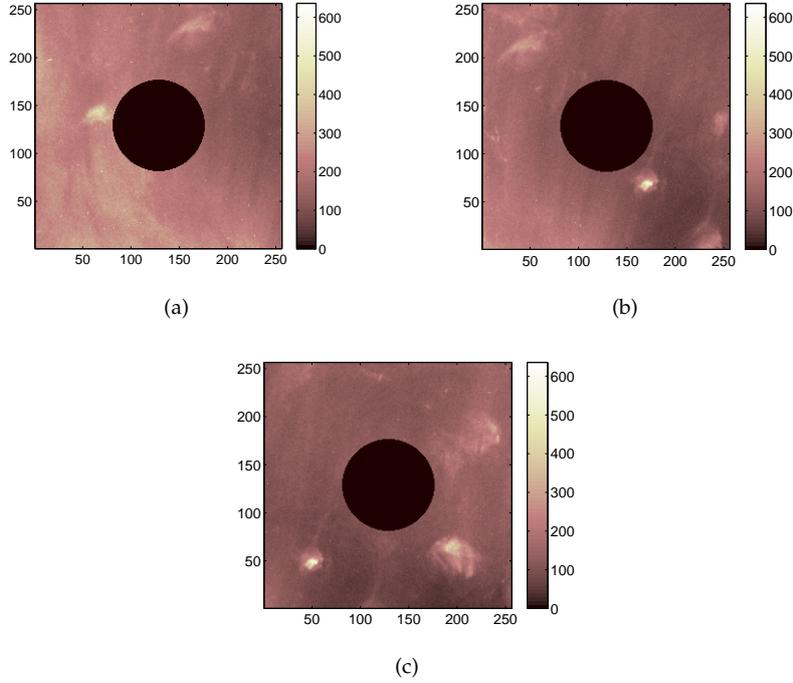
In this section, we first present some synthetic results that allow us to validate the effectiveness of the proposed blind deconvolution method to recover the correct filter. Later we present experimental results on images taken by the SDO/AIA and SECCHI/EUVI telescopes.

All algorithms were implemented in MATLAB and executed on a 3.2 GHz Intel i5-650 CPU with 3.7 GiB of RAM, running on a 64 Bit Ubuntu 14.04 LTS operating system.

### 5.6.1 Synthetic Data

Three synthetic images are selected to test the reconstruction ( $P = 3$ ). They are defined on a  $256 \times 256$  pixel grid ( $N = 256^2$ ). To simulate actual images from the celestial transit, we take an image of the Sun (the image observed by SDO/AIA at 00:00 UT on June 6th 2012) and select three different cutouts of  $N$  pixels each. The center of the cutouts are arbitrarily selected such that the corresponding regions are sufficiently distant from each other and correspond to a sample of a full Venus trajectory. In each image, we add a black disk of radius  $R = 48$  pixels, centered at the pixel  $[129, 129]$ , which represents the celestial body. Figure 5.3 depicts the images from the simulated transit.

During the solar transit, the position and size of the celestial body are known at all moments. However, when this event is imaged, the coordinates of the center represented in the discrete image grid may contain an error of maximum one pixel. This uncertainty is simulated in our synthetic experiments by considering the sets  $\Omega_j$  and  $\Theta$  to be defined using disks of radius

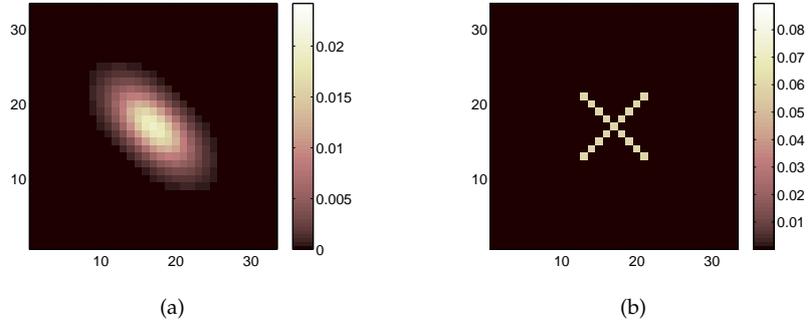


**Figure 5.3** Realistic images: (a) first image  $x_1$ , (b) second image  $x_2$  and (c) third image  $x_3$  from the simulated transit.

$r_\Omega$  and  $r_\Theta$ , respectively, with one pixel difference with respect to the actual object's radii, *i.e.*,  $r_\Omega = R - 1$  and  $r_\Theta = R + 1$ .

Two kinds of discrete filters are selected and they have a limited support on a  $33 \times 33$  pixel grid ( $b = 16$ ). The first filter is simulated by an anisotropic Gaussian function with standard deviation of 2 pixels horizontally and 4 pixels vertically, rotated by 45 degrees (see Figure 5.4-(a)). The second filter is simulated by a X shape (see Figure 5.4-(b)), hereafter called the X filter. These functions help to demonstrate the capacity of the proposed method to recover not only diffusion filters such as the anisotropic Gaussian, but also diffraction filters such as the X example. Let us note that, in these synthetic experiments, there is no need to test the long-range assumption on the PSF.

The measurements are simulated according to (5.2), where additive white Gaussian noise  $(N)_{ij} \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma^2)$  is added in order to simulate a realistic



**Figure 5.4** Realistic filters in a  $(2b + 1) \times (2b + 1)$  window with  $b = 16$ . (a) Anisotropic Gaussian Filter. (b) X Filter.

scenario. The noise variance ( $\sigma^2$ ) is set such that the blurred signal to noise ratio

$$\text{BSNR} = 10 \log_{10} \frac{\text{var}(\Phi(\mathbf{h}) \mathbf{X})}{\sigma^2}$$

is equal to 30 dB, which corresponds to a realistic BSNR in the actual observations.

During the experiments, the wavelet transformation used for the sparse image representation is the redundant Daubechies wavelet basis<sup>1</sup> with two vanishing moments and three levels of decomposition (see Section 2.C for more details). Other mother wavelets and more vanishing moments can be considered, however, due to the dictionary redundancy, the choice does not have a significant impact on the reconstruction results. Higher levels can also be considered but the computational time is notably higher and the reconstruction quality does not significantly improve.

As discussed in Section 5.3.1, the wavelet coefficients with support touching the boundary of the occulting body are not sparse as they spread over all the scales of the wavelet transform. Thus, we need to define the set  $\Theta$  containing the coefficients that are not affected by the disk. To determine this set, we first generate a set of images with constant background that contain disks of radius  $r_\Theta$  centered as the celestial body. All the elements inside the disks are set as random values with a standard uniform distribution ( $\mathcal{U}(0, 1)$ ). We then

<sup>1</sup>The wavelet operator and the other operators used in this work were numerically implemented using the SPARCO framework [132].

compute the wavelet coefficients of this image on the basis  $\Psi$ . Finally, the set  $\Theta$  is composed by the indices of the *detail* coefficients that are equal to zero.

The reconstruction quality of  $\tilde{\mathbf{X}}$  with respect to the true image  $\mathbf{X}$  is measured using the increase in SNR (ISNR) defined as

$$\text{ISNR} = 20 \log_{10} \frac{\|\mathbf{Y} - \mathbf{X}\|_F}{\|\tilde{\mathbf{X}} - \mathbf{X}\|_F}. \quad (5.14)$$

The reconstruction quality of  $\tilde{\mathbf{h}}$  with respect to the true filter  $\mathbf{h}$  is measured using the reconstruction SNR (RSNR), where

$$\text{RSNR} = 20 \log_{10} \frac{\|\mathbf{h}\|_2}{\|\mathbf{h} - \tilde{\mathbf{h}}\|_2}. \quad (5.15)$$

Let us first analyze the behavior of the algorithm when we increase the number of observations  $P$  used for the reconstruction. Table 5.1 presents the reconstruction quality of the results using the blind deconvolution problem for the anisotropic Gaussian filter of Figure 5.4-(a). The results correspond to an average over four trials and they are presented for  $P = 1, 2$  and 3 observations. For  $P = 2, 3$  we use a warm start by initializing the algorithm with the results obtained for  $P - 1$ . The whiteness measure  $W(\tilde{\mathbf{R}})$  presented in Table 5.1 has been computed using (5.12) with the residual obtained from resulting image and filter, *i.e.*,  $\tilde{\mathbf{R}} = \mathbf{Y} - \Phi(\tilde{\mathbf{h}})\tilde{\mathbf{X}}$ .

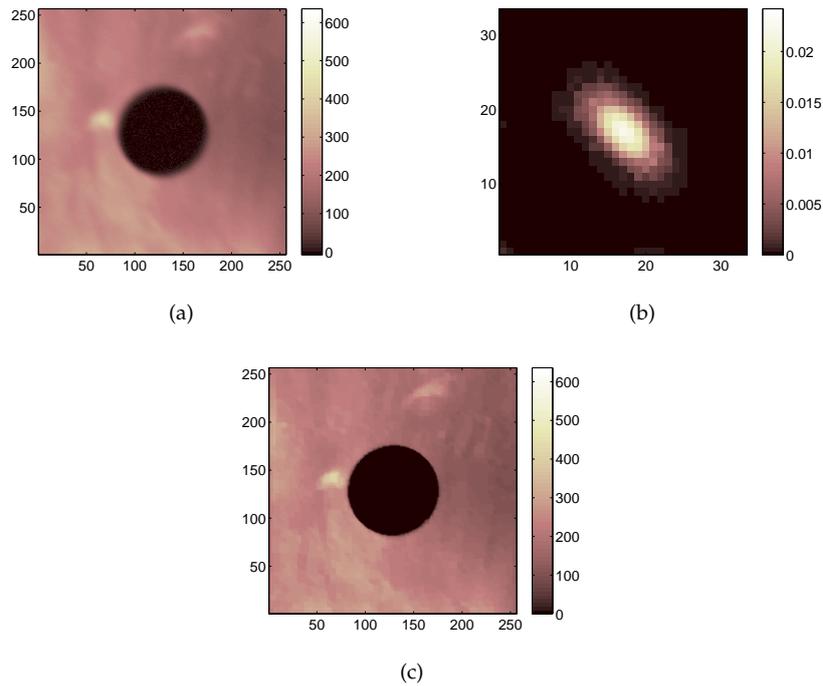
$P$	ISNR $\tilde{\mathbf{X}}$ [dB]	RSNR $\tilde{\mathbf{h}}$ [dB]	$W(\tilde{\mathbf{R}})$
1	1.27	9.63	-0.25
2	1.33	13.07	-0.16
3	1.36	14.05	-0.03

**Table 5.1** Reconstruction quality for different number of simulated transit observations using the synthetic anisotropic Gaussian filter.

We notice that the reconstruction quality of the filter significantly improves as the number of observations  $P$  increases. The whiteness measure  $W(\tilde{\mathbf{R}})$  presents the same behavior, which indicates that the residual image is whiter when more observations are considered. This is explained by an increase of the available information in the filter reconstruction problem for the same number of unknowns.

Regarding the numerical complexity, the algorithm requires an average of five iterations on the value of  $\rho$  and a total time of 1.5 hours. When the number of iterations on  $\rho$  increases, we observed that the estimated residual energy is closer to the actual noise energy.

Let us now show some reconstruction results of the different filters and how they reproduce the true zero pixel values inside the object. Figure 5.5-(a) depicts the first noisy observed patch and Figure 5.5-(b) presents the resulting PSF when reconstructing the anisotropic Gaussian filter for  $P = 3$ . We can observe how the algorithm is able to estimate the filter with a high quality, even when no hard constraints are introduced on the filter shape. The reconstruction quality can be further improved if stronger conditions are assumed on the filter.



**Figure 5.5** Results for the anisotropic Gaussian filter and  $P = 3$ . (a) Noisy observation of image 1 ( $y_1$ ). (b) Filter Reconstruction ( $\tilde{h}$ ) using blind deconvolution, RSNR = 14.05 dB. (c) Image Reconstruction ( $\tilde{x}_1$ ) using non-blind deconvolution, ISNR = 1.83 dB.

Figure 5.5-(c) presents the reconstructed image using the estimated filter from Figure 5.5-(b) in the non-blind deconvolution of Section 5.5. Note that the majority of the pixels inside the estimated disk are zero, except for some numerical errors. To quantify this validation, we use as measurement the disk intensity, *i.e.*, the sum of the pixel values inside the disk. We compare the ratio between the disk intensity for the deconvolved image, denoted by  $S_X$ , and the disk intensity for the observed patch, denoted by  $S_Y$ . We obtained a disk intensity ratio of  $S_X/S_Y = 8.16 \times 10^{-2}$ . This shows the effectiveness of the estimated filter to recover the true zero emissions inside the disk.

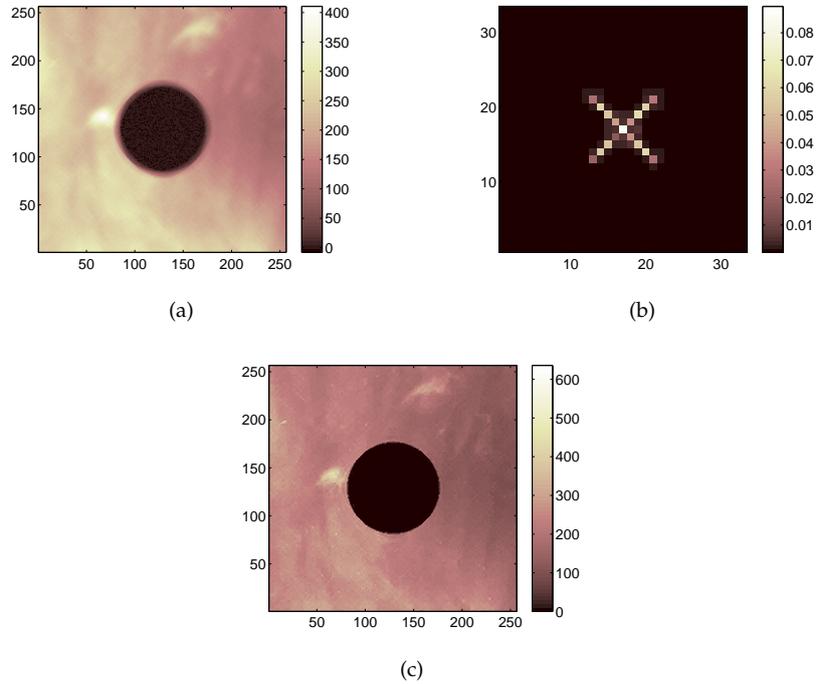
Figure 5.6 depicts the results for  $P = 3$  observations using the  $X$  filter of Figure 5.4-(b). Let us note that the reconstruction quality is lower than in the case of the Gaussian filter because the  $X$  filter is more complex and hence it is harder to estimate without extra information on its shape. Nevertheless, since it causes less diffusion on the observation, the image reconstruction quality is better. When comparing the values inside the disk between the observations and the estimated images, we observed that the disk intensity ratio is  $S_X/S_Y = 2.03 \times 10^{-2}$ , which validates the zero emissions inside the disk.

Finally, we also considered a case (not displayed here) where the observation is only corrupted by noise. As expected, the filter estimated by the algorithm is a high quality Kronecker delta function with  $\text{RSNR} = 66$  dB. This result shows the capability of the algorithm to recover highly localized filters of one pixel radius.

### 5.6.2 Experimental Data

The proposed blind deconvolution method was tested on two experimental sets of data<sup>2</sup>. A first data set corresponds to the observations of the Venus transit on June 5th - 6th 2012 by the Atmospheric Imaging Assembly (AIA) [160] on board the Solar Dynamics Observatory (SDO). The second data set corresponds to the observations of the Moon transit on February 25th 2007 by the Extreme Ultraviolet Imager (EUVI) [8], a part of the Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) instrument suite on board the STEREO-B spacecraft. These images are compressed using the RICE

<sup>2</sup>The data has been provided by the Royal Observatory of Belgium (ROB), Brussels, Belgium.



**Figure 5.6** Results for the X filter and  $P = 3$ . (a) Noisy observation of image 1 ( $y_1$ ). (b) Filter Reconstruction ( $\tilde{h}$ ) using blind deconvolution,  $\text{RSNR} = 7.52$  dB. (c) Image Reconstruction ( $\tilde{x}_1$ ) using non-blind deconvolution,  $\text{ISNR} = 2.55$  dB.

algorithm [161], which is lossless [36] and hence does not introduce additional errors in the PSF estimation.

