High Performance Reconstruction Framework for Straight Ray Tomography: From Micro to Nano Resolution Imaging

Masih Nilchian

Thèse N° 6621 (mai 2015)

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

Prof. Jean-Philippe Thiran, *président*
Prof. Michael Unser, *directeur de thèse*
Prof. Françoise Peyrin, *rapporteur*
Prof. Carlos Óscar Sánchez Sorzano, *rapporteur*
Prof. Dimitri Van De Ville, *rapporteur*

École polytechnique fédérale de Lausanne—2015

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

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

Abstract

We develop a high-performance scheme to reconstruct straight-ray tomographic scans. We preserve the quality of the state-of-the-art schemes typically found in traditional computed tomography but reduce the computational cost substantially. Our approach is based on 1) a rigorous discretization of the forward model using a generalized sampling scheme; 2) a variational formulation of the reconstruction problem; and 3) iterative reconstruction algorithms that use the alternating-direction method of multipliers. To improve the quality of the reconstruction, we take advantage of total-variation regularization and its higher-order variants. In addition, the prior information on the support and the positivity of the refractive index are both considered, which yields significant improvements.

The two challenging applications to which we apply the methods of our framework are grating-based x-ray imaging (GI) and single-particle analysis (SPA). In the context of micro-resolution GI, three complementary characteristics are measured: the conventional absorption contrast, the differential phase contrast, and the small-angle scattering contrast. While these three measurements provide powerful insights on biological samples, up to now they were calling for a large-dose deposition which potentially was harming the specimens (*e.g.*, in small-rodent scanners). As it turns out, we are able to preserve the image quality of filtered back-projection-type methods despite the fewer acquisition angles and the lower signal-to-noise ratio implied by a reduction in the total dose of *in-vivo* grating interferometry. To achieve this, we first apply our reconstruction framework to differential phase-contrast imaging (DPCI). We then add Jacobian-type regularization to simultaneously reconstruct phase and absorption. The experimental results confirm the power of our method. This is a crucial step toward the deployment of DPCI in medicine and biology. Our algorithms have been implemented in the TOMCAT laboratory of the Paul Scherrer Institute.

In the context of near-atomic-resolution SPA, we need to cope with hundreds or thou-

sands of noisy projections of macromolecules onto different micrographs. Moreover, each projection has an unknown orientation and is blurred by some space-dependent point-spread function of the microscope. Consequently, the determination of the structure of a macromolecule involves not only a reconstruction task, but also the deconvolution of each projection image. We formulate this problem as a constrained regularized reconstruction. We are able to directly include the contrast transfer function in the system matrix without any extra computational cost. The experimental results suggest that our approach brings a significant improvement in the quality of the reconstruction. Our framework also provides an important step toward the application of SPA for the *de novo* generation of macromolecular models. The corresponding algorithms have been implemented in Xmipp.

Keywords: Discretization, variational formulation, iterative reconstruction, alternating-direction method of multipliers, grating-based x-ray imaging, single-particle analysis, phase-contrast imaging.

Résumé

Dans ce travail, nous développons des méthodes de reconstruction de haute-performances pour la tomographie sans diffraction. Nous obtenons des résultats comparables à l'état-de-l'art en tomographie par ordinateur en termes de qualité, tout en réduisant significativement le coût de calcul. Notre approche comporte trois aspects : 1) une discrétisation rigoureuse du modèle d'analyse en suivant un schéma d'échantillonnage généralisé, 2) une formulation variationnelle du problème de reconstruction, et 3) des algorithmes de reconstruction itératifs fondés sur la méthode intitulée *alternating direction method of multipliers*. Afin d'augmenter la qualité de la reconstruction, nous utilisons les techniques de régularisation basées sur la variation totale ou des variantes à des ordres supérieurs. De plus, la prise en compte des informations *a priori* sur le support et la positivité de l'indice de réfraction de l'objet d'étude conduit à des améliorations significatives.

Les méthodes que nous avons développées sont mises en pratique sur deux applications importantes : l'imagerie par rayon X pour l'interférométrie à réseau (en anglais GI pour *Grating interferometry*) et l'analyse de particule unique (en anglais SPA pour *single particle analysis*). Dans le contexte de la micro-résolution GI, trois caractéristiques complémentaires sont mesurées : le contraste d'absorption, le contraste de phase et le contraste de diffusion. Bien que ces trois grandeurs apportent des informations très utiles pour l'étude biologiques *in vivo*, jusquà maintenant leur obtention nécessitait une longue exposition qui pouvait être toxique pour le spécimen étudié (par exemple lors des scanners réalisés sur de petits rongeurs). Il apparaît cependant que nous sommes capables de reproducer la qualité de l'image reconstruite par des méthodes de rétroprojection filtrée, malgré le faible nombre de projections durant l'acquisition, ce qui implique une réduction de la dose totale de radiation. Pour atteindre ce but, nous appliquons d'abord notre méthode de reconstruction à l'imagerie par contrast de phase differential. Nous effectuons ensuite une régularisation basée sur le Jacobien afin de reconstruire simultanément la phase et

l'absorption. Les résultats expérimentaux confirment les performances de notre méthode. Il s'agit d'une étape cruciale pour l'utilisation de la l'imagerie par contraste de phase differentiel en médecine et en biologie. Nos algorithmes ont été implémentés et sont utilisés au laboratoire TOMCAT de l'Institut Paul Scherrer.

Pour la SPA, réalisée à une résolution quasi atomique, nous devons alors composer avec des centaines ou des milliers de projections de macromolécules, qui sont corrompues par du bruit. Chaque projection a une orientation inconnue et elles sont floutées par une fonction d'étalement du point (en anglais PSF pour *point spread function*) qui peut être différente pour chaque orientation. Ainsi, la détermination de la structure de la macromolecule nımplique pas seulement une tâche de reconstruction, mais aussi la deconvolution de chaque image projetée, ce que nous formulons comme un problème de reconstruction régularisé sous contrainte. Nous sommes capables d'inclure directement la function du transfer de contraste dans le système matriciel sans coût de calcul supplémentaire. Les résultats expérimentaux confirment que notre approche augmente significativement la qualité de la reconstruction. Le cadre proposé est de plus une étape importante pour l'application de la SPA pour l'interprétation de modèles de macromolécules. Les algorithmes correspondants ont été implémentés sur Xmipp.

Mots-clés : Discrétisation, formulation variationnelle, des algorithmes de reconstruction itératifs, *alternating direction method of multipliers*, l'imagerie par rayon X pour l'interférométrie à réseau, l'analyse de particule unique, l'imagerie par contraste de phase.

Dedicated to my mother, father, brother and wife
for their unconditional and everlasting love, support and encouragements.

Acknowledgement

This thesis would not have been completed without the help and support of many people. I take this opportunity to express my gratitude to all of them.

I would like to express my deepest gratitude to my thesis advisor, Prof. Michael Unser. This thesis would not have been possible without his guidance and support. Michael is a great advisor and a fantastic group leader. His doors were always open for discussions. Working with him was a great opportunity for me and allowed me to grow, not only scientifically but also on a personal level. I am really indebted to him for what he taught me. I am also deeply indebted to my master thesis supervisor, Prof. Mohammad-Reza Aref, for his well-known wisdom, support, kindness, and generosity of heart. His invaluable pieces of advice have been extremely useful throughout my life.

I would like to express my sincere to the president of the thesis jury, Prof. Jean-Philippe Thiran and my committee members, Prof. Francoise Peyrin, Prof. Carlos Oscar Sánchez Sorzano, Prof. Dimitri Van De Ville, for accepting to assess my thesis and for their helpful comments and suggestions.

I have been very fortunate to collaborate with Prof. Marco Stampanoni, Dr. Zhential Wang and Dr. Peter Modregger from TOMCAT beamline in Paul Scherrer institute. Marco, Zhentian and Peter have a great personality, sense of humor, and a mind full of innovative ideas. Working with them was another unique experience, they taught me many interesting things about the problem of grating-based X-ray imaging. I hope that our collaboration will continue and we will finally step over the radiation dose limitation for *in vivo* imaging. In addition, I have been very fortunate to collaborate with Prof. Carlos Oscar Sánchez Sorzano for introducing and patiently teaching me the subject of single particle analysis. He has a great personality, and a mind full of innovative ideas. He is a super active and hardworking person and one of the best collaborators with who I have had experienced. I hope that our collaboration will continue and we will finally step over the reconstruction resolution of

2.5$\mathring{A}$.

During my PhD research, I was very fortunate to have been in the biomedical imaging group (BIG). I would like to offer my gratitude to my friends and colleagues in BIG for all the enjoyable moments and memories. I thank the fellow present and past lab members of the Biomedical Imaging Group (BIG), Prof. Arash Amini, Anais Badoual, Dr. Jean-Charles Baritaux, Ayush Bhandari, Emrah Bostan, Dr. Nicolas Chenouard, Dr. Ning Chu, Dr. Ricard Delgado Gonzalo, Prof. Adrien Depeursinge, Julien Fageot, Dr. Denis Fortun, Dr. Matthieu Guerquin-Kern, Dr. Ulugbek Kamilov, Dr. Hagai Kirshner, Dr. Florian Luisier, Junhong Min, Pedram Pad, Zsuzsanna Püspöki, Dr. Sathish Ramani, Dr. Daniel Sage, Daniel Schmitter, Prof. Chandra Sekhar Seelamantula, Dr. Tomáš Škovránek, Dr. Martin Storath, Dr. Pouya D. Tafti, Raquel Terrés Cristofani, Dr. Philippe Thévenaz, Virginie Uhlmann, Dr. Cédric Vonesch, and Dr. John Paul Ward. And more particularly, I thank my present and past office mates Dr. Jean-Charles Baritaux, Ayush Bhandari, Julien Fageot, and Zsuzsanna Püspöki. I am also grateful to Manuelle Mary, and Nadia Macor for helping me out with various administrative matters inside and outside EPFL.

I am grateful to all my Iranian friends in Lausanne who made the life in Switzerland so amazing and memorable. In particular, I am so thankful to my friends known among iranians as Tir-Federal crew: Mohammad Parhizkar, Reza Parhizkar, Hamed Hassani, Farid Movahedi Naini and Saeid Haghighatshoar. We shared an apartment for more than a year and we had so many memories together. All the time that we spent together and all the friendship and dedication that we had for one another is beyond imagination. I hope the best wishes for them all. I am sure I will miss forever these amazing friends (to be honest, they are more than friends, like brothers) and the unforgettable moments that I spent with them.

I would have not succeeded either in life or in my research without the unconditional love, support and patience of my family. My deepest gratitude goes to my father for being an amazing father, a true friend and a great support all throughout my life. He taught me how to follow the path of honor, how to devote myself to others and how to always live away from pride and ignorance. I can not stop crying. Today is the father day, and I missed him too much (divanevar dooset daram baba!). I am deeply thankful to my mother for her unconditional love and all the sacrifices that she made in her life for her children and specially for me. There are no words that can express my deepest gratitude to her. Undoubtedly, she is the strongest woman I know. She was my mother and my father for the last ten years (binahaiat mamnunetun hastam madaram va aasheghetunam!). I would like to be specially thankful to my grandmother for her continuous supports, even more when we missed my lovely father. I would like to deeply thank my brother and his lovely

wife. Having such a nice older brother is an exceptional experience. He was my brother, my friend and in one word everything for me. All the time that we spent together and dedication that we had for one another is beyond imagination. I am specially thankful to his wife who has always been like true sister to me. I am specially thankful to my aunt's family, Majid, Farzaneh, Tohid and Tolou. They were kindly hosting me during my study in Tehran. Without any doubt, they are the best aunt's family that anyone can have. They are like my second father and mother, and my lovely brother and sister. I am really indebted to my aunt's husband, Majid, for what he taught me. Specially, I would like to thank Tohid for sharing his room with me for more than four years. I am also thankful to all my aunts, uncles and cousins. My special gratitudes go to my father-in-law and mother-in-law who always treated me like a son and trusted me in crucial occasions of my married life. I also deeply thank my sister-in-law who has always been like true sister to me.

Special thanks and recognitions go to my beloved wife for her deep love and heart-filling affection. I truly thank her for sticking by my side and believing in me even in those moments that I did not believe in myself. There are no words that can truly express my gratitude and love for her.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Computerized tomography (CT) revolutionized diagnostic medicine by enabling physicians to view the internal structure of organs in 1970s. It aims at reconstructing the object using several images taken. The acquisition is performed by illuminating a specimen (or an organ) by an electromagnetic wave along different orientations. In more detail, a monochromatic wave is represented by complex wave function

$$u(\mathbf{y},t) = a(\mathbf{y})\exp\left(j\phi(\mathbf{y})\right)\exp\left(j2\pi\nu t\right), \tag{1.1}$$

where a, ϕ, and ν are the amplitude, phase, and frequency of the wave, respectively [1]. The parameter $\mathbf{y} \in \mathbb{R}^2$ is a coordinate on a plane that is perpendicular to the direction of propagation of the wave, and t is the time parameter. The interaction of the wave with the specimen can be described by the complex refractive index $n(\mathbf{x}) = 1 - \delta(\mathbf{x}) + j\beta(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^3$ specifies a spatial coordinate. Then, the wave function over the exiting curve of the specimen in the context of diffraction-less electromagnetic plane wave (parallel beam with extremely small wavelength) is given by

$$u_o(\mathbf{y},t) = u(\mathbf{y},t)\exp\left(j\frac{2\pi}{\lambda}\int_{\mathbb{R}} n(s\boldsymbol{\theta} + \mathbf{P}_{\boldsymbol{\theta}^\perp}^T\mathbf{y})\mathrm{d}s\right), \tag{1.2}$$

where λ is called the wavelength, $\boldsymbol{\theta}$ is the unit vector that specifies the direction of propagation of the wave, and $\mathbf{P}_{\boldsymbol{\theta}^\perp}^T \in \mathbb{R}^{3\times 2}$ is the transpose of the matrix $\mathbf{P}_{\boldsymbol{\theta}^\perp} \in \mathbb{R}^{2\times 3}$ with $\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}$ the orthogonal projection of $\mathbf{x}$ onto the plane perpendicular to $\boldsymbol{\theta}$. Equation (1.2) describes the phase shift and attenuation introduced by the specimen, as shown in Figure 1.1.

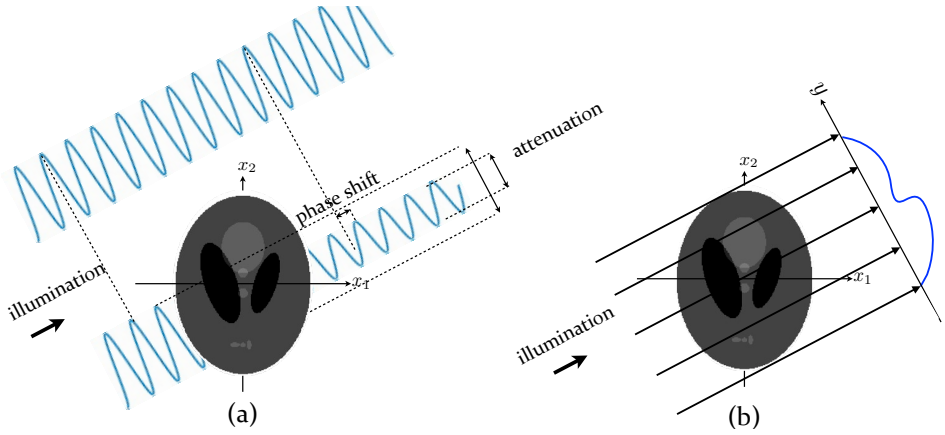


Figure 1.1: (a) The specimen introduces a phase shift and an attenuation. (b) Illustration of the intensity received by the detector.

In conventional CT, the decay of the wave intensity (the attenuation) is inferred for several orientations. This is linked to the X-ray transform of the attenuation coefficient $\mu(\mathbf{x}) = 4\pi\beta(\mathbf{x})/\lambda$ [2]. Since the mathematical model of CT is based on the x-ray transform, the specimen can be reconstructed using the analytical solution, particularly filtered back-projection-type algorithm (FBP). The drawback of FBP is the requirement of a large number of orientations with high signal-to-noise projection measurements for good-quality reconstruction which is equivalent to long exposure time. In order to reduce the radiation dose, several sophisticated algorithms have been developed for CT, including iterative coordinate descent (ICD) methods [3], block-based coordinate descent [4], ordered subset algorithms based on separable quadratic surrogates [5], preconditioned nonlinear conjugate-gradient methods [6], and alternating direction method of multipliers [7]. Precise modeling of the acquisition process in these methods offers a gain in quality with respect to FBP for the same data and degrades more gracefully than FBP when the data worsen; equivalently, it results in a notable radiation dose reduction. Moreover, specialized hardwares (CT vendors) have also been manufactured to implement these techniques. GE Healthcare started with ASIR (adaptive statistical iterative reconstruction) in 2008; the Siemens company provided SAFIRE (sinogram affirmed iterative reconstruction) in 2010;

the Toshiba company proposed AIDR (adaptive iterative dose reduction) in 2010.

The phase shift of the transmitted wave is given by the x-ray transform of the local phase shift per length, $\phi(\mathbf{x}) = 2\pi\delta(\mathbf{x})/\lambda$. Note that, in practice, intensity is the only measurable quantity. Therefore, it is necessary to find a mechanism that transforms the phase into an intensity. This fact motivates the development of various phase-contrast imaging modalities (PCI) including analyzer based [8–10], interferometric [11–13], and free-space propagation methods [14–16]. These methods differ substantially in terms of the physical signal that is measured and their experimental setup. They often show higher contrast over the conventional imaging of biological samples and soft tissues [17–23]. The iterative reconstruction scheme in PCI has not been developed as much as the conventional CT.

In this thesis, we develop a unified and high-performance reconstruction scheme for straight-ray tomography. We achieve the same level of sophistication as the state-of-the-art iterative schemes in conventional CT and take profit of recent developments in the specialized area of straight-ray tomography, but at a much lower computational cost. After successive application of conventional CT for the visualization of the specimen with the resolution of lower than micro meter, several imaging modalities have been developed from micro to nano resolution. In these modalities, the state-of-the-art reconstruction until very recently has been using direct methods. We demonstrate the proposed framework in the context of grating-based x-ray imaging for micro resolution and single-particle analysis for near-atomic-resolution imaging.

1.1 X-ray Grating Interferometry: Potentially *in vivo* Imaging Modality

Phase-sensitive x-ray imaging using grating interferometry (GI) is a tomographic technique that was first proposed by David et al. [24] and Momose et al. [19]. A unique property of GI is to provide simultaneously three complementary information about the object of interest: 1) The absorption contrast, 2) the differential phase contrast, and 3) the small scattering angle which is called dark field or visibility-reduction contrast as demonstrated in Figure 1.2. Additional advantages are its compatibility with regular laboratory sources of x-rays and its high sensitivity to variations in the density of electrons, which offers further opportunities to probe the specimen.

The data provided by differential phase-contrast imaging (DPCI) corresponds to the

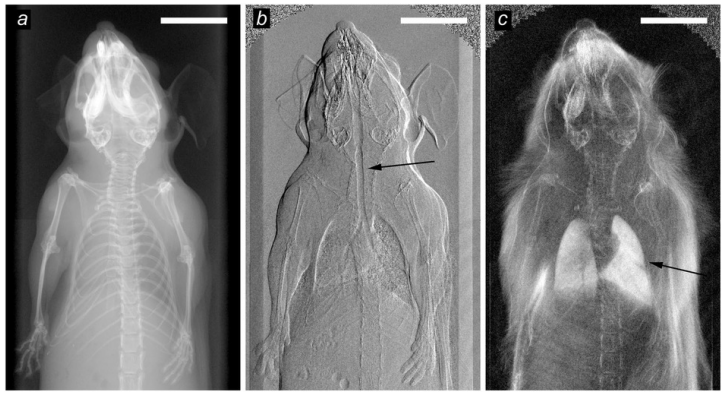


Figure 1.2: [25] (a) Conventional x-ray image based on attenuation. (b) Differential phase-contrast image based on x-ray refraction. (c) Dark-field image based on x-ray scattering. All three images are intrinsically perfectly registered as they are extracted from the same data recorded with a grating interferometer. Examples of regions of enhanced contrast are marked with arrows, showing (b) the refraction of the trachea and (c) the scattering of the lungs. The white bars correspond to 1 cm.

first derivative of the x-ray transform of the real part of the refractive index of the sample. Thus, in practical applications, the common reconstruction scheme for DPCI is based on a variant of the filtered back-projection (FBP) algorithm. While FBP is a fast (non-iterative) method, it typically requires a large number of projections with high signal-to-noise ratio to achieve a good reconstruction quality [26]. This implies long exposure times which could damage the specimen. High doses of x-ray radiation can lead to an increased risk of developing cancer and may cause the genetic deffects [27–32]. In order to be able to use this technique for *in vivo* imaging, one requires reducing the radiation dose significantly.

Recently, several authors have proposed iterative techniques that exploit prior knowledge on the specimen to significantly reduce the number of required projections [33–36] at no cost in the quality. Their approaches are all based on a penalized maximum-likelihood formulation, with a standard ℓ_2-norm data-fidelity term. In this thesis, we aim at further reducing the number of projections by proposing an improved iterative reconstruction algorithm for DPCI.

1.2 Single Particle Analysis: A Step Towards *de novo* Generation of Atomic Models

The purpose of single-particle analysis is to combine images of similar particles, typically proteins or viruses, often acquired from transmission electron microscopy. The proper combination then provides high-resolution 3D reconstruction of the sample. The first use of this technique dates back to the reconstruction of human wart virus and bushy stunt virus in the 1970s [37, 38]. This technique has progressed significantly since then, currently reaching to 0.5 nanometer resolution [39]. This motivates researchers to improve its resolution even further to be able to use this technique for *de novo* generation of an atomic model. The resolution improvement can be achieved by developing precise, fast and robust particle determination scheme besides high-end microscopes and careful data acquisition.

Since the linear image-formation model of the cryo-electron microscopy (Cryo-EM) is based on the x-ray transform, one needs to reform an inversion equation for reconstructing the particle. Given the size of the data and the fact that the angular distribution of the views is uneven, this is already a non-trivial reconstruction problem. In single-particle analysis, the problem is further complicated owing to the lack of information about the orientations besides the structural variability in the particle. Moreover, the Cryo-EM images are extremely noisy because of the low electron exposure to limit the radiation damage. Then,

the reconstruction problem suffers from overfitting, which means that the reconstruction express noise instead of the underlying particle details. In order to deal with this issue, It is convenient to introduce some prior information and formulate the problem as regularized inverse problem. In the literature, the smoothness concept is widely used as prior information in SPA [40–42]. This is performed by imposing a Gaussian distribution on the Fourier components of the particle in the context of maximum a posteriori estimation or using Tikhonov regularization in penalized likelihood estimation.

1.3 Main Contributions

The main focus of this thesis is on the development of a high-performance reconstruction framework for straight-ray computerized tomography. We then demonstrate the proposed framework in 1) grating-based x-ray imaging in order to reduce the radiation dose to increase its potential for being used in *in vivo* imaging, and in 2) single-particle analysis to improve the resolution such that one can potentially use it for *de novo* generation of atomic model. The five contributions of this thesis can be summarized as follows:

- **Discretization Scheme:** In order to formulate the reconstruction problem as an inverse problem, one needs first to discretize the forward operator. We specify the reconstruction space through the choice of the generating function. In this regard, we investigate an extended family of box splines and Kaiser-Bessel window functions. We first provide a general characterization of the x-ray transform for the extended family of box splines. We find that this level of generality simplifies the analysis because the family happens to be closed under the Radon/x-ray transform. Since all commonly used brands of B-splines are special instances of box splines [43], it makes sense to investigate these functions in more detail. Then, we consider the family of Kaiser-Bessel window functions. These are isotropic, and they involve several parameters that need to be adjusted [44,45]. We investigate approximation-theoretic properties of these basis functions and we show how to optimize the parameters for the best performance.

- **Reconstruction Framework:** We have designed new iterative reconstruction algorithms that take advantage of the proposed discretization and use a combination of non-quadratic regularizes. The regularization consists either of total variation (TV) for piecewise-constant images or a higher-order extension that is better matched to biological specimens and quadratic regularizations. Our method follows

an augmented-Lagrangian optimization principle and makes use of the conjugate-gradient method to solve the linear step in the alternating direction method of multipliers (ADMM). We propose a problem-specific preconditioner that considerably speeds up the convergence of the linear optimization step. Moreover, we impose support and positivity constraints in the reconstruction algorithm.

- **FFT-Cost Implementation of $\mathbf{H}^T\mathbf{H}$:** The computationally costly step in the proposed reconstruction scheme is the calculation of $\mathbf{H}^T\mathbf{H}$. We show theoretically that $\mathbf{H}^T\mathbf{H}$ is a digital convolution operator if the generating function satisfies the radial Nyquist criteria. We then show that, if we use B-spline functions or Kaiser-Bessel windows as generating function, then the proposed ADMM scheme converges to the fixed point of the problem. It improves the speed of reconstruction scheme significantly. We show that the use of the proposed digital convolution instead of the direct implementation of $\mathbf{H}^T\mathbf{H}$ makes the computational cost independent upon the number of orientations and the support of the generating function.

- **Grating-Based X-Ray Imaging (collaboration with the Paul Scherrer Institute (PSI)):** Up to now, in-vivo tomography with grating interferometry faces the challenge of large-dose deposition, which potentially harms the specimens (*e.g.*, in small rodent scanners). Grating-based x-ray imaging is a powerful modality to investigate biological samples. It measures three complementary characteristics of the imaged sample: the conventional absorption contrast (AC), the differential phase contrast (DPC), and a small-angle scattering contrast. To reduce the total scanning time, we apply the proposed reconstruction framework to the context of differential phase-contrast imaging. We present experimental results to validate the proposed discretization method and the corresponding iterative technique. Our findings confirm that the proposed reconstruction framework is quite competitive for solving regularized problems. Moreover, our method allows for a substantial dose reduction while preserving the image quality of FBP-type methods. This is a crucial step towards the diffusion of DPCI in medicine and biology. The codes have been implemented in the TOMCAT laboratory of PSI and are being used by scientists to visualize the internal structure of their samples.

  Unlike DPC tomography, where the phase information can be recovered effectively by a reconstruction algorithm, the retrieval of phase images from DPC projections remains challenging and reduces the advantages of the phase information in radiographic applications. We utilize the same algorithm proposed in the first part of

the thesis and deploy a novel discretization approach using B-spline calculus to establish the differential operator. The algorithm is evaluated with breast biopsy and mastectomy samples. Although it is predicted theoretically that the phase image can give higher contrast for breast tissue, this has not yet been demonstrated in a clinical environment. The present study constitutes the first practical demonstration that DPC is capable of providing a higher contrast in clinical by relevant features like spiculation. These results could help to improve the diagnosis of breast cancer.

- **Single Particle Analysis (collaboration with Centro Nacional de Biotecnologia, Spain (CSIC)):** Several critical difficulties arise in the context of single-particle analysis (SPA). They can be summarized as follows: 1) hundreds (thousands) of low-signal-to-noise-ratio micrographs with unknown orientations (too many projection images which are highly noisy), and 2) space-dependent contrast transfer function (CTF). More precisely, the measurements (micrographs) are the x-ray transform of specimens (macromolecules), filtered by the point-spread function of the microscope and varying from one projection image to the next. Consequently, the determination of the structure of macromolecule involves deconvolving of each projection image along with the reconstruction of the specimen. We apply the proposed reconstruction framework to the context of SPA. We formulate the problem as a constrained regularized reconstruction. We show that we can directly include the CTF in the system matrix $\mathbf{H}^T\mathbf{H}$ without any extra computational cost. The experimental results suggest that our approach improves significantly the resolution of the reconstruction. It is an important step towards the application of SPA for *de novo* generation of an atomic model. The corresponding codes have been already implemented in Xmipp [46, 47].

1.4 Thesis Outline

This thesis involves two main parts. The first part describes the development of a high-performance reconstruction framework for straight-ray computerized tomography (Chapters 2, 3, 4, 5). The linear mathematical model of straight-ray imaging is based on the Radon transform, the x-ray transform, and its differential. We discuss the mathematical properties of these operators in Chapter 2. We describe the discretization scheme of the imaging operator and discuss the properties that should be satisfied by the generating function to be used for discretizing the forward model in Chapter 3 as well as a fast and accurate implementation. In particular, we investigate the use of box splines and Kaiser-Bessel win-

dow functions (KBWF). We study the KBWFs from an approximation-theory point of view and propose new parameters for their use in the discretization scheme. In Chapter 4, we describe several reconstruction schemes that use the alternating-direction method of multipliers (ADMM) to solve constrained and regularized reconstruction problems. The costiest step in the proposed iterative methods is the computation of $\mathbf{H}^T\mathbf{H}$. We show that, under certain conditions, we can implement it as a digital filter with the help of the FFT.

The second part is the application of the proposed framework to the context of x-ray grating interferometry (Chapter 6, 7, 8) and single-particle analysis (Chapter 9). In Chapter 6, we briefly review the physical model of x-ray grating interferometry. The end-result is that differential phase-contrast imaging can be described mathematically in terms of derivative variants of the x-ray transform. We then use real data from the TOMCAT beam line to validate the proposed framework. We improve the performance of the reconstruction framework in the context of grating based imaging in Chapter 7 by developing new reconstruction schemes. Beside the tomography problem, we develop phase and absorption retrieval in the context of grating-based radiography which is important clinically in Chapter 8. The second important application is widely investigated in Chapter 9. We finally conclude this thesis in Chapter 10.

Chapter 2

Mathematical Preliminaries

[1] Let an object be characterized by its complex refractive index $n(\mathbf{x}) = 1 - \alpha(\mathbf{x}) + \mathrm{j}\beta(\mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^3$ specifies the object coordinate. The measurements in straight-ray imaging modalities are related to the x-ray transform and its differential variants, such as

- Conventional x-ray CT
$$g(\mathbf{y}, \boldsymbol{\theta}) = \mathscr{P}\{\beta\}(\mathbf{y}, \boldsymbol{\theta}). \tag{2.1}$$
- Propagation-based phase-contrast CT [48]
$$g(\mathbf{y}, \boldsymbol{\theta}) = \triangle_{\mathbf{y}} \mathscr{P}\{\alpha\}(\mathbf{y}, \boldsymbol{\theta}). \tag{2.2}$$
where $\triangle_{\mathbf{y}} = \partial^2/\partial y_1^2 + \partial^2/\partial y_2^2$ is the Laplacian operator with $\mathbf{y} = (y_1, y_2)$.
- Differential phase-contrast CT [13, 49, 50]
$$g(\mathbf{y}, \boldsymbol{\theta}) = \langle \mathbf{u}, \nabla_{\mathbf{y}} \mathscr{P}\{\alpha\}(\mathbf{y}, \boldsymbol{\theta}) \rangle_{\mathbf{y}}, \tag{2.3}$$
where $\mathbf{u}$ is a unit vector in a projection coordinate $\mathbf{y}$, $\nabla_{\mathbf{y}}$ is the gradient operator with respect to $\mathbf{y}$ and $\langle \cdot, \cdot \rangle_{\mathbf{y}}$ is the corresponding inner product,

There $\mathscr{P}\{f\}(\mathbf{y}, \boldsymbol{\theta})$ denotes the x-ray transform of a function f along a given orientation $\boldsymbol{\theta}$ with $\mathbf{y} \in \mathbb{R}^2$ is the projection coordinates. In this regard, we establish some higher-level mathematical properties of the x-ray transform and its differential variants in this chapter.

[1] A part of this chapter has been presented in [36]

2.1 X-ray transform

In order to specify the x-ray transform and its differential variants, we need to set the geometry of the problem. The spatial coordinates of the input function are denoted by $\mathbf{x} = [x_1\, x_2\, \ldots\, x_d]^T$ and the hyperplane projection coordinates are $\mathbf{y} = [y_1\, \ldots\, y_{d-1}]^T$. The unit vector $\boldsymbol{\theta} \in S^{d-1}$ (S^{d-1} is the unit sphere in $\mathbb{R}^d$) points along the direction of integration. The projection matrix $\mathbf{P}_{\boldsymbol{\theta}^\perp} \in \mathbb{R}^{(d-1)\times d}$ is constructed such that its rows specify the normal basis of the hyperplane perpendicular to the direction of integration $\boldsymbol{\theta}$. Then, a point $\mathbf{x}$ can be expressed in the rotated coordinate $[\boldsymbol{\theta}, \mathbf{P}^T_{\boldsymbol{\theta}^\perp}]$ as

$$\mathbf{x} = t\boldsymbol{\theta} + \mathbf{P}^T_{\boldsymbol{\theta}^\perp}\mathbf{y}, \tag{2.4}$$

where $\mathbf{P}^T_{\boldsymbol{\theta}^\perp}$ is the transpose of the matrix $\mathbf{P}_{\boldsymbol{\theta}^\perp}$.

2.1.1 Problem geometry

We now explicitly describe the geometry in the case of 2-D and 3-D input functions.

We start with two-dimensional functions. The unit vector $\boldsymbol{\theta} = (-\sin\theta, \cos\theta)$ lies along the line of integration as depicted in Figure 2.1. The spatial coordinates of the input function are denoted by $\boldsymbol{x} = (x_1, x_2)$. They are also expressed in a rotated coordinate system as $\boldsymbol{x} = t\boldsymbol{\theta} + y\boldsymbol{\theta}^\perp$, where $t \in \mathbb{R}$ and $\boldsymbol{\theta}^\perp = (\cos\theta, \sin\theta)$ is the unit vector orthogonal to the integral line ($\boldsymbol{\theta}^\perp$ specifies the direction of projection).

In the case of three-dimensional problem, the integral orientation can be determined by knowing two Euler angles φ and θ as depicted in Figures 2.2(a) and 2.2(b) and is given by

$$\boldsymbol{\theta} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\,, \tag{2.5}$$

In order to specify the projection matrix $\mathbf{P}_{\boldsymbol{\theta}^\perp}$, it is necessary to determine the unit vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ in the input function coordinates. The unit vector $\mathbf{y}_2$ is the projection of the unit vector $\mathbf{x}_3 = (0,0,1)$ along direction $\boldsymbol{\theta}$ and is computed as

$$\mathbf{y}_2 = \mathbf{x_3} - <\mathbf{x_3}, \boldsymbol{\theta}>\boldsymbol{\theta}\,. \tag{2.6}$$

This yields

$$\mathbf{y}_2 = \frac{1}{|\sin\theta|}\begin{bmatrix} -\cos\varphi\cos\theta\sin\theta \\ -\sin\varphi\cos\theta\sin\theta \\ \sin^2\theta \end{bmatrix}. \tag{2.7}$$

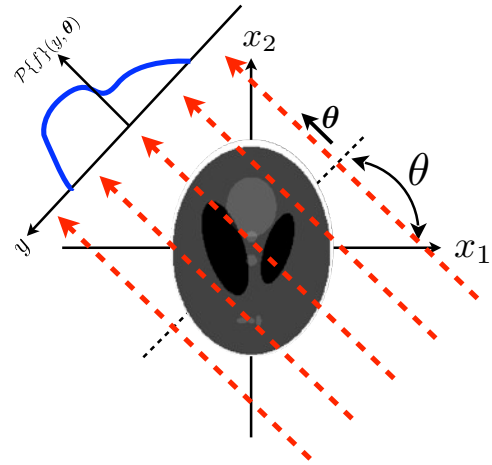


Figure 2.1: The object lies in a 2-D plane and is imaged along angle θ.

The unit vector $\mathbf{y}_1$ is the cross product of $\mathbf{y}_2$ and $\boldsymbol{\theta}$

$$\mathbf{y}_1 = \mathbf{y}_2 \times \boldsymbol{\theta} \tag{2.8}$$

which is in the form of

$$\mathbf{y}_1 = \frac{1}{|\sin\theta|} \begin{bmatrix} \sin\varphi \sin\theta \\ -\cos\varphi \sin\theta \\ 0 \end{bmatrix} \tag{2.9}$$

Then the projection matrix is

$$\mathbf{P}_{\boldsymbol{\theta}^\perp} = [\mathbf{y}_1, \mathbf{y}_2]^{\mathrm{T}}. \tag{2.10}$$

2.1.2 Definition and properties

The x-ray transform is the continuous-domain operator that maps a d-dimensional function into its line integrals; $\mathscr{P} : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^{d-1} \times \mathrm{S}^{\mathrm{d}-1})$ where $\mathrm{S}^{\mathrm{d}-1}$ is the unit sphere in $\mathbb{R}^d$. More specifically,

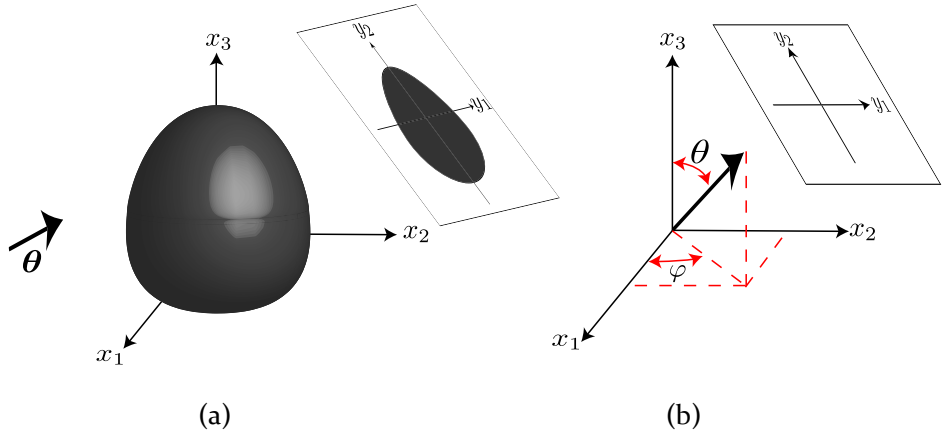


Figure 2.2: The object lies in a 2-D plane and is imaged along angle θ.

$$\mathscr{P}\{f\}(\mathbf{y},\theta) = \int_{\mathbb{R}} f(t\theta + \mathbf{P}_{\theta^\perp}^T \mathbf{y})\mathrm{d}t, \tag{2.11}$$

Note that when $f \in L_2(\mathbb{R}^2)$, the x-ray transform is equivalent to the Radon transform. In some cases, we use the notation $\mathscr{P}_\theta\{f\}(\mathbf{y})$ in place of $\mathscr{P}\{f\}(\mathbf{y},\theta)$.

Theorem 2.1. ***Fourier slice theorem in the context of x-ray transform*** *For $f \in L_2(\mathbb{R}^d)$, we have*

$$\mathscr{F}_{\mathbf{x}}\{f(\mathbf{x})\}(\mathbf{P}_{\theta^\perp}^T \boldsymbol{\omega}) = \mathscr{F}_{\mathbf{y}}\{\mathscr{P}\{f(\mathbf{x})\}(\mathbf{y},\boldsymbol{\theta})\}(\boldsymbol{\omega}), \tag{2.12}$$

where $\boldsymbol{\omega} \in \mathbb{R}^{d-1}$.

The Fourier slice theorem states that the $(d-1)$-dimensional Fourier transform of the x-ray transform of a function f is equal to the d-dimensional Fourier transform of the given function on the hyperplane parallel to the projection coordinate.

One can define the x-ray transform with the use of the Dirac impulse within the framework of distribution theory

$$\mathscr{P}\{f\}(\mathbf{y};\theta) = \langle f(\mathbf{x}), \delta(\mathbf{P}_{\theta^\perp}\mathbf{x}-\mathbf{y})\rangle_{\mathbf{x}\in\mathbb{R}^d}\,, \tag{2.13}$$

where $\delta(\mathbf{P}_{\theta^\perp}\mathbf{x}-\mathbf{y})$ is a line distribution that specifies a line that passes through the point $\mathbf{y}$ in the projection coordinate and is parallel to the orientation $\boldsymbol{\theta}$. Initially one assumes that f and its x-ray transform are in the Schwartz space, of smooth and rapidly decaying functions. One can extend the transformation to $L_2(\mathbb{R}^d)$ by applying a standard density argument. The main properties of the x-ray transform are as follows:

- **Linear**:

  It is a linear map by definition since it is an integral operator. Then, the x-ray transform adjoint (back-projection) is well-defined $\mathscr{P}^* : L_2(\mathbb{R}^{d-1}\times \mathrm{S}^{\mathrm{d}-1}) \to \mathrm{L}_2(\mathbb{R}^{\mathrm{d}})$ and is in the form of

  $$\mathscr{P}^*\{g(\mathbf{y},\boldsymbol{\theta})\}(\mathbf{x}) = \int_{\mathrm{S}^{\mathrm{d}-1}} g(\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x},\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}\,. \tag{2.14}$$

- **Scale invariance**

  For any $\alpha > 0$,

  $$\mathscr{P}\{f(\alpha\mathbf{x})\}(\mathbf{y},\boldsymbol{\theta}) = \alpha^{-1}\mathscr{P}\{f\}(\alpha\mathbf{y},\boldsymbol{\theta})\,. \tag{2.15}$$

  The proof is achieved with the use of the x-ray transform definition,

  $$\begin{aligned}\mathscr{P}\{f(\alpha\mathbf{x})\}(\mathbf{y},\boldsymbol{\theta}) &= \langle f(\alpha\mathbf{x}), \delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}-\mathbf{y})\rangle_{\mathbf{x}\in\mathbb{R}^d} \\ &\overset{(1)}{=} \left\langle f(\overline{\mathbf{x}}), \alpha^{-d}\delta(\alpha^{-1}\mathbf{P}_{\boldsymbol{\theta}^\perp}\overline{\mathbf{x}}-\mathbf{y})\right\rangle_{\overline{\mathbf{x}}\in\mathbb{R}^d} \\ &\overset{(2)}{=} \left\langle f(\overline{\mathbf{x}}), \alpha^{-1}\delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\overline{\mathbf{x}}-\alpha\mathbf{y})\right\rangle_{\overline{\mathbf{x}}\in\mathbb{R}^d} \\ &\overset{(3)}{=} \alpha^{-1}\left\langle f(\overline{\mathbf{x}}), \delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\overline{\mathbf{x}}-\alpha\mathbf{y})\right\rangle_{\overline{\mathbf{x}}\in\mathbb{R}^d}\,.\end{aligned} \tag{2.16}$$

  The change of variable $\overline{\mathbf{x}} = \alpha\mathbf{x}$ and consequently $\mathrm{d}\overline{\mathbf{x}} = |\alpha|^d\mathrm{d}\mathbf{x}$ yield the equality (1). The scaling property of the Dirac impulse and the linearity of the inner product with respect to the second term implies the other two equalities, respectively.

- **Projection translation invariance**

  $$\mathscr{P}\{f(\cdot-\mathbf{x}_0)\}(\mathbf{y},\boldsymbol{\theta}) = \mathscr{P}\{f\}(\mathbf{y}-\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}_0,\boldsymbol{\theta})\,. \tag{2.17}$$

Starting form the definition, we have

$$\begin{aligned}\mathscr{P}\{f(\mathbf{x}-\mathbf{x}_0)\}(\mathbf{y},\boldsymbol{\theta}) &= \langle f(\mathbf{x}-\mathbf{x}_0), \delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}-\mathbf{y})\rangle_{\mathbf{x}} \\ &\overset{(1)}{=} \langle f(\mathbf{x}), \delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}+\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}_0-\mathbf{y})\rangle_{\mathbf{x}} \\ &= \langle f(\mathbf{x}), \delta(\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}-(\mathbf{y}-\mathbf{P}_{\boldsymbol{\theta}^\perp}\mathbf{x}_0))\rangle_{\mathbf{x}}\,. \end{aligned} \tag{2.18}$$

The adjoint of a shift operator is exactly the same shift with opposite direction which implies equality (1).

- **Pseudo-distributivity with respect to convolution**

$$\mathscr{P}\{f*g\}(\mathbf{y},\boldsymbol{\theta}) = (\mathscr{P}f(\cdot,\boldsymbol{\theta}) * \mathscr{P}g(\cdot,\boldsymbol{\theta}))(\mathbf{y},\boldsymbol{\theta})\,. \tag{2.19}$$

2.2 Differential variants of the x-ray transform

Since in the parallel-beam geometry, the 3-D problem can be decomposed into a set of two dimensional problems, we focus on two dimensional function $f \in L_2(\mathbb{R}^2)$. The nth derivative of the x-ray transform of a function $f(\boldsymbol{x})$ is denoted by

$$\mathscr{P}^{(n)}f(y,\boldsymbol{\theta}) = \frac{\partial^n}{\partial y^n}\mathscr{P}f(y,\boldsymbol{\theta})\,. \tag{2.20}$$

The derivatives of the x-ray transform are linear operators with the following properties:

- Scale invariance

$$\mathscr{P}^{(n)}\{f(\alpha\mathbf{x})\}(y,\boldsymbol{\theta}) = \alpha^{n+1}\mathscr{P}^{(n)}f(\alpha y,\boldsymbol{\theta}), \quad \alpha \in \mathbb{R}^+\,. \tag{2.21}$$

- Pseudo-distributivity with respect to convolution

$$\mathscr{P}^{(n)}\{f*g\}(y,\boldsymbol{\theta}) = (\mathscr{P}^{(n)}f(\cdot,\boldsymbol{\theta}) * \mathscr{P}g(\cdot,\boldsymbol{\theta}))(y,\boldsymbol{\theta}) = (\mathscr{P}f(\cdot,\boldsymbol{\theta}) * \mathscr{P}^{(n)}g(\cdot,\boldsymbol{\theta}))(y,\boldsymbol{\theta})\,. \tag{2.22}$$

- Projected translation invariance

$$\mathscr{P}^{(n)}\{f(\cdot-\boldsymbol{x}_0)\}(y,\boldsymbol{\theta}) = \mathscr{P}^{(n)}f(y-\langle\boldsymbol{x}_0,\boldsymbol{\theta}\rangle,\boldsymbol{\theta})\,. \tag{2.23}$$

2.3 Direct inversion formula

To derive the necessary relations, we define a new operator, the Hilbert transform along the second coordinate.

Definition 2.1. *The Hilbert transform along the x_2 axis, $\mathscr{H}_2 : L_2(\mathbb{R}^2) \longrightarrow L_2(\mathbb{R}^2)$, is defined in the Fourier domain as*

$$\widehat{\mathscr{H}_2\{f\}}(\omega_1,\omega_2) = -j \cdot sgn(\omega_2)\widehat{f}(\omega_1,\omega_2)\,, \tag{2.24}$$

where (ω_1,ω_2) are spatial frequency coordinates.

Proposition 2.1. *The sequential application of the x-ray transform, the nth derivative operator and the adjoint of the x-ray transform on function $f \in L_2(\mathbb{R}^2)$ is*

$$\mathscr{P}^*\{\frac{\partial^n}{\partial y^n}\mathscr{P}\{f\}(y,\theta)\}(\boldsymbol{x}) = 2\pi(-1)^n\mathscr{H}_2^n(-\triangle)^{\frac{n-1}{2}}\{f\}(\boldsymbol{x})\,, \tag{2.25}$$

where $(-\triangle)^{\frac{1}{2}}$ is the fractional Laplace operator with transfer function $\|\boldsymbol{\omega}\|$, $(-\triangle)^{\frac{n}{2}}$ is n times application of this operator, and the adjoint of x-ray transform (2.14) *can be restated in two dimensional case in the form of*

$$\mathscr{P}^*\{g\}(\mathbf{x}) = \int_0^\pi g(x_1\cos\theta + x_2\sin\theta,\boldsymbol{\theta})\mathrm{d}\theta\,. \tag{2.26}$$

Proof. Let $g(y,\theta) = \mathscr{P}f(y,\theta)$. The Fourier Slice Theorem states that:

$$\widehat{g}(\omega,\theta) = \widehat{f}(\omega\cos\theta,\omega\sin\theta)\,. \tag{2.27}$$

The Fourier transform of the nth derivative of $g(y,\theta)$ with respect to y is $(i\omega)^n\widehat{g}(\omega,\theta)$. Thus,

$$(\mathrm{j}\omega)^n\widehat{g}(\omega,\theta) = \mathrm{j}^n \times \mathrm{sgn}^n(\omega)|\omega|^n\widehat{f}(\omega\cos\theta,\omega\sin\theta)\,. \tag{2.28}$$

where $\omega = \|\boldsymbol{\omega}\|$ with $\boldsymbol{\omega} = (\omega_1,\omega_2) = \omega(\cos\theta,\sin\theta)$. For $\theta \in (0,\pi)$, $\mathrm{sgn}(\omega) = \mathrm{sgn}(\omega\sin\theta) = \mathrm{sgn}(\omega_2)$. The space-domain equivalent is

$$\frac{\partial^n}{\partial y^n}\mathscr{P}f(y,\theta) = \mathscr{P}\{((-1)^n(\mathscr{H}_2)^n(-\triangle)^{\frac{n}{2}}\{f\})(\boldsymbol{x})\}(y,\theta)\,,\forall\theta\in(0,\pi)\,. \tag{2.29}$$

Therefore we have

$$\mathscr{P}^*\{\frac{\partial^n}{\partial y^n}\mathscr{P}\{f\}(y,\theta)\}(\boldsymbol{x}) = \mathscr{P}^*\mathscr{P}\{((-1)^n(\mathscr{H}_2)^n(-\triangle)^{\frac{n}{2}}\{f\})(\boldsymbol{x})\}(y,\theta), \forall\theta\in(0,\pi) \tag{2.30}$$

which, owing to the property that $\mathscr{P}^*\mathscr{P} = (-\triangle)^{1/2}$, yields the desired results. □

Equation(2.25) is a key equation. It implies that if one is interested in solving the inverse problem

$$\mathscr{P}^{(n)}f(y,\theta) = g(y,\theta), \tag{2.31}$$

the direct solution is to first apply the x-ray adjoint on the measured data and then apply the inverse of the operator $2\pi(-1)^n\mathscr{H}_2^n(-\triangle)^{(n-1)/2}$. The transfer function of this inverse is $\frac{1}{2\pi i^n}\mathrm{sgn}(\omega_2)^n\|\boldsymbol{\omega}\|^{-(n-1)}$.

An equivalent form of (2.25) using the fact that $(\frac{\partial}{\partial y})^* = -\frac{\partial}{\partial y}$ is

$$\mathscr{P}^{(n)^*}\{(q * \mathscr{P}^{(n)}f(\cdot,\theta))(y)\}(\boldsymbol{x}) = f(\boldsymbol{x}), \tag{2.32}$$

where $\mathscr{P}^{(n)}$ is the adjoint of the n-th derivative of the x-ray transform and the transfer function of $q(y)$ is

$$\widehat{q}(\omega_y) = \frac{1}{2\pi}\times\frac{1}{|\omega_y|^{2n-1}}. \tag{2.33}$$

Equation (2.32) is the basis for the generalized filtered back projection (GFBP). The full procedure is described in Algorithm 1.

Input: $g_\theta(y) = \mathscr{P}_\theta^{(n)} f(y)$ as data
Output: Reconstructed image $f(\mathbf{x})$
initialization $N_\theta =$ The number of angles;
for $i = 1 \to N_\theta$ **do**
    Filter the input data with the transfer function

$$\widehat{q}(\omega_y) = \frac{1}{2\pi} \frac{1}{|\omega_y|^{2n-1}} \tag{2.34}$$

end
Apply the adjoint of the n-th derivative of the x-ray transform on the output of the previous stage;
return $f(\mathbf{x})$.

Algorithm 1: GENERALIZED FILTERED BACK PROJECTION(GFBP) FOR THE INVERSE PROBLEM $\mathscr{P}_\theta^{(n)} f(y) = g_\theta(y)$

Chapter 3

Discretization Scheme

[1]

Since the mathematical model of straight-ray imaging modalities are based on the x-ray transform and its differential variants, the object can be reconstructed using direct methods such as filtered back-projection. These techniques require large number of projections with high signal-to-noise ratio in order to provide high quality reconstruction. It is equivalent to a long exposure time and high radiation dose. It is highly desirable to reduce the radiation dose in different imaging modalities. This can be achieved by either a reduction in the intensity of the photons or in the number of projection angles. The price to pay for this reduction is that the reconstruction problem becomes more ill-posed and its solution can no longer be well approximated using traditional direct methods. Instead the deployment of more sophisticated iterative schemes is needed. In order to specify such methods, one first discretizes the imaging operator, and then selects a reconstruction scheme that typically involves the choice of a cost functional to minimize.

In this chapter we concentrate on the first aspect: we use generalized sampling framework to discretize the forward operator. Therefore, the discretization problem is summarized in the choice of a suitable reconstruction space. This space is usually determined as a set of functions of the form

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c[\mathbf{k}]\varphi\left(\frac{\mathbf{x}}{T} - \mathbf{k}\right), \tag{3.1}$$

[1] A part of this chapter has been presented in [51]

where T is the sampling step. The reconstruction space is then specified through the choice of the generating function φ.

In computed tomography, where the mathematical model is based on the x-ray transform and its variants, it is beneficial to use a generating function that has 1) short support and 2) good approximation properties. Isotropy is an additional property that simplifies the implementation since the footprint (x-ray transform of the function) is independent of the orientation, but typically introduces some loss of accuracy.

Two favorable candidates are box splines (in particular B-splines) and Kaiser-Bessel window functions. B-splines are compactly supported functions with the best cost/quality trade-off for the interpolation of discrete data on uniform grid. Their tensor product is used in order to extend them to higher dimension. Note that high degree B-splines are approximately isotropic. They are a special case of the box-splines investigated in this chapter. Kaiser-Bessel window functions (KBWFs) are widely used in electron microscopy [52–54] and conventional and differential phase-contrast x-ray computed tomography [34, 44, 55–57]. KBWFs involve three parameters that need to be adjusted [44, 45]. In this chapter, we investigate approximation-theoretic properties of the basis functions and we show how to optimize the parameters for the best performance. We also present experimental results that corroborate our theoretical prediction.

3.1 Discretization Using Shift-Invariant Functional Spaces

We first explain how the discretization of the forward model is intimately connected with the choice of a given basis function. We then recall some fundamental results from approximation theory that ensure stability and allow one to predict the expected discretization error. This will point to the importance of the partition-of-unity property which, unfortunately as we shall prove for compactly support functions, is incompatible with isotropy properties.

