Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Kadri Aditya Mohan

Entitled

MODULAR FORWARD MODELS AND ALGORITHMS FOR REGULARIZED RECONSTRUCTION OF TIME-SPACE SCALAR AND VECTOR FIELDS

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Charles A. Bouman, Chair	William C. Karl
Jan P. Allebach	
Mary L. Comer	
Peter W. Voorhees	

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Charles A. Bouman

Approved by: Venkataramanan Balakrishnan 12/9/2016

Head of the Departmental Graduate Program Date

MODULAR FORWARD MODELS AND ALGORITHMS FOR REGULARIZED RECONSTRUCTION OF TIME-SPACE SCALAR AND VECTOR FIELDS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Kadri Aditya Mohan

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2017

Purdue University

West Lafayette, Indiana

Dedicated to my family.

ACKNOWLEDGMENTS

I thank my PhD advisor, Prof. Charles Bouman, for his guidance and support that helped me realize my life long dream of completing a PhD degree. The energy and passion with which he pursues research will always be an inspiration to me. He repeatedly stressed the need to maintain high standards for research and publications that significantly improved my research skills. He provided me ample opportunities to present my research at various conferences and workshops that enhanced my presentation skills and also helped me develop new ideas for research.

I thank Prof. Peter Voorhees from Northwestern University for providing us valuable data and sharing his expertise in materials imaging. I am indebted to Prof. Marc De Graef from Carnegie Mellon University (CMU) who was truly responsible for nurturing a collaborative environment for multi-disciplinary research involving signal processing and materials imaging. I am much obliged to Xianghui Xiao for providing us access to his beamline for our experiments at the advanced photon source (APS) in Argonne National Lab. I am also grateful to Charudatta Phatak from Argonne National Lab for helping me understand magnetism and for providing us real microscopy data. I will always cherish the sleepless yet memorable beam times at Argonne with John, Ashwin, Begum, and Yue. I thank the Air Force Office of Scientific Research (AFOSR) for the multi-disciplinary university research initiative (MURI) grant that supported me as a research assistant at Purdue University

I am very grateful to my PhD committee members, Prof. Peter Voorhees, Prof. Mary Comer, Prof. Jan Allebach, and Prof. Clem Karl for their valuable feedback and guidance. I would also like to express my heartfelt gratitude to my friend and mentor S. Venkatakrishnan for guiding me in my early days of research. Special thanks to Dilshan, Zeeshan, Suhas, Shruthi, Xiao, Usman, Venkatesh, Thilo, Soumendu, Casey, Ben, Ruoqiao, and Prabhat (CMU) for all the intellectually stimulating discussions

in the lab and outside. Also, thank you to Prof. Bouman's wife, Cristina Bouman, for organizing and inviting us to many wonderful picnics. Finally, my time at Purdue would not have been much fun without my wonderful friends Dilshan, Munadir, James, Ankush, Karthik, Keerthiraj, Praveen, Shwetha, Shruthi, Nikhil, Prashanth, Aditi, Ashesh, and Divya.

I thank Prof. Chandra Murthy from Indian Institute of Science and my previous mentors at Manipal Dot Net (MDN) who were truly responsible for my interest in research. I also thank my teachers at Sri Ramakrishna School and National Institute of Technology Karnataka, Surathkal.

This dissertation would not have been possible without the support of my mother, Dakshayini, and father, Chandra Mohan. I will be eternally grateful for all the hard work and sacrifices made by my parents to ensure a bright future for both myself and my sister. I am grateful to my sweet younger sister, Ashwitha, for being an awesome sibling who always believed in me. I am thankful to my wife Prajna and my parents-in-law who have always been very supportive and patient in all my endeavors. I am so lucky to have Prajna as my wife and this dissertation is a testimony for her trust in me. I will always remember my grandparents for their love, wisdom, guidance, and support. Finally, I would like to thank every friend, uncle, aunt, and cousin for supporting me in this wonderful journey.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

ABSTRACT

Mohan, Kadri Aditya Ph.D., Purdue University, May 2017. Modular Forward Models and Algorithms for Regularized Reconstruction of Time-Space Scalar and Vector Fields. Major Professor: Charles A. Bouman.

The framework of model-based iterative reconstruction (MBIR) is a versatile but powerful technique for solving a wide variety of inverse problems in imaging. In MBIR, image reconstruction works by finding the solution that minimizes a cost function consisting of the sum of a forward model and a prior model term. The physics of imaging is captured by the forward model and image sparsity is modeled by the prior model. The challenge lies in the design of physically accurate forward models and optimization algorithms for reconstruction.

First, we solve the inverse problem associated with traditional X-ray computed tomography imaging of time-varying samples that uses a linear and sparse model for the forward problem. We formulate a 4D MBIR algorithm for reconstruction that when combined with an interlaced view sampling strategy achieves dramatic improvements in spatial and temporal resolution of reconstructions. However, this algorithm cannot be used for more complex forms of non-sparse or non-linear tomographic imaging systems.

In the area of magnetic imaging using vector field electron tomography (VFET), traditional algorithms do not permit us to directly reconstruct the sample's magnetization due to the non-sparse and complex nature of the forward problem. We formulate the first algorithm ever to perform 3D reconstruction of the sample's magnetization using the MBIR framework. Next, we solve the inverse problem in X-ray phase contrast tomography (XPCT) using a non-linear forward model that accounts for X-ray refraction and diffraction. We show that our algorithm accurately recon-

structs the sample and does not adversely constrain the experimental conditions by using a linear approximation to the forward model.

Both the algorithms designed for VFET and XPCT use modular forward models that permit us to solve the original inverse problem by iteratively solving multiple but simpler optimization problems using efficient algorithms. We generalize this approach and call it the framework of plug-and-play forward models for MBIR. We show that plug-and-play forward modeling allows us to solve complex inverse problems using independent software modules that solve simpler inverse problems.

1. INTRODUCTION

Computational imaging uses computing resources to reconstruct the morphology of a sample, also referred to as the object, from measurements that does not provide direct information about the sample morphology. Typically, the sample is exposed to some form of radiation such as X-rays, electrons, or visible light and the radiation after interacting with the sample is measured by a detector. The forward problem that expresses the measurement data as a function of the sample morphology and its physical properties can be described using a forward model. The associated inverse problem attempts to reconstruct the unknown sample from the measurement data. In typical imaging systems, the forward model is usually known and can be formulated using the laws of physics. However, the inverse problem is typically ill-posed either due to missing information or sparse sampling of measurement data. Furthermore, measurement data almost always contains noise that might result in noisy or unstable reconstructions.

The common approach to solving an inverse problem is to use analytic inversion formulas that rely on direct analytic inversion of the forward model. This approach typically results in algorithms that are not only relatively straightforward to implement but also computationally fast. However, these algorithms are sensitive to noise and result in inaccurate reconstructions if the problem is ill-posed. A possible solution to this issue is to constrain the reconstruction to some sparse basis or use prior knowledge about the underlying object that will mitigate the effects of the ill-posed nature of the problem. One such simple but powerful constraint is that our reconstructed values for the sample must be continuous and smooth albeit the occasional discontinuities due to image edges. Such knowledge is typically incorporated in a prior model for the sample. However, such knowledge is also typically hard to incorporate using an analytic inversion algorithm. The solution is to use iterative reconstruction algo-

rithms that solve the inverse problem while taking into account the prior knowledge in the prior model.

Model-based iterative reconstruction (MBIR) is a versatile but powerful framework for solving a wide variety of inverse problems in computational imaging. It presents a systematic approach for solving inverse problems using a physically accurate forward model for the imaging system and a prior model to enforce sparsity in reconstruction. One application where MBIR algorithms have been extremely useful is in the field of X-ray computed tomography (XCT). In XCT, the sample is mounted on a rotary stage and exposed to X-ray radiation. The sample is then rotated along an axis and measurements are made at regular angular intervals [1–3]. In this modality, the logarithm of the measured image can be expressed as a 2D projection of the sample's 3D morphology along the direction of X-ray propagation. Thus, XCT is an example of a computational imaging system with a linear and sparse forward model. The forward model is sparse since every measured data point depends only the value at a few points within the sample. MBIR algorithms for XCT have found widespread applications in medical imaging [4–10], material imaging [1–3], and biological imaging. In medical XCT, MBIR algorithms have achieved significant improvements in dose reduction and reconstruction quality [6–8]. MBIR is also used to achieve dramatic improvements in reconstruction resolution, artifact reduction, and noise robustness in the area of materials and biological imaging [1–3].

The power of MBIR lies in its statistical models. The noise model in MBIR models the statistics of measurement noise and ensures noise robustness of the algorithm. The forward model ensures that the reconstruction agrees with the observed data. The prior model mitigates the adverse effects due to the ill-posed nature of the inverse problem by enforcing sparsity in the reconstructed values. It is therefore very important to meticulously design the various models depending on the physics of the imaging modality, the physical properties of the sample, and the statistics of measurement noise. The forward model can also be used to account for any undesired non-ideal behavior of the imaging system. For instance, corrupt measurements or

unknown calibration parameters in the detector can result in inaccurate reconstructions. These unknown parameters, also called nuisance parameters, can be accounted for in the forward model which will prevent undesired artifacts and inaccuracies in the reconstruction. Note that even though the original forward model in XCT is linear, the effective forward model after accounting for the nuisance parameters can be non-linear.

In MBIR, the reconstruction is formulated as the maximum a posteriori (MAP) estimate of the unknown sample, x, given the data, y. MBIR also allows us to estimate unknown calibration parameters, ϕ, related to the measurement system that cause artifacts if not accounted for in the forward model. The reconstruction is given by,

$$\left(\hat{x}, \hat{\phi}\right) = \underset{x,\phi}{\operatorname{argmin}} \left\{ -\log p(y|x,\phi) - \log p(x) \right\} \tag{1.1}$$

where $p(y|x,\phi)$ is the probability density function (pdf) for the data, y, given the sample, x, and calibration parameters, ϕ, and $p(x)$ is the probability density function (pdf) for the sample morphology, x. The forward model term, $\log p(y|x,\phi)$, uses measurement physics to express the likelihood of the data, y, given the sample, x, and calibration parameters, ϕ. The prior model term, $\log p(x)$, is used to model sparsity and prior knowledge about the sample, x.

First, we will use the MBIR framework to solve the inverse problem associated with X-ray computed tomography (XCT) imaging of time varying samples using synchrotron X-rays [1–3]. A synchrotron is a particle accelerator that is used to generate ultra-bright and high-energy X-ray beams for imaging experiments in a wide range of scientific disciplines. Next, we solve the inverse problem associated with vector field electron tomography (VFET) imaging of magnetic samples. We use MBIR to formulate an algorithm that reconstructs the 3D magnetization of a magnetic sample using a non-sparse forward model. We also solve the inverse problem in X-ray phase contrast tomography (XPCT) by formulating an algorithm that reconstructs the 3D distribution of absorption and refractive index within the sample using a non-linear forward model [11]. Both the algorithms designed for VFET and XPCT

solve the original inverse problem by iteratively solving multiple simpler inverse sub-problems. We generalize this approach and call it the framework of plug-and-play forward models for MBIR. Plug-and-play can be used to solve the inverse problem associated with complex tomographic imaging systems by breaking down the original inverse problem into an iterative solution of simpler inverse problems that can be solved using efficient algorithms. This allows for modular software architecture by permitting us to independently design efficient algorithmic solutions for solving the individual inverse sub-problems.

1.1 Inverse Problem in X-ray Computed Tomography

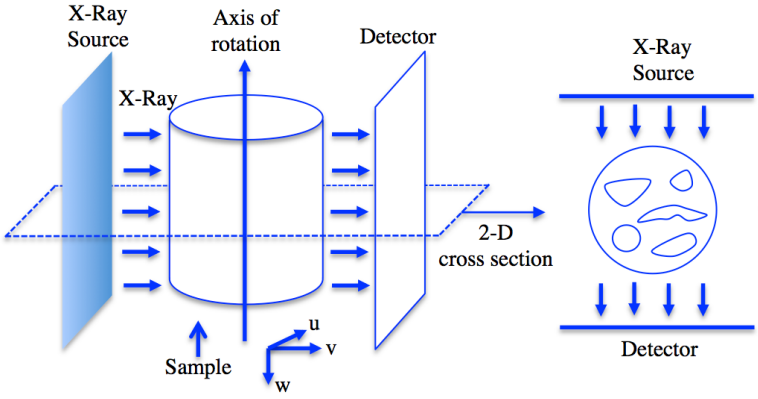


Fig. 1.1. Illustration of data acquisition in 4D-SXCT. A parallel beam of X-rays from a synchrotron is used to repeatedly image a sample rotating at a constant speed. The intensity of the attenuated X-ray beam after passing through the sample is measured by a 2D detector.

Material imaging to characterize the 3D morphology of materials at various length scales is very important for engineering new materials that solve modern challenges in clean energy, health care, and national security. Material characterization can be done at a wide-range of length scales ranging from centimeter to angstrom resolution. The desired resolution and material properties determine the imaging modality that can be used to visualize the sample morphology. For example, we can use synchrotron

X-ray computed tomography (SXCT) for imaging millimeter thick metallic samples such as an AlCu alloy. SXCT relies on X-ray absorption to perform both 3D and 4D (time-space) imaging studies of material structures at micrometer resolution. In SXCT, the varying absorption index within the sample modulates the amplitude of the X-ray electric field that results in intensity contrast in the detector measurements. The forward model is linear and sparse since the data, the logarithm of detector measurements, can be expressed as a projection of the 3D distribution of absorption index along the direction of X-ray propagation. This is the traditional absorption tomography forward model that is used in a wide range of X-ray imaging modalities including medical XCT imaging [1–4, 10]. It is also used in a variety of electron tomography (ET) imaging modalities for nanoscale imaging [12, 13].

In SXCT, the sample is rotated along an axis of rotation and measurements are made at different angular positions of the sample as shown in Fig. 1.1. The inverse problem then involves reconstruction of the absorption index distribution in the sample from the data i.e., logarithm of detector measurements. SXCT is increasingly being used for 3D imaging of material samples at micron and finer scales. The success of these techniques has increased interest in 4D reconstruction methods that can image a sample in both space and time. However, the temporal resolution of widely used 4D reconstruction methods is severely limited by the need to acquire a very large number of views for each reconstructed 3D volume. Consequently, the temporal resolution of current methods is insufficient to observe important physical phenomena. Furthermore, unknown measurement parameters also tend to introduce ring and streak artifacts into the 4D reconstructions.

In this thesis, we present a time-interlaced model-based iterative reconstruction (TIMBIR) method [3] which is a synergistic combination of two innovations. The first innovation, interlaced view sampling, is a novel method of data acquisition which distributes the view angles more evenly in time. The second innovation is a 4D model-based iterative reconstruction algorithm (MBIR) which can produce time-resolved volumetric reconstruction of the sample from the interlaced views. The forward model

in 4D MBIR also accounts for detector saturation and unknown gain offsets in the detector. These parameters if not modeled accurately result in streak and ring artifacts in the reconstruction. The prior model for the sample is used to model sparsity across space and time. In typical SXCT imaging experiments, the reconstruction problem could involve estimation of more than 10 billion variables (including voxels). The high dimensional nature of the problem necessitates the design of highly efficient and parallel algorithms to perform reconstruction. Our optimization algorithm makes the 4D reconstruction problem feasible by providing for efficient parallel computation on a super-computing cluster. We present reconstructions of both simulated and real X-ray synchrotron data which indicate that TIMBIR can improve temporal resolution by an order of magnitude relative to existing approaches.

The extensive research in MBIR algorithms for XCT [1, 3, 4] cannot be directly applied to more complex tomographic systems such as X-ray phase contrast tomography (XPCT) [11,14] and vector field electron tomography [15]. Such imaging systems address the shortcomings of traditional XCT and are widely used for imaging materials in the fields of material science, biology, and nanotechnology. However, it is challenging to solve the inverse problem associated with these imaging modalities due to the more complex non-sparse and/or non-linear forward model. Similar to XCT, these imaging modalities involve estimation of parameters in very high dimensional spaces and hence require efficient algorithms for reconstruction.

1.2 Inverse Problem in Vector Field Electron Tomography

The magnetic property of any material is characterized by its magnetization that expresses the density of magnetic dipole moments in the 3D sample. The magnetization results in a magnetic vector potential and magnetic field distribution that extends outside the sample. Vector field electron tomography (VFET) is extensively used for three dimensional imaging of magnetic materials at nanometer resolutions. In VFET, the sample is tilted across two orthogonal axes and repeatedly imaged at

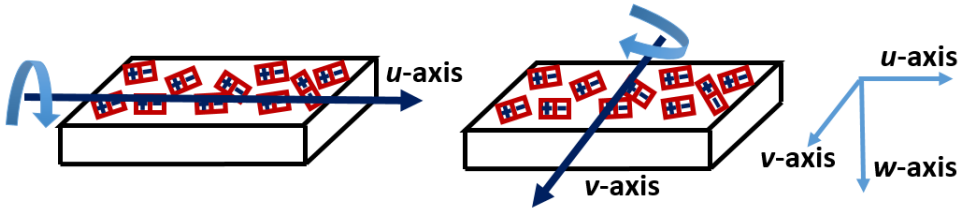


Fig. 1.2. Illustration of data acquisition in vector field electron tomography (VFET). In VFET, the sample is mounted on a rotary stage and exposed to electron radiation. The sample is then tilted across two orthogonal tilt axes ($u-$axis and $v-$axis) and measurements are made at multiple tilt angles.

several tilt angles as shown in Fig. 1.2. At each tilt angle, the phase of the electrons coming out of the sample is retrieved from measurements [15]. The inverse problem then involves reconstruction of the vector fields associated with the magnetic sample from the phase data at all the tilt angles and the two tilt axes. The conventional methods either reconstructs the magnetic vector potential or the magnetic field but not the magnetization [15]. The forward model that expresses the phase data as a function of the magnetic field or magnetic vector potential is linear and sparse but the relation between the phase data and magnetization is linear but is non-sparse and much more complex. However, the scientists are mainly interested in observing the 3D magnetization which can form complex textures that can be difficult to visualize by just observing the magnetic vector potential or the magnetic field. The magnetization is a fundamental property of the sample unlike the magnetic field and the magnetic vector potential which are derived from the magnetization. Furthermore, the conventional algorithms also result in noisy reconstruction artifacts in the reconstruction of the magnetic field and magnetic vector potential due to the ill posed nature of the inverse problem [16].

We present a model-based iterative reconstruction algorithm (MBIR) that reconstructs the magnetization by minimizing a cost function consisting of a forward model term and a prior model term. Our MBIR algorithm uses a physics based forward model and a prior model for the object. The forward model models the data as a

function of the magnetization. The prior model enforces sparsity in the reconstructed values. We then formulate the first algorithm that reconstructs the 3D distribution of magnetization from the phase data. Our algorithm solves the original inverse problem as an iterative solution to multiple but simpler inverse sub-problems that can be solved efficiently. It is based on variable splitting and the theory of alternate direction method of multipliers (ADMM) [17]. Using simulated and real data, we show that our algorithm accurately reconstructs both the magnetization and the magnetic vector potential. Furthermore, our MBIR algorithm also reduces noise and artifacts when compared to the conventional method.

1.3 Inverse Problem in X-ray Phase Contrast Tomography

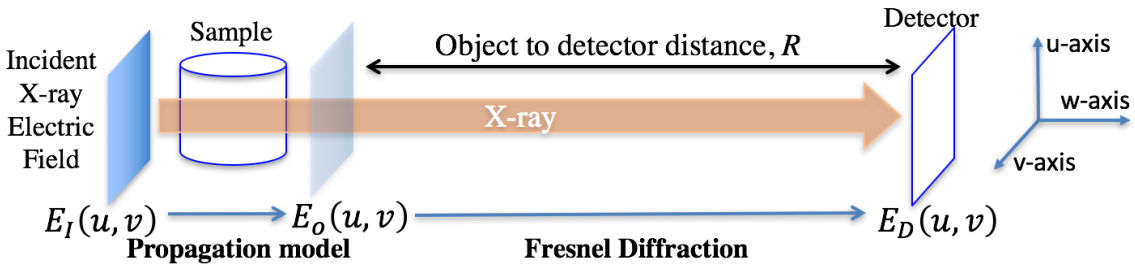


Fig. 1.3. Illustration of data acquisition in X-ray phase contrast tomography (XPCT). A parallel beam of X-rays from a synchrotron is used to repeatedly image a sample rotating at a constant speed. The X-rays undergo absorption and refraction as it propagates through the object and undergo Fresnel diffraction as it propagates towards the detector.

Since SXCT relies on X-ray absorption for contrast, it cannot be used to image materials with weak absorption contrast. Any material with similar atomic numbers such as Al-Si and biological tissues such as brain cells have weak contrast in absorption. To image such samples, we use synchrotron X-ray phase contrast tomography (XPCT) that relies on X-ray refraction rather than absorption for measurement. The 3D distribution of refractive index and absorption index can be conveniently repre-

sented as the real and imaginary parts of the complex refractive index. XPCT has become popular as a method for tomographically reconstructing the complex refractive index of a material from single distance measurements. In XPCT, the sample is rotated along an axis and detector measurements are made at regular angular intervals as shown in Fig. 1.3.

XPCT has two major advantages over traditional X-ray computed tomography: It can be used at higher cone-beam magnifications and it typically produces higher contrast. In XPCT, the varying refractive index modulates the X-ray phase that results in intensity contrast on the detector due to Fresnel diffraction [11]. The measurement contrast increases with the distance between the object and detector. The forward model in XPCT is non-linear due to the combined effects of X-ray absorption, refraction and diffraction that makes it very challenging to solve the inverse problem of reconstructing the sample directly from the measurements. The current XPCT reconstruction algorithms are limited to near-field diffraction or thin samples since they use a linear approximation to the forward model that limits both their use and the quality of reconstructions.

In this thesis, we present a model-based iterative reconstruction (MBIR) algorithm called complex refractive index tomographic iterative reconstruction (CRITIR). CRITIR is based on a non-linear physics based model for X-ray propagation and a prior model for the complex refractive index of the object being imaged. Our reconstruction algorithm solves the original inverse problem as an iterative solution to multiple but simpler inverse sub-problems that can be solved efficiently. Unlike conventional methods, CRITIR does not limit the sample thickness nor impose the near-field diffraction condition as it uses the non-linear forward model in its original form without resorting to linear approximations. We use simulation to show that our algorithm accurately reconstructs the object while the conventional methods result in inaccurate reconstructions with blurry edges beyond the near-field region. Furthermore, CRITIR directly reconstructs the 3D object from measurements at all the view

angles rather than the conventional two-step procedure involving 2D phase retrieval at each view followed by 3D tomographic reconstruction.

1.4 Framework of Plug-and-Play Forward Models

VFET and XPCT are two imaging modalities with forward models that are originally quite complex but can be expressed as the composition of simpler forward sub-models that facilitates the design of efficient reconstruction algorithms. We generalize this approach and present the framework of plug-and-play forward models for solving inverse problems associated with forward models that can be represented as the composition of multiple simpler forward sub-models.

The framework of plug-and-play forward model breaks down the original forward model used to describe a complex imaging system as a composition of multiple forward sub-models that represent simpler and more traditional imaging systems. The inverse sub-problem associated with each individual forward sub-model can typically be solved using efficient algorithms in a MBIR framework. Under certain conditions, there exist an iterative framework for solving the multiple inverse sub-problems so that the iterates solve the original inverse problem. Note that each forward sub-model is paired with a prior sub-model that incorporates knowledge from reconstructions associated with the other inverse sub-problems. The plug-and-play framework is based on variable splitting and the theory of alternate direction method of multipliers (ADMM) [17].

1.5 Organization of Dissertation

In chapter 2, we present the method of TIMBIR for high spatial and temporal resolution SXCT 4D imaging of time-varying samples. The material presented in

chapter 2 was previously published in [1][1] [2][2] [3][3]. In chapter 3, we present an algorithm to reconstruct the 3D magnetization from VFET data. The core ideas associated with the material in chapter 3 was originally published as an abstract in [18]. In chapter 4, we present a MBIR algorithm called CRITIR that reconstructs the 3D complex refractive index of a sample using XPCT. An abridged version of the material in chapter 4 was originally published in [11][4]. In chapter 5, we present the framework of plug-and-play forward models for solving inverse problems using MBIR.

[1]2014 ©IEEE. Reprinted with permission from [1]
[2]2015 ©IEEE. Reprinted with permission from [2]
[3]2015 ©IEEE. Reprinted with permission from [3]
[4]2016 ©IEEE. Reprinted with permission from [11]

2. TIMBIR: A METHOD FOR TIME-SPACE TOMOGRAPHIC RECONSTRUCTION FROM INTERLACED VIEWS

2.1 Introduction

Four-dimensional synchrotron X-ray computed tomography (4D-SXCT) is enabling scientists to study a wide variety of physical processes such as solidification and solid-state phase transformations in the field of material science [19,20]. In contrast to conventional CT, 4D-SXCT produces time-resolved three-dimensional volumetric reconstruction of the sample. The high intensity and strong collimation of synchrotron radiation makes it especially suitable for high speed imaging of a wide variety of samples at the micron scale [19–21]. However, in-situ 4D imaging using SXCT still remains a major challenge owing to limitations on the data acquisition speed [19]. Moreover, impurities in the scintillator and imperfections in the detector elements cause ring artifacts in the reconstruction [1, 22]. Furthermore, detector pixels occasionally get saturated by high energy photons (often called "*zingers*") which cause streak artifacts in the reconstruction [1, 23].

The traditional approach to 4D-SXCT is to acquire a sequence of parallel beam projections of the object, which is rotated at a constant speed, at progressively increasing equi-spaced view angles (henceforth called progressive view sampling) as shown in Fig. 2.1. Typically, the projections in each π radians rotation are grouped together and reconstructed into a single 3D volume using an analytic reconstruction algorithm such as filtered back projection (FBP) [24, 25] or a Fourier domain reconstruction method [26–28]. The time sequence of 3D reconstructions then forms the 4D reconstruction of the object.

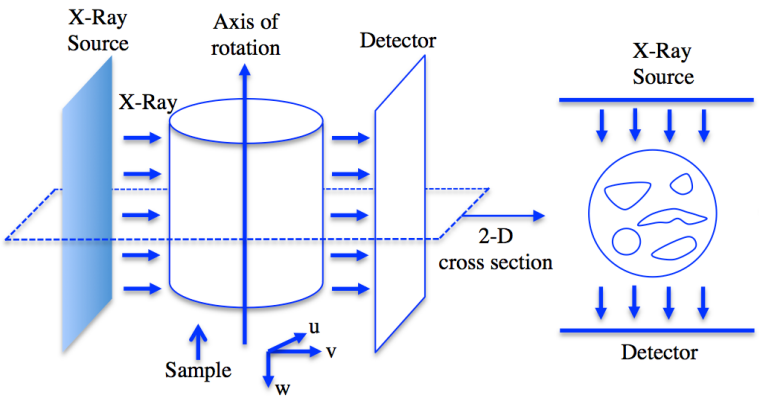


Fig. 2.1. Illustration of data acquisition for 4D-SXCT. A parallel beam of X-rays from a synchrotron is used to repeatedly image an object rotating at a constant speed. The intensity of the attenuated X-ray beam after passing through the object is measured by a 2D detector.