In the following, we present for each telescope the results obtained when reconstructing the filter using the blind deconvolution approach. Then, we validate the obtained filters with transit images that were not used for the PSF estimation. Finally, we demonstrate how the obtained filters work when deconvolving non-transit images. All results are compared with previously estimated PSFs.

### 5.6.2.1 SDO/AIA - Venus Transit

We consider three  $4096 \times 4096$  level 1 images from the transit ( $P = 3$ ) recorded by the 19.3 nm channel of AIA. The filter is assumed to have a limited

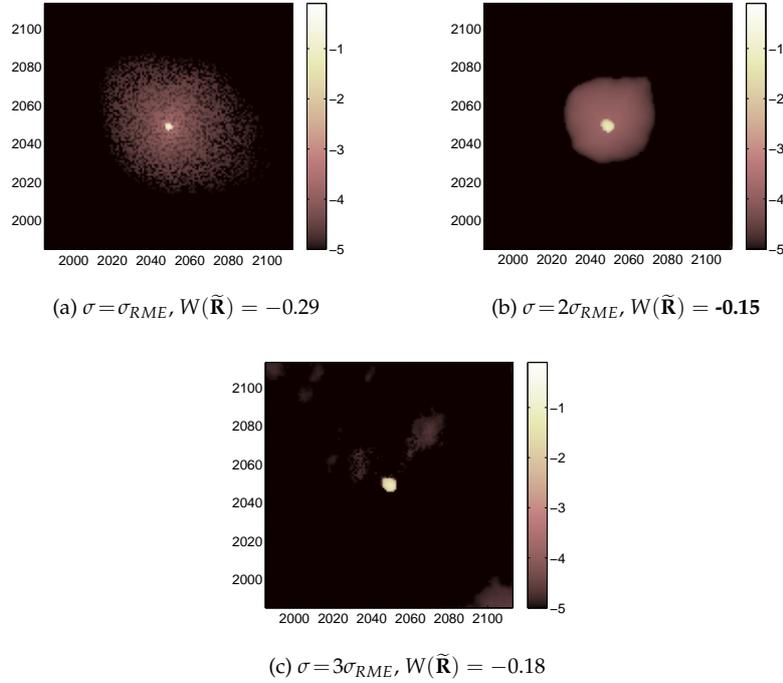
support inside a  $129 \times 129$  pixel grid ( $b = 64$ ), which allows encompassing 99% of the energy of previously estimated PSFs. The presence of a long-range PSF was verified in the observations by analyzing the pixels inside the disk of Venus. To estimate this effect, we considered a total of 10 patches and computed the mean intensity value on a disk of radius 10 pixels inside the disk of Venus. The obtained value  $\gamma = 43.3$  DN was removed from the observations to estimate the filter using the blind deconvolution scheme.

The radius of the Venus disk is known to be 49 pixels and hence, the disk represents a small area inside the high resolution image. Since the blind deconvolution takes benefit mainly from the black disk in the transit images, we select a patch of size  $N = 256^2$  centered on the Venus disk. As explained in Appendix 5.E, our method considers unknown border pixels based on the PSF support (completely defined by  $b$  after removing the value of  $\gamma$  from the observed patch). Therefore, cropping the observed image does not have any effects on the reconstruction results. Furthermore, considering smaller observations keeps the optimization problem computationally tractable.

As explained in Section 5.2.2, we use an *adaptive* noise estimation strategy to optimize the AWGN assumption in (5.3). For this, we start from  $\sigma = \sigma_{RME} = 2.81$  DN, the value estimated by the Robust Median Estimator (RME). Then, the blind deconvolution method is performed for different values of  $\sigma$  that are multiples of  $\sigma_{RME}$ . Finally, we take the value of  $\sigma$  that optimizes the whiteness of the residual  $W(\tilde{\mathbf{R}})$  as defined in (5.12), *i.e.*, the value of  $\sigma$  ensuring that the residual in (5.10) only contains white noise without any remaining of the signal features.

Figure 5.7 shows the resulting PSFs for different values of  $\sigma$ . We can observe that the whitest residual is obtained for  $\sigma = 2\sigma_{RME} = 5.62$  DN and we thus select this value for our non-parametric PSF estimate  $\tilde{\mathbf{h}}_{np}$  depicted in Figure 5.7-(b). Note that, if the noise is underestimated, we reconstruct part of the noise in the PSF and image, and the resulting PSF is too noisy (see Figure 5.7-(a)). Oppositely, if the noise is overestimated, the algorithm provides the trivial solution, *i.e.*, the PSF tends to a discrete delta and the image to the observations (see Figure 5.7-(c)).

The estimated non-parametric PSF in Figure 5.7-(b) is compared with the parametric PSF estimated by Poduval *et al.* [17], depicted in Figure 5.8-(a).



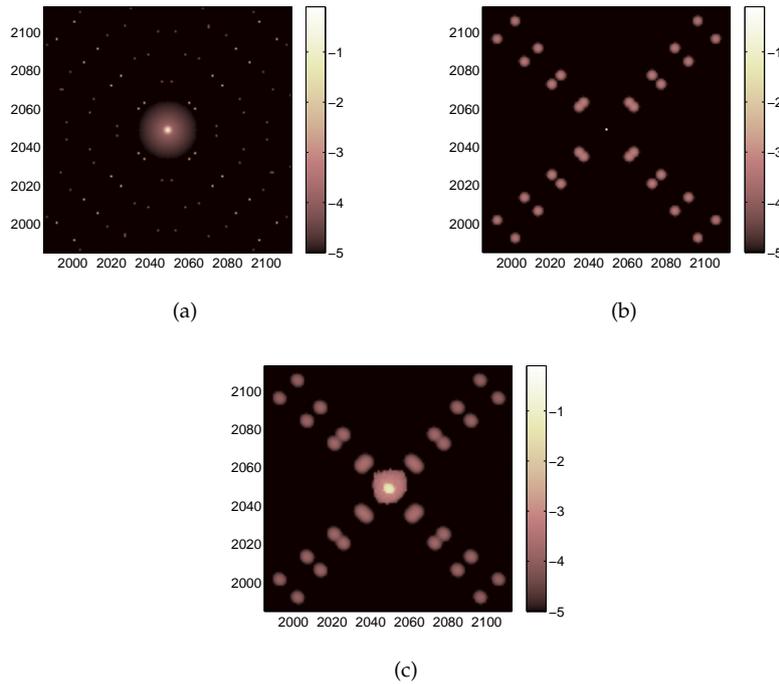
**Figure 5.7** Logarithm of the non-parametric filters estimated for the SDO/AIA telescope taken inside the set  $\Lambda$ . The filters are calculated for several values of  $\sigma$  providing different residual whiteness (measured by  $W(\tilde{\mathbf{R}})$ ).

This PSF, denoted  $\mathbf{h}_{p_1}$ , was obtained by fitting a parametric model based on the optical characterization of the telescope. We observe that the obtained non-parametric filter is highly localized in the center of the discrete grid and presents a limited support. Let us note that the observation noise level prevents the estimation of the diffraction peaks present in  $\mathbf{h}_{p_1}$ . These can only be obtained by constraining the shape of the filter in the reconstruction process, as implicitly done by parametric deconvolution methods.

To solve the blind deconvolution problem, the algorithm requires an average of four iterations on the value of  $\rho$  and a total time of 8 hours on our computer. In order to reduce the computation time, some functions such as the proximity operators may be computed in parallel, as many of the convex functions on which they are defined are separable [51]. The computation time

can be further reduced if, for instance, the algorithms are implemented in C instead of MATLAB.

In order to obtain a more accurate PSF estimation containing the diffraction peaks, let us incorporate a parametric PSF inside the acquisition model in (5.2) by considering a combined parametric/non-parametric PSF defined as  $\tilde{\mathbf{h}}_{\text{p-np}} = \mathbf{h}_{\text{p}} \otimes \tilde{\mathbf{h}}_{\text{np}}$  [16, 151]. The parametric part is obtained by considering only the mesh diffraction components in the PSF estimated by Poduval *et al.* [17]. This modified PSF, denoted  $\mathbf{h}_{\text{p}_2}$ , is depicted in Figure 5.8-(b). The non-parametric part ( $\tilde{\mathbf{h}}_{\text{np}}$ ) is estimated using the proposed blind deconvolution approach. Figure 5.8-(c) presents the resulting parametric/non parametric PSF, *i.e.*,  $\tilde{\mathbf{h}}_{\text{p}_2\text{-np}}$ .



**Figure 5.8** Logarithm of filters for the SDO/AIA telescope taken inside the set  $\Lambda$ . (a) Parametric PSF estimated by Poduval *et al.* [17], *i.e.*,  $\mathbf{h}_{\text{p}_1}$ . (b) Diffraction mesh pattern modeled using the parameters estimated by Poduval *et al.* [17], *i.e.*,  $\mathbf{h}_{\text{p}_2}$ . (c) Combined parametric/non-parametric filters using the parametric PSF in (b), *i.e.*,  $\tilde{\mathbf{h}}_{\text{p}_2\text{-np}}$ .

We can observe that the resulting parametric/non-parametric PSF is a corrected version of the parametric model from Figure 5.8-(b). Its central shape is also more diffused similarly to what is observed in the parametric PSF estimated by Poduval *et al.* [17].

The estimated filters from Figure 5.7-(b), Figure 5.8-(a) and Figure 5.8-(c) were validated using the non-blind deconvolution formulation from Section 5.5. Notice that, similarly to what has been done for the blind deconvolution, the value of  $\rho$  has been selected adaptively by optimizing the residual whiteness defined in (5.11). We also observed empirically that the value of  $\rho$  must be kept smaller than the one obtained in Algorithm 5.2. The validation was done on transit images using the modified observations with  $\gamma = 43.3$  DN. A first validation was performed on the Venus transit image taken by SDO/AIA at 00:02 UT on June 6th 2012 (see Figure 5.9-(a)), a patch that has not been used before for estimating  $\tilde{\mathbf{h}}$ . A second validation of the filters was done on the Moon transit image taken by SDO/AIA at 13:00 UT on March 4th 2011 (see Figure 5.10-(a)).

We quantify the apparent disk emissions by summing the pixel values inside the disk. Table 5.2 displays the disk intensity ratio for the images deconvolved using the following filters: (1) the parametric PSF estimated by Poduval *et al.* [17], *i.e.*,  $\mathbf{h}_{p_1}$ ; (2) the non-parametric PSF of Figure 5.7-(b), *i.e.*,  $\tilde{\mathbf{h}}_{np}$ ; and (3) the parametric/non-parametric PSF of Figure 5.8-(c), *i.e.*,  $\tilde{\mathbf{h}}_{p_2-np}$ .

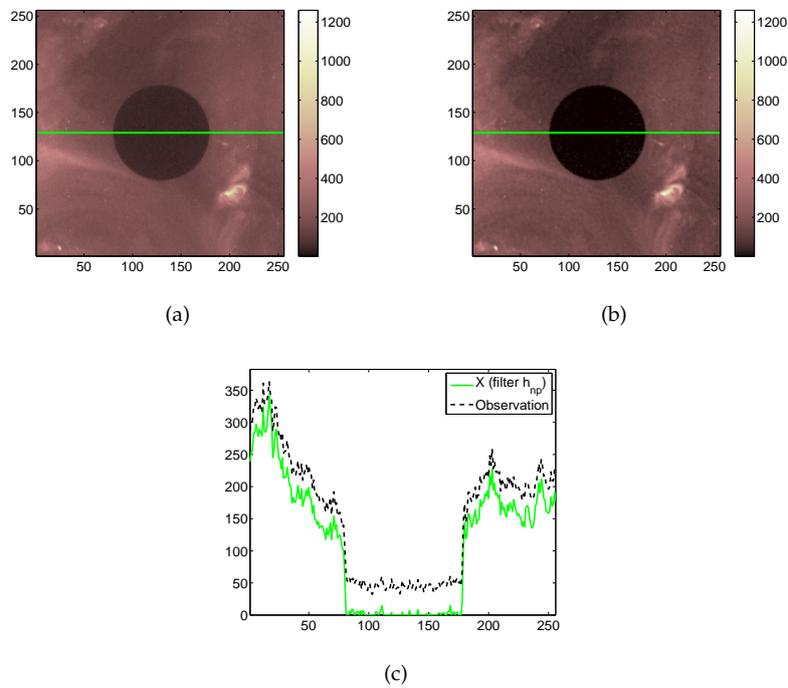
Filter	$S_X/S_Y$ Venus	$S_X/S_Y$ Moon
$\mathbf{h}_{p_1}$	$0.74 \times 10^{-2}$	$1.30 \times 10^{-2}$
$\tilde{\mathbf{h}}_{np}$	$3.65 \times 10^{-2}$	$3.64 \times 10^{-2}$
$\tilde{\mathbf{h}}_{p_2-np}$	$0.28 \times 10^{-2}$	$0.65 \times 10^{-2}$

**Table 5.2** Disk intensity ratio for the different filters.

We observe that, for both transits, the parametric/non-parametric model reaches lower disk apparent emissions. We also note that, compared to the non-parametric filter estimation, the parametric model provide slightly better results in terms of reducing the disk apparent emissions. However, the non-parametric scheme presents the advantage of being generally applicable for

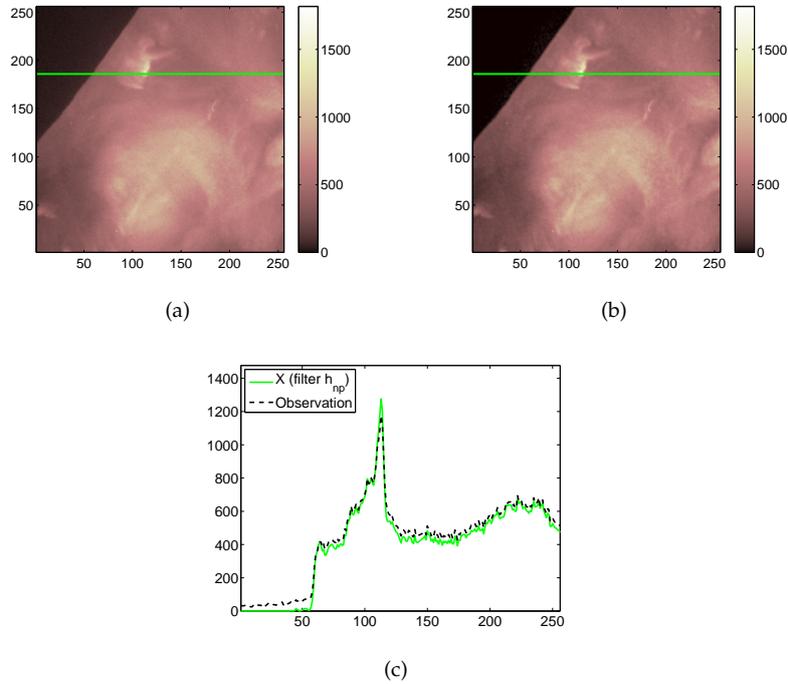
any optical instrument without the need of an exact model of the imaging process.

Figures 5.9 and 5.10 illustrate the image reconstruction results when using the non-parametric PSF, *i.e.*,  $\tilde{\mathbf{h}}_{\text{np}}$ . Figures 5.9-(a) and 5.10-(a) present the observed patches; Figures 5.9-(b) and 5.10-(b) depict the 2-D estimated images; and Figures 5.9-(c) and 5.10-(c) present the 1-D profiles that allow the observation of the disks of Venus and the Moon, respectively.



**Figure 5.9** Image reconstruction results for the Venus transit observed by SDO/AIA at 00:02 UT on June 6th 2012. Results are shown for the non-parametric filter, *i.e.*,  $\tilde{\mathbf{h}}_{\text{np}}$ . (a) Observed image, (b) 2-D reconstruction result, and (c) 1-D profile along  $y = 129$ . The figures on the top contain a line indicating the 1-D profiles shown in (c).