3.1.1 Matrix formulation

Reconstruction is usually formulated as a linear inverse problem. To solve it, it is convenient to introduce discrete representations of the object and the imaging operator. Here, we consider an object in two dimensions. The model of the object, from the perspective of the generalized sampling theory [58], is obtained by specifying a suitable reconstruction space. Specifically, we select $V_T(\varphi)$ as the principal shift-invariant space generated by the

function $\varphi \in L_2(\mathbb{R}^2)$. This space is defined by

$$V_T(\varphi) = \left\{ \sum_{\mathbf{k}\in\mathbb{Z}^2} c[\mathbf{k}]\varphi\left(\frac{\mathbf{x}}{T} - \mathbf{k}\right) : c \in \ell_2(\mathbb{Z}^2) \right\}, \tag{3.2}$$

where $\mathbf{x} \in \mathbb{R}^2$. The corresponding orthogonal projection operator $\mathbf{P}_T : L_2(\mathbb{R}^2) \to V_T(\varphi)$ is defined as

$$\mathbf{P}_T f = \underset{g\in V_T(\varphi)}{\operatorname{argmin}} \|f - g\|_{L_2}. \tag{3.3}$$

In practice, however, the values of c in (3.2) are determined based on the solution of an inverse problem.

As the derivative variants of x-ray transform are linear, pseudo shift-invariant operators, their application on a function $f \in V_T$ is

$$\begin{aligned} \mathscr{P}^{(n)}\{f\}(y,\theta) &= \sum_{\mathbf{k}\in\mathbb{Z}^2} c[\mathbf{k}]\mathscr{P}^{(n)}\{\varphi_T(\cdot - T\mathbf{k})\}(y,\theta) \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^2} c[\mathbf{k}]\mathscr{P}^{(n)}\{\varphi_T\}(y - T\langle \mathbf{k}, \boldsymbol{\theta}\rangle, \theta), \end{aligned} \tag{3.4}$$

where $\varphi_T(\mathbf{x}) = \varphi(\mathbf{x}/T)$, $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$, and

$$\mathscr{P}^{(n)} f(y,\theta) = \frac{\partial^n \mathscr{P} f}{\partial y^n}(y,\theta),$$

with $\mathscr{P} : L_2(\mathbb{R}^2) \to L_2(\mathbb{R}\times[0,\pi])$ being the x-ray transform operator.

The formulation of the reconstruction as a linear inverse problem is then restated as the matrix equation

$$\mathbf{g} = \mathbf{H}\mathbf{c}, \tag{3.5}$$

where $\mathbf{g}$ is the measurement vector, $\mathbf{H}$ is the system matrix, and $\mathbf{c}$ is the discrete representation of the object of interest. Using (3.4), the matrix formulation can be obtained as follows: The measurement vector $\mathbf{g}$ contains values of the imaging transform $\mathscr{P}^{(n)}\{f\}(y,\theta)$ at the sampled points $y_j = j\Delta y$ and $\theta_i = i\Delta\theta$, where $i, j \in \mathbb{Z}$. The object f is represented with its coefficients c within the space V_T. The system matrix $\mathbf{H}$ is given by

$$[\mathbf{H}]_{(i,j),\mathbf{k}} = \mathscr{P}^{(n)}\{\varphi_T\}(y_j - T\langle \mathbf{k}, \boldsymbol{\theta}_i\rangle, \theta_i). \tag{3.6}$$

Note that, in order to compute the imaging operator, there is no need to store the whole system matrix because it is sufficient to have access to a lookup table that contains the projection of one basis function along every direction. For an isotropic basis function, storing its footprint along one orientation is enough since its footprint is independent of the orientation.

3.1.2 Fast implementation

The calculation of the x-ray transform (or its variants) involves the application of the system matrix on the coefficients of the object. However the system matrix $\mathbf{H}$ cannot be stored explicitly due to its size. To circumvent this problem we exploit the translation-invariance of the x-ray transform. This property implies that all the matrix entries in (3.6) can be derived from a single derivative of the x-ray transform, namely that of the generating function φ:

$$\mathbf{H}_{(i,j),\mathbf{k}} = \mathscr{P}^{(n)}\varphi_T(y_j - T\langle \boldsymbol{k}, \boldsymbol{\theta}_i\rangle, \theta_i).$$

To improve the speed, we oversample $\mathscr{P}^{(n)}\varphi(y,\theta)$ on a fine grid $Y \times \Theta$ with for example 100 samples along each angular direction and store the values in a lookup table $\mathbf{L}$. To compute the matrix entries we define a mapping

$$\begin{aligned} I:\ \mathbb{R}\times[0,\pi] &\longrightarrow \{1,2,\cdots,K\}\times\{1,2,...,P\} \\ (y,\theta) &\longmapsto (j,i)\,, \end{aligned} \tag{3.7}$$

with K is the number of samples along each direction, P is the number of projections and $(Y(j),\Theta(i))$ is the sample in $Y\times\Theta$ that is nearest to (y,θ). Therefore, we have

$$[\mathbf{H}]_{(i,j),\boldsymbol{k}} = \mathbf{L}_{I(y_j - \langle \boldsymbol{k},\boldsymbol{\theta}_i\rangle,\theta_i)}\,. \tag{3.8}$$

In the case of isotropic basis functions, it is sufficient to store its footprint along one orientation in the look-up table since its x-ray transform is independent of the orientation. Note that the algorithm can easily be parallelized, since projections corresponding to different angles are completely independent of each other (Figure 3.1). We designed a multithreaded implementation for an 8-core workstation which allows for 8 simultaneous projection computations. Similarly, for the adjoint of the forward model, the computation can be parallelized with respect to each object point.

In summary, our implementation is based on an accurate continuous-to-discrete model. Moreover it is fast thanks to the use of look-up tables and multi-threading. Note that our

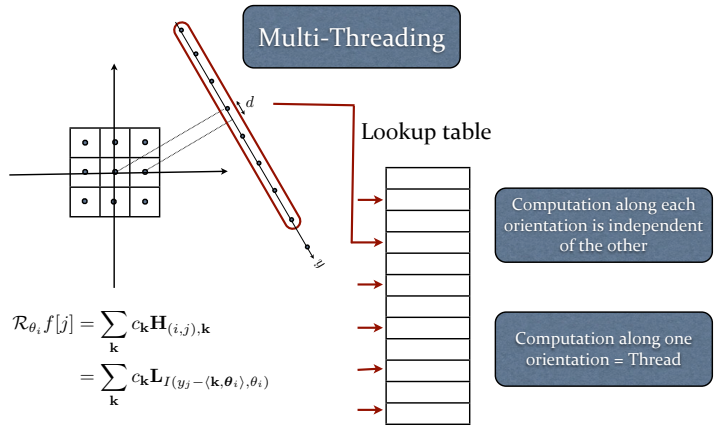


Figure 3.1: A simple demonstration of the implementation of the projection operator using lookup table in multithread scenario.

method could be also adapted to fan beam geometry by mapping it back to the parallel beam geometry. This would lead to a non-uniform sampling pattern but our method can account for this at no additional cost (thanks to our look-up-table-based implementation).

3.1.3 Desirable properties of the basis functions

We require the basis function φ to satisfy the four following properties:

1) **Riesz basis.** Every object $f \in V(\varphi)$ must be uniquely specified by its coefficients c. This requires the existence of a positive constant A such that

$$\forall c \in \ell_2, \quad A \cdot \|c\|_{\ell_2}^2 \leq \left\| \sum_{\mathbf{k} \in \mathbb{Z}^2} c[\mathbf{k}] \varphi \left(\frac{\mathbf{x}}{T} - \mathbf{k} \right) \right\|_{L_2} . \tag{3.9}$$

In addition, the representation should be stable. This requires the existence of a positive constant B such that

$$\forall c \in \ell_2, \quad \left\| \sum_{\mathbf{k} \in \mathbb{Z}^2} c[\mathbf{k}] \varphi \left(\frac{\mathbf{x}}{T} - \mathbf{k} \right) \right\|_{L_2} \leq B \cdot \|c\|_{\ell_2}^2 . \tag{3.10}$$

Together, these two conditions are equivalent to φ being a Riesz basis of $V(\varphi)$.

2) **Partition of unity.** It is constructive for such a discretization scheme that the model approximate any input function as closely as desired by choosing a sufficiently small sampling step. More precisely, the approximation error should vanish whenever the sampling step T tends to zero. We thus require that

$$\lim_{T\to 0}\left\{\|f-\mathbf{P}_T f\|_{L_2}\right\}=0\,. \tag{3.11}$$

Theorem 3.1 ([59]). *Let f be a continuously defined function. The L_2-approximation error of the operator $\mathbf{P}_T : L_2 \to V_T(\varphi)$ can be written as*

$$\begin{aligned}\varepsilon_f(T) &= \|f-\mathbf{P}_T\{f\}\|_{L_2}\\ &= \left(\int_{\mathbb{R}^2} E_\varphi(T\boldsymbol{\omega})|\widehat{f}(\boldsymbol{\omega})|^2\frac{\mathrm{d}\boldsymbol{\omega}}{2\pi}\right)^{1/2}+\varepsilon_{\text{corr}}\,,\end{aligned} \tag{3.12}$$

where ε_{corr} is a correction term and E_φ is the error kernel defined in the least-squares case as

$$E_\varphi(\boldsymbol{\omega})=1-\frac{|\widehat{\varphi}(\boldsymbol{\omega})|^2}{\sum_{\mathbf{k}\in\mathbb{Z}^2}|\widehat{\varphi}(\boldsymbol{\omega}+2\mathbf{k}\pi)|^2}\,, \tag{3.13}$$

where $\widehat{\varphi}$ is the Fourier transform of φ. Specifically, if $f\in W_2^r$ (Sobolev space of order r) with $r>1/2$, then $|\varepsilon_{corr}|<\gamma T^r\|f^{(r)}\|_{L_2}$, where γ is some constant.

The asymptotic convergence

$$\lim_{T\to 0}\varepsilon_f(T)=0 \tag{3.14}$$

is achieved if and only if the basis function φ satisfies the partition-of-unity condition [58]

$$\sum_{\mathbf{k}\in\mathbb{Z}^2}\varphi(\mathbf{x}+\mathbf{k})=1,\ \forall\mathbf{x}\in\mathbb{R}^2\,. \tag{3.15}$$

The equivalent formulation of the partition of unity in the frequency domain is

$$\widehat{\varphi}(2\pi\mathbf{n})=\delta[\mathbf{n}],\ \forall\mathbf{n}\in\mathbb{Z}^2\,, \tag{3.16}$$

where δ is the two-dimensional Kronecker delta function.

3) **Compact support.** The basis function φ should be compactly supported in order to reduce the computational cost and also for localization in the spatial domain.

4) **Isotropy.** For the implementation of the imaging operator, it is required to store the values of its application on the basis function along different directions. If the basis function is isotropic, its projections do not depend on the direction, which leads to simplicity and efficiency of implementation.

3.1.4 Revisiting optimality in the projection domain

We now bound the error of approximation incurred by $\mathscr{P}_T f = \mathscr{P}\{\mathbf{P}_T f\}$. It can be extended to any derivative of the x-ray transform through the Fourier-slice theorem since $||\mathscr{P}^{(n)} f||_{L_2} = ||\mathscr{P}\mathscr{F}^{-1}\{|\boldsymbol{\omega}|^n \widehat{f}(\boldsymbol{\omega})\}||_{L_2}$. To this end, we use the Sobolev norm $\|\cdot\|^2_{W_2^{1/2}}$ in the projection domain. If $g \in L_2(\mathbb{R}^2)$, then

$$\|g\|^2_{W_2^{1/2}} = \int_0^{2\pi} \int_0^{\infty} (1+\omega^2)^{\frac{1}{2}} |\widehat{g}(\omega,\theta)| \mathrm{d}\theta \mathrm{d}\omega, \tag{3.17}$$

where $\widehat{g}(\omega,\theta)$ is the polar form of the Fourier transform of g.

Theorem 3.2. *Let $\varepsilon_{\mathscr{P}f}(T) == \|\mathscr{P}f - \mathscr{P}_T f\|_{W_2^{1/2}}$ be the Sobolev approximation error of the operator $\mathscr{P}_T$. Then, there exist positive constants $r_1, R_1 > 0$ such that*

$$r_1 \varepsilon_f(T) \leq \varepsilon_{\mathscr{P}f}(T) \leq R_1 \varepsilon_f(T). \tag{3.18}$$

Lemma 3.1. *Let $\Omega \subset \mathbb{R}^2$ be a compact domain. Then, there exist positive constants r_2 and R_2 such that, for any $L_2(\mathbb{R}^2)$ function f that is supported on Ω, it holds that*

$$r_2 \|f\|_{L_2} \leq \|\mathscr{P}f\|_{W_2^{1/2}} \leq R_2 \|f\|_{L_2}. \tag{3.19}$$

Proof. The Fourier-slice theorem implies that

$$\widehat{\mathscr{P}f}(\omega,\theta) = \widehat{f}(\omega,\theta), \quad \forall \omega \in [0,+\infty), \theta \in [0,2\pi). \tag{3.20}$$

To show the left-hand-side inequality, we write that

$$\begin{aligned}
\|f\|_{L_2}^2 &= \left\|\widehat{f}\right\|_{L_2}^2 \\
&= \int_0^{2\pi}\int_0^{\infty} |\widehat{f}(\omega,\theta)|^2 |\omega| \mathrm{d}\omega \mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^{\infty} |\widehat{\mathscr{P}f}(\omega,\theta)|^2 |\omega| \mathrm{d}\omega \mathrm{d}\theta \\
&\leq \int_0^{2\pi}\int_0^{\infty} |\widehat{\mathscr{P}f}(\omega,\theta)|^2 (1+\omega^2)^{1/2} \mathrm{d}\omega \mathrm{d}\theta \\
&= \|\mathscr{P}f\|_{W_2^{1/2}}^2 \,. \qquad (3.21)
\end{aligned}$$

For the right-hand side, we decompose the integral into an integral over $|\omega| \geq 1$ and an integral over $|\omega| \leq 1$. In the first one, we have that $2|\omega| \geq (1+|\omega|^2)^{1/2}$. So,

$$\begin{aligned}
\int_0^{2\pi}\int_1^{\infty} |\widehat{\mathscr{P}f}(\omega,\theta)|^2 (1+\omega^2)^{1/2} \mathrm{d}\omega \mathrm{d}\theta &\leq 2\int_0^{2\pi}\int_1^{\infty} |\widehat{\mathscr{P}f}(\omega,\theta)|^2 |\omega| \mathrm{d}\omega \mathrm{d}\theta \\
&\leq 2\,\|f\|_{L_2}^2 \,. \qquad (3.22)
\end{aligned}$$

The integral over $|\omega| \leq 1$ is estimated using

$$\begin{aligned}
\int_0^{2\pi}\int_0^{1} |\widehat{\mathscr{P}f}(\omega,\theta)|^2 (1+\omega^2)^{1/2} \mathrm{d}\omega \mathrm{d}\theta &\leq \int_0^{2\pi}\int_0^{1} |\widehat{f}(\omega,\theta)|^2 (1+\omega^2)^{1/2} \mathrm{d}\omega \mathrm{d}\theta \\
&\leq \sup_{\theta\in[0,2\pi),\omega\in[0,1)} |\widehat{f}(\omega,\theta)|^2 \int_0^{2\pi}\int_0^{1} (1+\omega^2)^{1/2} \mathrm{d}\omega \mathrm{d}\theta \\
&\leq \tilde{C}\,\|f\|_{L_2}^2 \,. \qquad (3.23)
\end{aligned}$$

Details concerning the last inequality can be found in [60, Section II.5]. Together, these inequalities yield the desired result. □

Proof of Theorem 3.2. By letting $f \leftarrow (f - \mathbf{P}_T\{f\})$ in (3.19), we obtain (3.18). □

This theorem implies that the average error over all angles is small in the transform domain when the error of approximation is small in the object domain. While the theorem is an average result that involves a continuum of angles, it is still useful practically because it gives us the approximation error in the transform domain over a family of images that would correspond to all rotated versions of a given reference image.

3.1.5 Incompatible properties

There is an inconvenient result that is expressed in Theorem 3.3:

Theorem 3.3. *The following properties are mutually exclusive for an isotropic basis function:*

1. *compact support;*
2. *partition of unity.*

Proof. Here, we first provide a sketch of the argument. The partition-of-unity condition implies the configuration (3.16) of zeros of the Fourier transform of the basis function. At the same time, the Hankel transform of an even compactly supported function is an entire function of finite exponential type. Jensen's theorem provides a contradiction between these two properties.

We prove Theorem 3.3 using a proof by contradiction. We suppose that there is a compactly supported isotropic function ϕ that satisfies the partition-of-unity condition. Then, using Jensen's theorem, we obtain a contradiction.

Theorem 3.4 ([61]). *(J.L. Griffith) Let $\nu > -1/2$ and $1/p + 1/q = 1$. Let f be an even entire function of exponential type 1. If $1 < p \leq 2$ and $t^{\nu+1/2} f(t) \in L^p(0,\infty)$, then f can be represented by*

$$f(z) = \int_0^1 (xz)^{-\nu} J_\nu(xz)\phi(x)dx \; (z \in \mathbb{C}), \tag{3.24}$$

with $x^{-\nu-1/2}\phi(x) \in L^q(0,1)$. Conversely, if f has this representation and $x^{-\nu-1/2}\phi(x) \in L^p(0,1)$, $1 < p \leq 2$, then f is an even entire function of exponential type 1 such that $t^{\nu+1/2} f(t) \in L^q(0,\infty)$.

Without loss of generality, let us assume that $\phi(\mathbf{x}) = 0$, for $\|\mathbf{x}\| \geq 1$. We have the following:

- The function ϕ is isotropic, so its Fourier transform is the Hankel transform of the function $\overline{\phi}(x) = \phi(\|\mathbf{x}\|)$ with $x = \|\mathbf{x}\|$. We write that

$$\mathscr{F}\{\phi\}(\omega) = 2\pi \int_0^\infty x\phi(x)J_0(\|\omega\|x)\mathrm{d}x. \tag{3.25}$$

- We define

$$f(z) = 2\pi \int_0^\infty x\phi(x) J_0(zx) \mathrm{d}x, \tag{3.26}$$

so $f(\|\boldsymbol{\omega}\|) = \mathscr{F}\{\phi\}(\boldsymbol{\omega})$. According to Theorem 3.4 (with $\nu = 0$),

$$f(z) = \int_0^\infty \psi(x) J_0(zx) \mathrm{d}x, \tag{3.27}$$

where $\psi(x) = 2\pi x\overline{\phi}(x)$. Since $x^{-\frac{1}{2}}\psi(x) \in L^2(0,1)$, f is an even entire function of exponential type 1.

- Satisfying the partition of unity is equivalent to having the equality in the Fourier domain

$$\widehat{\phi}(2\pi\mathbf{n}) = \delta[\mathbf{n}], \tag{3.28}$$

where $\mathbf{n} \in \mathbb{Z}^2$ and δ is the two-dimensional Kronecker-delta function. It means that the set of zeros of $f(z)$ is $\{z = 2\pi\|\mathbf{n}\|, \forall \mathbf{n} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\}\}$. Therefore,

$$n(R) \geq cR^2, \tag{3.29}$$

where $n(R)$ is the number of zeros in the circle with radius R and c is a positive constant.

- Jensen's theorem implies the inequality

$$\int_0^R \frac{n(t)}{t} \mathrm{d}t \leq \max_{|z|=R} \log|f(z)|. \tag{3.30}$$

This inequality restricts the number of zeros inside the disc. We have that

$$\begin{aligned} n(R/2)log2 &= \int_{R/2}^R \frac{n(R/2)}{t} \mathrm{d}t \\ &\leq \int_{R/2}^R \frac{n(t)}{t} \mathrm{d}t \\ &\leq \max_{|z|=R} \log|f(z)|. \end{aligned} \tag{3.31}$$

- Since f is of exponential type 1, it implies that $|f(z)| \leq Ae^{|z|}$. Therefore,

$$\max_{|z|=R} \log|f(z)| \leq CR, \tag{3.32}$$

  where C is a positive constant.

- Equations (3.29), (3.31), and (3.32) imply that

$$\begin{aligned} c(R/2)^2 log2 &\leq n(R/2)log2 \\ &\leq \max_{|z|=R} \log|f(z)| \\ &\leq CR. \end{aligned} \tag{3.33}$$

  Taking R sufficiently large, we reach a contradiction.

□

3.2 Basis functions

Here, we investigate two favorable Basis functions, box splines and Kaise-Bessel window functions in order to discretize the projection operator. We first discuss box-splines, particularly B-spline functions which satisfy all the desirable properties of basis functions for tomographic application except the isotropy one. Note that high degree B-splines are approximately isotropic. We show that the space of these functions are close under the x-ray transform and we derive the analytical formula for their x-ray projection. We then present Kaiser-Bessel window functions which are compactly supported and isotropic. As implied by Theorem 3.3, these functions do not satisfy the partition of unity condition. Subsequently, we propose an optimal parameter selection based on approximation theory to have minimal deviation from the partition of unity condition.

The main interest of basis functions is to provide an effective and consistent way to discretize the forward model of a computed-tomography reconstruction problem. The basis for such an approach is to characterize one image by its coefficients $\mathbf{c} = (c_{\boldsymbol{k}})_{\boldsymbol{k}\in\Omega}$ where Ω denotes the domain of the image and to apply (3.4) to obtain the simulated line-integral measurements (x-ray transform). The image reconstruction is then formulated as a regularized least-square optimization problem that is solved iteratively. Specifically, the

reconstructed image is determined as

$$\mathbf{c}^{\star} = \arg\min_{\mathbf{c}} \left\{ \|\mathbf{g} - \mathbf{H}\mathbf{c}\|^2 + \lambda R(\mathbf{c}) \right\}, \tag{3.34}$$

where $\mathbf{g}$ is the input measurement vector and $\mathbf{H}$ the matrix-representation of the forward model. The quantity $R(\mathbf{c})$ is a regularization functional (e.g., the energy of the gradient of the image or its total variation) that penalizes non-desirable solutions; it is a way of introducing prior information on the solution to make the problem well-posed. The scalar parameter $\lambda \geq 0$ is a tradeoff factor that balances the fitting accuracy versus the amount of regularization.

The success of such a reconstruction algorithm depends on two factors: 1) the quality and accuracy of the forward model, and 2) the constraints that are imposed by the regularization. The latter is very much application-dependent and becomes especially relevant when the reconstruction problem is ill-posed (e.g., limited angle tomography). Since the appropriate choice of the regularization is a whole field of investigation in itself, we focus here on the assessment of the quality of the forward model.

To that end, we consider a well-conditioned scenario where the measurement noise is negligible and the number of projection angles is sufficient to reconstruct the image by numerical inversion of the forward model. Our series of experiments is set-up such that the number of degrees of freedom of the image model (square grid of size $M \times M$) matches the number of measurements (M properly-sampled projections in an equiangular configuration). The reconstruction is performed by solving (3.34) with $\lambda = 0$ iteratively (least-squares solution) using the conjugate Gradient (CG) method. In the next sections, we investigate box splines and Kaiser-Bessel window functions, separately. We evaluate the performance of each one using the proposed scheme independently. We then conclude and discuss about their advantages and disadvantages.

3.3 Box splines

Box splines are smooth piecewise polynomial functions defined in $\mathbb{R}^d$ that are (non-separable) generalization of univariate B-splines to the multivariate setting. The definitive reference on the subject is the monograph by de Boor and Hölig [43], which is rather mathematically-oriented. Here, we briefly summarize the results of box spline theory that are pertinent to the derivations in this section. In particular, we emphasize the convolutional interpretation of these functions and their intimate connection with directional derivative operators.

3.3.1 Basic geometric definition

Geometrically, a box spline is the shadow (i.e., x-ray image) of a hypercube, in $\mathbb{R}^N$, when projected to a lower-dimensional space, $\mathbb{R}^d$ ($N \geq d$). A box spline is defined for a set of N vectors $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_N$ in $\mathbb{R}^d$. Each of these vectors is the shadow of an edge of the N-hypercube adjacent to its origin. The matrix of directions $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1 \boldsymbol{\xi}_2 \ldots \boldsymbol{\xi}_N]$ completely specifies the box spline in $\mathbb{R}^d$. Note that the vectors in this (multi-) set need not be distinct as they can appear with some multiplicity. When $N = d$, the box spline is simply the (normalized) indicator function of the parallelepiped formed by d vectors in $\mathbb{R}^d$:

$$M_{\boldsymbol{\Xi}}(\mathbf{x}) = \begin{cases} \frac{1}{|\det \boldsymbol{\Xi}|} & \mathbf{x} = \sum_{n=1}^{d} t_n \boldsymbol{\xi}_n \text{ for some } 0 \leq t_n \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

For $N > d$, box splines are defined recursively by a "directional" convolution which makes them particularly suitable for the Radon transform:

$$M_{\boldsymbol{\Xi} \cup \boldsymbol{\xi}}(\mathbf{x}) = \int_0^1 M_{\boldsymbol{\Xi}}(\mathbf{x} - t\boldsymbol{\xi}) \mathrm{d}t. \tag{3.35}$$

When the lower dimensional space is $\mathbb{R}$ (i.e., $d = 1$), the box splines coincide with univariate B-splines (**b**asic splines). When the distinct column vectors of $\boldsymbol{\Xi}$ are orthogonal to each other, box splines amount to tensor-product B-splines.

The shifts of $M_{\boldsymbol{\Xi}}$ on $\mathbb{Z}^d$ form the spline space

$$S_{M_{\boldsymbol{\Xi}}} = \operatorname{span} \left(M_{\boldsymbol{\Xi}}(\cdot - \mathbf{k}) \right)_{\mathbf{k} \in \mathbb{Z}^d}. \tag{3.36}$$

If κ is the minimal number of directions whose removal from $\boldsymbol{\Xi}$ makes the remaining directions not span $\mathbb{R}^d$, then all polynomials up to degree $(\kappa - 1)$ are contained in $S_{M_{\boldsymbol{\Xi}}}$ [43]; also, the *approximation order* of $S_{M_{\boldsymbol{\Xi}}}$ is κ. Furthermore, the continuity of the box spline is at least

$$M_{\boldsymbol{\Xi}} \in C^{\kappa - 2}(\mathbb{R}^d). \tag{3.37}$$

3.3.2 Elementary box spline constituents

Another way of constructing box splines, which is probably more transparent to engineers, is by repeated convolution of elementary line-segment-like distributions. Specifically, we have

$$M_{\boldsymbol{\Xi}}(\mathbf{x}) = \left(M_{\boldsymbol{\xi}_1} * \cdots * M_{\boldsymbol{\xi}_N} \right)(\mathbf{x}) \tag{3.38}$$

where the elementary box splines, $M_{\boldsymbol{\xi}_n}$, are Dirac-like line distributions supported over $\mathbf{x} = t\boldsymbol{\xi}_n$ with $t \in [0,1]$ with a unit integral. These elementary box splines are in direct geometric correspondence (via a rotation and a proper scaling) with the primary box spline

$$M_{\vec{e}_1}(\mathbf{x}) = \text{box}(x_1)\delta(x_2, \cdots, x_d)$$

where $\delta(x_2, \cdots, x_d)$ is the $(d-1)$-dimensional Dirac distribution and

$$\text{box}(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Moreover, they integrate to 1 which is a property that is shared by all box splines (and also preserved through convolution).

Based on (3.38), one directly infers that the box splines are positive, compactly-supported functions. Their support is a zonotope, which is the Minkowski sum of N vectors in $\boldsymbol{\Xi}$. The center of the support of $M_{\boldsymbol{\Xi}}(\mathbf{x})$ is given by $\mathbf{c}_{\boldsymbol{\Xi}} = \frac{1}{2}\sum_{n=1}^{N} \boldsymbol{\xi}_n$. The Fourier transform of the box spline is therefore given by:

$$\begin{aligned} \hat{M}_{\boldsymbol{\Xi}}(\boldsymbol{\omega}) &= \prod_{n=1}^{N} \frac{1 - \exp\left(-\mathrm{j}\left\langle \xi_n, \boldsymbol{\omega} \right\rangle\right)}{\mathrm{j}\left\langle \xi_n, \boldsymbol{\omega} \right\rangle} \\ &= \exp\left(-\mathrm{j}\left\langle \mathbf{c}_{\boldsymbol{\Xi}}, \boldsymbol{\omega} \right\rangle\right) \prod_{n=1}^{N} \text{sinc}\left(\frac{\left\langle \boldsymbol{\xi}_n, \boldsymbol{\omega} \right\rangle}{2\pi}\right), \end{aligned} \tag{3.39}$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_d)$ is the multivariate frequency vector. Alternatively, a **centered box spline**, denoted by $M^c_{\boldsymbol{\Xi}}$ which is shifted to the origin has the simple Fourier transform

$$\hat{M}^c_{\boldsymbol{\Xi}}(\boldsymbol{\omega}) = \prod_{n=1}^{N} \text{sinc}\left(\frac{\left\langle \boldsymbol{\xi}_n, \boldsymbol{\omega} \right\rangle}{2\pi}\right). \tag{3.40}$$

3.3.3 x-ray projection of box splines

We now turn to our main objective, which is the derivation of an explicit formula for $\mathscr{P}_{\boldsymbol{\theta}}\{M_{\boldsymbol{\Xi}}\}(\mathbf{y})$ where $M_{\boldsymbol{\Xi}}$ is a given box spline generator specified by N direction vectors $\boldsymbol{\xi}_n \in \boldsymbol{\Xi}$. In the following discussion, $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_{d-1})$ denotes the $(d-1)$-variate frequency vector corresponding to the projection-domain spatial coordinate vector $\mathbf{y} \in \mathbb{R}^{d-1}$, while the projection geometry is the specified in Section 2.1.

Theorem 3.5 (Sinogram-domain Parameterization)**.** *The x-ray transform of a d-variate box spline specified by the direction set, Ξ, is a $(d-1)$-variate box spline whose direction set, $\mathbf{Z} = [\zeta_1\, \zeta_2\, \ldots\, \zeta_N]$, is the geometric projection of the former. Specifically,*

$$\mathscr{P}_{\boldsymbol{\theta}}\{M_{\Xi}\}(\mathbf{y}) = M_{\mathbf{P}_{\boldsymbol{\theta}^\perp}\Xi}(\mathbf{y}).$$

where $\mathbf{P}_{\boldsymbol{\theta}^\perp}$ is the transformation matrix that geometrically projects the canonical system onto the coordinate system perpendicular to $\boldsymbol{\theta}$.

Proof. We start with the derivation of the x-ray transform of the elementary (Dirac-type) box spline $M_{\boldsymbol{\xi}}$ whose distributional Fourier transform is

$$\hat{M}_{\boldsymbol{\xi}}(\boldsymbol{\omega}) = \frac{1-\exp(\langle \boldsymbol{\xi}, \boldsymbol{\omega}\rangle)}{\mathrm{j}\langle \boldsymbol{\xi}, \boldsymbol{\omega}\rangle}.$$

We can proceed geometrically by determining the "shadow" of the direction vector $\boldsymbol{\xi}$ since the latter specifies the support of the elementary box spline as a line segment in $\mathbb{R}^d$. The alternative is to apply the central slice theorem which states that the Fourier transform of $\mathscr{P}_{\boldsymbol{\theta}}\varphi(\mathbf{y})$ corresponds to the restriction of $\hat{\varphi}(\boldsymbol{\omega})$ to the hyperplane perpendicular to $\boldsymbol{\theta}$. Specifically, we have that

$$\widehat{\mathscr{P}_{\boldsymbol{\theta}} M_{\boldsymbol{\xi}}}(\nu) = \hat{M}_{\boldsymbol{\xi}}(\boldsymbol{\omega})\big|_{\boldsymbol{\omega}=\mathbf{P}_{\boldsymbol{\theta}^\perp}^T \nu} = \hat{M}_{\mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}}(\nu).$$

Since $\left\langle \boldsymbol{\xi}, \mathbf{P}_{\boldsymbol{\theta}^\perp}^T \nu \right\rangle = \langle \mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}, \nu\rangle$, we can define the *projected directions* by $\boldsymbol{\zeta} = \mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}$. This allows us to deduce that

$$\mathscr{P}_{\boldsymbol{\theta}}\{M_{\boldsymbol{\xi}}\}(\mathbf{y}) = M_{\mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}}(\mathbf{y}) = M_{\boldsymbol{\zeta}}(\mathbf{y}). \tag{3.41}$$

This proves the theorem for $N = 1$. By defining

$$\mathbf{Z} = [\mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}_1 \cdots \mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}_N] = \mathbf{P}_{\boldsymbol{\theta}^\perp}\Xi, \tag{3.42}$$

we are then able to transfer the result to the general case using convolution properties of X-ray transform and (3.38). □

The theorem is illustrated in Figure 3.2. The box spline on the right is a trivariate tensor-product B-spline (first order) whose direction vectors are $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$. When projected to the plane orthogonal to $\boldsymbol{\theta}$, it yields a bivariate, three-direction,

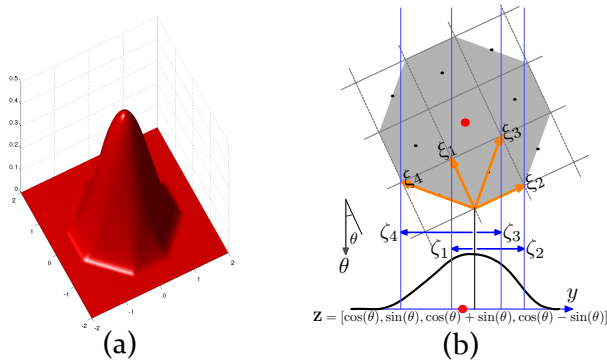


Figure 3.2: The x-ray transform of a box spline is a box spline whose directions are projections of the directions of the original box spline onto the projected plane. On the right: a trivariate box spline (a tensor-product B-spline) projected to 2-D.

box spline that is a hat function with hexagonal support. Likewise, the x-ray transform of the trilinear B-spline (second order) is again a three-direction box spline, but with multiplicity of 2. The concept carries over to higher-order tensor-product B-splines which are transformed into three-direction box splines with repeated directions, the main point being that these can be evaluated efficiently.

Corollary 3.1 (Image-domain Parameterization)**.** *The x-ray transform of an s-variate box spline specified by the direction set, Ξ, along a direction θ is an $(s-1)$-variate box spline. The directions of the latter $(s-1)$-variate box spline are obtained by geometric projection of the directions $\xi \in \Xi$ into the (hyper) plane orthogonal to the projection direction θ:*

$$\zeta = \xi - \langle \xi, \theta \rangle \, \theta. \tag{3.43}$$

Corollary 3.2 (The Radon Transform of Box Splines)**.** *The Radon transform of an s-variate box spline specified by the direction set, Ξ, along a direction θ is a univariate box spline (i.e., a B-spline along θ). The directions (i.e., knots) of the latter univariate box spline are obtained by geometric projection of the directions $\xi \in \Xi$ onto the projection*

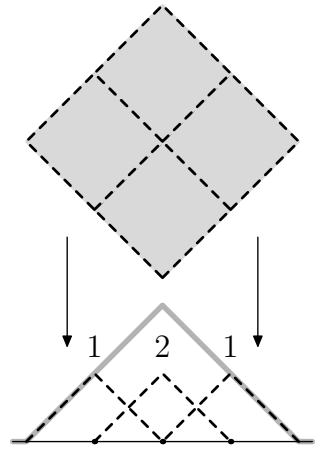


Figure 3.3: The natural multiscale relationship for box splines by projection.

direction θ*:*

$$\zeta = \langle \xi, \theta \rangle \, \theta. \tag{3.44}$$

Since any box spline is (geometrically) constructed by the shadow (x-ray) transform of a hypercube, these results establish that the space of box splines are closed under x-ray transform. These results suggest that box splines are suitable basis functions for problems involving tomographic reconstruction.

Another feature of box splines that is particularly useful in the context of tomography is their multi-scale property. Since box splines are obtained by projecting a hypercube down to a lower-dimensional space, the subdivision of the hypercube leads to a natural formula for a box spline that is written as a sum of scaled versions of itself where the scaled versions are projections of subdivided hypercubes, see Figure 3.3.

$$M_{[1\ 1]}(x) = M_{[1\ 1]}(2x-1) + 2M_{[1\ 1]}(2x) + M_{[1\ 1]}(2x+1).$$

Since this is based on the subdivision of the N-hypercube, this multiscale relationship exists in *any dimension*, d, and can be used to develop non-separable wavelets that use box

splines as the scaling functions [62, 63].

3.3.4 Explicit formulae in 2-D

For $d = 2$, we now show that the x-ray transforms of box splines are polynomial splines of degree $(N-1)$. The geometric configuration is the one shown in Figure 3.4 with the projection matrix given by $\mathbf{P}_{\boldsymbol{\theta}^\perp} = [\cos\theta \quad \sin\theta]$. The application of Theorem 3.5 together with the convolution formula (3.38) yields

$$\mathscr{P}_{\boldsymbol{\theta}} M_{\boldsymbol{\Xi}}(y) = \left(M_{\zeta_1} * M_{\zeta_2} * \cdots * M_{\zeta_N}\right)(y) \tag{3.45}$$

with $\zeta_n = \mathbf{P}_{\boldsymbol{\theta}^\perp}\boldsymbol{\xi}_n = [\boldsymbol{\xi}_n]_1 \cos\theta + [\boldsymbol{\xi}_n]_2 \sin\theta$, and

$$M_{\zeta_n}(y) = \frac{1}{\zeta_n} \mathrm{box}\left(\frac{y}{\zeta_n}\right),$$

which is a rectangular box of width ζ_n when $\zeta_n \neq 0$. Note that the convolution factors with $\zeta_n = 0$ may be eliminated from (3.45) since $M_0(y) = \delta(y)$. To evaluate the above convolution product, we write $M_{\zeta_n}(y)$ as

$$M_{\zeta_n}(y) = \Delta_{\zeta_n} u(y) \tag{3.46}$$

where $\Delta_h f(y) = \frac{f(y) - f(y-h)}{h}$ is the finite-difference operator with step h, and where u is the unit-step (or Heaviside) function. By substituting (3.46) in (3.45), we find that

$$\begin{aligned} \mathscr{P}_{\boldsymbol{\theta}} M_{\boldsymbol{\Xi}}(y) &= (\Delta_{\zeta_1} u * \cdots * \Delta_{\zeta_N} u)(y) \\ &= \frac{\Delta_{\zeta_1} \cdots \Delta_{\zeta_N} y_+^{N-1}}{(N-1)!} \end{aligned} \tag{3.47}$$

where we have used the fact that the $(n-1)$-fold convolution of a step function is $\frac{y_+^n}{n!}$ with $y_+^n = \max(y, 0)^n$. Finally, we may expand the finite-difference operators which yield a linear expansion of $\mathscr{P}_{\boldsymbol{\theta}} M_{\boldsymbol{\Xi}}(y)$ in terms of some shifted versions of y_+^{N-1}. The result therefore implies that $\mathscr{P}_{\boldsymbol{\theta}} M_{\boldsymbol{\Xi}}(y)$ is a non-uniform polynomial spline of degree $(N-1)$, or less if some ζ_n vanishes. We can also infer that this box spline function is bell-shaped and that its support is $\sum_{n=1}^N \zeta_n$.

A case of special interest is when the 2-D basis function (or generator) is the tensor-product B-spline of degree n within $\varphi(\mathbf{x}) = \beta^n(x_1)\beta^n(x_2)$ [64]. In the present formalism,

this corresponds to a box spline with direction vectors $\boldsymbol{\xi}_1 = (1,0)$ and $\boldsymbol{\xi}_2 = (0,1)$, each having a multiplicity $(n+1)$ so that $N = 2n+2$. The specialization of (3.47) for these particular values yields an explicit formula for the Radon transform of a separable B-spline of degree n:

$$\mathscr{P}_\theta\{\beta^n(x_1)\beta^n(x_2)\}(y) = \frac{\Delta_{\cos\theta}^{n+1}\,\Delta_{\sin\theta}^{n+1}\,y_+^{2n+1}}{(2n+1)!} \tag{3.48}$$

which corresponds to the spline bikernel identified by Horbelt et al. in [65]. A MATLAB routine for computing the centered versions of these functions is provided in Appendix B.

The general result (3.47), which is valid for any 2-D box spline, is new to the best of our knowledge. For instance, the Zwart-Powell element [66] is represented by the box spline directions:

$$\boldsymbol{\Xi} = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{bmatrix}. \tag{3.49}$$

According to Theorem 3.5, its projection along the angle θ provides the univariate box spline that is specified by the directions:

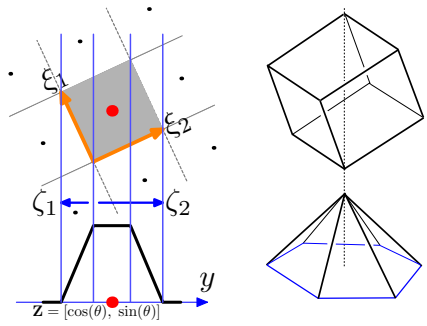
$$\mathbf{Z} = \begin{bmatrix} \cos\theta & \sin\theta & \cos\theta+\sin\theta & \cos\theta-\sin\theta \end{bmatrix}.$$

This integration process is illustrated in Figure 3.4. By applying (3.47), we find that the Radon transform of the Zwart-Powell box spline has the explicit closed-form representation:

$$\mathscr{P}_\theta\{M_{\boldsymbol{\Xi}}\}(y) = \frac{\Delta_{\cos\theta}\,\Delta_{\sin\theta}\,\Delta_{\cos\theta+\sin\theta}\,\Delta_{\cos\theta-\sin\theta}\,y_+^3}{3!}. \tag{3.50}$$

3.4 Optimized Kaiser-Bessel window function

We showed that isotropy is incompatible with the partition of unity condition for compactly support functions. Here, we investigate to which extent the problem can be deviated by adjusting the parameters of KBWF. The generalized family of KBWFs is isotropic, which makes it advantageous for the representation of the imaging operator. Our goal here is to review briefly this family and then to determine the optimal set of parameters to best attempt to recover an approximation of the partition-of-unity condition.



(a) (b)

Figure 3.4: The (non-separable) Zwart-Powell element which is a box spline associated with the directions in (3.49). The Radon transform of the Zwart-Powell box spline can be derived, exactly, using our approach.

3.4.1 Generalized Kaiser-Bessel window functions

The generalized KBWF, defined as

$$\varphi(\mathbf{x}) = \begin{cases} \frac{\left(\sqrt{1-(\|\mathbf{x}\|/a)^2}\right)^m I_m\left(\alpha\sqrt{1-(\|\mathbf{x}\|/a)^2}\right)}{I_m(\alpha)} & 0 \le \|\mathbf{x}\| \le a \\ 0 & \text{otherwise}, \end{cases} \tag{3.51}$$

is specified by three parameters: 1) the order m of the modified Bessel function I_m; 2) the window taper α; 3) the support radius a of the function. The parameter m allows us to control the smoothness of the function and the parameter α determines its shape. This function is isotropic, which makes the computation of the imaging operator significantly faster. However, it is worth noting that this function does not satisfy the partition of unity (see Theorem 3.2).

3.4.2 Measure of optimality of a basis function

If a basis function satisfies the partition-of-unity condition, then, as the sampling step vanishes, the error of approximation tends to zero. For those bases that do not satisfy the

partition of unity, we define the residual error

$$A_\varphi = \sup_{f\in L_2} \|f\|_{L_2}^{-1} \lim_{T\to 0} \varepsilon_f(T) \tag{3.52}$$

for $f \in L_2(\mathbb{R}^d)$, which shows the deviation from the partition of unity. A basis function φ with lower residual error is more desirable as generating function for the reconstruction space.

Theorem 3.6. *The residual error of a function $\varphi \in L_2(\mathbb{R}^d)$ is the quantity*

$$A_\varphi = \frac{\sum_{\mathbf{n}\neq\mathbf{0}} |\widehat{\varphi}(2\pi\mathbf{n})|^2}{|\widehat{\varphi}(\mathbf{0})|^2}. \tag{3.53}$$

Proof. Let $f \in L_2(\mathbb{R}^d)$. From (3.12), we have a formula for ε_f in terms of E_φ as defined in (3.13). We represent E_φ using its Taylor series

$$E_\varphi(T\boldsymbol{\omega}) = \sum_{|\mathbf{n}|=0}^{N} \frac{\partial^{\mathbf{n}} E_\varphi(\mathbf{0})}{\mathbf{n}!} (T\boldsymbol{\omega})^{\mathbf{n}} + o(\|\boldsymbol{\omega}\|^{N+1}), \tag{3.54}$$

where $\mathbf{n} = (n_1, n_2, ..., n_d)$ with nonnegative integer values, $|\mathbf{n}| = \sum_{i=1}^d n_i$, $\boldsymbol{\omega} = (\omega_1, \omega_2, ..., \omega_d)$, $\mathbf{n}! = n_1!n_2!...n_d!$, $\boldsymbol{\omega}^{\mathbf{n}} = \omega_1^{n_1}\omega_2^{n_2}...\omega_d^{n_d}$, and

$$\partial^{\mathbf{n}} E_\varphi(\mathbf{0}) = \frac{\partial^{n_1}}{\omega_1}\frac{\partial^{n_2}}{\omega_2}\cdots\frac{\partial^{n_d}}{\omega_d} E_\varphi(\mathbf{0}). \tag{3.55}$$

Therefore, we can rewrite the approximation error ε_f as

$$\begin{aligned}\varepsilon_f(T) &= \|f - \mathbf{P}_T\{f\}\|_{L_2} \\ &= \left(\int_{\mathbb{R}^d} \left(\sum_{|\mathbf{n}|=0}^{N} \frac{\partial^{\mathbf{n}} E_\varphi(\mathbf{0})}{\mathbf{n}!} (T\boldsymbol{\omega})^{\mathbf{n}}\right) |\widehat{f}(\boldsymbol{\omega})|^2 \frac{\mathrm{d}\boldsymbol{\omega}}{2\pi}\right)^{1/2} + \varepsilon,\end{aligned} \tag{3.56}$$

where $\varepsilon = o(\|\boldsymbol{\omega}\|^{N+1}) + \varepsilon_{\text{corr}}$. Then, Fubini's theorem implies that

$$\begin{aligned}\varepsilon_f(T) &= \left(\sum_{|\mathbf{n}|=0}^{N} \frac{\partial^{\mathbf{n}} E_\varphi(\mathbf{0})}{\mathbf{n}!} T^{|\mathbf{n}|} \int_{\mathbb{R}^d} \boldsymbol{\omega}^{\mathbf{n}} |\widehat{f}(\boldsymbol{\omega})|^2 \frac{\mathrm{d}\boldsymbol{\omega}}{2\pi}\right)^{1/2} + \varepsilon \\ &= \left(\sum_{|\mathbf{n}|=0}^{N} \frac{\partial^{\mathbf{n}} E_\varphi(\mathbf{0})}{\mathbf{n}!} T^{|\mathbf{n}|} \left\| f^{(\mathbf{n}/2)} \right\|_{L_2}^2\right)^{1/2} + \varepsilon,\end{aligned} \tag{3.57}$$

where

$$f^{(\mathbf{n})} = \frac{\partial^{n_1}}{\partial x_1}\frac{\partial^{n_2}}{\partial x_2}\cdots\frac{\partial^{n_d}}{\partial x_d} f\,. \tag{3.58}$$

We now have that

$$\lim_{T\to 0} \varepsilon_f(T) = E_\varphi(\mathbf{0})^{1/2}\,\|f\|_{L_2}\,. \tag{3.59}$$

Therefore,

$$\begin{aligned}\sup_{f\in L_2} \|f\|_{L_2}^{-2}\left(\lim_{T\to 0}\varepsilon_f(T)\right)^2 &= \sup_{f\in L_2} E_\varphi(\mathbf{0}) \\ &= \frac{\sum_{\mathbf{n}\neq\mathbf{0}}|\widehat{\varphi}(2\pi\mathbf{n})|^2}{|\widehat{\varphi}(\mathbf{0})|^2}.\end{aligned} \tag{3.60}$$

□

3.4.3 Optimal parameters for the Kaiser-Bessel window function

There are three parameters that describe KBWFs. The radius parameter a determines its support. We set it to $a = 2$; this allows us to compare the optimal KBWF with the cubic B-spline. The order of the modified Bessel function is set to $m = 2$.

In the context of 3-D imaging, Matej and Lewitt [45] empirically tune the window taper parameter α to improve the quality of reconstructed constant images. In contrast, we base our analysis on approximation-theoretic properties and determine α to minimize the residual error A_φ. Interestingly, this leads to a condition similar to the complicated criterion of [45]. But we go one step farther and provide a simplified equivalent condition in (3.53). The measure for different values α is depicted in Fig. 3.5(b). This plot indicates that values of α in the range $[6, 11.2]$ are good choices for reconstruction, with two local optima of $\alpha = 7.05, 10.45$ of comparable magnitude. The latter value is very close to 10.4, which is the value proposed in [45].

There are modalities where the reconstruction problem is separable into a set of independent 2-D problems: x-ray parallel-beam tomography, transmission electron microscopy with single-axis tilting, two-dimensional positron emission tomography systems with septa, and single-photon emission computed tomography with parallel or fan-beam collimators. Then, it is worthwhile to consider the problem in dimension two. We illustrate in Fig. 3.5(a) the residual error with respect to the parameter α in a two-dimensional space. Again, it appears that values of α in the range $[7, 11.5]$ are good choices for 2-D reconstruction, with $\alpha = 7.91, 10.83$ being the two best choices.

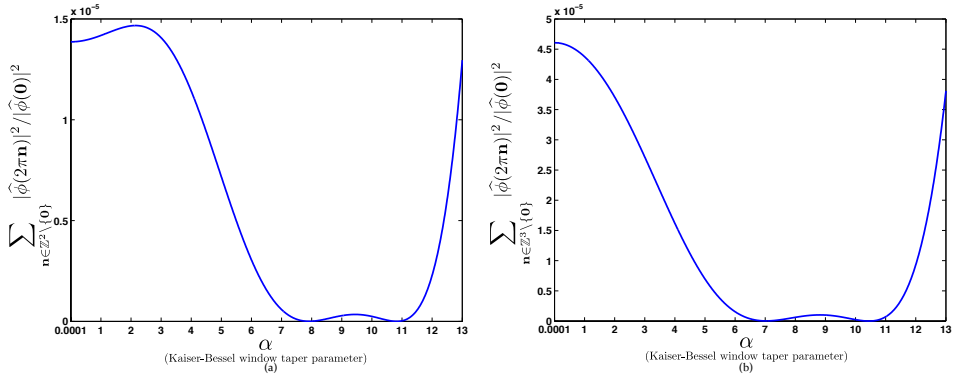


Figure 3.5: Optimality measure with respect to different values of α in the 2-D (a) and 3-D (b) domains.

3.5 Numerical evaluation

In this section, we first compare different members of the family of Box-splines family with each other and with the MATLAB radon implementation. Then, We separately investigate the performance of Kaiser-Bessel window functions with different α parameters. The experimental results validate that the Kaiser-Bessel window function with the proposed parameter has the best performance. We then compare the cubic B-spline function with the Kaiser-Bessel window function with optimal parameter. Finally, we discuss and conclude when one should use which family.

3.5.1 Box splines and Kaiser-Bessel window functions

We study the accuracy of the family of separable B-spline models as well as the non-separable Zwart-Powell box spline and compare our method with the traditional implementation of the x-ray transform. We concentrate on the piecewise linear and cubic B-spline solutions with $n = 1, 3$ in (3.48), respectively.

Our reference algorithm is the function `radon` in Matlab, which proceeds in a hierarchical fashion. It first subdivides each pixel into four sub-pixels of equal intensity and then projects each of the subcomponents using a triangular profile function (splatting). This

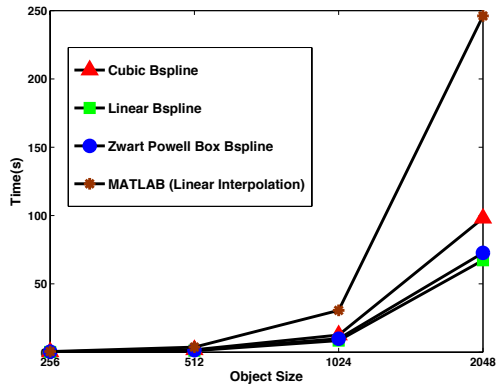


Figure 3.6: Comparison of different methods with respect to processing time.

implies that the Matlab function is at least 4 times more demanding than the first-order version (linear B-spline) of our method.

Algorithm Speed

The present family of forward models (B-spline Radon transform) was coded in C and linked to Matlab as a mex file. The adjoint operator is implemented similar to the forward model by changing the direction of the flow graph. The Radon transform of the B-splines were precomputed once and stored in a 2-D lookup table for best efficiency. To measure the speed, we computed the Radon transforms of a series of images of increasing size M. The expected computational cost is $O(M^3)$ with a proportionality factor that depends on the choice of algorithm and the size of the underlying basis function. The results are documented in Figure 3.6 and are consistent with the prediction. The Matlab implementation is the slowest, while the cubic B-spline version of our algorithm is approximatively 8 times more costly than the piecewise-linear version. The Zwart-Powell box spline is notably faster than cubic B-spline and is very close to the performance of linear B-splines.

Consistency of the Forward Model

In order to examine the ability of the proposed discretization method to capture sharp image details, we considered an analytical phantom whose Radon transform is known analytically.

Proposition 3.1. *The Radon transform of the function*

$$f(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\|^2 & \|\boldsymbol{x}\| < a \\ 0 & \text{otherwise,} \end{cases} \tag{3.61}$$

where $a \in \mathbb{R}^+$, is

$$\mathscr{R}_\theta\{f\}(y) = \begin{cases} \frac{2}{3}\sqrt{a^2-y^2}(a^2+2y^2) & |y| < a \\ 0 & \text{otherwise.} \end{cases} \tag{3.62}$$

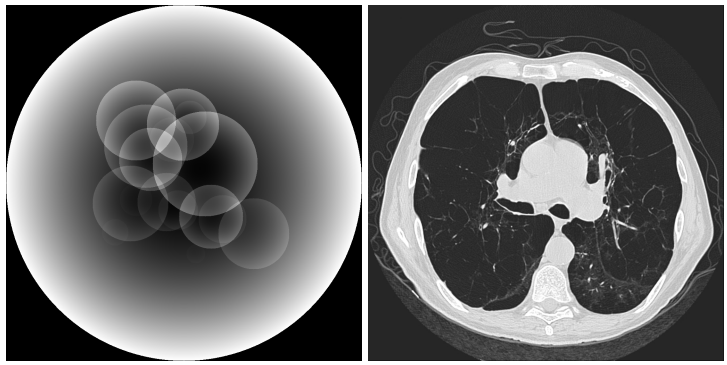
This can be readily verified by evaluating the integral

$$\mathscr{R}_\theta\{f\}(y) = 2 \times \int_0^{\sqrt{a^2-y^2}} (x^2+y^2)\mathrm{d}x.$$

Using the linearity and the projected shift-invariance property of the x-ray transform, we use this result to determine the Radon transform of the object

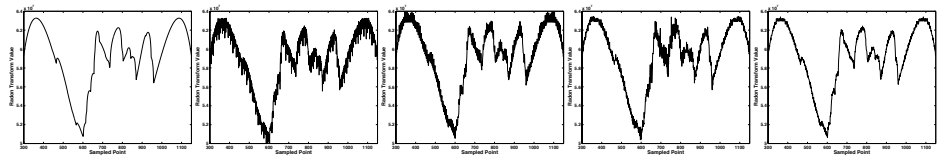
$$I(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \alpha_{\boldsymbol{k}} f(\boldsymbol{x} - \boldsymbol{x}_{\boldsymbol{k}}),$$

where $\alpha_{\boldsymbol{k}} \in \mathbb{R}$ and $\boldsymbol{x}_{\boldsymbol{k}} \in \mathbb{R}^2$ are some prescribed parameters. For our experiments, we considered the analytical phantom shown in Figure 3.7. Starting from a 1024×1024 representation, we calculated its Radon transform along 1024 directions with the help of the different algorithms and compared the output with the analytical one. Examples of projections are shown in Figure 3.8. The higher-order versions of our spline models produce the sinograms that are the most faithful to the analytical ones. The Matlab results in 3.8(c) are not quite as favorable as the cubic B-spline 3.8(e) and Zwart-Powell box spline 3.8(d), although they oscillate less than the linear spline version 3.8(b). This ranking is confirmed by the global signal-to-noise ratio (SNR) presented in Table 3.1.