Unfortunately, this traditional approach based on progressive view sampling and analytic reconstruction severely limits the temporal resolution of 4D reconstructions. The number of 3D volumes (henceforth called time samples) of the 4D reconstruction per unit time is given by $F_s = F_c/N_\theta$ where N_θ is the number of views used to reconstruct each time sample, and F_c is the data acquisition rate i.e., the rate at which projection images are collected. The maximum data acquisition rate, F_c, depends on a wide variety of hardware constraints such as the camera frame rate, the data transfer rate, buffer memory sizes, etc. The number of views, N_θ, required for spatial Nyquist sampling of the projection data is $\pi/2$ times the number of detector pixels, N_p, in the sensor's field of view perpendicular to the rotation axis [24]. However, in practice, for Nyquist sampling the number of views, N_θ, is chosen to be approximately equal to the number of cross-axial sampled pixels, N_p [24]. This means that in order to reconstruct a single time sample from a sensor with a frame size of 1024×1024 (i.e., a 1 mega-pixel sensor) one needs $N_\theta = 1024$ views of the object, resulting in a temporal reconstruction rate that is reduced by a factor of 1024 relative to the data rate, F_c. Consequently, the data rate becomes the limiting factor on the temporal resolution in a typical 4D-SXCT imaging system.

In order to increase the temporal reconstruction rate, either the number of projections per rotation can be reduced or the range of view angles used for reconstruction can be decreased. However, if the number of views per rotation is reduced the signal is under-sampled and analytic reconstruction algorithms produce substantial artifacts due to aliasing [24, 29–31]. If the number of view angles used for a single reconstruction is reduced, the Fourier space of the object is not fully sampled which results in a missing wedge of spatial frequencies [32]. With analytic reconstruction algorithms, this missing wedge results in time-varying non-isotropic resolution that will typically produce severe reconstruction artifacts. Therefore, using conventional reconstruction algorithms with a traditional progressive view sampling approach presents a fundamental limitation on the temporal resolution that can be achieved for a given spatial resolution.

In order to improve the quality of reconstruction, several new sampling strategies have been proposed for other tomographic applications. Using 2D Nyquist sampling theory, it has been shown that using a hexagonal sampling strategy [33–35] for the Radon transform the number of sample points can be reduced. Alternatively, by formulating the data acquisition as a time-sequential process, Willis et al. show that for an object with localized temporal variation the sampling rate can be reduced using an optimally scrambled angular sampling order [36, 37]. Another novel angular sampling strategy called equally sloped tomography [38], has been shown to produce superior quality reconstructions when reconstructing from a set of projections spaced equally in a slope parameter rather than angle. In [39], Zheng et al. present a method of reducing the number of projections by identifying favorable views based on prior knowledge of the object. In [40–44], the authors discuss different compressed sensing approaches for tomography. In general, the methods which change the angular order are optimal when the time interval between successive views is independent of their angular separation. Thus, optimal view sampling strategies have been shown to improve the reconstruction quality, even with conventional reconstruction algorithms.

An alternate approach to improving reconstruction quality is to use more advanced model-based iterative reconstruction (MBIR) methods, which are based on the estimation of a reconstruction which best fits models of both the sensor measurements (i.e., the forward model) and the object (i.e., prior model) [1, 9, 10, 12, 45]. These 3D reconstruction methods have been shown to be very effective when the angular range is limited [46] and also when the number of views is less than that required by Nyquist sampling criterion [1]. In the context of medical CT, several authors have also shown that modeling the temporal correlations [47–51] in addition to modeling the spatial correlations improves the quality of 4D MBIR reconstructions.

In this chapter, we propose an approach to 4D reconstruction of time varying objects, which we call time interlaced model based iterative reconstruction (TIMBIR). TIMBIR is the synergistic combination of a novel interlaced view sampling technique with an innovative model-based iterative reconstruction (MBIR) algorithm. In [2] we present preliminary results using the TIMBIR method and in [52] we use TIMBIR to determine the morphology of a growing metallic dendrite in 4D. In the new interlaced view sampling method, all the views typically acquired over half a rotation using progressive view sampling are instead acquired over multiple half-rotations. Intuitively, interlaced view sampling spreads the view angles more evenly in time as opposed to the conventional progressive view sampling method that groups views at nearby angles together in time. Nonetheless, interlaced view sampling when used with conventional analytic reconstruction methods does not result in any gains since the number of views in each half-rotation is insufficient to achieve Nyquist sampling for a single reconstruction of the object. Consequently, analytic reconstruction methods produce severe artifacts when used to reconstruct the interlaced views at higher temporal rates.

In order to reconstruct the data acquired using the interlaced view sampling method, we propose a new 4D MBIR algorithm. In addition to modeling the measurement noise and spatio-temporal correlations in the 4D object, the MBIR algorithm reduces ring and streak artifacts by modeling the detector non-idealities [22, 23] and

measurement outliers caused by high energy photons (called zingers) [1]. We adapt our forward model for 3D MBIR introduced in [1] to the current 4D framework and combine it with a modified q-generalized Gaussian Markov random field (qGGMRF) [4] based prior model, which models the spatio-temporal correlations in the reconstruction, and formulate the MBIR cost function. We note that our qGGMRF prior model is similar to the total-variation (TV) prior used in compressive sensing (CS) methods [41, 44]. The parameters of our qGGMRF prior model can be adjusted to result in the TV prior that is widely used in CS methods. We then present a fast distributed parallel multi-resolution algorithm based on surrogate functions to minimize this cost function.

2.2 Interlaced View Sampling for Tomography

In order to satisfy the spatial Nyquist sampling requirement for each 3D time sample of a 4D reconstruction, it is typically necessary to collect approximately $N_\theta = N_p$ progressive views, where N_p is the number of sampled pixels perpendicular to the axis of rotation [24]. In the traditional approach, these N_θ progressive views are taken in sequence while the sample is rotated continuously over π radians. The object is then reconstructed at a temporal rate of $F_s = F_c/N_\theta$ where F_c is the data acquisition rate.

In contrast to this approach, we propose an interlaced view sampling method where each frame of data consisting of N_θ distinct views are acquired over K interlaced sub-frames (see Fig. 2.2). Each sub-frame of data then consists of N_θ/K equally spaced views, but together the full frame of data contains all N_θ distinct views of the object. For a continuously rotating object, the formula which gives the view angle as a function of the discrete sample index, n, is given by

$$\theta_n = \left[nK + \mathcal{B}_r \left(\left\lfloor \frac{nK}{N_\theta} \right\rfloor \bmod K \right) \right] \frac{\pi}{N_\theta}, \tag{2.1}$$

where K is a power of 2, $b = \mathcal{B}_r(a)$ is the bit-reverse function which takes the binary representation of the integer a and reverses the order of the bits to form the output

b [53]. It is interesting to note that there exist ordered subset methods [54] for 3D tomography which use a similar bit-reversal technique to determine the grouping of projection views. If we write a in base 2 expansion as $a = \sum_{i=0}^{l-1} b_i 2^i$ where $l = \log_2(K)$, then $b = \sum_{i=0}^{l-1} b_i 2^{l-1-i}$. If the angular range of the projections is limited to 2π radians, then the interlaced view angles are given by

$$\theta_n = \left[\left(n \bmod \frac{2N_\theta}{K} \right) K + \mathcal{B}_r \left(\left\lfloor \frac{nK}{N_\theta} \right\rfloor \bmod K \right) \right] \frac{\pi}{N_\theta}. \tag{2.2}$$

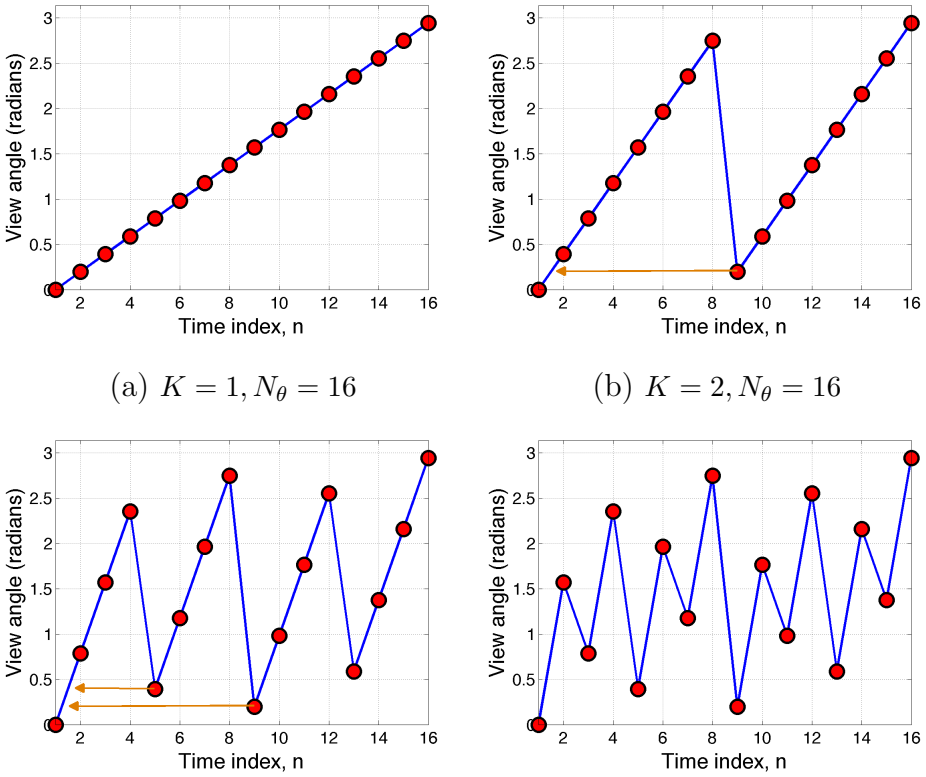
If the angular range of the projections is limited to π radians, then the interlaced view angles are given by

$$\theta_n = \left[\left(n \bmod \frac{N_\theta}{K} \right) K + \mathcal{B}_r \left(\left\lfloor \frac{nK}{N_\theta} \right\rfloor \bmod K \right) \right] \frac{\pi}{N_\theta}. \tag{2.3}$$

It is important to note that for the same index, n, the view angles generated by equations (2.1), (2.2), and (2.3) are all separated by an integer multiple of π radians. Hence, the projections obtained using equations (2.1), (2.2), and (2.3) are essentially the same. Fig. 2.2 compares progressive views with interlaced views and also highlights the interlacing of view angles across sub-frames. The object is then reconstructed at a temporal rate of $F_s = rF_c/N_\theta$ where r is the number of time samples of the 4D reconstruction in a frame. In TIMBIR, we reconstruct the object at a temporal rate which is r times the conventional rate of $F_s = F_c/N_\theta$. We typically set the parameter r equal to the number of sub-frames, K.

In TIMBIR, the conventional approach of reconstructing using progressive views is obtained when $r = K = 1$. We will show empirically that we can significantly improve the spatial and temporal reconstruction quality by increasing the value of K and r while using the same value of N_θ. Thus, we show that we can get significantly improved reconstruction quality by changing the view sampling method without increasing the amount of input data.

In progressive view sampling, the entire set of N_θ distinct views is acquired over a π radians rotation of the object. However, in interlaced view sampling all the distinct N_θ views are acquired over a $K\pi$ radians rotation of the object. Thus, if N_θ and F_c



(a) $K = 1, N_\theta = 16$ (b) $K = 2, N_\theta = 16$

(c) $K = 4, N_\theta = 16$ (d) $K = 16, N_\theta = 16$

Fig. 2.2. Illustration of interlaced view sampling pattern for different values of K. (a-d) are plots of $\theta_n \bmod (\pi)$ vs. time index, n, for $K = 1, 2, 4,$ and 16. The arrows show the relative difference between the angular values across sub-frames.

are fixed, then increasing the value of K increases the rotation speed of the sample. We will show that increasing the value of K can also improve the reconstruction quality. However, in some cases, the increased rotation speed of the sample may not be desirable. So in these cases, the parameter K can be adjusted to balance the need for improved image quality with the need to limit the sample's rotation speed. While increasing the rotation speed and reducing N_θ appears to be an intuitive step for increasing the temporal resolution of the reconstructions, we will empirically demonstrate that interlacing the views will give us significant improvements in the temporal reconstruction quality.

2.3 Formulation of MBIR cost function

The goal of SXCT is to reconstruct the attenuation coefficients of the sample from the acquired data. We reconstruct the attenuation coefficients of the object from the acquired data using the MBIR framework. The MBIR reconstruction is given by

$$\left(\hat{x}, \hat{\phi}\right) = \underset{x,\phi}{\operatorname{argmin}}\{-\log p(y|x,\phi) - \log p(x)\}, \tag{2.4}$$

where $p(y|x,\phi)$ is a pdf for the projection data, y, given the object, x, and the unknown system parameters, ϕ, and $p(x)$ is a pdf for the 4D object.

2.3.1 Measurement Model

We begin by deriving a likelihood function $p(y|x,\phi)$ for the projections, y, from a time varying object, x. We model each voxel of the object as an independent piecewise constant function in time such that there are r equi-length reconstruction time samples in each frame. Thus, the projections ranging from $(j-1)N_\theta/r + 1$ to jN_θ/r are assumed to be generated from the j^{th} time sample. The vector of attenuation coefficients of the object at the j^{th} time sample is denoted by x_j.

A widely used model for X-ray transmission measurements is based on Beer's law and Poisson counting statistics for the measurement [55]. Using this model, if $\lambda_{n,i}$ is the measurement at the i^{th} detector element and n^{th} view and if $\lambda_{D,i}$ is the measurement in the absence of the sample, then an estimate of the projection integral is given by $y_{n,i} = \log\left(\frac{\lambda_{D,i}}{\lambda_{n,i}}\right)$. If we denote y to be the vector of projections $y_{n,i}$ and x to be the vector of attenuation coefficients at all time steps, then using a Taylor series approximation to the Poisson log-likelihood function [56] it can be shown that,

$$-\log p\left(y|x\right) \approx \frac{1}{2}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M}\left(\left(y_{n,i} - A_{n,i,*}x_j\right)\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma}\right)^2 + f(y), \tag{2.5}$$

where $n_j = j\frac{N_\theta}{r} + 1$, L is the total number of time samples in the reconstruction, $A_{n,i,*}$ is the i^{th} row of the forward projection matrix A_n, Λ_n is a diagonal matrix modeling the noise statistics, M is the total number of detector elements, and $f(y)$

is a constant that will be ignored in the subsequent optimization. The variance of the projection measurement, $y_{n,i}$, is inversely proportional to the mean photon count and hence we set $\Lambda_{n,i,i} = \lambda_{n,i}$ [55]. Since $\lambda_{n,i}$ is not equal to the photon count but is proportional to the photon count, there exists a constant of proportionality σ such that $\frac{\lambda_{n,i}}{\sigma^2}$ is the inverse variance of the projection measurement, $y_{n,i}$.

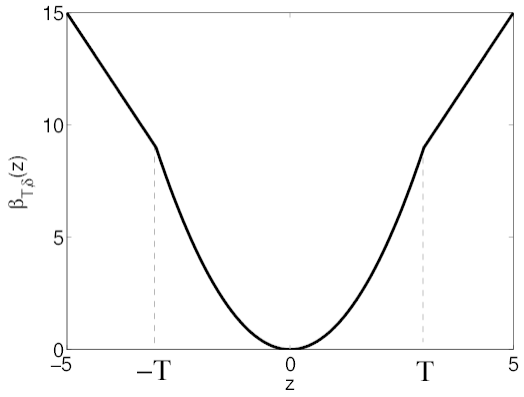


Fig. 2.3. Plot of the generalized Huber function $\beta_{T,\delta}$ used in the likelihood term with $T = 3$ and $\delta = \frac{1}{2}$. Projections with large data mismatch error are penalized thereby reducing their influence in the overall cost function.

While this model is useful in several applications, it does not account for the non-idealities in the synchrotron measurement system. In particular, the log-likelihood term in (2.5) corresponds to a quadratic penalty on the weighted data mismatch error and does not account for the occurrence of zingers [1]. The zinger measurements correspond to a distribution with heavier tails than that corresponding to (2.5). Hence we change the quadratic penalty in (2.5) to a generalized Huber penalty (see Fig. 2.3) of the form [1, 13]

$$\beta_{T,\delta}(z) = \begin{cases} z^2 & |z| < T \\ 2\delta T\,|z| + T^2(1-2\delta) & |z| \geq T, \end{cases} \tag{2.6}$$

where T and δ are parameters of the generalized Huber function. In our model, the parameters of the generalized Huber function are chosen such that $0 < \delta < 1$ and

$T > 0$. The generalized Huber function is non-convex in this range of parameter values. This penalty implies that if the ratio of the data mismatch error to the noise standard deviation is greater than the parameter T, then the measured projection corresponds to a zinger.

Next, we model the non-idealities in the measurement system that cause ring artifacts. It has been shown [25] that the non-idealities that cause ring artifacts can be modeled via an additive detector dependent offset error, d_i, in the projection measurements, $y_{n,i}$. Hence, if we assume an unknown offset error d_i in the projection, $y_{n,i}$, then an estimate of the line integral is given by,

$$\tilde{y}_{n,i} = y_{n,i} - d_i. \tag{2.7}$$

The offset error, d_i, is typically not known from the measurements and hence we jointly estimate it during reconstruction. Thus, the new likelihood function that models the offset error and the zinger measurements is given by

$$p(y|x,d,\sigma) = \frac{1}{Z(\sigma)} \exp\left\{-U(y,x,d,\sigma)\right\}, \tag{2.8}$$

$$\text{where } U(y,x,d,\sigma) = \frac{1}{2}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M}\beta_{T,\delta}\left((y_{n,i} - A_{n,i,*}x_j - d_i)\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma}\right), \tag{2.9}$$

$Z(\sigma)$ is a normalizing constant and $d = [d_1 \cdots d_M]$ is the vector of all offset error parameters. Since (2.8) is a pdf, we can show that $Z(\sigma) = Z(1)\sigma^{ML\frac{N_\theta}{r}}$ using the property that (2.8) integrated over y should be equal to one [57]. Thus, the log-likelihood function is given by

$$-\log p(y|x,d,\sigma) = U(y,x,d,\sigma) + ML\frac{N_\theta}{r}\log(\sigma) + \tilde{f}(y), \tag{2.10}$$

where $\tilde{f}(y)$ is a constant which is ignored in the subsequent optimization. We note that when $\delta = 0$, $p(y|x,d,\sigma)$ is not a density function since it does not integrate to 1 and hence we assume $\delta > 0$ in the rest of the chapter.

2.3.2 Prior Model

We use a q-generalized Gaussian Markov random field (qGGMRF) [4] based prior model for the voxels. The prior is used to model the 4D object in time as well as space. Using this model, the logarithm of the density function of x is given by

$$-\log p(x) = \sum_{j=1}^{L} \sum_{\{k,l\}\in\mathcal{N}} w_{kl}\rho_s(x_{j,k} - x_{j,l}) + \sum_{k=1}^{P} \sum_{\{j,i\}\in\mathcal{T}} \tilde{w}_{ji}\rho_t(x_{j,k} - x_{i,k}) + \text{constant}, \tag{2.11}$$

$$\text{where } \rho_s(\Delta) = \frac{\Delta_t\Delta_s^3 \left|\frac{\Delta}{\Delta_s\sigma_s}\right|^2}{c_s + \left|\frac{\Delta}{\Delta_s\sigma_s}\right|^{2-p}},\ \rho_t(\Delta) = \frac{\Delta_t\Delta_s^3 \left|\frac{\Delta}{\Delta_t\sigma_t}\right|^2}{c_t + \left|\frac{\Delta}{\Delta_t\sigma_t}\right|^{2-p}},$$

and $x_{j,k}$ is the k^{th} voxel of the object at time sample j, P is the total number of voxels in each time sample, $\mathcal{N}$ is the set of all pairwise cliques in 3D space (all pairs of neighbors in a 26 point spatial neighborhood system), $\mathcal{T}$ is the set of all pairs of indices of adjacent time samples (two point temporal neighborhood system), p, c_s, c_t, σ_s and σ_t are qGGMRF parameters, Δ_s is a parameter proportional to the side length of a voxel and Δ_t is a parameter proportional to the duration of each time sample in the reconstruction. The weight parameters are set such that $w_{kl} \propto |k-l|^{-1}$, $\tilde{w}_{ji} \propto |j-i|^{-1}$, and normalized such that $\sum_{l\in\mathcal{N}_k} w_{kl} + \sum_{i\in\mathcal{T}_j} \tilde{w}_{ji} = 1$, where $\mathcal{N}_k$ is the set of all spatial neighbors and $\mathcal{T}_j$ is the set of all temporal neighbors of voxel $x_{j,k}$. The terms Δ_s and Δ_t in the prior model ensures invariance of the prior to changing voxel sizes [58].

2.3.3 MBIR Cost Function

By substituting (2.10) and (2.11) into (2.4), we get the following MBIR cost function,

$$c(x,d,\sigma) = \frac{1}{2}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M}\beta_{T,\delta}\left((y_{n,i} - A_{n,i,*}x_j - d_i)\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma}\right)$$
$$+\sum_{j=1}^{L}\sum_{\{k,l\}\in\mathcal{N}} w_{kl}\rho_s(x_{j,k} - x_{j,l}) + \sum_{k=1}^{P}\sum_{\{j,i\}\in\mathcal{T}} \tilde{w}_{ji}\rho_t(x_{j,k} - x_{i,k}) + ML\frac{N_\theta}{r}\log(\sigma). \tag{2.12}$$

The reconstruction is obtained by jointly minimizing the cost, $c(x, d, \sigma)$, with respect to x, d and σ. Additionally, we impose a linear constraint of the form $Hd = 0$ to minimize any shift in the mean value of the reconstruction. The form of the matrix H can be adjusted depending on the application. The form of the matrix H used in our application is described in appendix A.

2.4 Optimization Algorithm

The cost function (2.12) is non-convex in x, d and σ. Minimizing the current form of the cost function given by (2.12) is computationally expensive. So, instead we use the functional substitution approach [59, 60] to efficiently minimize (2.12). Our method also ensures monotonic decrease of the cost function (2.12). A substitute cost function $c_{\text{sub}}(x, d, \sigma; x', d', \sigma')$ to the cost function $c(x, d, \sigma)$ at the point (x', d', σ') is a function which bounds the cost function from above such that minimizing the substitute cost function results in a lower value of the original cost function.

2.4.1 Construction of Substitute Function

To derive a substitute function to the overall cost we find a substitute function to each term of the cost (2.12) and sum them together to derive an overall substitute function. In particular, we will use quadratic substitute functions, as they make the subsequent optimization computationally simple.

A sufficient condition for a function $q(z; z')$ to be a substitute function to $g(z)$ at the point z' is that $\forall z$, $q(z; z') \geq g(z)$ and $q(z'; z') = g(z')$. We can prove that

$$Q_{T,\delta}(z; z') = \begin{cases} z^2 & |z'| < T \\ \frac{\delta T}{|z'|} z^2 + \delta T \, |z'| + T^2(1 - 2\delta) & |z'| \geq T \end{cases}$$

is a substitute function to $\beta_{T,\delta}(z)$ at the point z' by showing that it satisfies the sufficiency condition [13]. If the error sinogram is defined as $e_{j,n,i} = y_{n,i} - A_{n,i,*}x_j - d_i$ and $e'_{j,n,i} = y_{n,i} - A_{n,i,*}x'_j - d'_i$ is the error sinogram at the current values of (x', d', σ'),

function $[x', e', b'] \leftarrow$ UPDATEVOX(x', e', b', σ', j, m)

for each voxel k in voxel-line m at time step j **do**

Compute $\tilde{\theta}_1$ and $\tilde{\theta}_2$ as in equation (2.19)

Compute substitute function parameters

for each $l \in \mathcal{N}_k$ **do**

Compute a^s_{jkl} using (2.14)

end for

for each $i \in \mathcal{T}_j$ **do**

Compute a^t_{kji} using (2.16)

end for

Compute $\hat{x}_{j,k}$ as in equation (2.21)

$u^* \leftarrow x'_{j,k} + \alpha(\hat{x}_{j,k} - x'_{j,k})$

Update all variables affected by update of $x'_{j,k}$

$e'_{j,n,i} \leftarrow e'_{j,n,i} - A_{n,i,j}(u^* - x'_{j,k}) \quad \forall n, i$

$x'_{j,k} \leftarrow u^*$

Using (2.18), update $b'_{j,n,i} \quad \forall n, i$

end for

end function

Fig. 2.4. Pseudo code to update a voxel. The voxel update is obtained by minimizing a quadratic substitute function in $x_{j,k}$ while the other variables are treated as constants.

then a substitute function to $\beta_{T,\delta}\left(e_{j,n,i}\sqrt{\Lambda_{n,i,i}}/\sigma\right)$ in the original cost is given by $Q_{T,\delta}\left(e_{j,n,i}\sqrt{\Lambda_{n,i,i}}/\sigma; e'_{j,n,i}\sqrt{\Lambda_{n,i,i}}/\sigma'\right)$ by Property 7.9 in [61].