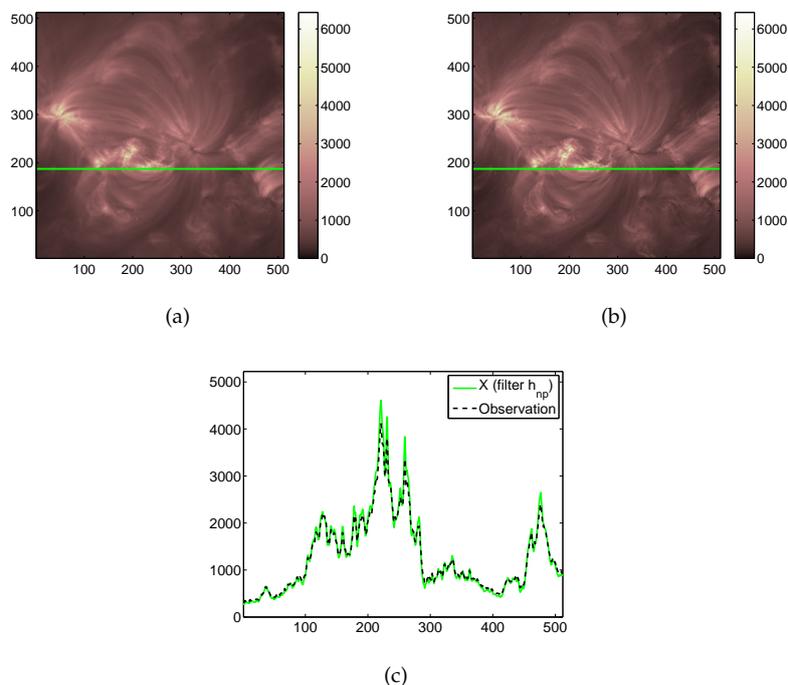
Let us note that, if the long-range effect  $\gamma$  is not removed from the observations, the parametric PSFs taken with a larger support of  $2048 \times 2048$  pixels are not able to eliminate the offset and, hence, are not able to remove the apparent emissions inside the disk of Venus.



**Figure 5.10** Image reconstruction results for the Moon transit observed by SDO/AIA at 13:00 UT on March 4th 2011. Results are shown for the non-parametric filter, *i.e.*,  $\tilde{h}_{np}$ . (a) Observed image, (b) 2-D reconstruction result, and (c) 1-D profile along  $y = 186$ . The figures on the top contain a line indicating the 1-D profiles shown in (c).

Finally, the estimated filters were also used to deconvolve non-transit images. For this, the non-blind deconvolution formulation from Section 5.5, using an adaptive  $\rho$ , was applied to the modified observations with  $\gamma = 43.3$  DN. The results are shown for a non-transit image taken by SDO/AIA at 10:00 UT on August 8th 2011. We selected a portion of the original image of  $512 \times 512$  pixels around the active region (see Figure 5.11-(a)). Figure 5.11-(b) depicts the 2-D estimated image using the non-parametric PSF, *i.e.*,  $\tilde{h}_{np}$ , and Figure 5.11-(c) shows a 1-D profile along  $y = 187$ .

We observe that the non-parametric PSF enhances the image, providing brighter active regions and coronal loops, and darker regions of lower intensity than in the original image. To compare those results with the other filters mentioned above, we take the ratio between the deconvolved images and the

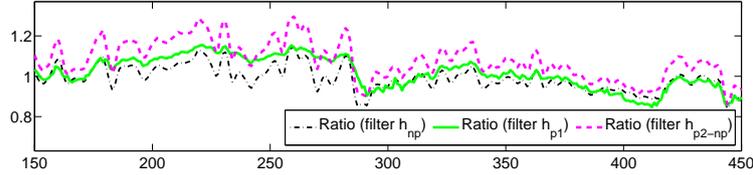


**Figure 5.11** Reconstruction results for a SDO/AIA image containing an active region. The image has been captured at 10:00 UT on August 8th 2011. Results are shown for the non-parametric filter, *i.e.*,  $\tilde{h}_{np}$ . (a) Observed image, (b) 2-D reconstruction result, and (c) 1-D profile along  $y = 187$ . The figures on the top contain a line indicating the 1-D profiles shown in (c).

observation (computing a pixel by pixel division), as shown in Figure 5.12. We notice that the deconvolution resulting from the combined parametric/non-parametric PSF presents a higher correction from the observations than the ones obtained using the other filters. In dark regions, the non-parametric PSF provides results similar to those obtained by the parametric PSF, however, in brighter areas the non-parametric PSF seems to be able to recover more details.

### 5.6.2.2 SECCHI/EUVI - Moon Transit

We consider three  $2048 \times 2048$  images from the transit ( $P = 3$ ) recorded by the 17.1 nm channel of EUVI. The images are calibrated with the `secchi_prep.pro` procedure available within the IDL *SolarSoft* library. Fol-

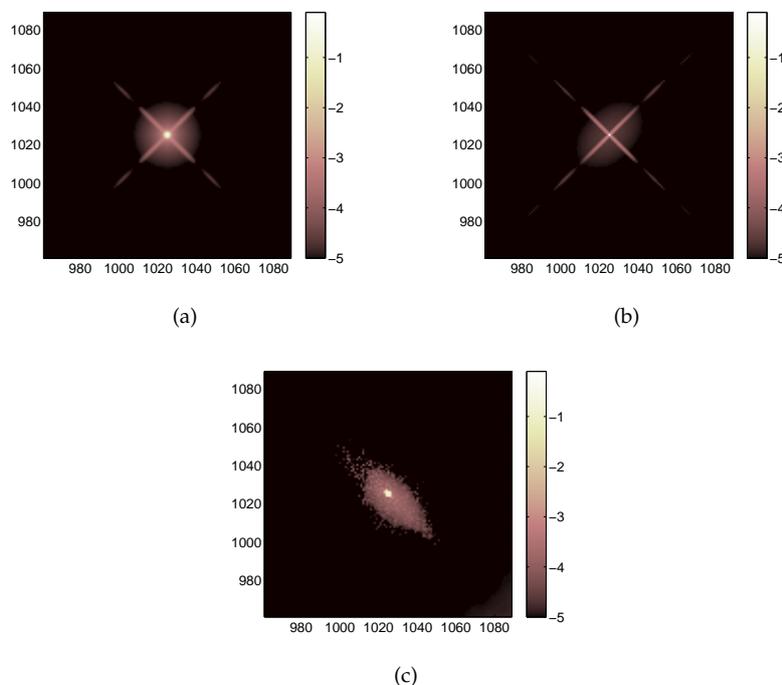


**Figure 5.12** Reconstruction results for the image in Figure 5.11-(a). The figure shows the ratio (computed using a pixel by pixel division) between the deconvolved and the observed images along  $y = 187$ , for the different considered filters:  $\tilde{\mathbf{h}}_{np}$ ,  $\mathbf{h}_{p1}$  and  $\tilde{\mathbf{h}}_{p2-np}$ . The horizontal axis is in correspondence with the horizontal axis of Figure 5.11-(c).

Following the practice described for the SDO/AIA Venus transit, each image was cropped using a  $512 \times 512$  window ( $N = 512^2$ ) centered around the Moon disk. The filter is assumed to have a limited support inside a  $129 \times 129$  pixel grid ( $b = 64$ ). This allows one to obtain the core of the PSF of around  $100 \times 100$  pixels, as observed by Shearer *et al.* [16], and encompass 99% of the energy of previously estimated PSFs.

The long-range PSF is modeled by a constant  $\gamma = 12.5$  DN, estimated by computing the mean intensity value (over five patches) on a disk of radius 10 pixels inside the Moon disk. The estimated noise variance using the RME provided  $\sigma_{RME}^2 = 2.21$  DN<sup>2</sup>. Following the same procedure as described before for SDO/AIA, we obtained that the value of  $\sigma$  that allows obtaining the whitest residual is  $\sigma = 2 \sigma_{RME} = 4.04$  DN.

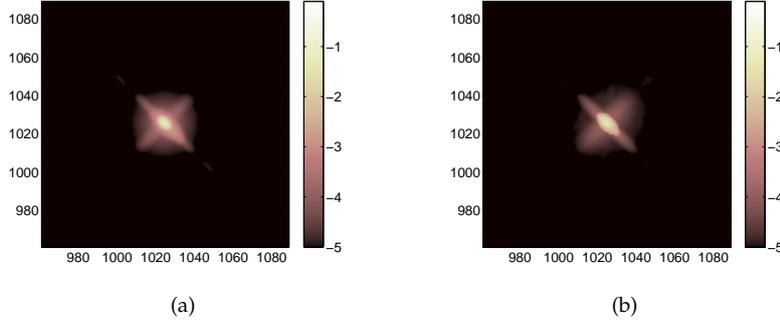
Hereafter, we provide a comparison between the following filters: (1) the parametric PSF given by the `euvi_psf.pro` procedure of *SolarSoft*, *i.e.*,  $\mathbf{h}_{p1}$ , the standard PSF used for SECCHI/EUVI image analysis; (2) the parametric PSF estimated by Shearer *et al.* [16], *i.e.*,  $\mathbf{h}_{p2}$ , and given by the `euvi_deconvolve.pro` procedure of *SolarSoft*; (3) the non-parametric PSF, *i.e.*,  $\tilde{\mathbf{h}}_{np}$ ; (4) the parametric/non-parametric PSF obtained by incorporating in the acquisition model the parametric PSF given by `euvi_psf.pro`, *i.e.*,  $\tilde{\mathbf{h}}_{p1-np}$ ; and (5) the parametric/non-parametric PSF obtained by incorporating in the acquisition model the parametric PSF given by `euvi_deconvolve.pro`, *i.e.*,  $\tilde{\mathbf{h}}_{p2-np}$ . The core of the parametric and non-parametric filters are presented in Figure 5.13 and the core of the resulting parametric/non-parametric filters are depicted in Figure 5.14.



**Figure 5.13** Logarithm of the filters for the SECCHI/EUVI telescope taken inside the set  $\Lambda$ . (a) Parametric PSF given by the `euvi_psf.pro` procedure of *SolarSoft*, i.e.,  $\mathbf{h}_{p1}$ . (b) Parametric PSF given by the `euvi_deconvolve.pro` procedure of *SolarSoft*, i.e.,  $\mathbf{h}_{p2}$ . (c) Non-parametric PSF, i.e.,  $\tilde{\mathbf{h}}_{np}$ .

We can observe that, while the two parametric PSFs favor both diagonals, the one given by the `euvi_deconvolve.pro` presents a slight dominance of that at 45 degrees. The estimated non-parametric PSF clearly favors one of the two diagonals. Nevertheless, it presents an orientation at  $-45$  degrees. The source of this difference in orientation is unknown and it should be further investigated. This, however, is beyond the scope of this work. Both parametric/non-parametric PSFs present a similar behavior, favoring both diagonals but with a slight dominance of that at  $-45$  degrees.

The filters from Figure 5.13 and Figure 5.14 were validated using the non-blind deconvolution described in Section 5.5. Similarly to what has been done for SDO/AIA, the value of  $\rho$  has been selected adaptively by optimizing the value of the residual whiteness. Let us note that, as observed empirically,



**Figure 5.14** Logarithm of the combined parametric/non-parametric filters for the SECCHI/EUVI telescope taken inside the set  $\Lambda$ . (a) Using the parametric PSF given by the `euvi_psf.pro` procedure of *SolarSoft*, i.e.,  $\tilde{\mathbf{h}}_{p_1-np}$ . (b) Using the parametric PSF given by the `euvi_deconvolve.pro` procedure of *SolarSoft*, i.e.,  $\tilde{\mathbf{h}}_{p_2-np}$ .

the value of  $\rho$  must be kept smaller than the one obtained in Algorithm 5.2. The filter validation was performed on the Moon transit image taken by SECCHI/EUVI at 08:02 UT on February 25th 2007 (see Figure 5.15-(a)). Let us note that this patch has not been used before for estimating  $\tilde{\mathbf{h}}$  and that the non-blind deconvolution is applied to the modified observation using  $\gamma = 12.5$  DN. We quantify the apparent Moon emissions by summing the pixel values inside the disk. Table 5.3 displays the disk intensity ratio for the image deconvolved using the different filters presented above.

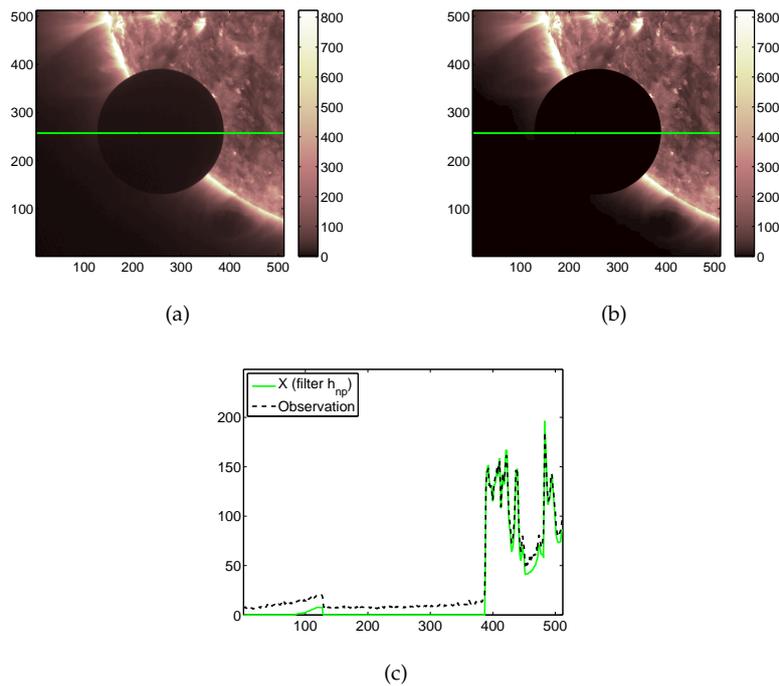
Filter	$S_X/S_Y$
$\mathbf{h}_{p_1}$	$5.72 \times 10^{-3}$
$\mathbf{h}_{p_2}$	$5.80 \times 10^{-3}$
$\tilde{\mathbf{h}}_{np}$	$2.02 \times 10^{-3}$
$\tilde{\mathbf{h}}_{p_1-np}$	$1.81 \times 10^{-3}$
$\tilde{\mathbf{h}}_{p_2-np}$	$1.29 \times 10^{-3}$

**Table 5.3** Moon disk intensity ratio for the different filters.

We notice that, as for the Venus transit images, when the long-range effect  $\gamma$  is not removed from the observations, the parametric PSFs considered in a larger support of  $1024 \times 1024$  pixels are not able to remove the offset

and to recover zero emissions inside the lunar disk. If the long-range effect is considered through the parameter  $\gamma$ , the non-parametric PSF presents a better behavior and is able to reduce more of the disk emissions as compared to the parametric PSFs (see Table 5.3). We also observe that the parametric/non-parametric PSFs reach lower disk apparent emissions.

In Figure 5.15, we illustrate the reconstruction results when using the non-parametric PSF, *i.e.*,  $\tilde{\mathbf{h}}_{\text{np}}$ . Figure 5.15-(a) presents the observed patch, Figure 5.15-(b) depicts the 2-D estimated image and Figure 5.15-(c) depicts a 1-D profile along  $y = 257$  (to allow the observation of the lunar disk).



**Figure 5.15** Image reconstruction results for the Moon transit observed by SECCHI/EUVI at 08:02 UT on February 25th 2007. Results are shown for the non-parametric filter, *i.e.*,  $\tilde{\mathbf{h}}_{\text{np}}$ . (a) Observed image, (b) 2-D reconstruction result, and (c) 1-D profile along  $y = 257$ . The figures on the top contain a line indicating the 1-D profiles shown in (c).