(a) (b)

Figure 3.7: (a) The analytical phantom includes 30 circles with different quadratic intensity distributions. (b) Lung image used for the evaluation of forward model and reconstruction. These datasets at the resolution of (1024×1024) serve as the ground truth for our experiments.



(a) Analytical Radon transform (b) Linear B-spline (c) Matlab Radon operator (d) Zwart-Powell box spline (e) Cubic B-spline

Figure 3.8: Radon transform of the phantom along $\theta = \frac{\pi}{4}$. Both the Zwart-Powell box spline and the cubic B-spline outperform Matlab's Radon operator.

Table 3.1: Comparison of different discrete models of X-ray Transform.

Method	MATLAB	Degree1	Zwart-Powell	Degree3
SNR(dB)	43.85	39.88	44.65	52.75

Reconstruction error as a function of grid size

In a well-conditioned scenario with a sufficient number of measurements, we can expect the quality of the reconstruction to be depend on the grid size (the degrees of freedom of the reconstruction model). Yet, we also know from approximation theory that not all basis functions are equally good at representing arbitrary signals at a fixed resolution. From a sampling point of view, B-splines are optimal in the sense that they have the maximal order of approximation for a given support.

To investigate the dependence upon the sampling rate, we conducted a series of experiments using the framework described in Section 3.2 where the grid size is progressively reduced. The reference object and signal-to-noise computations are defined with respect to the fine grid (e.g., 1024×1024). The coarse grid measurements are obtained by suitable angular and spatial resampling of the fine-grid Radon transform of the object. An ideal lowpass filter is applied in the spatial domain prior to downsampling to avoid aliasing. The object is reconstructed on the coarse grid using the different flavors of the forward model. The reconstruction is calculated iteratively and corresponds to the least squares solution. Finally, the result is interpolated back to the finer grid (resampling of the spline model) for quality assessment. For the Matlab version, we used a linear interpolation which was found to give better results than a cubic interpolation (for upsampling).

Analytical phantom for analysis of accuracy

In the case of the phantom in Figure 3.7(a), we used the analytical calculation of the Radon transform as initial fine-grid measurements. We then performed the various signal reconstructions for $M = 1024$, 512, 256, and 128 using the corresponding down-sampled versions of the input data. The evolution of the signal-to-noise ratio as a function of the down-sampling factor is shown in Figure 3.9. Figure 3.10 compares the reconstruction results for a central region of the phantom that is reconstructed from 256 projections (down-sampling by $(4,4)$). The specific region of the phantom is depicted in Figure 3.10. Note that the best results are obtained with the cubic B-spline (3.10(e)) and Zwart-Powell box spline

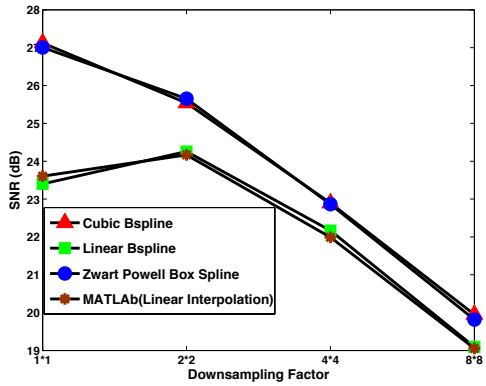


Figure 3.9: Evolution of the signal-to-noise ratio for the least-squares reconstruction of the analytical phantom as both the grid size and the number of measurements are reduced.

(3.10(d)) models. In particular, we can distinguish some of the fainter circles (e.g., at the bottom left) that are barely visible in the other reconstructions. The differences between the Matlab model and the linear B-spline reconstruction are less significant, although there may be a slight preference towards the former because of its smoother appearance.

Biomedical data

Next, we considered the cross section of the Lung image shown in Figure 3.7(b) as representative example of a medical image. Its Radon transform was calculated using the different forward models and the results averaged to specify a fine-scale set of measurements that is not biased towards one of the methods. We then performed the same experiments as in the previous case. The corresponding evolution of the SNR is shown in Figure 3.11. Figure 3.12 presents a region of interest that was reconstructed from 256 projections. The conclusion that can be drawn are essentially the same as in the previous experiment; namely, that the cubic B-spline (3.12(e)) and Zwart-Powell box spline (3.12(d)) basis functions outperform the others. This is significant specially when considering the computational performance offered by these two algorithms.

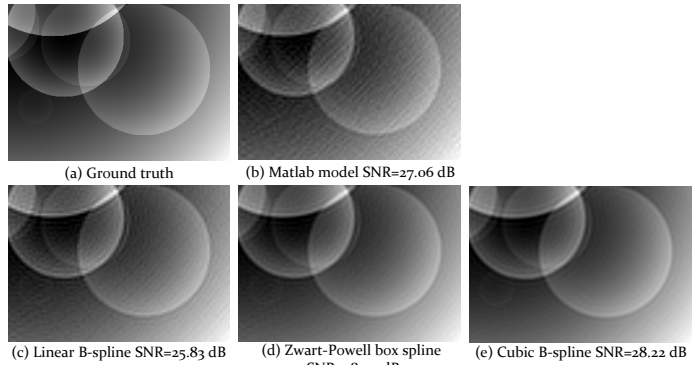


Figure 3.10: Reconstructed phantom from 256 projections using different discrete forward models. Both Zwart-Powell box spline and cubic B-spline outperform the Matlab's reconstruction. Besides having smaller artifacts in the reconstruction, the faint circle in the bottom left is better reconstructed in (d) and (e).

In addition, we did also reconstruct images using the quintic B-spline version of the method which is computationally more expensive, but did not observe any significant improvement over the cubic spline reconstruction which appears to offer an excellent trade-off in terms of cost/quality. We believe that the present cubic B-spline and Zwart-Powell box spline versions of the Radon transform are to be preferred over the standard Matlab implementation because they consistently yield better quality results while being computationally quite competitive.

We also performed experiments with real biomedical data (i.e., Human bicuspid calcific heart valve derived nodule) acquired from a CT scanner. The computational advantages of our spline model as well as improvements in the accuracy of reconstruction, afforded by higher order basis functions, were similar to the Lung data experiments reported in Figure 3.12. Moreover, we have performed an experiment to reconstruct edges of a test image to evaluate the impact of higher order basis functions on the sharpness of reconstruction. This experiment illustrates, numerically, that the higher order basis functions do provide a more accurate reconstruction for preserving the edges.

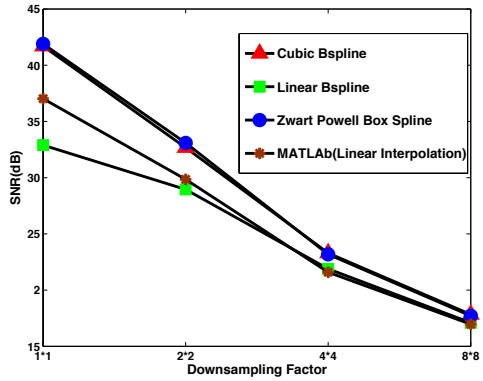


Figure 3.11: Evolution of the signal-to-noise ratio for the least-squares reconstruction of the lung as both the grid size and the number of measurements are reduced.

3.5.2 Optimality of the proposed taper parameter for KBWFs

We now present experiments where we numerically evaluate the discretization scheme based on KBWFs, with the parameters suggested in the Section 3.4.

Influence of the discretization step

It is clear that the optimal reconstruction in the least-square sense is the orthogonal projection of the sample on the reconstruction space, independently of the chosen algorithm. To investigate the dependence upon the grid size, we compute the optimal reconstruction with respect to different grid sizes when the generating function of the reconstruction space is a KWBF with different parameters. The reference object and signal-to-noise (SNR) computations are defined with respect to the fine grid. The SNR is defined as the relative mean-square with respect to the reference (oracle). The grid size is progressively increased, which shows the dependence upon the sampling rate.

We choose two medical samples: a coronal section of a human lung and a coronal section of a rat brain. Also, a region of interest has been chosen as shown in Figs. 3.13(a) and 3.13(b). We first tested the KBWF with $\alpha = 2$, which is well outside of the optimal

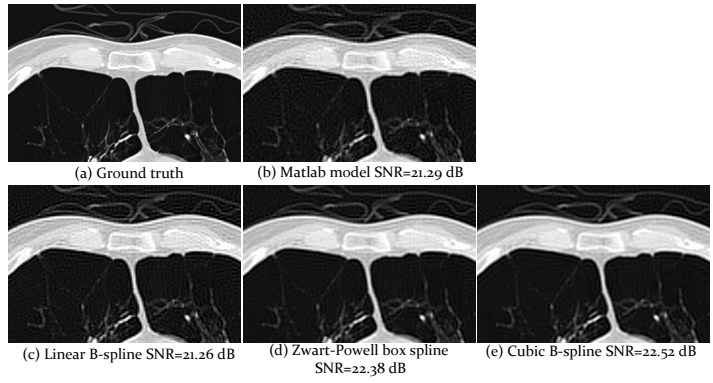


Figure 3.12: Central region of the reconstructed lung (a) from 256 projections using different discrete forward models. The linear B-spline (b) is comparable to the Matlab's approach , while both Zwart-Powell box spline (d) and Cubic B-spline (e) provide the most accurate reconstructions. The difference in the reconstructions are mostly visible in areas close to edges as well as the white structures within the gray areas.

interval $[7, 11.5]$, and the results were very poor (SNR= 4 dB). We then compared the performance for the value $\alpha = 5$ and $\alpha = 7.91$. The former is close to, but outside of the optimal interval, while the latter is the first of our proposed choices. Their performances are depicted in Figs. 3.13(c) and 3.13(d). It confirms that using KBWFs with the proposed parameter has better optimal reconstruction compared to $\alpha = 5$ for different grid sizes. This experiment shows that the "optimal choice" based on the asymptotical behavior (see (3.52)) is also always better for different grid sizes.

Reconstruction of an analytical phantom

As data, we use the two-dimensional synthetic phantom presented in [36] and shown in Fig. 3.14. (a). The analytical formula for computing imaging transforms of the phantom is given in [36, Section 4.4].

Conventional tomography

The first experiment is as follows: The size of the phantom for this experiment is $(2{,}048 \times 2{,}048)$ pixels. The sinogram of the phantom is computed analytically with 1,800 viewing angles that are chosen uniformly between 0 and π; we consider it as the measurements. We use the framework suggested in Section 3.2.

The object is reconstructed on a grid that is (4×4) times coarser than the discretization grid. Then, the basis function helps us to resample the object on a finer grid. We use the conjugate-gradient algorithm for the minimization. As the number of directions is on the order of the size of the object, we do not use any regularization. The signal-to-noise ratio (SNR) of the reconstructions and the projection versus different values of the window taper of KBWFs are shown in Fig. 3.14(c) and 3.14(d). The best performance is obtained by using a KBWF with $\alpha = 7.75$, which is very close to the first minimum of our criterion function in Fig. 3.5. However, values of α in the range $[7, 11.5]$ do also perform reasonably well, which is consistent with the theoretical analysis of Section 3.4.3.

Differential phase-contrast tomography

We evaluate the performance of KBWFs with the proposed parameters in x-ray differential phase-contrast tomography. The mathematical model of this imaging modality is based on the derivative of the Radon transform.

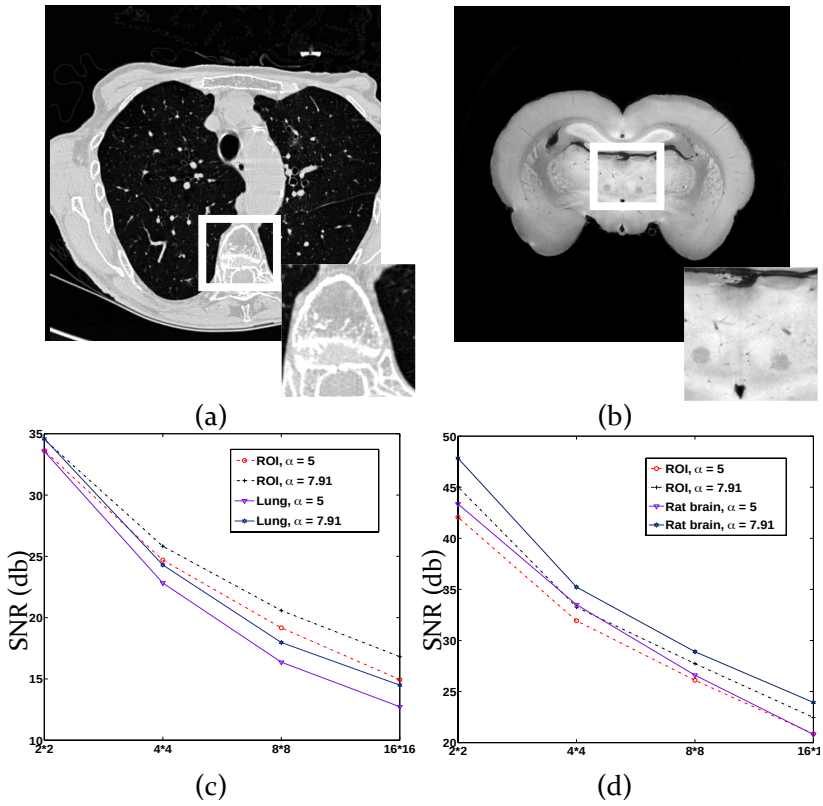


Figure 3.13: (a) Coronal section of a human lung and region of interest. (b) Coronal section of a rat brain and region of interest. The performance of the optimal solution with respect to the grid size is depicted in (c) and (d).

The differentiated sinogram of the phantom with size (512×512) pixels is again computed analytically with 1,800 viewing angles that are chosen uniformly between 0 and π; we consider it as the measurements. As there is a large number of views for the reconstruction, we minimize the least-squares error for the reconstruction. This is done for different

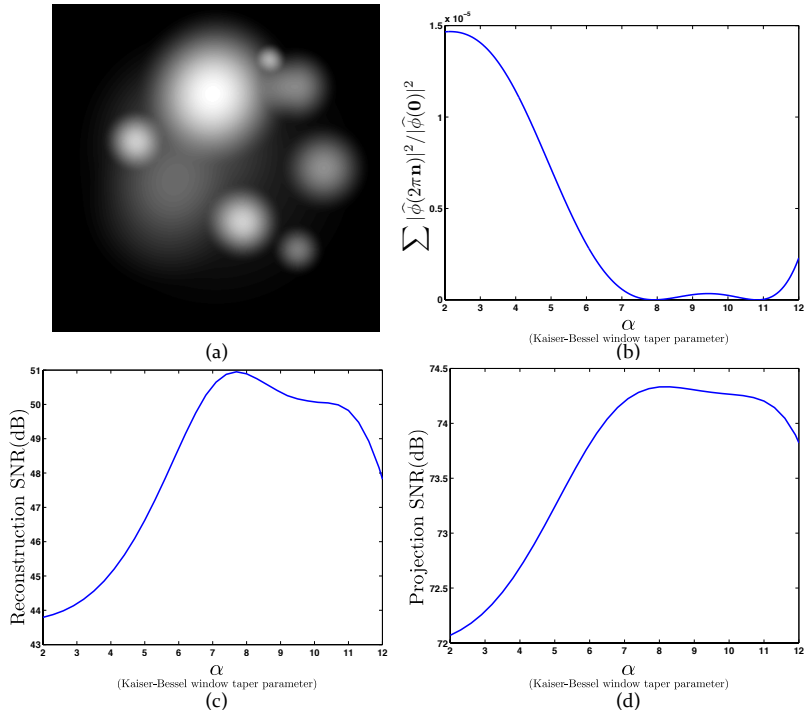


Figure 3.14: 2-D analytical phantom with isotropic elements (a). Zoomed version of the proposed measure (b). The accuracy of the reconstruction of the analytical phantom versus the window taper parameter of KBWFs is shown in (c). Its Radon- transform error in the same coarse grid is depicted in (d).

discretizations of the forward model using KBWFs with different taper parameters. Therefore, the quality of the reconstructed image depends on how well the discretization scheme represents the imaging operator, as shown in Fig. 3.15(a). We also compute the SNR in the transform domain (Fig. 3.15(b)). The results validate the importance of using KBWFs with optimized parameters in order to improve the reconstruction performance. We also repeated those experiments with measurements corrupted by additive Gaussian noise with

different noise levels (10 dB, 20 dB, 30 dB). The results suggest that using KBWF with the proposed parameters results in better performance. The SNR of the reconstructions was improve by close to 3 dB with respect to $\alpha = 5$.

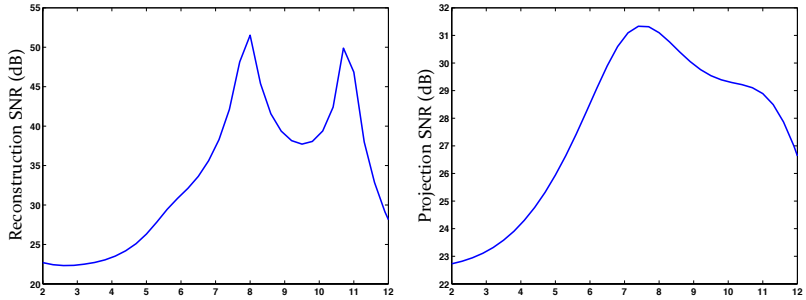


Figure 3.15: Performance of the reconstruction (a) and projection (b) using KBWFs for differential phase-contrast tomography versus the window taper parameter.

3.5.3 B-splines vs Kaiser-Bessel

KBWFs with the optimal parameter converge to cubic B-spline

We first compute the mean square error of the tensor product of two cubic B-splines (see Figure 3.16 (a)) with the Kaiser-Bessel window functions of the same support with respect to different window taper parameters α as shown in Figure 3.16 (c). Interestingly, Kaiser-Bessel window functions with the optimal window taper parameter are the closset one to the cubic B-spline based on the mean square error. It shows that the KBWF with optimal parameter resembles to cubic B-spline function. Kaiser-Bessel window function with optimal parameter and its difference with cubic B-spline are depicted in Figures 3.16 (b) and 3.16 (d), respectively. On the other hand, it suggests that the tensor product of two cubic B-spline functions is approximately isotropic.

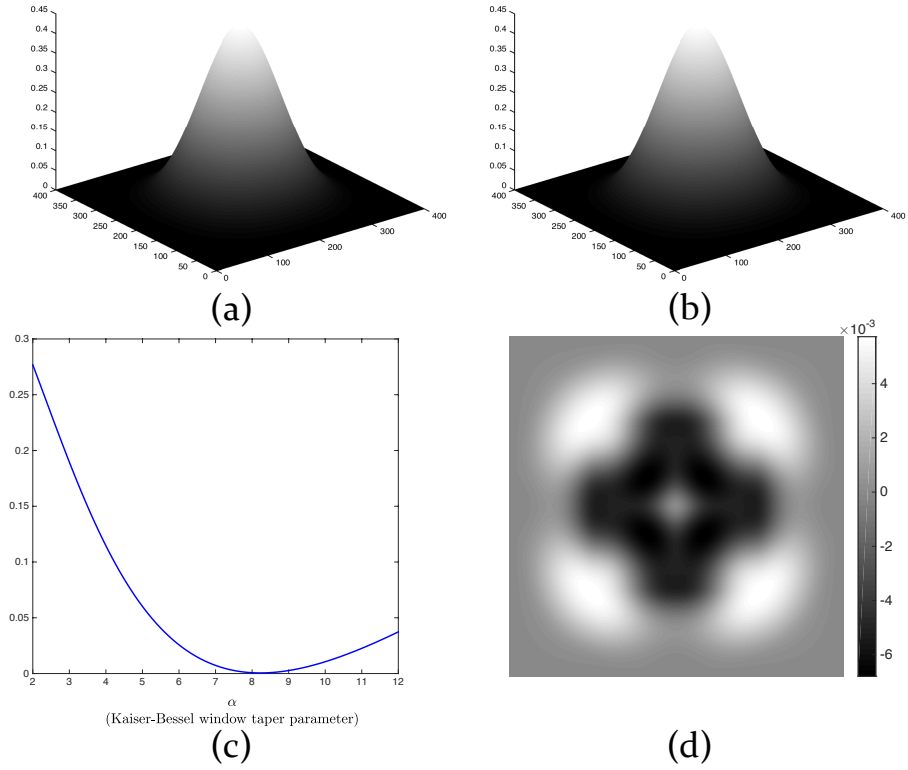


Figure 3.16: (a) Tensor product of two cubic B-splines. (b) Kaiser-Bessel window function with the proposed $\alpha = 8$. Mean-square error between the tensor product of two cubic B-splines and Kaiser-Bessel window functions versus different α values. (d) Difference of functions in (a) and (b).

Table 3.2: Comparison of the projection and reconstruction accuracy using cubic B-splines and KBWFs with the parameters proposed in [57].

SNR (dB)	KBWF $a = 2$, $m = 2$,			Cubic B-spline
	$\alpha = 2$	$\alpha = 10.4$	$\alpha = 7.95$	
Projection	23.59	27.15	29.19	29.26
Reconstruction	37.31	43.42	48.59	48.54

performance comparison

For the reconstruction of x-ray differential phase-contrast tomograms, it was shown in [36] that the use of cubic B-splines results in better performance than using KBWFs with the parameters chosen as in [45]. Here, we compare the performance of three basis functions for the phantom with size $(2{,}048 \times 2{,}048)$ pixels. The projection operator is computed using KBWFs with the parameter proposed in [44] ($\alpha = 10.4$) and with the parameter suggested in Fig. 3.15 ($\alpha = 7.95$); furthermore, we also perform the comparison with cubic B-splines. The computed SNR shown in Table 3.2 suggests that the proposed parameter provides a significantly better performance in computing the projection operator.

We conclude that a KBWF with the proposed parameters improves the performance of the discretization scheme in comparison with [44, 45]. In addition, its performance is as good as that of cubic B-splines in terms of quality, while its isotropy allows for a drastic reduction in its computational costs.

3.6 Discussion and Conclusion

We investigated two favorable generating function families, box splines (in particular B-splines), and Kaiser-Bessel windows. For a fixed support, B-spline functions have a better order of approximation than Kaiser-Bessel window functions (KBWFs). We showed that, by adjusting the taper parameter of KBWF using the proposed approximation-theoretic framework, these functions perform almost as well as B-splines. In two-dimensional tomography or three-dimensional tomography with fixed rotation axis, B-splines are preferable owing to their order of approximation. In addition, the separability of the tensor product of cubic B-spline functions allows one to decompose the different three-dimensional

reconstruction in parallel-beam geometry with fixed rotation axis into some easier two-dimensional and one-dimensional subproblems. It results in the development memory-efficient reconstruction framework. In the three-dimensional problem with random orientations, the implementation of the x-ray transform using B-splines is too complicated and it is therefore more practical to use KBWFs.

Chapter 4

Reconstruction Algorithms

[1] We have already presented a direct method for objet reconstruction in the context of straight-ray imaging by inverting (2.31), namely the generalized filtered back-projection (GFBP) described in Algorithm 1. To apply the Radon adjoint in step 2 of this algorithm, we use the method proposed in Section 3.1. However, for large images with a limited number of measurements, direct methods such as FBP are not accurate enough. In order to improve the reconstruction quality, one requires to model the imaging process precisely and to formulate the problem as an inverse problem. This leads to the application of iterative reconstruction schemes; This is a hot topic for all clinical CT for the past few years, owing to the large computational capacities of normal work stations besides the ongoing efforts towards lower dose in CT.

We have chosen to categorize iterative techniques into three types: 1) iterative methods that are purely algebraical without any statistical modeling, 2) statistical methods that model the photon-counting statistics, and 3) model-based methods that go beyond statistical modeling (we refer the reader to a nice review paper [68] for more information). The first step in most of these methods is to formulate the reconstruction as an optimization problem and then to develop a proper iterative algorithm to minimize the given cost function in order to reconstruct the object of interest.

Model-based methods have been widely used in the context of absorption-based computed tomography [68], including iterative coordinate descent (ICD) methods [3], block-

[1] A part of this chapter has been presented in [36] and [67]

based coordinate descent [4], ordered-subset algorithms based on separable quadratic surrogates [5], preconditioned nonlinear conjugate-gradient methods [6] and alternating-direction method of multipliers (ADMM) [7]. These methods have been also recently developed for different types of phase-contrast tomography [33–35, 67, 69].

We develop a unified approach to the reconstruction problem in straight-ray imaging modalities. The reconstruction problem is formulated as a penalized likelihood estimator in Section 4.1. The estimator takes the form of the minimization of penalized weighted least squares. We develop an iterative reconstruction scheme to solve the minimization problem based on alternating direction method of multipliers using two variable-splitting scenarios in Subsection 4.2.1. This yields a fast (practically reasonable) reconstruction scheme. The first variable splitting is a classical approach similar to the one proposed in [7]. We then present a novel variable splitting that improves the reconstruction quality as its last step is a denoising operator. A problem-specific preconditioner is proposed that speeds-up the convergence of the linear optimization step considerably. In Subsection 4.3, we modify the proposed reconstruction scheme to promote memory efficiency in parallel-beam three-dimensional reconstruction with fixed rotation axis. We then generalize the proposed ADMM scheme for a general noise model, fan-beam or cone-beam geometry, and limited field of view in Subsection 4.2.2.

4.1 Reconstruction as an optimization problem

Considering the statistical behavior of systematic noise besides the deterministic model of the imaging operator, one can formulate the reconstruction as a maximum-likelihood estimator. The deterministic model of the imaging system was described in Section 3.1 and is given by

$$\mathbf{g} = \mathbf{H}\mathbf{c}\,, \tag{4.1}$$

where $\mathbf{g} \in \mathbb{R}^M$ is the measurement vector, $\mathbf{H} \in \mathbb{R}^{M\times N}$ is the system matrix, and $\mathbf{c} \in \mathbb{R}^N$ is the discrete representation of the unknown object. The measurement vector $\{\mathbf{g}[k]\}_{k=1}^M$ is typically modeled as a random vector with conditional probability $p(\mathbf{g}|\mathbf{c})$. The maximum-likelihood estimator of the unknown object of interest $\mathbf{c}$ is that given by

$$\begin{aligned}\widehat{\mathbf{c}} &= \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmax}}\,\{p(\mathbf{g}|\mathbf{c})\} \\ &= \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} - \sum_{k=1}^{M} \log\, p(\mathbf{g}[k]|\mathbf{c}[k])\,. \end{aligned} \tag{4.2}$$

The conditional probability $p(\mathbf{g}|\mathbf{c})$ is linked to the model of noise in the imaging system. Two main ingredients are Gaussian electronic noise and the number of photons that hit the detector which are usually described by a Poisson distribution. Most researchers approximate it as either i.i.d. zero-mean Gaussian with variance σ^2,

$$-\log\left(p(\mathbf{g}|\mathbf{c})\right) = \frac{1}{2\sigma^2}\left\|\mathbf{g}-\mathbf{H}\mathbf{c}\right\|^2, \tag{4.3}$$

or as a Poisson distributed where the mean is the average of the received intensity in that pixel that is proportional to $\mathrm{e}^{-\{\mathbf{H}\mathbf{c}\}_k}$. Then, the likelihood estimator is given by

$$-\log\left(p(\mathbf{g}|\mathbf{c})\right) = C + \sum_k \left(-\{\mathbf{H}\mathbf{c}\}_k + \mathbf{g}[k]\log\{\mathbf{H}\mathbf{c}\}_k\right), \tag{4.4}$$

where $\mathbf{g}[k]$ and $\{\mathbf{H}\mathbf{c}\}_k$ are the k th entry of $\mathbf{g}$ and $\mathbf{H}\mathbf{c}$, respectively.

A second-order Tailor approximation of the latter equation results in the square weighted-norm formulation [70, 71],

$$-\log\left(p(\mathbf{g}|\mathbf{c})\right) \approx \left\|\mathbf{g}-\mathbf{H}\mathbf{c}\right\|_{\mathbf{D}}^2, \tag{4.5}$$

with diagonal weight matrix $\mathbf{D} = \mathrm{diag}\left(\mathrm{e}^{-\mathbf{g}[k]}\right)$.

As the reconstruction problem with a limited number of views is ill-posed, one typically includes additional prior information in order to resolve the ambiguity associated with the non-empty null space of the forward imaging operator. Then the estimator can be written in the form of

$$\begin{aligned}\widehat{\mathbf{c}} &= \underset{\mathbf{c}\in\mathscr{C}}{\mathrm{argmin}}\left\{-\log\left(p(\mathbf{g}|\mathbf{c})\right) + \Psi(\mathbf{c})\right\} \\ &= \underset{\mathbf{c}\in\mathscr{C}}{\mathrm{argmin}}\left\{\frac{1}{2}\left\|\mathbf{g}-\mathbf{H}\mathbf{c}\right\|_{\mathbf{D}}^2 + \Psi(\mathbf{c})\right\},\end{aligned} \tag{4.6}$$

where Ψ is the so-called regularization or potential function. The solution of (4.6) is called penalized likelihood estimator. There are different types of regularization. One of the popular regularization in the context of image reconstruction is the ℓ_1 penalty. The idea of ℓ_1 was introduced by Huber's [72] work on robust statistics and a work of Claerbout [73] in geophysics. Some classical papers on the subject are total variation [74], compressed sensing [75–77] , soft thresholding [78], basis pursuit [79] , lasso [80], and structural learning of sparse graphical models [81].

By selecting the regularization term as the potential function which is $(-\log(p(\mathbf{c})))$ where $p(\mathbf{c})$ is the statistical model of the object, the penalized likelihood can be equivalent to maximum a posteriori estimator. Briefly, the maximum a posteriori estimator is given by

$$\widehat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmax}} \left\{ \log\left(p(\mathbf{c}|\mathbf{g})\right) \right\}, \tag{4.7}$$

where $p(\mathbf{c}|\mathbf{g})$ is the conditional probability. Using Bayes's rule, (4.8) can be rewritten as

$$\widehat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \left\{ -\log\left(p(\mathbf{g}|\mathbf{c})\right) - \log\left(p(\mathbf{c})\right) \right\}. \tag{4.8}$$

The first part is the likelihood term. Note that the a priori information, $-log(p(\mathbf{c}))$, can have different forms. In the simplest case, it is equivalent to an ℓ_2-norm with a Gaussian assumption. Moreover, it has the form of total-variation regularization when a p-variation distribution is assumed [82,83]. Consequently, the reconstruction problem is formulated as a penalized weighted-norm least square. In the next section we aim at developing iterative techniques in order to solve the optimization problem (4.6). Due to the explosion in the size and complexity of modern imaging systems, it is really important to be able to do the reconstruction and solve the corresponding optimization problem with a very large dimension in a reasonable computational time. In the next section, we use the alternating-direction method of multipliers and demonstrate that it is well suited to reconstruct real-world images.

4.2 Reconstruction algorithm

We formulate the reconstruction as a constrained optimization problem with a generalized weighted ℓ_2-norm data term. Specifically, we aim at finding the vector $\mathbf{c}_0$ such that

$$\mathbf{c}_0 = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} \left\{ \underbrace{\frac{1}{2}\|\mathbf{H}\mathbf{c}-\mathbf{g}\|_{\mathbf{D}}^2 + \lambda_1\Psi_1(\mathbf{c}) + \lambda_2\Psi_2(\mathbf{c})}_{J(\mathbf{c})} \right\}, \tag{4.9}$$

where $\mathscr{C}$ is a convex set that enforces support and positivity constraints. The regularization parameters λ_1 and $\lambda_2 \in \mathbb{R}$ control the strength of the regularization. We separate the regularization to the quadratic term Ψ_1 and non-quadratic term Ψ_2. Typically

$\Psi_2(\mathbf{c}) = \sum_i \Psi^{(i)}(\mathbf{L}^{(\mathbf{i})}\mathbf{c})$. Here, we consider only one term that we note it by $\Psi_2(\mathbf{Lc})$. The functions Ψ_1 and Ψ_2 are called potential functions and $\mathbf{L}$ is the regularization operator.

There has been a considerable development of optimization algorithms to solve such problems including gradient projection [84, 85], proximal gradient [86–88], augmented-Lagrangian methods [89], interior-point methods [90], Bergman iterative algorithms [91], and alternating-direction method of multipliers [92].

4.2.1 Alternating direction method of multipliers

High-dimensional regularized problems with a non-smooth potential function are typically solved using the family of iterative shrinkage/thresholding algorithms (ISTA). The convergence speed of these methods depend on the conditioning of $\mathbf{H}^T\mathbf{DH}$. In our case, the x-ray transform and its differential variants are particularly ill-conditioned operators leading to very slow convergence when $\mathbf{D}$ is the identity matrix. To overcome this difficulty, we use a variable-splitting scheme to map our general optimization problem into simpler ones [93–97]. We present two different variable-splitting scenarios:

A) **Alternating-direction method of multipliers using a preconditioned conjugate gradient (ADMM-PCG):** Specifically, we introduce the auxiliary variable $\mathbf{u} = \mathbf{Lc}$ and reformulate the reconstruction problem (4.9) as a linear equality-constrained problem

$$\begin{aligned} \mathbf{c} = \underset{\mathbf{c},\mathbf{u}}{\operatorname{argmin}} &\left\{ \frac{1}{2} \|\mathbf{Hc}-\mathbf{g}\|_{\mathbf{D}}^2 + \lambda_1 \Psi_1(\mathbf{c}) + \lambda_2 \Psi_2(\mathbf{u}) \right\} \\ &\text{subject to } \mathbf{u} = \mathbf{Lc}\,. \end{aligned} \tag{4.10}$$

We can map this into an unconstrained problem by considering the augmented-Lagrangian functional

$$\begin{aligned} \mathscr{L}_\mu(\mathbf{c},\mathbf{u},\boldsymbol{\alpha}) = &\frac{1}{2} \|\mathbf{Hc}-\mathbf{g}\|_{\mathbf{D}}^2 + \lambda_1 \Psi_1(\mathbf{c}) + \lambda_2 \Psi_2(\mathbf{u}) \\ &+ \boldsymbol{\alpha}^T(\mathbf{Lc}-\mathbf{u}) + \frac{\mu}{2} \|\mathbf{Lc}-\mathbf{u}\|^2\,, \end{aligned} \tag{4.11}$$

where $\boldsymbol{\alpha}$ is the vector of Lagrange multipliers. The advantage of using an augmented Lagrangian instead of the Lagrangian (including the quadratic term) is to bring robustness to the dual-ascent method for updating the Lagrangian multipliers and, in particular, to relax the assumption of strict convexity on the main cost function. The method of multipliers for

minimizing (4.11) has the form of

$$\begin{cases} \left(\mathbf{c}^{\mathbf{k+1}}, \mathbf{u}^{\mathbf{k+1}}\right) \leftarrow \underset{\mathbf{c},\mathbf{u}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}, \mathbf{u}, \boldsymbol{\alpha}^k) \\ \boldsymbol{\alpha}^{k+1} \leftarrow \boldsymbol{\alpha}^k + \mu(\mathbf{L}\mathbf{c}^{k+1} - \mathbf{u}^{k+1}). \end{cases}$$

The first step is a joint minimization which is costly to solve. In this regard, we adopt the cyclic update scheme known as alternating-direction method of multipliers, which consists of the iterations

$$\begin{cases} \mathbf{c}^{k+1} \leftarrow \underset{\mathbf{c}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^k, \boldsymbol{\alpha}^k) \\ \mathbf{u}^{k+1} \leftarrow \underset{\mathbf{u}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}^{k+1}, \mathbf{u}, \boldsymbol{\alpha}^k) \\ \boldsymbol{\alpha}^{k+1} \leftarrow \boldsymbol{\alpha}^k + \mu(\mathbf{L}\mathbf{c}^{k+1} - \mathbf{u}^{k+1}). \end{cases}$$

Hence, $\mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^k, \boldsymbol{\alpha}^k)$ is a quadratic function with respect to $\mathbf{c}$. Its gradient is

$$\boldsymbol{\nabla}\mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^k, \boldsymbol{\alpha}^k) = \underbrace{\left(\mathbf{H}^T\mathbf{D}\mathbf{H} + \mu\mathbf{L}^T\mathbf{L} + \lambda_1\nabla\Psi_1\right)}_{\mathbf{A}}\mathbf{c} - \underbrace{\left(\mathbf{H}^T\mathbf{g} + \mu\mathbf{L}^T\left(\mathbf{u}^k - \frac{\boldsymbol{\alpha}^k}{\mu}\right)\right)}_{\mathbf{b}}.$$

We minimize $\mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^k, \boldsymbol{\alpha}^k)$ iteratively using the conjugate-gradient (CG) algorithm to solve for $\mathbf{A}\mathbf{c} = \mathbf{b}$. For speeding up the convergence of the CG algorithm, we introduce two different strategies.

1. Since $\mathbf{A}$ has a large condition number, it is helpful to introduce a preconditioning matrix $\mathbf{M}^{-1}$. This matrix is chosen such that $\mathbf{M}^{-1}\left(\mathbf{H}^T\mathbf{D}\mathbf{H} + \mu\mathbf{L}^T\mathbf{L} + \lambda_1\mathrm{I}\right)$ has a condition number close to 1. To design this problem-specific preconditioner, we use the following proposition.

   Proposition 4.1. *The successive application of the derivatives of the Radon transform and their adjoint is a highpass filter with frequency response* $\|\boldsymbol{\omega}\|^{2n-1}$ *such that*

   $$\mathscr{R}^{(n)*}\mathscr{R}^{(n)}\{f\}(\boldsymbol{x}) = 2\pi\left(-\triangle\right)^{\frac{2n-1}{2}}\{f\}(\boldsymbol{x}). \tag{4.12}$$

   Proof. This follows from $(\frac{\partial}{\partial y})^* = -\frac{\partial}{\partial y}$ and Proposition 2.1. □

This shows that $\mathscr{R}^{(n)^*}\mathscr{R}^{(n)}$ has a Fourier transform that is proportional to $\|\boldsymbol{\omega}\|^{2n-1}$ while $\mathbf{L}^T\mathbf{L}$ is a discretized Laplace operator whose continuous-domain frequency response is $\|\boldsymbol{\omega}\|^2$. Thus, the preconditioner $\mathbf{M}^{-1}$ that we use in the discrete domain in the case of identity weight $\mathbf{D}$ is the discrete filter that approximates the frequency response $\frac{1}{\|\boldsymbol{\omega}\|^{2n-1}+\mu\|\boldsymbol{\omega}\|^2+\lambda_1}$.

2. In the case of a non statistical formulation of the cost function, we propose to use a weighting matrix that is the discrete counterpart of the convolution operator q in (2.32), with a slight modification of the frequency-domain singularity at zero. We modify it with the frequency response $\frac{1}{|\omega|^{2n-1}+\beta}$, where β is an appropriate positive parameter; it is a positive-definite operator.

The solution of the minimization of $\mathscr{L}_\mu(\mathbf{c}^{k+1},\mathbf{u},\boldsymbol{\alpha}^k)$ with respect to $\mathbf{u}$ is the proximal map linked to the regularization Ψ_2,

$$\text{prox}_{\Psi_2,\lambda,\mathscr{P}_\mathscr{C}}(\mathbf{z}) = \underset{\mathbf{c}\in\mathscr{C}}{\text{argmin}}\left\{\frac{1}{2}\|\mathbf{z}-\mathbf{v}\|_2^2+\lambda\Psi_2(\mathbf{v})\right\}, \tag{4.13}$$

where $\mathbf{z}=\mathbf{L}\mathbf{c}^{k+1}+\frac{\boldsymbol{\alpha}^k}{\mu}$ and $\lambda=\lambda_2/\mu_2$.

The complete reconstruction method is summarized in Algorithm 2 below. Note that normally the initial estimate of a CG procedure would be zero. Here we use a warm initialization in the sense that the starting point of each inner PCG iteration is the outcome of the previous PCG iteration.

B) **Constrained regularized weighted-norm reconstruction (CRWN):** We solve the nonlinear regularized problem by defining another variable splitting $\mathbf{u}=\mathbf{c}$ and using an augmented-Lagrangian (AL) scheme. This in turn is equivalent to finding critical point of the augmented Lagrangian (AL)

$$\begin{aligned}\mathscr{L}_\mu(\mathbf{c},\mathbf{u},\boldsymbol{\alpha}) =& \frac{1}{2}\|\mathbf{H}\mathbf{u}-\mathbf{g}\|_\mathbf{D}^2+\lambda_1\Psi_1(\mathbf{u})+\lambda_2\Psi_2(\mathbf{c})\\ &+\boldsymbol{\alpha}^T(\mathbf{u}-\mathbf{c})+\frac{\mu}{2}\|\mathbf{u}-\mathbf{c}\|_2^2,\end{aligned} \tag{4.14}$$

where $\boldsymbol{\alpha}$ is the vector of Lagrange multipliers that imposes the constraint $\mathbf{u}=\mathbf{c}$. The classical AL scheme alternates between a joint minimization step and an update step, so that

$$\begin{cases}(\mathbf{c}^{k+1},\mathbf{u}^{k+1}) \leftarrow \underset{\mathbf{c}\in\mathscr{C},\mathbf{u}}{\text{argmin}}\,\mathscr{L}_\mu(\mathbf{c},\mathbf{u},\boldsymbol{\alpha}^k)\\ \boldsymbol{\alpha}^{k+1}\leftarrow\boldsymbol{\alpha}^k+\mu(\mathbf{u}^{k+1}-\mathbf{c}^{k+1}).\end{cases} \tag{4.15}$$

Input: phase measurements $g(y_j, \theta_i)$, $\forall i, j$.
Output: reconstructed image $f(\boldsymbol{x})$.
initialization λ_1, λ_2, μ, $\mathbf{c}^{(\mathbf{0})}$, $\mathbf{u}^{(\mathbf{0})}$, and B-spline degree m;
while *stopping criterion is not satisfied* **do**
  $\mathbf{c}^{k+1} \leftarrow \underset{\mathbf{c}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^k, \boldsymbol{\alpha}^k)$, using the preconditioned CG method with an initial estimate $\mathbf{c}^{(0)}$;
  $\mathbf{u}^{k+1} \leftarrow \operatorname{prox}_{\Psi_2, \lambda, \mathscr{P}_\mathscr{C}}\left(\mathbf{L}\mathbf{c}^{k+1} + \frac{\boldsymbol{\alpha}^k}{\mu}\right)$;
  $\boldsymbol{\alpha}^{k+1} \leftarrow \boldsymbol{\alpha}^k + \mu(\mathbf{L}\mathbf{c}^{k+1} - \mathbf{u}^{k+1})$;
  $\mathbf{c}^{(\mathbf{0})} \leftarrow \mathbf{c}^{\mathbf{k+1}}$;
  $k \leftarrow k+1$;
end
return $f(\boldsymbol{x}) = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}} \beta^m(\boldsymbol{x} - \boldsymbol{k})$.

Algorithm 2: ADMM-PCG WITH WARM INITIALIZATION RECONSTRUCTION METHOD

Moreover, we use ADMM [93] to separate the joint minimization into the succession of simpler partial problems

$$\begin{cases} \mathbf{u}^{k+1} \leftarrow \underset{\mathbf{u}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}^k, \mathbf{u}, \boldsymbol{\alpha}^k) & \text{(Step 1)} \\ \mathbf{c}^{k+1} \leftarrow \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}, \mathbf{u}^{k+1}, \boldsymbol{\alpha}^k) & \text{(Step 2)} \\ \boldsymbol{\alpha}^{k+1} \leftarrow \boldsymbol{\alpha}^k + \mu(\mathbf{u}^{k+1} - \mathbf{c}^{k+1}). & \text{(Step 3)} \end{cases} \tag{4.16}$$

Since the zero frequency is in the nullspace of the forward operator, we use the Tikhonov regularization term $\Psi_1(\mathbf{u}) = 1/2\,\|\mathbf{u}\|^2$.

In Step 1, $\mathbf{c}^k$ and $\boldsymbol{\alpha}^k$ are fixed, therefore $\mathscr{L}_\mu(\mathbf{c}^k, \mathbf{u}, \boldsymbol{\alpha}^k)$ is a quadratic function of $\mathbf{u}$ with gradient

$$\begin{aligned} \boldsymbol{\nabla}\mathscr{L}_\mu(\mathbf{c}^k, \mathbf{u}, \boldsymbol{\alpha}^k) = &\left(\mathbf{H}^T\mathbf{D}\mathbf{H} + (\mu + \lambda_1)\mathbf{I}\right)\mathbf{u} \\ &- \left(\mathbf{H}^T\mathbf{D}\mathbf{g} - \left(\boldsymbol{\alpha}^k - \mu\mathbf{c}^k\right)\right). \end{aligned} \tag{4.17}$$

We use the CG method to solve this step. One can choose the weight $\mathbf{D}$ such that the

condition number of the matrix $\mathbf{H}^T\mathbf{W}\mathbf{H}+(\mu+\lambda_1)\mathbf{I}$ becomes quite small, Then the corresponding iterative algorithm converges rapidly.

Step 2 of ADMM, which minimizes $\mathscr{L}_\mu(\mathbf{c},\mathbf{u}^k,\boldsymbol{\alpha}^k)$ with respect to $\mathbf{c}$, is the constrained denoising problem

$$\begin{aligned}\underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}}\ \{\mathscr{L}_\mu(\mathbf{c},\mathbf{u}^{k+1},\boldsymbol{\alpha}^k) &= {\boldsymbol{\alpha}^k}^T(\mathbf{u}^{k+1}-\mathbf{c})+\frac{\mu}{2}\left\|\mathbf{u}^{k+1}-\mathbf{c}\right\|_2^2+\lambda_2\psi_2(\mathbf{c})\}\\ &= \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}}\left\{\frac{1}{2}\left\|\mathbf{u}^{k+1}+\frac{\boldsymbol{\alpha}^k}{\mu}-\mathbf{c}\right\|_2^2+\frac{\lambda_2}{\mu}\Psi_2(\mathbf{c})\right\}. \end{aligned} \tag{4.18}$$

The common expression for the regularizer is

$$\Psi_2(\mathbf{c}) = \|\mathbf{R}\mathbf{c}\|\,, \tag{4.19}$$

where $\|\cdot\|$ is a non-quadratic norm and $\mathbf{R}:\mathbb{R}^N\to\mathbb{R}^{(NK)}$ is the regularization operator (e.g., gradient with $K=2$ or Hessian with $K=2\times 2$). For the identity regularization operator $\mathbf{R}=\mathbf{I}$, (4.18) typically admits a direct threshold-based solution.

For the general case of the regularization operator, we aim at solving the denoising problem. This is equivalent to the proximal map

$$\operatorname{prox}_{\|\cdot\|,\lambda,\mathscr{P}_\mathscr{C}}(\mathbf{z}) = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}}\left\{\frac{1}{2}\|\mathbf{z}-\mathbf{c}\|_2^2+\lambda\|\mathbf{R}\mathbf{c}\|\right\}, \tag{4.20}$$

where $\mathscr{P}_\mathscr{C}$ is the convex projection that corresponds to the constraint. In order to find the solution of (4.20), we use the Fenchel duality to rewrite the regularization term as

$$\|\mathbf{R}\mathbf{c}\| = \max_{\mathbf{p}\in\mathscr{B}}\left\langle\mathbf{R}^T\mathbf{p},\boldsymbol{u}\right\rangle, \tag{4.21}$$

where $\mathbf{R}^T:\mathbb{R}^{(NK)}\to\mathbb{R}^N$ is the adjoint of the operator $\mathbf{R}$, $\mathbf{p}\in\mathbb{R}^{(NK)}$, and $\mathscr{B}=\{\mathbf{p}\in\mathbb{R}^{(NK)}|\,\|\mathbf{p}\|_*\le 1\}$ with $\|\cdot\|_*$ the dual norm.

It can be shown that the solution of (4.20) is $\mathscr{P}_\mathscr{C}(\mathbf{z}-\lambda\mathbf{R}^T\mathbf{p}^*)$, where

$$\mathbf{p}^* = \underset{\mathbf{p}}{\operatorname{argmin}}\, f(\mathbf{p})+\mathbf{1}_\mathscr{B}, \tag{4.22}$$

with $\nabla f(\mathbf{p}) = -\lambda\mathbf{R}\mathscr{P}_\mathscr{C}(\mathbf{z}-\lambda\mathbf{R}^T\mathbf{p})$. We apply the fast iterative-shrinkage-thresholding algorithm (FISTA) [87] to solve (4.22). The step size is constrained by the Lipschitz constant

Input: $\mathbf{z}$, λ, $\tau \leq L^{-1}$, $\mathscr{P}_{\mathscr{B}}$, $\mathscr{P}_{\mathscr{C}}$
Output: $\mathbf{c}$ (optimal solution of (4.32))
initialization $\mathbf{p}_0$, $t_1 = 1$;
while *stopping criterion is not satisfied* **do**
  $\mathbf{p}_k \leftarrow \mathscr{P}_{\mathscr{B}}(\mathbf{y}_k + \tau\lambda\mathbf{R}\mathscr{P}_{\mathscr{C}}(\mathbf{z} - \lambda\mathbf{R}^T\mathbf{p}_k))$;
  $t_{k+1} \leftarrow \frac{1+\sqrt{1+4t_k^2}}{2}$;
  $y_{k+1} \leftarrow \mathbf{p}_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{p}_k - \mathbf{p}_{k-1})$;
  $k \leftarrow k + 1$;
end
return $\mathbf{c} = \mathscr{P}_{\mathscr{C}}(\mathbf{z} - \lambda\mathbf{R}^T\mathbf{p})$.

Algorithm 3: DENOISING ALGORITHM

L of $\nabla f(\mathbf{p})$ that depends on the regularization operator $\mathbf{R}$. The other important component is the orthogonal projection onto the set $\mathscr{B}$ that is specified by the chosen norm. Let us denote it by $\mathscr{P}_{\mathscr{B}}$. Algorithm 3 describes the denoising algorithm.

The benefits of the proposed splitting are: 1) the transformation of a complex reconstruction problem into a sequence of simpler optimizations where the constraint is applied as a simple projection in each iteration of the denoising step. Note that there is no simple way to impose the convex constraint to the linear step of the ADMM-PCG method; 2) any regularization term can be handled by knowing its corresponding denoising function; 3) the output of the algorithm is the solution of the denoising step, which results in an improved quality of reconstruction. The reconstruction method is summarized in Algorithm 4. Here, the starting point of each inner CG iteration is the outcome of the previous CG iteration called as warm initialization.

4.2.2 Generalization of the proposed reconstruction scheme

In some applications such as fan-beam and cone-beam imaging modalities, we do not have access to the whole projection along each direction because of the limited field of view of the imaging system. For this reason, we express the forward operator as **MH** where **M** is a suitable mask. Thus, the general form of the estimator is

Input: $\mathbf{g}$, $\mathbf{H}$, λ_1, λ_2, $\mathscr{P}_{\mathscr{C}}$
Output: ($f(\boldsymbol{x})$ reconstructed image)
set λ_1, λ_2, μ, and B-spline degree m;
initialization $\mathbf{c}^0$, $\mathbf{u}^0$ and $\boldsymbol{\alpha}^0$;
while *stopping criterion is not satisfied* **do**
  $\mathbf{u}^{k+1} \leftarrow \underset{\mathbf{u}}{\operatorname{argmin}}\, \mathscr{L}_\mu(\mathbf{c}^k, \mathbf{u}, \boldsymbol{\alpha}^k)$, using CG method with initial estimate $\mathbf{u}^k$ ("warm initialization");
  $\mathbf{c}^{k+1} \leftarrow \operatorname{prox}_{\|\cdot\|, \frac{\lambda_2}{\mu}, \mathscr{P}_{\mathscr{C}}}(\mathbf{u}^{k+1} + \boldsymbol{\alpha}^k/\mu)$;
  $\boldsymbol{\alpha}^{k+1} \leftarrow \boldsymbol{\alpha}^k + \mu(\mathbf{u}^{k+1} - \mathbf{c}^{k+1})$;
  $k \leftarrow k+1$;
end
return $f(\boldsymbol{x}) = \sum_{\boldsymbol{k}} c_{\boldsymbol{k}} \beta^m(\boldsymbol{x} - \boldsymbol{k})$

Algorithm 4: CONSTRAINED REGULARIZED RECONSTRUCTION WITH WEIGHTED NORM (CRWN).

$$\widehat{\mathbf{c}} = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} \left\{ \underbrace{\frac{1}{2}\|\mathbf{MHc} - \mathbf{g}\|_{\mathbf{D}}^2 + \lambda_1 \Psi_1(\mathbf{c}) + \lambda_2 \Psi_2(\mathbf{Lc})}_{\mathscr{J}(\mathbf{c})} \right\}, \tag{4.23}$$

To solve (4.23), we extend ADMM-PCG by defining the new auxiliary variables $\mathbf{u} = \mathbf{Hc}$ and $\mathbf{v} = \mathbf{Lc}$ and reformulate the minimization (4.23) as the linear equality-constrained problem

$$\begin{aligned} \mathbf{c} = \underset{\mathbf{c},\mathbf{u},\mathbf{v}}{\operatorname{argmin}} & \left\{ \frac{1}{2}\|\mathbf{Mu} - \mathbf{g}\|_{\mathbf{D}}^2 + \lambda_1 \Psi_1(\mathbf{c}) + \lambda_2 \Psi_2(\mathbf{v}) \right\} \\ & \text{subject to } \mathbf{u} = \mathbf{Hc}, \mathbf{v} = \mathbf{Lc}. \end{aligned} \tag{4.24}$$

When the matrix $\mathbf{D}$ is a circulant matrix and the mask matrix $\mathbf{M}$ does not exist, there is no need to define the auxiliary variable $\mathbf{u}$.