A quadratic substitute function for the prior term $\rho_s(x_{j,k} - x_{j,l})$ can be shown to be [9],

$$\rho_s(x_{j,k} - x_{j,l}; x'_{j,k} - x'_{j,l}) = \frac{a^s_{jkl}}{2}(x_{j,k} - x_{j,l})^2 + b^s_{jkl}, \tag{2.13}$$

function $[\hat{x}, \hat{d}, \hat{\sigma}] \leftarrow$ RECONSTRUCT(y, x', d', σ')

%**Inputs**: Projections y, Initial reconstruction x', Initial estimate of offset error d', Initial estimate of variance parameter σ'

%**Outputs**: Reconstruction $\hat{x}$, estimate of offset error $\hat{d}$, estimate of variance parameter $\hat{\sigma}$

$e'_{j,n,i} = y_{n,i} - A_{n,i,*}x'_j - d'_i \quad \forall i, j, n$

Set b' using (2.18)

while Stopping criteria is not met **do**

Iteratively update all voxel lines in a random order

for each voxel-line m at each time step j **do**

$(x', e', b') \leftarrow$ UPDATEVOX($x', e', b', \sigma', j, m$)

end for

$d' \leftarrow$ Update d' using (2.24)

Update σ' using (2.25)

end while

$\hat{x} \leftarrow x'$, $\hat{d} \leftarrow d'$, $\hat{\sigma} \leftarrow \sigma'$

end function

Fig. 2.5. Pseudo code of the 4D MBIR algorithm. The algorithm works by alternatively minimizing a substitute cost function (2.17) with respect to all the voxels, $x_{j,k}$, offset error, d_i, and variance parameter σ.

$$\text{where } a^s_{jkl} = \begin{cases} \frac{\rho'_s(x'_{j,k} - x'_{j,l})}{(x'_{j,k} - x'_{j,l})} & x'_{j,k} \neq x'_{j,l} \\ \rho''_s(0) & x'_{j,k} = x'_{j,l} \end{cases}, \tag{2.14}$$

and b^s_{jkl} is a constant. A quadratic substitute function for the prior term $\rho_t(x_{j,k} - x_{i,k})$ can be shown to be [9],

$$\rho_t(x_{j,k} - x_{i,k}; x'_{j,k} - x'_{i,k}) = \frac{a^t_{kji}}{2}(x_{j,k} - x_{i,k})^2 + b^t_{kji}, \tag{2.15}$$

$$\text{where } a^t_{kji} = \begin{cases} \frac{\rho'_t(x'_{j,k}-x'_{i,k})}{(x'_{j,k}-x'_{i,k})} & x'_{j,k} \neq x'_{i,k} \\ \rho''_t(0) & x'_{j,k} = x'_{i,k} \end{cases}, \tag{2.16}$$

and b^t_{kji} is a constant. Thus, a substitute function to (2.12) is given by,

$$c_{\text{sub}}(x,d,\sigma;x',d',\sigma') = \frac{1}{2}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M} Q_{T,\delta}\left(e_{j,n,i}\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma}; e'_{j,n,i}\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma'}\right)$$
$$+ML\frac{N_\theta}{r}\log(\sigma)+\sum_{j=1}^{L}\sum_{\{k,l\}\in\mathcal{N}} w_{kl}\rho_s(x_{j,k}-x_{j,l};x'_{j,k}-x'_{j,l})+\sum_{k=1}^{P}\sum_{\{j,i\}\in\mathcal{T}}\tilde{w}_{ji}\rho_t(x_{j,k}-x_{i,k};x'_{j,k}-x'_{i,k}). \tag{2.17}$$

2.4.2 Parameter Updates used in Optimization

To minimize the cost function given by (2.12), we repeatedly construct and minimize (2.17) with respect to each voxel, $x_{j,k}$, the offset error parameters, d_i, and the variance parameter, σ. To simplify the updates, we define $b'_{j,n,i}$ to be a indicator variable which classifies measurements as anomalous based on the current value of the error sinogram, $e'_{j,n,i}$, and the current value of the variance parameter, σ', as shown below,

$$b'_{j,n,i} = \begin{cases} 1 & \left|e'_{j,n,i}\sqrt{\Lambda_{n,i,i}}/\sigma'\right| < T \\ 0 & \left|e'_{j,n,i}\sqrt{\Lambda_{n,i,i}}/\sigma'\right| \geq T \end{cases} \tag{2.18}$$

and let $\tilde{b}'_{j,n,i} = 1 - b'_{j,n,i}$.

Voxel Update

In order to minimize (2.17) with respect to a voxel k at time step j, we first rewrite (2.17) in terms of $x_{j,k}$ in the form of the following cost function,

$$\tilde{c}_{\text{sub}}(x_{j,k}) = \tilde{\theta}_1 x_{j,k} + \frac{\tilde{\theta}_2}{2}\left(x_{j,k}-x'_{j,k}\right)^2$$
$$+\sum_{l\in\mathcal{N}_k} w_{kl}\rho_s(x_{j,k}-x'_{j,l};x'_{j,k}-x'_{j,l})+\sum_{i\in\mathcal{T}_j}\tilde{w}_{ji}\rho_t(x_{j,k}-x'_{i,k};x'_{j,k}-x'_{i,k}), \tag{2.19}$$

$$\text{where } \tilde{\theta}_1 = -\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M} A_{n,i,k}u'_{j,n,i},\ \tilde{\theta}_2 = \sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M} A^2_{n,i,k}v'_{j,n,i},$$

$$u'_{j,n,i} = \frac{\sqrt{\Lambda_{n,i,i}}}{\sigma'}\left[b'_{j,n,i}e'_{j,n,i}\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma'} + \tilde{b}'_{j,n,i}\delta T \text{sgn}\left(e'_{j,n,i}\right)\right],$$

$$v'_{j,n,i} = \frac{\sqrt{\Lambda_{n,i,i}}}{\sigma'}\left[b'_{j,n,i}\frac{\sqrt{\Lambda_{n,i,i}}}{\sigma'} + \tilde{b}'_{j,n,i}\frac{\delta T}{\left|e'_{j,n,i}\right|}\right], \tag{2.20}$$

and sgn is the *signum* function. Then, the optimal update for $x_{j,k}$ is obtained by minimizing (2.19) with respect to $x_{j,k}$ and is given by

$$\hat{x}_{j,k} = \frac{\sum_{l\in\mathcal{N}_k} w_{kl}a^s_{jkl}x'_{j,l} + \sum_{i\in\mathcal{T}_j}\tilde{w}_{ji}a^t_{kji}x'_{i,k} + \tilde{\theta}_2 x'_{j,k} - \tilde{\theta}_1}{\sum_{l\in\mathcal{N}_k} w_{kl}a^s_{jkl} + \sum_{i\in\mathcal{T}_j}\tilde{w}_{ji}a^t_{kji} + \tilde{\theta}_2}. \tag{2.21}$$

In order to speed up the computation of voxel updates, we update certain groups of voxels sequentially. In particular, let w be the axis of rotation and $u-v$ axes be in the plane perpendicular to w. A voxel-line [9] consists of all voxels along w-axis which share the same $u-v$ coordinate. Since all voxels along a voxel-line share the same geometry computation in the $u-v$ plane, we update all the voxels along a voxel-line sequentially. To further speed up convergence of the algorithm, we use the over-relaxation method [9] which forces larger updates of voxels. The update for voxel $x_{j,k}$ is then given by

$$x'_{j,k} \leftarrow x'_{j,k} + \alpha\left(\hat{x}_{j,k} - x'_{j,k}\right), \tag{2.22}$$

where α is the over-relaxation factor which is set to be equal to 1.5. Typically, α is chosen in the range of $(1,2)$ which ensures decreasing values of (2.12) [9]. The pseudo code to update a voxel-line is shown in Fig. 2.4.

Offset Error Update

To minimize (2.17) with respect to d, subject to the constraint $Hd = 0$, we first rewrite (2.17) in terms of the offset error d as,

$$\tilde{c}_{sub}(d) = \frac{1}{2}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\left(d-\bar{d}_{j,n}\right)^t V'_{j,n}\left(d-\bar{d}_{j,n}\right) \tag{2.23}$$

where $\bar{d}_{j,n,i} = e'_{j,n,i} + d'_i$ and $V'_{j,n}$ is a diagonal matrix such that $V'_{j,n,i,i} = v'_{j,n,i}$ (given by (2.20)). We use the theory of Lagrange multipliers [62] to minimize (2.23) with respect to the offset error parameters, d_i, subject to the constraint $Hd = 0$. The optimal update for d is given by

$$d' \leftarrow \Omega^{-1}\left(\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1} V'_{j,n}\bar{d}_{j,n} \; - H^t\left(H\Omega^{-1}H^t\right)^{-1} H\Omega^{-1}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1} V'_{j,n}\bar{d}_{j,n}\right) \tag{2.24}$$

where $\Omega = \sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1} V'_{j,n}$.

Variance Parameter Update

The update for the variance parameter, σ^2, is obtained by taking the derivative of (2.17) with respect to σ^2 and setting it to zero. The update is then given by,

$$\sigma'^2 \leftarrow \frac{r}{N_\theta LM}\sum_{j=1}^{L}\sum_{n=n_{j-1}}^{n_j-1}\sum_{i=1}^{M}\left[e'^2_{j,n,i}\Lambda_{n,i,i}b'_{j,n,i} + \tilde{b}'_{j,n,i}\delta T\left|e'_{j,n,i}\right|\sigma'\sqrt{\Lambda_{n,i,i}}\right]. \tag{2.25}$$

Thus, the pseudo code of the optimization algorithm to minimize (2.12) is shown in Fig. 2.5. The sequence of costs is convergent since the surrogate function approach ensures monotonic decrease of the original cost function (2.12). We have also empirically observed that the reconstructions converge. In some cases, theoretical convergence proofs exist for majorization techniques with alternating minimization [63, 64]. However, due to the complicated nature of our cost function we have no theoretical proof of convergence for our algorithm.

Finally, we implemented non-homogeneous iterative coordinate descent (NHICD) [9] to improve the convergence speed of the algorithm. NHICD works by more frequently updating those voxels which have a greater need for updates. We also use

multi-resolution initialization [13, 65] in which reconstructions at coarser resolutions are used to initialize reconstructions at finer resolution to improve convergence speed. The multi-resolution method transfers the computational load to the coarser scales where the optimization algorithm is faster due to the reduced dimensionality of the problem. Furthermore, we use bilinear interpolation to up-sample reconstructions at coarser resolutions to finer resolutions.

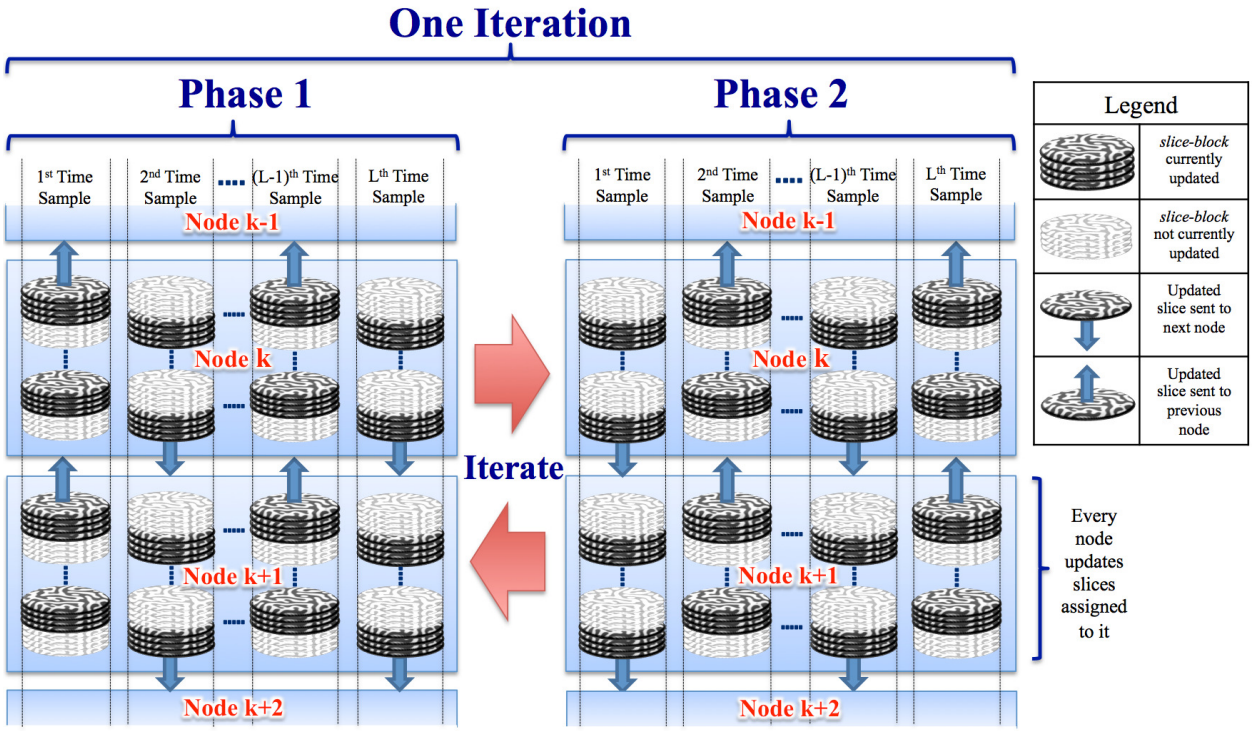


Fig. 2.6. Block diagram describing the distributed parallel MBIR algorithm. The data y is split axially among multiple nodes. Each node then updates only those slices of the reconstruction that depend on its share of the data. In phase 1, all the even numbered slice-blocks are updated at even time samples and all the odd numbered slice blocks are updated at odd time samples. In phase 2, all the odd numbered slice-blocks are updated at even time samples and all the even numbered slice blocks are updated at odd time samples. The algorithm iterates between the two phases until convergence is achieved.

2.4.3 Algorithm Initialization

Since the MBIR cost function is non-convex, using a reasonable initial estimate is important to obtain a reasonable solution. At the coarsest scale of the multi-resolution algorithm, the variance parameter, σ, is initialized to one, the offset error, d_i, is set to zero and all voxels are initialized to zero. Furthermore, to prevent the algorithm from converging to a local minimum, we do not update the offset error, d_i, and variance parameter, σ, at the coarsest scale. At the i^{th} multi-resolution stage, the prior model parameter Δ_s is set to 2^{S-i} where S is the total number of multi-resolution stages and the parameter Δ_t is set to N_θ/r.

The algorithm stops when the ratio of average magnitude of voxel updates to the average magnitude of voxel values is less than a pre-defined threshold (convergence threshold). We use different values for the convergence threshold at different multi-resolution stages. If the convergence threshold at the finest multi-resolution stage is T, then the convergence threshold at the k^{th} multi-resolution stage is chosen to be $T/(S-k+1)$.

2.5 Distributed parallelization of MBIR

To enable high spatial and temporal resolution reconstructions typically required for 4D-SXCT, we propose a distributed parallel MBIR algorithm optimized to run on a high performance cluster (HPC) consisting of several supercomputing nodes. Each node consists of one or more multi-core processors with a single shared memory. In the distributed MBIR algorithm, the computation is distributed among several nodes so that the data dependency across nodes is minimized. If a node requires information from another node, then the information is explicitly communicated using a message passing interface (MPI). The multiple cores in each node are used to further speed up computation using OpenMP based shared memory parallelization.

Each iteration of the distributed MBIR algorithm consists of two update phases as shown in Fig. 2.6. All the voxels updated in the first phase in Fig. 2.6 are not

updated in the second phase and all the voxels updated in the second phase are not updated in the first phase. The algorithm alternates between the two update phases until convergence is achieved.

Due to the specific form of the voxel updates in (2.21), we need to ensure that only those voxels that do not share sinogram[1] entries or spatio-temporal neighbors are updated in parallel to ensure that the cost function decreases with each update. We equally divide all the slices in each time sample of the reconstruction (along w-axis) across multiple nodes as shown in Fig. 2.6. Each node only updates voxels in its share of the 3D volume at all the time samples. The projection data, y, is distributed such that each node only stores those projection slices in its local memory which depend on its share of the reconstruction. Such a distribution ensures that voxel updates in a node do not require sinogram entries from another node. Thus, there is no sinogram dependency across nodes. However, there is a voxel neighborhood dependency across nodes since the voxel slices which are on the "edge" of a 3D volume within a node (along w-axis) contain spatial neighbors which are stored in other nodes. Additionally, within a given node there are spatial and temporal neighbors which we need to ensure are not updated in parallel.

To address this problem, we first equally divide the slices in each node into an even number of blocks which we call "slice-blocks" as shown in Fig. 2.6. In each phase, the different darkly shaded slice-blocks in Fig. 2.6 are updated in parallel and the lightly shaded slice-blocks are updated in the next phase. The voxels within a slice-block are updated serially taking into advantage the same geometry computation along a "voxel-line" [9] as explained in the previous section. In phase 1, all the odd numbered slice-blocks at odd time samples and even numbered slice-blocks at even time samples are updated in parallel. In node k, after the updates of phase 1, the first slice of each odd time sample is communicated to node $k-1$ and the last slice of each even time sample is communicated to node $k+1$. In phase 2, all the even numbered slice-blocks

[1]A sinogram image is a column wise stack of a single projection slice (i.e., projection values along a single detector line) across all the views.

at the odd time samples and the odd numbered slice-blocks at the even time samples are updated in parallel. In node k, after the updates of phase 2, the last slice of each odd time sample is communicated to node $k+1$ and the first slice of each even time sample is communicated to node $k-1$. All the slice-blocks within a node are updated in parallel using a OpenMP based shared memory parallelization scheme. This update scheme ensures that the cost-function (2.12) decreases monotonically with every iteration.

After updating the slice-blocks in each phase, only those elements of the error sinogram, e, and offset error, d, dependent on the last updated slice-blocks are updated. To update the variance parameter, σ, each node computes the summation in equation (2.25) corresponding to its share of the data and communicates the result to a master node. The master node then updates σ by accumulating the result from all the nodes and broadcasts the updated value back to the nodes. The updated values of slices from node $k-1$ and node $k+1$ required by node k are communicated using MPI calls. The offset error, d, is updated when the slices are communicated across nodes. Thus, the compute nodes are not idle during communication. In all our experiments, the time taken for computing the offset error update was found to be greater than the time taken for communicating the slices. Thus, the algorithm is limited by computational speed and not by the time taken for communication and synchronization.

2.6 Experimental Results

2.6.1 Simulated Data Set

In this section, we compare FBP and MBIR reconstructions of simulated datasets using both the traditional progressive view sampling and the proposed interlaced view sampling methods. The simulated dataset is generated from a time-varying phantom by accurately modeling the data acquisition in a real physical system. First, a 2D phantom is generated using the Cahn-Hilliard equation which models the process of

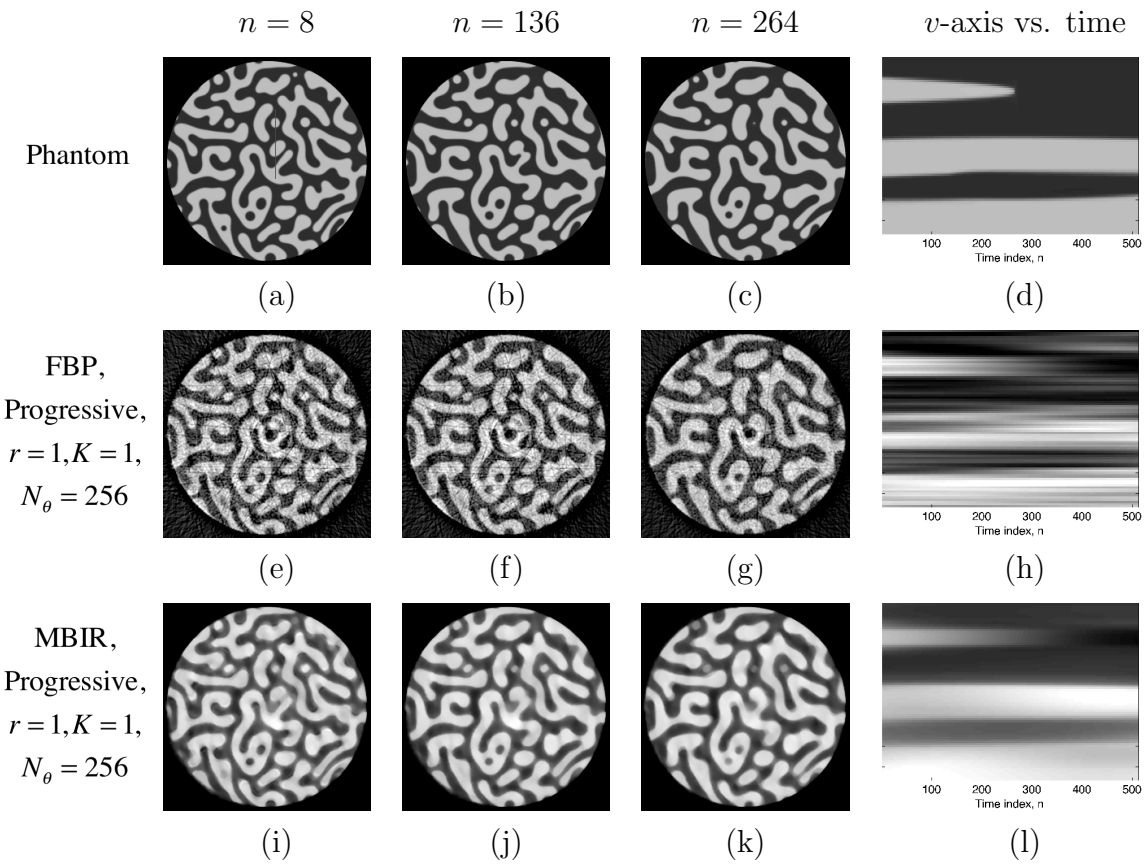


Fig. 2.7. The conventional approach to reconstruction using progressive views. (a-d) shows the time-varying phantom with two attenuation phases. Reconstruction of progressive views with $N_\theta = 256$ using FBP is shown in (e-h) and using MBIR is shown in (i-l) ($K = 1, N_\theta = 256, r = 1$). The first three columns show a $u - v$ slice of the object sampled at different time instants. The last column shows the time evolution of a v-axis slice of the object (along the red line in (a)). Reconstruction of rapidly time-varying objects using conventional methods results in poor temporal resolution and spatial artifacts.

phase separation in the cross-axial plane ($u - v$ axes) [66]. The two phases of the object have attenuation coefficients of 2.0 mm^{-1} and 0.67 mm^{-1} respectively. The 3D phantom is then generated by repeatedly stacking the 2D phantom along the axial dimension ($w-$axis). This phantom is representative of the phenomenon that we are interested in studying in 4D. A $u - v$ slice of the phantom at different time instants is

shown in Fig. 2.7 (a - c) and Fig. 2.7 (d) shows a v-axis slice of the phantom (along the red line in Fig. 2.7 (a)) as a function of time. The phantom is assumed to have a voxel resolution of $0.65 \times 0.65 \times 0.65$ μm^3 and a size of $N_w \times N_v \times N_u = 16 \times 1024 \times 1024$. The phantom is sampled in time at the data acquisition rate, F_c, and the projections are generated by forward projecting the sampled phantom at the appropriate angles. To simulate the detector non-idealities in SXCT, we add an offset error d_i to the projection $y_{n,i}$ at every n^{th} view. To simulate the effect of zingers, we randomly set 0.1% of the projections, y, to zero. We also simulate the measurement data $\lambda_{n,i}$ to have Poisson statistics. The simulated value of the variance parameter is $\sigma^2 = 10$.

The simulated sensor has a resolution of $N_p = 256$ pixels in the cross-axial direction and 4 pixels in the axial direction. A 3D time sample of the 4D reconstruction has a voxel resolution of $2.6 \times 2.6 \times 2.6$ μm^3 and a size of $N_w \times N_v \times N_u = 4 \times 256 \times 256$. The temporal reconstruction rate is $F_s = rF_c/N_\theta$ where r is the number of time samples of the reconstruction in a frame, F_c is the data acquisition rate, and N_θ is the number of distinct views acquired. Since the temporal reconstruction rate varies with r, the reconstructions are up sampled in time to the data acquisition rate, F_c, using cubic interpolation and then compared with the phantom. Also, since the phantom has higher spatial resolution than the reconstructions, the phantom is down-sampled by averaging over blocks of pixels to the reconstruction resolution before comparison. The regularization parameters in the prior model are chosen such that they minimize the root mean square error (RMSE) between the reconstruction and the phantom. The parameter p of the qGGMRF model is set to 1.2 and the parameters of the generalized Huber function are set to be $\delta = 0.5$ and $T = 4$. We use a convergence threshold of $T = 0.01$.

The traditional approach to 4D-SXCT is to use progressive view sampling and an analytic reconstruction algorithm such as FBP. So, we first generate a dataset of progressive views satisfying the Nyquist spatial sampling criterion and reconstruct it using FBP and MBIR algorithms. For analytic algorithms, the Nyquist view sampling requirement for a sensor of size $N_p = 256$ in the cross-axial dimension is to acquire

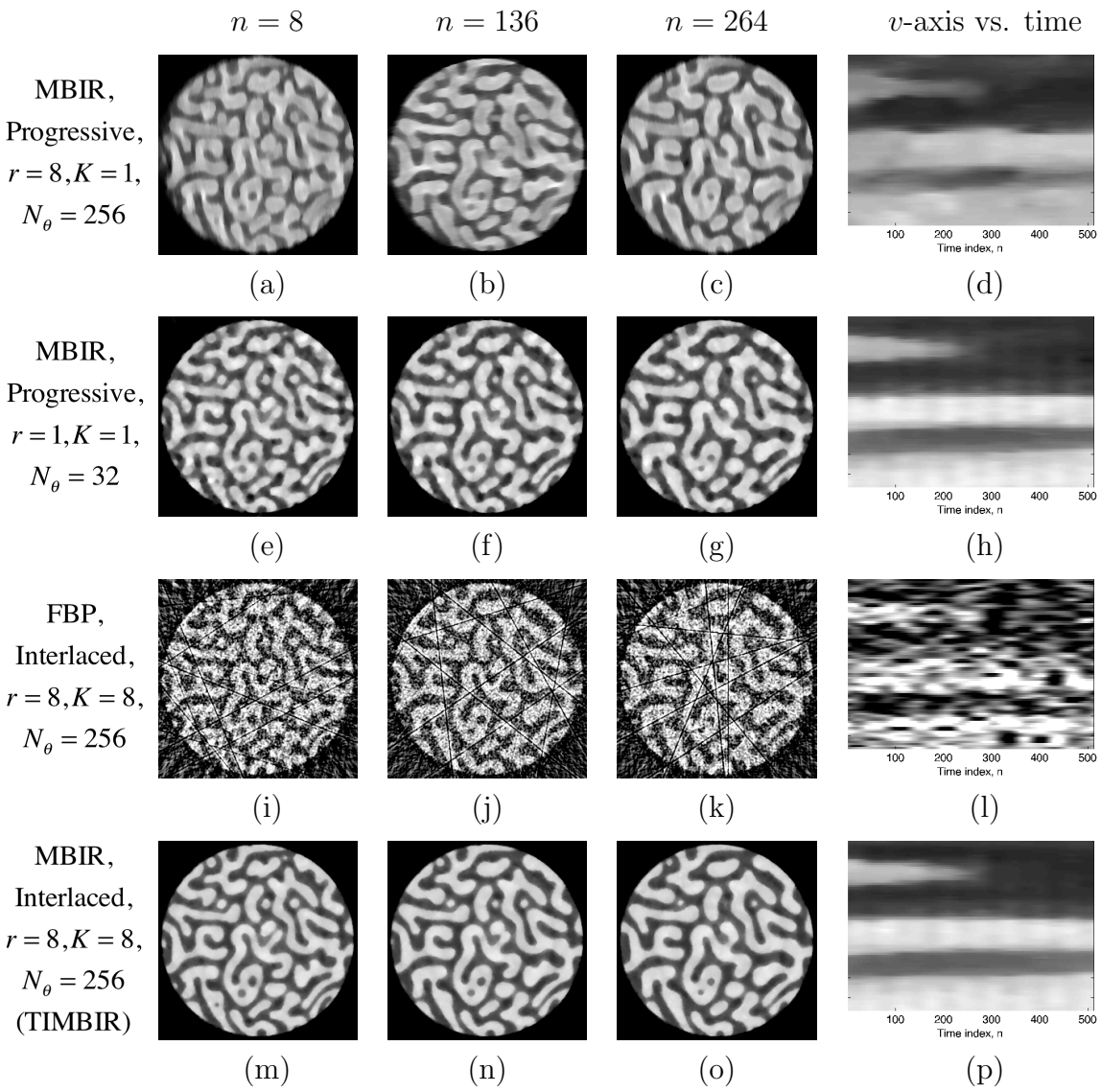


Fig. 2.8. Comparison of TIMBIR with other approaches to high temporal resolution reconstruction. The first three columns show a $u - v$ slice of the 4D reconstruction at different times and the last column shows a v-axis slice of the reconstruction versus time. All reconstructions have a temporal reconstruction rate of $F_s = F_c/32$. (a-d) is MBIR reconstruction ($r = 8$) of progressive views with $N_\theta = 256$. (e-h) is MBIR reconstruction ($r = 1$) of progressive views with $N_\theta = 32$. The reconstruction ($r = 8$) of interlaced views with $K = 8, N_\theta = 256$ using FBP is shown in (i-l) and using MBIR (TIMBIR) is shown in (m-p). Progressive view sampling results in poor reconstruction quality in time and/or space. Moreover, interlaced view sampling combined with FBP still causes severe artifacts due to under-sampling of view angles. However, TIMBIR produces high quality reconstructions in both time and space.