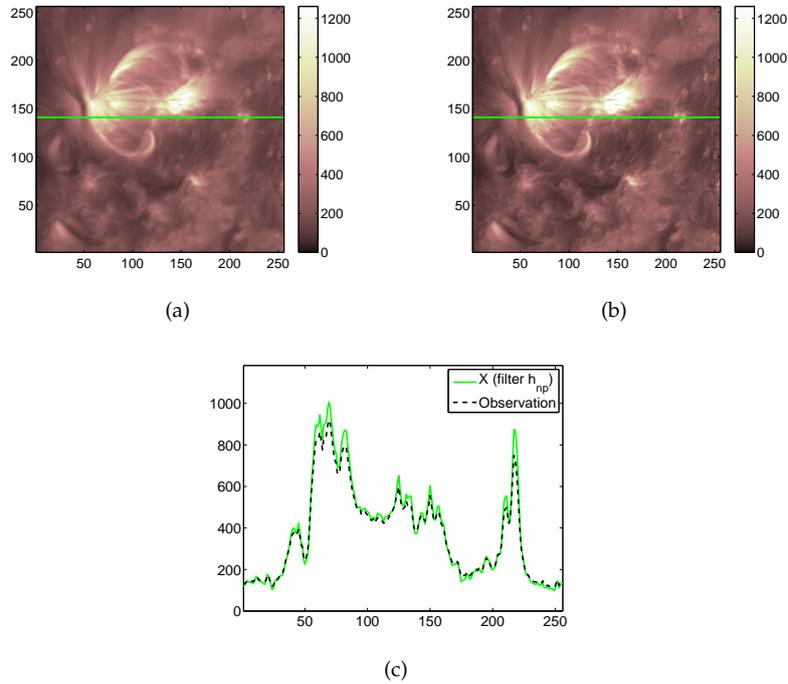
We observe that, in Figure 5.15, part of the off-limb portion of the image is forced to zero. Since this part of the image is fainter than the lunar disk, the proposed non-blind deconvolution method in Section 5.5 is not able to preserve such low intensities. However, finding a deconvolution method that would preserve these low intensities in the off-limb region (*e.g.*, by selecting other sparsity priors) is beyond the scope of this work.

The estimated filters were also used to deconvolve non-transit images. For this, the non-blind deconvolution formulation from Section 5.5, using an adaptive  $\rho$ , was applied to the modified observations with  $\gamma = 12.5$  DN. The results are shown for a non-transit image taken by SECCHI/EUVI at 04:02 UT on February 25th 2007. We selected a portion of the original image of  $256 \times 256$  pixels around the active region (see Figure 5.16-(a)). Figure 5.16-(b) depicts the 2-D estimated image using the non-parametric PSF, *i.e.*,  $\tilde{\mathbf{h}}_{\text{np}}$ , and Figure 5.16-(c) shows a 1-D profile along  $y = 141$ .

Figure 5.16 shows that the estimated non-parametric PSF is able to enhance the image and provide more details. Similarly to SDO/AIA, we compare these results with the ones from the other PSFs by taking the ratio between the deconvolved images and the observation (computing a pixel by pixel division). Results are depicted in Figure 5.17. We observe that the non-parametric PSF provides similar results to those obtained by the parametric PSF given by the `euvi_psf.pro` procedure. These two PSFs are able to provide higher details than the parametric PSF given by the `euvi_deconvolve.pro` procedure. The deconvolution resulting from the combined parametric/non-parametric PSFs present a higher correction from the observations than the ones obtained using the other three filters. This corresponds to what was observed for the Moon transit images in Table 5.3.

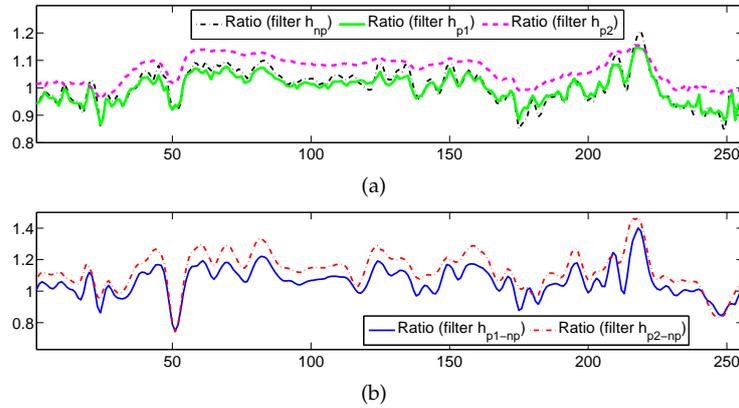
## 5.7 Conclusion

We have demonstrated how a non-parametric blind deconvolution technique is able to estimate the core of the PSF of an optical instrument with high quality. The quality of the estimated PSF core is comparable with the one provided by parametric models based on the optical characterization of the imaging instrument. We also demonstrate that, if the parametric PSF is in-



**Figure 5.16** Reconstruction results for a SECCHI/EUVI image containing an active region. The image has been captured at 04:02 UT on February 25th 2007. Results are shown for the non-parametric filter, *i.e.*,  $\tilde{h}_{np}$ . (a) Observed image, (b) 2-D reconstruction result, and (c) 1-D profile along  $y = 141$ . The figures on the top contain a line indicating the 1-D profiles shown in (c).

incorporated in the acquisition model, the blind deconvolution approach is able to provide a “corrected” PSF such that most of the apparent emissions inside the disk of a celestial body during a solar transit can be removed. Let us note that non-parametric techniques cannot outperform the accuracy of parametric methods, however, in situations where the telescope imaging model cannot be obtained due to some instrument’s properties (*e.g.*, PICARD/SODISM [158]), the non-parametric method presents a great advantage. Moreover, the proposed method is not specific to a given instrument but can be applied to any optical instrument provided that we have strong knowledge on some image pixels values and their exact location, such as the information available during the transit of the Moon or a planet. We have also shown that the use of the proximal alternating minimization technique proposed by Attouch *et al.* [104],



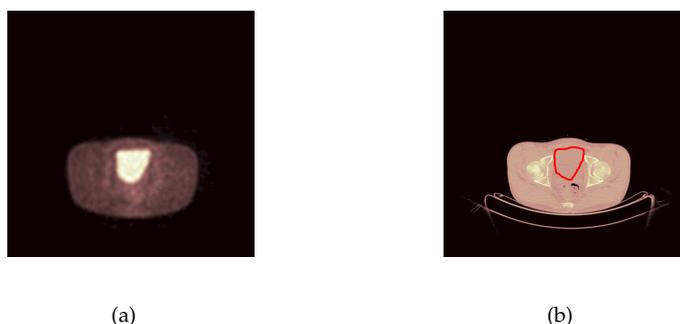
**Figure 5.17** Reconstruction results for the image in Figure 5.16-(a). The figure shows the ratio (computed using a pixel by pixel division) between the deconvolved and the observed images along  $y = 141$ . (a) Ratio for filters  $\tilde{h}_{np}$ ,  $h_{p1}$ , and  $h_{p2}$ . (b) Ratio for filters  $\tilde{h}_{p1-np}$  and  $\tilde{h}_{p2-np}$ . The horizontal axis is in correspondence with the horizontal axis of Figure 5.16-(c).

allows efficiently solving the non-convex problem of blind deconvolution with theoretical convergence guarantees. Furthermore, we show the importance of considering multiple observations for the same filter in order to provide a better conditioning to the filter estimation problem.

We are aware that the proposed method presents some limitations in terms of the considered noise model and the estimation of the PSF long-range effects. Fortunately, as discussed in Chapter 6, the optimization techniques we use in this chapter are flexible enough to open perspectives of improvement in future works.

It is important to note that the blind deconvolution method investigated in this chapter, does not exclusively apply to solar telescopes that have observed a solar transit. There exists many imaging applications where we have some knowledge about the intensity of a certain area of pixels in the image and its exact location. Positron emission tomography (PET) is a good example of an imaging modality where the proposed blind deconvolution method could be applied [162]. This imaging modality aims at obtaining metabolic information of the human body from measurements of gamma rays emitted by tracers that have been previously injected in the body [4, 163]. In PET imaging, the

recovered image containing the metabolic information is affected by the PSF of the instrument, which needs to be accurately estimated in order to obtain useful information from the images [163, Chapter 3]. Interestingly, the urinary bladder can be used as an anatomical landmark, since in a PET image the bladder is observed as an area where the pixels intensity remains constant (see Figure 5.18-(a)) [163, Chapter 14], *i.e.*, where the gradient of the image is known to be zero. The location of the urinary bladder in the image can be known with high precision by performing image registration of the available PET image with a computed tomography (CT) image (see Figure 5.18-(b)).

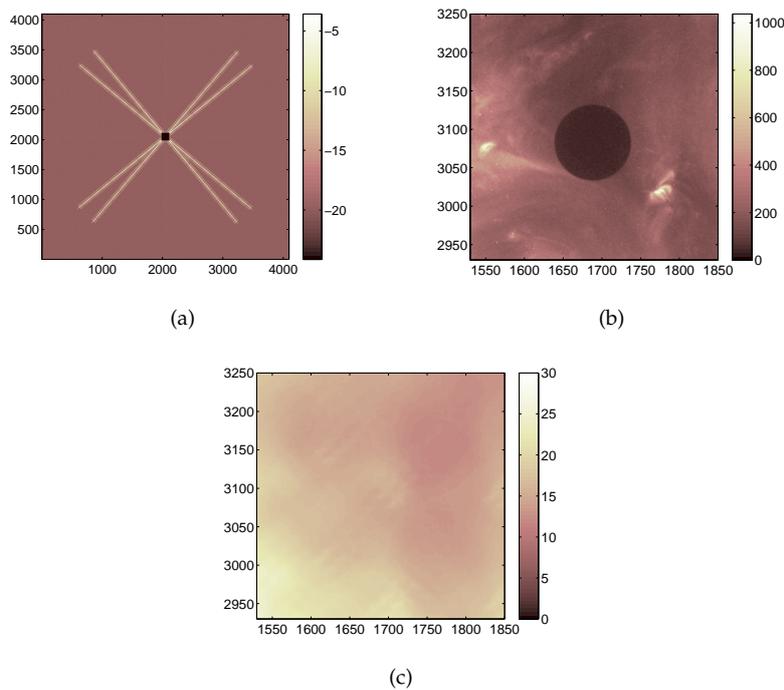


**Figure 5.18** Transaxial (a) PET and (b) CT images of a patient. Activity in the bladder in PET image is high and uniform. The bladder can be delineated from the high resolution CT image. Source: Cliniques Universitaires Saint-Luc.

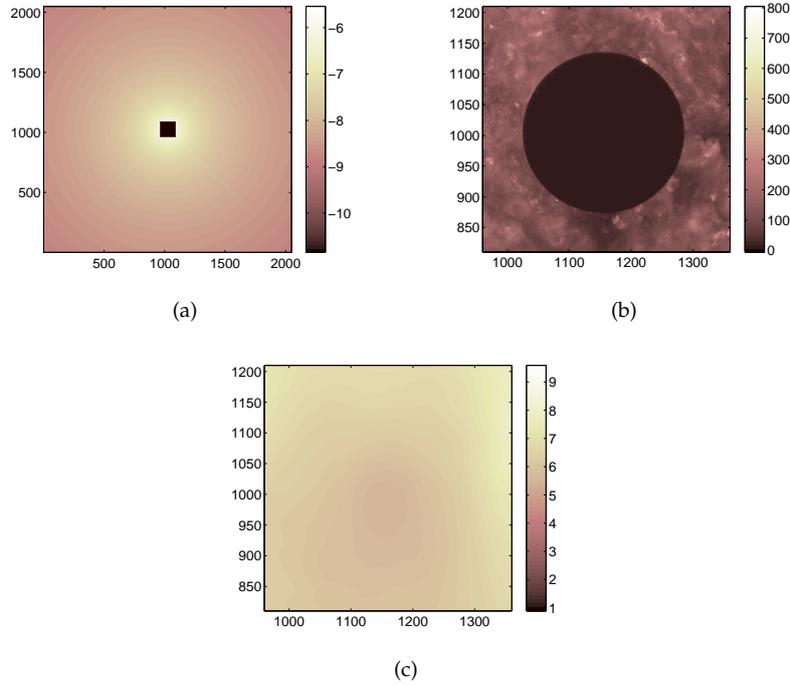
## Appendix 5.A On the Approximation of the Long-Range PSF Impact by a Constant

In order to compensate the restriction of the estimated filter to its core, in this chapter we assume that, for moderate solar intensities, the long-range effect of the PSF can be approximated by a constant inside a limited patch. This assumption is compatible with the observations made on the convolution of a solar image with a parametric PSF that only contains the long-range effect, *i.e.*, a PSF built by taking a standard parametric PSF containing long-range patterns and setting to zero the central pixels inside a window of size  $(2b + 1) \times (2b + 1)$  with  $b = 64$ .

For SDO/AIA, we use the parametric PSF estimated by Grigis *et al.* [18]. The validation was performed using a  $4096 \times 4096$  level 1 image from the Venus transit recorded by the 19.3 nm channel of AIA at 00:00:01 UT on June 6th 2012. These results are presented in Figure 5.19. For SECCHI/EUVI, we use the parametric PSF given by the `euvi_psf.pro` procedure of *SolarSoft*. The validation was performed using a  $2048 \times 2048$  image from the Moon transit recorded by the 17.1 nm channel of EUVI at 14:00:00 UT on February 25th 2007. The convolution results are presented in Figure 5.20. The resulting low frequency images in Figure 5.19-(c) and Figure 5.20-(c) validate the use of a constant  $\gamma$  to approximate the long-range effect.



**Figure 5.19** Validating the approximation of the long-range PSF impact by a constant inside the disk of Venus. (a) Logarithm of the filter used for the validation: parametric filter estimated by [18] with the central pixels set to zero inside a window of size  $(2b + 1) \times (2b + 1)$ , with  $b = 64$ . (b) Observed image inside a window containing the disk of Venus. (c) Convolved image inside a window containing the disk of Venus. The intensity within the window is almost constant.



**Figure 5.20** Validating the approximation of the long-range PSF impact by a constant inside the lunar disk. (a) Logarithm of the filter used for the validation: parametric filter given by the `euvi_psf.pro` procedure of *SolarSoft* with the central pixels set to zero inside a window of size  $(2b + 1) \times (2b + 1)$ , with  $b = 64$ . (b) Observed image inside a window containing the lunar disk. (c) Convolved image inside a window containing the lunar disk. The window intensity is almost constant.

## Appendix 5.B On the Convergence of the Proximal Alternating Algorithm

In this appendix, we demonstrate that Algorithm 5.1 satisfy Hypotheses 2.1, 2.2 and 2.3 required for the algorithm to converge to a critical point of (5.7).

To match the formulation in (2.39) with the blind deconvolution problem in (5.7), we set  $F(\mathbf{X}) = \rho \|\mathbf{S}_\Theta \mathbf{Y}^* \mathbf{X}\|_1 + \iota_{\mathcal{P}_0}(\mathbf{X})$ ,  $G(\mathbf{h}) = \iota_{\mathcal{D}}(\mathbf{h})$  and  $Q(\mathbf{X}, \mathbf{h}) = \frac{1}{2} \|\mathbf{Z} - \Phi(\mathbf{h})\mathbf{X}\|_{\mathbb{F}}^2$ . Note that both  $F(\mathbf{X})$  and  $G(\mathbf{h})$  belong to the

class  $\Gamma_0(\mathbb{R}^D)$ , for  $D \in \{N \times P, N\}$ . The function  $Q(\mathbf{X}, \mathbf{h})$  is a  $C^2$  function<sup>3</sup>, and hence its gradient  $\nabla Q$  is Lipschitz continuous on bounded subsets of  $\mathbb{R}^{N \times P} \times \mathbb{R}^N$  [82]. This demonstrates that the objective function in (5.7) satisfies Hypothesis 2.1.

Hypothesis 2.2 is also satisfied, given the definitions of the functions  $F$ ,  $G$  and  $Q$  provided above and the choice of cost-to-move parameters from Section 5.4.2.

The satisfaction of Hypothesis 2.3 is not immediately clear since we need to prove that the objective function in (5.7) satisfies the KL property. As stated in [82,104], all semi-algebraic functions (see Definition 2.7) satisfy the KL property. Therefore it suffices to prove that  $L(\mathbf{X}, \mathbf{h})$  is semi-algebraic.