We then map (4.24) to an unconstrained problem by considering the augmented- La-

grangian functional

$$\begin{aligned}\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u},\mathbf{v},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2) = &\frac{1}{2}\|\mathbf{M}\mathbf{u}-\mathbf{g}\|_{\mathbf{D}}^2+\lambda_1\Psi_1(\mathbf{c})+\lambda_2\Psi_2(\mathbf{v})\\ &+\boldsymbol{\alpha}_1^T(\mathbf{u}-\mathbf{H}\mathbf{c})+\frac{\mu_1}{2}\|\mathbf{u}-\mathbf{H}\mathbf{c}\|_2^2\\ &+\boldsymbol{\alpha}_2^T(\mathbf{v}-\mathbf{L}\mathbf{c})+\frac{\mu_2}{2}\|\mathbf{v}-\mathbf{L}\mathbf{c}\|_2^2 \end{aligned} \tag{4.25}$$

where $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are the vectors of Lagrange multipliers that impose the constraints $\mathbf{u}=\mathbf{H}\mathbf{c}$ and $\mathbf{v}=\mathbf{L}\mathbf{c}$. The advantage of using an augmented Lagrangian instead of the Lagrangian (including the quadratic term) is to bring robustness to the dual-ascent method for updating the Lagrangian multipliers and, in particular, to relax the assumption o strict convexity on the main cost function.

The classical method of multipliers alternates between a joint minimization step and update steps, so that

$$\begin{cases} (\mathbf{c}^{k+1},\mathbf{u}^{k+1},\mathbf{v}^{k+1}) \leftarrow \underset{\mathbf{c}\in\mathscr{C},\mathbf{u},\mathbf{v}}{\operatorname{argmin}}\ \mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u},\mathbf{v},\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k)\\ {\boldsymbol{\alpha}_1}^{k+1} \leftarrow {\boldsymbol{\alpha}_1}^k+\mu_1(\mathbf{u}^{k+1}-\mathbf{H}\mathbf{c}^{k+1})\\ {\boldsymbol{\alpha}_2}^{k+1} \leftarrow {\boldsymbol{\alpha}_2}^k+\mu_2(\mathbf{v}^{k+1}-\mathbf{L}\mathbf{c}^{k+1}). \end{cases} \tag{4.26}$$

The first step is a joint minimization which is costly to solve. In this regard, we adopt a cyclic update scheme also known as ADMM, which consists of the iterations

$$\begin{cases} \mathbf{c}^{k+1} \leftarrow \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}}\ \mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k) & \text{(Step 1)}\\ \mathbf{u}^{k+1} \leftarrow \underset{\mathbf{u}}{\operatorname{argmin}}\ \mathscr{L}_{\mu_1,\mu_2}(\mathbf{c}^{k+1},\mathbf{u},\mathbf{v},\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k) & \text{(Step 2)}\\ \mathbf{v}^{k+1} \leftarrow \underset{\mathbf{v}}{\operatorname{argmin}}\ \mathscr{L}_{\mu_1,\mu_2}(\mathbf{c}^{k+1},\mathbf{u}^{k+1},\mathbf{v},\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k) & \text{(Step 3)}\\ {\boldsymbol{\alpha}_1}^{k+1} \leftarrow {\boldsymbol{\alpha}_1}^k+\mu_1(\mathbf{u}^{k+1}-\mathbf{H}\mathbf{c}^{k+1}) & \text{(Step 4)}\\ {\boldsymbol{\alpha}_2}^{k+1} \leftarrow {\boldsymbol{\alpha}_2}^k+\mu_2(\mathbf{v}^{k+1}-\mathbf{L}\mathbf{c}^{k+1}) & \text{(Step 5).} \end{cases} \tag{4.27}$$

The solutions of the steps are as follows:

1. The criterion $\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k)$ is a quadratic cost function with respect to $\mathbf{c}$ whose gradient is

$$\nabla\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k) = \underbrace{\left(\mu_1\mathbf{H}^T\mathbf{H}+\mu_2\mathbf{L}^T\mathbf{L}+\lambda_1\nabla\Psi_1\right)}_{\mathbf{A}}\mathbf{c} - \underbrace{\left(\mu_1\mathbf{H}^T\left(\mathbf{u}-\frac{\boldsymbol{\alpha}_1}{\mu_1}\right)+\mu_2\mathbf{L}^T\left(\mathbf{v}^k-\frac{\boldsymbol{\alpha}^k}{\mu_2}\right)\right)}_{\mathbf{b}}. \quad (4.28)$$

   We minimize $\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k)$ iteratively using the conjugate-gradient algorithm to solve for $\mathbf{Ac}=\mathbf{b}$. For speeding up the convergence of CG algorithm, we introduce two different strategies:

2. The second step is the minimization with respect to $\mathbf{u}$ while the other variables are fixed. This is in the form of

$$\mathbf{u}^{k+1} = \underset{\mathbf{u}}{\operatorname{argmin}}\left\{\frac{1}{2}\|\mathbf{Mu}-\mathbf{g}\|_{\mathbf{D}}^2+\boldsymbol{\alpha}_1^T(\mathbf{u}-\mathbf{Hc})+\frac{\mu_1}{2}\|\mathbf{u}-\mathbf{Hc}\|_2^2\right\}. \quad (4.29)$$

   Its gradient with respect to $\mathbf{u}$ is given by

$$\nabla\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c}^{k+1},\mathbf{u},\mathbf{v},\boldsymbol{\alpha}_1^k,\boldsymbol{\alpha}_2^k) = \left(\mathbf{M}^T\mathbf{DM}+\mu_1\mathbf{I}\right)\mathbf{u} - \left(\mathbf{M}^T\mathbf{Dg}+\mu_1\left(\mathbf{Hc}^{k+1}+\frac{{\boldsymbol{\alpha}_1}^k}{\mu_1}\right)\right).$$

   In order to retrieve the critical point, one requires to find the zeros of the gradient function. Since $\mathbf{M}$ is a mask matrix and $\mathbf{D}$ is diagonal, the solution is a point-wise operator that takes the form

$$\mathbf{u}^{k+1}[n] = \frac{1}{\mu_1+\{\mathbf{D}\}_{n,n}\{\mathbf{M}\}_{n,n}}\left\{\mathbf{M}^T\mathbf{Dg}+\mu_1\left(\mathbf{Hc}^{k+1}+\frac{{\boldsymbol{\alpha}_1}^k}{\mu_1}\right)\right\}_{n,n}. \quad (4.30)$$

3. Step 3, the minimization with respect to $\mathbf{v}$, is

$$\mathbf{v}^{k+1} = \underset{\mathbf{v}}{\operatorname{argmin}}\left\{\boldsymbol{\alpha}_2^T(\mathbf{v}-\mathbf{Lc}^{k+1})+\frac{\mu_2}{2}\left\|\mathbf{v}-\mathbf{Lc}^{k+1}\right\|_2^2+\lambda_2\Psi_2(\mathbf{c})\right\}. \quad (4.31)$$

This is the denoising step. The solution is the proximal map linked to the regularization Ψ_2,

$$\text{prox}_{\Psi_2,\lambda,\mathscr{P}_{\mathscr{C}}}(\mathbf{z}) = \underset{\mathbf{c}\in\mathscr{C}}{\text{argmin}} \left\{ \frac{1}{2} \|\mathbf{z}-\mathbf{v}\|_2^2 + \lambda \Psi_2(\mathbf{v}) \right\}, \tag{4.32}$$

where $\mathbf{z} = \left(\mathbf{L}\mathbf{c}^{k+1} - \alpha_2/\mu_2\right)$ and $\lambda = \lambda_2/\mu_2$.

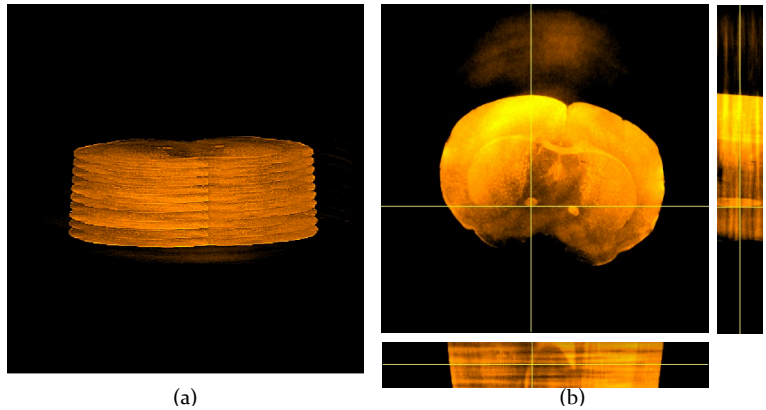
4. The variables α_1 and α_2 are updated in Steps 4 and 5.

Among all steps, the most computationally costly is the second. It involves the computation of $\mathbf{H}^T\mathbf{H}$ and $\mathbf{H}^T\mathbf{g}$ (see (4.28)). The value of $\mathbf{H}^T\mathbf{g}$ is only computed once and is used in all iterations, while $\mathbf{H}^T\mathbf{H}$ of $\mathbf{v}$ is computed in each iteration. Accordingly, in order to speed up the proposed algorithm, we require to derive a fast implementation of $\mathbf{H}^T\mathbf{H}$.

4.3 Memory efficient and fast 3D reconstruction in parallel-beam tomography

Typically to discover the 3D object in parallel-beam imaging modalities with single axis tilting (the object is rotated along a fixed axis), its 2D slices are reconstructed separately. In practice, there are two crucial drawbacks to this technique. The aliasing effects along the z direction and grating interferometer drifting during the imaging process can introduce artifacts in the reconstructed image, for instance the horizontal stripes on the vertical coronal section of the sample shown in Figure 4.1.

One can also formulate the 3-D reconstruction as an inverse problem equipped with three-dimensional total-variation regularization. The main drawback is the size of the 3-D specimen which is often extremely large. Therefore, the 3-D forward operator requires a drastically large memory space which is completely inefficient. To avoid this state of affairs, we reformulate the problem as a combination of simpler optimization problems in lower dimensions. Concerning the discretization framework, the forward imaging operator



(a) (b)

Figure 4.1: (a) In parallel-beam geometry, 3-D reconstruction is decomposed into several 2-D slice reconstruction. (b) Horizontal artifacts appear on the vertical coronal section of the sample.

with x_3 as rotation axis is

$$\begin{aligned}\mathbf{g}[i,j,k] &= [\mathbf{H}_{3D}\mathbf{c}]_{i,j,k_3} \\ &= \sum_{\mathbf{k}} c[\mathbf{k}]\mathscr{R}^{(n)}\left\{\beta(x_1)\beta(x_2)\right\}(y_j - k_1\cos\theta_i - k_2\sin\theta_i)\beta(k-k_3) \\ &= \sum_{k_1,k_2}\left(\sum_{k_3} c[k_1,k_2,k_3]\beta(k-k_3)\right)\mathscr{R}^{(n)}\left\{\beta(x_1)\beta(x_2)\right\}(y_j - k_1\cos\theta_i - k_2\sin\theta_i)\,.\end{aligned} \tag{4.33}$$

Thus, the forward imaging operator is written in the form of

$$\begin{aligned}\mathbf{g}[i,j,k] &= [\mathbf{H}_{3D}\mathbf{c}]_{i,j,k} \\ &= \left[\mathbf{H}_{2D}\left\{\mathbf{u}[\cdot,\cdot,k]\right\}\right]_{i,j}\,,\end{aligned} \tag{4.34}$$

where

$$\left[\mathbf{H}_{2D}\left\{\mathbf{u}[\cdot,\cdot,k_3]\right\}\right]_{i,j} = \sum_{k_1,k_2}\mathbf{u}[k_1,k_2,k_3]\mathscr{R}^{(n)}\left\{\beta(x_1)\beta(x_2)\right\}(y_j - k_1\cos\theta_i - k_2\sin\theta_i)\,, \tag{4.35}$$

and

$$\mathbf{u}[k_1,k_2,k_3] = (\mathbf{c}[k_1,k_2,\cdot] * \beta[\cdot])[k_1,k_2,k_3], \tag{4.36}$$

where $*$ denotes the discrete convolution operator, $\beta[\cdot]$ is the B-spline function sampled at the integer values. The convolution operator can be written in the matrix form

$$\mathbf{u}[k] = [\mathbf{B}\{\mathbf{c}[k_1,k_2,\cdot]\}]_k, \forall k_1,k_2, \tag{4.37}$$

where $\mathbf{B}$ is the interpolating matrix.

Therefore, the reconstruction problem is expressed as

$$\underset{\mathbf{c}}{\operatorname{argmin}} \left\{ \underbrace{\frac{1}{2}\|\mathbf{H}_{3D}\mathbf{c}-\mathbf{g}\|^2_{\mathbf{D}} + \lambda_1\Psi_1(\mathbf{c}) + \lambda_2\Psi_2(\mathbf{c})}_{J(\mathbf{c})} \right\}. \tag{4.38}$$

We separate the regularization on each 2D horizontal plane from the regularization along the rotation axis. We introduce total variation along with Tikhonov regularization as prior information. Then, the minimization (4.38) can be rewritten in the form of minimizing

$$J(\mathbf{c}) = \frac{1}{2}\sum_k \|\mathbf{H}_{2D}\{\{\mathbf{B}\mathbf{c}\}[\cdot,\cdot,k]\} - \mathbf{g}[\cdot,\cdot,k]\|^2_D + \frac{\lambda_1}{2}\|\mathbf{c}\|^2 + \lambda_2 \sum_k \|\mathbf{L}_{2D}\{\mathbf{c}[\cdot,\cdot,k]\}\|_1 + \lambda_3 \sum_{k_1,k_2} \|L_{x_3}\{\mathbf{c}[k_1,k_2,\cdot]\}\|_1, \tag{4.39}$$

where $\mathbf{L}_{2D}$ and $\mathbf{L}_{x_3}$ are the gradient operators on the horizontal plane and along the vertical coordinate, respectively. These operators are computed consistently with the splined-based discretization framework.

In order to decompose the three-dimensional reconstruction (4.39) into a set of optimization problems in lower dimensions, first, we use variable-splitting techniques by introducing

$$\mathbf{u}_1 = \mathbf{c}$$

$$\mathbf{u}_2[k_1,k_2,\cdot] = \mathbf{B}\{\mathbf{u}_1[k_1,k_2,\cdot]\}, \forall k_1,k_2 \tag{4.40}$$

$$\mathbf{u}_3[k_1,k_2,\cdot] = \mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}, \forall k_1,k_2. \tag{4.41}$$

These auxiliary variables yield the constrained optimization problem

$$\underset{\mathbf{c},\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3}{\operatorname{argmin}} \{\frac{1}{2}\sum_k \{\|\mathbf{H}_{2D}\{\mathbf{u}_2[\cdot,\cdot,k]\}-\mathbf{g}[\cdot,\cdot,k]\|_D^2+\lambda_2\|\mathbf{L}_{2D}\{\mathbf{c}[\cdot,\cdot,k]\}\|_1\} + \frac{\lambda_1}{2}\|\mathbf{u}_1\|^2+\lambda_3\sum_{k_1,k_2}\|\mathbf{u}_3[k_1,k_2,\cdot]\|_1\}$$
$$\text{subject to } \mathbf{u}_1=\mathbf{c}\,,\mathbf{u}_2[\cdot,\cdot,k]=\{\mathbf{B}\mathbf{u}_1\}[\cdot,\cdot,k] \text{ and } \mathbf{u}_3[k_1,k_2,\cdot]=\mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}\,,\forall k_1,k_2\,,\forall k\,. \tag{4.42}$$

The augmented-Lagrangian function of the constrained optimization problem is

$$\begin{aligned}\mathscr{L}_{\mu_1,\mu_2,\mu_3}(\mathbf{c},\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\mathbf{d}_1,\mathbf{d}_2,\mathbf{d}_3)=&\\ &\frac{1}{2}\sum_k\left\{\|\mathbf{H}_{2D}\{\mathbf{u}_2[\cdot,\cdot,k]\}-\mathbf{g}[\cdot,\cdot,k]\|_D^2+\lambda_2\|\mathbf{L}_{2D}\{\mathbf{c}[\cdot,\cdot,k]\}\|_1\right\}+\\ &\lambda_3\sum_{k_1,k_2}\|\mathbf{u}_3[k_1,k_2,\cdot]\|_1+\frac{\mu_1}{2}\|\mathbf{u}_1-\mathbf{c}+\mathbf{d}_1\|+\frac{\lambda_1}{2}\|\mathbf{u}_1\|^2+\\ &\frac{\mu_2}{2}\sum_{k_1,k_2}\|\mathbf{u}_2[k_1,k_2,\cdot]-\mathbf{B}\{\mathbf{u}_1[k_1,k_2,\cdot]\}+\mathbf{d}_2[k_1,k_2,\cdot]\|^2+\\ &\frac{\mu_3}{2}\sum_{k_1,k_2}\|\mathbf{u}_3[k_1,k_2,\cdot]-\mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}+\mathbf{d}_3[k_1,k_2,\cdot]\|^2\,,\end{aligned} \tag{4.43}$$

where μ_1, μ_2, and μ_3 are penalty parameters of the augmented Lagrangian and where $\mathbf{d}_1$, $\mathbf{d}_2$ and $\mathbf{d}_3$ are the augmented-Lagrangian variables.

Since (4.42) is strictly convex, the joint minimization of (4.43) converges to the unique solution of the constrained optimization problem (4.42). To solve it, we separate the prob-

lem into basic optimizations by taking advantage of the ADMM

$$\begin{cases} \mathbf{u}_2^{k+1} \leftarrow \underset{\mathbf{u}_2}{\operatorname{argmin}}\, \mathscr{L}_{\mu_1,\mu_2,\mu_3}(\mathbf{c}^k, \mathbf{u}_1^k, \mathbf{u}_2, \mathbf{u}_3^k, \mathbf{d}_1^k, \mathbf{d}_2^k, \mathbf{d}_3^k) \\ \mathbf{u}_1^{k+1} \leftarrow \underset{\mathbf{u}_1}{\operatorname{argmin}}\, \mathscr{L}_{\mu_1,\mu_2,\mu_3}(\mathbf{c}^k, \mathbf{u}_1, \mathbf{u}_2^{k+1}, \mathbf{u}_3^k, \mathbf{d}_1^k, \mathbf{d}_2^k, \mathbf{d}_3^k) \\ \mathbf{u}_3^{k+1} \leftarrow \underset{\mathbf{u}_3}{\operatorname{argmin}}\, \mathscr{L}_{\mu_1,\mu_2,\mu_3}(\mathbf{c}^k, \mathbf{u}_1^{k+1}, \mathbf{u}_2^{k+1}, \mathbf{u}_3, \mathbf{d}_1^k, \mathbf{d}_2^k, \mathbf{d}_3^k) \\ \mathbf{c}^{k+1} \leftarrow \underset{\mathbf{c}}{\operatorname{argmin}}\, \mathscr{L}_{\mu_1,\mu_2,\mu_3}(\mathbf{c}, \mathbf{u}_1^{k+1}, \mathbf{u}_2^{k+1}, \mathbf{u}_3^{k+1}, \mathbf{d}_1^k, \mathbf{d}_2^k, \mathbf{d}_3^k) \\ \mathbf{d}_1^{k+1} \leftarrow \mathbf{d}_1^k + \mathbf{u}_1^{k+1} - \mathbf{c}^{k+1} \\ \mathbf{d}_2^{k+1}[k_1,k_2,\cdot] \leftarrow \mathbf{d}_2^k[k_1,k_2,\cdot] + \mathbf{u}_2^k[k_1,k_2,\cdot] - \mathbf{B}\left\{\mathbf{u}_1^k[k_1,k_2,\cdot]\right\}, \forall k_1,k_2 \\ \mathbf{d}_3^{k+1}[k_1,k_2,\cdot] \leftarrow \mathbf{d}_3[k_1,k_2,\cdot] + \mathbf{u}_3[k_1,k_2,\cdot] - \mathbf{L}_{x_3}\left\{\mathbf{u}_1[k_1,k_2,\cdot]\right\}, \forall k_1,k_2\,. \end{cases}$$

For simplicity in notation, we denote the concatenation of two-dimensional matrices $\mathbf{B}\{\mathbf{u}_1[\cdot,\cdot,k]\}, \forall k$ by the three-dimensional matrix $\mathbf{B}\mathbf{u}_1$. As first step, the minimization with respect to $\mathbf{u}_2$,

$$\underset{\mathbf{u}_2}{\operatorname{argmin}} \sum_k \left\{ \frac{1}{2}\|\mathbf{H}_{2D}\{\mathbf{u}_2[\cdot,\cdot,k]\} - \mathbf{g}[\cdot,\cdot,k]\|_D^2 + \frac{\mu_2}{2}\|\mathbf{u}_2[\cdot,\cdot,k] - \mathbf{B}\mathbf{u}_1[\cdot,\cdot,k] + \mathbf{d}_2[\cdot,\cdot,k]\|^2 \right\}, \tag{4.44}$$

can be parallelized into a set of two-dimensional quadratic problems. We use CG to minimize each quadratic function

$$J_{1,k}(\mathbf{u}_{2,k}) = \frac{1}{2}\|\mathbf{H}_{2D}\left\{\mathbf{u}_{2,k}\right\} - \mathbf{g}[\cdot,\cdot,k]\|_D^2 + \frac{\mu_2}{2}\|\mathbf{u}_{2,k} - \{\mathbf{B}\mathbf{u}_1\}_k + \mathbf{d}_{2,k}\|^2, \tag{4.45}$$

where $\mathbf{u}_{2,k} = \mathbf{u}_2[\cdot,\cdot,k]$ and $\{\mathbf{B}\mathbf{u}_1\}_k = \mathbf{B}\mathbf{u}_1[\cdot,\cdot,k]$. Its gradient with respect to $\mathbf{u}_{2,k}$ is

$$\begin{aligned} \nabla J_{1,k}(\mathbf{u}_{2,k}) = &\underbrace{(\mathbf{H}_{2D}^T\mathbf{D}_k\mathbf{H}_{2D} + \mu_2\mathbf{I})\mathbf{u}_{2,k}}_{\mathbf{A}} - \\ &\underbrace{(\mathbf{H}_{2D}^T\mathbf{D}_k\{\mathbf{g}[\cdot,\cdot,k]\} + \mu_2(\{\mathbf{B}\mathbf{u}_1\}_k + \mathbf{d}_2[\cdot,\cdot,k]))}_{\mathbf{b}}. \end{aligned}$$

We can choose the matrix $\mathbf{D}_k$ such that the condition number of $\mathbf{A}$ becomes close to one, and then it can be expected that the corresponding iterative scheme converges reasonably fast.

As second step, the minimization with respect to $\mathbf{u}_1$ is a set of 1D optimization problems for each pair (k_1,k_2)

$$\begin{aligned}\underset{\mathbf{u}_1[k_1,k_2,\cdot]}{\operatorname{argmin}}\{&\frac{\lambda_1}{2}\|\mathbf{u}_1[k_1,k_2,\cdot]\|^2+\frac{\mu_1}{2}\|\mathbf{u}_1[k_1,k_2,\cdot]-\mathbf{c}[k_1,k_2,\cdot]+\mathbf{d}_1[k_1,k_2,\cdot]\|+\\&\frac{\mu_2}{2}\sum_{k_1,k_2}\|\mathbf{u}_2[k_1,k_2,\cdot]-\mathbf{B}\{\mathbf{u}_1[k_1,k_2,\cdot]\}+\mathbf{d}_2[k_1,k_2,\cdot]\|^2+\\&\frac{\mu_3}{2}\sum_{k_1,k_2}\|\mathbf{u}_3[k_1,k_2,\cdot]-\mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}+\mathbf{d}_3[k_1,k_2,\cdot]\|^2\},\end{aligned}\tag{4.46}$$

whose direct solution is

$$\begin{aligned}\mathbf{u}_1[k_1,k_2,.]=&((\mu_1+\lambda_1)\mathbf{I}+\mu_2\mathbf{B}^T\mathbf{B}+\mu_3\mathbf{L}_{x_3}^T\mathbf{L}_{x_3})^{-1}\\&(\mu_1(\mathbf{c}[k_1,k_2,\cdot]-\mathbf{d}_1[k_1,k_2,\cdot])+\mu_2\mathbf{B}^T(\mathbf{u}_2[k_1,k_2,\cdot]+\mathbf{d}_2[k_1,k_2,\cdot])+\\&\mu_3\mathbf{L}_{x_3}^T(\mathbf{u}_3[k_1,k_2,\cdot]+\mathbf{d}_3[k_1,k_2,\cdot])).\end{aligned}\tag{4.47}$$

This can be implemented using recursive filters since the kernel is symmetric. Moreover, the inversion can also be performed in the FFT domain with a point-wise division.

As third step, the minimization with respect to $\mathbf{u}_3$,

$$\underset{\mathbf{u}_3[k_1,k_2,\cdot]}{\operatorname{argmin}}\left\{\lambda_3\|\mathbf{u}_3[k_1,k_2,\cdot]\|_1+\frac{\mu_3}{2}\sum_{k_1,k_2}\|\mathbf{u}_3[k_1,k_2,\cdot]-\mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}+\mathbf{d}_3[k_1,k_2,\cdot]\|^2\right\},\tag{4.48}$$

is the point-wise soft-thresholding operator

$$\mathbf{u}_3[k_1,k_2,\cdot]=\operatorname{prox}_{\lambda_2/\mu_3}\left(\mathbf{L}_{x_3}\{\mathbf{u}_1[k_1,k_2,\cdot]\}-\mathbf{d}_3[k_1,k_2,\cdot]\right).\tag{4.49}$$

In Step 4, the minimization with respect to $\mathbf{c}$ is a set of two-diminutional total-variation denoising problems. To solve the nonlinear TV problem, we develop a modified version of the gradient-based fast iterative-shrinkage-thresholding algorithm [98], which requires the repeated evaluation of the proximal map of the non-smooth part $J_2(\mathbf{c})=\sum_{\boldsymbol{k}}\|\{\mathbf{L}\mathbf{c}\}_{\boldsymbol{k}}\|_1$. The last steps are the updates of the Lagrangian variables.

Chapter 5

FFT-cost implementation of $\mathbf{H}^T\mathbf{H}$

In chapter 4, we presented unified reconstruction frameworks for straight-ray tomography using the alternating direction method of multipliers. Their main advantage is that one can obtain a reasonable solution in a few number of iterations. There is a strong incentive to reduce the computational cost of each iteration to speed up the reconstruction procedure even further. The dominating computational cost is the matrix-vector multiplications $\mathbf{H}^T\mathbf{g}$ where $\mathbf{g}$ is the measurement vector and the multiplication of the matrix $\mathbf{H}^T\mathbf{H}$ with an updated vector. The vector $\mathbf{H}^T\mathbf{g}$ is precomputed once and then is used in each iteration. Then, in this chapter, we propose fast and efficient implementation of the matrix multiplication of $\mathbf{H}^T\mathbf{H}$ with a vector. This way of speeding-up the reconstruction procedure is well-known in magnetic resonance imaging and has been widely used [99, 100], while in the context of computed tomography, it has not much been developed [101]. In this regard, we first derive the necessary conditions to make it equivalent to a digital convolution operator. We then present the corresponding kernel, so that one can apply this operator at minimal cost (in the order of FFT).

5.1 Notations

Our formulation uses infinite matrices and vectors. Obviously, in practice these are truncated to finite length. Matrices and vectors are denoted by bold letters. The entries of a matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_{\mathbf{p},\mathbf{k}}$ where $\mathbf{p}$ and $\mathbf{k}$ specify the position of the entry of interest.

The $\mathbf{k}$-th entry of a vector $\mathbf{c}$ is denoted by $c[\mathbf{k}]$ or $[\mathbf{c}]_{\mathbf{k}}$.

Thus, the matrix formulation of (3.5) is

$$g[\mathbf{p}] = \sum_{\mathbf{k}} [\mathbf{H}]_{\mathbf{p},\mathbf{k}} c[\mathbf{k}]\,, \tag{5.1}$$

where $\mathbf{p} = (i, j)$ for simplicity in notation.

5.2 Computation of $\mathbf{H}^T\mathbf{H}$

5.2.1 Review

As discussed in Chapter 3, the classical discretization approach considers basically pixel values of the object as its discrete representation, and the projection is approximated by a discrete line integral over each pixel. In the case of differential variants of the x-ray transform, the derivative operators are implemented with finite-differences. The other way to reduce the complexity of the implementation and improve its speed is to rely on Fourier-based techniques which are applicable in a parallel geometry [102, 103]. The Fourier slice theorem relates the Fourier transform of the projection to the Fourier transform of the object in polar coordinates. In the direct Fourier implementation, the polar frequency data are interpolated on a Cartesian grid. Afterwards, the inverse of FFT is applied. To achieve acceptable reconstruction quality, the interpolation step relies on sampling. When the number of projection views is not large enough, the interpolation steps perform poorly, which distributes a significant error in the reconstruction.

An alternative technique is the pseudo polar Fourier transform [104–106]. Its drawback is that it requires the number of orientations to be in the order of the image size. Moreover, all these approaches are valid in the parallel-beam geometry only. In the context of divergent beams, measurements are interpolated to fill the whole sinogram which is typically performed using rebinning techniques. Since the Fourier transform is a non-local operator, the interpolation error is distributed in the space domain representation of the reconstructed object.

5.2.2 Generalized sampling based implementation

Based on the discretization scheme described in Chapter 3, the computation of $\mathbf{H}^T\mathbf{H}\mathbf{c}$ can be decomposed as the computation over each projection angle

$$\mathbf{H}^T\mathbf{H}\mathbf{c} = \sum_i \mathbf{H}_{\theta_i}^T \mathbf{H}_{\theta_i}\mathbf{c} \tag{5.2}$$

where $\mathbf{H}_\theta$ denotes the system matrix corresponding to the orientation θ. We start with two dimensional images and then extend the approach to higher dimensions. Let $\mathbf{g}_\theta$ be the projection vector along the direction $\boldsymbol{\theta}$. Then, the backprojection of $\mathbf{g}_\theta$ is in the form of

$$[\mathbf{H}_\theta^T \mathbf{g}_\theta]_{\mathbf{k}} = \sum_j g_\theta[j]\, \mathscr{P}\{\varphi_h\}\left(j - h\left\langle \mathbf{k}, \boldsymbol{\theta}\right\rangle\right). \tag{5.3}$$

Lemma 5.1. *The entries of the vector* $\tilde{\mathbf{c}}_\theta = \mathbf{H}_\theta^T\mathbf{H}_\theta\mathbf{c}$ *are given by*

$$\tilde{c}_\theta[\mathbf{l}] = \sum_{\mathbf{k}} c[\mathbf{k}]\mathbf{R}_\theta[\mathbf{k},\mathbf{l}], \tag{5.4}$$

where $\mathbf{R}_\theta = \mathbf{H}_\theta^T\mathbf{H}_\theta$ *is the bi-infinite normal matrix whose entries are*

$$\mathbf{R}_\theta[\mathbf{k},\mathbf{l}] = \sum_j \mathscr{P}\{\varphi_h\}\left(j - h\left\langle \mathbf{k}, \boldsymbol{\theta}\right\rangle, \boldsymbol{\theta}\right)\mathscr{P}\{\varphi_h\}\left(j - h\left\langle \mathbf{l}, \boldsymbol{\theta}\right\rangle, \boldsymbol{\theta}\right). \tag{5.5}$$

The computational cost of the calculation of $\tilde{c}$,

$$\tilde{c} = \sum_{i=1}^{K} \tilde{c}_{\theta_i} = \sum_{i=1}^{K}\sum_{\mathbf{k}} c[\mathbf{k}]\mathbf{R}_{\theta_i}[\mathbf{k},\mathbf{l}], \tag{5.6}$$

is on the order of $\left(N^2 \times a \times h \times K\right)$, where the number of entries of $\mathbf{c}$ is $(N \times N)$, a is the support of φ, and K is the number of orientations. Thus, the number of multipliers is directly proportional to the number of orientations, the scale ratio h, the support of the basis function, and the number of entries of $\mathbf{c}$.

5.3 FFT-cost implementation of $\mathbf{H}^T\mathbf{H}$ in parallel beam geometry

In this section, we derive the necessary conditions to make $\mathbf{H}^T\mathbf{H}$ translation invariant so that we can speed up its computational cost with the help of FFT.

Definition 5.1 (Radial Nyquist criterion)**.** *The function φ satisfies the radial Nyquist criterion with respect to the grid $\mathbb{Z}^d$ if $\widehat{\varphi}(\omega) = 0$ for all $\|\omega\| \geq \pi$.*

Proposition 5.1. *For any pair of function (f,g) satisfying the Nyquist condition, it holds that*

$$\sum_{\mathbf{n}\in\mathbb{Z}^2} f(\mathbf{n})g(\mathbf{n}) = \int_{\mathbb{R}^2} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}, \tag{5.7}$$

where $\mathbf{n} = (n_1,n_2)$ and $\mathbf{x} = (x_1,x_2)$.

Proof. Since the functions f and g satisfy the Nyquist condition, we can apply Shannon's theorem and expand them using the sinc functions.

$$\begin{aligned} f(\mathbf{x}) &= \sum_{\mathbf{k}} f(\mathbf{k})\mathrm{sinc}(\mathbf{x}-\mathbf{k}), \\ g(\mathbf{x}) &= \sum_{\mathbf{k}} g(\mathbf{k})\mathrm{sinc}(\mathbf{x}-\mathbf{k}), \end{aligned} \tag{5.8}$$

where $\mathrm{sinc}(\mathbf{x}) = \mathrm{sinc}(x_1)\mathrm{sinc}(x_2)$. The orthonormality of the sinc function and its shifts yields the desired result. □

Theorem 5.1. *if φ satisfies the radial Nyquist criterion for all $h \geq h_0$, then for all $h \geq h_0$*

1. *$\mathbf{H}_\theta^T\mathbf{H}_\theta$ is a discrete convolution matrix with $[\mathbf{H}_\theta^T\mathbf{H}_\theta]_{\mathbf{k},\mathbf{l}} = r_\theta[\mathbf{k}-\mathbf{l}]$ where*

$$\begin{aligned} r_\theta[\mathbf{k}] &= (\mathscr{P}_\theta\varphi_h(\cdot) * \mathscr{P}_\theta\varphi_h(-\cdot))(h\langle\mathbf{k},\theta\rangle) \\ &= \mathscr{P}_\theta\left\{\varphi_h(\cdot) * \varphi_h(-\cdot)\right\}(h\langle\mathbf{k},\theta\rangle) \end{aligned} \tag{5.9}$$

2. *As $\mathbf{H}^T\mathbf{H} = \sum_i \mathbf{H}_{\theta_i}^T\mathbf{H}_{\theta_i}$, $\mathbf{H}^T\mathbf{H}$ is a discrete convolution whose impulse response is*

$$r[\mathbf{k}] = \sum_i r_{\theta_i}[\mathbf{k}] \tag{5.10}$$

Proof. Starting from (5.5),

$$\begin{aligned} \mathbf{R}_\theta[\mathbf{k},\mathbf{l}] &= \sum_j \mathscr{P}_\theta\left\{\varphi_h\right\}(j-h\langle\mathbf{k},\theta\rangle)\mathscr{P}_\theta\left\{\varphi_h\right\}(j-h\langle\mathbf{l},\theta\rangle) \\ &\overset{(1)}{=} \int \mathscr{P}_\theta\left\{\varphi_h\right\}(y-h\langle\mathbf{k},\theta\rangle)\mathscr{P}_\theta\left\{\varphi_h\right\}(y-h\langle\mathbf{l},\theta\rangle)\mathrm{d}y \\ &\overset{(2)}{=} \int \mathscr{P}_\theta\left\{\varphi_h\right\}(y-h\langle\mathbf{l}-\mathbf{k},\theta\rangle)\mathscr{P}_\theta\left\{\varphi_h\right\}(y)\mathrm{d}y \\ &= r_\theta[\mathbf{l}-\mathbf{k}]. \end{aligned} \tag{5.11}$$

As φ_h satisfies the radial Nyquist criterion, the Fourier slice theorem implies that $\mathscr{P}_\theta\{\varphi_h\}$ satisfies the Nyquist condition. Then, Proposition 5.1 yields the equality (1), and while (2) is the result of the change of variable $\bar{y} = y - h\langle \mathbf{k}, \boldsymbol{\theta}\rangle$ and $\mathrm{d}\bar{y} = \mathrm{d}y$. □

This theorem shows that $\mathbf{H}_\theta^T\mathbf{H}_\theta$ is a circulant matrix that can be implemented using FFT operator. It is conveniently summarized by its impulse response

$$\begin{aligned} r_\theta[\mathbf{k}] &= \int \mathscr{P}_\theta\{\varphi_h\}(y - h\langle \mathbf{k}, \boldsymbol{\theta}\rangle)\mathscr{P}_\theta\{\varphi_h\}(y)\mathrm{d}y \\ &= (\mathscr{P}_\theta\{\varphi_h\}(\cdot) * \mathscr{P}_\theta\{\varphi_h\}(-\cdot))(h\langle \mathbf{k}, \boldsymbol{\theta}\rangle), \end{aligned} \tag{5.12}$$

which is obtained by resampling the autocorrelation of the continuous-domain function $\mathscr{P}_\theta\{\varphi_h\}$ which is the x-ray transform of φ_h. In practice, the number of entries of $\mathbf{c}$ is limited, typically $(N \times N)$, but the support of the introduced kernel is the whole space. In order to use the FFT-cost implementation of $\mathbf{H}^T\mathbf{H}$, first the kernel is computed using (5.29) in a window with size $((2N-1)\times(2N-1))$ as summarized in Algorithm 5. In the cases that is not possible to analytically compute the autocorrelation function, one can numerically calculate (interpolate) its value in the sampling points. The matrix-vector multiplication $\mathbf{H}^T\mathbf{H}\mathbf{c}$ can be computed by convolving the kernel and $\mathbf{c}$. This can be done using a FFT operator. For more details, we refer the reader to Algorithm 6. Accordingly, the computational cost is on the order of $(2N)^2\log(2N)$. The cost does not depend on the number of orientations, scale ratio, and the support of the basis function φ. Note that in iterative reconstruction scheme, the kernel $\mathbf{r}$ is computed once and then is being used in each iteration.

5.3.1 Error of approximation

The operator $\mathbf{H}^T\mathbf{H}$ can be computed at a computational cost comparable to one FFT along with inverse of FFT operations when the support of the generating function φ in the Fourier domain is bounded to the closed circle around the origin with radius π. This introduces a significant improvement in computational performance. For example, if the image size is (1024×1024), and 512 different orientations exist with $h = 2$ and $a = 4$, then

$$\text{ratio of speed improvement} = \frac{1024^2 \times 512 \times 4 \times 2}{(2\times1024)^2 \times \log 2048} \approx 310\,. \tag{5.13}$$

Input: φ, $\Theta = \{\theta_1, \theta_2, \ldots, \theta_K\}$ (set of all orientations)
Output: $\mathbf{r}$ (the kernel)
initialization $\mathbf{r} = 0$ and $i = 1$;
while $i \leq$ *the number of orientations* **do**
  $\mathbf{r}_{\theta_i}[\mathbf{k}] = \left(\mathscr{P}_{\theta_i}\{\varphi_h\}(\cdot) * \mathscr{P}_{\theta_i}\{\varphi_h\}(-\cdot)\right)(h\langle \mathbf{k}, \boldsymbol{\theta}_i\rangle)$ where
  $\mathbf{k} = \{-(N-1), \ldots, (N-1)\} \times \{-(N-1), \ldots, (N-1)\}$;
  $\mathbf{r} \leftarrow \mathbf{r} + \mathbf{r}_{\theta_i}$;
  $i \leftarrow i + 1$;
end
return $\widehat{\mathbf{r}} = \mathscr{F}\{\mathbf{r}\}$, discrete Fourier transform of the kernel $\mathbf{r}$.

Algorithm 5: COMPUTATION OF THE KERNEL CORRESPONDING TO $\mathbf{H}^T\mathbf{H}$

Input: $\mathbf{c}$, $\widehat{\mathbf{r}}$ (discrete Fourier transform of the kernel $\mathbf{r}$)
Output: $\tilde{\mathbf{c}} = \mathbf{H}^T\mathbf{H}\mathbf{c}$
zero-pad $\mathbf{c}$ and extend it to size $(2N-1 \times 2N-1)$;
$\tilde{\mathbf{c}} = iFFT(FFT(\mathbf{c}) * \widehat{\mathbf{r}})$;
return cropped version of $\tilde{\mathbf{c}}$ with the same size as the input $\mathbf{c}$.

Algorithm 6: FFT-COST IMPLEMENTATION OF $\mathbf{H}^T\mathbf{H}\mathbf{c}$

The generating functions that are typically used are compactly supported in the spatial domain and consequently are not bandlimited; in other words, they only satisfy the condition in Theorem 5.1 approximately).

Definition 5.2. *We define a π-energy concentration measure for the basis function φ as*

$$E_\pi(\varphi) = \frac{\int_{\|\boldsymbol{\omega}\| \geq \pi} |\widehat{\varphi}_h(\boldsymbol{\omega})|^2 \mathrm{d}\boldsymbol{\omega}}{\int_{\mathbb{R}^2} |\widehat{\varphi}_h(\boldsymbol{\omega})|^2 \mathrm{d}\boldsymbol{\omega}}\,. \tag{5.14}$$

A basis function φ satisfies the necessary conditions in Theorem 5.1 if and only if $E_\pi(\varphi) = 0$. The measure $E_\pi(\varphi)$ shows how suitable a given basis function φ is for being used in Theorem 5.1 (lower values are more preferable). In other words, this measure computes the relative part of the energy of a given basis function that is outside the Nyquist region.

In Chapter 3, box splines (particularly B-splines) and Kaiser-Bessel windows were presented as two favorable families of basis functions. The Fourier transform of polynomial

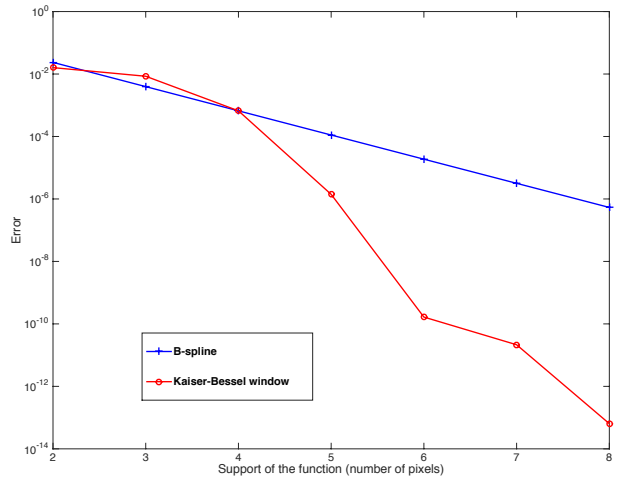


Figure 5.1: The values of the error measure (5.14) for different basis function supports.

B-splines with degree n is in the form of

$$\widehat{\beta^{(n)}}(\omega_1,\omega_2)=\left(\frac{\sin(\omega_1/2)}{\omega_1/2}\right)^{n+1}\left(\frac{\sin(\omega_2/2)}{\omega_2/2}\right)^{n+1}, \tag{5.15}$$

and the Fourier transform of Kaiser-Bessel window functions with parameters a, α, and m is given by

$$\widehat{\varphi}_{\mathrm{KB}}(\omega_1,\omega_2)=\begin{cases}\dfrac{(2\pi)^{n/2}a^n\alpha^m I_{n/2+m}\left(\sqrt{\alpha^2-(a\|\boldsymbol{\omega}\|)^2}\right)}{I_m(\alpha)\left(\sqrt{\alpha^2-(a\|\boldsymbol{\omega}\|)^2}\right)^{n/2+m}} & a\|\boldsymbol{\omega}\|\le\alpha\\ \dfrac{(2\pi)^{n/2}a^n\alpha^m J_{n/2+m}\left(\sqrt{(a\|\boldsymbol{\omega}\|)^2-\alpha^2}\right)}{I_m(\alpha)\left(\sqrt{(a\|\boldsymbol{\omega}\|)^2-\alpha^2}\right)^{n/2+m}} & a\|\boldsymbol{\omega}\|\ge\alpha\,.\end{cases} \tag{5.16}$$

The values of the measure (5.14) for B-splines and Kaiser-Bessel windows with different supports are depicted in Figure 5.1. Note that the parameters of Kaiser-Bessel functions are computed by minimizing the residual error (3.53). The parameters are reported

in Table 5.1. It suggests that it is more desirable to use Kaiser-Bessel windows with large supports than B-splines.

a and m: support and smoothness parameters	2	3	4	5	6	7	8
α: window taper	7.95	17.25	19.63	16.36	14.89	19.67	17.69

Table 5.1: Optimal taper parameter of Kaiser-Bessel window functions with different supports.

For simplicity in the notation, we denote $\mathbf{H}_\theta^T\mathbf{H}_\theta$ by $\mathbf{R}_\theta$ and the convolution kernel by $\mathbf{r}_\theta$, respectively. Then the error between the entries of these two kernels is

$$e[\mathbf{l},\mathbf{k}] = \mathbf{R}_\theta[\mathbf{l},\mathbf{k}] - \mathbf{r}_\theta[\mathbf{l}-\mathbf{k}]\,. \tag{5.17}$$

Proposition 5.2. *The approximation constant between the kernel $\mathbf{r}_\theta$ and the exact calculation of $\mathbf{H}_\theta^T\mathbf{H}_\theta$ is given by*

$$C_{kernel} = \|e\|_\infty = \max_{\mathbf{k},\mathbf{l}} |\mathbf{R}_\theta[\mathbf{k},\mathbf{l}] - \mathbf{r}_\theta[\mathbf{l}-\mathbf{k}]| \leq C \int_{\|\boldsymbol{\omega}\|\geq\pi} |\widehat{\varphi}(\boldsymbol{\omega}\, h)| d\boldsymbol{\omega}\,, \tag{5.18}$$

which is zero when the used basis function satisfies the radial Nyquist criteria.

Proof. Equation (5.33) can be written in the form of

$$\begin{aligned} C_{\text{kernel}} =& \max_{\mathbf{k},\mathbf{l}} |\mathbf{R}_\theta[\mathbf{k},\mathbf{l}] - \mathbf{r}_\theta[\mathbf{l}-\mathbf{k}]| \\ =& \max_{\mathbf{k},\mathbf{l}} |\sum_j \mathscr{P}\{\varphi_h\}(j - h\langle\mathbf{k},\boldsymbol{\theta}\rangle)\mathscr{P}\{\varphi_h\}(j - h\langle\mathbf{l},\boldsymbol{\theta}\rangle) - \\ & (\mathscr{P}_\theta\{\varphi_h\}(\cdot) * \mathscr{P}_\theta\{\varphi_h\}(-\cdot))(h\langle\mathbf{l}-\mathbf{k},\boldsymbol{\theta}\rangle)| \\ =& \max_{\mathbf{k},\mathbf{l}} |\sum_{n\in\mathbb{Z}} s_k(n)s_l(n) - \int_{\mathbb{R}} s_k(x)s_l(x)\mathrm{d}x|\,, \end{aligned} \tag{5.19}$$

where $s_k(x) = \mathscr{P}\{\varphi_h\}(x - h\langle \mathbf{k}, \boldsymbol{\theta}\rangle)$. So we have

$$
\begin{aligned}
|\sum_{n\in\mathbb{Z}} s_k(n)s_l(n) - \int_{\mathbb{R}} s_k(x)s_l(x)\mathrm{d}x| = & \\
& |\int_{-\pi}^{\pi} \sum_k \widehat{s_k}(\omega + 2\pi k) \sum_p \widehat{s_l}(-\omega + 2\pi p)\mathrm{d}\omega - \int_{\mathbb{R}} \widehat{s_k}(\omega)\widehat{s_l}(-\omega)\mathrm{d}\omega| \\
& = |\int_{\mathbb{R}} \widehat{s_k}(\omega) \sum_p \widehat{s_l}(-\omega + 2\pi p)\mathrm{d}\omega - \int_{\mathbb{R}} \widehat{s_k}(\omega)\widehat{s_l}(-\omega)\mathrm{d}\omega| \\
& = |\int_{\mathbb{R}} \widehat{s_k}(\omega) \left(\sum_p \widehat{s_l}(-\omega + 2\pi p) - \widehat{s_l}(-\omega) \right) \mathrm{d}\omega| \\
& \leq C|\int_{\pi}^{\infty} \widehat{s_0}(\omega)\mathrm{d}\omega| \\
& \leq C\int_{\pi}^{\infty} |\widehat{s_0}(\omega)|\mathrm{d}\omega \\
& \leq C\int_{\|\boldsymbol{\omega}\|\geq\pi} |\widehat{\varphi}(\boldsymbol{\omega} h)|\mathrm{d}\boldsymbol{\omega}\,. \qquad (5.20)
\end{aligned}
$$

It yields the desired result. □

The derived FFT-cost implementation of $\mathbf{H}^T\mathbf{H}$ is going to be used in the proposed reconstruction frameworks. Interestingly, ADMM converges to the fixed point of the problem even if each subproblems are solved approximately under some conditions specified in [107]. It implies that, the proposed iterative reconstruction framework using the derived kernel can converge to the same solution as using the exact formulation.

5.4 General cases

5.4.1 Differential variants of x-ray projection

The derivatives of the x-ray transform are also used to describe the mathematical model of some imaging modalities. For instance, the mathematical model of differential phase contrast tomography is based on the first derivative of the x-ray projection.

Let us denote the derivatives of the projection by

$$\mathscr{P}^{(n)}\{f\}(y, \boldsymbol{\theta}) = \frac{\partial^n}{\partial y^n} \mathscr{P}^{(n)}\{f\}(y, \boldsymbol{\theta})\,. \qquad (5.21)$$

We recall (3.4) for derivations of the projection as

$$\mathscr{P}^{(n)}\{f\}(j\Delta y,\theta_i)=\sum_{\mathbf{k}}c[\mathbf{k}]\mathscr{P}^{(n)}\{\varphi_h\}(j\Delta y-h\langle\mathbf{k},\boldsymbol{\theta}_i\rangle). \tag{5.22}$$

Therefore $\left[\mathbf{H}^{(n)}\right]_{(i,j),\mathbf{k}}=\mathscr{P}^{(n)}\{\varphi_h\}(j\Delta y-\langle\mathbf{k},\boldsymbol{\theta}_i\rangle,\boldsymbol{\theta})$.

Corollary 5.1. ***Extension to differential variants of x-ray projection***

For every function $f(\mathbf{x})=\sum_{\mathbf{k}\in\mathbb{Z}^2}c[\mathbf{k}]\varphi(\mathbf{x}/h-\mathbf{k})\in L(\mathbb{R}^2)$, if φ satisfies the radial Nyquist criterion for all $h\geq h_0$, then for all $h\geq h_0$, $\tilde{\mathbf{c}}=\mathbf{H}^{(n)^T}\mathbf{H}^{(n)}\mathbf{c}$ can be computed using

$$\tilde{c}[\mathbf{k}]=(c[\cdot]*r[\cdot])[\mathbf{k}], \tag{5.23}$$

with $r[\mathbf{k}]=\sum_i r_{\theta_i}[\mathbf{k}]$ and

$$r_{\theta_i}[\mathbf{k}]=\left(\mathscr{P}^{(n)}_{\theta_i}\{\varphi_h\}(\cdot)*\mathscr{P}^{(n)}_{\theta_i}\{\varphi_h\}(-\cdot)\right)(h\langle\mathbf{k},\boldsymbol{\theta_i}\rangle). \tag{5.24}$$

Proof. Since the derivative operator does not modify the bandwidth of the input function, the function $\mathscr{P}^{(n)}_{\theta}\{\varphi_h\}$ satisfies the Nyquist condition. So the proof is the same as that of Theorem 5.1. □

5.4.2 Weighted norm and speed of convergence

The weighted norm formulation of the data fidelity term results in the computation of $\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{c}$ for deducing the gradient in each iteration. In order to speed up the convergence rate, typically $\mathbf{W}$ is a circulant matrix. It is the filter which is used in filtered back projection (FBP) method, for example, the frequency response of the filter used in conventional computed tomography (CT) is $\|\omega\|$ in the projection domain. Then, the digital filter which is used in the practical application of CT has the same frequency response as the continuous filter in one period as shown in Figure 5.2.