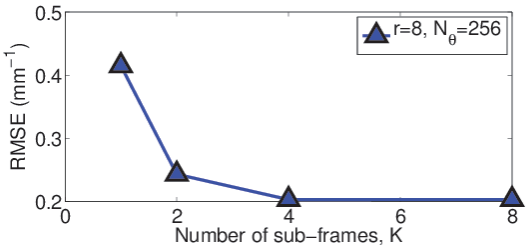


Fig. 2.9. Illustration of the effect of K on the RMSE between the reconstruction and the phantom when $r = 8$ and $N_\theta = 256$. Since the RMSE reduces as K is increased, interlacing of views is vital when reconstructing at a higher temporal rate of $F_s = F_c/32$ ($r = 8$).

projections at $N_\theta = 256$ distinct angles using progressive view sampling. Thus, we generate a dataset of progressive views with $N_\theta = 256$ and then reconstruct it at a temporal rate of $F_s = F_c/256$ by reconstructing $r = 1$ time sample every frame. The FBP reconstruction of this dataset is shown in Fig. 2.7 (e - h) and the MBIR reconstruction is shown in Fig. 2.7 (i - l). When compared to FBP, MBIR produces lower noise reconstructions while preserving the spatial resolution (Fig. 2.7 (e - g) and Fig. 2.7 (i - k)). Furthermore, FBP reconstructions suffer from strong ring and streak artifacts. From Fig. 2.7 (h) & (l), we can also see that neither FBP nor MBIR are able to reconstruct temporal edges accurately with progressive views.

Next, we investigate different methods of increasing temporal resolution using progressive views. First, we reconstruct a dataset of progressive views satisfying the Nyquist criterion at a temporal rate faster than the conventional method. Thus, we reconstruct a progressive view dataset with parameter $N_\theta = 256$ using MBIR at a rate of $F_s = F_c/32$ by reconstructing $r = 8$ time samples every frame. As shown in Fig. 2.8 (a - c), the reconstruction obtained using this method has strong artifacts due to a missing wedge of view angles in the data used to reconstruct every time sample of the reconstruction. In the next approach, we reduce the number of distinct angles in the progressive view dataset to $N_\theta = 32$ and reconstruct at a rate of $F_s = F_c/32$

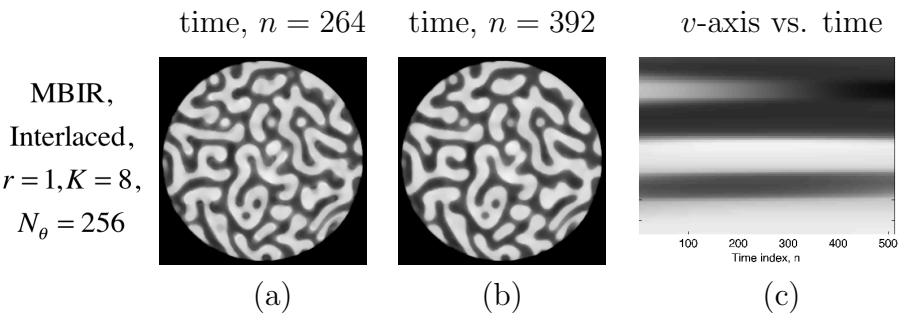


Fig. 2.10. Effect of interlacing of views on the reconstruction quality when reconstructing at the conventional rate of one time sample every frame ($r = 1$). (a) and (b) show a $u-v$ slice at two different time instants and (c) shows a v-axis slice of the reconstruction versus time. Interlacing of views improves the reconstruction quality even when reconstructing at the conventional rate using MBIR.

by reconstructing $r = 1$ time sample every frame. In this case, due to the severe under sampling of views, the MBIR reconstruction suffers from severe loss in quality as shown in Fig. 2.8 (e - h). This illustrates that merely reducing the number of views every π radians and using an advanced reconstruction algorithm such as MBIR is insufficient for our problem. Thus, there is an inherent sub-optimality in using progressive view sampling to achieve high temporal resolution reconstructions.

Finally, we investigate the effect of interlaced views on the spatial and temporal reconstruction quality when reconstructing using FBP and MBIR algorithms. We reconstruct a dataset of interlaced views in which each frame of $N_\theta = 256$ distinct angles is interlaced over $K = 8$ sub-frames. The object is then reconstructed at a rate of $F_s = F_c/32$ by reconstructing $r = 8$ time samples every frame. The reconstruction using FBP is shown in Fig. 2.8 (i - l) and using MBIR is shown in Fig. 2.8 (m - p). We can see that reconstructing the interlaced views with FBP results in extremely poor quality reconstructions. In contrast, MBIR with interlaced views (TIMBIR) results in a substantially better reconstruction of the object with minimal artifacts as shown in Fig. 2.8 (m - p). Furthermore, we can see that TIMBIR is able to more accurately reconstruct temporal edges (Fig. 2.8 (p)) than other methods (Fig. 2.8 (d, h & l) and

Fig. 2.7 (h,l)). Thus, by comparing Fig. 2.8 (m - p) with Fig. 2.8 (a - h) and Fig. 2.7 (i - l), we can conclude that interlaced view sampling is superior to progressive view sampling when reconstructing using MBIR. Furthermore, by comparing Fig. 2.8 (i - l) with Fig. 2.7 (e - h), we can conclude that combining interlaced view sampling with FBP does not result in any improvements. In Fig. 2.11, we plot a single voxel as a function of time for different values of r, K, and N_θ. We can see that TIMBIR produces the most accurate reconstruction of the voxel as a function of time among all the methods. The root mean squared errors (RMSE) between the reconstructions and the phantom ground-truth shown in Table 2.1 support these visual conclusions. Thus, TIMBIR with its synergistic combination of interlaced sampling and MBIR reconstruction results in a much higher quality reconstruction than either method can achieve by itself.

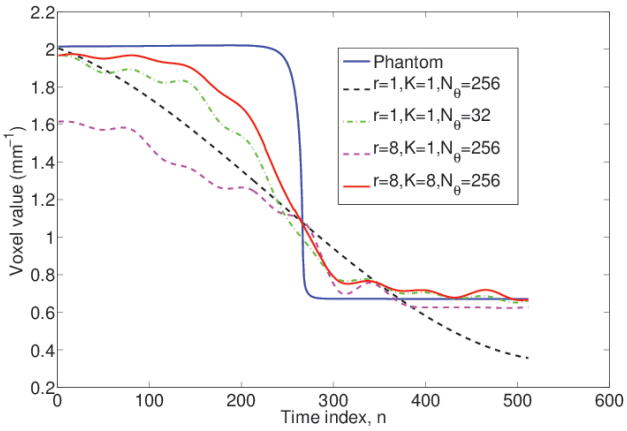


Fig. 2.11. Plot of a voxel as a function of time for different values of r, K, and N_θ. The most accurate reconstruction of the voxel as a function of time is obtained for the case of TIMBIR with parameters $r = 8, K = 8$, and $N_\theta = 256$.

Fig. 2.9 shows the RMSE between the MBIR reconstructions and the phantom as a function of K when $N_\theta = 256$ and $r = 8$. We can see that the RMSE reduces as

Table 2.1
Root Mean Square Error between the reconstruction and the phantom. TIMBIR has the lowest RMSE among all the methods

Description	Parameters	RMSE (mm^{-1})
Nyquist Progressive, FBP ($F_s = F_c/256$)	$K = 1, N_\theta = 256$ $r = 1$	0.3376
Nyquist Progressive, MBIR ($F_s = F_c/256$)	$K = 1, N_\theta = 256$ $r = 1$	0.2182
Nyquist Progressive, MBIR ($F_s = F_c/32$)	$K = 1, N_\theta = 256$ $r = 8$	0.3302
Sub-Nyquist Progressive, MBIR ($F_s = F_c/32$)	$K = 1, N_\theta = 32$ $r = 1$	0.2393
Interlaced, FBP ($F_s = F_c/32$)	$K = 8, N_\theta = 256$ $r = 8$	0.6035
Interlaced, MBIR ($F_s = F_c/32$) (TIMBIR)	$K = 8, N_\theta = 256$ $r = 8$	0.2119

the number of sub-frames, K, is increased from 1 to 8. This shows that interlacing of views is vital to achieving high temporal resolution reconstructions.

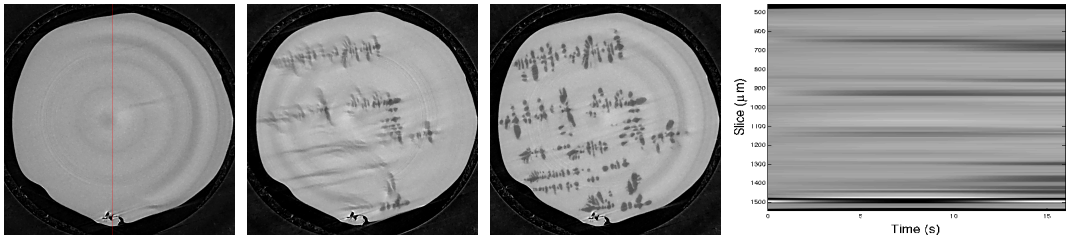
Even while reconstructing at the conventional rate of $F_s = F_c/256$ using $r = 1$, we can see that the RMSE reduces as K is increased. The reconstruction of interlaced views with parameters $K = 16, N_\theta = 256$ using $r = 1$ is shown in Fig. 2.10. By comparing Fig. 2.7 (g - i) with Fig. 2.10, we can see that interlacing of views reduces artifacts and improves reconstruction quality even when $r = 1$. Furthermore, the better spatial resolution at $r = 1$ essentially causes the RMSE curve corresponding to $r = 1$ to be lower to that corresponding to $r = 16$. Thus, interlacing of views improves the reconstruction quality when reconstructing at the conventional rate ($r = 1$) as well as at the improved rate ($r = 16$).

2.6.2 Real Data Set

FBP, Progressive Views $(r = 1, K = 1, N_\theta = 1536)$

(a) time $= 0s$ (b) time $= 8s$ (c) time $= 16s$ (d) v-axis vs. time

MBIR, Progressive Views $(r = 1, K = 1, N_\theta = 1536)$

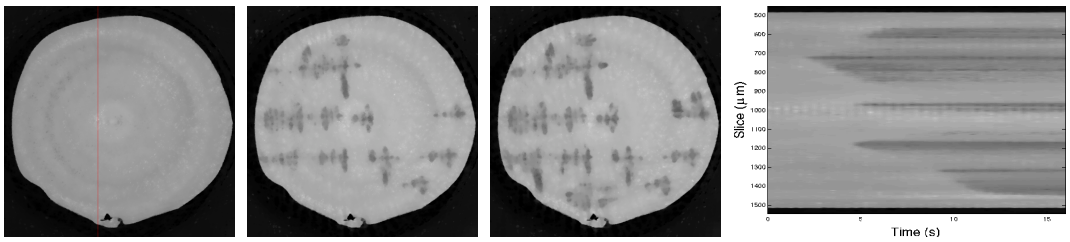


(e) time $= 0s$ (f) time $= 8s$ (g) time $= 16s$ (h) v-axis vs. time

Fig. 2.12. Reconstructions of dendritic growth using Nyquist progressive views. The first three columns show a $u-v$ slice of the sample at different times. The last column shows a v-axis slice of the sample (along the red line in (a, e)) as a function of time. The reconstruction ($r = 1$) from progressive views with $N_\theta = 1536$ (Nyquist criteria) using FBP is shown in (a - d) and using MBIR is shown in (e - h). The reconstruction shown in (a - h) has a low temporal reconstruction rate of $F_s = 0.125$ Hz which is insufficient to temporally resolve the growing dendrites.

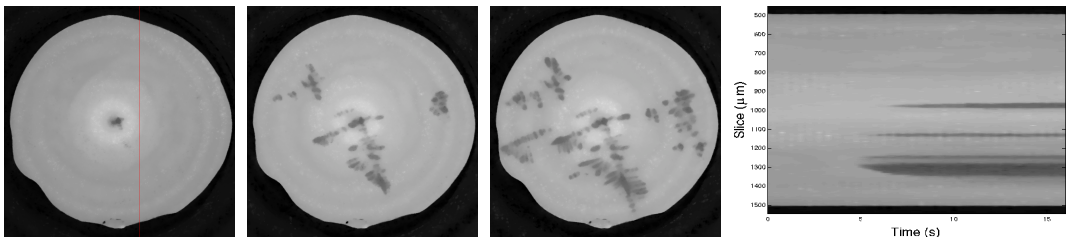
To demonstrate the performance gains achieved by TIMBIR in a real physical system, we reconstruct the dendritic growth in an Al-Cu alloy [19] in 4D. Dendrites are complex tree like structures which form as liquids are cooled from a sufficiently high temperature. It is of great interest to study dendritic growth since the morphology of the growing dendrites determine the properties of many materials. Thus, by studying the dendritic growth we can better understand the processes controlling the morphology of these tree like structures.

MBIR, Progressive Views ($r = 1, K = 1, N_\theta = 48$)



(a) time = $0s$ (b) time = $8s$ (c) time = $16s$ (d) v-axis vs. time

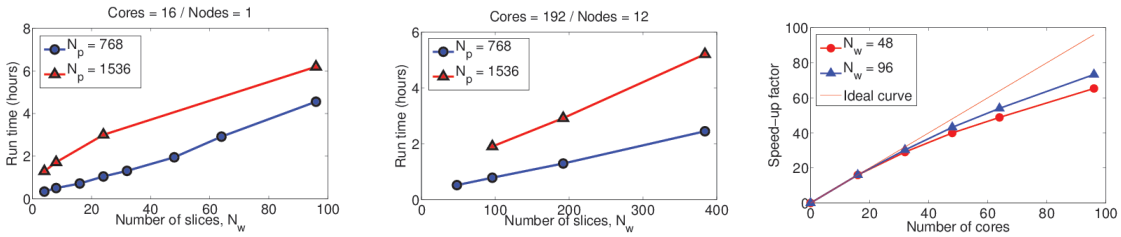
TIMBIR : MBIR, Interlaced Views ($r = 32, K = 32, N_\theta = 1536$)



(e) time = $0s$ (f) time = $8s$ (g) time = $16s$ (h) v-axis vs. time

Fig. 2.13. Comparison of reconstructions of dendritic growth using TIMBIR with other approaches to high temporal resolution reconstructions. The first three columns show a $u-v$ slice of the sample at different times. The last column shows a v-axis slice of the sample (along the red line in (a, e)) as a function of time. The MBIR reconstruction ($r = 1$) from progressive views with $N_\theta = 48$ (Sub-Nyquist) is shown in (a - d). The MBIR reconstruction ($r = 8$) from interlaced views with $K = 32, N_\theta = 1536$ is shown in (e - h) (TIMBIR). Even though the reconstruction in (a - d) has a rate of $F_s = 4$ Hz, the spatial reconstruction quality is very poor. However, the TIMBIR reconstruction in (e - h) not only has a rate of $F_s = 4$ Hz to temporally resolve the growing dendrites, it also has good spatial reconstruction quality. Note that results in (a-d) and (e-h) correspond to data from different experiments.

In our experiments, the data acquisition rate is limited to $F_c = 192$ Hz due to limitations on the camera frame rate, the data transfer rate, and buffer sizes. The detector width is $N_p = 1600$ pixels in the cross-axial direction and 1080 pixels along the axial direction with a pixel resolution of 0.65 μm $\times$ 0.65 μm. However, we only reconstruct a window of 4 pixels in the axial direction and $N_p = 1536$ pixels in the cross-axial direction. The reconstructions have a size of $N_w \times N_v \times N_u = 4 \times 1536 \times 1536$



(a) (b) (c)

Fig. 2.14. Run time analysis of the distributed parallel MBIR algorithm. (a) is a plot of the run time using 16 cores (or 1 node) vs. the total number of the slices, N_w, at different cross-axial detector resolutions, N_p. (b) is a plot of the run time using 192 cores (or 12 nodes) vs. the total number of slices, N_w. (c) is a plot of the run time speed-up factor vs. the number of cores. The speed-up factor with n cores is computed as the ratio of the average run time per iteration with 1 core to that with n cores. We can see that the algorithm speed up factor gradually deviates from the ideal behavior when the number of cores is increased. Furthermore, the speed up behavior improves when the number of slices, N_w, is increased.

and a voxel resolution of $0.65 \times 0.65 \times 0.65$ μm^3. The exposure time of the detector is set to 1 ms. The regularization parameters, σ_s and σ_t, are chosen to provide the best visual reconstruction quality. The parameter p of the qGGMRF model is set to 1.2 and the parameters of the generalized Huber function are set to be $\delta = 0.5$ and $T = 4$. We use a convergence threshold of $T = 0.05$. We increased the threshold from $T = 0.01$ used for simulated data to $T = 0.05$ for real data to reduce run time.

Imaging is done using polychromatic X-ray radiation from a synchrotron. The attenuated X-rays from the object are converted to visible light using a 25 μm thick LuAG:Ce scintillator which is then imaged by a PCO Edge CMOS camera. We also use a 10$\times$ magnifying objective to get an effective pixel resolution of 0.65 μm $\times$ 0.65 μm. The polychromatic radiation results in beam hardening which causes cupping artifacts in the reconstruction. To correct for these artifacts, we use a simple quadratic correction polynomial of the form $y = az^2 + z$ where z is the measured value of the projection and y is the projection after correction [67]. The corrected projections are then used to do the reconstruction. To find the optimal value of a,

we observed the reconstruction for cupping artifacts for increasing values of a. Based on this empirical analysis, we used a value of $a = 0.5$ which was found to minimize cupping artifacts.

First, we reconstruct the 4D object using a progressive view dataset satisfying the Nyquist criterion. To acquire N_θ progressive views over a rotation of π radians at a data acquisition rate of F_c, the object has to be rotated by an angle of π radians every $F_c^{-1}N_\theta$ seconds. Thus, the angular rotation speed of the object is fixed at $R_s = \pi F_c/N_\theta = 0.3925$ radians per second and data is acquired at a rate of $F_c =$ 192 Hz using progressive view sampling with parameter $N_\theta = 1536$. The object is then reconstructed at a rate of 0.125 Hz by reconstructing $r = 1$ time sample every frame. With this method, we get a temporal resolution of $F_s^{-1} = 8s$. The FBP and MBIR reconstructions of data acquired using this technique are shown in Fig. 2.12 (a - d) and Fig. 2.12 (e - h) respectively. From Fig. 2.12 (d, h), we can see that a temporal resolution of $8s$ is inadequate to temporally resolve the growing dendrites. Furthermore, it also causes blur artifacts in the reconstructions as seen in Fig. 2.12 (b, f). The strong ring artifacts in the FBP reconstructions results in additional distortion. However, by modeling the measurement non-idealities, MBIR is able to substantially reduce ring artifacts. Thus, using the conventional method we cannot reconstruct the dendritic growth with sufficient resolution in time and space.

To increase the temporal resolution with progressive views, we reduce the number of distinct angles to $N_\theta = 48$ (Sub-Nyquist) per π radians, increase the rotation speed to $R_s = 12.56$ radians per second, and acquire data at the maximum rate of $F_c =$ 192 Hz. The object is then reconstructed at a rate of $F_s = 4$ Hz by reconstructing $r = 1$ time sample every frame. The reconstructions using this method are shown in Fig. 2.13 (a - d). Note that the dendritic structure is different than in Fig. 2.12 (a - h) because it is not possible to replicate such physical phenomenon in an exact manner with each experimental run. With this method, we get a temporal resolution of $F_s^{-1} = 0.25s$. In this case, even though the temporal resolution is high (Fig. 2.13 (d)), the spatial reconstruction quality is very poor (Fig. 2.13 (a - c)).

Next, we study the effect of TIMBIR on the reconstruction quality when reconstructing at a higher temporal rate of $F_s = 4$ Hz. To acquire N_θ interlaced views over a rotation of $K\pi$ radians at a rate of F_c, the object has to be rotated by an angle of π radians every $F_c^{-1}N_\theta/K$ seconds. Thus, the angular rotation speed of the object is fixed at $R_s = \pi K F_c/N_\theta = 12.56$ radians per second and data is acquired at a rate of $F_c = 192$ Hz using interlaced view sampling with parameters $K = 32$ and $N_\theta = 1536$. The object is then reconstructed at a rate of $F_s = 4$ Hz by reconstructing $r = 32$ time samples every frame. The reconstructions using this method are shown in Fig. 2.13 (e - h). We can see that the reconstruction using TIMBIR has very good spatial (Fig. 2.13 (e - g)) and temporal resolution (Fig. 2.13 (h)). Thus, TIMBIR with its synergistic combination of interlaced view sampling and MBIR algorithm is able to reconstruct the dendritic evolution at a high spatial and temporal resolution. We have empirically found that using a value of K less than but closest to $\sqrt{N_\theta}$ gives us good results.

2.6.3 Computational Cost

We study the variation of run time of the MBIR algorithm with the number of cores when reconstructing one frame of projections with parameters $K = 32, N_\theta = 1536, r = 32$. The algorithm run time is determined for two different cross-axial detector resolutions of $N_p = 768$ and 1536 and for different number of axial slices, N_w. The cross-axial reconstruction resolution is $N_v \times N_u = N_p \times N_p$ and the number of slices in the reconstruction is N_w. To determine the run time, we run the algorithm on the Conte supercomputing cluster at Purdue University. Each node on Conte consists of two 8 core Intel Xeon-E5 processors. Fig. 2.14 (a, b) shows the run time as a function of the number of slices, N_w, for two different cross-axial detector resolutions of $N_p = 768$ and 1536. The run time using 16 cores (or 1 node) is shown in Fig. 2.14 (a) and using 192 cores (or 12 nodes) is shown in Fig. 2.14 (b). Thus,

we can see that increasing the number of cores from 16 to 192 significantly reduces the run time of the MBIR algorithm.

In Fig. 2.14 (c), we plot the speed-up factor as a function of the number of cores for different number of slices when $N_p = 768$. The speed-up factor with n cores is computed as the ratio of the average run time per iteration with 1 core to that with n cores at the finest multi-resolution stage. From the figure, we can see that initially the speed-up improvement is almost linear with the number of cores and gradually deviates from the ideal behavior when the number of cores is increased. Furthermore, when the number of slices, N_w, is increased, the speed-up behavior improves since the computation can be more efficiently distributed among the different nodes. Thus, using the distributed MBIR algorithm we can efficiently reconstruct large volumes using multiple cores. In the future, we believe the run time can be significantly reduced using better parallel architectures such as in [68].

2.7 Conclusion

In this chapter, we propose a novel interlaced view sampling method which when combined with our 4D MBIR algorithm is able to achieve a synergistic improvement in reconstruction quality of 4D-SXCT. In addition to accounting for spatial and temporal correlations in the object, the MBIR algorithm also accounts for the measurement non-idealities encountered in a SXCT system. Using the new interlaced view sampling strategy with the 4D MBIR algorithm (TIMBIR), we were able to achieve a 32× improvement in temporal resolution by reconstructing 32× more time samples while preserving spatial reconstruction quality. We also present a distributed parallel MBIR algorithm to enable reconstructions of large datasets. In the future, by using better spatial and temporal prior models we expect to achieve much better reconstruction quality than presented in this paper. Furthermore, the TIMBIR method can be extended to any tomographic technique involving sampling of projections at different angles.

3. MODEL-BASED ITERATIVE RECONSTRUCTION OF MAGNETIZATION USING VECTOR FIELD ELECTRON TOMOGRAPHY

3.1 Introduction

Imaging of magnetic materials at nanometer resolutions has wide-spread applications in the field of nanotechnology [69]. Magnetic particles with nanometer dimensions have unique quantum mechanical properties that are enabling new technologies in the fields of healthcare, life sciences, and material science [69]. The magnetic property of a magnetic material is characterized its magnetization, a vector field that expresses the density of magnetic dipole moment within the material [70]. The magnetization results in a vector field called the magnetic vector potential that extends outside the sample [70]. The magnetic field, also called magnetic induction, is then given by the curl of the magnetic vector potential [70]. Advances in nanotechnology rely on the ability to observe the various 3D vector fields of magnetization, magnetic vector potential, and magnetic field.

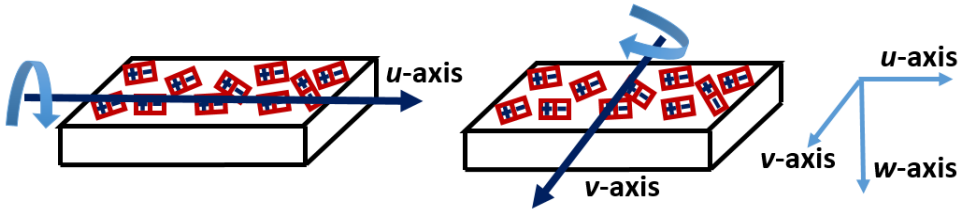


Fig. 3.1. Illustration of data acquisition in vector field electron tomography (VFET). In VFET, the sample is mounted on a rotary stage and exposed to electron radiation. The sample is then tilted across two orthogonal tilt axes ($u-$axis and $v-$axis) and measurements are made at multiple tilt angles.