Since the finite sum of semi-algebraic functions is semi-algebraic [81, 82], we need to prove that  $F(\mathbf{X})$ ,  $G(\mathbf{h})$  and  $Q(\mathbf{X}, \mathbf{h})$  are semi-algebraic. The function  $Q(\mathbf{X}, \mathbf{h}) = \frac{1}{2} \|\mathbf{Z} - \Phi(\mathbf{h})\mathbf{X}\|_{\mathbb{F}}^2$  is a real polynomial function, hence it is semi-algebraic [82]. The function  $G(\mathbf{h}) = \iota_{\mathcal{D}}(\mathbf{h})$  is also semi-algebraic since the set

$$\begin{aligned} \mathcal{D} &= \{\mathbf{v} \in \mathbb{R}_+^N : \|\mathbf{v}\|_1 = 1, \text{supp } \mathbf{v} = \Lambda\} \\ &= \left\{ \mathbf{v} \in \mathbb{R}^N : \sum_{i=1}^N a_i v_i - 1 = 0 \right\} \cap \left\{ \mathbf{v} \in \mathbb{R}^N : \sum_{i=1}^N v_i - 1 = 0 \right\} \cap \mathbb{R}_+^N, \\ &\quad \text{with } a_i = \begin{cases} 1 & \text{if } i \in \Lambda \\ 0 & \text{if } i \notin \Lambda \end{cases} \end{aligned}$$

is semi-algebraic by definition (see Definition 2.6).

Finally, for the function  $F(\mathbf{X}) = \rho \|\mathbf{S}_{\Theta} \Psi^* \mathbf{X}\|_1 + \iota_{\mathcal{P}_0}(\mathbf{X})$ , we need to analyze each term independently. Bolte *et al.* [82, Example 3.5] demonstrated that the  $\ell_p$  norm is semi-algebraic if  $p > 0$  is a rational number. Being  $\mathbf{S}_{\Theta} \Psi^*$  a real matrix, then the function  $\rho \|\mathbf{S}_{\Theta} \Psi^* \mathbf{X}\|_1$  is semi-algebraic. We also have that the set

$$\begin{aligned} \mathcal{P}_0 &= \{\mathbf{U} \in \mathbb{R}_+^{N \times P} : (\mathbf{u}_j)_i = 0 \text{ if } i \in \Omega_j\} \\ &= \left\{ \bigcap_{j=1}^P \bigcap_{i=1}^N \{\mathbf{U} \in \mathbb{R}_+^{N \times P} : a_{ij}(\mathbf{u}_j)_i = 0\} \right\} \cap \mathbb{R}_+^{N \times P}, \text{ with } a_{ij} = \begin{cases} 1 & \text{if } i \in \Omega_j \\ 0 & \text{if } i \notin \Omega_j \end{cases} \end{aligned}$$

<sup>3</sup> $C^2$  is the class of twice continuously differentiable functions.

is semi-algebraic by definition (see Definition 2.6) and, hence, the convex indicator on that set is semi-algebraic. Therefore, since  $F$  consists of the sum of two semi-algebraic functions, it is semi-algebraic.

## Appendix 5.C Algorithm Initialization

In this appendix, we describe in detail how we proceed with the initialization of the image ( $\mathbf{X}_0$ ), the filter ( $\mathbf{h}_0$ ) and the regularization parameter ( $\rho_0$ ) for Algorithm 5.1.

### 5.C.1 Regularization Parameter

The regularization parameter has an important role in solving the first step in Algorithm 5.1, since it determines the trade-off between the image regularization and the fidelity to the observations. Several works have studied the choice of this parameter when solving general inverse problems. In basis pursuit denoising, Donoho and Johnstone [41] proposed a universal value for  $\rho$  given by  $\sigma\sqrt{2\log(N)}$ . Since this value increases with the amount of samples  $N$ , it tends to provide overly smoothed images. Another state-of-the-art choice, proposed by Donoho and Johnstone [105], is based on the minimization of Stein's Unbiased Risk Estimate (SURE) [164]. This method provides accurate denoising results, however it requires complete knowledge of the degradation model which makes it unsuitable for our blind deconvolution problem. In this chapter, we use the Bayesian approach proposed by Chang *et al.* [42] to estimate the parameter  $\rho$  using the noise characteristics and the probability distribution function of the wavelet coefficients:

$$\rho = \frac{\sqrt{2} \sigma^2}{\tau}, \quad (5.16)$$

where  $\tau$  is the standard deviation of the wavelet coefficients of the signal to reconstruct. When we assume that they follow a Laplacian probability distribution with zero mean, the value of  $\tau$  can be estimated as:  $\frac{\sqrt{2}}{PQ} \sum_{j=1}^P \|(\mathbf{A})_j\|_1$ , where  $\mathbf{A} = \mathbf{S}_\Theta \Psi^* \mathbf{X}$  is the matrix containing the image wavelet coefficients inside the set  $\Theta$ .

Since the ground truth image  $\mathbf{X}$  is usually not available, the value of  $\tau$  in (5.16) is determined using the modified observations  $\mathbf{Z}$ , *i.e.*,

$\tau = \tau_Z = \frac{\sqrt{2}}{PQ} \sum_{j=1}^P \|(\mathbf{A}')_j\|_1$ , with  $\mathbf{A}' = \mathbf{S}_\Theta \mathbf{\Psi}^* \mathbf{Z}$ . This value is higher than the actual  $\rho$ , hence it will lead to over-regularized images. Therefore, we propose to refine the value of  $\rho$  by an iterative update based on the discrepancy principle between the actual noise energy in (5.4) and the energy of the residual image (5.10). By performing a maximum of five updates of the parameter  $\rho$ , we assure an appropriate image regularization. The iterative estimation of  $\rho$  is summarized in Algorithm 5.2.

### 5.C.2 Image and Filter

For the first value of  $\rho$ , the alternated algorithm (Algorithm 5.1) is initialized with the trivial solution, where the image is given by the modified observations  $\mathbf{Z}$  and the filter is given by the Kronecker function  $\delta_0$ . For the subsequent iterations on  $\rho$ , the image and the filter are initialized using the value of the previous iteration.

---

#### Algorithm 5.2 Iterative estimation of $\rho$

---

**Initialization:**  $\mathbf{X}^{(1)} = \mathbf{Z}$ ;  $\mathbf{h}^{(1)} = \delta_0$ ;  $\rho^{(1)} = (\sqrt{2}\sigma^2 / \tau_Z)$ ;  $\varepsilon = \sigma \sqrt{NP + 2\sqrt{NP}}$ ; MaxIter = 5

1: **for**  $l = 1$  to MaxIter **do**

*Images and Filter estimation:*

2: Estimate  $\mathbf{X}^{(l+1)}$  and  $\mathbf{h}^{(l+1)}$  using Algorithm 5.1 with  $\mathbf{X}_0 = \mathbf{X}^{(l)}$ ,  $\mathbf{h}_0 = \mathbf{h}^{(l)}$  and  $\rho_0 = \rho^{(l)}$

*Compute residual and whiteness measure:*

3:  $\mathbf{R}^{(l)} = \mathbf{Z} - \Phi(\mathbf{h}^{(l+1)}) \mathbf{X}^{(l+1)}$

4: Compute  $W(\mathbf{R}^{(l)})$  using (5.12)

*Parameter update:*

5:  $\varepsilon^{(l)} = \|\mathbf{R}^{(l)}\|_F$

6:  $\rho^{(l+1)} = \rho^{(l)} (\varepsilon / \varepsilon^{(l)})$

*Stop when residual is spectrally whiter:*

7: **if**  $W(\mathbf{R}^{(l+1)}) < W(\mathbf{R}^{(l)})$  **then** break.

8: **end if**

9: **end for**

10: Return  $\tilde{\mathbf{X}} = \mathbf{X}^{(l+1)}$  and  $\tilde{\mathbf{h}} = \mathbf{h}^{(l+1)}$

---

## Appendix 5.D Numerical Reconstruction Algorithms

For the sake of completeness, in this appendix, we briefly describe how to apply the algorithms from Section 2.4.2 to solve the two subproblems of Algorithm 5.1.

### 5.D.1 First Subproblem: Image Estimation

In the first subproblem, we are interested in finding the image candidate that minimizes

$$\begin{aligned} \mathbf{X}^{(k+1)} = \operatorname{argmin}_{\bar{\mathbf{X}} \in \mathbb{R}^{N \times P}} & \rho \|\mathbf{S}_\Theta \Psi^* \bar{\mathbf{X}}\|_1 + \frac{1}{2} \|\mathbf{Z} - \Phi(\mathbf{h}^{(k)}) \bar{\mathbf{X}}\|_{\mathbb{F}}^2 \\ & + \frac{\lambda_x^{(k)}}{2} \|\bar{\mathbf{X}} - \mathbf{X}^{(k)}\|_{\mathbb{F}}^2 + \iota_{\mathcal{P}_0}(\bar{\mathbf{X}}), \end{aligned} \quad (5.17)$$

a problem that contains a sum of four functions belonging to  $\Gamma_0(\mathbb{R}^D)$ , for  $D \in \{QP, NP\}$ . For this, we use the expanded Chambolle-Pock (CP) algorithm for  $L = 3$  in (2.25) (see Section 2.4.2.2).

To match the formulation in (2.25), we set  $F_1(\mathbf{V}_1) = \rho \|\mathbf{V}_1\|_1$  for  $\mathbf{V}_1 \in \mathbb{R}^{Q \times P}$  and  $F_2(\mathbf{V}_2) = \frac{1}{2} \|\mathbf{Z} - \mathbf{V}_2\|_{\mathbb{F}}^2$ ,  $F_3(\mathbf{V}_3) = \frac{\lambda_x^{(k)}}{2} \|\mathbf{V}_3 - \mathbf{X}^{(k)}\|_{\mathbb{F}}^2$  and  $H(\mathbf{U}) = \iota_{\mathcal{P}_0}(\mathbf{U})$  for  $\mathbf{V}_2, \mathbf{V}_3, \mathbf{U} \in \mathbb{R}^{N \times P}$ . Then, the expanded CP algorithm in (2.37) becomes:

$$\begin{cases} \mathbf{V}_1^{(t+1)} &= \operatorname{prox}_{\nu F_1^*} \left( \mathbf{V}_1^{(t)} + \nu \mathbf{S}_\Theta \Psi^* \bar{\mathbf{U}}^{(t)} \right) \\ \mathbf{V}_2^{(t+1)} &= \operatorname{prox}_{\nu F_2^*} \left( \mathbf{V}_2^{(t)} + \nu \Phi(\mathbf{h}^{(k)}) \bar{\mathbf{U}}^{(t)} \right) \\ \mathbf{V}_3^{(t+1)} &= \operatorname{prox}_{\nu F_3^*} \left( \mathbf{V}_3^{(t)} + \nu \bar{\mathbf{U}}^{(t)} \right) \\ \mathbf{U}^{(t+1)} &= \operatorname{prox}_{\frac{\mu}{3} H} \left\{ \mathbf{U}^{(t)} - \frac{\mu}{3} \left[ \Psi \mathbf{S}_\Theta^* \mathbf{V}_1^{(t+1)} + \left( \Phi(\mathbf{h}^{(k)}) \right)^* \mathbf{V}_2^{(t+1)} + \mathbf{V}_3^{(t+1)} \right] \right\} \\ \bar{\mathbf{U}}^{(t+1)} &= 2 \mathbf{U}^{(t+1)} - \mathbf{U}^{(t)} \end{cases}, \quad (5.18)$$

with  $\mathbf{U}^{(t)}$  tending to a minimizer  $\mathbf{X}^{(k+1)}$  of (5.17) for  $t \rightarrow +\infty$ .

Functions  $F_j^*$  are the convex conjugates of functions  $F_j$ , for  $j = \{1, 2, 3\}$ , and their proximal operators are computed via the proximal operator of  $F_j$  by

means of the conjugation property  $\text{prox}_{\nu F_j^*} \zeta = \zeta - \nu \text{prox}_{\frac{1}{\nu} F_j} \left( \frac{1}{\nu} \zeta \right)$  [80], for  $\zeta \in \mathbb{R}^{D \times P}$  and  $D \in \{Q, N\}$ . The proximal operator of  $F_1$  is given by the soft-thresholding operator in (2.24):

$$\text{prox}_{\frac{1}{\nu} F_1} \mathbf{V}_1 = \text{SoftTh} \left( \mathbf{V}_1, \frac{\rho}{\nu} \right),$$

while the ones of  $F_2$  and  $F_3$  are computed as follows:

$$\text{prox}_{\frac{1}{\nu} F_2} \mathbf{V}_2 = \frac{\mathbf{V}_2 + \frac{1}{\nu} \mathbf{Z}}{1 + \frac{1}{\nu}}, \quad \text{prox}_{\frac{1}{\nu} F_3} \mathbf{V}_3 = \frac{\mathbf{V}_3 + \frac{\lambda_x^{(k)}}{\nu} \mathbf{X}^{(k)}}{1 + \frac{\lambda_x^{(k)}}{\nu}}.$$

Finally, the proximal operator of  $H$  is given by the projection onto the convex set  $\mathcal{P}_0$ :

$$\left( \text{prox}_{\mu H} \mathbf{U} \right)_{ij} = \begin{cases} 0 & \text{if } i \in \Omega_j \\ (\mathbf{u}_j)_i & \text{otherwise} \end{cases}.$$

The step sizes  $\nu$  and  $\mu$  in (5.18) are adaptively selected using the procedure described in Section 3.7.2. Matrices  $\bar{\mathbf{U}}^{(1)}$  and  $\mathbf{V}_j$  with  $j \in \{1, 2, 3\}$  are initialized to zero, and matrix  $\mathbf{U}^{(1)}$  is initialized as explained in Section 5.C.2 for the image. Algorithm (5.18) stops when the relative error between two successive iterations reaches the value  $10^{-5}$ , *i.e.*, when  $\|\mathbf{U}^{(t+1)} - \mathbf{U}^{(t)}\|_F / \|\mathbf{U}^{(t)}\|_F \leq 10^{-5}$ .

### 5.D.2 Second Subproblem: Filter Estimation

In the second subproblem, we are interested in finding the filter candidate that minimizes

$$\mathbf{h}^{(k+1)} = \underset{\bar{\mathbf{h}} \in \mathbb{R}^N}{\text{argmin}} \frac{1}{2} \sum_{j=1}^P \|\mathbf{z}_j - \Phi(\mathbf{x}_j^{(k+1)}) \bar{\mathbf{h}}\|_2^2 + \frac{\lambda_h^{(k)}}{2} \|\bar{\mathbf{h}} - \mathbf{h}^{(k)}\|_2^2 + \iota_{\mathcal{D}}(\bar{\mathbf{h}}), \quad (5.19)$$

with  $\Phi(\mathbf{x}_j^{(k+1)}) \in \mathbb{R}^{N \times N}$  a *block-circulant with circulant blocks* matrix of kernel  $\mathbf{x}_j^{(k+1)}$  (see Section 5.2.1).

This problem has the shape of (2.26) with  $G(\mathbf{u}) = \frac{1}{2} \sum_{j=1}^P \|\mathbf{z}_j - \Phi(\mathbf{x}_j^{(k+1)}) \mathbf{u}\|_2^2 + \frac{\lambda_h^{(k)}}{2} \|\mathbf{u} - \mathbf{h}^{(k)}\|_2^2$  and  $F_1(\mathbf{u}) = \iota_{\mathcal{D}}(\mathbf{u})$ , for  $\mathbf{u} \in \mathbb{R}^N$ . Since  $G : \mathbb{R}^N \rightarrow \mathbb{R}$  is convex and differentiable with gradient  $\nabla G$

and  $F_1 \in \Gamma_0(\mathbb{R}^N)$ , the problem can be solved using the accelerated proximal gradient (APG) algorithm described in Section 2.4.2.1.

The APG iterations in (2.28) provide a solution  $\mathbf{u}^{(t)}$  that tends to a minimizer  $\mathbf{h}^{(k+1)}$  of (5.19) for  $t \rightarrow +\infty$ . In order to apply the APG iterations in (2.28), we must compute the gradient of  $G$  and the proximal operator of  $F_1$ . The gradient of  $G$  is given by:

$$\nabla G(\mathbf{u}) = - \sum_{j=1}^P \left( \Phi(\mathbf{x}_j^{(k+1)}) \right)^* \left( \mathbf{z}_j - \Phi(\mathbf{x}_j^{(k+1)})\mathbf{u} \right) + \lambda_h^{(k)} \left( \mathbf{u} - \mathbf{h}^{(k)} \right).$$

The proximal operator of  $F_1$ , computed via (2.23), is given by the projection on the convex set  $\mathcal{D}$ :

$$\text{prox}_{\nu F_1} \mathbf{u} = \text{proj}_{\mathcal{D}} \mathbf{u} = \begin{cases} \text{proj}_{\mathcal{P}_S} \mathbf{u}_\Lambda & \text{if } \mathbf{u} \in \Lambda \\ \mathbf{0}_{|\Lambda^c|} & \text{otherwise} \end{cases},$$

with  $\text{proj}_{\mathcal{P}_S}$  denoting the projection on the probability simplex, defined as  $\text{proj}_{\mathcal{P}_S} \boldsymbol{\zeta} = (\boldsymbol{\zeta} - \lambda \mathbf{1}_{|\Lambda|})_+$  for  $\boldsymbol{\zeta} \in \mathbb{R}^{|\Lambda|}$  [51]. The parameter  $\lambda$  can be found by solving the following equation  $\sum_{i=1}^{|\Lambda|} \max\{0, |\zeta_i| - \lambda\} = 1$ , using, for instance, the bisection method [149].