Proposition 5.3 (Filtered Back-Projection)**.**

$$\mathscr{P}^{(n)^*}\{(q*\mathscr{P}^{(n)}f(\cdot,\theta))(y)\}(\boldsymbol{x})=f(\boldsymbol{x}), \tag{5.25}$$

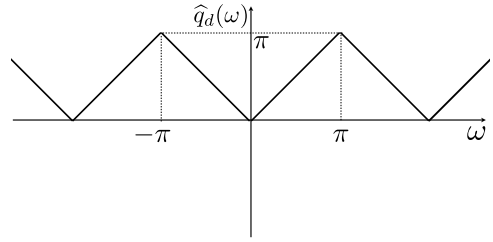


Figure 5.2: Frequency response of the digital filter used in the FBP method for conventional CT.

where $\mathscr{P}^{(n)*}$ is the adjoint of the n-th derivative of the x-ray projection and the transfer function of $q(y)$ is

$$\widehat{q}(\omega_y) = \frac{1}{2\pi} \times \frac{1}{|\omega_y|^{2n-1}}.$$

Then with the assumption $\tilde{\mathbf{c}} = \mathbf{H}_\theta^T \mathbf{W} \mathbf{H}_\theta \mathbf{c}$, we have

$$\tilde{c}_\theta[\mathbf{l}] = \sum_{\mathbf{k}} c[\mathbf{k}] r_\theta[\mathbf{k},\mathbf{l}], \tag{5.26}$$

where

$$r_\theta[\mathbf{k},\mathbf{l}] = \sum_j \mathscr{P}_\theta\{\varphi_h\}(j - h\langle \mathbf{k},\boldsymbol{\theta}\rangle)\left(q_{d,\theta}[\cdot] * \mathscr{P}\{\varphi_h\}(\cdot - h\langle \mathbf{l},\boldsymbol{\theta}\rangle)\right)[j]. \tag{5.27}$$

Theorem 5.2. ***Extension to*** $\mathbf{H}^T\mathbf{W}\mathbf{H}$

If φ satisfies the radial Nyquist criterion for all $h \geq h_0$ and $\mathbf{W}_\theta$ is a discrete convolution matrix that corresponds to the continuous kernel q, then $\mathbf{H}_\theta^T \mathbf{W}_\theta \mathbf{H}_\theta$ is a convolution matrix as well and

$$[\mathbf{H}_\theta^T \mathbf{W}_\theta \mathbf{H}_\theta]_{\mathbf{k},\mathbf{l}} = r_\theta[\mathbf{k}-\mathbf{l}], \tag{5.28}$$

where

$$r_\theta[\mathbf{k}] = \left(\mathscr{P}_\theta^{(n)}\{\varphi_h\}(\cdot) * q(\cdot) * \mathscr{P}_\theta^{(n)}\{\varphi_h\}(-\cdot)\right)(h\langle \mathbf{k},\boldsymbol{\theta}\rangle). \tag{5.29}$$

Proof. The generating function φ is band-limited, $\text{supp}\{\widehat{\varphi}\} \subset \Omega_h^2$. It implies that its x-ray transform is also band-limited so that it satisfies the Nyquist condition. Therefore

$$\left(q_{d,\theta}[\cdot] * \mathscr{P}_\theta \varphi[\cdot]\right)[j] = \left(q_\theta(\cdot) * \mathscr{P}_\theta \varphi(\cdot)\right)(j), \tag{5.30}$$

where $f[\cdot]$ defines the sequence of the sampled point values of the continuous function f. The rest of the proof is the same since $\left(q_\theta(\cdot) * \mathscr{P}_\theta \varphi(\cdot)\right)(y)$ is band-limited. □

5.4.3 Extension to higher dimension

Theorem 5.3. ***Extension to higher order dimension*** *For every function $f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c_\mathbf{k}\varphi(\mathbf{x}/h - \mathbf{k}) \in L(\mathbb{R}^d)$, if φ satisfies the radial Nyquist criterion for all $h \geq h_0$, then for all $h \geq h_0$, $\tilde{\mathbf{c}} = \mathbf{H}_\theta^T\mathbf{H}_\theta\mathbf{c}$ can be computed using*

$$\tilde{c}[\mathbf{k}] = (c[\cdot] * r[\cdot])[\mathbf{k}], \tag{5.31}$$

with

$$r[\mathbf{k}] = \left(\mathscr{P}_\theta\{\varphi_h\}(\cdot) * \mathscr{P}_\theta\{\varphi_h\}(-\cdot)\right)(h\mathbf{P}_{\boldsymbol{\theta}_\perp}\mathbf{k}). \tag{5.32}$$

The matrix $\mathbf{P}_{\boldsymbol{\theta}_\perp} \in \mathbb{R}^{d\times d-1}$ is the projection matrix described in (2.4)*.*

5.5 Experimental validation

5.5.1 One-by-one comparison

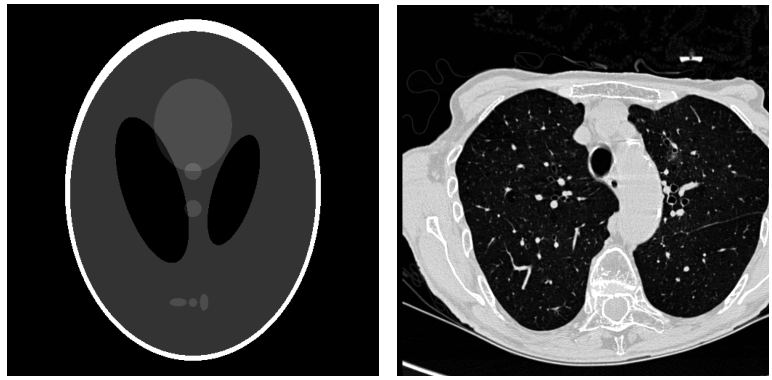
X-ray transform

In order to validate the digital convolution implementation of $\mathbf{H}^T\mathbf{H}$, we test it on two different objects as shown in Figure 5.3. The first one is a Shepp-Logan phantom and the next one is a coronal section of a human lung. We first compute $\mathbf{H}^T\mathbf{H}\mathbf{c}$ where $\mathbf{H}$ is the system matrix corresponding to conventional CT. Equation (5.5) is used for exact computation. Different scale parameters $h = 1, 2, 4, 8$ for 180 orientations which have been uniformly distributed between 0 and π are investigated. We then use the proposed digital filter implementation with the kernel suggested in (5.29). The signal-to-noise ratio between two different implementations and their computation times are given in Table 5.2 and 5.3, respectively. The results of two different implementations are depicted in Figures 5.4 (a) and 5.4 (b). Note that the speed of the implementation using the proposed digital kernel

is independent of the support of the given basis function, the number of orientations, and the scale ratio h. Its computational cost is proportional to $N^2 \log N$ where the object size is $(N \times N)$. On the other hand, the computational cost of the typical implementation is proportional to $\big($number of orientations $\times\, h \times N^2 \times$ support of the basis function$\big)$.

	$h=1$	$h=2$	$h=4$	$h=8$
t1 (second): kernel implementation (5.29)	.3	.08	.02	.005
t2 (second): exact implementation using (5.5)	93.03	39.8	23.8	11.8
t2/t1	310	497.5	1190	2360

Table 5.2: Time ratio between the kernel implementation in (5.29) and (5.5).



(a) (b)

Figure 5.3: (a) Shepp-Logan phantom with size (1024×1024). (b) A coronal section of a lung with size (751×751).

	$h=1$	$h=2$	$h=4$	$h=8$
phantom (cubic B-spline) (dB)	94.2	102	105	105
phantom (dB) (KBWF $a=2$, $m=2$, $\alpha=8$)	95	102.3	105	105
lung (cubic B-spline) (dB)	95.8	100	100	100
lung (dB) (KBWF $a=2$, $m=2$, $\alpha=8$)	96	100	100	100

Table 5.3: Signal-to-noise ratio between the kernel implementation in (5.29) and (5.5) in conventional CT.

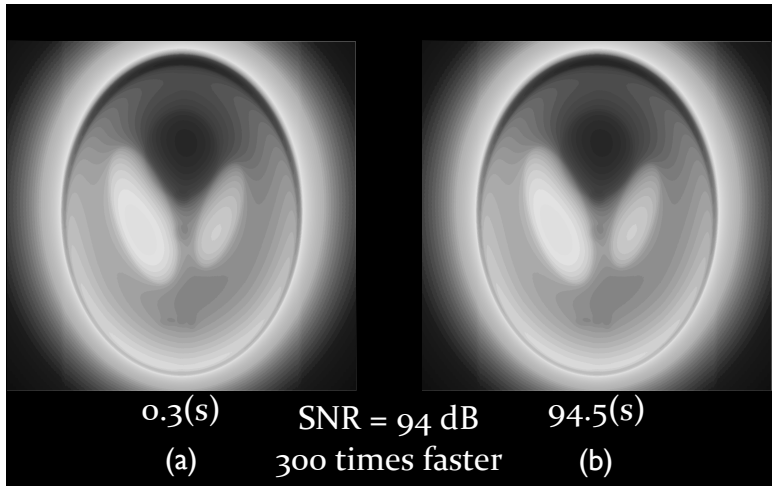


Figure 5.4: Comparison of the $\mathbf{H}^T\mathbf{H}$ using the kernel implementation and direct application

The first derivative of the x-ray transform

We then conduct the same experiment in the case of the first derivative of the Radon transform (mathematical model of differential phase-contrast imaging). The approximation

constant in the context of differential variants of the x-ray transform is given in the following proposition:

Proposition 5.4. *The approximation constant between the kernel* $\mathbf{r}_\theta^{(n)}$ *and the exact calculation of* ${\mathbf{H}^{(n)}}_\theta^T \mathbf{H}_\theta^{(n)}$ *is given by*

$$C_{kernel} = \|e\|_\infty = \max_{\mathbf{k},\mathbf{l}} |\mathbf{R}_\theta[\mathbf{k},\mathbf{l}] - \mathbf{r}_\theta[\mathbf{l}-\mathbf{k}]| \leq C \int_{\|\boldsymbol{\omega}\|\geq\pi} \|\boldsymbol{\omega}\|^{2n} |\widehat{\varphi}(\boldsymbol{\omega} h)| d\boldsymbol{\omega}\,, \tag{5.33}$$

which is zero when the basis function satisfies the radial Nyquist criteria.

This proposition implies that, in the context of differential phase contrast tomography, the approximation constant is linked to the relative energy outside the radial Nyquist criteria of the Laplacian of the basis function. The signal-to-noise ratio values are given in Table 5.4.

	$h=1$	$h=2$	$h=4$	$h=8$
phantom (cubic B-spline) (dB)	18.7	19.4	28.4	48.9
phantom (dB) (B-spline degree 5) (dB)	22.6	33.7	43.5	57.7
lung (cubic B-spline) (dB)	20	25	27	47
lung (dB) (B-spline degree 5)	20.8	32.4	38.3	58

Table 5.4: Signal-to-noise ratio between the kernel implementation in (5.29) and (5.5) in the case of the first derivative of the x-ray transform.

5.5.2 Performance evaluation of the kernel in different reconstruction frameworks

The main reason of the proposed kernel is to improve the speed of the reconstruction by proposing a convolution framework instead of the exact calculation of $\mathbf{H}_\theta^T \mathbf{H}$. In this part, we evaluate the performance of different reconstruction schemes in two different implementations. In order to evaluate their performance, we plot the cost function with respect

to the iteration number using the exact formulation and the kernel implementation. We first use conjugate gradient. The cost function values with respect to the iteration number is shown in Figure 5.5(a). Then, the performance of ADMM-PCG and CRWN techniques described in Chapter 3 are depicted in Figure 5.5(b) and Figure 5.5(c), respectively. It substatiates claim that, although the kernel implementation is not exact when we use cubic B-spline, it still yields a cost function that promotes convergence to the same value.

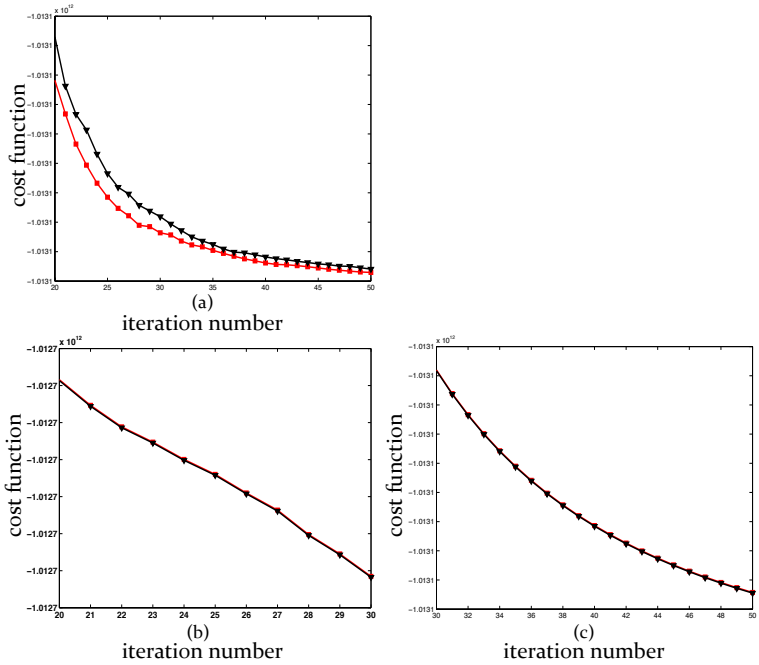


Figure 5.5: Cost function versus iteration number when either the exact formulation of $\mathbf{H}^T\mathbf{H}$ is used or the proposed kernel formulation for (a) conjugate gradient (b) ADMM-PCG (c) CRWN reconstruction frameworks.

Chapter 6

X-ray Grating Interferometry: potentially *in vivo* imaging modality

[1] In this chapter, we first briefly explain the physical setup of grating-based x-ray imaging (GI) and derive its physical model based on the wave equations. Our basic approach (ADMM-PCG) is developed in the context of GI to reconstruct the real part of the refractive index of the imaged sample. We finally evaluate the performance of our basic methodology with state-of-the-art techniques by conducting experiments on real data.

6.1 Motivation

Weitkamp *et al.* [13] and Momose *et al.* [19] developed a new X-ray imaging method based on grating interferometry (GI). It has attracted increasing interest in a variety of fields owing to its unique combination of imaging characteristics. First, GI provides a high sensitivity to electron-density variations, down to $0.18e/\mathrm{nm}^3$ [26]. This makes the technique particularly suitable for soft-tissue specimens. It has been applied successfully

[1] A part of this chapter has been presented in [36]

to biological samples such as insects [13, 26], rat-brain tissue, and even human as breast tissue [108]. Second, GI produces three complementary types of information; namely, attenuation, phase-shift, and dark-field measurements. In differential phase-contrast imaging (DPCI), one focuses exclusively on the phase information, which in principle allows one to reconstruct the real part of the refractive index distribution of the object. Third, GI does not require a highly monochromatic source, which means that conventional laboratory X-ray tubes can be used. The combination of the aforementioned characteristics makes GI suitable for a broad range of applications, such as material sciences (e.g., material testing), biomedical research (e.g., monitoring drug effects), or even clinical diagnostics (e.g., mammography).

6.2 Physical Model

An X-ray plane wave can be characterized by its intensity and phase. However the intensity is the only directly measurable part. Therefore, to extract the phase it is necessary to map this information into intensity patterns. In DPCI, this is achieved by using grating interferometers. The principle has been described in detail in [13,109,110], and we briefly review it here.

The physical setup of GI is depicted in Figure 6.1. It consists of an object on a rotation stage, a phase grating (G1), and an analyzer absorption grating (G2) which are behind the sample. Typically the phase grating produces a phase shift of π on the incident wave. After passing through the object, the X-ray reaches the phase grating which introduces a periodic phase modulation. It essentially splits the beam into its first two diffraction orders. A periodic interference pattern perpendicular to the optical axis is formed. To measure this pattern with high resolution, one then uses a phase stepping technique (PST) [13]. To that end, an absorption grating is placed in front of the detector with the same periodicity and orientation as the interference pattern. In this technique, the absorption grating is moved perpendicular to the optical axis and the intensity signal in each pixel in the detector plane is recorded as a function of the grating position x_g. The recorded intensity is called the phase stepping curve (PSC).

When an object is placed on the optical path, the illumination wave is attenuated and refracted. The refraction causes a local shift of the interference pattern. This displacement is given by

$$\triangle x_g(y,\theta) = d\gamma(y,\theta),$$

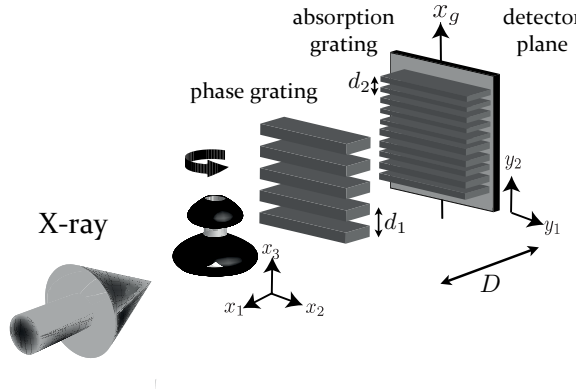


Figure 6.1: Grating-based X-ray imaging setup. The phase grating introduces a phase shift of π in the transmitted wave. The absorption grating is necessary for measuring the received wave given the limited resolution of the detector.

where d is the distance between the two gratings. The refraction angle γ is proportional to the derivative of the phase of the output wave with respect to y, like in

$$\gamma(y,\theta) = \frac{\lambda}{2\pi}\frac{\partial\phi(y,\theta)}{\partial y}\,, \tag{6.1}$$

where $\gamma(y,\theta)$ is the refraction angle in radians and λ is the wavelength. The phase shift induced by the object on the transmitted beam is given by

$$\phi(y,\theta) = \mathscr{P}\{\alpha\}(y,\theta)\,. \tag{6.2}$$

For a given photon energy E, α takes the form

$$\alpha = \frac{r_0\hbar^2c^2}{2\pi E^2}\rho_e\,, \tag{6.3}$$

with r_0 the classical electron radius, $\hbar$ the reduced Planck's constant, and c the speed of light. These relations yield

$$g_1(y,\theta) = \frac{1}{d}\triangle x_g(y,\theta) = \frac{\partial\mathscr{P}\{\alpha\}}{\partial y}(y,\theta)\,. \tag{6.4}$$

For our purpose, (6.4) characterizes the forward model of DPCI: α is the quantity of interest and g_1 is the physically measurable signal.

In addition to the introduced displacement on the PSC, the object causes the average and the amplitude of the PSC vary. The average of the PSC is linked to the average of the intensity on that pixel. This measurement corresponds to the conventional CT whose mathematical model is the x-ray transform of the imaginary part of the refractive index of the object,

$$g_2(y,\theta) = \mathscr{P}\{\beta\}(y,\theta)\,. \tag{6.5}$$

The β can be expressed as

$$\beta = \left(\overline{m}Z^{\overline{n}} + \overline{b}\right)\rho_e \tag{6.6}$$

where ρ_e is the electron density, Z the effective atomic number, and $\overline{m}$, $\overline{n}$, $\overline{b}$ parameters that depend only on the details of the experimental implementation and not on sample properties [111].

The variation of the amplitude is known as the visibility. Its formulation is given by

$$g_3(y,\theta) = \frac{a_{\max}(y,\theta) - a_{\min}(y,\theta)}{a_{\max}(y,\theta) + a_{\min}(y,\theta)}\,, \tag{6.7}$$

where $a_{\max}(y,\theta)$ and $a_{\min}(y,\theta)$ are the maximum and the minimum value of the PSC for the the index (y,θ). This is modeled by

$$g_3(y,\theta) = \mathscr{P}\{s\}(y,\theta)\,, \tag{6.8}$$

where s is a generalized scattering parameter that quantifies the local scatter strength [93].

6.3 Differential phase-contrast imaging

X-ray phase-contrast imaging (PCI) is a promising alternative to absorption-based computed tomography (CT) for visualizing many structures in biological samples, in particular soft-tissues. PCI is based on the phase shift induced by the propagation of a coherent wave through the investigated object. Various PCI methods have been developed including analyzer based [8–10], interferometric [11–13], and free space propagation methods [14–16]. These methods differ substantially in terms of the physical signal that is measured and the required experimental setup.

DPCI essentially yields the derivative of the Radon transform of the real part of the refractive-index map of the object. This map can therefore be retrieved using a suitable variant of the filtered back-projection (FBP) algorithm known from conventional tomography. However the FBP method requires a large number of viewing angles, and thus the acquisition time is very long. The long radiation time can also damage biological samples. This provides a strong motivation for developing algorithms that can handle fewer views, so as to reduce the exposure time.

6.3.1 Mathematical Consideration

Available descriptions of X-ray DPCI only provide a qualitative explanation of this interferometric system [13, 112, 113]. However we are not aware of a self-contained theoretical description of X-ray DPCI. We have thus developed a concise mathematical model that is solely based on wave optics. Such a characterization is necessary for quantitative imaging and for recasting the computerized reconstruction as a linear inverse problem. At the same time it clarifies which approximations are used to model the system and under which conditions imaging can be performed. Our main contribution is a concise physical derivation of the relationship between the displacement of the interference fringes and the derivative of the phase of the incoming wave. In addition we give several conditions that need to be fulfilled to make imaging possible. We rely exclusively on standard wave optics so as to make the presentation accessible. For simplicity we consider only one lateral dimension, but our description can easily be extended to 2 dimensions since the geometry of the interferometer is separable.

A simplified X-ray DPCI setup is represented in Figure 6.2 and the corresponding notations are summarized in Table 6.1.

Regarding the propagation of waves within the interferometer we make the following two assumptions:

Assumption 6.1 (Fresnel diffraction)**.** $(y-x)^2 \ll D^2$.

In addition we assume that the phase of the incoming wave varies slowly.

Assumption 6.2 (Taylor approximation)**.** *For $l \in \mathbb{Z}$ and $x \in [(l-1/2)d, (l+1/2)d]$,*

$$\phi(x) \simeq \phi(ld) + (x-ld)\phi'(ld)\,. \tag{6.9}$$

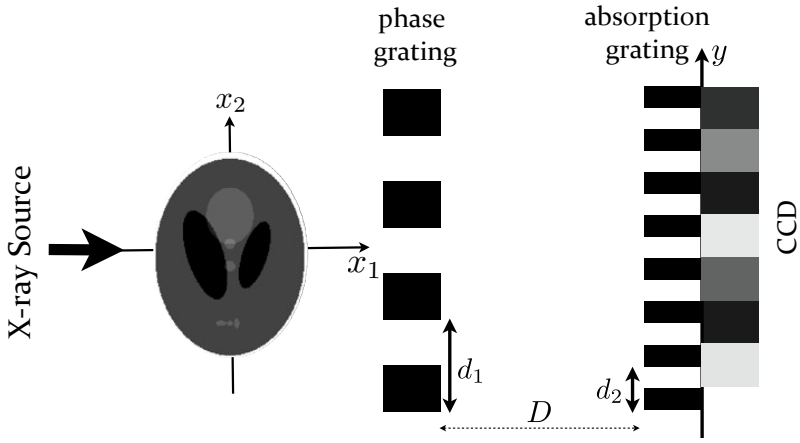


Figure 6.2: Setup of grating-based x-ray imaging in 2D.

In other words, we are approximating the phase with a piecewise-linear function. We can then show that, at the imaging plane, the diffracted wave is a sum of shifted versions of the Fourier transform of the grating profile (up to some phase factors).

Proposition 6.1. *The wave field at the imaging plane is proportional to*

$$\sum_{l\in\mathbb{Z}} \exp\left(i\phi(ld)\right) \exp\left(i\overline{\phi}(y-ld)\right) \overline{\psi}\left(y-ld-\lambda D\phi'(ld)/2\pi\right), \tag{6.10}$$

where

$$\begin{aligned} \overline{\phi}(y) &= \exp\left(i\pi\frac{y^2}{\lambda D}\right) \\ \overline{\psi}(y) &= \mathscr{F}\left\{\psi(x)\exp\left(i\pi\frac{x^2}{\lambda D}\right)\right\}\left(\frac{2\pi}{\lambda D}y\right). \end{aligned} \tag{6.11}$$

Most importantly, the position of these patterns depends linearly on the derivative of the phase.

Notation	Description	Typical value/Definition
x	Input-plane coordinate	
y	Imaging-plane coordinate	
z	Axial coordinate	
λ	Wavelength	$1.38(\text{Å})$
d_1	Phase-grating period	$4(\mu\text{m})$
D	Distance between input plane and imaging plane	$2.32(\text{cm})$
$\phi(x)$	Phase of the input-plane wave field	
$\psi(x)$	Phase-grating profile	$\psi(x) = \text{rect}(2x/d) - 2\text{rect}(x/d)$
$\mathscr{F}$	Fourier transform operator	$\mathscr{F}\{f\}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}\text{d}x$

Table 6.1: Description of notations.

Proof. From Assumption 6.1 we can apply the Fresnel diffraction formula (see [114])

$$\int_{-\infty}^{+\infty} \exp\left(i\phi(x)\right)\left(\sum_{l\in\mathbb{Z}} \psi(x-ld)\right)\exp\left(i\pi\frac{(y-x)^2}{\lambda D}\right)\text{dx}\,. \tag{6.12}$$

From there we use Assumption 6.2 (Taylor approximation) to get

$$\sum_{l\in\mathbb{Z}} \exp(i\phi(ld))\int_{-\infty}^{+\infty} \exp\left(i(x-ld)\phi'(ld)\right)\psi(x-ld)\exp\left(i\pi\frac{(y-x)^2}{\lambda D}\right)\text{dx}\,. \tag{6.13}$$

Next we rewrite the integral using the change of variable $x' = x - ld$ and obtainn

$$\sum_{l\in\mathbb{Z}} \exp(i\phi(ld))\int_{-\infty}^{+\infty} \exp\left(ix\phi'(ld)\right)\psi(x)\exp\left(i\pi\frac{(y-ld-x)^2}{\lambda D}\right)\text{dx}\,. \tag{6.14}$$

Developing the quadratic term then yields

$$\sum_{l\in\mathbb{Z}} \exp(i\pi(ld))\int_{-\infty}^{+\infty} \exp\left(ix\phi'(ld)\right)\psi(x)\exp\left(i\pi\frac{(y-ld)^2+x^2-2(y-ld)x}{\lambda D}\right)\text{dx}\,. \tag{6.15}$$

Finally separating the exponential into individual factors leads to

$$\sum_{l\in\mathbb{Z}} \exp(i\phi(ld)) \exp\left(i\pi\frac{(y-ld)^2}{\lambda D}\right) \int_{-\infty}^{+\infty} \psi(x) \exp\left(i\pi\frac{x^2}{\lambda D}\right) \exp\left(i2\pi\frac{y-ld-\lambda D\phi'(ld)/2\pi}{\lambda D}\right) \mathrm{dx}\,. \tag{6.16}$$

It remains to observe that the integral is actually a Fourier transform.

□

6.3.2 Imaging requirements

In this section we discuss under which conditions one can extract the derivative of the phase from the diffraction pattern.

Separable fringe patterns

The first requirement is that the individual diffraction patterns in the sum of Property 1 do not overlap too much, so that we can separate them. To obtain a quantitative condition, we use the simplifying assumption that the grating profile is

$$\psi(x) = \mathrm{rect}(x/d) \tag{6.17}$$

and we will also neglect the additional phase factors. Then

$$\mathscr{F}\{\psi\}\left(\frac{2\pi y}{\lambda D}\right) = d\,\mathrm{sinc}\left(\frac{dy}{\lambda D}\right) \tag{6.18}$$

and the Rayleigh criterion for separating two such functions is given in the Condition 6.3.1.

Condition 6.3.1 (Separability)**.**

$$\lambda D/d < \Delta y\,. \tag{6.19}$$

Using the values in Table 1, $\lambda D/d \simeq 0.8\times 10^{-6}$.

Measurable displacement

For the diffraction patterns to be separable, the displacement due to the derivative of the phase should only be a fraction of the Rayleigh distance. At the same time, it should be large enough to be measurable using a phase-stepping technique (that is, moving the intensity grating along the y axis). This motivates the arbitrary condition $\lambda D\phi'(ld)/2\pi \simeq \lambda D/10d$.

Condition 6.3.2 (Measurability)**.** $\phi'(ld) \simeq 2\pi/10d$.

This means that the practical phase gradients that can be measured with this technique are on the order of one cycle (2 radians) per $40\mu m$ when the values of Table 6.1 are taken into account.

Optimal contrast

Another issue is how to best choose the distance D between the phase grating and the imaging plane. We address this problem in the case where there is no sample before the phase grating ((x) = constant). Introducing an object in the beam path amounts to perturbing this reference case.

Proposition 6.2. *If the wave arriving at the phase grating is plane, the field at the imaging plane is proportional to*

$$\int_{-\infty}^{\infty} \psi(x) \sum_{l\in\mathbb{Z}} \exp\left(-i\pi \frac{l^2 d^2}{\lambda D}\right) \exp\left(i2\pi(y-x)/d\right) dx\,. \tag{6.20}$$

In particular, if $D = d^2/2N\lambda$ where $N \in \mathbb{N}$, the field reduces to $\sum_{l\in\mathbb{Z}} \psi(x-ld)$.

The fact that the grating profile is repeated at distances that are integer multiples of $2d^2/\lambda$ is known as the Talbot effect. Note that when using a phase grating such configurations are not desired because the contrast vanishes (constant intensity).

Proof. In the case of a plane wave, the Fresnel diffraction formula (1) reduces to

$$\int_{-\infty}^{+\infty} \psi(x) \sum_{l\in\mathbb{Z}} \exp\left(i\pi \frac{(y-x-ld)^2}{\lambda D}\right) \mathrm{d}x\,. \tag{6.21}$$

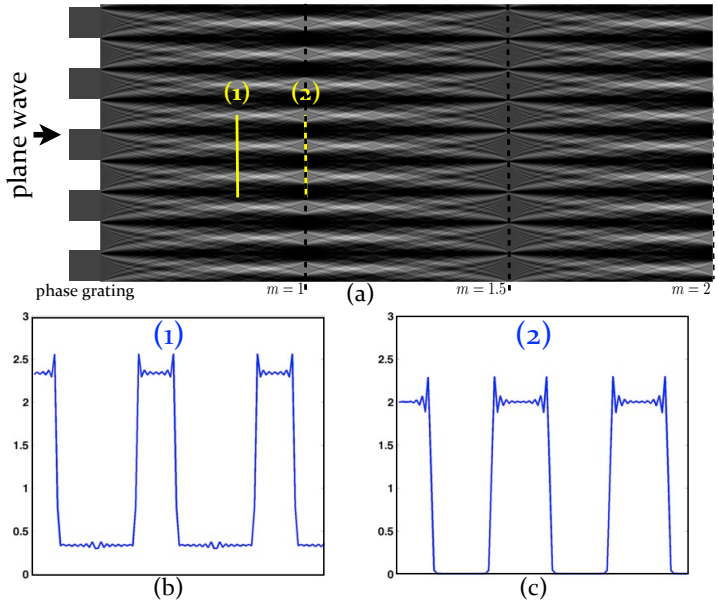


Figure 6.3: (a) Talbot carpet for phase grating. The intensity curves correspond to two different distances from the grating as specified in (a) by vertical lines. The intensity of (1) and (2) is depicted in (b) and , respectively.

The integral (6.21) is a convolution. To compute the second function involved in this convolution (the sum) we use the Poisson summation formula

$$\sum_{l\in\mathbb{Z}} f(t-ld) = \frac{1}{d}\sum_{l\in\mathbb{Z}} \mathscr{F}\{f\}(2\pi\lambda/d)\exp(i2\pi lt/d) \tag{6.22}$$

with $t = y-x$. Here $f(t) = \exp(i\pi t^2/\lambda D)$ and thus $\mathscr{F}\{f\}(\omega) = \sqrt{i\lambda D}\exp(-i\lambda D\omega^2/4\pi)$, which yields the general result. If $D = d^2/2N\lambda$ we have $\mathscr{F}\{f\}(2\pi l/d) \simeq 1$. Noting that this also happens when $f(t)$ is the Dirac distribution $\delta(t)$ and applying the Poisson summation formula in the other direction yields the particular case (convolution with a stream of Diracs). □

6.4 Reconstruction framework using ADMM with preconditioned CG

We develop the proposed reconstruction framework in the context of GI. In this regard, we formulation the reconstruction as an optimization problem 4.9.

6.4.1 Discretization of the Forward model

The mathematical model of differential phase contrast tomography is based on the first derivative of the x-ray transform of the real part of the refractive index of the object. The model of the dark field and the attenuation part are linked to the x-ray transform. In order to discretize the forward imaging operator, we use the discretization scheme introduced in Section 3.1.

6.4.2 problem specific regularization

The regularization term is chosen based on the following considerations: First, we observe that the null space of the derivatives of the Radon transform contain the zero frequency (constant). To compensate for this we incorporate the energy of the coefficients into the regularization term. Second, to enhance the edges in the reconstructed image, we impose a sparsity constraint in the gradient domain. This leads to

$$\Psi_1(\mathbf{c}) = \|\mathbf{c}\|^2 ,$$
$$\Psi_2(\mathbf{c}) = \sum_{\boldsymbol{k}} \|\{\mathbf{L}\mathbf{c}\}_{\boldsymbol{k}}\|_1 , \tag{6.23}$$

where λ_1 and λ_2 are regularization parameters and $\{\mathbf{L}\mathbf{c}\}_{\boldsymbol{k}} \in \mathbb{R}^2$ denotes the gradient of the image at position $\boldsymbol{k}$. More precisely, the gradient is computed based on our continuous-domain image model

$$\frac{\partial f}{\partial x_1}[k_1,k_2] = (h_1[\cdot,\cdot] * c[\cdot,\cdot])[k_1,k_2]$$
$$\frac{\partial f}{\partial x_2}[k_1,k_2] = (h_2[\cdot,\cdot] * c[\cdot,\cdot])[k_1,k_2], \tag{6.24}$$

where $h_1 = (\partial/\partial x_1)\,\varphi$ and $h_2 = (\partial/\partial x_2)\,\varphi$.

If the basis function is the tensor product of B-splines, the following property holds:

Proposition 6.3. *Let $f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{Z}^2} c_{\boldsymbol{k}}\beta^n(\boldsymbol{x}-\boldsymbol{k})$. The gradient of f on the Cartesian grid is*

$$\frac{\partial f}{\partial x_1}[k_1,k_2] = ((h_1[\cdot,k_2] * c[\cdot,k_2])[k1,\cdot] * b_2[k_1,\cdot])[k_1,k_2]$$
$$\frac{\partial f}{\partial x_2}[k_1,k_2] = ((h_2[k_1,\cdot] * c[k_1,\cdot])[\cdot,k_2] * b_1[\cdot,k_2])[k_1,k_2], \tag{6.25}$$

where $k_1,k_2 \in \mathbb{Z}$, $h_i[k_1,k_2] = \beta^{n-1}(k_i+\frac{1}{2}) - \beta^{n-1}(k_i-\frac{1}{2})$, and $b_i[k_1,k_2] = \beta^n(k_i)$ for $i = 1,2$.

This property allows one to compute the gradient operator for the case of cubic B-splines at the cost of simple finite difference. It is consistent with the continuous-domain image model.

A standard result for FISTA-type algorithms is that the solution of the minimization of $\mathscr{L}_\mu(\mathbf{c}^{k+1},\mathbf{u},\boldsymbol{\alpha}^k)$ with respect to $\mathbf{u}$ is given by the shrinkage function [87, 115]

$$\mathbf{u}^{k+1} = \max\left\{\left|\mathbf{L}\mathbf{c}^{k+1} + \frac{\boldsymbol{\alpha}^k}{\mu}\right| - \frac{\lambda_2}{\mu}, 0\right\} \operatorname{sgn}(\mathbf{L}\mathbf{c}^{k+1} + \frac{\boldsymbol{\alpha}^k}{\mu}). \tag{6.26}$$

6.4.3 Parameter selection

From a statistical point of view, the Baysian estimator for the additive Gaussian white noise inverse problem $\mathbf{g} = \mathbf{Hc} + \mathbf{n}$ can be written as

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{Hc} - \mathbf{g}\|^2 - \sigma_n^2 \log p(\mathbf{c}) \right\}, \tag{6.27}$$

where σ_n^2 is the noise variance and $p(\mathbf{c})$ is the prior density of the signal. Assuming that the gradient sample values $\mathbf{u}_k = \{\mathbf{Lc}\}_k$ are i.i.d. Laplace distributed, we have

$$-\log p(\mathbf{c}) \propto \frac{1}{\sigma_u} \sum_k \|\mathbf{u}_k\|_1 ,$$

where σ_u^2 is the variance of the gradient values. On the other hand, by definition $\text{SNR} = 10\log(\|\mathbf{Hc}\|_2^2 / N\sigma_n^2)$ where N is the number of pixels in the image. Therefore $\sigma_n^2 = 10^{-.1\text{SNR}} \|\mathbf{Hc}\|_2^2 / N$. In practice $\|\mathbf{Hc}\|_2^2$ can be approximated by $\|\mathbf{g}\|_2^2$ and σ_u^2 is proportional to $\sum_k \|\mathbf{u}_k\|_2^2$ which is empirically estimated using $\|\mathbf{g}\|_2^2$. This leads to the following rule of thumb for setting the TV regularization parameter: $\lambda_2 \propto 10^{-.1\text{SNR}} \|\mathbf{g}\|_2$. Therefore, the TV parameter is proportional to the norm of the measurement. The proportionality constant is related to its signal-to-noise ratio.

The Tikhonov parameter is chosen as small as possible and is set to $\lambda_1 = 10^{-5}$. Here we use $\lambda_2 = ||\mathbf{g}||_2 \times 10^{-3}$. Based on our experience, the parameter μ can be chosen ten times larger than λ_2. The number of inner iterations for solving the linear step plays no role in the convergence of the proposed technique, but affects speed; we suggest to choose it as small as possible, typically 2 or 3. We use cubic B-splines with $m = 3$ as the basis functions.

6.4.4 Convergence and inexact minimization

In ADMM-PCG, we use conjugate gradient in order to solve the first step of ADMM because there is no direct solution. One main advantage of ADMM is that it can converge even when the minimization steps have been solved approximately using an iterative method, if the minimizations satisfy an appropriate condition, such as being summable [92, 107].

6.4.5 Stopping criteria

Is has been shown in [92] that there is a bound on the objective function of the current point which involves primal and dual residuals. The primal and dual residuals are denoted by **r** and **s**, respectively. Their definitions are

$$\begin{aligned} \mathbf{r}^k &= \mathbf{L}\mathbf{c}^k - \mathbf{u}^k \\ \mathbf{s}^k &= \mathbf{L}^T \left(\mathbf{u}^k - \mathbf{u}^{k-1} \right). \end{aligned} \tag{6.28}$$

This suggests that a reasonable termination criteria is when the norm of primal and dual residuals is small.

6.4.6 Convergence speed in practical problems

In general, ADMM can be slow for high accuracy convergence. However, it converges to a sufficiently accurate solution for many applications within a few iterations. This makes it practically valuable for the cases when modest accuracy is sufficient. Fortunately, this is usually the case of the large-scale reconstruction problem. Note that typically there is not enough prior information of the object of interest, then the proposed cost function is based on some approximations [83].

6.5 Experimental Validation

6.5.1 Performance metrics

We use the structural similarity measure (SSIM) [116,117] and signal-to-noise ratio (SNR) for measuring the quality of the reconstructed image. SSIM is a similarity measure proposed by Z. Wang *et al.* which compares the luminance, contrast, and structure of images. SSIM is computed for a window of size $(R \times R)$ around each image pixel. The SSIM measure for two images $\mathbf{x}$ and $\widehat{\mathbf{x}}$ for the specified window is

$$\mathrm{SSIM}(\mathbf{x}, \widehat{\mathbf{x}}) = \frac{(2\mu_x \mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)}, \tag{6.29}$$

where C_1 and C_2 are small constant values to avoid instability. μ_x and $\mu_{\hat{x}}$ denote the empirical mean of the images $\mathbf{x}$ and $\widehat{\mathbf{x}}$ in the specified window, respectively. The empirical variance of the corresponding images are σ_x and $\sigma_{\hat{x}}$. The covariance of two images is denoted

by $\sigma_{x\hat{x}}$ for the corresponding window. In our experiments, we choose $C_1 = C_2 = (.001 * L)^2$ where L is the dynamic range of the image pixel values. SSIM for the total image is obtained as the average of SSIM over all pixels. It takes values between 0 and 1 with 1 corresponding to the highest similarity.

Our other quality measure is the SNR. If $\mathbf{x}$ is the oracle and $\widehat{\mathbf{x}}$ is the reconstructed image we have

$$\mathrm{SNR}(\mathbf{x}, \widehat{\mathbf{x}}) = \max_{a,b \in \mathbb{R}} 20 \log \frac{||\mathbf{x}||_2}{||\mathbf{x} - a\widehat{\mathbf{x}} + b||_2}. \tag{6.30}$$

Higher values of the SNR correspond to a better match between the oracle and the reconstructed image.

6.5.2 Experimental result

To validate our reconstruction method, we conducted experiments with real data acquired using the TOMCAT beam line of the Swiss Light Source at the Paul Scherrer Institut in Villigen, Switzerland. The synchrotron light is delivered by a 2.9 T super-bending magnet. The energy of the X-ray beam is 25keV [118]. We used nine phase steps over two periods to measure the displacement of the diffraction pattern described in Section 2. For each step a complete tomogram was acquired around 180 degrees; we used 721 uniformly distributed projection angles. Image acquisition was performed with a CCD camera whose pixel size was 7.4μm.

For our experiments we used a rat brain sample. The sample is embedded in liquid paraffin at room temperature. This is necessary to match the refractive index of the sample with its environment, so that the small-refraction-angle approximation holds. Finally the projections were post-processed, including flat-field and dark-field corrections, for the extraction of the phase gradient.

Figure 6.4 contains a comparison of the performance of the proposed algorithm against FISTA. For comparison, the convergence in [35] requires at least 65 iterations to converge. This allows for a rough yet informative comparison of computational costs, in terms of number of evaluations of the forward model or its adjoint (which are the most expensive operations in the schemes discussed here). In [35], it is required to compute the forward operator twice per iteration. Therefore, the cost estimate is $65 \times 2 = 130$ evaluations of the forward operator. Meanwhile, our reconstruction scheme converges after 5 outer iterations with 2 conjugate-gradient inner iterations. Each conjugate gradient step requires one application of the forward operator and one application of its adjoint. Since the number of

viewing angles is typically less than the size of the object, the cost of the adjoint operator is less than the computation of the forward operator. If we neglect this fact, we need $(2 \times 2 \times 5) = 20$ evaluations of the forward operator. Based on these considerations, we expect our algorithm to be substantially faster. It demonstrates the benefits of using a warm initialization as well as a problem-specific preconditioner for the linear optimization step.

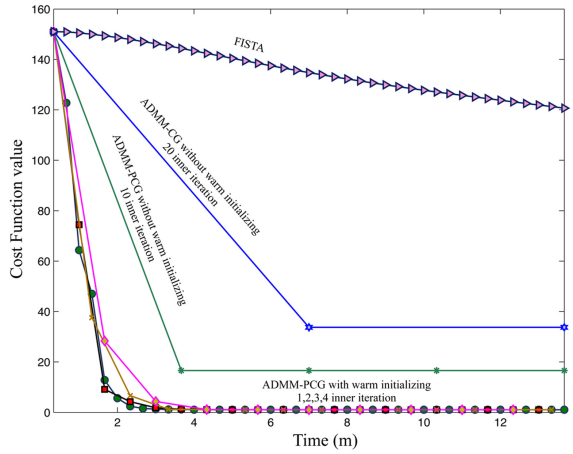


Figure 6.4: Speed of convergence of different iterative techniques for solving our regularized problem. Significant gains over the standard FISTA algorithm can be obtained using our ADMM-based scheme. Observe that the number of inner iterations in the ADMM-PCG method does not significantly influences its convergence.

We further investigated the performance of the direct filtered back-projection and of the proposed iterative reconstruction techniques on a coronal section of the rat brain. The reconstructed images with 721 angles using GFBP and those produced by our method are shown in Figure 6.5. In terms of quality, for real data, our results are more less equivalent to using the Kaiser Bessel Window functions proposed in [35].

The GFBP reconstruction contains artifacts at the boundary of the image and at specific anatomical features. For example, the bottom-right sub images in the reconstructions of

Figure 6.5 show the mammal-thalamic tract in this coronal section. One can clearly see oscillatory artifacts in the GFBP reconstruction. This could confuse the biologist or automated diagnostic systems for determining the nucleus and immoneurins in that region. The middle-right and top-right images show a part of the thalamus and the region between the thalamus and the hippocampus, respectively. To reduce the artifacts in the GFBP technique, we also implemented a smoothed version of the GFBP algorithm. Specifically, we modified the filter in the third step of Algorithm 3 as

$$\widehat{q}(\omega_y) = \frac{1}{2\pi}\frac{1}{|\omega_y|} \times h^k(\omega_y)\,, \tag{6.31}$$

where $h(\omega_y)$ is a lowpass filter and k is an exponent that acts as a smoothing parameter. We chose $h(\omega)$ to be the standard Hamming window. The reconstructions are shown in the bottom row of Figure 6.5. They suggest that with GFBP there is a tradeoff between artifacts and image contrast. Note that for these experiments the parameter k was optimized so as to achieve the best SNR. The figure of merits (SNR and SSIM) are indicated below each image in Figure 6.5. Visually the reconstructed image with 721 angles using our proposed technique is more faithful in comparison with the GFBP approach. Therefore, we consider it as our gold standard for investigating the dependence of our algorithm on the number of views as shown in Figure 6.6. The SNR and SSIM values are computed for the main region of the sample which includes the brain. they suggest that we can reduce the number of views with our method at least fourfold while essentially maintaining the quality of the standard reconstruction method (FBP with a complete set of views).

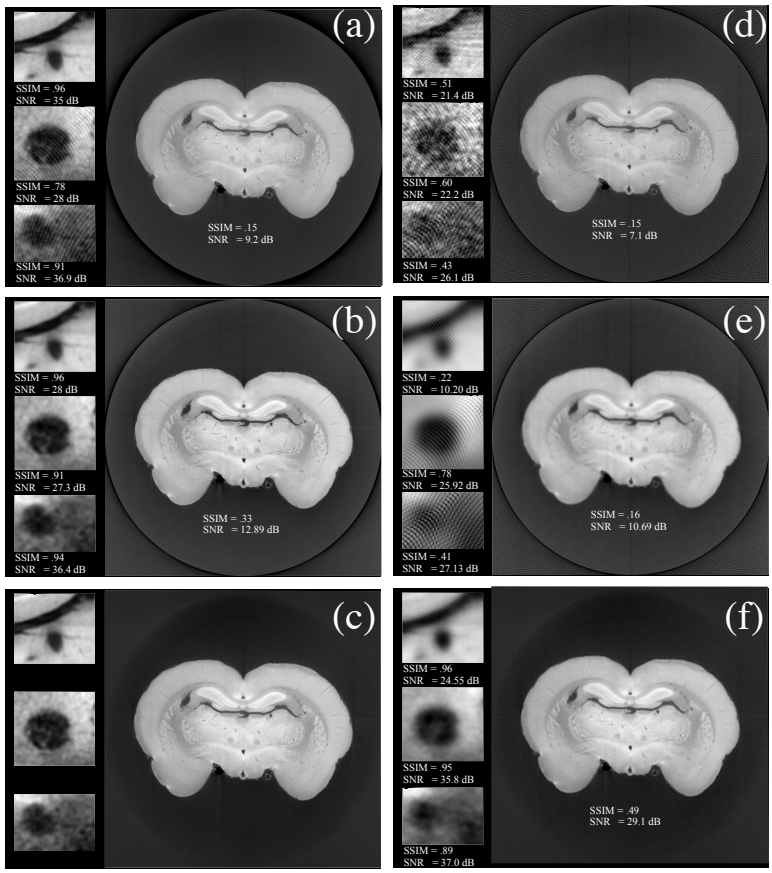
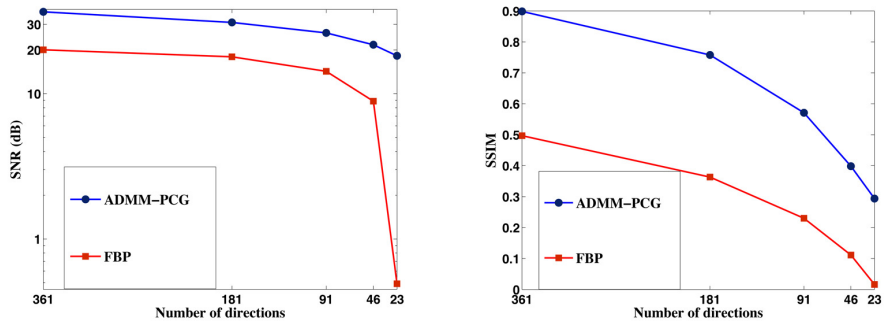


Figure 6.5: Comparison of the reconstruction results for 721 viewing angles (first column) and 181 viewing angles (second column). (a,d) GFBP, (b,e) GFBP with smoothing kernel, (c,f) the iterative ADMM. The sub-images correspond to the region between the thalamus and the hippocampus (top), a part of the talamus (middle), and the Fornix (bottom). Notice the oscillatory artifacts produced by GFBP. Applying a smoothing kernel reduces the artifacts but also blurs the reconstruction.



(a) (b)

Figure 6.6: The SNR and SSIM metrics for images reconstructed from a subset of projections.

Chapter 7

Improved Reconstruction Scheme for X-ray Grating Interferometry

[1] In Chapter 6, we applied our basic reconstruction framework (ADMM-PCG) in the context of GI. Our experimental results show the feasibility of our framework. They suggest that the proposed technique outperforms the-state-of-the-art methods. The drawback of ADMM-PCG is that there is not a clear way to impose the convex constraints such as the positivity of the refractive index on the solution. Moreover, some line artifacts were visible on the reconstruction. In this chapter, we consider a GI problem with the same physical model as discussed in Chapter 6. We first use the constrained regularized weighted norm (CRWN) to improve the reconstruction. In the context of differential phase contrast imaging, we then investigate the situations in which the measurements are wrapped. We develop a reconstruction framework in order to simultaneously unwrap and reconstruct the object of interest.

One advantage of GI is that it provides simultaneously an absorption-contrast and a phase-contrast information. In this regard, we develop a reconstruction framework to simultaneously retrieve the real and the imaginary part of the refractive index (complex refractive index reconstruction).

[1] A part of this chapter has been presented in [67]

7.1 Constrained regularized weighted norm

The CRWN algorithm is described in Section 4.2.2. The main step is the choice of the regularization and its proximal map.

7.1.1 Problem specific regularization

As for the regularization, we consider two distinct options.

1) Our first option is the use of a total-variation (TV) regularization term to enhance the edges in the reconstructed image. Therefore, we set

$$\Psi_2(\mathbf{c}) = \|\mathbf{L}\mathbf{c}\|_{1,1} \tag{7.1}$$

with $\|\mathbf{L}\mathbf{c}\|_{1,1} = \sum_i \|\{\mathbf{L}\mathbf{c}\}_i\|_1$, where the sum is computed on all B-spline coefficients and $\{\mathbf{L}\mathbf{c}\}_i \in \mathbb{R}^2$ is the gradient vector of the image at position i. The discrete gradient operator $\mathbf{L} : \mathbb{R}^N \to \mathbb{R}^{N\times 2}$ is computed according to Proposition 2 in [36]. Here, the regularization operator is the discrete gradient operator and the mixed $\ell_1 - \ell_1$ norm is chosen as the potential function. As the dual norm of the ℓ_1 norm is ℓ_∞, the dual ball is defined as

$$\begin{aligned}\mathscr{B}_{\infty,\infty} = \{\mathbf{p} = [\mathbf{p}_1^T, \mathbf{p}_2^T, ..., \mathbf{p}_N^T]^T \in \mathbb{R}^{N\times 2} : \\ \|\mathbf{p}_i\|_\infty \leq 1, \forall i = 1,2,...,N\}.\end{aligned} \tag{7.2}$$

Therefore, the orthogonal projection of $\mathbf{y} \in \mathbb{R}^{N\times 2} = [\mathbf{y}_1^T, \mathbf{y}_2^T, ..., \mathbf{y}_N^T]^T$ onto this ball is $\widetilde{\mathbf{y}} = \mathscr{P}_{\mathbb{B}_{\infty,\infty}}(\mathbf{y})$ with

$$\begin{aligned}[\widetilde{\mathbf{y}}_i]_j = \mathrm{sgn}([\mathbf{y}_i]_j)\min(|[\mathbf{y}_i]_j|, 1), \\ \forall i = 1,2,...,N, j = 1,2,\end{aligned} \tag{7.3}$$

where $[\cdot]_j$ is the jth entry of the corresponding vector and $\widetilde{\mathbf{y}} = [\widetilde{\mathbf{y}}_1^T, \widetilde{\mathbf{y}}_2^T, ..., \widetilde{\mathbf{y}}_N^T]^T$. This regularization is well-matched to piecewise-constant images.

2) Owing to the fact that biological and medical specimens consist mostly of filament-like and complicated structures, we investigate higher-order extensions of total variation. We apply the Hessian Schatten-norm regularization (HS) as our second option. It can eliminate the staircase effect of TV regularization and results in piecewise-smooth variations of intensity in the reconstructed image. We set

$$\Psi_2(\mathbf{c}) = \|\mathbf{H}\mathbf{c}\|_{1,\mathscr{S}_1}\,, \tag{7.4}$$

where $\mathbf{H}:\mathbb{R}^N \to \mathbb{R}^{N\times 2\times 2}$ is the discrete Hessian operator and $\|\mathbf{Hc}\|_{1,\mathscr{S}_1}$ is the mix of ℓ_1 and nuclear norm. The norm can be computed with $\|\mathbf{Hc}\|_{1,\mathscr{S}_1} = \sum_i(\sigma_{1,i}+\sigma_{2,i})$, where $\sigma_{1,i}$ and $\sigma_{2,i}$ are the singular values of the Hessian matrix at position i. Therefore, the corresponding unit-norm dual ball is defined as

$$\begin{aligned}\mathscr{B}_{\infty,\mathscr{S}_\infty} = \{\mathbf{p} = [\mathbf{p}_1^T, \mathbf{p}_2^T, ..., \mathbf{p}_N^T]^T \in \mathbb{R}^{N\times 2\times 2} : \\ \|\mathbf{p}_i\|_{\mathscr{S}_\infty} \leq 1, \forall i = 1,2,...,N\},\end{aligned} \tag{7.5}$$

where $\|\cdot\|_{\mathscr{S}_\infty}$ is the ℓ_∞-norm of the singular values of the corresponding matrix (for more details, we refer the reader to [119]).

7.1.2 Parameter setting

The proposed algorithm has several parameters.

- Parameters λ_1 and λ_2: We use the approach proposed in [36]; $\lambda_1 = 10^{-5}$ and $\lambda_2 = 10^{-4}\|\mathbf{g}\|$. The experimental results suggest that this choice of parameters yields the optimal performance.

- Parameter μ: This parameter affects the convergence speed of ADMM. Since the algorithm is not too sensitive to it, we use a fixed value ($\mu = 1$).

- Parameter λ: This is a parameter of the proximal map operator in (4.32). Since the second step of ADMM is solving (4.18), we have $\lambda = \lambda_2/\mu$.

- Lipschitz constant L: The Lipschitz constant of $\nabla f(\mathbf{p}) = -\lambda\mathbf{R}\mathscr{P}_{\mathscr{C}}(\mathbf{z} - \lambda\mathbf{R}^T\mathbf{p})$ is approximated by the Lipschitz constant of the same operator without the convex projection $\mathscr{P}_{\mathscr{C}}$ since the projection on the convex set is firmly non-expansive. Thus,

$$L \approx \lambda^2 \times \lambda_{\max}(\mathbf{RR}^T), \tag{7.6}$$

  where $\lambda_{\max}(\mathbf{A})$ is the maximum eigenvalue of the matrix $\mathbf{A}$. For our regularization scheme $\lambda_{\max}(\mathbf{RR}^T) \leq \gamma$ where $\gamma = 8$ for the TV regularization for two-dimensional problems, and its value is 64 for the HS regularization as computed in [119].

- Parameter τ: We set it to $\tau = 1/10 \times L^{-1}$.