Vector field tomography (VFT) has been widely used to reconstruct vector fields in 2D [71–75] and 3D [15,76] in a wide range of applications. Since our objective is to reconstruct the vector fields associated with a magnetic sample, we will use vector field electron tomography (VFET) [15] implemented using a conventional transmission electron microscope (TEM). In a typical experiment, electrons are focused on the sample and the transmitted electrons after passing through the sample are measured by a planar detector. In VFET, the sample is tilted across two orthogonal rotation axes and measurements are made at multiple tilt angles using a TEM as shown in Fig. 3.1. The information about the vector fields associated with the magnetic sample is contained in the phase shift of the electrons exiting the sample. At each tilt angle, the electron phase shift is retrieved from detector measurements before performing a tomographic reconstruction [77]. This is achieved using either electron holography methods [78, 79] or through-focal series measurements in Lorentz TEM mode [77]. If the measurement is made in the presence of electrostatic fields, then the retrieved phase contains contributions from the electrostatic field in addition to the magnetic field [80]. In this case, the magnetic component of the phase can be extracted using the method presented in [80].

There exist a wide variety of analytical inversion methods that reconstruct the magnetic vector potential or the magnetic field [15, 77, 78] from the electron phase shift data. These methods rely on approximately inverting the linear relation that expresses the electron phase shift as a projection of the magnetic vector potential component along the direction of electron propagation [15]. The conventional approach to reconstruct the magnetic vector potential from the phase shift data relies on an analytical back-projection method that also assumes that the divergence of the magnetic vector potential is zero [15, 77]. The magnetic field can then be computed as the curl of the magnetic vector potential. There also exist analytical methods to reconstruct the magnetic field directly from gradients of electron phase shift images [15, 78] while assuming zero divergence for the magnetic field. These analytical methods cause noisy artifacts due to the so called singular surfaces [16] that occur

in the inverse formulas used for reconstruction. The reconstruction can be fixed by acquiring a third set of tilt series where the rotation axis is perpendicular to the axes of the first two tilt series [16]. However, this technique is difficult to implement in practice. Note that none of these methods reconstruct the magnetization which is a fundamental material property that does not extend outside the sample. Typically, the magnetization may form complex textures that can be difficult to visualize by just observing the magnetic vector potential or the magnetic field.

There also exist a variety of iterative least-squares and regularized inversion methods for reconstruction of vector fields in applications such as fluid flow fields [75] and optical flow [81,82]. These methods fall under the general framework of model-based iterative reconstruction (MBIR). In MBIR, the reconstruction is given by the solution to the minimization of a cost function consisting of a forward model term and a prior model term. The forward model uses the physics of imaging to expresses the measured data as a function of the unknown vector field. The prior model regularizes the reconstruction of the vector field using different models of sparsity. Algebraic reconstruction methods without a prior model that rely on finding a least squares solution to the vector field from measurements of projections of the vector field is given in [75, 83]. Prior models using penalty functions such as the L2-norm or L1-norm for regularizing the magnitude of gradient of vector field components is presented in [81, 82]. There also exist algorithms that use angular regularization for vector fields [84,85]. More advanced algorithms that rely on regularizing the divergence and curl of the vector fields are presented in [86, 87]. In [87], the authors discuss the geometric invariance properties of various regularization functions for vector fields.

There is a lack of MBIR algorithms for reconstruction of magnetic vector fields from VFET data. An algebraic reconstruction technique for unregularized reconstruction of magnetic field from gradient of electron phase data is presented in [83]. The regularized reconstruction algorithms in [81, 82] can also be extended using the MBIR framework to reconstruct the magnetic vector potential or the magnetic field. An MBIR algorithm for imaging magnetic materials can dramatically improve re-

construction quality, increase noise robustness, and reduce reconstruction artifacts. MBIR algorithms have resulted in significant gains in a wide range of scalar imaging applications such as X-ray computed tomography [1,3,10,11,56,88], bright field electron tomography [13], and HAADF-STEM tomography [12]. However, none of these algorithms can be extended to reconstruct magnetization since the electron phase cannot be represented as a projection of the magnetization. Rather, the electron phase has a global and more complex dependence on the magnetization [89].

In this thesis, we will use the framework of model-based iterative reconstruction (MBIR) to formulate an algorithm that will reconstruct the magnetization directly from the electron phase data [18]. In the forward model, we rely on physics to express the measured electron phase as a function of the magnetization. We do this by first expressing the magnetic vector potential as a convolution of the magnetization with a vector form of the Green's function [70]. Then, the electron phase at each view is given by the projection of the magnetic vector potential component along the direction of electron propagation [70]. In the prior model, we use a Gaussian Markov Random Field (GMRF) [61] potential function to regularize the magnitude of gradient of the magnetization. Since the resulting cost function is difficult to minimize directly, we will use variable splitting and the theory of alternate direction method of multipliers (ADMM) to solve the original minimization problem as an iterative solution to two simpler minimization problems. We show that the simpler minimization problems can be solved using standard optimization algorithms such as iterative coordinate descent (ICD) [3] and gradient descent techniques [62].

Our algorithm enables the reconstruction of the magnetic vector potential in addition to the magnetization. We show that our algorithm does not result in noisy artifacts that is typically seen in reconstructions of magnetic vector potential using the conventional method [16]. Furthermore, our algorithm is the first algorithm to accurately reconstruct the 3D distribution of the magnetization vector field directly from the phase shift data. Note that our algorithm also does not enforce the zero divergence or the zero curl condition used by conventional algorithms [15, 16]. We

validate our algorithm by presenting reconstructions of both simulated and real experimental data.

3.2 Measurement Model

To formulate the reconstruction algorithm, we have to first express the phase shift of the electron wave propagating through the object as a function of the magnetization. We do this by first expressing the magnetic vector potential as a function of the magnetization followed by expressing the electron phase shift as a function of the magnetic vector potential.

3.2.1 Relation between magnetic vector potential and magnetization

Let $r = (u, v, w)$ and $r' = (u', v', w')$ be position vectors in 3D space spanned by the mutually orthogonal $u-$axis, $v-$axis and $w-$axis. Then, the magnetic vector potential, $A(r)$, is given by the convolution cross-product of the magnetization, $M(r)$, with a vector form of Green's function, $h_C(r) = r/|r|^3$, as shown below [70, 90],

$$A(r) = \frac{\mu_0}{4\pi} \int_{\mathcal{R}^3} M(r') \times h_C(r - r') dr', \tag{3.1}$$

where $\times$ denotes vector cross-product and μ_0 is the permeability of vacuum. The magnetic field, $B(r)$, is then given by the curl of the magnetic vector potential i.e., $B(r) = \nabla \times A(r)$.

We will now express the individual vector components of $A(r)$ in terms of the vector components of $M(r)$ by expanding the vector cross-product in the convolution relation shown in (3.1). Let $M(r) = \left(M^{(u)}(r), M^{(v)}(r), M^{(w)}(r)\right)$ represent the three orthogonal vector components of magnetization, $M(r)$, along each of the orthogonal

(u, v, w) coordinate axes. Similarly, the Green's function in (3.1) can be expressed in terms of its components as $h_C(r) = \left(h_C^{(u)}(r), h_C^{(v)}(r), h_C^{(w)}(r)\right)$ where

$$h_C^{(u)}(u,v,w) = \frac{u}{\left|u^2+v^2+w^2\right|^{3/2}} \tag{3.2}$$

$$h_C^{(v)}(u,v,w) = \frac{v}{\left|u^2+v^2+w^2\right|^{3/2}} \tag{3.3}$$

$$h_C^{(w)}(u,v,w) = \frac{w}{\left|u^2+v^2+w^2\right|^{3/2}} \tag{3.4}$$

Then, the corresponding orthogonal components of the magnetic vector potential denoted by $A^{(u)}(r)$, $A^{(v)}(r)$, and $A^{(w)}(r)$ can be expressed as,

$$A^{(u)}(r) = \frac{\mu_0}{4\pi}\int_{\mathcal{R}^3}\left[M^{(v)}(r')h_C^{(w)}(r-r') - M^{(w)}(r')h_C^{(v)}(r-r')\right]dr' \tag{3.5}$$

$$A^{(v)}(r) = \frac{\mu_0}{4\pi}\int_{\mathcal{R}^3}\left[M^{(w)}(r')h_C^{(u)}(r-r') - M^{(u)}(r')h_C^{(w)}(r-r')\right]dr' \tag{3.6}$$

$$A^{(w)}(r) = \frac{\mu_0}{4\pi}\int_{\mathcal{R}^3}\left[M^{(u)}(r')h_C^{(v)}(r-r') - M^{(v)}(r')h_C^{(u)}(r-r')\right]dr'. \tag{3.7}$$

Next, we will express the above convolution relations in a discrete form that is suitable for implementation on a computer. A detailed derivation of the discrete approximations to the above continuous space relations are provided in appendix B. Let $x^{(u)}$, $x^{(v)}$, and $x^{(w)}$ be vector arrays containing all voxel values of $M^{(u)}(r)$, $M^{(v)}(r)$, and $M^{(w)}(r)$ respectively in raster order. We can then express the vectors arrays $z^{(u)}, z^{(v)}$, and $z^{(w)}$ containing the voxel values of $A^{(u)}(r)$, $A^{(v)}(r)$, and $A^{(w)}(r)$ respectively as,

$$z^{(u)} = H^{(w)}x^{(v)} - H^{(v)}x^{(w)} \tag{3.8}$$

$$z^{(v)} = H^{(u)}x^{(w)} - H^{(w)}x^{(u)} \tag{3.9}$$

$$z^{(w)} = H^{(v)}x^{(u)} - H^{(u)}x^{(v)} \tag{3.10}$$

where $H^{(u)}$, $H^{(v)}$, and $H^{(w)}$ are matrices that implement 3D convolution with point spread functions given by,

$$h_D^{(u)}[i,j,k] = w[i,j,k]\frac{i\Delta}{|i^2+j^2+k^2|^{3/2}}, \tag{3.11}$$

$$h_D^{(v)}[i,j,k] = w[i,j,k]\frac{j\Delta}{|i^2+j^2+k^2|^{3/2}}, \tag{3.12}$$

$$h_D^{(w)}[i,j,k] = w[i,j,k]\frac{k\Delta}{|i^2+j^2+k^2|^{3/2}} \tag{3.13}$$

respectively, where $[i,j,k]$ are discrete coordinates, Δ is the voxel width, $h_D^{(u)}[0,0,0] = h_D^{(v)}[0,0,0] = h_D^{(w)}[0,0,0] = 0$, and $w[i,j,k]$ is a 3D Hamming window (appendix B).

Thus, we can express $z = \left[z^{(u)t}, z^{(v)t}, z^{(w)t}\right]^t$ in terms of $x = \left[x^{(u)t}, x^{(v)t}, x^{(w)t}\right]^t$ as,

$$z = Hx \text{ where } H = \begin{bmatrix} 0 & H^{(w)} & -H^{(v)} \\ -H^{(w)} & 0 & H^{(u)} \\ H^{(v)} & -H^{(u)} & 0 \end{bmatrix}. \tag{3.14}$$

3.2.2 Relation between electron phase shift and magnetic vector potential

In VFET, the sample is first tilted across the $u-$axis and measurements are made at several tilt angles. The procedure is then repeated by tilting the sample across the $v-$axis and making measurements at multiple tilt angles. At each tilt angle, the phase shift of the electrons exiting the sample is recovered from measurements. The phase shift of the electrons propagating along the positive $w-$axis is given by [15,91],

$$\phi(r_\perp) = \frac{2\pi e}{h}\int A(r_\perp + l\hat{w})\cdot\hat{w}dl \tag{3.15}$$

where $r_\perp$ is a 2D position vector on the projection plane, $\hat{w}$ is a unit vector directed along the positive w-axis, h is Planck's constant, and e is electron charge.

Let $y_i^{(u)}$ be a vector array containing all the pixel values of the electron phase image at the i^{th} tilt angle for tilt across the u-axis. In this case, the dot product in (3.15) ensures that only the $z^{(v)}$ and $z^{(w)}$ components of the magnetic vector potential will have a influence on the phase shift. Note that the direction of the vector component $z^{(u)}$ will always be perpendicular to the propagation direction (negative w-axis). If

$P_i^{(u)}$ denote the projection matrix that implements the line integral in (3.15) at the i^{th} tilt angle for tilt across the u-axis, we can show that,

$$y_i^{(u)} = -P_i^{(u)} z^{(v)} \sin\left(\theta_i^{(u)}\right) + P_i^{(u)} z^{(w)} \cos\left(\theta_i^{(u)}\right). \tag{3.16}$$

Similarly, only the $z^{(u)}$ and $z^{(w)}$ components will influence the phase shift for tilt across the v-axis. In this case, the vector component $z^{(v)}$ will always be perpendicular to the propagation direction. If $P_j^{(v)}$ denote the projection matrix that implements the line integral in (3.15) at the j^{th} tilt angle for tilt across the v-axis, we can show that,

$$y_j^{(v)} = -P_j^{(v)} z^{(u)} \sin\left(\theta_j^{(v)}\right) + P_j^{(v)} z^{(w)} \cos\left(\theta_j^{(v)}\right). \tag{3.17}$$

Note that our framework allows for different number of tilt angles across the u-axis and v-axis.

Then, we can express the relations in (3.16) and (3.17) in the form of matrix-vector products as,

$$y_i^{(u)} = F_i^{(u)} \begin{bmatrix} z^{(u)} \\ z^{(v)} \\ z^{(w)} \end{bmatrix} \text{ and } y_j^{(v)} = F_j^{(v)} \begin{bmatrix} z^{(u)} \\ z^{(v)} \\ z^{(w)} \end{bmatrix} \tag{3.18}$$

where

$$F_i^{(u)} = \left[0, \quad -P_i^{(u)} \sin\left(\theta_i^{(u)}\right), \quad P_i^{(u)} \cos\left(\theta_i^{(u)}\right)\right] \tag{3.19}$$

$$F_j^{(v)} = \left[-P_j^{(v)} \sin\left(\theta_j^{(v)}\right), \quad 0, \quad P_j^{(v)} \cos\left(\theta_j^{(v)}\right)\right]. \tag{3.20}$$

Let y be a vector array containing every element from the vectors $y_i^{(u)}$ and $y_j^{(v)}$ for all the tilt angles indexed by i and j. We can then express y in terms of z as,

$$y = Fz + w \tag{3.21}$$

where w is the noise vector and F is a matrix that implements the linear relation in (3.18) for all indices i and j by appropriately stacking the matrices $F_i^{(u)}$ and $F_j^{(v)}$.

By substituting (3.14) in (3.21), we get the forward model given by,

$$y = FHx + w. \tag{3.22}$$

3.3 Model-Based Iterative Reconstruction (MBIR) of Magnetization

In MBIR, the reconstruction is given by the solution to the following optimization problem,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ -\log p(y|x) - \log p(x) \right\} \tag{3.23}$$

where $\log p(y|x)$ is the forward log-likelihood function for the data, y, given the object x and $\log p(x)$ is the prior log-likelihood function for the object, x.

The forward log-likelihood function under the Gaussian noise assumption in (3.22) is given by,

$$-\log p(y|x) = \frac{1}{2\sigma^2} \left|\left| y - FHx \right|\right|^2 + \text{ constant} \tag{3.24}$$

where σ^2 is the variance of noise. The prior model is given by,

$$-\log p(x) = \sum_{\{k,l\} \in \mathcal{N}} \frac{w_{kl}}{2\sigma_x^2} \left[\left(x_k^{(u)} - x_l^{(u)} \right)^2 + \left(x_k^{(v)} - x_l^{(v)} \right)^2 + \left(x_k^{(w)} - x_l^{(w)} \right)^2 \right] + \text{constant}$$

where σ_x is the regularization parameter and $\mathcal{N}$ is the set of all pairwise cliques in 3D space (set of all pairwise indices of neighboring voxels). The weight parameter w_{kl} is set such that it is inversely proportional to the spatial distance between voxels indexed by k and l and $\sum_{l \in \mathcal{N}_k} w_{kl} = 1$, where $\mathcal{N}_k$ is the set of all indices of neighbors of voxel x_k. We can then express the prior model as,

$$-\log p(x) = \frac{1}{2} \left[x^{(u)t} \tilde{B} x^{(u)} + x^{(v)t} \tilde{B} x^{(v)} + x^{(w)t} \tilde{B} x^{(w)} \right] + \text{constant} \tag{3.25}$$

where $\tilde{B}$ is a matrix such that

$$\tilde{B}_{k,l} = \begin{cases} 1/\sigma_x^2 & \text{if} \quad k = l \\ -w_{kl}/\sigma_x^2 & \text{if} \quad l \in \mathcal{N}_k \end{cases}. \tag{3.26}$$

We can then express the prior model in terms of $x = \left[x^{(u)t}, x^{(v)t}, x^{(w)t} \right]^t$ as,

$$-\log p(x) = \frac{1}{2} x^t B x + \text{constant} \tag{3.27}$$

where

$$B = \begin{bmatrix} \tilde{B} & 0 & 0 \\ 0 & \tilde{B} & 0 \\ 0 & 0 & \tilde{B} \end{bmatrix}. \tag{3.28}$$

The reconstruction is then obtained by solving the following optimization problem,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} ||y - FHx||^2 + \frac{1}{2} x^t B x \right\}. \tag{3.29}$$

In (3.29), the matrix FH is dense since the convolution matrix H is dense even though the projection matrix F is sparse. Since (3.29) is difficult to solve directly, we will use variable splitting to express it as a constrained optimization problem of the form shown below,

$$(\hat{x}, \hat{z}) = \underset{x,z}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} ||y - Fz||^2 + \frac{1}{2} x^t B x \right\} \ s.t. \ z = Hx \tag{3.30}$$

The augmented Lagrangian function for this constrained optimization problem is,

$$L(x, z; t) = \frac{1}{2\sigma^2} ||y - Fz||^2 + \frac{\mu}{2} ||Hx - z + t||^2 + \frac{1}{2} x^t B x \tag{3.31}$$

where z is the auxiliary vector, t is the scaled dual vector, and $\mu > 0$ is the augmented Lagrangian parameter.

3.4 Optimization Algorithm

To solve the augmented Lagrangian formulation of the optimization problem, we use the theory of alternate direction method of multipliers (ADMM). Using the ADMM method [17], we can show that the optimization algorithm that solves (3.30) is given by algorithm 1. Thus, the original minimization problem in (3.29) can be solved by iteratively solving simpler minimization problems shown in (3.32) and (3.33). Our modular framework allows us to develop independent software modules to solve (3.32) and (3.33). Both (3.32) and (3.33) can be solved efficiently using a variety of well known optimization algorithms. We solve (3.32) using the steepest gradient descent algorithm [62] and (3.33) using a variant of the iterative coordinate descent algorithm [3].

In algorithm 1, we use an adaptive update strategy [17, 92, 93] for the augmented Lagrangian parameter μ. We update μ based on the primal and dual residuals at each ADMM iteration. The primal residual is defined as $r = \hat{z} - H\hat{x}$ and the dual

Algorithm 1 RECONSTRUCTION

while not converged **do**

$z^{(old)} \leftarrow \hat{z}$

DECONVOLUTION -

$$\hat{x} \leftarrow \underset{x}{\operatorname{argmin}} \left\{ \frac{\mu}{2} ||Hx - \hat{z} + t||^2 + \frac{1}{2} x^t B x \right\} \tag{3.32}$$

TOMOGRAPHIC INVERSION -

$$\hat{z} \leftarrow \underset{z}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} ||y - Fz||^2 + \frac{\mu}{2} ||H\hat{x} - z + t||^2 \right\} \tag{3.33}$$

$t \leftarrow t + (H\hat{x} - \hat{z})$

$r \leftarrow (\hat{z} - H\hat{x})$ ▷ compute primal residual

$s \leftarrow \left(\hat{z} - z^{(old)}\right)$ ▷ compute dual residual

if $||s||_2 > \gamma ||r||_2$ **then**

$\mu \leftarrow \mu / \tau$

$t \leftarrow \tau t$

end if

if $||r||_2 > \gamma ||s||_2$ **then**

$\mu \leftarrow \tau \mu$

$t \leftarrow t / \tau$

end if

end while

residual is defined as $s = \hat{z} - z^{(old)}$, where $\hat{x}$ and $\hat{z}$ are the estimates after performing the updates in (3.32) and (3.33) and $z^{(old)}$ is the estimate before performing the update. If the primal residual is greater than γ times the dual residual, we increase the parameter μ by a factor of τ, where $\gamma > 1$ and $\tau > 1$. Similarly, if the dual residual is greater than γ times the primal residual, we decrease the parameter μ by a factor of τ. The idea here is to keep the norms of the primal and dual residuals within a factor of μ from one and another. As the iterations progress, both the primal

and dual residuals converge to zero. Note that the scaled dual vector t must also be updated appropriately after updating μ [17].

3.4.1 Tomographic Inversion

The solution to the minimization problem in (3.33) is the value of z that minimizes the cost function given by,

$$c_z(z) = \frac{1}{2\sigma^2} ||y - Fz||^2 + \frac{\mu}{2} ||Hx - z + t||^2 . \tag{3.34}$$

To minimize (3.34), we will use a variant of the iterative coordinate descent (ICD) algorithm [3]. In this algorithm, we sequentially minimize (3.34) with respect to the magnetic vector potential value $z_j = \left[z_j^{(u)}, z_j^{(v)}, z_j^{(w)}\right]^t$ at each voxel location j while keeping the voxel values at other locations fixed. We repeat this minimization procedure for each voxel chosen in a random order until the algorithm converges.

To minimize (3.34) with respect to z_j while keeping other voxel values constant, we will reformulate (3.34) in terms of just z_j while ignoring all terms that does not depend on z_j. Let z' denote the current estimate for z before performing the update. The cost function with respect to z_j is given by,

$$c_z^{vox}(z_j) = \tilde{\theta}^t z_j + \frac{1}{2} \left(z_j - z'_j\right)^t \tilde{\Theta} \left(z_j - z'_j\right) + \frac{\mu}{2} ||\tilde{z}_j - z_j + t_j||^2 \tag{3.35}$$

where $\tilde{\theta}$ and $\tilde{\Theta}$ are the respective gradient and Hessian of $\frac{1}{2\sigma^2} ||y - Fz||^2$ with respect to z_j and

$$\tilde{z}_j = \begin{bmatrix} H_{j,*}^{(w)} x^{(v)} - H_{j,*}^{(v)} x^{(w)} \\ H_{j,*}^{(u)} x^{(w)} - H_{j,*}^{(w)} x^{(u)} \\ H_{j,*}^{(v)} x^{(u)} - H_{j,*}^{(u)} x^{(v)} \end{bmatrix} \tag{3.36}$$

where $H_{j,*}^{(u)}$, $H_{j,*}^{(v)}$, and $H_{j,*}^{(w)}$ denote the elements along the j^{th} row of the matrices $H^{(u)}$, $H^{(v)}$, and $H^{(w)}$ respectively. Note that minimization of the cost function (3.35) is equivalent to minimizing (3.34) with respect to z_j while assuming constant values for other voxels.

Next, we will derive the gradient $\tilde{\theta}$ and Hessian $\tilde{\Theta}$ parameters in (3.35). Let the error sinogram vectors be defined as,

$$e_i^{(u)} = y_i^{(u)} - P_i^{(u)} \left(-z^{(v)} \sin\left(\theta_i^{(u)}\right) + z^{(w)} \cos\left(\theta_i^{(u)}\right) \right)$$
$$e_i^{(v)} = y_i^{(v)} - P_i^{(v)} \left(-z^{(u)} \sin\left(\theta_i^{(v)}\right) + z^{(w)} \cos\left(\theta_i^{(v)}\right) \right).$$

If $\tilde{\theta}_i$ is the i^{th} element of the 3×1 vector $\tilde{\theta}$, we can show that,

$$\tilde{\theta}_1 = \frac{1}{\sigma^2} \sum_{i=1}^{M_v} e_i^{(v)t} P_{i,*,j}^{(v)} \sin\left(\theta_i^{(v)}\right), \tag{3.37}$$

$$\tilde{\theta}_2 = \frac{1}{\sigma^2} \sum_{i=1}^{M_u} e_i^{(u)t} P_{i,*,j}^{(u)} \sin\left(\theta_i^{(u)}\right), \tag{3.38}$$

$$\tilde{\theta}_3 = -\frac{1}{\sigma^2} \sum_{i=1}^{M_v} e_i^{(v)t} P_{i,*,j}^{(v)} \cos\left(\theta_i^{(v)}\right) - \frac{1}{\sigma^2} \sum_{i=1}^{M_u} e_i^{(u)t} P_{i,*,j}^{(u)} \cos\left(\theta_i^{(u)}\right) \tag{3.39}$$

where M_u and M_v are the total number of tilt angles across the $u-$axis and $v-$axis respectively and $P_{i,*,j}^{(u)}$ and $P_{i,*,j}^{(v)}$ represent the elements of the j^{th} column of the projection matrices $P_i^{(u)}$ and $P_i^{(v)}$ respectively (defined in (3.16) and (3.17)). Similarly, if $\tilde{\Theta}_{i,j}$ denotes the element at the i^{th} row and j^{th} column of the 3×3 Hessian matrix $\tilde{\Theta}$, then,

$$\Theta_{1,1} = \frac{1}{\sigma^2} \sum_{i=1}^{M_v} \left|\left| P_{i,*,j}^{(v)} \right|\right|^2 \sin^2\left(\theta_i^{(v)}\right), \tag{3.40}$$

$$\Theta_{1,3} = \Theta_{3,1} = -\frac{1}{\sigma^2} \sum_{i=1}^{M_v} \left|\left| P_{i,*,j}^{(v)} \right|\right|^2 \sin\left(\theta_i^{(v)}\right) \cos\left(\theta_i^{(v)}\right), \tag{3.41}$$

$$\Theta_{2,2} = \frac{1}{\sigma^2} \sum_{i=1}^{M_u} \left|\left| P_{i,*,j}^{(u)} \right|\right|^2 \sin^2\left(\theta_i^{(u)}\right), \tag{3.42}$$

$$\Theta_{2,3} = \Theta_{3,2} = -\frac{1}{\sigma^2} \sum_{i=1}^{M_u} \left|\left| P_{i,*,j}^{(u)} \right|\right|^2 \sin\left(\theta_i^{(u)}\right) \cos\left(\theta_i^{(u)}\right), \tag{3.43}$$

$$\Theta_{3,3} = \frac{1}{\sigma^2} \sum_{i=1}^{M_v} \left|\left| P_{i,*,j}^{(v)} \right|\right|^2 \cos^2\left(\theta_i^{(v)}\right) + \frac{1}{\sigma^2} \sum_{i=1}^{M_u} \left|\left| P_{i,*,j}^{(u)} \right|\right|^2 \cos^2\left(\theta_i^{(u)}\right), \tag{3.44}$$

$$\Theta_{1,2} = \Theta_{2,1} = 0. \tag{3.45}$$

Then, the value of z_j that minimizes (3.35) is given by,

$$\hat{z}_j = \left(\tilde{\Theta} + \mu I\right)^{-1} \left(-\tilde{\theta} + \tilde{\Theta} z_j' + \mu \tilde{z}_j + \mu t_j \right) \tag{3.46}$$

Thus, we can minimize the cost function in (3.34) using algorithm 2.