The step size  $\nu$  in (2.28) is adaptively selected using the line search of Beck and Teboulle [89]. The vector  $\mathbf{x}^{(1)}$  is initialized as explained in Section 5.C.2 for the filter and the relaxation parameter  $\beta^{(1)}$  is initialized with zero such that the value of  $\mathbf{x}^{(0)}$  (a zero vector) does not have any importance. The algorithm stops when the relative error between two successive iterations is  $10^{-5}$ .

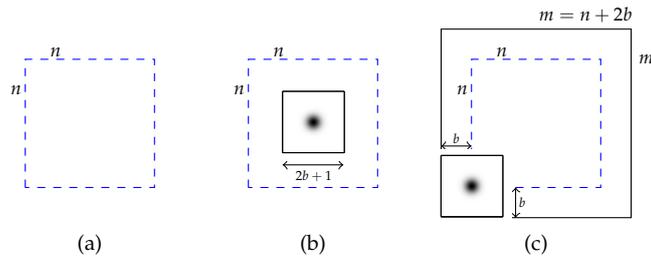
## Appendix 5.E Convolutions with Unknown Boundary Conditions

As explained in Section 5.2, our blind deconvolution approach is realized in a restricted FoV of size  $n \times n$ . For all our computations, as those realized in Algorithm 5.1, fast convolution methods exploiting the FFT must be performed. Therefore, it is important to properly handle the frontiers of our images and avoid both implicit FFT frontier periodizations and the influence of unobserved image values integrated by the filter extension. Inspired by the

work of Almeida and Figueiredo [61], we implement their *unknown boundary conditions*, where instead of expanding the observation using zero-padding as in [25, 165], the boundaries are considered unknown in the convolution process and then, by properly selecting the pixels inside the known boundaries, the observed image is obtained. In a nutshell, the unknown boundary conditions method proceeds by considering every convolution in a bigger space obtained by expanding the original one by the radius of the filter in all directions, *i.e.*, by a border of width  $b$ . Later, the optimization methods can freely set the values in this border. We only impose that the correct convolution values are those selected in the former domain.

Mathematically, we consider a  $m \times m$  pixel grid of  $M = m^2$  samples, with  $m = n + 2b$ . As illustrated in Figure 5.21, we use this grid to define a larger image  $\mathbf{x}_{L,j} \in \mathbb{R}^M$  and a larger PSF  $\mathbf{h}_L \in \mathbb{R}^M$ . The observation is obtained by selecting the pixels inside the boundaries of the convolved image using a selection operator  $\mathbf{S}_N \in \mathbb{R}^{N \times M}$ . The problem in (5.2) can be rewritten as:

$$\mathbf{Y} = \mathbf{S}_N \Phi(\mathbf{h}_L) \mathbf{X}_L + \mathbf{N}. \quad (5.20)$$



**Figure 5.21** Illustration of the dimensions of the deconvolution problem when considering *unknown boundary conditions*. (a)  $n \times n$  pixel grid where the observations are defined. (b) PSF defined in a  $n \times n$  pixel grid and with a limited support inside the set  $\Lambda$  of size  $(2b + 1)^2$  (see Section 5.3.2). (c)  $m \times m$  pixel grid where the larger image and PSF are defined for the reconstruction.

The regularized blind deconvolution problem in (5.6) transforms as follows:

$$\begin{aligned}
& \min_{\bar{\mathbf{X}}, \bar{\mathbf{h}}} \rho \|\mathbf{S}_\Theta \mathbf{\Psi}^* \bar{\mathbf{X}}\|_1 + \frac{1}{2} \|\mathbf{Z} - \mathbf{S}_N \Phi(\bar{\mathbf{h}}) \bar{\mathbf{X}}\|_{\mathbb{F}}^2 \\
& \text{s.t. } (\bar{x}_j)_i = 0 \text{ if } i \in \Omega_j; (\bar{x}_j)_i \geq 0 \text{ otherwise} \\
& \bar{\mathbf{h}} \in \mathcal{PS}; \text{supp } \bar{\mathbf{h}} \in \Lambda
\end{aligned} \tag{5.21}$$

with  $\bar{\mathbf{X}} \in \mathbb{R}^{M \times P}$  and  $\bar{\mathbf{h}} \in \mathbb{R}^M$ . The actual image  $\tilde{\mathbf{X}}$  and filter  $\tilde{\mathbf{h}}$  are obtained by applying the operator  $\mathbf{S}_N$  to both the estimated image  $\tilde{\mathbf{X}}_L$  and filter  $\tilde{\mathbf{h}}_L$ , respectively, *i.e.*,  $\tilde{\mathbf{X}} = \mathbf{S}_N \tilde{\mathbf{X}}_L$ ,  $\tilde{\mathbf{h}} = \mathbf{S}_N \tilde{\mathbf{h}}_L$ .

The unknown boundaries are also taken into consideration in the set  $\Theta$  by excluding from this set the wavelet coefficients that are affected by the image boundaries.

# Conclusions and Future Work

# 6

In this chapter, we summarize the contributions of the thesis and provide some perspectives for future work.

## 6.1 Summary of the Thesis

In the thesis, we study several inverse problems in imaging where we are interested in enhancing or generating images from deteriorated observations made by digital imaging devices.

In Chapter 2, we review the literature related to discrete inverse problems in imaging. We provide a formal definition of these problems and how to generate a linear forward model that describes the physical acquisition process. Then, we present a common method of solving inverse problems based on the minimization of the distance between the observations and the linear forward model, taking into account the different distortions present in the acquisition and modeling processes. Since this formulation often yields several possible solutions, we describe different functions and constraints that can be included in the minimization problem in order to reduce the amount of possible solutions to those that satisfy certain conditions imposed by the available information on the original image. Finally, we provide an overview of the optimization tools used in the thesis.

In Chapter 3, we analyze an inverse problem that exists in a novel imaging modality called optical deflectometric tomography (ODT). In this imaging modality, we aim at generating an image of the refractive index spatial distribution from light deflection measurements under multiple orientations. This would allow the optical characterization of transparent materials in industrial manufacturing processes. We provide first an overview of the physical acquisition process using a phase-shifting Schlieren deflectometer and we explain in detail how to build the associated linear discrete forward model. Then, we provide a thoughtful analysis on the estimation of the different sources of noise coming from the acquisition and modeling processes, which is later used to compute a bound on the distance between the available observations and the forward model. Assuming the refractive index map follows a bounded-variation model, its reconstruction is driven by the minimization of the total-variation norm under a fidelity constraint with the observations. In order to improve the stability of our method and obtain a unique solution for the refractive index map, two other realistic assumptions are added to the minimization problem: the image non-negativity and a frontier constraint. Numerically, the proposed approach relies on an accurate sensing model and on a primal-dual optimization scheme that can include the sensing operator and the constraints imposed on the image. The results on synthetic and experimental data show that the proposed method outperforms traditional tomographic reconstruction methods, providing a high quality image of the refractive index map from noisy and incomplete observations. This demonstrates that the ODT imaging modality can be made robust and compressive.

In Chapter 4, we study the 2-D phase unwrapping problem. This inverse problem affects many optical applications (*e.g.*, optical deflection tomography, radar imaging and interferometry) where the phase image cannot be directly observed but must be extracted from complex or modulated data recorded by the imaging device. In those cases, the phase is “wrapped” inside a certain interval (*e.g.*,  $[-\pi, \pi)$ ) through a modulo operation. We propose a robust convex optimization approach that, in contrast to existing unwrapping methods, aims at simultaneously unwrap and denoise the phase image. The proposed approach relies on a forward model formulated in the phase derivative domain, which allows relaxing this highly non-linear and non-convex problem.

The unwrapping inverse problem is formulated as a constrained optimization problem that enforces the proximity of the observations to the forward model, taking into account the characteristics of the noise in the phase image and the relaxation errors, while promoting a sparse phase prior. Using synthetically generated data, we demonstrate that, under different noise levels, the proposed approach compares favorably with state-of-the-art methods that perform the unwrapping and denoising in two separate steps. We also illustrate the performance of the proposed unwrapping method on a more realistic scenario by using experimental data from interferometric synthetic aperture radar (InSAR) imaging.

In Chapter 5, we analyze the well-known blind deconvolution problem. Many imaging applications, such as photography, astronomy and microscopy, are concerned by this inverse problem since the acquired image corresponds to the convolution of the ideal image and the response of the imaging device (generally unknown). Blind deconvolution aims at simultaneously retrieving the original enhanced image and the response of the instrument (also called point spread function or PSF) from corrupted observations. In this chapter, we focus on the astronomical imaging application and the study of solar telescopes. However, the proposed techniques can be extended to other instruments and applications. Most common methods in the literature impose parametric models on the PSF, making them specific for a given instrument. In contrast to those parametric methods, our method is based on the image regularization while imposing only soft constraints on the PSF and can thus be applied to any instrument. Our scheme relies on a wavelet analysis prior on the image and on the knowledge of specific values of well localized areas in the image (*e.g.*, information provided by solar transit images, dark calibration patterns or anatomical landmarks in the human body). The blind deconvolution problem is solved through a recently proposed alternating algorithm, which, in opposition to common alternating approaches, allows us to guarantee that a solution of the problem is always reached. The method also makes use of an innovative stopping criterion that is based on the spectral characteristics of the residual image, given by the difference between the observation and the model at each iteration. The non-parametric blind deconvolution method proposed in this chapter is validated on synthetic and experimental data. Results show that our method is able to estimate the core of the PSF with a similar

quality to parametric methods proposed in the literature. Moreover, we show that the parametric estimations can be incorporated in the acquisition model, which results in a “corrected” PSF that outperforms both the parametric and non-parametric methods.

## 6.2 Future Work

In Chapter 3, we assumed that the light rays traversing the object follow a straight line trajectory. This approximation allowed us to build a fast linear forward model based on the deflectometric Fourier slice theorem. From the light ray equation we know that the actual trajectory of the light when traversing an object is curved and depends itself on the refractive index map. Following the work of Antoine *et al.* [37], it would be interesting to study the improvement of the ODT sensing model by tracing the actual light trajectories. Remark that this would lead to non-convex optimization procedures and it would also break the fast computability of the forward sensing operator through the Fourier domain. An alternative strategy could consist in replacing the additive model used for the modeling error by a function that allows representing its true nature. Using the ray tracing method, we noticed that the modeling error depends on the value of the observed deflection angle, which reveals a multiplicative behavior. We also observed that the error points are limited, *i.e.*, the modeling error in the deflection angle measurements is sparse. Therefore, it would be interesting either to consider an  $\ell_1$  norm fidelity constraint or to remove from the fidelity constraint the points where the adopted approximation generates a higher error. A special attention must be paid to this last point, since this information is not available in the Fourier domain but in the domain of the deflection angle measurements.

In addition to diminishing the impact of the modeling error present in the simplified ODT sensing model of Chapter 3, we could also investigate additional constraints that can be applied to the refractive index map (RIM) in order to avoid losses in the RIM dynamics. For instance, we could envisage adding a constraint that aims at bounding the mean value of the RIM.

Chapter 3 was dedicated to the estimation of the 2-D RIM under the assumption that the refractive index is constant along the third dimension. In

order to widen the amount of possible objects that ODT can image, it would be interesting to address the 3-D RIM estimation. In this case, the whole (high dimensional) optimization driving the RIM reconstruction must be considered in a discretized three dimensional space. Additionally, the new acquisition process would have to be carefully analyzed in order to develop an accurate fast forward model. The work of Sudhakar *et al.* [27] can be used for a compressive acquisition of the deflection map that would be later used for recovering the 3-D RIM.

In Chapter 4, we adopted a relaxed forward model in the phase derivative domain in order to avoid working with the actual non-linear model. The resulting sensing model is only satisfied by a sufficiently smooth phase image. The proposed approach is able to consider the errors committed by these approximations through an  $\ell_1$  norm fidelity term. In the future, a stronger fidelity constraint should be built by removing from the reconstruction these few pixels where the noisy phase image is not smooth enough. However, the question remains in how to obtain a good estimation on the position of those discontinuities since it depends on the original phase image. This may be dependent on the application we are treating and the available prior information. For instance, in interferometric synthetic aperture radar (InSAR) imaging, we know that there is additional information such as coherence and quality maps that could be used, *e.g.*, to estimate the discontinuity points and the noise in the phase image.

The wavelet sparsity prior used in Chapter 4 should be analyzed more carefully, since it may prevent the algorithm from retrieving discontinuities that are present in the original phase image. Different prior models should be analyzed based on the expected characteristics of the phase image. We could envisage a sparsity basis that can be used for any imaging application or maybe one that is specifically adapted to a given application. Additionally, the problem of consistency of the unwrapped solution exposed in [72, 148] should be further investigated, since it seems possible that the reconstruction program requires additional constraints that would force the solution to be consistent with the observations.

In Chapter 5, we considered a simplified Gaussian average noise model. However, in reality, solar images are corrupted by many sources of noise. The

actual acquisition process can be summarized as follows. The incidence of photon flux on an extreme ultraviolet (EUV) telescope is converted to digital numbers (DN) through a series of steps, each potentially introducing some noise. The beam of photons impinges the optical system where the PSF acts as a blurring operator. Simultaneously, a spectral selection is performed on the signal before it reaches the detector, which has an heterogeneous response across its surface. Finally, the camera electronics convert electrons into DN, adding the read-out noise. The pixels in the resulting image can be modeled as the realization from a random variable, whose noise part can be decomposed into additive, Poissonian, and multiplicative degradations. In order to take into account the ensemble of degradations, we can generalize our additive white Gaussian noise (AWGN) model and include, as done in the literature [16], a Variance Stabilization Transform (VST) in the  $\ell_2$ -fidelity term, on both the observations and the convolution result. This can be done provided that we know the conversion between DN and the photon counts (*e.g.*, from instrument specifications). The method proposed in Chapter 5 is adaptable to this stabilized fidelity term through specific proximal operators or gradient descent methods [4, 53]. Notice, however, that such a stabilization cannot be applied only on the observations as a mere preprocessing of the data, while using afterwards other deconvolution methods assuming AWGN corruption. As an alternative to using a VST, the true Poisson distribution could also be used to design a specific fidelity term based on the negative log-likelihood of the posterior distribution induced by the observation model, *i.e.*, the Kullback-Leibler divergence [4, 34, 155]. However, it is unclear how to adapt the proximal alternating method to the resulting framework.

The blind deconvolution technique presented in Chapter 5 aimed at preventing strong assumptions on the shape of the core PSF. Nonetheless, additional PSF regularizations could be included in our scheme by, for instance, promoting the sparsity (in synthesis or in analysis) of this PSF in an appropriate basis (*e.g.*, wavelet basis), or by enforcing a certain decaying law of its amplitude in function of the radial distance. Both kind of priors can be expressed as convex costs (*e.g.*, with a  $\ell_1$  norm or a weighted  $\ell_\infty$  ball constraint) with *simple* proximal operators. These adaptations can thus be integrated in the algorithm with additional efforts for limiting the computational time of the more complex deconvolution procedure. Moreover, such additional priors

can help in enlarging the support of the PSF, where this one is truly estimated, hence providing an estimation of the long-range PSF. Additionally, the PSF could be estimated in a multi-scale fashion, with a finer scale for the center (core PSF) and a coarser scale for the long-range PSF. This would allow the latter to be included in the deconvolution problem rather than estimated in a preprocessing step.