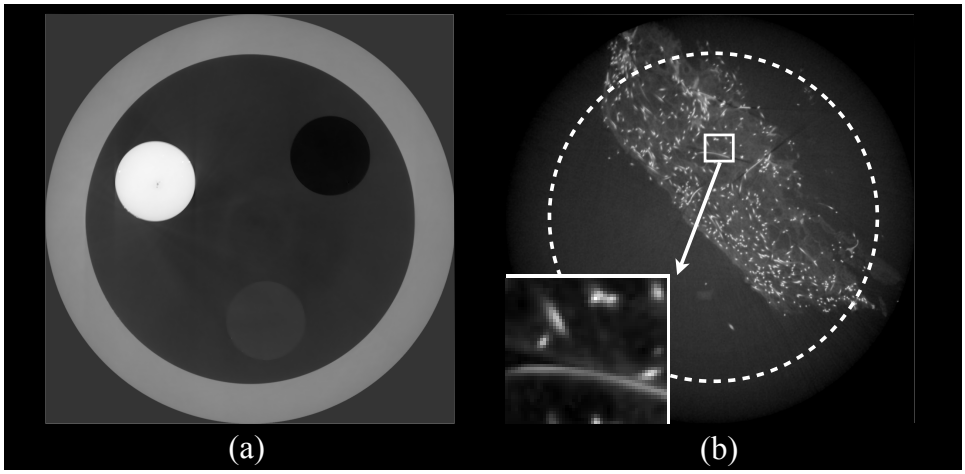


Figure 7.1: Two reference samples (a) and (b).

7.1.3 Experimental result

We compared the proposed algorithm to FBP and to ADMM-PCG, which appears to be the current state of the art for the reconstruction of X-ray-DPCI tomograms [36].

All experiments involved real data acquired at the TOMCAT beam line of the Swiss Light Source at the Paul Scherrer Institut in Villigen, Switzerland. The common approach for these experiments is to use a reconstruction from a large number of projections as a reference for evaluating results obtained with significantly fewer projections. In addition, the convex constraints that we apply are the positivity of the refractive index combined with the support-related constraint that the solution should be zero outside the tube that contains the specimen.

In order to identify the benefits of the proposed algorithm (CRWN), we first tested the algorithms under extreme conditions: We used only 72 projections as input, while the reference was reconstructed from 1,200 projections. For this first experiment we used a phantom that was composed of a tube and three cylinders containing liquids with different refractive indices as shown in Figure 7.1(a).

The performance of different algorithms are compared in Table 7.1. Clearly, the new method outperforms ADMM-PCG [36]. Applying the convex constraint improves the

	The proposed algorithm (CRWN)					
	Constrained		Unconstrained			
	TV	HS	TV [120]	HS	ADMM-PCG [36]	FBP
Phantom						
SNR(dB)	**27.49**	23.91	25.89	21.82	17.62	2.177
SSIM	**0.509**	0.369	0.339	0.196	0.145	0.07
Scaffold						
SNR(dB)	25.34	**25.58**	22.91	22.25	20.09	6.45
SSIM	0.673	**0.699**	0.574	0.566	0.512	0.186
Scaffold ROI						
SNR(dB)	26.51	**27.05**	23.78	23.75	23.58	23.09
SSIM	0.968	**0.974**	0.944	0.958	0.852	0.516

Table 7.1: Performance of different reconstruction techniques that have been applied on Phantom and Scaffold samples.

signal-to-noise ratio (SNR) and the structural similarity index measure (SSIM) [116] even further. The result of the algorithm proposed in [120] is the same as CRWN-TV without CC, but it is slower since it uses FISTA. As expected, owing to the piecewise-constant structure of the sample, TV outperforms HS regularization.

We conducted another experiment with a coronal section of a scaffold that is used for surgery. The reference image was built from 2,000 projections as depicted in Figure 7.1(b). The algorithms were then benchmarked on a subset of 250 projections. Although these conditions are less severe, FBP still produces high-frequency patterns that are visible in Figure 7.2(a). ADMM-PCG almost completely suppresses these artifacts, at the expense of light smoothing as shown in Fig. 7.2(b). Overall, CRWN yields sharper images, as shown in Figure 7.2(c) and 7.2(d), which is also reflected by the quality metrics. In addition, Hessian type regularization eliminates the staircase effect of TV which is more visible in the selected region of interest.

It is seen in Figure 7.3(a) that CRWN is significantly faster at minimizing the cost functional than the standard FISTA algorithm. In addition, it appears that the convergence speed is not very sensitive to the number of inner iterations as we use a warm initialization. We illustrate in Figure 7.3(b) the robustness of CRWN with respect to the number of projections in terms of SNR. Owing to the poor performance of FBP in reconstructing boundaries, we compute the SNR for the region specified by a dashed circle in Figure 7.1(b).

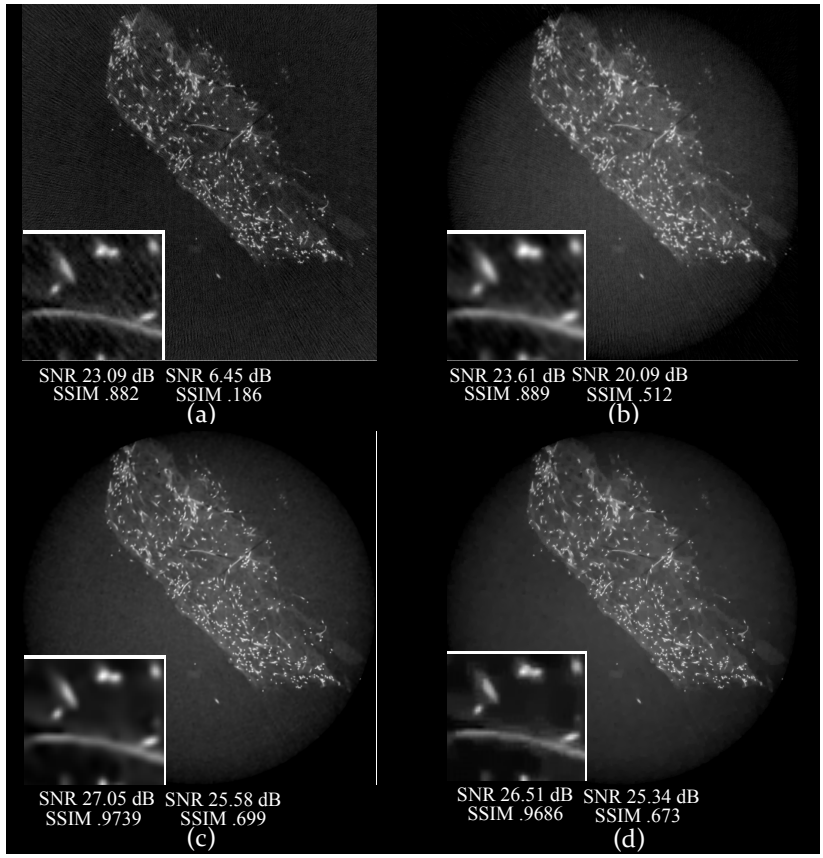


Figure 7.2: Scaffold reconstruction with 250 projections using (a) FBP, (b) ADMM-PCG, (c) CRWN with HS regularization and (d) CRWN with TV regularization.

7.2 Joint phase unwrapping and radiation dose reduction in DPCI

Phase measurements are ambiguous. In the case of X-ray differential phase-contrast tomography, the measured phase is the wrapped version of the physical phase shift, as shown

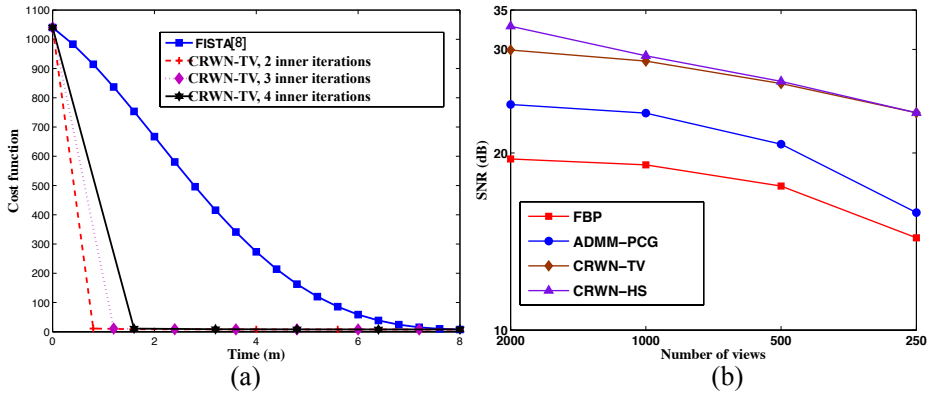


Figure 7.3: The reconstruction performances concerning speed and quality is shown in (a) and (b), respectively.

in Figure 7.6 (a) and Figure 7.6 (b). Note that the wrapped measurements introduce strong line artifacts in the reconstructed object.

The wrapping operator can be defined as (the principal domain for the phase is chosen as $]-\pi,\pi]$ here)

$$W(\varphi) = \varphi + 2k\pi, \text{ with } k \in \mathbb{Z} \text{ such that } W(\varphi) \in]-\pi,\pi] \,. \tag{7.7}$$

Observe that (7.7) implies that

$$W(a) = a \Leftrightarrow |a| \leqslant \pi \,. \tag{7.8}$$

The ambiguity comes from the fact that W is a non-injective operator. Finding the original phase φ from the wrapped phase $\phi = W(\varphi)$ is known as the *phase unwrapping problem*. Phase unwrapping is an ill-posed problem because infinitely many original phase solutions ϕ exist for any given phase measurement φ. The discrete space of solutions $S_c = \{\phi + 2k\pi \mid k \in \mathbb{Z}\}$ contains φ. There is no general way of finding the best solution to phase unwrapping and problems should be studied in a case by case manner. Additional knowledge has to be provided in order to determine a particular solution. One possibility

is to use side measurements such as attenuation contrast [121] to correct pixels that are suspected to be incorrectly wrapped. Another procedure is to use prior knowledge about the data. Typically, phase shift values could theoretically take any value, but it can reasonably be assumed that they span a finite range for objects with finite dimensions. Moreover, refractive indices are bounded by the highest and lowest values of the materials present in the sample and the type of objects being imaged.

Itoh pioneered phase unwrapping theory by proposing a solution that minimizes the phase differences [122]. His solution is obtained by computing the finite differences of the observed phase, wrapping them, and then summing them again to recover the unwrapped phase values (see Figure 7.4). The original phase is recovered under the condition that $|\Delta\varphi_i| < \pi \ \forall i$, which is known as *Itoh's condition*. This is a central condition for several phase unwrapping algorithms, as it combines a condition on sampling rate and phase evolution. Note that applying the algorithm on the reversed signal leads to the same solution up to a constant, because of the integration step. The principle of the algorithm can be seen as using the previous point as an estimator for the current point and to take the solution in S that is the closest to the estimator (while the wrapped phase can be seen as using 0 as an estimator). Based on this observation, other algorithms for direct phase unwrapping can be built, by simply choosing a different estimator.

High sensitivity to electron density variations in DPCI provides on one hand the possibility to measure small phase shifts. On the other hand, it implies that a phase wrapping on small deflection angle can lead to strong artifacts in the differential phase-contrast images and their reconstructions. In order to eliminate the artifacts, one can first unwrap the phase measurements. Note that DPCI measurements are the spatial gradient of the phase which makes the unwrapping problem more challenging. Recently, Epple formulated the unwrapping procedure as a maximum likelihood estimator by taking advantage of the polychromatic spectrum provided by conventional X-ray sources. Thus, one can first apply the proposed iterative scheme [123] to unwrap the phase measurements and then reconstruct the object of interest. This technique is computationally costly.

Here, we aim at developing a reconstruction framework such that it 1) removes the line artifacts on the reconstructed object and 2) reduces the radiation dose by decreasing the required number of orientations for the reconstruction. The goal is to avoid the introduction of extra computational cost in the reconstruction algorithm. We develop two distinct algorithms. We validate each one using real data measurements. The sample is an aortic heart valve with two leaflets. Normally, the aortic heart valve has three leaflets. However around 1% to 2% of the human population have a genetic defect that results in the formation of only two leaflets [124]. This can have severe effects on the quality of life so that

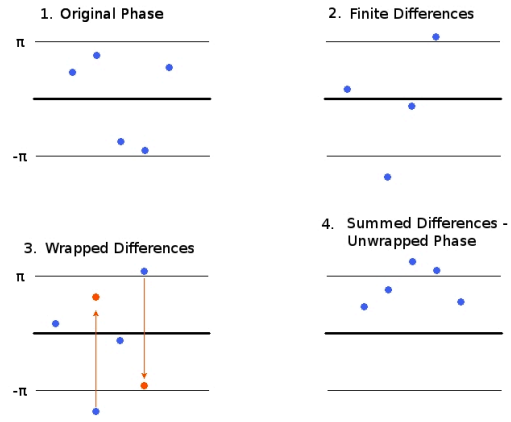
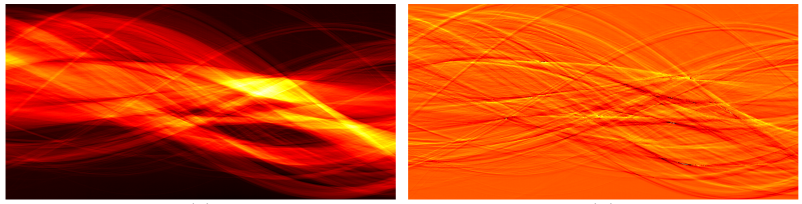


Figure 7.4: Itoh's Phase Unwrapping; Illustration of Itoh's algorithm for phase unwrapping.

surgery is often recommended. The main problem is the performance reduction of the heart due to the inflammation around and calcification of the valves, which we want to monitor. Therefore, it is required to be able to segment different tissues and calcifications in the 3-D reconstructed heart. GI, by providing three different complementary information, shows significant potential in this case. Modregger [125] showed that effective atomic number by deviding the imaginary part of the refractive index by its real part show better contrast to segment different parts. Unfortunately, on the given measurements, the phase information (the real part of the refractive index) suffers from wrapping problem as shown in Figure 7.6 (a). In the sequel, we validate the proposed methods using this real data measurements. The sinogram corresponds to the absorption (imaginary part of the refractive index) and the phase are depicted in Figure 7.5.

1) Unlike filtered back projection type methods, iterative reconstruction algorithms do not need the whole sinogram. Therefore, we first detect the wrapped points on the sinogram using the Itoh's rule, and then we remove them from the measurements. This can be performed by introducing a mask $\mathbf{M}$ in front of the system matrix $\mathbf{H}$ or, equivalently,



(a) (b)

Figure 7.5: The GI setup provides the absorption and the phase measurements simultaneously. Sinogram of (a) the absorption and (b) the phase measurements.

formulating the reconstruction problem as the minimization

$$\mathbf{c}_0 = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} \left\{ J(\mathbf{c}) = \frac{1}{2} \|\mathbf{H}\mathbf{c} - \mathbf{g}\|_{\mathbf{M}}^2 + \lambda_1 \Psi_1(\mathbf{c}) + \lambda_2 \Psi_2(\mathbf{c}) \right\} \tag{7.9}$$

with a weighted norm data fidelity term where $\mathbf{M}$ is the mask matrix whose entries are one except at the unwrapped points discovered by Itoh's rule where they take zero values. Total variation is chosen as regularization. We use the generalization of the ADMM reconstruction algorithm to solve the optimization problem.

To validate the method, we use an aortic heart valve real data measurement. Since there is a correlation between the real and the imaginary part of the refractive index of this sample, we compare visually the reconstructions with the imaginary part as depicted in Figure 7.7 (a).

First we use Itoh's unwrapping operator on the real data phase measurements and then we apply a standard iterative scheme with total variation regularization. The reconstructed object is shown in Figure 7.7 (b). The strong artifacts are pointed out by yellow arrows. It suggests that the typical regularization scheme with Itoh's unwrapping operator is not sufficient to remove the wrapping artifacts.

Second, we detect the wrapped points by Itoh's rule, and then we construct the mask as shown in Figure 7.6(c). We minimize the cost (7.9) using the generalized ADMM introduced in Section 4.2.2. The reconstruction result is shown in Figure 7.7 (c). It suggests that the proposed approach mostly eliminates the line artifacts on the reconstructed object. In addition, the number of views is reduced fourfold.

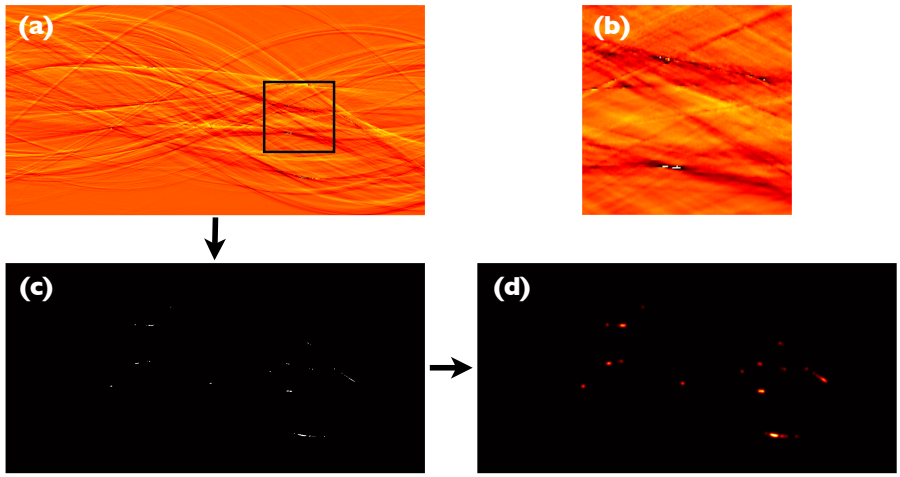


Figure 7.6: (a) The wrapped DPCI sinogram. (b) A zoomed region of the wrapped sinogram in which the wrapped points are more visible. (c) Detected wrapped points using Itoh's rule. (d) Gaussian weighted norm around the wrapped points in the sinogram.

2) Although the wrapped points of the sinogram suggested by Itoh's rule are removed, some line artifacts are still visible in the reconstructed object. It implies that there are still wrapped points which have not been detected. Instead of removing points from the measurements, one can assign a weight to each point of the sinogram. The weight specifies how suspected this point is for being wrapped. The points determined by Itoh's rule are wrapped with high probability. Moreover, the points around them are nominated with probability inversely linked to their distance to these points. Therefore, we dedicate a Gaussian curve around the suggested points by Itoh's rule as weights. The weighted operator is depicted in Figure 7.6(d). The reconstructed object is shown in Figure 7.7(d). In order to compare different reconstruction schemes more precisely, the region of interest specified by a yellow rectangle in Figure 7.7 (a) is shown in Figure 7.8. The specified edge is well shaped using the reconstruction framework with the suggested Gaussian weight.

7.3 Complex refractive index reconstruction

The GI setup provides simultaneous information about the phase shift and the attenuation introduced on the transmitted wave. In this regard, we develop a reconstruction scheme to simultaneously reconstruct phase and absorption distribution. The forward model, as stated in Section 6.2, is in the form of

$$\begin{cases} g_1(y,\theta) &= \frac{\partial \mathscr{P}\{\alpha\}}{\partial y}(y,\theta) + n_1(y,\theta) \\ g_2(y,\theta) &= \mathscr{P}\{\beta\}(y,\theta) + n_2(y,\theta), \end{cases} \tag{7.10}$$

where n_1 and n_2 represent the absorption and phase noise that is present in the measurements g_1 and g_2. Then, the forward discretization of (8.18) is

$$\mathbf{g} = \mathbf{H}\mathbf{c} + \mathbf{n}, \tag{7.11}$$

where $\mathbf{g}$ is the vector of measurements, and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^{(0)} \end{bmatrix}. \tag{7.12}$$

There, $\mathbf{H}^{(n)}$ is the system matrix that corresponds to the nth derivative of the x-ray transform given in (3.6). Moreover, $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2)$ collects in a single column vector each B-spline coefficient of the real and the imaginary part of the refractive index. We say that $\mathbf{c}$ is the discrete representation of the object.

7.3.1 Model Fitting

We formulate the joint-retrieval problem of $\mathbf{c}$ in the framework of a penalized least-square estimation. Our goal is to find the minimizer of

$$\mathscr{J}(\mathbf{c}) = \frac{1}{2} \|\mathbf{H}\mathbf{c} - \mathbf{g}\|^2 + \Psi(\mathbf{c}), \tag{7.13}$$

where Ψ is a regularizing term made of two components.

Letting the first component be Ψ_1, we promote the enhancement of the edges by encouraging sparsity in the gradient of the real and the imaginary part of the refractive index images. We define

$$\Psi_1(\mathbf{c}) = \|\mathbf{L}\mathbf{c}\|_1, \tag{7.14}$$

with $\mathbf{L}\mathbf{c} = [\mathbf{L}\mathbf{c}_1\, \mathbf{L}\mathbf{c}_2]$.

As second component Ψ_2, we choose the nuclear total variation, which is a vectorial extension of TV [126]. Its purpose is to strengthen the correlation of AC and DPC over edges. It is given by the point-wise sum of Schatten 1-norms of the Jacobian matrix $\mathbf{J}$ and is expressed as

$$\begin{aligned} \Psi_2(\mathbf{c}) &= \|\mathbf{J}\mathbf{c}\|_{1,\mathscr{S}^1} \\ &= \sum_{\mathbf{k}\in\mathbb{Z}^2} \left\| \begin{bmatrix} [\mathbf{D}_{x_1}\mathbf{c}_1]_{\mathbf{k}} & [\mathbf{D}_{x_2}\mathbf{c}_1]_{\mathbf{k}} \\ [\mathbf{D}_{x_1}\mathbf{c}_2]_{\mathbf{k}} & [\mathbf{D}_{x_2}\mathbf{c}_2]_{\mathbf{k}} \end{bmatrix} \right\|_{\mathscr{S}^1}. \end{aligned} \tag{7.15}$$

To be consistent with (8.18), the directional derivative operator must be such that

$$[\mathbf{D}_{x_1}]_{\mathbf{p}} = \frac{\partial\varphi}{\partial x_1}(\mathbf{p})\,,\, [\mathbf{D}_{x_2}]_{\mathbf{p}} = \frac{\partial\varphi}{\partial x_2}(\mathbf{p})\,, \tag{7.16}$$

where $\mathbf{p} = (p_1, p_2)$. There, $[\cdot]_{\mathbf{p}}$ specifies the $\mathbf{p}$th entry of a vector. The Schatten 1-norm $\|\cdot\|_{\mathscr{S}^1}$ is the ℓ_1 norm of the singular values of its matrix argument. Note that the Schatten 1-norm is a convex operator. The discrete Jacobian operator $\mathbf{J}$ maps $\mathbb{R}^N$ to $\mathbb{R}^{N\times 2\times 2}$, where N is the total number of coefficients within one image. All directional derivatives are computed in accordance with the discretization scheme suggested by (8.22).

7.3.2 Optimization

We aim at determining $\hat{\mathbf{c}}$ such that

$$\begin{aligned} \hat{\mathbf{c}} &= \operatorname*{argmin}_{\mathbf{c}} \Big\{ \frac{1}{2}\,\|\mathbf{H}\mathbf{c}-\mathbf{g}\|^2 \\ &\qquad + \lambda_1\,\|\mathbf{L}\mathbf{c}\|_1 + \lambda_2\,\|\mathbf{J}\mathbf{c}\|_{1,\mathscr{S}^1} \Big\}. \end{aligned} \tag{7.17}$$

We reformulate this unconstrained optimization problem as a problem constrained over the auxiliary variables $\mathbf{u} \in \mathbb{R}^N$ and $\mathbf{v} \in \mathbb{R}^{N\times 2\times 2}$ such that

$$\begin{aligned} \hat{\mathbf{c}} &= \operatorname*{argmin}_{\mathbf{c},\mathbf{u}=\mathbf{L}\mathbf{c},\,\mathbf{v}=\mathbf{J}\mathbf{c}} \Big\{ \frac{1}{2}\,\|\mathbf{H}\mathbf{c}-\mathbf{g}\|^2 \\ &\qquad + \lambda_1\,\|\mathbf{u}\|_1 + \lambda_2\,\|\mathbf{v}\|_{1,\mathscr{S}^1} \Big\}. \end{aligned} \tag{7.18}$$

The scaled form of its augmented-Lagrangian formulation is

$$\begin{aligned}\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u},\mathbf{v},\mathbf{d}_1,\mathbf{d}_2) &= \frac{1}{2}\left\|\mathbf{H}\mathbf{c}-\mathbf{g}\right\|^2 \\ &+ \lambda_1\left\|\mathbf{u}\right\|_1 + \frac{\mu_1}{2}\left\|\mathbf{L}\mathbf{c}-\mathbf{u}+\mathbf{d}_1\right\|^2 \\ &+ \lambda_2\left\|\mathbf{v}\right\|_{1,\mathscr{S}^1} + \frac{\mu_2}{2}\left\|\mathbf{J}\mathbf{c}-\mathbf{v}+\mathbf{d}_2\right\|^2. \end{aligned} \tag{7.19}$$

We use the alternating-direction method of multipliers (ADMM) to minimize (7.19). This results in

$$\left\{\begin{array}{lcl} \mathbf{c}^{k+1} & \leftarrow & \underset{\mathbf{c}}{\operatorname{argmin}}\left\{\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\mathbf{d}_1^k,\mathbf{d}_2^k)\right\} \\ \mathbf{u}^{k+1} & \leftarrow & \underset{\mathbf{u}}{\operatorname{argmin}}\left\{\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c}^{k+1},\mathbf{u},\mathbf{v}^k,\mathbf{d}_1^k,\mathbf{d}_2^k)\right\} \\ \mathbf{v}^{k+1} & \leftarrow & \underset{\mathbf{v}}{\operatorname{argmin}}\left\{\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c}^{k+1},\mathbf{u}^{k+1},\mathbf{v},\mathbf{d}_1^k,\mathbf{d}_2^k)\right\} \\ d_1^{k+1} & \leftarrow & d_1^k + \mathbf{L}\mathbf{c}^{k+1} - \mathbf{u}^{k+1} \\ d_2^{k+1} & \leftarrow & d_2^k + \mathbf{J}\mathbf{c}^{k+1} - \mathbf{v}^{k+1}. \end{array}\right. \tag{7.20}$$

The first minimization step is a quadratic problem. Its gradient is given by

$$\begin{aligned}&\nabla\mathscr{L}_{\mu_1,\mu_2}(\mathbf{c},\mathbf{u}^k,\mathbf{v}^k,\mathbf{d}_1^k,\mathbf{d}_2^k) = \\ &\quad\left(\mathbf{H}^*\mathbf{H}+\mu_1\mathbf{L}^*\mathbf{L}+\mu_2\mathbf{J}^*\mathbf{J}\right)\mathbf{c}-\mathbf{H}^*\mathbf{g} \\ &\quad-\mu_1\mathbf{L}^*\left(\mathbf{u}^k-\mathbf{d}_1^k\right)-\mu_2\mathbf{J}^*\left(\mathbf{v}^k-\mathbf{d}_2^k\right),\end{aligned} \tag{7.21}$$

where the superscript symbol $*$ indicates the adjoint operator. We take advantage of a conjugate-gradient approach to solve this first step. The second step of (7.20) is a minimization in terms of the auxiliary variable $\mathbf{u}$. It takes the form

$$\begin{aligned}\mathbf{u}^{k+1} &= \underset{\mathbf{u}}{\operatorname{argmin}}\left\{\lambda_1\left\|\mathbf{u}\right\|_1\right. \\ &\qquad\left.+\frac{\mu_1}{2}\left\|\mathbf{L}\mathbf{c}^{k+1}-\mathbf{u}+\mathbf{d}_1^k\right\|^2\right\}.\end{aligned} \tag{7.22}$$

Therefore, the solution involves a point-wise soft-thresholding operation (*i.e.*, the proximal map of the ℓ_1 norm), which leads to

$$\mathbf{u}^{k+1} = \operatorname{prox}_{\ell_1,\lambda_1/\mu 1}\left(\mathbf{L}\mathbf{c}^{k+1}+\mathbf{d}_1^k\right). \tag{7.23}$$

The third step of (7.20) is

$$\begin{aligned}\mathbf{v}^{k+1} &= \underset{\mathbf{v}}{\operatorname{argmin}} \Big\{ \lambda_2 \left\| \mathbf{v} \right\|_{1,\mathscr{S}_1} \\ &\quad + \frac{\mu_2}{2} \left\| \mathbf{J}\mathbf{c}^{k+1} - \mathbf{v} + \mathbf{d}_2^k \right\|^2 \Big\}, \end{aligned} \tag{7.24}$$

which involves the N optimization problems indexed by m

$$\begin{aligned}\{\mathbf{v}\}_m &= \underset{\mathbf{v}}{\operatorname{argmin}} \Big\{ \lambda_2 \left\| \mathbf{v} \right\|_{\mathscr{S}_1} \\ &\quad + \frac{\mu_2}{2} \left\| \left\{ \mathbf{J}\mathbf{c}^{k+1} - \mathbf{v} + \mathbf{d}_2^k \right\}_m \right\|^2 \Big\}, \end{aligned} \tag{7.25}$$

where $\{\cdot\}_m \in \mathbb{R}^{2\times 2}$ is the mth (2×2) matrix. Solving (7.25) requires the proximal map of the Schatten 1-norm

$$\operatorname{prox}_{\mathscr{S}_1,\lambda}(\mathbf{z}) = \underset{\mathbf{u}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{u} - \mathbf{z} \right\|^2 + \lambda \left\| \mathbf{u} \right\|_{\mathscr{S}_1} \right\}. \tag{7.26}$$

It has been shown in [126] that the solution of (7.26) can be obtained by considering the singular-value decomposition (SVD)

$$\mathbf{z} = \mathbf{P}\Lambda\mathbf{Q} \tag{7.27}$$

and by rewriting Eq. 7.26 as

$$\operatorname{prox}_{\mathscr{S}_1,\lambda}(\mathbf{z}) = \mathbf{P}\operatorname{prox}_{\ell_1,\lambda}(\Lambda)\mathbf{Q}. \tag{7.28}$$

Since $\mathbf{z} \in \mathbb{R}^{2\times 2}$, the determination of SVD is easy. Consequently, the third minimization step of (7.20) is summarized as

$$\{\mathbf{v}\}_m^{k+1} = \mathbf{P}\operatorname{prox}_{\ell_1,\lambda}(\Lambda_m)\mathbf{Q}, \tag{7.29}$$

where Λ_m is the singular-value matrix of $\left\{ \mathbf{J}\mathbf{c}^{k+1} + \mathbf{d}_2^k \right\}_m$. The remaining two steps of (7.20) are updates of the augmented Lagrangian parameters. More precisely, λ_1 is chosen proportional to the variance of the phase measurement and λ_2 is adjusted according to the side information of how much correlation there is between phase and absorption. The parameters μ_1 and μ_2 are fixed to 1.

7.3.3 Experimental result

We use the heart data described in Section 7.2. The reconstructions are depicted in Figure 7.9. The background of the reconstructed phase is more homogenous than with the scheme in Section 7.2. In addition, the retrieved edges are sharper. It shows the advantage of simultaneously reconstructing phase and absorption.

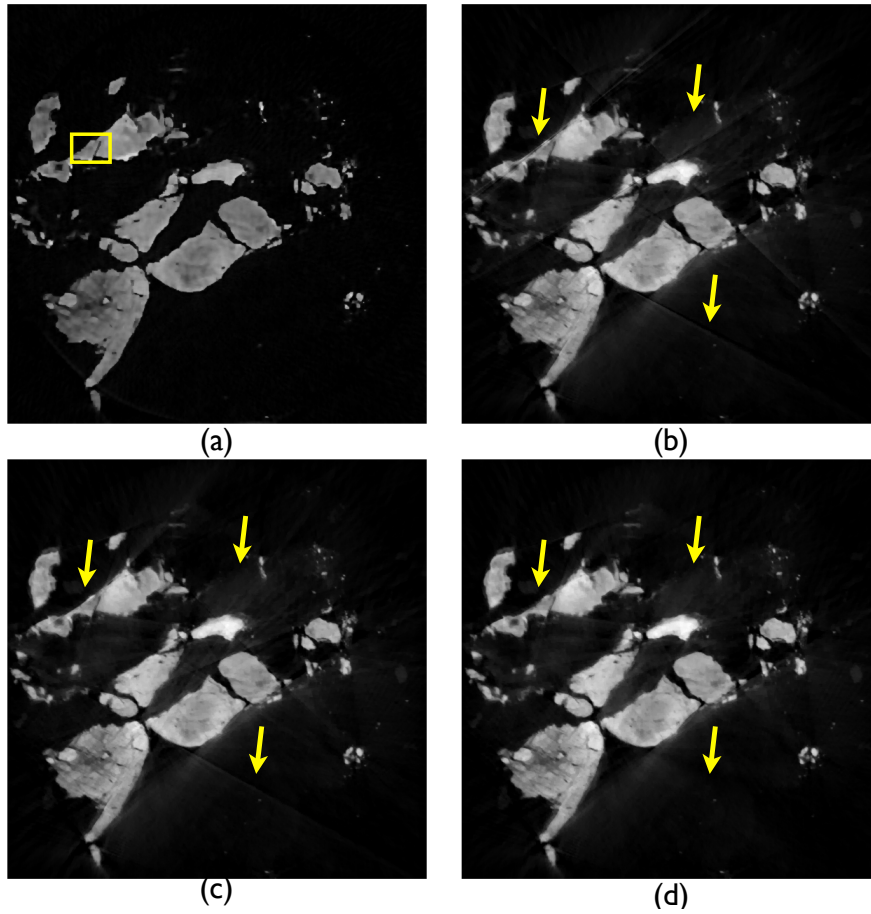


Figure 7.7: (a) The imaginary part of the refractive index of one coronal section of an aortic heart valve with two leaflets. The real part of the refractive index is reconstructed using (a) CRWN without unwrapping; (b) a weighted norm formulation whose weight is the mask given by Itoh's rule and taking advantage of a generalized ADMM; (d) a weighted norm whose weight is the Gaussian weight around the wrapped points and the use of a generalized ADMM.

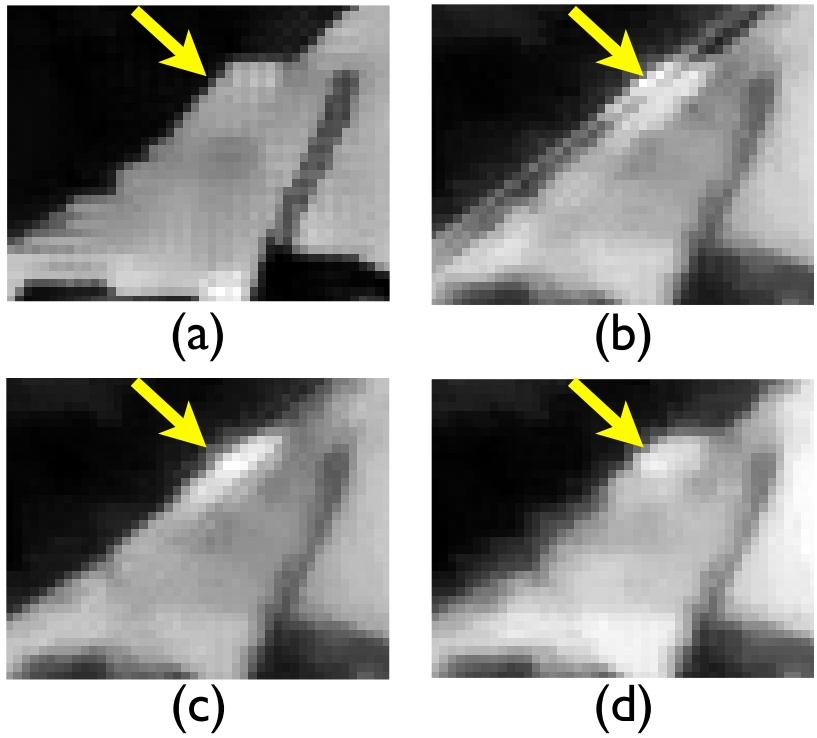


Figure 7.8: The ROI of Figures 7.7. (b) The edges are totally eliminated owing to the wrapping of the measurement. (c) Removing the wrapped points from the sinogram improves the quality of the reconstruction and the edges have a higher contrast. (d) Adding a Gaussian weight improves the contrast significantly.

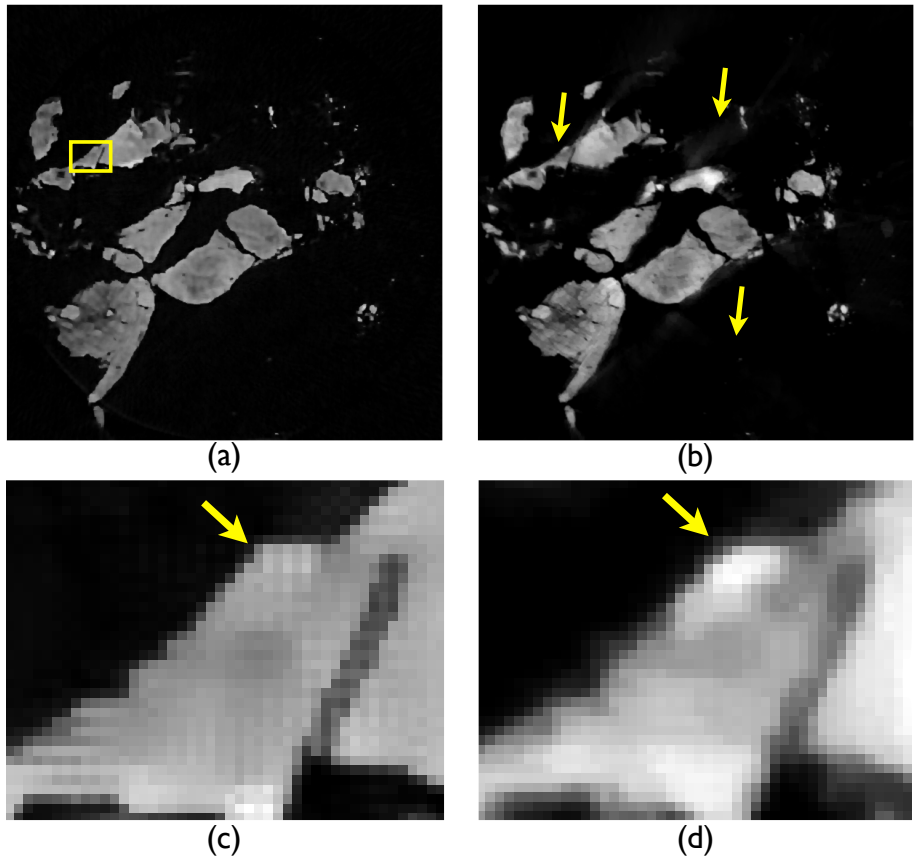


Figure 7.9: Reconstructed (a) imaginary part and (b) real part of the refractive index with one fourth number of orientations. The nice background and enhanced reconstructed edges are more visible in the depicted ROI (c) and (d).

Chapter 8

Grating-Based Radiography: Enhanced Contrast Radiographs in Mamography

[1] The radiography is not a tomography problem, but it is relevant clinically, and can benefit from the general inverse problem modeling we described in Chapter 4. In this chapter, we develop the proposed discretization and reconstruction scheme in the context of Grating-based radiography.

8.1 Phase Retrieval in Differential Phase contrast Imaging

Differential phase contrast imaging using X-ray grating interferometer is a promising tool to revolutionize conventional radiography. With an incremental modification to the conventional X-ray imaging apparatuses, this technology is able to yield three different physical contrasts of the underlying sample simultaneously: the conventional absorption contrast, the differential phase contrast (DPC), and the small-angle scattering contrast. This re-

[1] A part of this chapter has been presented in [127]

sults in much richer information than the traditional absorption-based imaging approaches [24, 110, 112, 128]. Successful experiments have been demonstrated for mammography [108] and hand imaging [129, 130] using an X-ray tube-based configuration.

The phase contrast, which is obtained by detecting the phase shifts of the X-ray waves when passing through the sample, has many advantages in imaging soft tissue (low-absorption materials) compared to the absorption contrast [12, 16, 110]. The optical properties of a tissue can be characterized by its complex refractive index $n = 1 - \alpha + i\beta$. The quantity α is the decrement of the real part of the refractive index responsible for the phase shift, whereas the imaginary part β describes the attenuation properties of the materials. At diagnostically relevant photon energies (i.e., between 10 and 150 keV), the phase shift plays a more prominent role than the attenuation for soft tissue because α is typically three orders of magnitude larger than β. The phase shift induces refraction of the X-ray wave traveling through the tissue while the grating interferometry is designed to detect the resulting overall refraction angle efficiently. In clinical applications such as mammography and computed tomography, when comparing various tissue samples, such relative differences in angular deviations are larger than the corresponding relative changes in intensity (related to the X-ray absorption). As a result, the phase contrast is expected to yield improved contrast when compared with conventional methods [131]. Moreover, for many tissue types, the phase differences drop less than their corresponding absorption differences as the X-ray energies go higher [131, 132]. Therefore, phase imaging can be applied at higher energy while still keeping the same contrast as the absorption-based approach, but with a lower dose deposition.

A potential shortcoming is that the grating interferometer does not measure the phase shift of X-rays directly, but only its first derivative. That is why this technique is called "differential" phase contrast imaging. Although the DPC image provides significant information on the edges of the sample structure, it doesn't allow a quantitative analysis of the phase profile, for instance, giving the contrast between different tissues or comparing with absorption contrast directly. Theoretically, the phase retrieval problem can be solved by a trivial one-dimensional direct integration in the spatial domain, which is equivalent to dividing by spatial frequency in the Fourier domain. However in practice, the direct integration fails to generate phase images of satisfactory quality because the noises are also cumulated during the integration. The resulting phase image suffers from stripe artifacts. The problem becomes especially severe when the image size is very large and a lower dose is used, which means a lower signal-to-noise ratio (SNR), as is common in the medical imaging.

For phase retrieval in grating-based DPC imaging, two noteworthy methods are the

bidirectional method [49] and the non-linear regularization method proposed by our group [133]. The bidirectional approach has shown to be well-suited for reducing stripe artifacts; however, it requires longer scanning time and possibly higher dose deposition. Besides, the two-scan approach requires one to rotate the gratings or the sample and to register two images precisely, which is hard to be implemented in clinical situation. The non-linear regularization method is more practical because the earned benefits are actually from the mathematic tools; therefore no additional efforts (e.g., more image acquisitions or system modifications) are needed, and it shows promising results [133]. Sperl *et al.* showed that the regularization can also be done in the Fourier domain [134].

Our major interests is to apply the differential phase contrast imaging to clinical applications, especially to mammography. In mammography, the image is usually quite large, with a typical size of (6000×4000). Such an application requires fast phase retrieval algorithms which can deliver accurate and quantitative phase information for diagnostic purpose. Our contributions are:

- The proposal of a more accurate discretization scheme for the directional derivative operator based on the B-spline calculus.

- The proposal of a better regularization strategy that penalizes discontinuities in all directions and also addresses the issue of the undetermination of the zero frequency part of the solution.

- The application of constrained regularization weighed norm (CRWN) algorithm in order to have more robust and faster phase retrieval scheme in comparison to [133]. We use the positivity and boundary condition to improve the performance of our retrieval algorithm more which is one main advantage of using CRWN.

- The experimental evaluation of the method using real data. We evaluated the algorithm with the phase contrast mammography data and directly compared the retrieved phase image with the absorption image. We demonstrated that the proposed method is able to produce high-quality phase image which can provide higher contrast of different tissues than conventional absorption-based image.

8.1.1 Methods

Phase retrieval in differential phase contrast imaging

The phase retrieval problem in DPC imaging has been described thoroughly in [133]. Briefly, the grating interferometer measures the phase shift $\varphi(x,y)$ of two phase-stepping curves (PSC), from an object scan and a reference scan without object, for each detector pixel at the spatial position (x,y). On one hand, the phase shift $\varphi(x,y)$ is linked to the refraction angle $\alpha(x,y)$ by

$$\varphi(x,y) = 2\pi \frac{d}{p_2} \alpha(x,y), \tag{8.1}$$

where p_2 is the pitch of the analyzer grating, and d is the distance between the two gratings. On the other hand, the refraction angle $\alpha(x,y)$ (namely, the DPC signal) can be expressed by [16]

$$\alpha(x,y) = \frac{\lambda}{2\pi} \frac{\partial \phi(x,y)}{\partial x}, \tag{8.2}$$

where λ is the wavelength of the x-ray photons and $\phi(x,y)$ is the target quantity (phase profile) that we want to obtain.

By combing (8.1) and (8.2), the phase retrieval problem is expressed as

$$\varphi(x,y) = g \frac{\partial \phi(x,y)}{\partial x}, \tag{8.3}$$

where $g = \frac{\lambda d}{p_2}$ is a constant that depends on the system design parameters. A direct integration then yields the phase image

$$\phi(x,y) = g \int_0^x \varphi(x',y) dx'. \tag{8.4}$$

However, due to the presence of noise, a more realistic form of the phase retrieval is

$$\tilde{\varphi}(x,y) = \varphi(x,y) + n(x,y) = g \frac{\partial \phi(x,y)}{\partial x} + n(x,y), \tag{8.5}$$

where $\tilde{\varphi}(x,y)$ denotes the actual measured signal and $n(x,y)$ represents the noise. In most cases, the direct integration fails because of the noise accumulation and lack of knowledge of the boundary conditions.

Regularization-based phase retrieval method

Retrieving the phase from (8.5) can be considered as an inverse problem and can be often solved by regularized optimization [133]. That is, to minimize the cost-function

$$J(\phi) = \|D_x\phi - \varphi\|_{l_2}^2 + \Psi(\phi), \tag{8.6}$$

where D_x is the derivative operator along the x-direction. The first term of the right side is the fidelity term and $\Psi(\phi)$ is the regularization term. In [133], we chose the regularization term to be

$$\Psi(\phi) = \lambda \left\|D_y\phi\right\|_{l_1} \tag{8.7}$$

in order to suppress the strips along the y-direction which is perpendicular to the x-direction. The l_1 norm is preferred because it generates more quantitative results compared to l_2 norm. Benefiting from the regularization term, it was shown that the regularization-based method suppresses the strip artifacts significantly. The resulting phase images provide an improved contrast-to-noise ratio (CNR) and reveal more details of the sample.

To discretize the derivative operator D_x (also D_y), the finite difference model is used which is given by

$$D_x\phi(i,j) = \begin{cases} \phi(i+1,j) - \phi(i,j) & if\ 1 \le i < N_i, \\ 0 & if\ i = N_i \end{cases} \tag{8.8}$$

where i and j are discrete coordinates (pixel coordinates) and N_i is the image size in x-direction.

The discretization model in (8.8) is very simple. In practice, our major concerns are to get quantitative information and diagnosis-level high quality image. This requirement calls for an accurate discretization model, and a corresponding algorithm for solving (8.6) and handling large data effectively. Here, we replace (8.8) with a more reliable model that uses polynomial B-splines and propose an efficient algorithm to solve (8.6) using the constrained regularized weighted-norm algorithm which will be described in the next two sections.

Discretization of the derivative operator using B-spline calculus

To formulate (8.5) as an inverse problem, a necessary step is to discretize the forward model. In this regard, we use a generalized sampling scheme and represent the object ϕ on

the principal shift-invariant space whose generating function is the tensor product of two centered B-spline functions $\beta(\mathbf{x})$,

$$\phi(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}} c_{\mathbf{k}} \beta^n(\mathbf{x}-\mathbf{k}), \tag{8.9}$$

where $\mathbf{k} = (k_1, k_2)$, $\mathbf{x} = (x_1, x_2)$, and $\beta(\mathbf{x})$ is the tensor product of two centered B-spline functions

$$\beta^n(\mathbf{x}) = \beta^n(x_1, x_2) = \beta^n(x_1)\beta^n(x_2). \tag{8.10}$$

Definition 8.1. *The centered B-spline is* $\beta^n(x) = \frac{\Delta_1^{n+1} x_+^n}{n!}$*, where* Δ_1^{n+1} *is the n times application of finite difference* $\Delta_h f = \frac{f(x+h/2)-f(x-h/2)}{h}$ *and* $x_+ = \max\{x, 0\}$.

The polynomial B-splines are well-known to offer the best cost/quality trade-off among interpolators. Since the mathematical model of our problem is based on the derivative along the direction x_1, we have

$$\begin{aligned} \varphi(\mathbf{x}) &= g \frac{\partial}{\partial x_1} \phi(\mathbf{x}) \\ &= g \sum_{\mathbf{k}\in\mathbb{Z}} c_{\mathbf{k}} \frac{\partial}{\partial x_1} \beta^n(x_1 - k_1)\beta^n(x_2 - k_2). \end{aligned} \tag{8.11}$$

. This equation leads to the matrix formulation of (8.3),

$$\boldsymbol{\varphi} = \mathbf{H}\mathbf{c}, \tag{8.12}$$

where $\mathbf{c}$ is a vector of B-spline coefficients in lexical order, $\boldsymbol{\varphi}$ is the measurement vector, and $\mathbf{H}$ is the system matrix with

$$[\mathbf{H}]_{\mathbf{l},\mathbf{k}} = g \frac{\partial}{\partial x_1} \beta^n(\mathbf{l} - \mathbf{k}). \tag{8.13}$$

Phase retrieving method

Now, we formulate the phase retrieving task as an inverse problem. We reformulate the method proposed in [67] for our problem. We aim at finding $\mathbf{c}$ where

$$\mathbf{c} = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} \{ \underbrace{\frac{1}{2} \|\mathbf{H}\mathbf{c} - \mathbf{g}\|_{\mathbf{W}}^2 + \lambda_1 \sum_i \|\{\mathbf{L}\mathbf{c}\}_i\|_1 + \lambda_2 \|\mathbf{c}\|^2}_{\mathscr{J}(\mathbf{c})} \}, \tag{8.14}$$

where the sum is computed on all the B-spline coefficients and $\{\mathbf{Lc}\}_i \in \mathbb{R}^2$ is the gradient vector of the image at position i. We constrain the solution through the convex set $\mathscr{C}$. It can be the restriction on the positivity of the refractive index and the information of the boundary of the object. Here we choose the weight $\mathbf{W}$ as identity operator along with the usual ℓ_2-norm.

Since the phase retrieval problem is ill-posed, we use the total variation (TV) regularization term to enhance the edges in the phase image. Since the null space of the imaging operator contains zero frequency, we also use the Tikhonov regularization term. In order to be consistent with the discretization scheme, the discrete gradient operator is computed using the following equations,

$$\begin{aligned}\frac{\partial f}{\partial x_1}[k_1,k_2] &= ((h_1[.,k_2] * c[.,k_2])[k1,.] * b_2[k_1,.])[k_1,k_2] \\ \frac{\partial f}{\partial x_2}[k_1,k_2] &= ((h_2[k_1,\cdot] * c[k_1,\cdot])[\cdot,k_2] * b_1[\cdot,k_2])[k_1,k_2] \end{aligned} \tag{8.15}$$

where $k_1,k_2 \in \mathbb{Z}$, $h_i[k_1,k_2] = \beta^{n-1}(k_i+\frac{1}{2}) - \beta^{n-1}(k_i-\frac{1}{2})$ and $b_i[k_1,k_2] = \beta^n(k_i)$ for $i = 1,2$.

In order to solve (8.14), we use a constrained regularized weighted norm (CRWN) algorithm with total variation regularization. We also choose parameters of the algorithm based on the suggested parameter selection scenario in Chapter 6.

Retrieval of the object boundary

In DPC image, the background of the image is meant to be zero. It does not contribute to quantitative information but only to noise when retrieving the phase. Therefore delineating the boundary of the object and masking the background out can significantly reduce the noises in the phase retrieving procedure. One advantage of grating-based imaging is the access to the information of the intensity and phase simultaneously. In order to retrieve the boundary of the object, we use both information.

We apply Canny edge detection on both measurements (absorption and phase). As we are interested in the boundary of the object, we choose the first and the last detected edge positions along each horizontal and vertical line inside the object. Afterwards, we apply median filter along each of the four boundary curves separately in order to remove some sudden jumps owing to not detecting some edges. We finally obtain two masks: one on the absorption measurements and the other on the phase information. We choose their intersection to mask the object. The constraint that is applied to the algorithm is that the reconstructed image should be zero outside this mask.

8.1.2 Experiments

Quantitative evaluation with tomographic dataset

It has been demonstrated in [133] that the regularization-based method works well with simulated data. In this section we directly evaluated the proposed method using experimental data. As a first step, we evaluated the quantitativeness of the proposed method. To establish a ground-truth for comparison, we generated a phase projection from phase tomographic dataset of a plastic phantom by forward-projection. The phase tomographic dataset was acquired at the TOMCAT beamline of the Swiss Light Source (SLS) using a Talbot interferometer. The differential phase signal was obtained by taking the one-dimension derivative of the phase projection. The retrieved phase from this differential phase projection was then compared with the ground-truth and the results were shown in Fig.8.1. By generating the differential phase signal in this way, we included the experimental noise in the image. Due to presence of the noise, quantitative recovery of the phase is very challenging, however as shown in Fig.8.1(d), the profile comparison indicates that the agreement between the retrived phase and the ground-truth is fair.

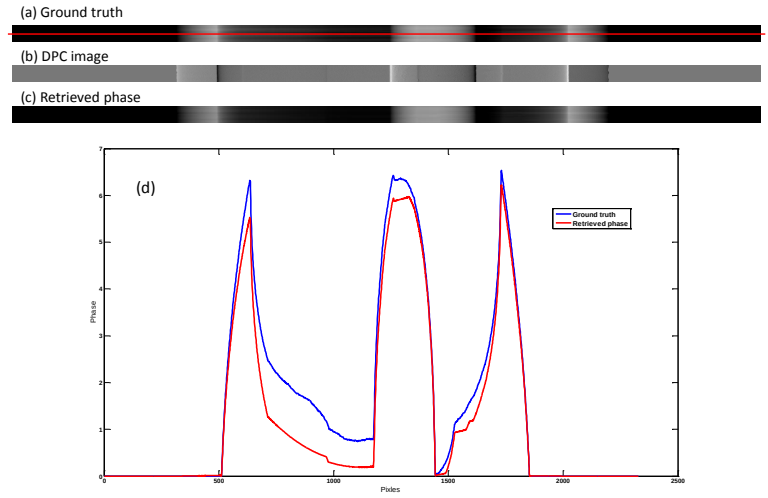


Figure 8.1: Quantitative comparison on phase tomographic data. (a) The phase projection obtained by forward-projection on a phase tomographic dataset. The projection image was used as the ground-truth in this experiment. (b) The differential phase contrast image obtained by taking the one-dimension derivative from the ground-truth. (c) The retrieved phase using the proposed method. (d) The profile (red line in (a)) comparison between the retrieved phase and the ground-truth.