Algorithm 2 TOMOGRAPHIC INVERSION

while not converged **do**
  for all voxel indices j **do** $//z'_j$ is current estimate of z_j
    $z'_j \leftarrow \hat{z}_j$
    Compute $\tilde{\theta}$ by substituting z'_j for z_j in (3.37), (3.38), and (3.39).
    Compute $\tilde{\Theta}$ using (3.40), (3.41), (3.42), (3.43), (3.44), and (3.45).
    Compute $\tilde{z}_j$ by substituting $\hat{x}$ for x in (3.36).
    $\hat{z}_j \leftarrow \left(\tilde{\Theta} + \mu I\right)^{-1} \left(-\tilde{\theta} + \tilde{\Theta} z'_j + \mu \tilde{z}_j + \mu t_j\right)$
    $e_i^{(u)} \leftarrow e_i^{(u)} + P_{i,*,j}^{(u)} \sin\left(\theta_i^{(u)}\right) \left(\hat{z}_j^{(v)} - z_j^{(v)\prime}\right) - P_{i,*,j}^{(u)} \cos\left(\theta_i^{(u)}\right) \left(\hat{z}_j^{(w)} - z_j^{(w)\prime}\right)$
    $e_i^{(v)} \leftarrow e_i^{(v)} + P_{i,*,j}^{(v)} \sin\left(\theta_i^{(v)}\right) \left(\hat{z}_j^{(u)} - z_j^{(u)\prime}\right) - P_{i,*,j}^{(v)} \cos\left(\theta_i^{(v)}\right) \left(\hat{z}_j^{(w)} - z_j^{(w)\prime}\right)$
  end for
end while

3.4.2 Deconvolution

The solution to the minimization problem in (3.32) is given by the value of x that minimizes the cost function shown below,

$$f(x) = \frac{\mu}{2} ||Hx - z + t||^2 + \frac{1}{2} x^t B x. \tag{3.47}$$

We will use steepest gradient descent algorithm [62] to iteratively minimize (3.47) with respect to the magnetization, x, until the algorithm converges. Gradient descent is based on the idea that the cost function (3.47) reduces in the direction of its negative gradient. By choosing an optimal step size that minimizes the cost in that direction, we will get closer to the global minimum of (3.47). Since the cost function in (3.47) is convex, our algorithm converges to the global minimum [62].

To formulate the optimization algorithm, we will derive expressions for the gradient, g, and Hessian, Q, of the cost function $f(x)$. In every iteration, we will update x in the direction of the negative gradient, $-g$. Since $f(x)$ is quadratic in x, there exists

a closed form expression for the stepsize, α, that minimizes $f(x)$ in the direction of $-g$. The gradient of $f(x)$ is given by,

$$g = \nabla f(x) = \mu H^t \left(Hx - z + t\right) + Bx \tag{3.48}$$

and the Hessian of $f(x)$ is given by,

$$Q = \mu H^t H + B. \tag{3.49}$$

Since (3.47) is quadratic, the optimal stepsize α can be computed as [62],

$$\alpha = \frac{g^t g}{g^t Q g} \tag{3.50}$$

Thus, the optimization algorithm that minimizes (3.49) is given in algorithm 3.

Algorithm 3 DECONVOLUTION

1: Compute Q from (3.49).
2: **while** not converged **do**
3: $x' \leftarrow \hat{x}$
4: Compute $g \leftarrow \mu H^t \left(Hx' - \hat{z} + t\right) + Bx'$. $//x'$ is current estimate for x
5: Compute $\alpha \leftarrow \frac{g^t g}{g^t Q g}$
6: $\hat{x} \leftarrow x' - \alpha g$
7: **end while**

Note that directly computing the expressions in (3.48) and (3.50) using matrix-vector multiplication is computationally intensive. Instead, by formulating these expressions in terms of $H^{(u)}, H^{(v)}$, and $H^{(w)}$ that represent linear space invariant filtering matrices, we can efficiently solve (3.48) and (3.50) [61]. The gradient vector $g = \left[g^{(u)t}, g^{(v)t}, g^{(w)t}\right]$ can be reformulated in terms of $H^{(u)}, H^{(v)}$, and $H^{(w)}$ as,

$$g^{(u)} = -\mu \left[H^{(w)t} f^{(v)} - H^{(v)t} f^{(w)}\right] + B^{(u)} x^{(u)}, \tag{3.51}$$

$$g^{(v)} = -\mu \left[H^{(u)t} f^{(w)} - H^{(w)t} f^{(u)}\right] + B^{(v)} x^{(v)}, \tag{3.52}$$

$$g^{(w)} = -\mu \left[H^{(v)t} f^{(u)} - H^{(u)t} f^{(v)}\right] + B^{(w)} x^{(w)}. \tag{3.53}$$

where

$$f^{(u)} = \left(H^{(w)}x^{(v)} - H^{(v)}x^{(w)} - z^{(u)} + t^{(u)}\right), \tag{3.54}$$

$$f^{(v)} = \left(H^{(u)}x^{(w)} - H^{(w)}x^{(u)} - z^{(v)} + t^{(v)}\right), \tag{3.55}$$

$$f^{(w)} = \left(H^{(v)}x^{(u)} - H^{(u)}x^{(v)} - z^{(w)} + t^{(w)}\right). \tag{3.56}$$

Thus, the gradient can be computed efficiently since matrix multiplication with $H^{(u)}, H^{(v)}$, and $H^{(w)}$ is equivalent to linear space invariant convolution with the point spread functions in (3.11), (3.12), and (3.13) respectively. And matrix multiplication with $H^{(u)t}, H^{(v)t}$, and $H^{(w)t}$ is equivalent to linear space invariant convolution with the space reversed version of the point spread functions in (3.11), (3.12), and (3.13) respectively [61].

Similarly, the Hessian can be expressed in block-matrix representation as,

$$Q = \begin{bmatrix} Q^{(uu)} & Q^{(uv)} & Q^{(uw)} \\ Q^{(vu)} & Q^{(vv)} & Q^{(vw)} \\ Q^{(wu)} & Q^{(wv)} & Q^{(ww)} \end{bmatrix} \tag{3.57}$$

where

$$Q^{(uu)} = H^{(w)t}H^{(w)} + H^{(v)t}H^{(v)} + B^{(u)} \tag{3.58}$$

$$Q^{(vv)} = H^{(w)t}H^{(w)} + H^{(u)t}H^{(u)} + B^{(v)} \tag{3.59}$$

$$Q^{(ww)} = H^{(v)t}H^{(v)} + H^{(u)t}H^{(u)} + B^{(w)} \tag{3.60}$$

$$Q^{(uv)} = -H^{(u)t}H^{(v)},\ Q^{(vu)} = -H^{(v)t}H^{(u)} \tag{3.61}$$

$$Q^{(uw)} = -H^{(u)t}H^{(w)},\ Q^{(wu)} = -H^{(w)t}H^{(u)} \tag{3.62}$$

$$Q^{(vw)} = -H^{(v)t}H^{(w)},\ Q^{(wv)} = -H^{(w)t}H^{(v)}. \tag{3.63}$$

The stepsize α can then be computed as,

$$\alpha = \frac{\sum_{d\in\{u,v,w\}} g^{(d)t}g^{(d)}}{\sum_{d_1\in\{u,v,w\}}\sum_{d_2\in\{u,v,w\}} g^{(d_1)t}Q^{(d_1d_2)}g^{(d_2)}} \tag{3.64}$$

Similar to the computation of gradient, the computation of α can be implemented using filtering operations.

3.4.3 Algorithm Initialization

Since the cost function in (3.29) is convex, our optimization algorithm (algorithm 1) is convergent. We use multi-resolution initialization [3,65] to improve convergence speed. In this method, we reconstruct the magnetic vector potential and magnetization vector fields at coarse resolution scales and use the solution to initialize the reconstruction at finer scales. We use $N = 3$ number of multi-resolution stages during reconstruction such that the voxel width at the n^{th} multi-resolution stage is $2^{N-n}\Delta$, where $1 \leq n \leq N$ and Δ is the voxel width at the finest resolution scale. Hence, $n = 1$ corresponds to reconstruction at the lowest voxel resolution and $n = N$ corresponds to reconstruction at the highest voxel resolution. In algorithm 1, we set $\mu = 10$ and $\tau = 2(N - n + 1)$ at the n^{th} multi-resolution stage of reconstruction. At the coarsest resolution scale, we start by initializing the magnetization and magnetic vector potential reconstructions to zeros. The voxel width of the reconstruction at the finest resolution scale is equal to the pixel width of the input phase shift images. The prior model regularization parameter σ_x is set empirically such that we get the best visual quality of reconstruction.

3.5 Experimental Results

3.5.1 Simulated Data Results

In this section, we will present reconstructions of simulated data using the new MBIR algorithm and compare it to reconstructions using traditional methods. We will use a phantom that is representative of the typical magnetization states commonly observed in magnetic materials. The magnetization of a sample is described using the concept of magnetic domains i.e., regions in the sample where the neighboring magnetization vectors are oriented in the same direction. The 3D magnetization ground truth of a Ni_2MnGa simulated phantom is shown in Fig. 3.2 (a-f). Fig. 3.2 (a-c) shows a slice through the magnetization ground-truth in the $u - v$ plane

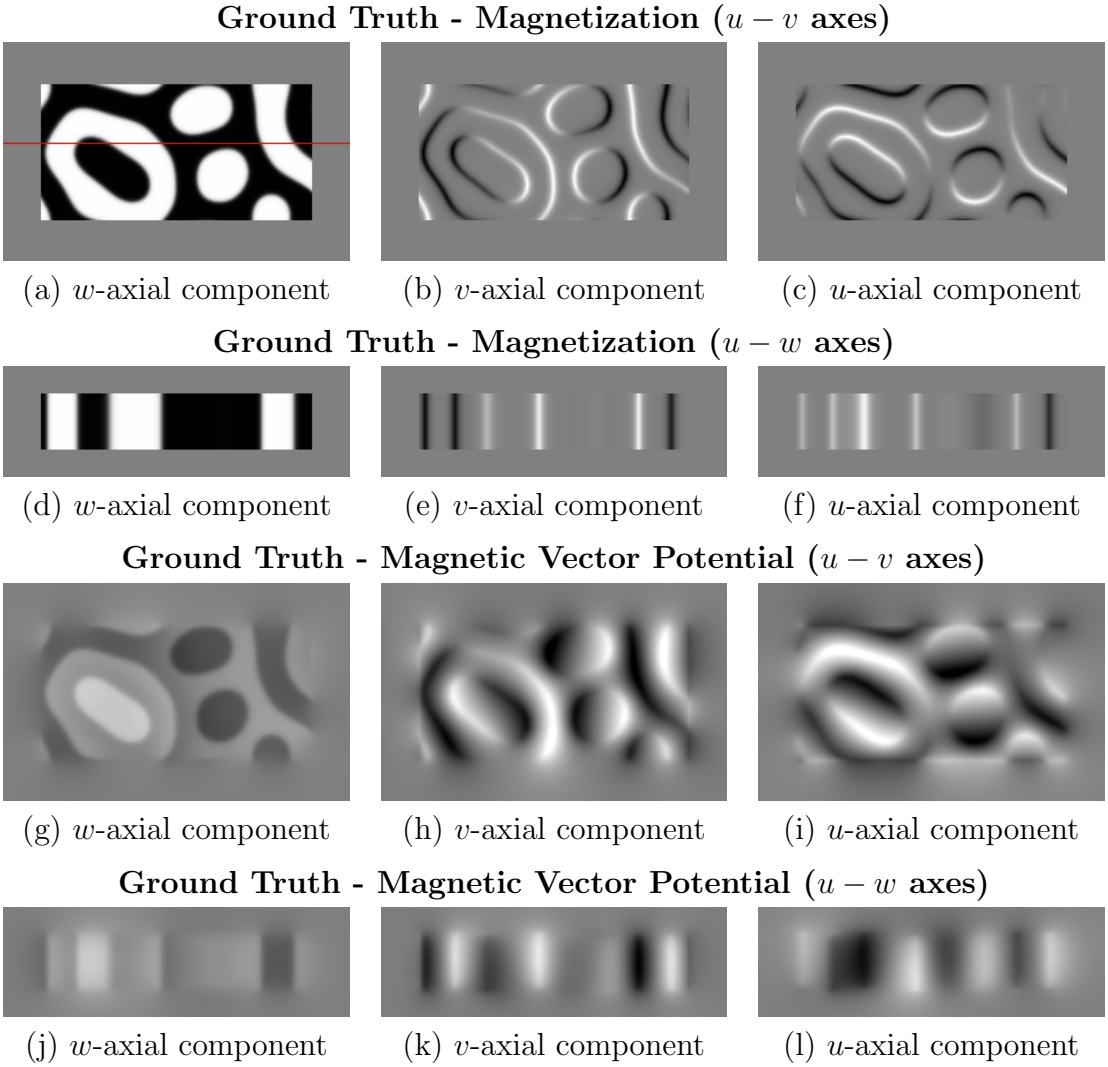
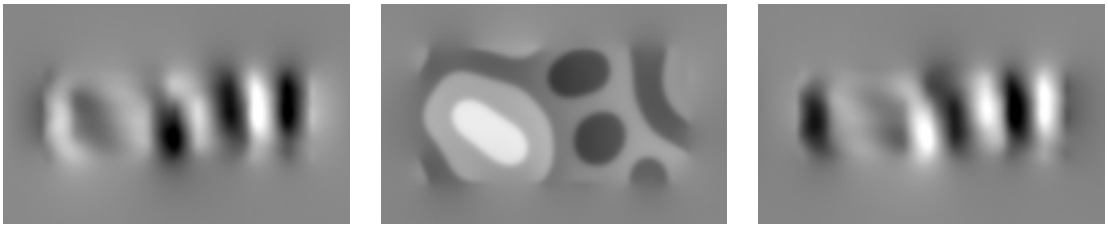


Fig. 3.2. Ground-truth of magnetization and magnetic vector potential. (a-c) and (d-f) show slices of the magnetization vector field along the $u-v$ plane (perpendicular to electron propagation) and $u-w$ plane (parallel to electron propagation) respectively. Similarly, (g-i) and (j-l) show slices of the magnetic vector potential along the $u-v$ plane and $u-w$ plane respectively. The 1^{st}, 2^{nd}, and 3^{rd} columns show the vector field components oriented along the $w-$axis, $v-$axis, and $u-$axis respectively. Note that (d-f) and (j-l) show slices in the $u-w$ plane that lie along the red line in (a).

perpendicular to the direction of electron propagation. Similarly, Fig. 3.2 (d-f) show a slice through the magnetization ground-truth in the $u-w$ plane parallel to the

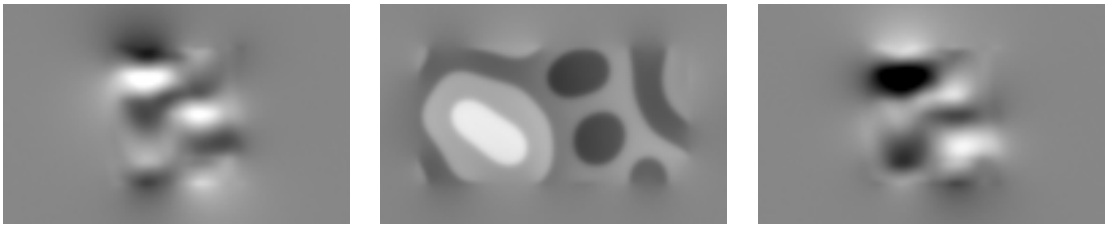
direction of electron propagation. The bright regions in Fig. 3.2 (a,d) correspond to magnetic domains where the magnetization direction is oriented along the positive w-axis. Similarly, the dark regions in Fig. 3.2 (a,d) correspond to magnetic domains oriented along the negative w-axis. Near the interface between the bright and dark regions, we can see the domain walls where the magnetization changes direction. This change in direction can be visualized by observing the v, u-axial components of magnetization shown in Fig. 3.2 (b,e) and Fig. 3.2 (c,f). Note that the magnetization has a constant magnitude of $4 \times 10^{-5} nm^{-2}$ within the sample (the inner rectangular region in Fig. 3.2 (a-f)).

Electron Phase Images for Tilt across u-axis



(a) -70^0 tilt angle (b) 0^0 tilt angle (c) 70^0 tilt angle

Electron Phase Images for Tilt across v-axis



(d) -70^0 tilt angle (e) 0^0 tilt angle (f) 70^0 tilt angle

Fig. 3.3. Simulated phase shift data generated from magnetization ground-truth phantom. (a-c) and (d-f) show the simulated electron phase shift images for various tilt angles when the sample is tilted across the $u-$axis and $v-$axis respectively. The 1^{st}, 2^{nd}, and 3^{rd} columns correspond to the tilt angles -70^0, 0^0, and 70^0 respectively.

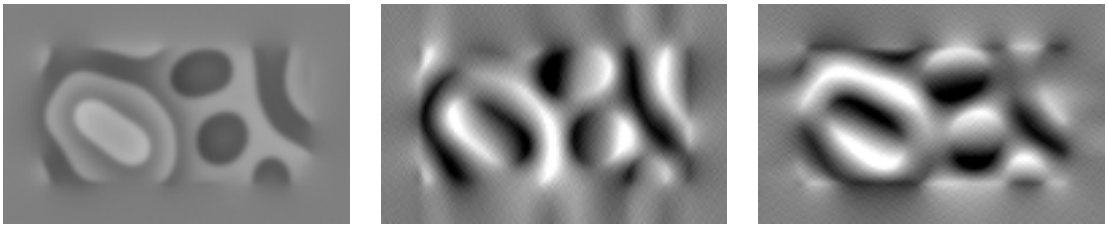
The magnetization results in a magnetic vector potential distribution extending both inside and outside the sample that is simulated using (3.14). Both the magnetization and magnetic vector potential ground truths have a size of $256 \times 256 \times 256$

and a voxel width of 2.5nm. The magnetic vector potential ground truth is shown in Fig. 3.2 (g-l). Fig. 3.2 (g-i) show a slice through the magnetic vector potential ground-truth in the $u-v$ plane perpendicular to the direction of electron propagation. Similarly, Fig. 3.2 (j-l) shows a slice through the magnetic vector potential ground truth in the $u-w$ plane parallel to the direction of electron propagation. In Fig. 3.2 (a-c) and (g-i), the horizontal direction is the $u-$axis and the vertical direction is the $v-$axis. In Fig. 3.2 (d-f) and (j-l), the horizontal direction is the $u-$axis and the vertical direction is the $w-$axis. Note that the slices in Fig. 3.2 (d-f) and (j-l) goes through the red colored line in Fig. 3.2 (a).

As the electron wave in a TEM propagates through the magnetic sample it undergoes a change in phase. We forward project the magnetic vector potential using the expression in (3.21) to generate the electron phase shift images at multiple tilt angles for both the u-axis tilt series and v-axis tilt series. We simulate electron phase images at tilt angles ranging from -70^0 to 70^0 at steps of 2^0 for tilt across both the $u-$axis and $v-$axis. Each simulated phase image has a size of 128×128 and a pixel size of 5nm. We also add Gaussian noise to the simulated phase images such that the average SNR is 56.85 dB. The simulated electron phase images at tilt angles of -70^0, 0^0, and 70^0 across the $u-$axis is shown in Fig. 3.3 (a-c). Similarly, the electron phase images at the same angles but for tilt across the $v-$axis is shown in Fig. 3.3 (d-f). Note that even though the number of tilt angles across the two axes are the same in this simulation, it is not a pre-requisite for our reconstruction algorithm. Also, the simulated angular range of -70^0 to 70^0 is reminiscent of typical TEM microscopy setups that do not permit data acquisition over a complete 180^0 angular range. All images in Fig. 3.3 (a-f) show the phase shift images such that the horizontal direction is the $u-$axis and the vertical direction is the $v-$axis.

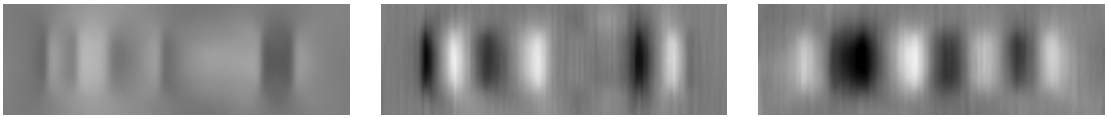
The simulated phase images at various tilt angles is then reconstructed using the new MBIR algorithm and the conventional method presented in [15]. The algorithm presented in [15] is an analytic technique for reconstruction of the 3D magnetic vector potential using a variation of the filtered back-projection algorithm. Note that there

Conventional Reconstruction of Magnetic Vector Potential ($u-v$ axes)



(a) w-axial component (b) v-axial component (c) u-axial component

Conventional Reconstruction of Magnetic Vector Potential ($u-w$ axes)



(d) w-axial component (e) v-axial component (f) u-axial component

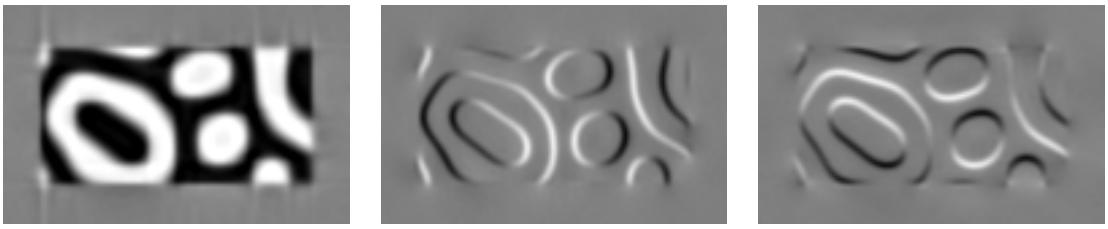
MBIR Reconstruction of Magnetic Vector Potential ($u-v$ axes)

(g) w-axial component (h) v-axial component (i) u-axial component

MBIR Reconstruction of Magnetic Vector Potential ($u-w$ axes)

(j) w-axial component (k) v-axial component (l) u-axial component

MBIR Reconstruction of Magnetization ($u-v$ axes)



(m) w-axial component (n) v-axial component (o) u-axial component

MBIR Reconstruction of Magnetization ($u-w$ axes)

(p) w-axial component (q) v-axial component (r) u-axial component

Fig. 3.4. Comparison of reconstructions of magnetization and magnetic vector potential using the conventional method and MBIR. (a-f) show the reconstruction of magnetic vector potential using the conventional method. (g-l) show the reconstruction of magnetic vector potential using MBIR. (m-r) show the reconstruction of magnetization using MBIR.

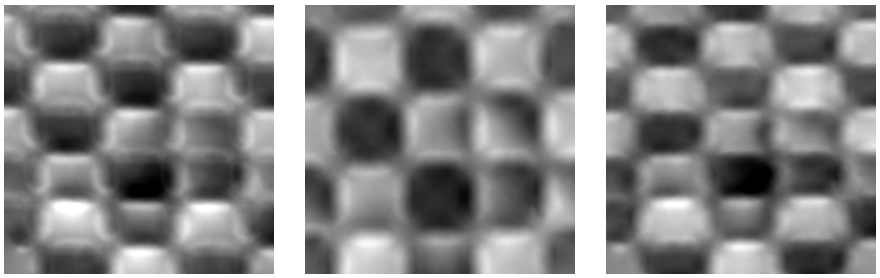
does not exist any algorithm for 3D reconstruction of magnetization from TEM phase images. All reconstructions have a size of $128 \times 128 \times 128$ in 3D space and a voxel size of 5nm. The reconstruction of magnetic vector potential using the conventional method and MBIR is shown in Fig. 3.4 (a-f) and Fig. 3.4 (g-l) respectively. The reconstruction of magnetization using MBIR is shown in Fig. 3.4 (m-r). The images in Fig. 3.4 (a-c), (g-i), and (m-o) show a slice of the corresponding vector field in the $u - v$ axial plane perpendicular to the direction of electron propagation. Similarly, the images in Fig. 3.4 (d-f), (j-l), and (p-r) show a slice of the corresponding vector field in the $u-w$ axial plane parallel to the direction of electron propagation. We can see that the reconstruction of the magnetic vector potential using the conventional method results in artifacts in the $v-$axis component shown in Fig. 3.4 (b,e) and the $u-$axis component shown in 3.4 (c,f). In contrast, our MBIR algorithm accurately reconstructs the magnetic vector potential without any artifacts as shown in Fig. 3.4 (g-l). Furthermore, the reconstruction of magnetization using MBIR in Fig. 3.4 (m-r) clearly show the various magnetic domains and the domain walls in the sample that cannot be visualized by just observing the magnetic vector potential in Fig. 3.4 (a-l).

The images showing the magnetization in Fig. 3.2 (a-f) and Fig. 3.4 (m-r) are scaled from a minimum of $-4 \times 10^{-5} nm^{-2}$ to a maximum of $4 \times 10^{-5} nm^{-2}$. The images showing the magnetic vector potential in Fig. 3.2 (g-l) and Fig. 3.4 (a-l) are scaled from a minimum of $-11 \times 10^{-3} nm^{-1}$ to a maximum of $11 \times 10^{-3} nm^{-1}$. The phase shift images in Fig. 3.3 (a-f) are scaled from a minimum of -1.2 radians to a maximum of 1.1 radians.