To further stabilize the non-convexity of the blind deconvolution problem, the non-parametric method proposed in Chapter 5 was complemented by a convolutive combination of a known parametric PSF with an unknown non-parametric one. However, this construction is limited since, from the convolution theorem, no correction of the parametric PSF can be made in the part of its spectrum where it vanishes. We believe that future works should consider additive correction of the parametric PSF as suggested by Poduval *et al.* [17], a modification that the proposed algorithm could also include.

An important feature investigated in the thesis was the stopping criteria of the different algorithms. In Chapters 3 and 4, we analyzed and implemented different stopping criteria related to the properties of the primal-dual algorithm used for the reconstruction, while in Chapter 5 we exploited the characteristics of the residual image, given by the difference between the observations and the forward model. The latter has been proven particularly important when the spectral properties of the noise corrupting the model are known. Also, it can be used in any application regardless of the algorithm chosen for the reconstruction.

Finally, we would like to address the issues of algorithm convergence and speed that we encountered throughout the thesis. In Chapter 3, we observed a great improvement of the Chambolle-Pock (CP) algorithm when the step-size parameters are adapted through the iterations. The same adaptive algorithm was used in Chapter 5 for the image estimation. However, for the phase unwrapping problem described in Chapter 4, no advantage was observed due to the difficulty of setting the parameters in the adaptive method. Further studies with respect of these parameters should be addressed in the future. In addition to the adaptivity, the CP algorithm (as other algorithms) has been proved to work slowly when the problem is badly conditioned. Therefore,

it would be of benefit to perform a preconditioning of the global operator in order to stabilize the algorithms and improve the convergence results [96].

We note that in Chapter 3 the use of the CP algorithm to solve the ODT inverse problem is necessary due to the characteristics of the sensing operator. However, in the cases where the involved operators allow it, faster algorithms such as the alternating direction method of multipliers (ADMM) should be used in order to improve the convergence time. This is the case, for instance, of the phase unwrapping problem in Chapter 4. We also note that most of the algorithms in the thesis are implemented in Matlab, except for the NFFT algorithm and the wavelet transformations, which are compiled in C++. A faster implementation could be reached by adopting other programming languages in addition to parallel (GPU) processing techniques, a framework particularly adapted to proximal optimization.

# List of Publications

## Journal Papers

- **A. González**, V. Delouille and L. Jacques, "Non-parametric PSF estimation from celestial transit solar images using blind deconvolution," *J. Space Weather Space Clim.*, vol. 6, no. A1, January 2016, DOI: 10.1051/swsc/2015040.
- **A. González**, L. Jacques, C. De Vleeschouwer and P. Antoine, "Compressive optical deflectometric tomography: a constrained total-variation approach," *Inverse Prob. Imaging*, vol. 8, no. 2, pp. 421-457, May 2014, DOI: 10.3934/ipi.2014.8.421.

## Conference Papers

- S. Guérit, **A. González**, A. Bol, J. A. Lee and L. Jacques, "Blind Deconvolution of PET Images using Anatomical Priors," submitted in *3rd International Traveling Workshop for Interacting Sparse models and Technology (iTWIST16)*, March 2016.
- **A. González**, H. Jiang, G. Huang and L. Jacques, "Multi-Resolution Compressive Sensing Reconstruction," submitted in *Image Processing (ICIP), 2016 23rd IEEE International Conference on*, January 2016, <http://arxiv.org/abs/1602.05941>.
- **A. González** and L. Jacques, "Robust Phase Unwrapping by Convex Optimization," in *Image Processing (ICIP), 2014 21st IEEE International Conference on*, October 2014, DOI: 10.1109/ICIP.2014.7025343.

- P. Sudhakar, L. Jacques, **A. González**, X. Dubois, P. Antoine, and L. Joannes, "Compressive Acquisition of Sparse Deflectometric Maps," in *Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA)*, July 2013.
- **A. González**, L. Jacques and P. Antoine, "TV- $\ell_2$  Refractive Index Map Reconstruction from Polar Domain Deflectometry," in *1st International Traveling Workshop for Interacting Sparse models and Technology (iTWIST12)*, May 2012.
- P. Antoine, E. Foumouo, J.-L. Dewandel, D. Beghuin, **A. González** and L. Jacques, "Sparsity-driven optical deflection tomography with non-linear light ray propagation," in *Proceedings of the 5th International Conference on Optical Measurement Techniques for Structures & Systems (OPTIMESS 2012)*, April 2012.
- J. Stamatakis, **A. González**, B. Caby, S. Lefebvre, Y. Vandermeeren and B. Macq, "Kinematic features of reach and grasp movements in stroke rehabilitation using accelerometers," in *Bio-inspired Systems and Signal Processing (BIOSIGNALS), 2012 International Conference on*, February 2012.
- L. Jacques, **A. González**, E. Foumouo and P. Antoine, "Refractive Index Map Reconstruction in Optical Deflectometry Using Total-Variation Regularization," in *Wavelets and Sparsity XIV, SPIE Optical Engineering and Applications*, August 2011.
- **A. González**, L. Jacques, E. Foumouo and P. Antoine, "Primal-Dual TV Reconstruction in Refractive Deflectometry," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS11)*, June 2011.

# Bibliography

- [1] F. Pfeiffer, C. Kottler, O. Bunk, and C. David, "Hard X-ray phase tomography with low-brilliance sources," *Phys. Rev. Lett.*, vol. 98, p. 108105, March 2007.
- [2] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, December 2007.
- [3] A. Hormati, I. Jovanović, O. Roy, and M. Vetterli, "Robust ultrasound travel-time tomography using the bent ray model," in *Proc. SPIE, Medical Imaging*, vol. 7629, March 2010, p. 76290J.
- [4] S. Anthoine, J.-F. Aujol, Y. Boursier, and C. Mélot, "Some proximal methods for Poisson intensity CBCT and PET," *Inverse Prob. Imaging*, vol. 6, no. 4, pp. 565–598, November 2012.
- [5] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.*, vol. 19, no. 1, p. 010901, January 2014.
- [6] L. Joannes, F. Dubois, and J.-C. Legros, "Phase-shifting schlieren: high-resolution quantitative schlieren that uses the phase-shifting technique principle," *Appl. Opt.*, vol. 42, no. 25, pp. 5046–5053, September 2003.
- [7] J. Tepe, T. Schuster, and B. Littau, "A modified algebraic reconstruction technique taking refraction into account with an application in terahertz tomography, preprint," 2016. [Online]. Available: <http://arxiv.org/abs/1601.04496>

- [8] R. A. Howard, J. D. Moses, A. Vourlidas, J. S. Newmark, D. G. Socker, and et al., "Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI)," *Space Sci. Rev.*, vol. 136, pp. 67–115, April 2008.
- [9] C. E. DeForest, P. C. H. Martens, and M. J. Wills-Davey, "Solar coronal structure and stray light in TRACE," *ApJ*, vol. 690, pp. 1264–1271, January 2009.
- [10] W. D. Pesnell, B. J. Thompson, and P. C. Chamberlin, "The solar dynamics observatory (SDO)," *Sol. Phys.*, vol. 275, no. 1–2, pp. 3–15, January 2012.
- [11] R. E. Carrillo, J. D. McEwen, and Y. Wiaux, "Sparsity Averaging Reweighted Analysis (SARA): a novel algorithm for radio-interferometric imaging," *Mon. Not. R. Astron. Soc.*, vol. 426, no. 2, pp. 1223–1234, 2012.
- [12] J.-B. Sibarita, "Deconvolution microscopy," *Adv. Biochem. Eng. Biot.*, vol. 95, pp. 201–243, 2005.
- [13] F.-X. Dupé, M. Fadili, and J. L. Starck, "Deconvolution of confocal microscopy images using proximal iteration and sparse representations," in *Biomedical Imaging: From Nano to Macro (ISBI 2008)*, 2008 5th IEEE International Symposium on, May 2008, pp. 736–739.
- [14] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, July 2006.
- [15] M. S. C. Almeida and L. B. Almeida, "Blind and semi-blind deblurring of natural images," *Image Process., IEEE Trans.*, vol. 19, no. 1, pp. 36–52, January 2010.
- [16] P. Shearer, R. A. Frazin, A. O. Hero, III, and A. C. Gilbert, "The first stray light corrected extreme-ultraviolet images of solar coronal holes," *ApJ*, vol. 749, p. L8, April 2012.
- [17] B. Poduval, C. E. DeForest, J. T. Schmelz, and S. Pathak, "Point-spread functions for the extreme-ultraviolet channels of SDO/AIA telescopes," *Astrophys. J.*, vol. 765, no. 2, p. 144, March 2013.

- [18] P. Grigis, S. Yingna, and M. Weber, "AIA PSF characterization and image deconvolution," AIA team, Tech. Rep., 2013. [Online]. Available: <http://hesperia.gsfc.nasa.gov/ssw/sdo/aia/idl/psf/DOC/psfreport.pdf>
- [19] T. Lan, D. Erdogmus, S. J. Hayflick, and J. U. Szumowskil, "Phase unwrapping and background correction in MRI," in *Machine Learning for Signal Processing (MLSP), IEEE Workshop on*, October 2008.
- [20] J. Mei, A. Kirmani, A. Colaco, and V. K. Goyal, "Phase unwrapping and denoising for time-of-flight imaging using generalized approximate message passing," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, September 2013, pp. 364–368.
- [21] J. Idier, *Bayesian Approach to Inverse Problems*. Wiley, June 2008.
- [22] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society for Industrial and Applied Mathematics, 2001.
- [23] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, no. 10, pp. 3065–3095, May 2012.
- [24] G. Puy and P. Vandergheynst, "Robust image reconstruction from multiview measurements," *SIAM J. Imaging Sci.*, vol. 7, no. 1, pp. 128–156, January 2014.
- [25] M. Prato, R. Cavicchioli, L. Zanni, P. Boccacci, and M. Bertero, "Efficient deconvolution methods for astronomical imaging: algorithms and IDL-GPU codes," *A&A*, vol. 539, p. A133, March 2012.
- [26] E. Fomouo, J.-L. Dewandel, L. Joannes, D. Beghuin, L. Jacques, and P. Antoine, "Optical tomography based on phase-shifting schlieren deflectometry," in *Proc. SPIE*, vol. 7726, April 2010, pp. 77 260U–77 260U–8.
- [27] P. Sudhakar, L. Jacques, X. Dubois, P. Antoine, and L. Joannes, "Compressive imaging and characterization of sparse light deflection maps," *SIAM J. Imaging Sci.*, vol. 8, no. 3, pp. 1824–1856, September 2015.
- [28] A. González, L. Jacques, C. D. Vleeschouwer, and P. Antoine, "Compressive optical deflectometric tomography: a constrained total-variation

- minimization approach," *Inverse Prob. Imaging*, vol. 8, no. 2, pp. 421–457, May 2014.
- [29] A. González and L. Jacques, "Robust phase unwrapping by convex optimization," in *Image Processing (ICIP), 2014 21st IEEE International Conference on*, October 2012.
- [30] A. González, V. Delouille, and L. Jacques, "Non-parametric PSF estimation from celestial transit solar images using blind deconvolution," *J. Space Weather Space Clim.*, vol. 6, no. A1, January 2016.
- [31] J. Hadamard, "Sur les problèmes aux dérivées partielles et leur signification physique. (French) [On the partial derivatives problems and their physical meaning]," *Princeton University Bull.*, vol. 13, no. 4, pp. 49–52, April 1902.
- [32] J. Starck and M. Fadili, "An overview of inverse problem regularization using sparsity," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, November 2009, pp. 1453–1456.
- [33] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J-STARS Special Issue on Hyperspectral Image and Signal Processing*, vol. 5, no. 2, pp. 354–379, 2012.
- [34] M. Prato, A. L. Camera, S. Bonettini, and M. Bertero, "A convergent blind deconvolution method for post-adaptive-optics astronomical imaging," *Inverse Prob.*, vol. 29, no. 6, p. 065017, June 2013.
- [35] B. Zhang, M. Fadili, J. Starck, and J.-C. Olivo-Marin, "Multiscale variance-stabilizing transform for mixed-Poisson-Gaussian processes and its applications in bioimaging," in *Image Processing (ICIP), 2007 14th IEEE International Conference on*, vol. 6, September 2007, pp. VI – 233–VI – 236.
- [36] W. D. Pence, R. Seaman, and R. White, "A comparison of lossless image compression methods and the effects of noise," in *Astronomical Data Analysis Software and Systems XVIII. ASP Conference Series*, vol. 411, 2009, p. 25.

- [37] P. Antoine, E. Fomouo, J.-L. Dewandel, D. Beghuin, A. González, and L. Jacques, "Sparsity-driven optical deflection tomography with non-linear light ray propagation," in *Optical Measurement Techniques for Structures & Systems (OPTIMESS 2012), Proceedings of the 5th International Conference on*, April 2012.
- [38] F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3–4, pp. 246–254, December 1948.
- [39] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, August 2004.
- [40] D. L. Donoho, "De-noising by soft-thresholding," *Inf. Theory, IEEE Trans.*, vol. 41, no. 3, pp. 613–627, May 1995.
- [41] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [42] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *Image Process., IEEE Trans.*, vol. 9, no. 9, pp. 1532–1546, September 2000.
- [43] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," in *Signal Processing Conference (EUSIPCO), 2006 14th European*, September 2006, pp. 1–5.
- [44] M. J. Fadili and J.-L. Starck, "Sparse representation-based image deconvolution by iterative thresholding," in *Astronomical Data Analysis (ADA'06)*, September 2006.
- [45] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. Springer Publishing Company, Incorporated, 2010.
- [46] J. N. Franklin, "On Tikhonov's method for ill-posed problems," *Math. Comp.*, vol. 28, no. 128, pp. 889–907, October 1974.
- [47] M. Fuhry and L. Reichel, "A new Tikhonov regularization method," *Numer. Algor.*, vol. 59, no. 3, pp. 433–445, September 2011.

- [48] J. P. Oliveira, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Blind estimation of motion blur parameters for image deconvolution," in *Pattern Recognition and Image Analysis*, ser. Lecture Notes in Computer Science, J. Martí, J. M. Benedí, A. M. Mendonça, and J. Serrat, Eds. Springer Berlin Heidelberg, 2007, vol. 4478, pp. 604–611.
- [49] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational bayesian blind deconvolution using a total variation prior," *Image Process., IEEE Trans.*, vol. 18, no. 1, pp. 12–26, January 2009.
- [50] Y. Wiaux, L. Jacques, G. Puy, A. M. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Mon. Not. R. Astron. Soc.*, vol. 395, no. 3, pp. 1733–1742, May 2009.
- [51] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, January 2014.
- [52] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM J. Sci. Comput.*, vol. 31, pp. 890–912, 2009.
- [53] F.-X. Dupé, J. M. Fadili, and J.-L. Starck, "A proximal iteration for deconvolving Poisson noisy images using sparse representations," *Image Process., IEEE Trans.*, vol. 18, no. 2, pp. 310–321, February 2009.
- [54] T. Goldstein and S. Osher, "The Split Bregman method for L1 regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
- [55] M. S. C. Almeida and M. A. T. Figueiredo, "Parameter estimation for blind and non-blind deblurring using residual whiteness measures," *Image Process., IEEE Trans.*, vol. 22, no. 7, pp. 2751–2763, July 2013.
- [56] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, vol. 58, no. 301, pp. 13–30, 1963.
- [57] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 629–654, December 2008.
- [58] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

- [59] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Appl. Comput. Harmon. Anal.*, vol. 34, no. 1, pp. 30–56, January 2013.
- [60] J.-L. Starck and F. Murtagh, "Image restoration with noise suppression using the wavelet transform," *A&A*, vol. 288, no. 1, pp. 342–348, August 1994.
- [61] M. S. C. Almeida and M. A. T. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *Image Process., IEEE Trans.*, vol. 22, no. 8, pp. 3074–3086, August 2013.
- [62] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comput. Harmon. Anal.*, vol. 31, no. 1, pp. 59–73, July 2011.
- [63] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Process., IEEE Trans.*, vol. 54, no. 11, pp. 4311–4322, November 2006.
- [64] S. Gleichman and Y. C. Eldar, "Blind compressed sensing," *Inf. Theory, IEEE Trans.*, vol. 57, no. 10, pp. 6958–6975, October 2011.
- [65] J.-L. Starck, F. Murtagh, and J. Fadili, *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. New York, NY, USA: Cambridge University Press, 2010.
- [66] L. Jacques, L. Duval, C. Chaux, and G. Peyré, "A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity signal processing," *Signal Process.*, vol. 91, no. 12, pp. 2699–2730, December 2011.
- [67] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.
- [68] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *Inf. Theory, IEEE Trans.*, vol. 50, no. 10, pp. 2231–2242, October 2004.