Evaluation with mammographic data

In this section, we evaluated the potentials of the proposed method in medical imaging with mammographic data. The goal here is to show that the proposed method can yield phase images with useful diagnostic information, which are not accessible with conventional direct integration approach. All the data were obtained from the differential phase contrast mammography study conducted at Paul Scherrer Institut, Switzerland and Kantonsspital Baden, Switzerland [108, 135] and the details of the experimental system can be found in [108].

Biopsy sample experiment

The biopsy sample experiment was to evaluate how the proposed method works with high SNR signals. It is considered as a benchmark experiment which can evaluate the method without being too much intruded by the noise. The sample was a biopsy breast tissue with a carcinoma mass. It was first fixed in 4% formalin solution and then put into liquid paraffin for imaging in order to suppress the phase-wrapping effect at the tissue-air interface. A 32-step phase stepping scan was performed to generate a relative high quality image. In the whole breast sample experiment of next section, only 8 steps were used. The relationship between the number of steps and the noise variance in DPC image is linear [136]. Therefore the experiment here generated DPC images which were four times less noisy than in the ex-vivo whole breast experiment.

The results are shown in Fig.8.2. Visually the DPC image in Fig.8.2(b) showes the edges clearly; however, no quantitative comparison can be done from it. The direct integrated phase image in Fig.8.2(c) contained obvious stripe artifacts. Although some structures of the sample can be distinguished due to the high SNR in this case, the artifacts still degraded the image significantly, making it impossible to interpret. On the contrary, the proposed method gave a strip-artifact-free phase image. The image had a similar appearance as the absorption image (Fig.8.2(a)) but with enhanced contrast and details. Profile comparisons of the horizontal and vertical red lines in Fig.8.2(a) was shown in the Fig.8.2(e) and (f), respectively. The contrast-to-noise ratio (CNR) is used as figure of merit:

$$C = \frac{\bar{I}_s - \bar{I}_b}{\sigma_b}, \tag{8.16}$$

where $\bar{I}_s$ and $\bar{I}_b$ were the mean values of selected ROIs in Fig.8.2(e) and (f) which represent the signal and background, respectively, and σ_b is the standard deviation of the background ROI. Note that the scales of the absorption and phase images are different since they are completely different physical quantities.

It is clear that the phase image gives a higher contrast compared to absorption image. Quantitatively, for the horizontal line (Fig.8.2(e), the phase CNR is 28 while the absorption CNR is 13; and for the vertical line (Fig.8.2(f), the phase CNR is 25 while the absorption CNR is 37. The results show that the phase image indeed can provide superior contrast than the absorption image. Clinically it would help increase the diagnostic accuracy.

Ex-vivo whole breast experiment

It is worth mentioning that the whole breast DPC image was formed by stitching several smaller acquisitions together due to the limited field of view (FOV) of the current system. Stitching artifacts were hard to avoid especially because of the sample deformation happened from time to time. Moreover, the background of each acquisition was not uniform because of the systemic drifts during the scans. Therefore the whole breast DPC data was "inconsistent" in the sense that not only noise corrupted the images but also the inhomogeneity of each block. Considering those factors as well as the large image size, the direct integrated phase image was uninterpretable for radiologists. The structure features were concealed in severe strip artifacts as the common case showed in Fig.8.3(c); therefore no useful clinical information was available from the integrated phase image.

A phase retrieval example of a whole breast that would be obtained inspire of those difficulties is shown in Fig.8.3. The sample selected here contained a large tumor mass and many spiculations. Spiculations were strong indications of the existence of malignant mass. For this particular sample, we want to explore what we can gain from the phase image compared to the absorption image, therefore they are compared directly. We applied the new algorithm to the whole image as well as a selected ROI where the carcinoma located.

The stitching artifacts are clearly seen in Fig.8.3(b) and are especially obvious at the background region. Those artifacts also cause inhomogeneous background in the retrieved phase. To suppress this effect, a background mask was generated using the absorption image as described in Section.8.1.1 and applied to the retrieved phase to create a clear background as showed in Fig.8.3(d). To the best of our knowledge, this is the first time ever that an image of the line integral of the phase signal from a whole breast sample obtained with a conventional X-ray tube has been presented.

The results of the selected ROI are shown in Fig.8.4. In principle, when operating on an ROI, the boundary conditions and thus the starting wave front profile $\phi(x = 0, y)$ for the integration are unknown. The loss of the boundary information will worsen the strip artifacts and even cause shadow artifacts. An equivalent effect is the "phase wrapping" happening at the border between the skin and air. The refraction angle there is so large that the induced fringe shift is larger than one detector pixel and therefore it cannot be measured by the grating interferometer. This effect also causes the loss of information at the boundaries. Usually the phase wrapping effect can be avoid by immersing the breast in a liquid solution. However, this option is not feasible in practice and, at the end, it will also reduce the contrast in the absorption image. Nevertheless the resulting phase image

(Fig.8.4(b)) is quite good and the spiculations are better seen. Three lines (indicated in Fig.8.4) across the spiculations were selected and their profiles were given in Fig.8.4(c) (d). The profiles clearly showes that the phase image provides much higher contrast of the spiculations compared to the absorption image.

It is worth mentioning that the contrast of the calcifications is higher in the absorption image than in the phase image (see Fig.8.4(a) and (b)). This is because the complex refractive index of the calcium is similar to the soft tissue. Detecting microcalcifications is important for early breast cancer screening. To this extent, the phase image will not be able to replace the absorption image completely but be a very useful complement. The third contrast, scattering contrast, on the other hand could play an important role in microcalcification detection. The parameters only need to be be adjusted once for a given protocol and the algorithm can then be run on the whole data set.

Comparison with former method

We also compared the new algorithm with the state-of-the-art algorithm proposed in [133] and the comparison are shown in Fig.8.4(b) (c). The corresponding regularization parameter was $\lambda = 10^{-3}$ and this parameter was decided by trial and error to find the best result. Visually the new algorithm gave a superior results. The details were more visible due to fewer artifacts.

8.2 Joint absorption and phase retrieval in Grating-based X-ray radiography

A refinement upon the previous scheme is to take advantage of the high correlation between DPC (where the contrast between features is pronounced) and AC (which provides structural details) to simultaneously retrieve PC and denoise AC. We therefore propose to jointly retrieve the phase from the differential-phase image and to attenuate noise in the absorption image. This is the first attempt of the nature, to the best of our knowledge.

By employing a grating interferometer, one gains direct access to the spatial derivative

$$g(\mathbf{x}) = \frac{\partial \phi(\mathbf{x})}{\partial x_2} \tag{8.17}$$

of the phase ϕ, where $\mathbf{x} \in \mathbb{R}^2$ specifies spatial coordinates. The simplest approach to retrieve ϕ from g is line integration. Unfortunately, the phase image thus recovered exhibits

artifacts due to noise amplification at low frequencies. These problems are exacerbated in the low-dose regime typically necessitated by biological tissues, which worsens the signal-to-noise ratio.

At least two advanced phase-retrieval approaches have been proposed in the literature. The first one, called the bidirectional method, is able to reduce stripe artifacts at the cost of longer scanning times and high radiation doses; it needs the precise registration of two images, along with the mechanical rotation of gratings, two requirements that are hard to achieve [137]. Meanwhile, the second method, called nonlinear regularization, is easier to deploy because no modification of the practical setup is required. Although it shows promising results when a total-variation (TV) regularization is applied either in the spatial domain [129] or in the frequency domain [134], it sometimes fails to achieve a satisfying attenuation of the stripe artifacts and occasionally results in an insufficient contrast between regions.

We propose here an inverse-problem approach inspired by nonlinear regularization; more precisely, we complement TV by another regularizer that leverages the correlation between DPC and AC. Thus, our main contributions are

- joint denoising and retrieval of the absorption and phase contrasts, respectively;
- simultaneous use of total-variation regularizer and Schatten-norm Jacobian regularizer to attenuate reconstruction artifacts.

We develop an alternating-direction method of multipliers to achieve our goals, guided by B-spline calculus to finesse the necessary discretization schemes. Finally, we illustrate the application of our method to breast tissue from real mammography data.

8.2.1 Joint absorption and phase retrieval

Forward Model

Letting ρ denote the ideal AC image, our problem is to retrieve an approximation of ρ and ϕ from the forward model

$$\begin{cases} g_1(\mathbf{x}) &= \rho(\mathbf{x}) + n_1(\mathbf{x}) \\ g_2(\mathbf{x}) &= \frac{\partial \phi(\mathbf{x})}{\partial x_2} + n_2(\mathbf{x}), \end{cases} \tag{8.18}$$

where n_1 and n_2 represent the AC and DPC noise that is present in the measurements g_1 and g_2.

For convenience, we discretize the forward model as

$$\begin{aligned}\rho(\mathbf{x}) &= \sum_{\mathbf{k}\in\mathbb{Z}^2} c_1[\mathbf{k}]\,\beta^n(\mathbf{x}-\mathbf{k}) \\ \phi(\mathbf{x}) &= \sum_{\mathbf{k}\in\mathbb{Z}^2} c_2[\mathbf{k}]\,\beta^n(\mathbf{x}-\mathbf{k}),\end{aligned}$$

where $\mathbf{x}\in\mathbb{R}^2$ and where β^n is the tensor product of two centered B-splines of integer degree n. (Polynomial B-splines are known to offer the best cost/quality tradeoff among many interpolators.) Then, the forward discretization of Eq. 8.18 is

$$\mathbf{g} = \mathbf{H}\mathbf{c} + \mathbf{n}, \tag{8.19}$$

where $\mathbf{g}$ is the vector of measurements, and

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{x_2} \end{bmatrix}. \tag{8.20}$$

There, $\mathbf{B}$ encodes the interpolation process, and $\mathbf{D}_{x_2}$ is the directional (by convention, vertical) derivative operator. Moreover, $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2)$ collects in a single column vector each B-spline coefficient of Eq. 8.19 and Eq. 8.19. We say that $\mathbf{c}$ is the discrete representation of the object. To be consistent with Eq. 8.18, the directional derivative operator must be such that

$$\frac{\partial\phi}{\partial x_2}(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^2} c_2[\mathbf{k}]\,\frac{\partial\beta^n}{\partial x_2}(\mathbf{x}-\mathbf{k}), \tag{8.21}$$

which amounts to

$$[\mathbf{D}_{x_2}]_{\mathbf{p}} = \frac{\partial\beta^n}{\partial x_2}(\mathbf{p}) = \beta^n(p_1)\frac{\mathrm{d}\beta^n}{\mathrm{d}x_2}(p_2),, \tag{8.22}$$

where $\mathbf{p} = (p_1, p_2)$. There, $[\cdot]_{\mathbf{p}}$ specifies the $\mathbf{p}$th entry of a vector.

Then we formulate the retrieval procedure as an optimization problem similar to what we explained in Section 7.3.2. We use the Jacobian regularization along with total variation term.

Experimental Results

The purpose of this section is to provide a case where the proposed method yields absorption and phase images that contain useful diagnostic information. In particular, we want to explore the extent to which AC and DPC are complementary and/or correlated.

Data

We illustrate the potential of the proposed method on real mammography data that were obtained from a differential-phase-contrast mammography study conducted at Paul Scherrer Institut and at Kantonsspital Baden, both in Switzerland [108]. The sample contained many spiculations that clearly hint at the presence of a malignant mass.

Individual acquisitions had a limited field of view; this was counterbalanced by an imaging stitching process. Unfortunately, not only did stitching artifacts arise because of deformations in the breast tissue, but also systemic drifts during the scan resulted in a nonuniform background.

Results

We applied the proposed method to data that contains a carcinoma, as revealed in Figure 8.5. Because of the limited field of view, the integration process for a line can be determined only up to the (unknown) integration constant that specifies the boundary conditions of the wavefront $\phi(x_1 = 0, x_2)$. However, one of the roles of our regularization is to promote a consistent choice of this constant across lines, thus avoiding the worsening of stripe and shadow artifacts. We see in Figure 8.5.b that the PC image we retrieved out of DPC is quite good; the spiculations are easier to recognize there than in the AC image shown in Figure 8.5.a.

At the same time, however, the micro-calcifications are better contrasted in AC than in PC. (The early detection of micro-calcifications is an important tool for oncologists.) This difference in the quality of contrast is to be expected since the complex refractive index of calcium is similar to that of soft tissue, thus leading to an absence of contrast in the phase image. It then follows that AC and PC offer truly complementary modalities. To illustrate one possible way to take full advantage of this complementarity, we have fused AC and PC in Figure 8.5.c, where AC corresponds to the green channel and PC to the red channel of a color representation.

Comparison to the State-of-the-Art

We now focus on the region of interest (ROI) delineated in Figure 8.5. This allows for a detailed comparison of the outcome of our method with that of [129]. We present the corresponding visual results in Figure 8.6. In the PC image, we observe that the state-of-the-art method (top right) fails to remove the stripe artifacts, while our proposed method

(bottom right) is clearly more successful. In addition, a substantial amount of noise has disappeared from our version of the AC image (bottom left), at no cost in the visibility of structural details.

8.3 Conclusion

Grating-based x-ray imaging provides simultaneously absorption and differential-phase images. Thus, the (integrated) phase image—which is the quantity of interest—must be retrieved computationally. This retrieval is challenging, which has limited the use of the phase in practical radiographic applications. In this work, we propose a new iterative method to retrieve the phase image and to jointly denoise the absorption image. The present study is a proof of concept that demonstrates that it is indeed possible to improve the clarity of phase features that have a clinical relevance (*e.g.*, spiculations). At the same time, the contrast of complementary features in the absorption image (*e.g.*, micro-calcifications) is improved, too. This suggests that a wider deployment of joint retrieval absorption and phase could have a beneficial clinical impact in the early diagnosis of breast cancers.

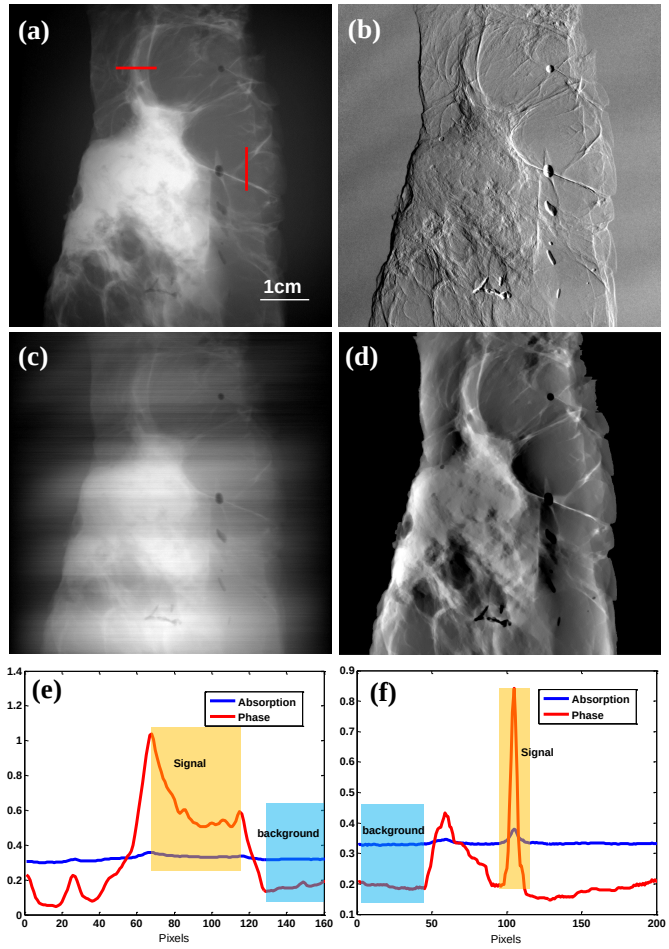


Figure 8.2: The phase retrieval results of the biopsy sample (High SNR case). (a) The absorption image; (b) the differential phase contrast image; (c) the phase image obtained by direct integration; (d) the iterative-retrieved phase image; (e) (f) are the profile comparisons of the vertical and horizontal lines in (a), respectively.

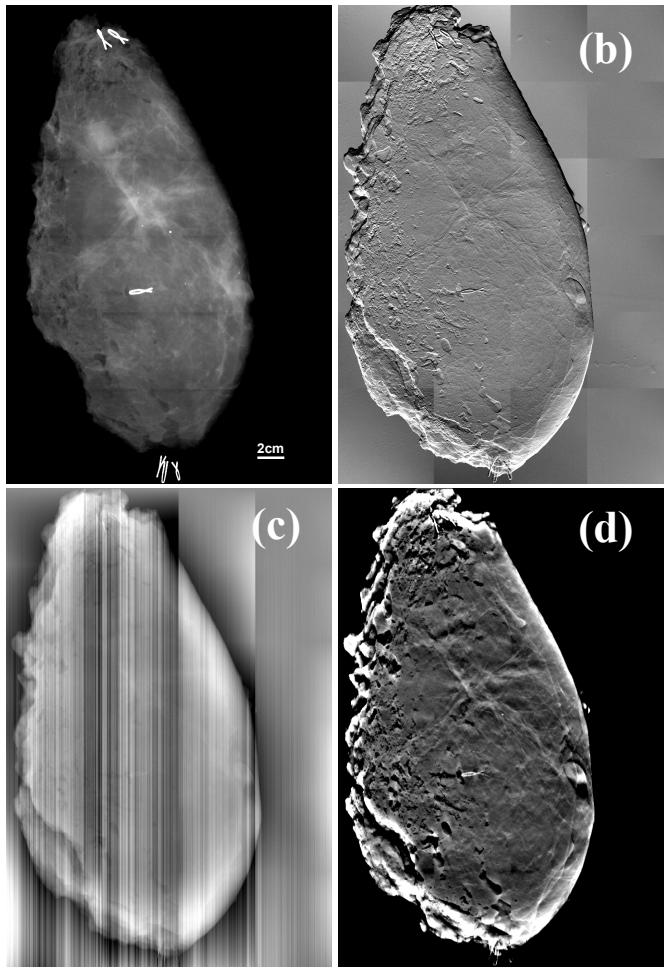


Figure 8.3: The phase retrieval results of the mastectomy breast sample (Low SNR case). (a) The absorption image; (b) the differential phase contrast image; (c) the phase image obtained by direct integration; (d) the phase image obtained by the proposed method. The images are showed in the best display window selected manually to reveal as many structures as possible.

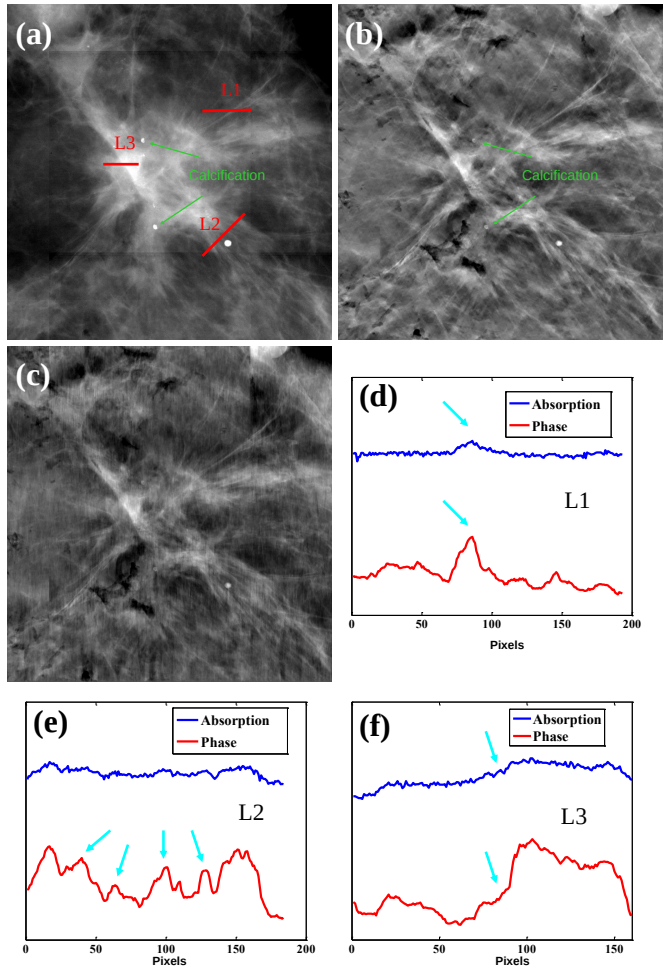


Figure 8.4: The results of a selected ROI from the whole breast phase retrieval. (a) The absorption image; (b) the phase image obtained by the proposed method, (c) the phase image obtained by our former method proposed in [133]. (d)-(e), the line profile comparisons of the lines indicated in (a).

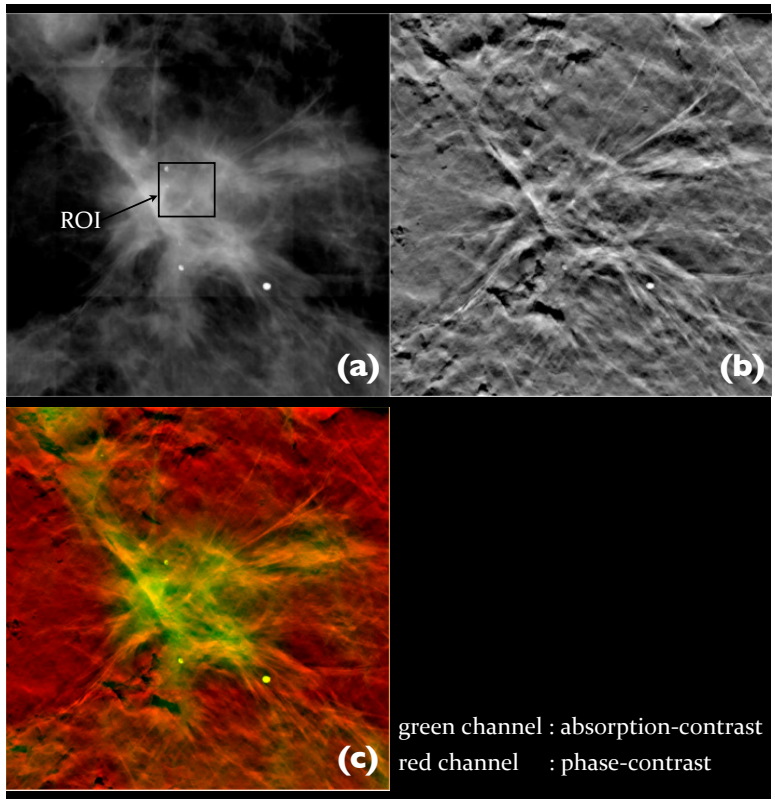


Figure 8.5: Top-left: denoised absorption image whose location corresponds to ROI in Figure 8.4. Top-right: retrieved phase. Bottom-left: multichannel visualization. Bottom-right: caption.

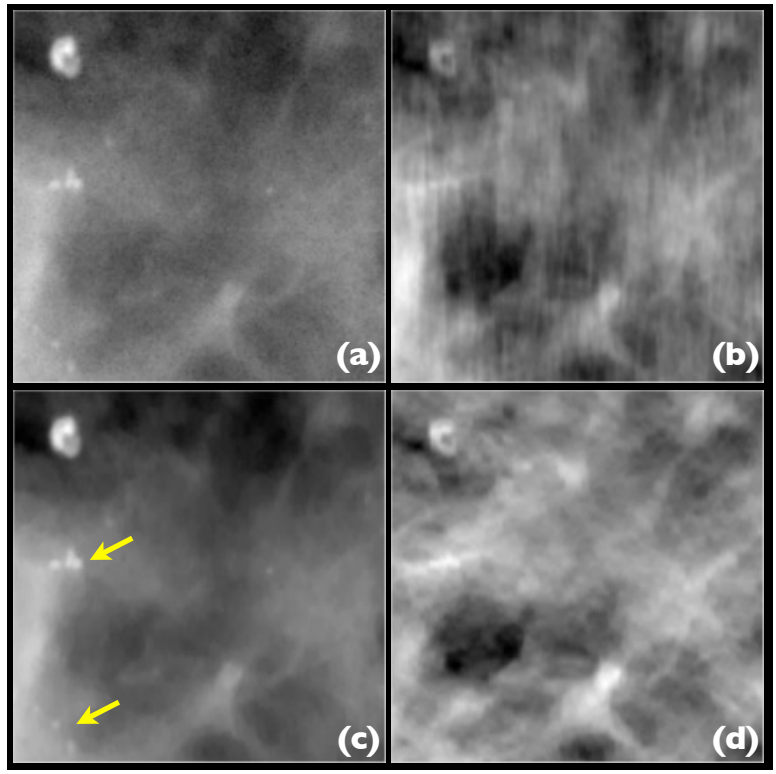


Figure 8.6: Top-left: raw absorption image whose location corresponds to the ROI in Figure 8.5. Top-right: phase retrieved by [129]. Bottom-left: absorption image after our joint denoising and retrieval process. Bottom-right: phase image that we retrieved from the differential phase.

Chapter 9

Single Particle Analysis: A Step Towards Interpretation of Atomic Models

X-ray crystallography and NMR spectroscopy allow scientists to determine 3-D structure of the biomolecules at high resolution ($1-5\mathring{A}$). These techniques are limited to small molecules. In contrast, single particle analysis (SPA) from cryo electron microscopy (cryo-EM) can provide the structure of molecules with sizes in wide range $10-150\mathring{A}$, e.g., ribosomes, proteins, and viruses. The main challenge in SPA is to improve the resolution of 3-D reconstruction to $4\mathring{A}$ or better to allow for the interpretation of its atomic structure. Single particle analysis deals with thousands of x-ray projections of identical samples (EM images) which have been taken in several micrographs. In addition, each EM images have been also modulated by the contrast transfer function of the microscope (CTF). In order to improve the resolution, it is necessary to develop more sophisticated reconstruction scheme which can also include the effect of the CTF. In this regard, we demonstrate the proposed discretization and reconstruction framework in the context of SPA. We show that using the proposed FFT-cost implementation of $\mathbf{H}^T\mathbf{H}$ improves the speed of reconstruction, significantly. The main advantage is that the computational cost of the application of the proposed kernel is independent of the number of orientations. Moreover, we include the CTF into the derivation of the kernel with no additional computational cost which results

in significant resolution improvements.

9.1 Physical Model

In transmission electron microscopy, high energy electrons (e.g. 100 kev) are emitted toward the specimen which can be highly scattered. For example, the mean free path of 120 kev electron in vitreous ice is around $2800\mathring{A}$. This is the reason why one can only image thin specimens using transmission electron microscopy. The de Broglieu wavelength associated to an accelerated electron with 100 key energy is $\lambda = 0.037\mathring{A}$. Although imaging with short wavelength particles (electron or electromagnetic waves) can provide high resolution images, the lens aberration limits the resolution. Today's highest resolution transmission electron microscopy are on the order of $1-2\mathring{A}$.

Electrons scatter in air. Consequently, the transmission electron microscope should operate in a vacuum media. However, naturally hydrated state of the biological specimens is incompatible with being in vacuum media. The cryo electron microscopes solve this incompatibility. In cryo-EM the sample is embedded in vitreous ice. This makes it possible to obtain images of fully hydrated macromolecules.

9.1.1 The Weak-Phase Object Approximation in cryo-EM

In our analysis, we assume that the object of interest is a phase object, so that it does not attenuate the transmitted wave. From the wave-optical point of view, the object introduces a phase shift $\Phi(\mathbf{y})$ in the interacted electron beam. The phase shift is linked to the x-ray transform of the real part of the refractive index of the object,

$$\Phi(\mathbf{y}) = \frac{2\pi}{\lambda}\int \alpha(\mathbf{y},z)\mathrm{d}z, \tag{9.1}$$

where λ the wavelength, α the imaginary part of the refractive index, z is the optical axis, and $\mathbf{y}$ specifies a coordinate on a plane perpendicular to the optical axis. If the incoming field is a plane wave, $u_0(\mathbf{y},z) = \exp(\mathrm{j}kz)$, then the wave function exiting the specimen is

$$u(\mathbf{y},z) = u_0 \exp(\mathrm{j}\Phi(\mathbf{y}))\,. \tag{9.2}$$

Using the Taylor series of the exponential function, one can rewrite it as

$$u(\mathbf{y},z) = u_0\left[1+\mathrm{j}\Phi(\mathbf{y}) - \frac{1}{2}\Phi^2(\mathbf{y}) + \cdots\right]. \tag{9.3}$$

Under the weak-phase object approximation, $\Phi(\mathbf{y}) \ll 1$, the wave function (9.2) is approximated by its first two taylor terms,

$$u(\mathbf{y},z) = u_0\left[1+\mathrm{j}\Phi(\mathbf{y})\right]. \tag{9.4}$$

The Equation (9.4) interprets the received wave function as the sum of an unmodified wave function and a scattered wave. The scattered wave is 90° out of phase with respect to the unmodified wave.

In the imaging mode of the transmission electron microscopy, the image plane is placed far from the object and the observation is obtained close to the optical axis. These conditions are well-matched to the Fraunhofer approximation of the diffraction theory. Consequently, the measured wave function is the Fourier transform of the wave equation (9.4). The imaged plane is placed in the back-focal plane of the objective lens. Since the applied lens is not ideal, it introduces a deflection on the transmitted wave. This aberration can be modeled as a frequency-dependent phase shift on the Fourier transform of the wave function u. Moreover, the aperture performs as a low-pass filter whose cut-off frequency depends on its size and the wavelength of the transmitted wave. Its frequency response $A(\boldsymbol{\omega})$ is

$$A(\boldsymbol{\omega}) = \begin{cases} 1 & \text{for } \|\boldsymbol{\omega}\| \leq \frac{2\pi}{\lambda}\theta_1\,, \\ 0 & \text{elsewhere}\,. \end{cases} \tag{9.5}$$

where θ_1 is the angle that corresponds to the radius of the objective aperture. Then, the wave function on the imaging plane in the Fourier domain can be written as

$$\widehat{\overline{u}}_z(\boldsymbol{\omega}) = \mathscr{F}\{u(\cdot,z)\}(\boldsymbol{\omega})A(\boldsymbol{\omega})\exp\{\mathrm{j}\gamma(\boldsymbol{\omega})\}\,, \tag{9.6}$$

where $\boldsymbol{\omega}$ the frequency coordinate, and $\gamma(\boldsymbol{\omega}) = 2\pi\chi(\boldsymbol{\omega})$. The wave aberration function χ in a polar coordinate system is

$$\begin{aligned} \chi(\omega,\phi) = -\frac{1}{2}\lambda\left[\Delta z + \frac{z_a}{2}\sin(2(\phi-\phi_0))\right]\omega^2 + \frac{1}{4}\lambda^3 C_s\omega^4 \\ \omega = \|\boldsymbol{\omega}\|\,,\ \phi = tan^{-1}(\omega_1,\omega_2)\,. \end{aligned} \tag{9.7}$$

Note that the detector measures the intensity of the received wave,

$$I(\mathbf{y}) = |\overline{u}_z(\mathbf{y})|^2\,. \tag{9.8}$$

Investigating the relation of the measured intensity with the phase function Φ and the properties of the optical systems is the subject of contrast transfer theory [138–142]. Here,

we briefly derive the mathematical relation between the object of interest and the received intensity. The Fourier transform of the wave function (9.4) is

$$\widehat{u}(\boldsymbol{\omega}) = \exp(\mathrm{j}kz)(\delta(\boldsymbol{\omega}) + \mathrm{j}\widehat{\Phi}(\boldsymbol{\omega}))\,. \tag{9.9}$$

By replacing Eq. 9.9 in Eq. 9.6, the wave function on the imaging plane in the Fourier domain is

$$\begin{aligned}\widehat{\overline{u}}_z(\boldsymbol{\omega}) &= \exp(\mathrm{j}kz)(\delta(\boldsymbol{\omega}) + \mathrm{j}\widehat{\Phi}(\boldsymbol{\omega}))A(\boldsymbol{\omega})\exp\{j\gamma(\boldsymbol{\omega})\} \\ &= \exp(\mathrm{j}kz)\left(\delta(\boldsymbol{\omega}) - \widehat{\Phi}(\boldsymbol{\omega})A(\boldsymbol{\omega})\sin(\gamma(\boldsymbol{\omega})) + \mathrm{j}\,\widehat{\Phi}(\boldsymbol{\omega})A(\boldsymbol{\omega})\cos(\gamma(\boldsymbol{\omega}))\right)\,. \end{aligned} \tag{9.10}$$

Accordingly,

$$\begin{aligned}\overline{u}_z(\mathbf{y}) &= \exp(\mathrm{j}kz) \\ &\left(1 - \left(\Phi(\cdot) * \mathscr{F}^{-1}\left\{A(\cdot)\sin(\gamma(\cdot))\right\}(\cdot)\right)(\mathbf{y}) + \mathrm{j}\left(\Phi(\cdot) * \mathscr{F}^{-1}\left\{A(\cdot)\cos(\gamma(\cdot))\right\}(\cdot)\right)(\mathbf{y})\right)\,. \end{aligned} \tag{9.11}$$

By making the weak phase object assumption, we neglect the higher order terms of the phase shift Φ so that

$$\widehat{I}(\boldsymbol{\omega}) = \delta(\boldsymbol{\omega}) - 2\widehat{\Phi}(\boldsymbol{\omega})A(\boldsymbol{\omega})\sin\gamma(\boldsymbol{\omega})\,. \tag{9.12}$$

The function $\sin\gamma(\boldsymbol{\omega})$ is called the phase contrast transfer function (CTF). According to (9.7), the rate of oscillation of the function γ increases with respect to the frequency. In summary, the measurement is linked to the convolution of the x-ray transform of the real part of the refractive index of the object with the point spread function of the microscope (aberration of the lens and the aperture limitation),

$$g(\mathbf{y}) = \frac{I(\mathbf{y}) - 1}{2} = \left(\mathscr{P}\{\alpha\}(\cdot,\theta) * h_i(\cdot)\right)(\mathbf{y})\,, \tag{9.13}$$

where h_i is the point spread function of the microscope whose Fourier transform is the microscope is CTF ($\widehat{h}_i$). Since the CTF can vary from one particle to the other, we indexed h by the particle number i. The CTF is typically a plateau followed by a rapid oscillation. Its main effect on the measured information is a combination of limiting the resolution by low-pass filtering and high-pass distortion by its rapid oscillations. Owing to the fact that the CTF has several zeros whose frequencies depend on the defocus position, typically several micrographs are imaged in different defocus positions such that the measurements

cover the whole frequency. Consequently, the CTF can be different from one micrograph to the other. In practical situation where the electron beam is partially coherent, the CTF dampens in the high spatial frequencies. It can be modeled as a factor in the CTF function:

$$\widehat{h}(\omega) = E(\omega)A(\omega)\sin(\gamma(\omega))\,, \tag{9.14}$$

where the dampen function E is modeled as an exponential function; we refer the reader to these references [143, 144] for more information.

9.2 General Overview

Single particle analysis deals with 2-D images taken from 3-D frozen-hydrated identical particles with random orientations. Note that in SPA, the radiation dose is limited in order to reduce the damage on the particles. Consequently, the measurements have extremely low signal-to-noise ratio. It makes the 3-D structure determination of the specimen a complicated and challenging process.

9.2.1 Iterative Refinement

In order to reconstruct the 3-D structure of the macromolecules, it is demanded to estimate the position and orientation of each image. The positions and orientations are estimated by random-conical tilt techniques or common-lines based approaches. The images are classified and partitioned based on their similarity in the viewing angle and then averaged over whole class to enhance the signal-to-noise-ratio (SNR). The enhanced SNR images are used to reconstruct the initial volume as shown in Figure 9.1.

Then the resolution of the initial volume is improved through a series of refinement iteration. The projections of the reconstructed volume along a given set of orientations are computed. These projections are used to reestimate the positions and orientations of the collected images. Then, the refined volume is reconstructed. The iteration continues up to convergence. The convergence is when the mean square error between the projection of the reconstructed volume along the estimated orientations and the measurements is minimized. According to the above explanation, a fast and accurate reconstruction scheme is crucial component for initial volume and refinement reconstruction, as depicted in Figure 9.2.

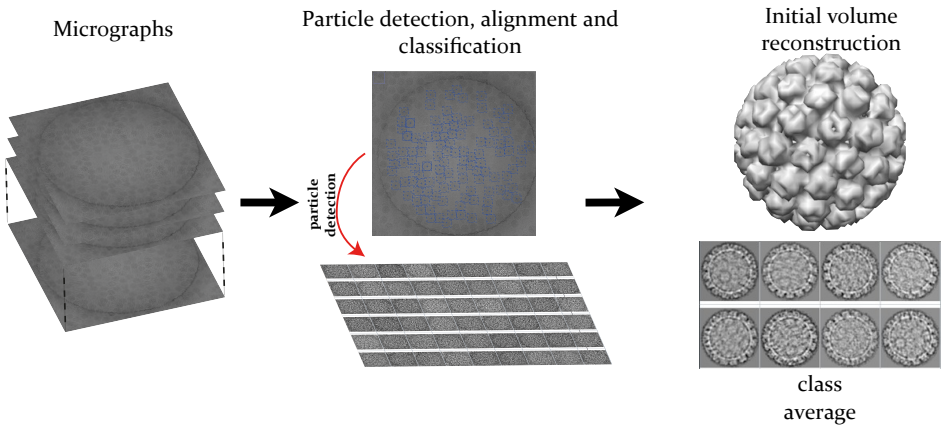


Figure 9.1: First, several micrographs are collected in single particle analysis. The particles are detected, aligned and classified. Each class average is computed to provide images with sufficient signal-to-noise-ratio. The initial volume is reconstructed.

9.2.2 Literature review

Several iterative and non-iterative reconstruction techniques have been developed in the context of SPA. Among the iterative techniques, algebraic reconstruction technique (ART) [145, 146] and simultaneous iterative reconstuction technique (SIRT) [147] are the pioneers. These techniques find the optimal volume in the sense of the least square error. The advantage of iterative techniques is that they allow us to incorporate of prior information and positivity constraints. Their disadvantage is the high computational cost which makes them extremely slow in comparison with the direct reconstruction techniques. Recently, more sophisticated iterative scheme have been developed in the context of SPA [148, 149].

From the computational time point of view, direct reconstruction techniques are more favorable. Among them, weighted back projection approach (WBP) [150–152], Gridding direct fourier reconstruction (GDFR) [153], and nearest neighbor direct inversion reconstruction algorithm (4NN) [154, 155] are used in the context of SPA. These techniques are limited to the cases where there is no major gap among the viewing directions of the

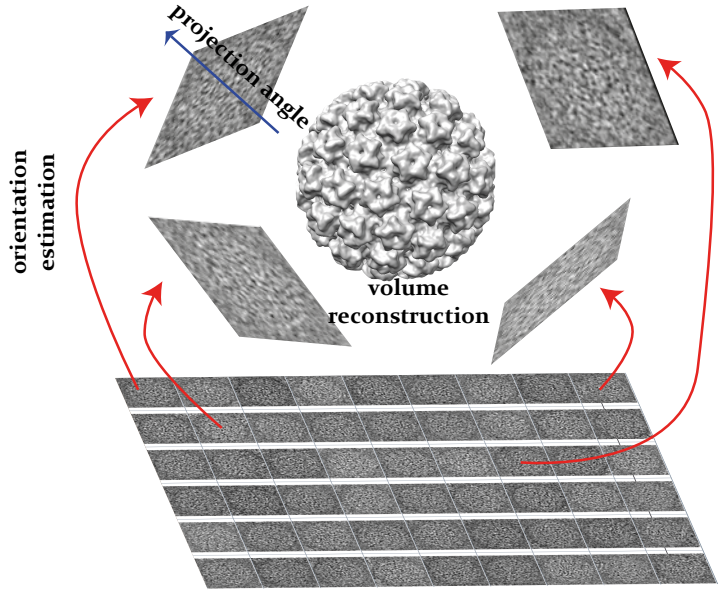


Figure 9.2: To improve the resolution of the initial reconstructed volume, several refinement iterations are performed. In each iteration, the orientations and positions of the particles are re-estimated and the volume of interest is reconstructed.

images.

Note that the measurements are not only the projections of the particles. They are also modulated by the contrast transfer function of the microscope whose functionality is highly dependent on the defocus distance. This degrades the performance of the reconstruction scheme. It is necessary to incorporate the contrast transfer function into the reconstruction procedure to enhance the resolution of the reconstruction.

In this regard, one can 1) first reconstruct the object using each defocus group (the measurements with the same defocus distance) and then correct the CTF effect by smartly combining the reconstructions, Or 2) first correct the CTF effect and then reconstruct the

object of interest. In order to correct the CTF effect, the Wiener filter is applied to find the least square solution. These techniques perform the reconstruction and the CTF correction separately [156, 157].

In contrast, Penczek *et al.* incorporated the CTF model into the forward imaging operator and then applied an ART in order to retrieve the object of interest [158]. Their simulation results suggest that the incorporation of the CTF into the forward imaging model outperforms the techniques which separate both steps. The computational time in this technique is extremely high which is its main drawback. Afterwards, the direct approach 4NN is modified such as to handle CTF correction and object reconstruction simultaneously [154, 155]. Recently, Wang *et al.* incorporate the contrast transfer function in the forward imaging operator. They formulate the reconstruction as an optimization problem without introducing any regularization. To speed-up the implementation process, they take advantage of the non-uniform FFT.

9.2.3 Contrast Transfer Function Correction

In order to improve the resolution of the determined 3-D structure of the specimen, it is necessary to correct the effect of CTF on the measurement. Since there are several zeros in the CTF, one cannot correct its effect by simple division. As discussed before, in the two-step reconstruction scenarios, the traditional way to correct the CTF effect can be a simple phase flipping or the application of a more complicated Wiener filter.

Phase Flipping

The CTF modifies the phase and the amplitude of the Fourier transform of the projection of the particle (9.14). Since the CTF oscillates between negative and positive values, a simple scheme can be correcting the phase with the following formulation

$$\widehat{\widetilde{g}}(\boldsymbol{\omega}) = \begin{cases} -\widehat{g}(\boldsymbol{\omega}) & \text{For } \widehat{h}(\boldsymbol{\omega}) \leq 0 \\ \widehat{g}(\boldsymbol{\omega}) & \text{For } \widehat{h}(\boldsymbol{\omega}) \geq 0\,. \end{cases} \tag{9.15}$$

Then $\widehat{\widetilde{g}}(\boldsymbol{\omega})$ has the correct phase information; however, the amplitude of the projection is still distorted.

Wiener Filtering

The idea behind the Wiener filter solution is to find the linear estimator that minimizes the mean square error. Let us denote the measurement corresponding to the i-th particle by g_i. It has been modulated by the CTF, $\mathbf{h}_i$. We assume that the measurements are corrupted by additive Gaussian noise and that the noise is independent of the signal. Then, the linear least square estimator in the frequency domain is in the form of

$$\widehat{\widehat{g}}_i(\boldsymbol{\omega}) = \frac{\widehat{h}_i^*(\boldsymbol{\omega})}{|\widehat{h}(\boldsymbol{\omega})|^2 + P_N(\boldsymbol{\omega})/P_g(\boldsymbol{\omega})} \widehat{g}_i(\boldsymbol{\omega}), \tag{9.16}$$

where P_N and P_g are the power spectra of noise and the projection, respectively. Note that the transcription of the proposed estimator in the object domain for merging the reconstructed objects from different defocus groups involves using the contrast transfer function in the object domain in combination with the power spectra of the noise in the object domain and the power spectra of the object.

9.3 Reconstruction framework

Here, we aim at formulating the reconstruction as an inverse problem in order to incorporate the CTF correction inside the reconstruction framework. We develop a fast and accurate reconstruction algorithm. Interestingly, one of our important finding is that the incorporation of the CTF in the forward imaging operator using the proposed framework is performed without additional computational cost for each iteration. In addition, the proposed framework allows to impose the prior information and positivity constraints on the reconstruction.

9.3.1 Discretization scheme

We assume that the orientations, positions and the corresponding contrast transfer functions are known for measured particles. We aim at reconstructing the volume using the given measurements

$$g_i(\mathbf{y}) = \left(\mathscr{P}\{f\}(\cdot, \theta_i) * h_i(\cdot)\right)(\mathbf{y}), \tag{9.17}$$

where i is the particle number. The measurements are the x-ray transform of the object of interest f along different orientations which have been also modulated by the contrast

transfer function. In order to formulate the reconstruction as an inverse problem, we first need to discretize the imaging operator. We use the discretization scheme presented in Section 3.1. The reconstruction space
$V = \{f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^3} \mathbf{c}[\mathbf{k}]\varphi(\mathbf{x}-\mathbf{k})\}$ is specified by the choice of the generating function φ. The coefficient vector $\mathbf{c}$ is considered as the discrete representation of the object.

To discretize the forward imaging operator, we recall the matrix formulation (3.6). The x-ray transform of function f using the pseudo shift-invariant property of the x-ray transform (2.17) is written as

$$\mathscr{P}\{f\}(\mathbf{y},\boldsymbol{\theta}) = \sum_{\mathbf{k}\in\mathbb{Z}^3} \mathbf{c}[\mathbf{k}]\mathscr{P}\{\varphi\}(\mathbf{y}-\mathbf{P}_{\boldsymbol{\theta}_i^{\perp}}\mathbf{k},\boldsymbol{\theta})\,. \tag{9.18}$$

In practice, the measurement g_i is acquired in the sampled points $\mathbf{y}_j = \mathbf{j}\Delta$ where Δ specifies the sampling step. For simplicity in the notation, we assume $\Delta = 1$. Based on the presented discretization scheme, the entries of the system matrix corresponds to the x-ray transform along the orientation $\boldsymbol{\theta}_i$ is

$$[\mathbf{H}_i]_{\mathbf{j},\mathbf{k}} = \mathscr{P}\{\varphi\}(\mathbf{j}-\mathbf{P}_{\boldsymbol{\theta}_i^{\perp}}\mathbf{k},\boldsymbol{\theta}_i)\,. \tag{9.19}$$

As described in the physical model, the measurements are the filtered version of the x-ray transform of the particles by the contrast transfer function of the microscope. Accordingly, the discrete model of the CTF operator can be in the form of a circulant matrix $\mathbf{K}_i$ in the discrete space. Then, the matrix form of the imaging operator is

$$\mathbf{K}_i\mathbf{H}_i\mathbf{c} = \mathbf{g}_i\,,\ \forall i = 1,\cdots,N\,. \tag{9.20}$$

In the classical approach, the reconstruction is done in two steps. First, each orientation is deconvolved to retrieve the x-ray transform of the object and then, the practical algorithms for inverting the x-ray transform is applied. The drawback of this approach is that there is no direct deconvolution step owing to the non-empty null space of the CTF operator and the noisy measurements. On the other side, using more sophisticated approaches (iterative algorithms) for doing deconvolution is very costly because there are thousands of orientations in single particle analysis. Our approach is to treat the two task simultaneously, as explained next.

9.3.2 Image Reconstruction

Linear inverse problem formulation

The discrete form of the imaging operator of the single particle analysis can be written in the form of

$$\mathbf{H}_{\text{tot}}\mathbf{c} = \mathbf{g}\,, \tag{9.21}$$

where

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_N \end{bmatrix}\,,\ \mathbf{H}_{\text{tot}} = \begin{bmatrix} \mathbf{K}_1\mathbf{H}_1 \\ \mathbf{K}_2\mathbf{H}_2 \\ \vdots \\ \mathbf{K}_N\mathbf{H}_N \end{bmatrix}\,. \tag{9.22}$$

We aim at finding $\mathbf{c}$ such that

$$\mathbf{c} = \underset{\mathbf{c}\in\mathscr{C}}{\operatorname{argmin}} \left\{ \frac{1}{2}\left\|\mathbf{H}_{\text{tot}}\mathbf{c}-\mathbf{g}\right\|^2 + \lambda\Psi(\mathbf{c}) \right\}\,, \tag{9.23}$$

where Ψ is the regularization and $\mathscr{C}$ is a convex constraint, e.g., positivity of the the object of interest. In order to minimize (9.28), we use the reconstruction algorithms ADMM-PCG and CRWN introduced Chapter 4. As we discussed in the Chapter 5, the main computational cost of each iteration is the matrix-matrix multiplication $\mathbf{H}_{\text{tot}}^T\mathbf{H}_{\text{tot}}\mathbf{c}$ that can be expanded as

$$\mathbf{H}_{\text{tot}}^T\mathbf{H}_{\text{tot}}\mathbf{c} = \sum_{i=1}^{N} \mathbf{H}_i^T\mathbf{K}_i^T\mathbf{K}_i\mathbf{H}_i\mathbf{c}\,. \tag{9.24}$$

Since $\mathbf{K}_i$ is a circulant matrix, the matrix $\mathbf{K}_i^T\mathbf{K}_i$ is also circulant. According to Theorem 5.3, if φ satisfies the radial Nyquist criteria, then the matrix-matrix multiplication $\mathbf{H}_i^T\mathbf{K}_i^T\mathbf{K}_i\mathbf{H}_i\mathbf{c}$ can be calculated using a discrete convolution $(\mathbf{c} * \text{kernel})$ where $*$ is a convolution operator. This is our crucial observation for our approach. The kernel is computed with the use of (5.29) and is given by

$$\mathbf{y}_{\theta_i}[\mathbf{k}] = (\mathscr{P}\varphi(\cdot,\theta_i) * k(\cdot) * \mathscr{P}\varphi(-\cdot,\theta_i) * k(-\cdot))(\mathbf{P}_{\theta_i^\perp}\mathbf{k})\,. \tag{9.25}$$

The fact that the matrix $\mathbf{H}_{\text{tot}}^T\mathbf{H}_{\text{tot}}$ is poorly conditioned, which makes the inverse problem challenging.

Conjugate-gradient based linear inverse problem formulation

The measurements are modulated by the contrast transfer function of the microscope. A simple CTF correction is flipping the phase of the measurement in the frequency space in which the CTF has negative values. The flipped phase CTF and the measurement vector are denoted by $\mathbf{K}_i^+$ and $\mathbf{g}_i^+$. Then, the matrix formulation of the of the forward model is

$$\mathbf{K}_i^+ \mathbf{H}_i \mathbf{c} = \mathbf{g}_i^+ \,,\ \forall i = 1, \cdots, N\,. \tag{9.26}$$

Since $\mathbf{K}_i^+$ is a semi-positive definite operator, one can reformulate (9.26)

$$\mathbf{H}_i^T \mathbf{K}_i^+ \mathbf{H}_i \mathbf{c} = \mathbf{H}_i^T \mathbf{g}_i^+ \,,\ \forall i = 1, \cdots, N\,. \tag{9.27}$$

We then propose the volume of interest is reconstructed by minimizing the modified cost function,

$$\mathscr{J}(\mathbf{c}) = \left\{ \underbrace{\frac{1}{2}\mathbf{c}^T \mathbf{A}\mathbf{c} - \mathbf{c}^T \sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{g}_i^+}_{\mathscr{D}(\mathbf{c})} + \lambda \Psi(\mathbf{c}) \right\}, \tag{9.28}$$

where

$$\mathbf{A} = \sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{K}_i^+ \mathbf{H}_i\,. \tag{9.29}$$

Now, the matrix $\mathbf{A}$ is positive definite, then the data fidelity term is a convex function. By using a convex regularization, one can use any convex optimization to reconstruct the volume. The Gradient of the data fidelity term $\mathscr{D}(\mathbf{c})$ is

$$\nabla \mathscr{D}(\mathbf{c}) = \mathbf{A}\mathbf{c} - \sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{g}_i^+\,. \tag{9.30}$$

Note that the condition number of the system matrix $\mathbf{A}$ is better than the condition number of $\mathbf{H}_{\text{tot}}^T \mathbf{H}_{\text{tot}}$. This yields better convergence speed. If φ satisfies the radial Nyquist criteria, then one can implement $\mathbf{A}\,\mathbf{c}$ using FFT s. The corresponding kernel is

$$\mathbf{y}[\mathbf{k}] = \sum_i \mathbf{y}_{\theta_i}[\mathbf{k}] = \sum_i (\mathscr{P}\varphi(\cdot, \boldsymbol{\theta}_i) * k^+(\cdot) * \mathscr{P}\varphi(-\cdot, \boldsymbol{\theta}_i))(\mathbf{P}_{\theta_i^\perp}\mathbf{k})\,. \tag{9.31}$$

We use ADMM-PCG to minimize the cost function (9.28) described in Algorithm 2.

9.4 Experimental result

To validate the proposed reconstruction scheme, we conduct two experiments in the context of single particle analysis. The first one is on a simulated dataset and the second one is on a real dataset.

9.4.1 Implementation details

In all experiments, we use total variation as regularization. The ADMM-PCG algorithm is applied to reconstruct the volume by minimizing the cost function (9.28). The positivity constraint is imposed on the solution. A kaiser Bessel window function with radius support 4 and parameters $\alpha = 19$ and $m = 2$ is used as generating basis function. We run the experiment for twenty outer iterations and three inner iterations.

We compare the reconstruction using our reconstruction framework with the standard Xmipp solution. The software Xmipp is one of the well-known software in the context of electron microscopy. It uses an optimized version of Fourier gridding techniques. We remark that our proposed framework has been also completely implemented in the Xmipp software.

9.4.2 Performance metric

To evaluate the performance of the proposed framework, we first compare the reconstructions visually by looking at the slices side by side. For the simulated dataset, we measure the resolution of the reconstructions using Fourier shell correlation. The Fourier shell correlation is defined by

$$\mathrm{FSC}(\mathbf{f}_1, \mathbf{f}_2) = \frac{\sum_{\mathbf{k}} \widehat{\mathbf{f}}_1[\mathbf{k}] \cdot \widehat{\mathbf{f}}_2^*[\mathbf{k}]}{\sqrt{\sum_{\mathbf{k}} |\widehat{\mathbf{f}}_1[\mathbf{k}]|^2 \cdot \sum_{\mathbf{k}} |\widehat{\mathbf{f}}_2^*[\mathbf{k}]|^2}}, \tag{9.32}$$

where $\widehat{\mathbf{f}}$ is the discrete Fourier transform of the volume $\mathbf{f}$. For more information, we refer the reader to [159].

9.4.3 Simulation-based analysis

The simulated dataset is produced as follows: The 3D density map of the 50S ribosome subunit bound with ObgE with pixel size $1.5\mathring{A}$ is generated using the structure 4csu.pdb

(Figure 9.3 (a-c)). The size of the density map is $256 \times 256 \times 256$. A set of 2-D images are obtained by computing the x-ray transform of the density map along 300 random orientations using Xmipp software. Each image is modulated by a realistic contrast transfer function and then corrupted by additive Gaussian noise. Example of collected images are depicted in Figure 9.3 (d).