3.5.2 Real Data Results

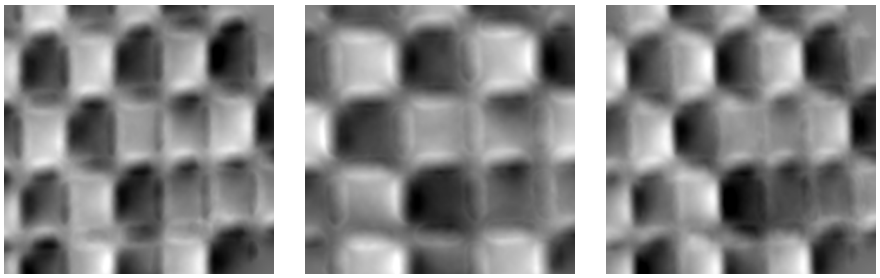
In this section, we will present reconstructions of real data using the new MBIR algorithm and compare it to reconstructions using traditional methods. A conventional TEM was used to image a NiFe sample in the Lorentz mode [15] using vector field electron tomography. At each tilt angle, the electron phase is recovered from mea-

Electron Phase Tilt Series Images across u-axis



(a) -50^0 tilt angle (b) 0^0 tilt angle (c) 50^0 tilt angle

Electron Phase Tilt Series Images across v-axis



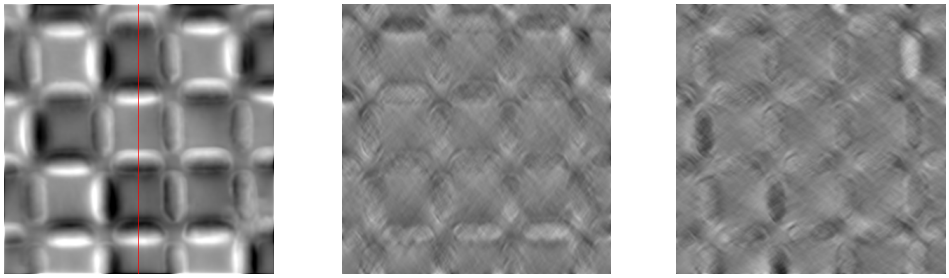
(d) -50^0 tilt angle (e) 0^0 tilt angle (f) 50^0 tilt angle

Fig. 3.5. Electron phase images of a Ni-Fe sample acquired using a TEM. (a-c) and (d-f) show the phase images at various tilt angles for tilt across the $u-$axis and $v-$axis respectively. (a,d), (b,e), and (c,f) are phase images at tilt angles of -50^0, 0^0, and 50^0 respectively.

surements using the transport-of-intensity phase retrieval algorithm presented in [80]. The data consists of electron phase images at tilt angles ranging from -50^0 to 50^0 at steps of 1^0 for tilt across both the $u-$axis and $v-$axis. Each phase image has a resolution of 3.0nm per pixel and a size of 384×384. The phase images at tilt angles of -50^0, 0^0, and 50^0 for tilt across the $u-$axis is shown in Fig. 3.5 (a-c). Similarly, the phase images at the same tilt angles but for tilt across the $v-$axis is shown in Fig. 3.5 (d-f). In Fig. 3.5 (a-f), the horizontal direction is the $u-$axis and the vertical direction is the $v-$axis.

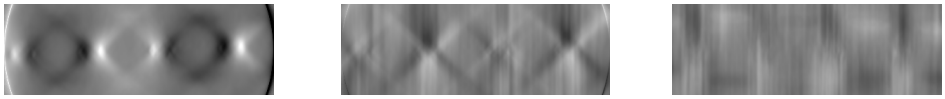
The phase data is then reconstructed using the proposed MBIR algorithm and the conventional method presented in [15]. All reconstructions of magnetization and magnetic vector potential have a voxel width of 3.0nm. The reconstruction of mag-

Conventional Reconstruction of Magnetic Vector Potential ($u-v$ axes)



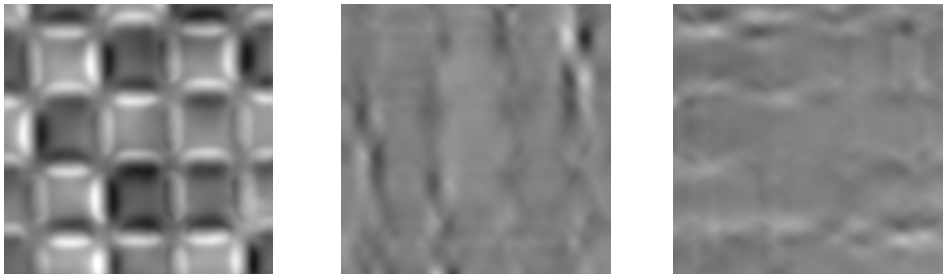
(a) w-axial component (b) v-axial component (c) u-axial component

Conventional Reconstruction of Magnetic Vector Potential ($v-w$ axes)



(d) w-axial component (e) v-axial component (f) u-axial component

Reconstruction of Magnetic Vector Potential using MBIR ($u-v$ axes)

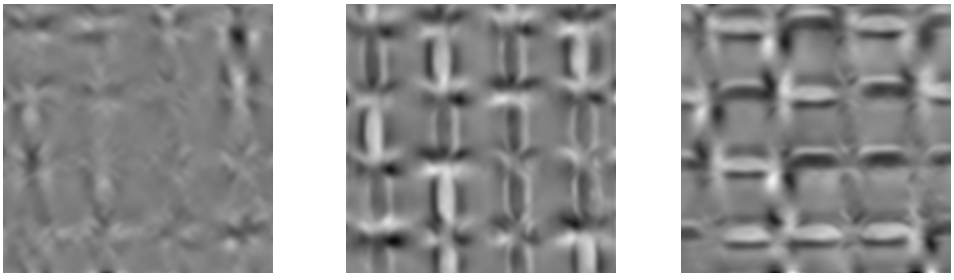


(g) w-axial component (h) v-axial component (i) u-axial component

Reconstruction of Magnetic Vector Potential using MBIR ($v-w$ axes)

(j) w-axial component (k) v-axial component (l) u-axial component

Reconstruction of Magnetization using MBIR ($u-v$ axes)



(m) w-axial component (n) v-axial component (o) u-axial component

Reconstruction of Magnetization using MBIR ($v-w$ axes)

(p) w-axial component (q) v-axial component (r) u-axial component

Fig. 3.6. Reconstructions of magnetization and magnetic vector potential of a Ni-Fe sample using the conventional method and MBIR.

netic vector potential using the conventional method and the MBIR algorithm is shown in Fig. 3.6 (a-f) and Fig. 3.6 (g-l) respectively. The conventional method results in prominent artifacts in the magnetic vector potential reconstruction of the $v-$axial component shown in Fig. 3.6 (b,e) and the $u-$axial component shown in Fig. 3.6 (c,f). In contrast, our MBIR algorithm significantly reduces the artifacts in the reconstruction of magnetic vector potential as shown in Fig. 3.6 (g-l). The reconstruction of magnetization using the MBIR algorithm is shown in Fig. 3.6 (m-r). The magnetization reconstructions clearly show the various magnetic domains in the sample. From Fig. 3.6 (n), we can see that the vertically aligned magnetic domains have magnetization oriented along the positive and negative $v-$axis. Similarly, we can see in Fig. 3.6 (o) that the horizontally aligned magnetic domains have magnetization oriented along the positive and negative $u-$axis. We could not have gained such an in-depth understanding of the magnetic domains in our Ni-Fe sample by just observing the magnetic vector potential in Fig. 3.6 (a-l). The images in Fig. 3.6 (a-c), (g-i), and (m-o) show a slice through the sample in the $u-v$ axial plane such that the horizontal direction is the $u-$axis and the vertical direction is the $v-$axis. Similarly, the images in Fig. 3.6 (d-f), (j-l), and (p-r) show a slice through the sample in the $v-w$ axial plane such that the horizontal direction is the $v-$axis and the vertical direction is the $w-$axis. Note that the slices in Fig. 3.6 (d-f), (j-l), and (p-r) go through the red colored line in Fig. 3.6 (a).

The phase shift images in Fig. 3.5 (a-f) are scaled from a minimum of -0.40 radians to a maximum of 0.43 radians. The images showing the magnetic vector potential in Fig. 3.6 (a-l) are scaled from a minimum of $-4\times10^{-3}nm^{-1}$ to a maximum of $4\times10^{-3}nm^{-1}$. The images showing the magnetization in Fig. 3.6 (m-r) are scaled from a minimum of $-1.5\times10^{-5}nm^{-2}$ to a maximum of $1.5\times10^{-5}nm^{-2}$.

4. CRITIR: ALGORITHM FOR DIRECT MODEL-BASED TOMOGRAPHIC RECONSTRUCTION OF THE COMPLEX REFRACTIVE INDEX

4.1 Introduction

Traditional X-ray computed tomography (CT) is widely used to reconstruct the X-ray absorption index (i.e., the imaginary part of the complex refractive index), and it assumes no X-ray diffraction. In practice, X-rays diffract whenever the object to detector distance is large, as happens in cone beam systems at high magnifications. So traditional CT is restricted to lower cone-beam magnifications. Alternatively, X-ray phase contrast tomography (XPCT) has become increasingly popular as a method for tomographic imaging when both diffraction and absorption occur. Moreover, since the contrast in the real part of the index of refraction that is responsible for diffraction is typically orders of magnitude greater than the imaginary part, XPCT typically has much higher contrast than traditional CT. In XPCT, a parallel beam of X-rays, typically from a synchrotron, is used to repeatedly image a sample rotating at a constant speed as shown in Fig. 4.1. The measured images from all the view angles are then reconstructed into a 3D map of the complex refractive index within the sample.

However, the widely used XPCT reconstruction algorithms have severe limitations [14]. First, all existing XPCT reconstruction algorithms either assume that the object is weakly absorbing [94, 95] or that the real part of the refractive index is linearly related to the imaginary part [95–98]. More importantly, existing XPCT algorithms generally assume that the diffraction is weak. For example, the widely used reconstruction algorithms presented in [94, 96–98] assume the near-field condition where only the first order Fresnel diffraction fringe is resolved. These methods result

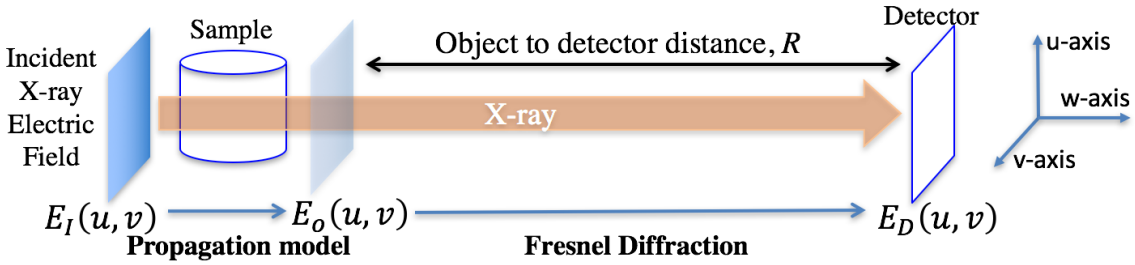


Fig. 4.1. Illustration of data acquisition in X-ray phase contrast tomography (XPCT). A parallel beam of X-rays from a synchrotron is used to repeatedly image a sample rotating at a constant speed. The X-rays undergo absorption and refraction as it propagates through the object and undergo Fresnel diffraction as it propagates towards the detector.

in inaccurate reconstructions with artifacts and excessively smooth edges beyond the near-field region. Other methods allow for more diffraction in the form of higher-order Fresnel diffraction fringes obtained beyond the near-field region [14,95]. However, they assume that the scattered field is weak in both amplitude and phase [14, 95], which severely limits the variety of samples that can be used for imaging. Nevertheless, these methods are widely used outside the scope of their validity due to the lack of better algorithms for XPCT. Note that all the above algorithms rely on a linear approximation to the inherently non-linear forward model that relates the measurement to the complex refractive index. Furthermore, they also rely on a two-step procedure that performs 2D phase retrieval at each view before performing a 3D tomographic reconstruction. There also exist iterative algorithms [99–101] that perform non-linear phase retrieval at each view before performing a tomographic reconstruction.

In this chapter, we present a model-based iterative reconstruction (MBIR) algorithm called complex refractive index tomographic iterative reconstruction (CRITIR) [11] that does not make simplifying approximations beyond the assumption of Fresnel diffraction. In CRITIR, the reconstruction of the complex refractive index is given by iteratively minimizing a cost function consisting of a forward model term and a prior model term that enforces sparsity in reconstruction. The forward model expresses the

measured data at all the view angles as a function of the complex refractive index distribution of the sample. The forward model accounts for both absorption and phase contrast in the measurements by modeling the physics of X-ray interaction with the object and Fresnel diffraction [102]. However, the forward model is highly non-linear due to the loss of phase during measurement and the complex exponential form of attenuation and refraction. To minimize the resulting non-convex cost function, we propose a novel algorithm using variable splitting and the method of alternating direction method of multipliers (ADMM) [17] to decompose the original problem into an iterative solution of multiple simpler optimization problems. Our framework allows for modular software architecture where we can independently optimize the algorithms used to solve the various simpler optimization problems. Unlike conventional methods, CRITIR does not make any linear approximations in the forward model and also directly reconstructs the complex refractive index from the measurements at all the view angles.

4.2 Propagation Model

In this section, we express the X-ray electric field in the detector plane as a function of the complex refractive index that is defined as [103],

$$n(u,v,w) = 1 - \delta(u,v,w) + j\beta(u,v,w) \tag{4.1}$$

where $j = \sqrt{-1}$, $\beta(u,v,w)$ is the absorption index, and $\delta(u,v,w)$ is the refractive index decrement of the object at the spatial coordinates (u,v,w). The relation between the X-ray electric field incident on the object, $f_I(u,v)$, and the electric field at the exit plane of the object, $f_O(u,v)$, is [102, 103],

$$f_O(u,v) = f_I(u,v) \exp\left\{ \frac{2\pi j}{\lambda} \int_w \left[1 - \delta(u,v,w) \quad + j\beta(u,v,w)\right] dw \right\} \tag{4.2}$$

where λ is the wavelength and w-axis is the direction of X-ray propagation. Note that the absorption index, β, is related to the linear attenuation coefficient, μ, as $\mu = (4\pi/\lambda)\,\beta$. From (4.2), we can see that the real part of the refractive index

modulates the phase whereas the imaginary part, the absorption index, modulates the amplitude of the incident electric field.

In XPCT, the scattering is mainly confined within a small angular range along the direction of propagation. Therefore, paraxial approximation can be safely used in the near-field diffraction regime and beyond. Thus, we use the Fresnel integral to express the electric field in the detector plane, $f_D(u,v)$, as a linear convolution of $f_O(u,v)$ with the Fresnel impulse response [102] as shown below,

$$f_D(u,v) = \frac{j\exp(jkR)}{\lambda R}\int_{u'}\int_{v'} f_O(u',v')\exp\left(\frac{j\pi}{\lambda R}\left[(u-u')^2+(v-v')^2\right]\right)du'dv'. \tag{4.3}$$

where R is the object to detector distance and $k = 2\pi/\lambda$ is the wavenumber. Note that it is sufficient to consider the case of parallel beam since the physics of cone-beam can be expressed in the form of an equivalent parallel beam configuration [102].

The partial coherence of the X-ray electric field limits the achievable spatial resolution. This effect is modeled in Fourier space by multiplying the Fresnel frequency response by a window such as a Gaussian [102]. Thus, the relation between the Fourier transform of $f_D(u,v)$ denoted by $F_D(\mu,\nu)$, and the Fourier transform of $f_O(u,v)$ denoted by $F_O(\mu,\nu)$ is [102],

$$F_D(\mu,\nu) = W(\mu,\nu)\exp(jkR)\exp\left((-j\pi\lambda R)\left(\mu^2+\nu^2\right)\right)F_O(\mu,\nu) \tag{4.4}$$

where (μ,ν) are the spatial frequency coordinates in inverse units of (x,y), $W(\mu,\nu) = \exp\left(-(\mu^2+\nu^2)/(2\gamma^2)\right)$, and γ is the standard deviation of the Gaussian window [102].

4.3 Problem Formulation

In MBIR, the reconstruction is given by solving the following optimization problem,

$$\hat{x} = \underset{x\geq 0}{\operatorname{argmin}}\left\{-\log p(y|x) - \log p(x)\right\} \tag{4.5}$$

where $\log p(y|x)$ is the forward log-likelihood function for the data, y, given the object x and $\log p(x)$ is the prior log-likelihood function for the object, x.

We first derive the forward likelihood function $p(y|x)$. Let the real part, $x^{(R)}$, of the vector array x give the refractive index decrement and the imaginary part, $x^{(I)}$, give the absorption index voxel values in discrete space. Then, the electric field in the exit plane of the object is given by numerically evaluating equation (4.2) as, $D\exp(-A_i x)$, where A_i is the matrix that implements the line integral in (4.2) at the i^{th} view, D is a diagonal matrix such that the diagonal elements give the incident X-ray electric field, and $\exp(z)$ is element-wise exponentiation of the vector z. Here, the object independent constant phase multiplier in (4.2) is ignored as it does not contribute to the measurement. In case of plane wave illumination, D is an identity matrix with a fixed multiplicative constant.

The detector response is proportional to the X-ray intensity, which is the magnitude square of the X-ray electric field after it propagates from the object to the detector. The variance in detector count measurements is typically modeled using Poisson statistics [3]. By the variance stabilizing property of the square root transformation for Poisson random variables, we can show that the square root of detector measurements, y_i, is well approximated by a constant variance Gaussian distribution with mean equal to $|HD\exp(-A_i x)|$ where H is the matrix that implements the Fresnel diffraction integral in Fourier space using (4.4) [104] and $|.|$ gives the element-wise magnitude of a vector. Thus, the relation between the data, y_i, at the i^{th} view and the object, x, is,

$$y_i = |HD\exp(-A_i x)| + n_i \tag{4.6}$$

where n_i is the noise in data, y_i, at the i^{th} view.

From (4.6), it follows that the forward log-likelihood function of y given x is,

$$-\log p(y|x) = \sum_{i=1}^{M} \frac{1}{2\sigma^2} \left|\left|y_i - |HD\exp(-A_i x)|\right|\right|^2 + g(y) \tag{4.7}$$

where σ^2 is the noise variance, M is the total number of views, and $g(y)$ is a constant.

Next, we derive the prior log-likelihood function, $\log p(x)$. Typically, there is a strong correlation between the refractive index decrement, $x_j^{(R)}$, and absorption index,

$x_j^{(I)}$, values of an object [96, 97] . The transform, T, that decorrelates the real part, $x_j^{(R)}$, and the imaginary part, $x_j^{(I)}$, of the voxel x_j is given by,

$$\begin{bmatrix} x_j^{(+)} \\ x_j^{(-)} \end{bmatrix} = \begin{bmatrix} 1 & \alpha \\ 1 & -\alpha \end{bmatrix} \begin{bmatrix} x_j^{(R)} \\ x_j^{(I)} \end{bmatrix} \tag{4.8}$$

where α is the ratio of standard deviation of $x_j^{(R)}$ to $x_j^{(I)}$.

In the prior model, we independently regularize the sum component, $x_j^{(+)}$, and the difference component, $x_j^{(-)}$, of x_j using a 26−point local neighborhood based qGGMRF prior model [4]. The prior model is given by,

$$R(x) = \sum_{i,j \in \mathcal{N}} w_{ij} \rho \left(Tx_i - Tx_j \right) \tag{4.9}$$

where $\mathcal{N}$ is the set of all pairwise cliques in 3D space and $\rho(\Delta)$ is the potential function given by,

$$\rho \left(\begin{bmatrix} \Delta^{(+)} \\ \Delta^{(-)} \end{bmatrix} \right) = \frac{\left| \Delta^{(+)} / \sigma_+ \right|^2}{c_+ + \left| \Delta^{(+)} / \sigma_+ \right|^{2-q}} + \frac{\left| \Delta^{(-)} / \sigma_- \right|^2}{c_- + \left| \Delta^{(-)} / \sigma_- \right|^{2-q}} \tag{4.10}$$

where $q = 1.2, \sigma_+, \sigma_-, c_+$, and c_- are qGGMRF parameters. The weight parameter w_{ij} is set such that it is inversely proportional to the spatial distance between voxels x_i and x_j and normalized such that $\sum_{j \in \mathcal{N}_i} w_{ij} = 1$, where $\mathcal{N}_i$ is the set of all neighboring voxel indices of voxel x_i. The decorrelating transform in (4.8) and the prior model in (4.9) can be used to enforce prior knowledge about the object. For instance, if we know that the refractive index decrement, δ, is approximately proportional to the absorption index, β, then we can set $\alpha = \delta/\beta$ and $\sigma_-/\sigma_+ << 1$.

By substituting (4.7) and (4.9) in (4.5), we can formulate the reconstruction as,

$$\hat{x} = \underset{x \geq 0}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \frac{1}{2\sigma^2} \left| \left| y_i - \left| HD \exp\left(-A_i x \right) \right| \right| \right|^2 + R(x) \right\} \tag{4.11}$$

where $-\log p(x) = R(x)$ is the prior model.

The above form of the cost function is non-convex and difficult to minimize directly. To simplify the problem, we will use variable splitting using the auxiliary vectors $w_i = \exp(-z_i)$ and $z_i = A_i x$, and formulate the problem as a constrained optimization problem of the form,

$$(\hat{x}, \hat{z}, \hat{w}) = \underset{x \geq 0, z, w}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} \sum_{i=1}^{M} ||y_i - |HDw_i|||^2 + R(x) \right\}$$

$$\text{subject to } w_i = \exp(-z_i) \text{ and } z_i = A_i x \; \forall \, i. \quad (4.12)$$

The above form of constrained optimization problem can be solved using the augmented Lagrangian and the method of alternate direction method of multipliers (ADMM) [17].

4.4 Optimization Algorithm

Algorithm 4 RECONSTRUCTION

1: **for all** k in 1 to K_{max} **do**

2: The Tomographic Inversion Step -

$$\hat{x} \leftarrow \underset{x \geq 0}{\operatorname{argmin}} \left\{ \frac{\mu}{2} \sum_{i=1}^{M} ||\hat{z}_i - u_i - A_i x||^2 + R(x) \right\} \quad (4.13)$$

3: **for all** l in 1 to L_{max} **do**

4: Estimating the complex exponential -

$$\hat{z} \leftarrow \underset{z}{\operatorname{argmin}} \left\{ \frac{\mu}{2} \sum_{i=1}^{M} ||A_i \hat{x} - z_i + u_i||^2 + \frac{\nu}{2} \sum_{i=1}^{M} ||\exp(-z_i) - \hat{w}_i + v_i||^2 \right\}. \quad (4.14)$$

5: Phase Retrieval Step -

$$\hat{w} \leftarrow \underset{w}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} \sum_{i=1}^{M} ||y_i - |HDw_i|||^2 + \frac{\nu}{2} \sum_{i=1}^{M} ||\exp(-\hat{z}_i) - w_i + v_i||^2 \right\}. \quad (4.15)$$

6: $v_i \leftarrow v_i + (\exp(-\hat{z}_i) - \hat{w}_i)$

7: **end for**

8: $u_i \leftarrow u_i + (A_i \hat{x} - \hat{z}_i)$

9: **end for**

Using ADMM [17], we can solve (4.12) by iteratively solving multiple but simpler optimization problems as shown in algorithm (4). Here u_i and v_i are the scaled dual

vectors, $\mu > 0$ and $\nu > 0$ are the augmented Lagrangian parameters, and K_{max} and L_{max} are the maximum number of outer and inner loop iterations respectively.

4.4.1 Tomographic Inversion

The solution to the minimization problem in (4.13) is the value of x that minimizes the cost function given by,

$$c_x(x) = \frac{\mu}{2} \sum_{i=1}^{M} ||\hat{z}_i - u_i - A_i x||^2 + \sum_{i,j \in \mathcal{N}} w_{ij} \rho \left(Tx_i - Tx_j \right). \tag{4.16}$$

The optimization problem in (4.13) is solved using an appropriate modification of the iterative coordinate descent (ICD) algorithm used in [1–3, 52]. We iteratively minimize (4.16) with respect to each voxel x_j sequentially until the algorithm converges. To achieve this, we will reformulate (4.16) in terms of just x_j by excluding terms that do not contain x_j. Let x' denote the current estimate for the vector x. Then, the cost function with respect to x_j is given by,

$$c_x^{(vox)}(x_j) = \tilde{\theta}^t x_j + \frac{1}{2}(x_j - x_j')^t \tilde{\Theta}(x_j - x_j') + \sum_{i \in \mathcal{N}_j} w_{ji} \rho \left(Tx_j - Tx_i' \right) \tag{4.17}$$

where $x_j = \left[x_j^{(R)}, x_j^{(I)} \right]^t$ is a vector representation of the real, $x_j^{(R)}$, and imaginary, $x_j^{(I)}$, parts of x_j, $\tilde{\theta}$ and $\tilde{\Theta}$ are the gradient and Hessian of $\frac{\mu}{2} \sum_{i=1}^{M} ||\hat{z}_i - u_i - A_i x||^2$ with respect to x_j.

Next, we will derive the gradient $\tilde{\theta}$ and the Hessian $\tilde{\Theta}$ parameters in (4.17). Let the complex valued error sinogram vector be defined as,

$$e_i = z_i - u_i - A_i x.$$

Then, the elements of the gradient vector $\tilde{\theta}$ are given by,

$$\tilde{\theta}_1 = \mu \sum_{i=1}^{M} e^{(R)t} A_{i,*,j} \tag{4.18}$$

$$\tilde{\theta}_2 = \mu \sum_{i=1}^{M} e^{(I)t} A_{i,*,j} \tag{4.19}$$

where $e^{(R)}$ and $e^{(I)}$ denote the real and imaginary parts of the error sinogram vector e and $A_{i,*,j}$ represent the elements along the j^{th} column of the projection matrix A_i. Similarly, the elements of the Hessian matrix $\tilde{\Theta}$ are given by,

$$\Theta_{1,1} = \Theta_{2,2} = \mu \sum_{i=1}^{M} ||A_{i,*,j}||^2 , \tag{4.20}$$

$$\Theta_{1,2} = \Theta_{2,1} = 0. \tag{4.21}$$

Since the prior model term in (4.17) is not quadratic, we will replace it with a quadratic substitute function [3,59] such that the new cost function upper-bounds the cost function in (4.17). We can then show that minimizing the new substitute cost function will reduce the original cost function (4.17). The substitute cost function with respect to x_j is given by,

$$c_x^{(sub)}(x_j) = \tilde{\theta}^t x_j + \frac{1}{2}(x_j - x_j')^t \tilde{\Theta}(x_j - x_j') + \sum_{i \in \mathcal{N}_j} \frac{w_{ji}}{2}(Tx_j - Tx_i')^t A_{ji}(Tx_j - Tx_i') \tag{4.22}$$

where A_{ji} is the substitute function parameter that will be derived later. Note that the substitute function must be reevaluated after every voxel update.