- [69] S. Chen and D. Donoho, "Basis pursuit," in *Signals, Systems and Computers. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on*, November 1994.
- [70] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, 2008.
- [71] H. Rauhut and R. Ward, "Interpolation via weighted  $\ell_1$  minimization, preprint," 2013. [Online]. Available: <http://arxiv.org/abs/1308.0759>
- [72] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, D. Psaltis, and M. Unser, "Isotropic inverse-problem approach for two-dimensional phase unwrapping," *J. Opt. Soc. Am. A*, vol. 32, no. 6, pp. 1092–1100, June 2015.
- [73] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1-4, pp. 259–268, November 1992.
- [74] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision*, vol. 40, no. 1, pp. 120–145, 2011.
- [75] W. Cong, J. Yang, and G. Wang, "Differential phase-contrast interior tomography," *Phys. Med. Biol.*, vol. 57, no. 10, pp. 2905–2914, May 2012.
- [76] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.
- [77] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," *Image Process., IEEE Trans.*, vol. 16, no. 12, pp. 2992–3004, December 2007.
- [78] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 492–526, 2010.
- [79] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magn. Reson. Med.*, vol. 65, no. 2, pp. 480–491, February 2011.

- [80] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ser. Springer Optimization and Its Applications, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. Springer New York, 2011, pp. 185–212.
- [81] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods," *Math. Program.*, vol. 137, no. 1-2, pp. 91–129, February 2013.
- [82] J. Bolte, S. Sabach, and M. Teboulle, "Proximal alternating linearized minimization for nonconvex and nonsmooth problems," *Math. Program.*, vol. 146, no. 1-2, pp. 459–494, August 2014.
- [83] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1997.
- [84] N. Komodakis and J.-C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," *Signal Processing Magazine, IEEE*, vol. 32, no. 6, pp. 31–54, November 2015.
- [85] D. P. Bertsekas, *Convex Optimization Theory*. Athena Scientific, 2009.
- [86] F.-X. Dupe, M. J. Fadili, and J.-L. Starck, "Inverse problems with Poisson noise: primal and primal-dual splitting," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, September 2011, pp. 1901–1904.
- [87] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien. (French) [Dual convex functions and proximal points in a Hilbert space]," *CR Acad. Sci.*, vol. 255, no. 3, pp. 2897–2899, November 1962.
- [88] J. Liang, J. Fadili, and G. Peyré, "Local linear convergence of forward-backward under partial smoothness," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2014, pp. 1970–1978.

- [89] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery problems," in *Convex Optimization in Signal Processing and Communications*, D. Palomar and Y. Eldar, Eds. Cambridge University Press, January 2010, pp. 42–88.
- [90] K. Bredies and D. A. Lorenz, "Linear convergence of iterative soft-thresholding," *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 813–837, 2008.
- [91] M. Figueiredo and R. Nowak, "An EM algorithm for wavelet-based image restoration," *Image Process., IEEE Trans.*, vol. 12, no. 8, pp. 906–916, August 2003.
- [92] Y. Nesterov, "Gradient methods for minimizing composite objective function," *CORE discussion paper, Catholic University of Louvain*, vol. 76, 2007.
- [93] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," in *SIAM J. Imaging Sci.*, 2009, vol. 2, no. 1, pp. 183–202.
- [94] H. Raguét, J. Fadili, and G. Peyré, "A generalized forward-backward splitting," *SIAM J. Imaging Sci.*, vol. 6, no. 3, pp. 1199–1226, 2013.
- [95] M. Zhu and T. Chan, "An efficient primal-dual hybrid gradient algorithm for total variation image restoration," *UCLA CAM Report*, pp. 08–34, 2008.
- [96] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, November 2011, pp. 1762–1769.
- [97] T. Goldstein, E. Esser, and R. Baraniuk, "Adaptive primal-dual hybrid gradient methods for saddle-point problems, preprint," 2013. [Online]. Available: <http://arxiv.org/abs/1305.0546>
- [98] E. Esser, X. Zhang, and T. F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM J. Imaging Sci.*, vol. 3, no. 4, pp. 1015–1046, 2010.

- [99] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 1-3, pp. 293–318, 1992.
- [100] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *Image Process., IEEE Trans.*, vol. 19, no. 9, pp. 2345–2356, September 2010.
- [101] M. K. Ng, P. Weiss, and X. Yuan, "Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2710–2736, 2010.
- [102] S. Ravishankar and Y. Bresler, "Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging," *SIAM J. Imaging Sci.*, vol. 8, no. 4, pp. 2519–2557, 2015.
- [103] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, June 2011, pp. 233–240.
- [104] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, May 2010.
- [105] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.
- [106] L. Jacques, "Ondelettes, repères et couronne solaire (french) [Wavelets, frames and solar corona]," Ph.D. dissertation, Université catholique de Louvain, 2004.
- [107] G. H. Golub and C. F. Van Loan, *Matrix Computations*. The John Hopkins University Press, 1996/2001.
- [108] R. M. Gray, "Toeplitz and circulant matrices: A review," *Found. Trends in Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.

- [109] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics, 2011.
- [110] F. Zhang and Q. Zhang, "Eigenvalue inequalities for matrix product," *Autom. Control, IEEE Trans.*, vol. 51, no. 9, pp. 1506–1509, September 2006.
- [111] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, 3rd ed. Academic Press, 2008.
- [112] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.
- [113] R. R. Coifman and D. L. Donoho, "Translation-invariant de-noising," in *Wavelets and Statistics*, ser. Lecture Notes in Statistics, A. Antoniadis and G. Oppenheim, Eds. Springer New York, 1995, vol. 103, pp. 125–150.
- [114] J.-L. Starck, M. Elad, and D. Donoho, "Redundant Multiscale Transforms and Their Application for Morphological Component Separation," *Adv. Imaging Electron Phys.*, vol. 132, pp. 287–348, January 2004.
- [115] M. Elad, M. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, June 2010.
- [116] D. Beghuin, J.-L. Dewandel, L. Joannes, E. Fomouuo, and P. Antoine, "Optical deflection tomography with the phase-shifting schlieren," *Opt. Lett.*, vol. 35, no. 22, pp. 3745–3747, November 2010.
- [117] G. W. Faris and R. L. Byer, "Beam-deflection optical tomography of a flame," *Opt. Lett.*, vol. 12, no. 3, pp. 155–157, March 1987.
- [118] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *Inf. Theory, IEEE Trans.*, vol. 52, no. 12, pp. 5406–5425, December 2006.
- [119] D. L. Donoho, "Compressed sensing," *Inf. Theory, IEEE Trans.*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [120] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß, "Improved total variation-based CT image reconstruction applied to clinical data," *Phys. Med. Biol.*, vol. 56, no. 6, p. 1545, March 2011.

- [121] W. Cong, A. Momose, and G. Wang, "Fourier transform-based iterative method for differential phase-contrast computed tomography," *Opt. Lett.*, vol. 37, no. 11, pp. 1784–1786, June 2012.
- [122] L. Jacques, A. González, E. Fomouou, and P. Antoine, "Refractive index map reconstruction in optical deflectometry using Total-Variation regularization," in *Proc. SPIE*, vol. 8138, 2011, pp. 813 807–813 807–10.
- [123] C. Zhang, J. Wu, and W. Sun, "Applications of pseudo-polar FFT in synthetic aperture radiometer imaging," *PIERS Online*, vol. 3, no. 1, pp. 25–30, 2007.
- [124] A. González, L. Jacques, and P. Antoine, "TV- $\ell_2$  refractive index map reconstruction from polar domain deflectometry," in *1st International Traveling Workshop for Interacting Sparse models and Technology (iTWIST'12)*, May 2012.
- [125] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.
- [126] P. Hariharan, B. F. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," *Appl. Opt.*, vol. 26, no. 13, pp. 2504–2506, 1987.
- [127] J. Keiner, S. Kunis, and D. Potts, "Using NFFT 3—a software library for various Nonequispaced Fast Fourier Transforms," *ACM Trans. Math. Softw.*, vol. 36, no. 4, pp. 19:1–19:30, August 2009.
- [128] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Inf. Theory, IEEE Trans.*, vol. 52, no. 2, pp. 489–509, February 2006.
- [129] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *Signal Process. Mag., IEEE*, vol. 25, no. 2, pp. 83–91, March 2008.
- [130] K. Morita, *Applied Fourier Transform*. IOS Press, 1995.

- [131] L. Jacques, D. Hammond, and J. M. Fadili, "Dequantizing compressed sensing: When oversampling and non-gaussian constraints combine," *Inf. Theory, IEEE Trans.*, vol. 57, no. 1, pp. 559–571, January 2011.
- [132] E. van den Berg, M. P. Friedlander, G. Hennenfent, F. Herrmann, R. Saab, and Ö. Yilmaz, "Sparco: A testing framework for sparse reconstruction," Dept. Computer Science, University of British Columbia, Vancouver, Tech. Rep. TR-2007-20, October 2007.
- [133] S. Becker and J. Fadili, "A quasi-newton proximal splitting method," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 2618–2626.
- [134] F. Denis, O. Basset, and G. Gimenez, "Ultrasonic transmission tomography in refracting media: reduction of refraction artifacts by curved-ray techniques," *Med. Imag., IEEE Trans.*, vol. 14, no. 1, pp. 173–188, March 1995.
- [135] G. Steidl, "A note on fast fourier transforms for nonequispaced grids," *Adv. Comput. Math.*, vol. 9, no. 3-4, pp. 337–352, November 1998.
- [136] L. Ying, Z.-P. Liang, D. C. J. Munson, R. Koetter, and B. J. Frey, "Unwrapping of MR phase images using a markov random field model," *IEEE Transactions on Medical Imaging*, vol. 25, no. 1, pp. 128–136, 2006.
- [137] R. Bürgmann, P. A. Rosen, and E. J. Fielding, "Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation," *Annu. Rev. Earth Planet. Sci.*, vol. 28, pp. 169–209, May 2000.
- [138] A. Ferretti, A. Monti-Guarnieri, C. Prati, F. Rocca, and D. Massonnet, "InSAR principles: Guidelines for SAR interferometry processing and interpretation," European Space Agency, Tech. Rep., 2007. [Online]. Available: [http://www.esa.int/About\\_Us/ESA\\_Publications/](http://www.esa.int/About_Us/ESA_Publications/)
- [139] L. Ying, *Phase unwrapping*. John Wiley & Sons, Inc., 2006.
- [140] M. A. Herráez, D. Burton, M. Lalor, and M. Gdeisat, "A fast two-dimensional phase unwrapping algorithm based on sorting by reliability following a non-continuous path," *Appl. Opt.*, vol. 41, no. 35, pp. 7437–7444, 2002.

- [141] Y. Lu, X. Wang, and X. Zhang, "Weighted least-squares phase unwrapping algorithm based on derivative variance correlation map," *Optik*, vol. 118, no. 2, pp. 62–66, February 2007.
- [142] D. C. Ghiglia and M. D. Pritt, *Two-dimensional phase unwrapping: theory, algorithms, and software*. Wiley, 1998.
- [143] J. Bioucas-Dias and G. Valadão, "Phase unwrapping via graph-cuts," *Image Process., IEEE Trans.*, vol. 16, no. 3, pp. 698–709, 2007.
- [144] H. Huang, L. Tian, Z. Zhang, Y. Liu, Z. Chen, and G. Barbastathis, "Path-independent phase unwrapping using phase gradient and total-variation (TV) denoising," *Opt. Express*, vol. 20, no. 13, pp. 14 075–14 089, June 2012.
- [145] J. Bioucas-Dias, V. V. Katkovnik, J. Astola, and K. Egiazarian, "Absolute phase estimation: adaptive local denoising and global unwrapping," *Appl. Opt.*, vol. 47, no. 29, pp. 5358–5369, October 2008.
- [146] G. Valadão and J. Bioucas-Dias, "CAPE: combinatorial absolute phase estimation," *J. Opt. Soc. Am. A*, vol. 26, pp. 2093–2106, September 2009.
- [147] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: isoperimetry and processes*. Springer, 1991, vol. 23.
- [148] M. Pritt, "Congruence in least-squares phase unwrapping," in *1997 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2, August 1997, pp. 875–877.
- [149] R. L. Burden and J. D. Faires, *Numerical Analysis*. Cengage Learning; 9 edition, 2010.
- [150] S. Jefferies, K. Schulze, C. Matson, K. Stoltenberg, and E. K. Hege, "Blind deconvolution in optical diffusion tomography," *Opt. Express*, vol. 10, no. 1, pp. 46–53, January 2002.
- [151] P. R. Shearer, "Separable inverse problems, blind deconvolution, and stray light correction for extreme ultraviolet solar images," Ph.D. dissertation, The University of Michigan, 2013.

- [152] W. Dong, H. Feng, Z. Xu, and Q. Li, "A piecewise local regularized Richardson-Lucy algorithm for remote sensing image deconvolution," *Opt. Laser Technol.*, vol. 43, no. 5, pp. 926–933, 2011.
- [153] G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Opt. Lett.*, vol. 13, no. 7, pp. 547–549, July 1988.
- [154] S. Gburek, J. Sylwester, and P. Martens, "The TRACE telescope point spread function for the 171 Å filter," *Sol. Phys.*, vol. 239, pp. 531–548, December 2006.
- [155] D. A. Fish, J. G. Walker, A. M. Brinicombe, and E. R. Pike, "Blind deconvolution by means of the Richardson-Lucy algorithm," *J. Opt. Soc. Am. A*, vol. 12, no. 1, pp. 58–65, January 1995.
- [156] P. C. Martens, L. W. Acton, and J. R. Lemen, "The point spread function of the soft X-ray telescope aboard YOHKOH," *Sol. Phys.*, vol. 157, no. 1–2, pp. 141–168, March 1995.
- [157] D. B. Seaton, A. De Groof, P. Shearer, D. Berghmans, and B. Nicula, "SWAP observations of the long-term, large-scale evolution of the extreme-ultraviolet solar corona," *Astrophys. J.*, vol. 777, no. 1, p. 72, November 2013.
- [158] M. Meftah, J.-F. Hochedez, A. Irbah, A. Hauchecorne, P. Boumier, and et al., "Picard SODISM, a space telescope to study the sun from the middle ultraviolet to the near infrared," *Sol. Phys.*, vol. 289, no. 3, pp. 1043–1076, March 2014.
- [159] M. Simões, L. B. Almeida, J. Bioucas-Dias, and J. Chanussot, "A general framework for fast image deconvolution with incomplete observations. Applications to unknown boundaries, inpainting, superresolution, and demosaicing, preprint," 2016. [Online]. Available: <http://arxiv.org/abs/1602.01410>
- [160] J. R. Lemen, A. M. Title, D. J. Akin, P. F. Boerner, C. Chou, and et al., "The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO)," *Sol. Phys.*, vol. 275, no. 1–2, pp. 17–40, January 2012.

- [161] R. W. Nightingale, "AIA/SDO FITS keywords for scientific usage and data processing at levels 0, 1.0, and 1.5," Lockheed-Martin Solar and Astrophysics Laboratory (LMSAL), Tech. Rep., 2011.
- [162] S. Guérit, A. González, A. Bol, J. A. Lee, and L. Jacques, "Blind deconvolution of PET images using anatomical priors, preprint," 2016.
- [163] D. L. Bailey, D. W. Townsend, P. E. Valk, and M. N. Maisey, *Positron Emission Tomography: Basic Sciences*. Springer, 2005.
- [164] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Stat.*, vol. 9, no. 6, pp. 1135–1151, November 1981.
- [165] B. Anconelli, M. Bertero, P. Boccacci, M. Carbillet, and H. Lanteri, "Reduction of boundary effects in multiple image deconvolution with an application to LBT LINC-NIRVANA," *A&A*, vol. 448, no. 3, pp. 1217–1224, 2006.