The volume is reconstructed using Xmipp with Fourier gridding and also with our proposed framework. The reconstructions are shown in Figure 9.4. The Fourier shell correlation is evaluated for the reconstructed volumes. The results show that the 0.5 FSC resolution of the reconstruction using ADMM-PCG ($3.57\mathring{A}$) is around 1.5 times better than the Fourier-based reconstruction ($5.55\mathring{A}$) when we assume the measurements are only the x-ray transform of the particle. Our method improves the performance even further when we consider CTF (around two times improvement of the resolution).

9.4.4 Real data experiment

We then run an experiment on a real SPA dataset. The dataset corresponds to the Bovine papilloma virus type 1. The human papilloma virus is a Baltimore Class I virus. The genome is double stranded circular DNA surrounded by an icosahedral capsid; it has 60 fold symmetry. The orientations, positions, and microscope contrast transfer functions are estimated by running the standard workflow of the Xmipp software. We considered three micrographs with three different defocuses. There are 245 EM-images. The projections are down-sampled by factor 2 and the volume is with $2.474\mathring{A}$ resolution. The size of the volume is $256 \times 256 \times 256$. Since the virus has 60 fold symmetry, the projection measurements are replicated 60 times considering its symmetry coordinates.

The volume is reconstructed using our method and the Xmipp traditional technique. The results are depicted in Figure 9.5. It suggests that the direct reconstruction frameworks performs poorly in high resolution reconstructions.

9.4.5 Implementation remarks

Symmetry volumes

In the case of volumes which is M-fold symmetry, one can replicate the measurements along its symmetry coordinates. This way of implementations increases the computational

cost of the kernel and the application of the adjoint of the x-ray transform on the measurement vector by factor M. Let us denote the volume by $\mathbf{v}$ and represent its symmetry properties through a set of rotation operators $\mathbf{R}_i$, $\forall i = 1, \cdots, M$; equivalently,

$$\mathbf{v} = \mathbf{R}_i \mathbf{v}, \ \forall i = 1, \cdots, M. \tag{9.33}$$

In order to improve the performance of the calculation of the kernel and the application of the adjoint of the x-ray transform on the measurement vector, Instead, one can compute them as follows:

$$\begin{aligned} \tilde{\mathbf{b}} &= \sum_{j=1}^{M} \mathbf{R}_j \mathbf{b} \\ \mathbf{r}_{\text{tot}} &= \sum_{j=1}^{M} \mathbf{R}_j \mathbf{r}, \end{aligned} \tag{9.34}$$

where $\mathbf{b} = \sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{g}_i^+$ and $\mathbf{r}$ is the kernel corresponds to the given set of orientations.

Application of the adjoint of the x-ray transform on the measurement vector

In the proposed framework, the main computational cost is the pre-computation of the kernel and the application of the adjoint of the x-ray transform on the measurement vector, $\sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{g}_i^+$. In order to apply the adjoint operator on the measurement vector, we use the algorithm given in Section 3.1.2. This is computationally heavy. To reduce its computational time, we reinvestigate the application of the adjoint of the x-ray transform for specific orientation, $\boldsymbol{\theta}$ on a given measurement vector $\mathbf{g}$. The operation can be written in the form of

$$\begin{aligned} \tilde{\mathbf{f}}[\mathbf{k}] &= \sum_{\mathbf{j} \in \mathbb{Z}^2} \mathbf{g}[j_1, j_2] \mathscr{P}\{\varphi\}(j_1 - \langle \mathbf{r}_1, \mathbf{k} \rangle, j_2 - \langle \mathbf{r}_2, \mathbf{k} \rangle, \boldsymbol{\theta}) \\ &= \left(\sum_{\mathbf{j} \in \mathbb{Z}^2} \mathbf{g}[j_1, j_2] \mathscr{P}\{\varphi\}(j_1 - i_1, j_2 - i_2, \boldsymbol{\theta}) \right)_{i_1 = \langle \mathbf{r}_1, \mathbf{k} \rangle, i_2 = \langle \mathbf{r}_2, \mathbf{k} \rangle}. \end{aligned} \tag{9.35}$$

In order to speed up its computational cost, one can first convolve the measurement vector $\mathbf{g}$ with the kernel $\mathscr{P}\{\varphi\}$, computes its values on a fine grid, and stores it in a look-up table. Then, its values on points $(\langle \mathbf{r}_1, \mathbf{k} \rangle, \langle \mathbf{r}_2, \mathbf{k} \rangle)$ are computed using a simple interpolation procedure.

9.5 Conclusion

We formulated the reconstruction from several noisy micrographs in single particle analysis as a constrained regularized optimization. We were able to directly include the contrast transfer function in the system matrix without any extra computational cost. The experimental results suggested that our approach brings a significant improvement in the quality of the reconstruction. Our framework also provided an important step toward the application of SPA for the atomic interpretation of macromolecular models. The corresponding algorithms have been implemented in Xmipp.

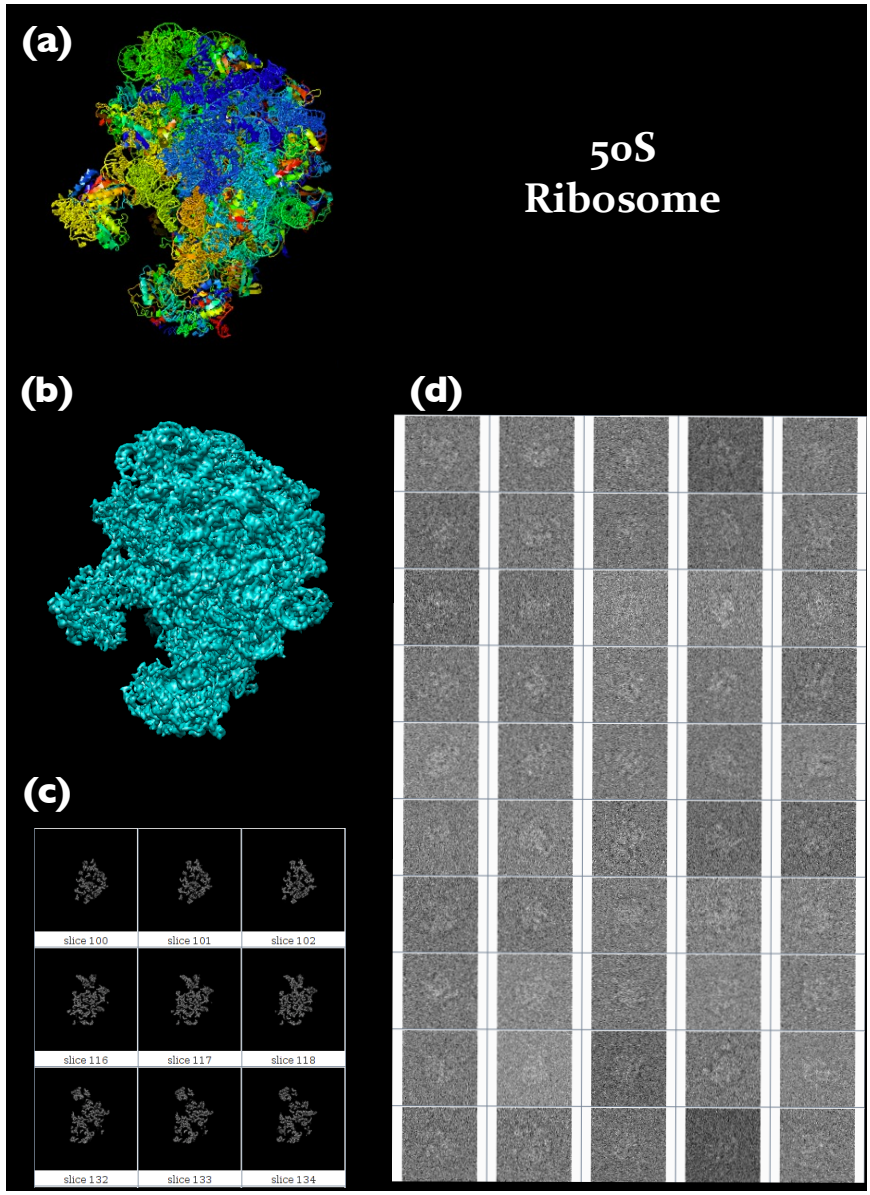


Figure 9.3: (a) 4CSU, 50S ribosome subunit count with ObgE with 1.5$\mathring{A}$, (b) 4CSU density map, (c) some slices of the volume, (d) the x-ray transform of the density map modulated by a realistic CTF is computed. It then corrupted by additive Gaussian noise.

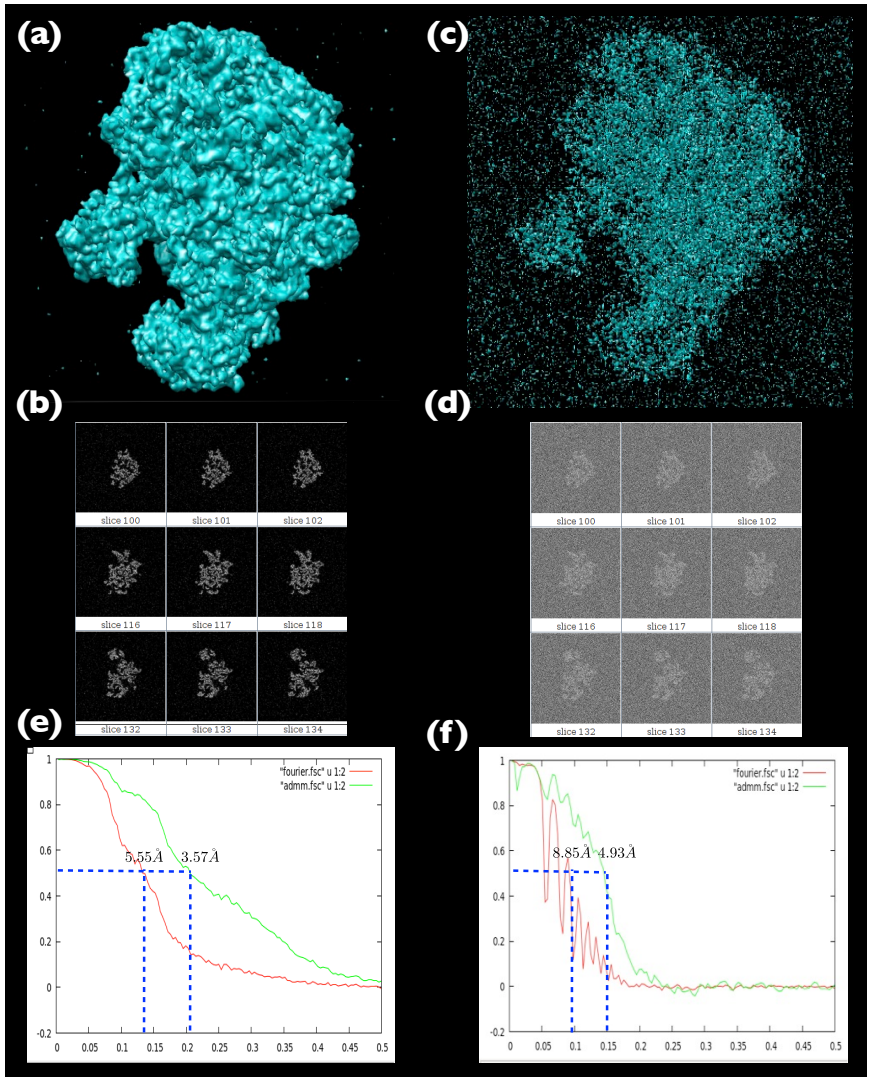


Figure 9.4: The reconstructed volume using ADMM-PCG and its slices are depicted in (a) and (b), respectively. The correspondent images using Xmipp technique are shown in (c) and (d). The Fourier shell correlation (e), (f) shows the significant improvement of using ADMM-PCG in comparison with the Xmipp technique in two different scenarios: 1) there is no CTF, 2) there is a CTF.

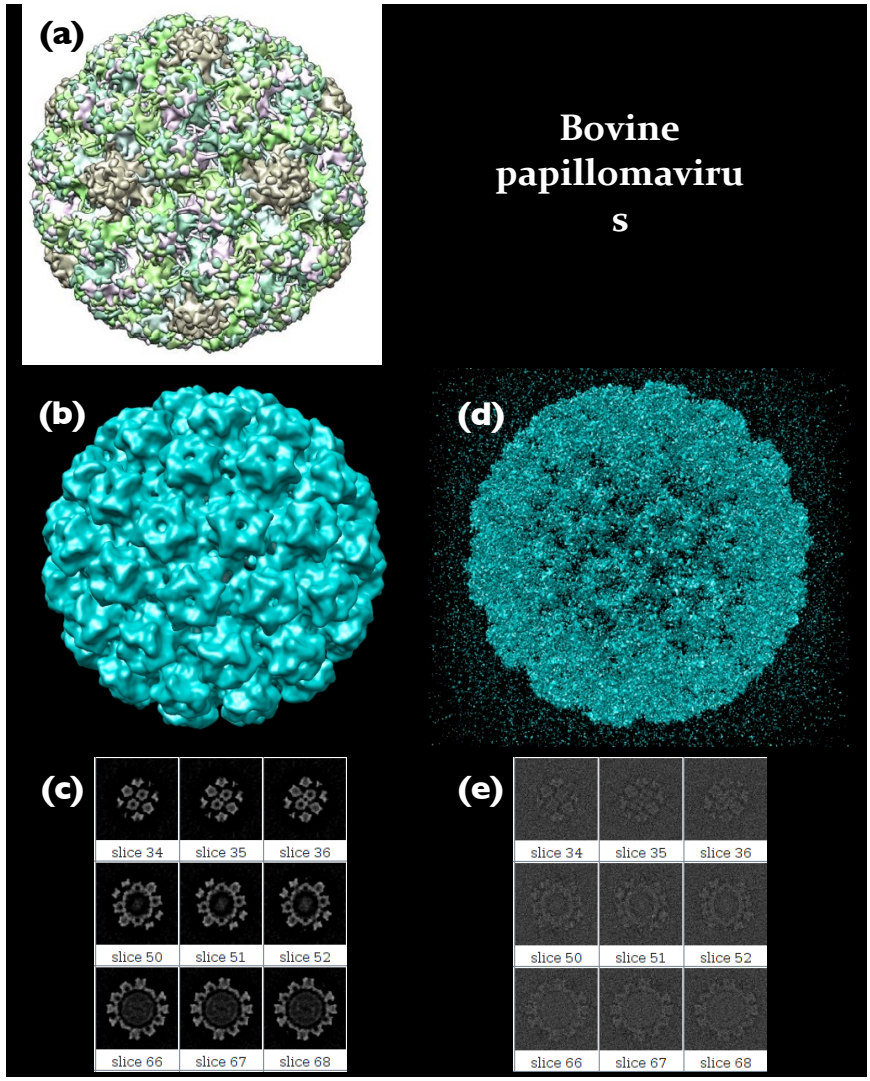


Figure 9.5: The Bovine papilloma virus type 1 (a). The reconstructed volume using ADMM-PCG and its slices are depicted in (b) and (c), respectively. The correspondent images using Xmipp technique are shown in (d) and (e).

Chapter 10

Conclusion

In this thesis, we developed high performance and competitive reconstruction frameworks in the context of straight-ray imaging modalities. Here, we first summarize the main research directions. We then present our results in the context of grating-based imaging (GI) and single particle analysis (SPA). Then, potential areas of interest for future research related to our work are listed in the last section.

10.1 Summary of Results

10.1.1 Reconstruction Framework

Our approach is based on 1) a rigorous discretization of the forward model using a generalized sampling scheme; 2) a variational formulation of the reconstruction problem; and 3) iterative reconstruction algorithms that use the alternating-direction method of multipliers.

- **Discretization Scheme:** In order to formulate the reconstruction as an inverse problem, one requires to discretize the imaging operator. In this regard, we specified the reconstruction space as $\mathbf{V} = \{f(\mathbf{x}) = \sum_{\mathbf{k}} c[\mathbf{k}]\varphi(\mathbf{x}-\mathbf{k})\}$. We considered the coefficient vector $\mathbf{c}$ as discrete representation of the object of interest. The discretization problem was summarized in the choice of basis function φ. We investigated box splines and kaiser-Bessel window functions as two main families.

The choice of the box spline generator should be determined by computational and approximation theoretic considerations. Basis functions with larger support and smoothness usually offer better approximation quality, but they also require more computations. This suggests the possibility of a tradeoff between approximation order and the density of the reconstruction grid. In particular, it demonstrated that it is computationally advantageous asymptotically to switch to a higher-order basis function than to increase the sampling rate. Tensor-product B-splines constitute a preferred set of basis functions because they are made up from univariate B-splines building blocks which are widely studied and efficient to evaluate. We should note, however, that the present box spline framework includes other non-separable basis functions with increased isotropy (e.g., Zwart-Powell box spline) and same approximation order, but lower polynomial degree and smaller support than their tensor-product counterparts.

Kaiser-Bessel window functions (KBWF) are isotropic. The projections of isotropic functions are independent of the viewing angle. Therefore, they are attractive candidates as generating functions of principal shift-invariant spaces for discretizing the imaging operators. The generalized Kaiser-Bessel window function is widely used in the context of straight-ray imaging. There are several parameters for KBWFs that should be adjusted. In this thesis, we proposed a measure to determine the performance of a basis function for the discretization scheme. Furthermore, we suggested a method to optimize the parameters of the KBWF based on this measure. By numerical experiments, we confirmed that using the proposed method improves the performance of the discretization scheme.

For a fixed support, B-spline functions have a better order of approximation than Kaiser-Bessel window functions (KBWFs). We showed that, by adjusting the taper parameter of KBWF using the proposed approximation-theoretic framework, these functions perform almost as well as B-splines. In two-dimensional tomography or three-dimensional tomography with fixed rotation axis, B-splines are preferable owing to their order of approximation. In addition, the separability of the tensor product of cubic B-spline functions allows one to decompose the different three-dimensional reconstruction in parallel-beam geometry with fixed rotation axis into some easier two-dimensional and one-dimensional subproblems. It results in the development memory-efficient reconstruction framework. In the three-dimensional problem with random orientations, the implementation of the x-ray transform using B-splines is too complicated and it is therefore more practical to use KBWFs.

Note that, the approach is applicable with minor adjustment to non-parallel geometries as long as the projections are line integrals or obtained by taking "point" measurements (i.e., ideal samples) of the Radon transform, or, with a very slight extension, that all the rays hitting one detector (pixel) are parallel. Since the Radon transform of the B-spline/box spline is available to us in closed-form, the handling of a non-parallel geometry then essentially amounts to a proper bookkeeping of the angles: For every ray angle in the non-parallel geometry, the proper ray direction in the parallel geometry can be looked up. The hypothesis of pure line integrals is implicit to all the discretization methods that we are aware of.

- **Variational formulation and iterative reconstruction:** We formulated the reconstruction of straight-ray tomograms as a constrained optimization problem. To improve the quality of the reconstruction, we took advantage of total-variation regularization and its higher-order variants. In addition, the prior information on the support and the positivity of the refractive index were both considered. We showed that side information such as the support-related constraints and positivity of the refractive index can significantly improve the quality of reconstruction. We introduced iterative algorithms that use the alternating direction method of multipliers to solve our constrained regularized reconstruction problem. In addition we derived an important practical twist which is the introduction of a problem-specific preconditioner. We showed that it significantly speeds up the quadratic optimization step of the algorithm. One can obtain a reasonable reconstruction in few number of iterations using the proposed reconstruction scheme. This is an important issue in practical applications. In order to speed-up the reconstruction even further, we reduced the computational cost of each step by exploring the structure of $\mathbf{H}^T\mathbf{H}$. We showed that under some conditions the operator $\mathbf{H}^T\mathbf{H}$ is indeed a digital convolution in which one can take advantage of FFT for implementation.

The two challenging applications to which we applied the methods of our framework are grating-based x-ray imaging (GI) and single-particle analysis (SPA).

10.1.2 Grating-based X-ray Imaging

Up to now, in-vivo tomography with grating interferometry faces the challenge of large dose deposition, which potentially harms the specimens e.g., in small rodent scanners. To reduce the total scanning time, we first demonstrated the reconstruction framework

in GI. We presented experimental results to validate the proposed discretization method and corresponding iterative technique. Our finding confirms that the proposed method is quite competitive for solving TV-regularized problems. Moreover, our method allows for a substantial dose reduction while preserving the image quality of FBP-type method. This is a crucial step towards the diffusion of DPCI in medicine and biology. We then improved the reconstruction framework even further by considering the correlation between the phase and absorption measurements. We added Jacobian-type regularization to simultaneously reconstruct phase and absorption. The experimental results confirmed the power of our method.

Unlike DPC tomography where the phase information can be recovered effectively by reconstruction algorithm, retrieving phase image from DPC projection remains challenging and limits the exploration of the advantages of the phase information in radiographic applications. We used the proposed iterative algorithm for differential phase contrast imaging. The algorithm utilized a novel discretization model of differential operator using B-spline calculus. The algorithm was evaluated with breast biopsy and mastectomy samples. The present study demonstrated that DPC signal is indeed capable of providing higher contrast in clinical relevant features, like the spiculation etc. These results could help to improve breast cancer diagnosis.

In addition, we proposed a new iterative method to retrieve the phase image and to jointly denoise the absorption image. The present study is a proof of concept that demonstrates that it is indeed possible to improve the clarity of phase features that have a clinical relevance (*e.g.*, spiculations). At the same time, the contrast of complementary features in the absorption image (*e.g.*, micro-calcifications) is improved, too. This suggests that a wider deployment of joint retrieval absorption and phase could have a beneficial clinical impact in the early diagnosis of breast cancers. Our algorithms have been implemented in the TOMCAT laboratory of the Paul Scherrer Institute.

10.1.3 Single Particle Analysis

In the context of near-atomic-resolution SPA, we need to cope with hundreds or thousands of noisy projections of macromolecules onto different micrographs. Moreover, each projection has an unknown orientation and is blurred by some space-dependent point-spread function of the microscope. Consequently, the determination of the structure of a macromolecule involves not only a reconstruction task, but also the deconvolution of each projection image. We formulated this problem as a constrained regularized reconstruction. We were able to directly include the contrast transfer function in the system matrix with-

out any extra computational cost. The experimental results suggested that our approach brings a significant improvement in the quality of the reconstruction. Our framework also provided an important step toward the application of SPA for the atomic interpretation of macromolecular models. The corresponding algorithms have been implemented in Xmipp.

10.2 Outlook

The research presented in this thesis opens several interesting directions for future investigation. Some of them are listed below.

- **Optimal scenario for radiation dose reduction:** The radiation dose reduction can be achieved by either a reduction in the intensity of the photons or in the number of projection angles. In the context of grating-based imaging, we investigated the scenario in which the number of projection angles is reduced. The question is what the optimal strategy to reduce the radiation dose is? What is the optimal intensity of the photons and the number of projection angles while using the proposed reconstruction frameworks in the context different straight-ray imaging modalities? It is worthwhile to investigate this problem mathematically and also to validate it by conducting some real experiments.

- **Optimal regularization parameters:** One of the main challenges in the regularized formulation of the reconstruction is the adjustment of regularization parameters. In the context of differential phase contrast tomography, we presented a candidate for the parameters. It is worthwhile to extend the proposed framework to straight-ray imaging modalities. In addition one can propose a new candidate for the optimal regularization parameter in the context of straight-ray imaging.

- **Generating function φ with radial Nyquist criteria:** We showed that if the basis function φ satisfies the radial Nyquist criteria, then $\mathbf{H}^T\mathbf{H}$ is a digital convolution. We then investigated B-splines and Kaiser-Bessel window functions as the ones which approximately satisfy the necessary condition. It is an interesting question to consider the shift invariant space introduced by a generation function φ which satisfies the necessary condition. The goal of this research direction is to investigate theoretically the space generated by a function φ that is a band-limitted function.

- **Extension to diffraction tomography:** We proposed a discretization scheme using generalized sampling. We showed that in the context of parallel beam imaging,

$\mathbf{H}^T\mathbf{H}$ is a digital filter under radial Nyquist condition. The next step is to extend this philosophy in the context of diffraction tomography. The derivation of efficient solutions to these problems are important in the context of optical tomography, based on the results obtained in this thesis.

- **Speeding up the matrix-vector multiplication of $\mathbf{H}^T\mathbf{g}$ (back-projection operator):** In order to speed-up the proposed reconstruction framework, we derived the necessary conditions to reduce the computational cost of the matrix-vector multiplication $\mathbf{H}^T\mathbf{H}$ in each iteration. Since the back projection of the measurement vector is precomputed and then is used in each iteration, we didn't investigate this operator in this thesis. It is likely that future research will investigate to reduce the computational cost of $\mathbf{H}^T\mathbf{g}$.

- **Orientation estimation and initial volume reconstruction in single particle analysis:** Single particle analysis deals with thousands of projection images of identical particles with random orientations. In order to reconstruct the 3-D structure of the macromolecules, it is demanded to estimate the position and orientation of each image. The positions and orientations are estimated by random-conical tilt techniques or common-lines based approaches. The images are classified and partitioned based on their similarity in the viewing angle and then averaged over whole class to enhance the signal-to-noise-ratio (SNR). The enhanced SNR images are used to reconstruct the initial volume. It is advantageous to involve the proposed reconstruction framework in orientation estimation and initial volume reconstruction.

Bibliography

[1] M. C. Teich and B. E. A.Saleh, *Fundamentals of photonics*, Wiley Interscience, 1991.

[2] D.Attwood, *Soft X-rays and extreme ultraviolet radiation: principles and applications*, Cambridge university press, 1999.

[3] C. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 5, no. 3, pp. 480–92, 1996.

[4] T. M. Benson, B. De Man, and J. Thibault, "Block-based iterative coordinate descent," in *IEEE Nuclear Science Symposuim & Medical Imaging Conference*. Oct. 2010, Ieee.

[5] F. J. Beekman and C. Kamphuis, "Ordered subset reconstruction for X-ray CT," *Physics in medicine and biology*, vol. 46, no. 7, pp. 1835–44, 2001.

[6] N. H. Clinthorne, T. S. Pan, P. C. Chiao, W. L. Rogers, and J. Stamos, "Preconditioning methods for improved convergence rates in iterative reconstructions.," *IEEE transactions on medical imaging*, vol. 12, no. 1, pp. 78–83, Jan. 1993.

[7] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Transactions on Medical Imaging*, vol. 31, no. 3, pp. 677–688, 2012.

[8] V. N. Ingal and E. A. Beliaevskaya, "X-ray plane-wave tomography observation of the phase contrast from a non-crystalline object," *Journal of Physics D: Applied Physics*, vol. 28, no. 11, pp. 2314–2317, 1995.

[9] T. J. Davis, D. Gao, T. E. Gureyev, A. W. Stevenson, and S. W.Wilkins, "Phase-contrast imaging of weakly absorbing materials using hard x-rays," *Nature*, vol. 373, pp. 595–598, 1995.

[10] D. Chapman, W. Thomlinson, R. E. Johnston, D. Washburn, E. Pisano, N. Gmr, Z. Zhong, R. Menk, F. Arfelli, and D. Sayers, "Diffraction enhanced X-ray imaging," *Physics, Medicine and Biology*, vol. 42, pp. 2015–2025, 1997.

[11] U. Bonse and M. Hart, "An X-ray interferometer," *Applied Physics Letters*, vol. 6, no. 8, pp. 155–156, 1965.

[12] A. Momose, T. Takeda, Y. itai, and K. Hirano, "Phase-contrast x-ray computed tomography for observing biological soft tissues," *Nature Medicine*, vol. 2, pp. 473–475, 1996.

[13] T. Weitkamp, A. Diaz, C. David, F. Pfeiffer, M. Stampanoni, P. Cloetens, and E. Ziegler, "X-ray phase imaging with a grating interferometer," *Opt. Express*, vol. 13, no. 16, pp. 6296–6304, 2005.

[14] A. Snigirev and I. Snigireva, "On the possibilities of X-ray phase contrast microimaging by coherent high-energy synchrotron radiation," *Review of Scientific Instruments*, vol. 66, no. 12, pp. 5486–5492, 1995.

[15] K. A. Nugent, T. E. Gureyev, D. F. Cookson, D. Paganin, and Z. Barnea, "Quantitative phase imaging using hard x rays," *Physical review letters*, vol. 77, pp. 2961–2964, Sep 1996.

[16] S. W. Wilkins, T. E. Gureyev, D. Gao, A. Pogany, and A. W. Stevenson, "Phase-contrast imaging using polychromatic hard X-rays," *Nature*, vol. 384, no. 6607, pp. 335–338, 1996.

[17] A. Bravin, "Exploiting the X-ray refraction contrast with an analyzer: the state of the art," *Journal of physics D: Applied physics*, vol. 36, pp. 24–29, 2003.

[18] P Suortti and W Thomlinson, "Medical applications of synchrotron radiation.," *Physics in medicine and biology*, vol. 48, no. 13, pp. R1–35, July 2003.

[19] A. Momose, "Phase-sensitive imaging and phase tomography using X-ray interferometers," *Optics express*, vol. 11, no. 19, pp. 2303–14, Sept. 2003.

[20] R. Lewis, "Medical phase contrast X-ray imaging: current status and future prospects," *Physics in Medicine and Biology*, vol. 49, no. 16, pp. 3573–3583, Aug. 2004.

[21] Shu-Ang Zhou and Anders Brahme, "Development of phase-contrast X-ray imaging techniques and potential medical applications.," *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB)*, vol. 24, no. 3, pp. 129–48, Sept. 2008.

[22] Yunzhe Zhao, Emmanuel Brun, and Paola Coan, "High-resolution, low-dose phase contrast X-ray tomography for 3-D diagnosis of human breast cancers," *Proceedings of the National Academy of Sciences*, 2012.

[23] A. Bravin, P. Coan, and P. Suortti, "X-ray phase-contrast imaging: from pre-clinical applications towards clinics," *Physics in medicine and biology*, vol. 58, no. 1, pp. R1–35, Jan. 2013.

[24] C. David, B. Nohammer, H. Solak, and E. Ziegler, "Differential X-ray phase contrast imaging using a shearing interferometer," *Applied physics letters*, vol. 81, no. 17, pp. 3287–3289, 2002.

[25] M. Bech, A. Tapfer, A. Velroyen, A. Yaroshenko, B. Pauwels, J. Hostens, P. Bruyndonckx, A. Sasov, and F. Pfeiffer, "In-vivo dark-field and phase-contrast X-ray imaging," *Scientific reports 3*, 2013.

[26] F. Pfieffer, O. Bunk, C. Kottler, and C. David, "Tomographic reconstruction of three-dimensional objects from hard X-ray differential phase contrast projection images," *Nuclear Instrument and Methods in Physics Research*, vol. 580, no. 2, pp. 925–928, 2007.

[27] D. Bharkhada, H. Yu, and S. Ge, "Cardiac computed tomography radiation dose reduction using interior reconstruction algorithm with the aorta and vertebra as known information," *Journal of computer assisted tomography*, vol. 33, no. 3, pp. 338–347, 2009.

[28] S. M. Hashemi, S. Beheshti, and P. R. Gill, "Efficient Low Dose X-ray CT Reconstruction through Sparsity-Based MAP Modeling," *arXiv preprint arXiv: 1402.1801*, pp. 1–10, 2014.

[29] L. Yu, X. Liu, and S. Leng, "Radiation dose reduction in computed tomography: techniques and future perspective," *Imaging in medicine*, vol. 1, no. 1, pp. 65–84, 2009.

[30] D. J. Brenner and E. J. Hall, "Computed tomography–an increasing source of radiation exposure," *The New England journal of medicine*, vol. 357, no. 22, pp. 2277–84, 2007.

[31] R. Smith-Bindman, J. Lipson, R. Marcus, K. Kim, M. Mahesh, R. Gould, A. Berrington de González, and D. L. Miglioretti, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer," *Archives of internal medicine*, vol. 169, no. 22, pp. 2078–86, 2009.

[32] R. L. Morin, T. C. Gerber, and C. H. McCollough, "Radiation dose in computed tomography of the heart," *Circulation*, vol. 107, no. 6, pp. 917–922, Feb. 2003.

[33] Z. Qi, J. Zambelli, N. Bevins, and G. Chen, "A novel method to reduce data acquisition time in differential phase contrast computed tomography using compressed sensing," *Proc. of SPIE*, vol. 7258, pp. 4A1–8, 2009.

[34] T. Köhler, B. Brendel, and E. Roessl, "Iterative reconstruction for differential phase contrast imaging using spherically symmetric basis functions," *Medical physics*, vol. 38, pp. 4542, 2011.

[35] Q. Xu, Emil Y. Sidky, X. Pan, M. Stampanoni, P. Modregger, and M. A. Anastasio, "Investigation of discrete imaging models and iterative image reconstruction in differential X-ray phase-contrast tomography," *Optics Express*, vol. 20, no. 10, pp. 10724–10749, 2012.

[36] M. Nilchian, C. Vonesch, P. Modregger, M. Stampanoni, and M. Unser, "Fast iterative reconstruction of differential phase contrast X-ray tomograms," *Optics Express*, vol. 21, no. 5, pp. 5511–5528, 2013.

[37] R. A. Crowther, "Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs," *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 261, no. 837, pp. 221–230, 1971.

[38] D. J. DeRosier and P. B. Moore, "Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry," *Journal of molecular biology*, vol. 52, no. 2, pp. 355–369, 1970.

[39] Y. Cong, M. L. Baker, J. Jakana, D. Woolford, E. J. Miller, S. Reissmann, and W. Chiu, "4.0- resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement," *Proceedings of the National Academy of Sciences*, vol. 107, no. 11, pp. 4967–4972, 2010.

[40] J. M. Carazo, "The fidelity of 3-D reconstructions from incomplete data and the use of restoration methods," *In Electron Tomography, Springer US*, pp. 117–164, 1992.

[41] J. Frank, *Three-dimensional electron microscopy of macromolecular assemblies*, Oxford University Press, New York, NY., 2006.

[42] F. J. Sigworth, P. C. Doerschuk, J. M. Carazo, and S. H. Scheres, "An introduction to maximum-likelihood methods in Cryo-EM," *Methods in enzymology*, vol. 482, pp. 263–294, 2010.

[43] C. de Boor, K. Höllig, and S. Riemenschneider, *Box Splines*, vol. 98 of *Applied Mathematical Sciences*, Springer-Verlag, New York, 1993.

[44] R. M.Lewit, "Multidimensional digital image representations using generalized kaiser-bessel window functions," *Optical Society of America*, vol. 7, no. 10, pp. 1834–1846, 1990.

[45] S. Matej and R. M. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements," *IEEE Transactions on Medical Imaging*, vol. 15, no. 1, pp. 68–78, 1996.

[46] C. O. S. Sorzano, R. Marabini, J. Velzquez-Muriel, J. R. Bilbao-Castro, S. H. Scheres, J. M. Carazo, and A. Pascual-Montano, "XMIPP: a new generation of an open-source image processing package for electron microscopy," *Journal of structural biology,*, vol. 148, no. 2, pp. 194–204, 2004.

[47] R. Marabini, I. M. Masegosa, M. C. San Martn, S. Marco, J. J. Fernandez, L. G. De la Fraga, C. Vaquerizo, and J. M. Carazo., "Xmipp: an image processing package for electron microscopy," *Journal of structural biology,*, vol. 116, no. 1, pp. 237–240, 1996.

[48] A. V. Bronnikov, "Theory of quantitative phase-contrast computed tomography," *Journal of Optical Society America A*, vol. 19, no. 3, pp. 472–480, 2002.

[49] C. Kottler, C. David, F. Pfeiffer, and O. Bunk, "A two-directional approach for grating based differential phase contrast imaging using hard X-rays," *Optics express*, vol. 15, no. 3, pp. 1175–81, Feb. 2007.

[50] H. Itoh, K. Nagai, and G. Sato, "Two dimensional grating-based X-ray phase-contrast imaging using Fourier transform phase retrieval," *Optics express*, vol. 19, no. 4, pp. 3339–3346, 2011.

[51] A. Entezari, M. Nilchian, and M. Unser, "A box spline calculus for the discretization of computed tomography reconstruction problems," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 8, pp. 1532 –1541, aug. 2012.

[52] R. Marabini, E. Rietzel, R. Schroeder, G. T. Herman, and J. M. Carazo, "Three-dimensional reconstruction from reduced sets of very noisy images acquired following a single-axis tilt schema: application of a new three-dimensional reconstruction algorithm and objective comparison with weighted backprojection," *Journal of structural biology*, vol. 120, no. 3, pp. 363–371, 1997.

[53] A. P. Leis, M. Beck, M. Gruska, C. Best, R. Hegerl, W. Baumeister, and J. W. Leis, "Cryo-electron tomography of biological specimens," *IEEE Signal Processing*, vol. 23, no. 3, pp. 95–103, 2006.

[54] J. J. Fernández, D. Gordon, and R. Gordon, "Efficient parallel implementation of iterative reconstruction algorithms for electron tomography," *Journal of Parallel and Distributed Computing*, vol. 68, no. 5, pp. 626–640, 2008.

[55] A. Ziegler, T. Kohler, T. Nielsen, and R. Proksa, "Efficient projection and backprojection scheme for spherically symmetric basis functions in divergent beam geometry," *Medical Physics*, vol. 33, no. 12, pp. 4653–4663, 2006.

[56] A. Zieglar, T. Nielsen, and M. Grass, "Iterative reconstruction of a region of interest for transmissiongrass tomography," *Medical physics*, vol. 35, no. 4, pp. 1317–1327, 2008.

[57] Q. Xu, E. Sidky, X. Pan, M. Stampanoni, P. Modregger, and M. Anastasio, "Investigation of discrete imaging models and iterative image reconstruction in differential

X-ray phase-contrast tomography.," *Optics express*, vol. 20, no. 10, pp. 10724–10749, 2012.

[58] M. Unser, "Sampling–50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569–587, 2000.

[59] T. Blu and M. Unser, "Approximation error for quasi-interpolators and (multi-) wavelet expansions," *Applied and Computational Harmonic Analysis*, vol. 6, no. 2, pp. 219–251, 1999.

[60] F. Natterer, *The mathematics of computerized tomography*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[61] N.B. Andersen, "On Fourier and Hankel sampling," *Australian Mathematical Society Gazette*, vol. 33, no. 1, pp. 44–46, 2006.

[62] S. Meignen, "Application of the convergence of the control net of box splines to scale-space filtering," *IEEE Transacions on Image Processing*, vol. 16, no. 11, pp. 2842–2848, 2007.

[63] W. He and M. Lai, "Construction of trivariate compactly supported biorthogonal box spline wavelets," *Journal of Approximation Theory*, vol. 120, no. 1, pp. 1–19, 2003.

[64] M. Unser, A. Aldroubi, and M. Eden, "B-Spline signal processing: Part I—Theory," *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821–833, Feb. 1993.

[65] S. Horbelt, M. Liebling, and M. Unser, "Discretization of the Radon transform and of its inverse by spline convolutions," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 363–376, 2002.

[66] M. Richter, "Use of box splines in computer tomography," *Computing*, vol. 61, no. 2, pp. 133–150, 1998.

[67] M. Nilchian, C. Vonesch, S. Lefkimmiatis, P. Modregger, M. Stampanoni, and M. Unser, "Constrained regularized reconstruction of X-ray DPCI tomograms with weighted-norm," *Optics express*, vol. 21, no. 26, pp. 32340–32348, 2013.

[68] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT.," *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB)*, vol. 28, no. 2, pp. 94–108, Apr. 2012.

[69] M. Nilchian and M. Unser, "Differential phase-contrast X-ray computed tomography: From model discretization to image reconstruction," in *Proceedings of the Ninth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'12)*, Barcelona, Spain, May 2-5, 2012, pp. 90–93.

[70] S. Maeda, W. Fukuda, A. Kanemura, and S. Ishii, "Maximum a posteriori X-ray computed tomography using graph cuts," *Journal of Physics: Conference Series*, vol. 233, pp. 012023, 2010.

[71] J. Thibault, K. D. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Medical Physics*, vol. 34, no. 11, pp. 4526–4544, 2007.

[72] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.

[73] J. F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, vol. 38, no. 5, pp. 826–844, 1973.

[74] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[75] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[76] E. J. Cands, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[77] E. J. Cands and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[78] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[79] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.

[80] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[81] N. Meinshausen and P. Bhlmann, "High-dimensional graphs and variable selection with the lasso," *The Annals of Statistics*, vol. 34, no. 3, pp. 1436–1462, 2006.

[82] M. Lassas and S. Siltanen, "Can one use total variation prior for edge-preserving Bayesian inversion?," *Inverse Problems*, vol. 20, no. 5, pp. 1537–1563, 2004.

[83] E. Bostan, U. S. Kamilov, M. Nilchian, and M. Unser, "Sparse stochastic processes and discretization of linear inverse problems," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2699–2710, July 2013.

[84] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[85] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.

[86] Y. Nesterov, "A method of solving a convex programming problem with convergence rate o$(1/k^2)$," *In Soviet Mathematics Doklady*, vol. 27, no. 2, pp. 372–376, 1983.

[87] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM imaging sciences*, vol. 2, pp. 183–202, 2009.

[88] S. Becker, J. Bobin, and E. J. Cands, "NESTA: a fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[89] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM journal on scientific computing*, vol. 33, no. 1, pp. 250–278, 2011.

[90] K. Koh, S. J. Kim, and S. P. Boyd, "An interior-point method for large-scale ℓ_1-regularized logistic regression," *Journal of Machine learning research*, vol. 8, no. 8, pp. 1519–1555, 2007.

[91] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.

[92] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[93] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[94] T. Goldstein and S. Osher, "The split bregman method for ℓ_1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[95] M. 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.

[96] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, March 2011.

[97] B. Vandeghinste, B. Goossens, J. De Beenhouwer, A. Pizurica, W. Philips, S. Vandenberghe, and Steven S.Staelens, "Split-Bregman-based sparse-view CT reconstruction," in *Fully 3-D 2011 proceedings*, 2011, pp. 431–434.

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

[99] F. T. A. W. Wajer and K. P. Pruessmann, "Major speedup of reconstruction for sensitivity encoding with arbitrary trajectories," *Proc. Intl. Soc. Mag. Res. Med.*, vol. 9, pp. 767, 2001.

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

[101] A. H. Delaney and Y. Bresler, "A fast and accurate Fourier algorithm for iterative parallel-beam tomography," *IEEE Transactions on Image Processing*, vol. 5, no. 5, pp. 740–753, 1996.

[102] H Stark and JW Woods, "Direct Fourier reconstruction in computer tomography," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 2, pp. 237–245, 1981.

[103] H Schomberg and J Timmer, "The gridding method for image reconstruction by Fourier transformation," *IEEE transactions on medical imaging*, vol. 14, no. 3, pp. 596–607, Jan. 1995.

[104] Jianwei Miao, Friedrich Förster, and Ofer Levi, "Equally sloped tomography with oversampling reconstruction," *Physical Review B*, vol. 72, no. 5, pp. 052103, Aug. 2005.

[105] Yu Mao and BP Fahimian, "Development and optimization of regularized tomographic reconstruction algorithms using equally-sloped tomography," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1259–1268, 2010.

[106] E. Lee, B. Fahimian, C. Iancu, C. Suloway, Gavin E. Murphy, E. Wright, D. Castaño Díez, G. Jensen, and J. Miao, "Radiation dose reduction and image enhancement in biological imaging through equally-sloped tomography," *Journal of Structural Biology*, vol. 164, no. 2, pp. 221–7, Nov. 2008.

[107] J. Eckstein and D. P. Bertsekas, "Sparse stochastic processes and discretization of linear inverse problems," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

[108] M. Stampanoni, Z. Wang, T. Thüring, C. David, E. Roessl, M. Trippel, R. A. Kubik-Huch, G. Singer, M. K. Hohl, , and N. Hauser, "The first analysis and clinical evaluation of native breast tissue using differential phase-contrast mammography," *Investigation Radiology*, vol. 46, pp. 801–806, 2011.

[109] A. Momose, W. Yashiro, Y. Takeda, Y. Suzuki, and T. Hattori, "Phase tomography by X-ray talbot interferometry for biological imaging," *Japanese journal of applied physics*, vol. 45, pp. 5254, 2006.

[110] F. Pfeiffer, C. Grünzweig, O. Bunk, G. Frei, E. Lehmann, and C. David, "Neutron phase imaging and tomography," *Physical Review Letter*, vol. 96, pp. 215505, Jun 2006.

[111] K. Itoh, "Quantitative imaging of electron density and effective atomic number using phase contrast CT," *Physics in medicine and biology*, vol. 55, no. 9, pp. 2669, 2010.

[112] A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki, "Demonstration of X-ray Talbot interferometry," *Japanese Journal of Applied Physics*, vol. 42, no. Part 2, No. 7B, pp. L866–L868, 2003.

[113] J. P. Guigay, "On Fresnel diffraction by one-dimensional periodic objects, with application to structure determination of phase objects," *Journal of Modern Optics*, vol. 18, no. 9, pp. 677–682, 1971.

[114] J. B. Develis, Parrent, G. B., and B. Thompson, "The new physical optics notebook: tutorials in Fourier optics," *Journal of Modern Optics*, vol. 61, 1989.

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

[116] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81 –84, March 2002.

[117] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[118] S. A. McDonald, F. Marone, C. Hintermuller, G. Mikuljan, C. David, F. Pfeiffer, and M. Stampanoni, "Advanced phase-contrast imaging using a grating interferometer," *Synchrotron Radiation*, vol. 16, pp. 562–572, 2009.

[119] S. Lefkimmiatis, J. P. Ward, and M. Unser, "Hessian schatten-norm regularization for linear inverse problems," *IEEE Transaction of Image Processing*, vol. 22, pp. 1873–1888, 2013.

[120] M. Nilchian, C. Vonesch, P. Modregger, M. Stampanoni, and M. Unser, "Iterative FBP for improved reconstruction of X-ray differential phase-contrast tomograms," *Proceeding of International Symposium on Biomedical Imaging*, pp. 1248–1251, 2013.

[121] I. Jerjen, V. Revol, P. Schuetz, C. Kottler, R. Kaufmann, T. Luethi, K. Jefimovs, C. Urban, and U. Sennhauser, "Reduction of phase artifacts in differential phase contrast computed tomography," *Opt. Express*, vol. 19, no. 14, pp. 13604–13611, Jul 2011.

[122] K. Itoh, "Analysis of the phase unwrapping algorithm," *Applied Optics*, vol. 21, no. 14, pp. 2470–2470, 1982.

[123] F. M. Epple, S. Ehn, P. Thibault, T. Koehler, G. Potdevin, J. Herzen, D. Pennicard, H. Graafsma, P. B. Nol, and F. Pfeiffer, "Phase unwrapping in spectral X-ray differential phase-contrast imaging with an energy-resolving photon-counting pixel detector," *IEEE Transaction on Medical Imaging*, pp. 816–823, 2014.

[124] N. Tzemos, J. Therrien, J. Yip, G. Thanassoulis, S. Tremblay, M. T. Jamorski, G. D. Webb, and S. C. Siu, "Outcomes in adults with bicuspid aortic valves," *JAMA*, vol. 300, no. 11, pp. 1317–1325, 2008.

[125] B. Pinzer, Z. Wang, M. Stampanoni, and P. Modregger, *Emerging Imaging Technologies in Medicine*, 2012.

[126] S. Lefkimmiatis, A. Roussos, M. Unser, and P. Maragos, "Convex generalizations of total variation based on the structure tensor with applications to inverse problems," in *Proceedings of the Fourth International Conference on Scale Space and Variational Methods in Computer Vision (SSVM'13)*, Seggauberg, Austria, 2013, pp. 48–60.

[127] M. Nilchian, Z. Wang, T. Thuering, M. Unser, and M. Stampanoni, "Spline based iterative phase retrieval algorithm for X-ray differential phase contrast radiography," *Optics express*, vol. 23, no. 8, pp. 10631–10642, 2015.

[128] F. Pfeiffer, M. Bech, O. Bunk, P. Kraft, E. F. Eikenberry, C. Brönnimann, C. Grünzweig, and C. David, "Hard-X-ray dark-field imaging using a grating interferometer," *Nature Mater*, vol. 7, pp. 134–137, 2008.

[129] T. Thüring, R. Guggenberger, H. Alkadhi, M. Vich, Z. Wang, C. David, and M. Stampanoni, "Human hand radiography using X-ray differential phase contrast combined with dark-field imaging," *Skeletal radiology*, vol. 42, pp. 827–835, 2013.

[130] J. Tanaka, M. Nagashima, K. Kido, Y. Hoshino, J. Kiyohara, C. Makifuchi, S. Nishino, S. Nagatsuka, and A. Momose, "Cadaveric and in vivo human joint imaging based on differential phase contrast by X-ray talbot-lau interferometry," *Zeitschrift fr Medizinische Physik*, vol. 23, no. 3, pp. 222–227, 2012.

[131] R. A. Lewis, "Medical phase contrast X-ray imaging: current status and future prospects," *Physics in medicine and biology*, vol. 49, no. 16, pp. 3573–3583, 2004.

[132] D. Rudolph, G. Schmahl, and B. Niemann, "Amplitude and phase contrast in X-ray microscopy," *In Modern Microscopies. Springer US*, pp. 59–67, 1990.

[133] T. Thring, P. Modregger, B. R. Pinzer, Z. Wang, and M. Stampanoni, "Non-linear regularized phase retrieval for unidirectional X-ray differential phase contrast radiography," *Optics express,*, vol. 19, no. 25, pp. 25545–25558, 2011.

[134] J. I. Sperl, D. Bequé, G. P. Kudielka, K. Mahdi, P. M. Edic, , and C. Cozzini, "A fourier-domain algorithm for total-variation regularized phase retrieval in differential X-ray phase contrast imaging," *Optics express*, vol. 22, no. 1, pp. 450–462, 2014.

[135] N. Hauser, Z. Wang, R. A. Kubik-Huch, M. Trippel, G. Singer, M. K. Hohl, E. Roessl, T. Koehler, U. van Stevendaal, N. Wieberneit, and M. Stampanoni, "A study on mastectomy samples to evaluate breast imaging quality and potential clinical relevance of differential phase contrast mammography," *Investigative Radiology*, vol. 49, pp. 131–137, 2014.

[136] V. Revol, C. Kottler, R. Kaufmann, U. Straumann, and C. Urban, "Noise analysis of grating-based X-ray differential phase contrast imaing," *Review of Scientific Instruments*, vol. 81, pp. 073709, 2010.

[137] C. Kottler, C. David, F Pfeiffer, and O. Bunk, "A two-directional approach for grating based differential phase contrast imaging using hard x-rays," *Optics Express*, vol. 15, no. 3, pp. 1175–1181, 2007.

[138] K. J. Hanszen, "The optical transfer theory of the electron microscope: fundamental principles and applications," *Advances in optical and electron microscopy*, vol. 4, pp. 1–84, 1971.

[139] J. C. Spence, *Experimental high-resolution electron microscopy*, Oxford University Press, 1988.

[140] P. W. Hawkes, "The electron microscope as a structure projector," *Springer New York*, pp. 83–111, 2006.

[141] P. W. Hawkes and E. Kasper, *Principles of electron optics: Wave Optics*, Academic Press, 1996.

[142] C. O. Sorzano, A. Otero, E. M. Olmos, and J. M. Carazo, "Error analysis in the determination of the electron microscopical contrast transfer function parameters from experimental power spectra," *BMC structural biology*, vol. 9, no. 1, pp. 18, 2009.

[143] K. H. Hanszen and L. Trepte, "The contrast transfer function of the electron microscope with partial coherent illumination. a. the ring condensor," *Optik*, vol. 33, no. 1, pp. 166–181, 1971.

[144] K. H. Hanszen and L. Trepte, "The contrast transfer function of the electron microscope with partial coherent illumination. b. disc-shaped source," *Optik*, vol. 33, no. 1, pp. 182–198, 1971.

[145] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and X-ray photography," *Journal of theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970.

[146] R. Marabini, G. T. Herman, and J. M. Carazo, "3-D reconstruction in electron microscopy using art with smooth spherically symmetric volume elements (blobs)," *Ultramicroscopy*, vol. 72, no. 1, pp. 53–65, 1988.

[147] P. Gilbert, "Iterative methods for the three-dimensional reconstruction of an object from projections," *Journal of theoretical biology*, vol. 36, no. 1, pp. 105–117, 1972.

[148] S. H. Scheres, "A bayesian view on cryo-em structure determination," *Journal of molecular biology*, vol. 415, no. 2, pp. 406–418, 2012.

[149] L. Wang, Y. Shkolnisky, and A. Singer, "Fourier-based approach for iterative 3-D reconstruction from cryo-em images," *arXiv preprint arXiv:1307.5824*, 2013.

[150] M. Radermacher, T. Wagenknecht, A. Verschoor, and J. Frank, "A new 3-D reconstruction scheme applied to the 50S ribosomal subunit of E. coli," *Journal of microscopy*, vol. 141, no. 1, pp. RP1–RP2, 1986.

[151] G. Harauz and M. van Heel, "Exact filters for general geometry three dimensional reconstruction," *Optik*, vol. 73, no. 4, pp. 146–156, 1986.

[152] M. Radermacher, "Weighted back-projection methods," *Springer New York*, pp. 245–273, 2006.

[153] P. A. Penczek, R. Renka, and H. Schomberg, "Gridding-based direct fourier inversion of the three-dimensional ray transform," *Journal of Optical Society of America A*, vol. 21, no. 4, pp. 499–509, 2004.

[154] P. A. Penczek, "Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms," *Journal of Structural Biology*, vol. 138, no. 1, pp. 34–46, 2002.

[155] W. Zhang, M. Kimmel, C. M. Spahn, and P. A. Penczek, "Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis," *Structure*, vol. 16, no. 12, pp. 1770–1776, 2003.

[156] Z. Huang, P. R. Baldwin, S. Mullapudi, and P. A. Penczek, "Automated determination of parameters describing power spectra of micrograph images in electron microscopy," *Journal of structural biology*, vol. 144, no. 1, pp. 79–94, 2003.

[157] J. Zhu, P. Penczek, R. Schroder, and J. Frank, "Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70s escherichia coli ribosome," *Journal of Structural Biology*, vol. 118, no. 3, pp. 197–219, 1997.

[158] P. A. Penczek, J. Zhu, R. Schrder, and J. Frank, "Three dimensional reconstruction with contrast transfer compensation from defocus series," *Scanning Microscope*, vol. 11, no. 1-10, pp. 147–154, 1997.

[159] M. van Heel and M. Schatz, "Fourier shell correlation threshold criteria. journal of structural biology," *Journal of structural biology,*, vol. 151, no. 3, pp. 250–262, 2005.