Now, we will derive the substitute function parameter A_{ji}. Let $\rho_+(\Delta^{(+)}) = \frac{\left|\Delta^{(+)}/\sigma_+\right|^2}{c_+ + \left|\Delta^{(+)}/\sigma_+\right|^{2-q}}$ and $\rho_-(\Delta^{(-)}) = \frac{\left|\Delta^{(-)}/\sigma_-\right|^2}{c_- + \left|\Delta^{(-)}/\sigma_-\right|^{2-q}}$ be the individual potential functions in (4.10). Then, we can show that,

$$A_{ji} = \begin{bmatrix} A_{ji}^{(+)} & 0 \\ 0 & A_{ji}^{(-)} \end{bmatrix} \tag{4.23}$$

where

$$A_{ji}^{(+)} = \begin{cases} \frac{\rho_+'(T_{1,*}x_j' - T_{1,*}x_i')}{T_{1,*}x_j' - T_{1,*}x_i'} & \text{if } T_{1,*}x_j' \neq T_{1,*}x_i' \\ \rho_+''(0) & \text{if } T_{1,*}x_j' = T_{1,*}x_i' \end{cases}, \tag{4.24}$$

$$A_{ji}^{(-)} = \begin{cases} \frac{\rho_-'(T_{2,*}x_j' - T_{2,*}x_i')}{T_{2,*}x_j' - T_{2,*}x_i'} & \text{if } T_{2,*}x_j' \neq T_{2,*}x_i' \\ \rho_-''(0) & \text{if } T_{2,*}x_j' = T_{2,*}x_i' \end{cases}, \tag{4.25}$$

$\rho_+'(\Delta^{(+)})$ and $\rho_+''(\Delta^{(+)})$ are the first and second derivatives of $\rho_+(\Delta^{(+)})$ respectively, $\rho_-'(\Delta^{(-)})$ and $\rho_-''(\Delta^{(-)})$ are the first and second derivatives of $\rho_-(\Delta^{(-)})$ respectively,

$T_{1,*}$ represents the elements along the first row of T and $T_{2,*}$ represents the elements along the second row of T.

Then, the value of x_j that minimizes (4.22) is given by,

$$\hat{x}_j = \left(\tilde{\Theta} + \sum_{i \in \mathcal{N}_j} w_{ji} T^t A_{ji} T \right)^{-1} \left(-\tilde{\theta} + \tilde{\Theta} x'_j + \sum_{i \in \mathcal{N}_j} w_{ji} T^t A_{ji} T x'_i \right) \tag{4.26}$$

Thus, the algorithm that solves the minimization problem in (4.13) by minimizing the cost function in (4.16) is shown in algorithm 5.

Algorithm 5 TOMOGRAPHIC INVERSION

while not converged **do**
  for all voxel indices j **do** $//x'_j$ is current estimate of x_j
    $x' \leftarrow \hat{x}$
    Compute $\tilde{\theta}$ by substituting x'_j for x_j in (4.18) and (4.19).
    Compute $\tilde{\Theta}$ using (4.20) and (4.21).
    Compute A_{ji} using (4.23).
    $\hat{x}_j \leftarrow \left(\tilde{\Theta} + \sum_{i \in \mathcal{N}_j} w_{ji} T^t A_{ji} T \right)^{-1} \left(-\tilde{\theta} + \tilde{\Theta} x'_j + \sum_{i \in \mathcal{N}_j} w_{ji} T^t A_{ji} T x'_i \right)$
    $e_i \leftarrow e_i - A_{i,*,j} \left(\hat{x}_j - x'_j \right)$
  end for
end while

4.4.2 Complex Phase Estimation

The minimization problem shown in (4.14) is separable in each element, $z_{i,k}$, of the vector z_i and reduces to multiple 2-dimensional minimization problems (in real and imaginary parts of $z_{i,k}$). The cost function in (4.14) with respect to $z_{i,k}$ is given by,

$$c_z^{(vox)}(z_{i,k}) \leftarrow \underset{z_{i,k}}{\operatorname{argmin}} \left\{ \frac{\mu}{2} \left| A_{i,k,*} \hat{x} - z_{i,k} + u_{i,k} \right|^2 + \frac{\nu}{2} \left| \exp(-z_{i,k}) - \hat{w}_{i,k} + v_{i,k} \right|^2 \right\}. \tag{4.27}$$

where $A_{i,k,*}$ represent the elements along the k^{th} row of the matrix $A_{i,k}$. The cost function (4.27) is minimized using the Nelder-Mead simplex algorithm [105]. Thus, (4.14) is solved by formulating (4.27) for each voxel $z_{i,k}$ and minimizing it.

4.4.3 Phase Retrieval

To solve the optimization problem in (4.15), we will use a diagonal phase matrix Ω_i such that $|\Omega_{i,k,k}| = 1$ to account for the unknown phase of y_i. We will minimize (4.15) by reformulating it as the following optimization problem,

$$\left(\hat{w}, \hat{\Omega}\right) = \underset{w,\Omega}{\operatorname{argmin}} \left\{ \frac{1}{2\sigma^2} \sum_{i=1}^{M} ||y_i - \Omega_i H D w_i||^2 + \frac{\nu}{2} \sum_{i=1}^{M} ||\exp(-\hat{z}_i) - w_i + v_i||^2 \right\} \text{ s.t. } |\Omega_{i,k,k}| = 1 \quad (4.28)$$

To solve (4.28), we will alternately minimize over w and Ω till convergence is achieved. Since (4.28) is quadratic in w, the minimization over w is solved using the steepest-descent algorithm [62]. The minimization over Ω_i is separable in each of its diagonal elements and reduces to the problem of computing projections onto a circle with unit radius.

First, we will consider the minimization of (4.28) with respect to Ω. Note that (4.28) is separable in each diagonal element $\Omega_{i,k,k}$ and hence we can rewrite the minimization as,

$$\hat{\Omega}_{i,k,k} = \underset{\Omega_{i,k,k}}{\operatorname{argmin}} \frac{1}{2\sigma^2} |y_{i,k} - \Omega_{i,k,k} H_{k,*} D w_i|^2 \text{ subject to } |\Omega_{i,k,k}| = 1 \quad (4.29)$$

The solution to this problem is the projection of $y_{i,k}(H_{k,*}Dw_i)^{-1}$ on a circle with unit radius and can be computed as,

$$\hat{\Omega}_{i,k,k} = \frac{y_{i,k}(H_{k,*}Dw_i)^{-1}}{|y_{i,k}(H_{k,*}Dw_i)^{-1}|} \quad (4.30)$$

Next, let us consider the minimization of (4.28) with respect to w_i. Since (4.28) is separable in i, we can minimize (4.28) with respect to each w_i independently. The cost function in (4.28) is quadratic in w_i and we use the steepest descent algorithm [62]

to minimize it. Thus, the cost function (4.28) with respect to w_i after ignoring terms that does not depend on w_i is,

$$c_w(w_i) = \frac{1}{2\sigma^2} ||y_i - \Omega_i H D w_i||^2 + \frac{\mu}{2} ||\exp(-\hat{z}_i) - w_i + v_i||^2 . \tag{4.31}$$

Algorithm 6 PHASE RETRIEVAL ALGORITHM

1: **while** algorithm has not converged **do**

2: Compute g_i and α_i by substituting $\hat{w}$ and $\hat{\Omega}$ in (4.34) and (4.35).

3: ESTIMATION OF w -

$$\hat{w}_i \leftarrow \hat{w}_i - \alpha_i g_i \quad \forall\, i \tag{4.32}$$

4: ESTIMATION OF Ω -

$$\hat{\Omega}_{i,k,k} \leftarrow \frac{y_{i,k}(H_{k,*} D \hat{w}_i)^{-1}}{|y_{i,k}(H_{k,*} D \hat{w}_i)^{-1}|} \quad \forall\, i,k \tag{4.33}$$

5: **end while**

Let the gradient of $c_w(w_i)$ with respect to the real and imaginary parts of $w_i = w_i^{(R)} + i w_i^{(I)}$ be represented in complex form as $g_i = g_i^{(R)} + i g_i^{(I)}$. Then, we can show that,

$$g_i = \frac{1}{\sigma^2} U_i^* U_i w_i - \frac{1}{\sigma^2} U_i^* y_i + \mu \left[w_i - \exp(-\hat{z}_i) - v_i \right] \tag{4.34}$$

where $U_i = \Omega_i H D$ and U_i^* represents the conjugate transpose of U_i. Each iteration of the steepest descent algorithm is then given by (4.32) where the step size is chosen such that the cost function is minimized in the direction of the negative gradient, $-g_i$. The value of the step size α_i can be derived to be,

$$\alpha_i = \frac{2\sigma^2 g_i^* g_i}{g_i^* U_i^* U_i g_i + (g_i^* U_i^* U_i g_i)^c + 2\sigma^2 \mu g_i^* g_i} \tag{4.35}$$

where g_i^* denotes the conjugate transpose of g_i and g_i^c denotes element wise conjugate of g_i. The algorithm that solves (4.15) is shown in algorithm 6. Since the matrix H and its conjugate transpose H^* represent filtering operators, the gradient and the step size can be efficiently computed using linear space invariant filtering operations implemented in the frequency domain using fast Fourier transforms (FFT) [61].

4.5 Simulation Results

In this section, we compare CRITIR with the conventional approach to reconstruction in the limit of the near-field approximation. Under the near-field assumption, if the width of the smallest feature that is resolved is a, then the Fresnel number defined as, $N_F = a^2/(\lambda R)$, is assumed to be much larger than unity. If this assumption does not hold, it means that even if the detector is capable of resolving a feature with width a, it cannot be resolved by an algorithm that makes the near-field approximation. One such widely used method in the near-field is the Paganin's phase retrieval method that assumes that the object is composed of a single homogeneous material embedded in a material with negligible attenuation. Here, the phase is retrieved from the measured image at each view of the object using Paganin's phase retrieval [96] and the object is reconstructed by filter back-projection [24] of the retrieved phase images.

We generated a phantom ground-truth of size $N_u \times N_v \times N_w = 128 \times 512 \times 512$ with a voxel width of 0.175μm. The phantom consists of three latex spheres [96] with average values of $\beta = 2.7218 \times 10^{-08}$ for the absorption and $\delta = 2.6639 \times 10^{-06}$ for the refractive index decrement embedded in a material with negligible absorption and refraction. A slice of the phantom along the $u - v$ axes is shown in Fig. 4.2 (a). The measurement data is generated for a detector of size $N_u \times N_v = 32 \times 128$ by simulating the physics of X-ray propagation at an energy of 3keV, object to detector distance of $R = 400$mm, and $\gamma = 0.5898\mu\text{m}^{-1}$ in (4.4). The noise variance is equal to the simulated measurement value, where we assume an average photon count of 2500 for unattenuated projections. This leads to an average SNR of 35dB. A measured image along the $u - v$ axes is shown in Fig. 4.2 (b).

The absorption and refractive index decrement values of the object are reconstructed from measurements at 128 view angles. The reconstructions have a size of $N_u \times N_v \times N_w = 32 \times 128 \times 128$ and a voxel width of 0.7μm. In Fig. 4.2 (c-e), we compare the reconstructions of the refractive index decrement using CRITIR with the traditional approach of Paganin's phase retrieval followed by FBP reconstruction. A

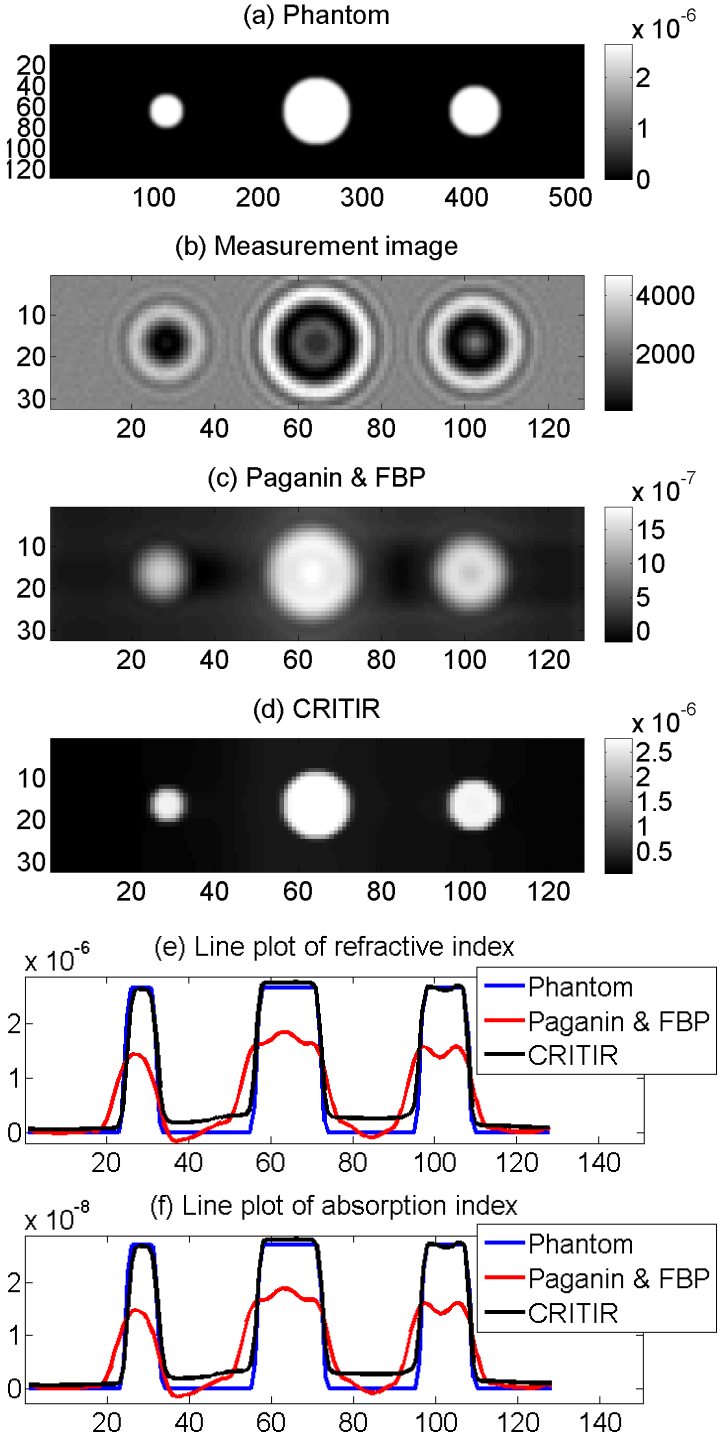


Fig. 4.2. (a) is the phantom ground-truth. (b) is the measured image at the detector. (c) is the FBP reconstruction of Paganin's phase retrieved images. (d) is the reconstruction using CRITIR. (e) and (f) are line plots of the refractive index decrement and absorption index respectively through the center of all three spheres. (a-d) correspond to the same 2D slice along the $u-$axis (vertical axis) and $v-$axis (horizontal axis). CRITIR accurately reconstructs the object while Paganin with FBP results in blurry edges and artifacts.

Table 4.1
Root mean square error comparison for different reconstruction methods

Description	Refractive index decrement, δ	Absorption index, β
Paganin & FBP	1.7850×10^{-7}	1.8100×10^{-9}
CRITIR	1.1705×10^{-7}	1.1938×10^{-9}

$u - v$ slice of the reconstruction using Paganin phase retrieval and FBP is shown in Fig. 4.2 (c). For CRITIR, we used the reconstruction obtained using Paganin and FBP as an initial estimate to avoid converging to a local minima due to the non-convex nature of (4.11). We set $\alpha = 98$ and $\sigma_-/\sigma_+ = 0.1$ in the prior model shown in (4.9). The regularization parameter σ_+ is adjusted to get the best subjective trade off between edge resolution and noise reduction. A $u-v$ slice of the reconstruction using CRITIR is shown in Fig. 4.2 (d). Fig. 4.2 (e, f) show a line plot of the refractive index decrement and absorption index through the center of the three spheres. For comparison purposes, the phantom line plot in Fig. 4.2 (e, f) is downsampled to the reconstruction resolution. We can see that the reconstruction of edges is much sharper and more accurate with CRITIR when compared to that with Paganin and FBP. CRITIR also dramatically improves the overall reconstruction quality by increasing resolution and reducing artifacts. The root mean square error between the reconstructions and the phantom for different algorithms is shown in table (4.1).

4.6 Conclusions

We presented a new algorithm called CRITIR that reconstructs the complex refractive index using a physics based model for X-ray propagation and a sparsity enforcing prior model for the object. Unlike traditional methods, CRITIR does not make simplifying approximations beyond the assumption of Fresnel diffraction. We generated simulated data beyond the near-field region and showed that CRITIR ac-

curately reconstructs the object while the conventional approach results in excessive blurring of edges and reconstruction artifacts. CRITIR will enable a new generation of experiments that was not possible until now due to various limitations imposed by the conventional reconstruction methods.

5. PLUG-AND-PLAY FORWARD MODELS FOR MODEL-BASED ITERATIVE RECONSTRUCTION

5.1 Introduction

In typical imaging systems, a sample is exposed to some form of radiation such as X-rays, electrons, or visible light and the radiation after interacting with the sample is measured by a detector. The sample morphology is then reconstructed from the measurements that contain information about the various structural and functional properties of the sample. The imaging system can be described by a forward model that expresses the measured data as a function of the unknown sample that must be reconstructed. The forward model is derived using the physics of imaging. In a MBIR framework, this forward model along with a prior model will enable us to solve the inverse problem that reconstructs the sample from the data. The prior model is used to enforce sparsity in reconstruction and also exploit prior knowledge about the sample. The complexity of the reconstruction algorithm heavily depends on the complexity of the forward model. In many computational imaging systems [11, 15], the forward model can be quite complex. Hence, solving the corresponding inverse problem using MBIR will be challenging and may not be computationally feasible.

In many cases, the function that implements the forward model can be expressed as a composition of several functions that represent simpler forward sub-models. To solve the MBIR reconstruction problem associated with such forward models, we present the framework of plug-and-play forward models. In plug-and-play, each forward sub-model is combined with a suitable prior sub-model and the associated inverse sub-problem is solved using the MBIR framework. Typically, the inverse sub-problems can be solved using practical and efficient algorithms. Then, we show that the original inverse problem can be solved using an iterative framework that repeatedly solves the

inverse sub-problems for different inputs using computationally efficient algorithms. Our solution also includes a combination step at each iteration that links the solutions from the different inverse sub-problems. The prior sub-models also contain information about the reconstructions from the other inverse sub-problems. This approach for solving inverse problems is the framework of plug-and-play forward models for MBIR. It is based on the variable splitting strategy and the theory of alternate direction method of multipliers [17].

The MBIR algorithm presented in chapter (3) that reconstructs the magnetization is based on the plug-and-play framework where the function that implements the forward model can be expressed as the composition of two simpler sub-functions. Similarly, the algorithm CRITIR that reconstructs the complex refractive index of a sample presented in chapter (4) is based on the plug-and-play framework where the function that implements the forward model can be expressed as the composition of three simpler sub-functions. Recently, the framework of plug-and-play prior models [106, 107] was introduced that enables us improve software modularity by independently combining inversion algorithms for system specific forward models with state-of-the-art denoising algorithms as prior models.

In this chapter, we will present the mathematical framework of plug-and-play forward models. First, we will formulate the plug-and-play framework of solving inverse problems when the forward model function can be expressed as the composition of several simpler sub-functions. Next, we present a proof based on mathematical induction that can be used to verify the validity of our approach.

5.2 Framework of Plug-and-Play Forward Models

In this section, we will solve the inverse problem associated with forward models that can be expressed as the composition of several sub-functions. The general form of the forward model is given by,

$$y = \left(f_0 \circ f_1 \circ \cdots f_{n-1}\right)(x) + w \tag{5.1}$$

where $f_i(.)\ \forall\, 0 \leq i \leq n-1$ are sub-functions, $\circ$ represents function composition, and w is the noise vector.

Then, we can formulate the reconstruction of x from y as the solution to the following optimization problem,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \left|\left| y - \left(f_0 \circ f_1 \circ \cdots f_{n-1} \right)(x) \right|\right|_{\Lambda}^2 + R(x) \right\} \tag{5.2}$$

where Λ is a diagonal matrix that models the variance of noise and $R(x)$ is a prior model for x. Using variable splitting, the above unconstrained optimization problem can be expressed as a constrained optimization problem of the form,

$$\begin{aligned} \left(\hat{x}, \hat{y}^{(1)}, \cdots, \hat{y}^{(n-1)}\right) = \underset{x}{\operatorname{argmin}} & \left\{ \frac{1}{2} \left|\left| y - f_0\left(y^{(1)}\right) \right|\right|_{\Lambda}^2 + R(x) \right\} \\ \text{s.t. } & y^{(1)} = f_1\left(y^{(2)}\right), y^{(2)} = f_2\left(y^{(3)}\right), \cdots, y^{(n-1)} = f_{n-1}(x) \end{aligned} \tag{5.3}$$

where $y^{(1)}, \cdots, y^{(n-1)}$ are auxiliary vectors.

Using ADMM, we can show that the above optimization problem can be solved using the INVERT function defined in algorithm 7. Note that algorithm 7 recursively calls the functions defined in algorithms 8 and 9. In algorithms 7, 8, and 9, we will assume $y^{(0)} = y$ and $y^{(n)} = x$. Also, $\left(u^{(0)}, \cdots, u^{(n)}\right)$ are the dual variable vectors and $(\mu_1, \cdots, \mu_{n-1})$ are the augmented Lagrangian parameters. Note that $\left(\hat{y}^{(0)}, \cdots, \hat{y}^{(n)}\right)$ represents the estimated value for the vectors $\left(y^{(0)}, \cdots, y^{(n)}\right)$. Thus, we can estimate $y^{(0)}$ (or x) from $y^{(n)}$ (or y) by calling the function defined in algorithm 7 as shown below,

$$\left(\hat{y}^{(0)}, \cdots, \hat{y}^{(n)}, u^{(0)}, \cdots, u^{(n)}\right) \leftarrow \text{INVERT}^{(n-1)}\left(\hat{y}^{(0)}, \cdots, \hat{y}^{(n)}, u^{(0)}, \cdots, u^{(n)}\right) \tag{5.4}$$

where the inputs to the function represent the initial estimates for all the various parameters.

5.3 Proof by Mathematical Induction

We can prove the validity of algorithms 7,8, and 9 using the theory of mathematical induction. First, we will prove the validity for $n = 3$. Then, we will assume the

Algorithm 7 INVERSION $(n-1)$

function $\textsc{Invert}^{(n-1)}(\hat{y}^{(0)}, \cdots, \hat{y}^{(n)}, u^{(0)}, \cdots, u^{(n)})$ ▷ where $\hat{y}^{(0)} = y$, $\hat{y}^{(n)} = \hat{x}$

  while not converged **do**

    INVERSION OF $f_{n-1}\left(y^{(n)}\right)$ ▷ where $y^{(n)} = x$

$$\hat{y}^{(n)} \leftarrow \underset{y^{(n)}}{\operatorname{argmin}} \left\{ \frac{\mu_{n-1}}{2} \left|\left| \hat{y}^{(n-1)} - f_{n-1}\left(y^{(n)}\right) + u^{(n-1)} \right|\right|^2 + R\left(y^{(n)}\right) \right\}$$

    INVERSE OF $\left(f_0 \circ \cdots \circ f_{n-2}\right)\left(y^{(n-1)}\right)$

$$\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(n-1)}, u^{(1)}, \cdots, u^{(n-1)}\right) \leftarrow \textsc{Invert}^{(n-2)}(\hat{y}^{(0)}, \cdots, \hat{y}^{(n)}, u^{(0)}, \cdots, u^{(n)})$$

    $u^{(n-1)} \leftarrow u^{(n-1)} + \left(\hat{y}^{(n-1)} - f_{n-1}(\hat{y}^{(n)})\right)$ ▷ $u^{(n-1)}$ is dual vector

  end while

  return $\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(n)}, u^{(1)}, \cdots, u^{(n)}\right)$

end function

Algorithm 8 INVERSION (i) $\forall\ 1 \leq i \leq n-2, n \geq 3$

function $\textsc{Invert}^{(i)}(\hat{y}^{(0)}, \cdots, \hat{y}^{(i+2)}, u^{(0)}, \cdots, u^{(i+2)})$ ▷ where $\hat{y}^{(0)} = y$, $\hat{y}^{(n)} = \hat{x}$

  while not converged **do**

    INVERSION OF $f_i\left(y^{(i+1)}\right)$

$$\hat{y}^{(i+1)} \leftarrow \underset{y^{(i+1)}}{\operatorname{argmin}} \left\{ \frac{\mu_i}{2} \left|\left| \hat{y}^{(i)} - f_i\left(y^{(i+1)}\right) + u^{(i)} \right|\right|^2 + \frac{\mu_{i+1}}{2} \left|\left| y^{(i+1)} - f_{i+1}\left(\hat{y}^{(i+2)}\right) + u^{(i+1)} \right|\right|^2 \right\}$$

    INVERSE OF $\left(f_0 \circ \cdots \circ f_{i-1}\right)\left(y^{(i)}\right)$

$$\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(i)}, u^{(1)}, \cdots, u^{(i)}\right) \leftarrow \textsc{Invert}^{(i-1)}(\hat{y}^{(0)}, \cdots, \hat{y}^{(i+1)}, u^{(0)}, \cdots, u^{(i+1)})$$

    $u^{(i)} \leftarrow u^{(i)} + \left(\hat{y}^{(i)} - f_i\left(\hat{y}^{(i+1)}\right)\right)$ ▷ $u^{(i)}$ is dual vector

  end while

  return $\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(i+1)}, u^{(1)}, \cdots, u^{(i+1)}\right)$

end function

algorithm is valid for an arbitrary value of $n = k$ and then show that the algorithm is also valid for $n = k + 1$.

Algorithm 9 INVERSION (0)

1: **function** INVERT$^{(0)}(\hat{y}^{(0)}, \hat{y}^{(1)}, \hat{y}^{(2)}, u^{(0)}, u^{(1)}, u^{(2)})$ ▷ where $\hat{y}^{(0)} = y$

2: INVERSION OF $f_0\left(y^{(1)}\right)$

$$\hat{y}^{(1)} \leftarrow \underset{y^{(1)}}{\operatorname{argmin}} \left\{ \left|\left|\hat{y}^{(0)} - f_0\left(y^{(1)}\right)\right|\right|_{\Lambda}^2 + \frac{\mu_1}{2} \left|\left|y^{(1)} - f_1(\hat{y}^{(2)}) + u^{(1)}\right|\right|^2 \right\} \tag{5.5}$$

3: **return** $\left(y^{(1)}, u^{(1)}\right)$

4: **end function**

5.3.1 Forward Model - Composition of Three Functions

We will solve the inverse problem associated with a forward model that can be expressed as a composition of $n = 3$ functions. Let $y^{(0)} = y$ be a vector array containing the measured data from the imaging system and $y^{(n)} = x$ be a vector array containing all the unknown parameters of the sample. Then, we can express the forward model as,

$$y = \left(f_0 \circ f_1 \circ f_2\right)(x) + w \tag{5.6}$$

where w is the noise vector and the forward model function $\left(f_0 \circ f_1 \circ f_2\right)(x) = f_0(f_1(f_2(x)))$ is the composition of three sub-functions, $f_0(.)$, $f_1(.)$ and $f_2(.)$.

Using MBIR, the reconstruction is given by solving the following optimization problem,

$$(\hat{x}) = \underset{x}{\operatorname{argmin}} \left\{ ||y - \left(f_0 \circ f_1 \circ f_2\right)(x)||_{\Lambda}^2 + R(x) \right\} \tag{5.7}$$

where Λ is a diagonal matrix such that each diagonal element Λ_{ii} gives the variance of the noise parameter w_i and $R(x)$ is a prior model for x. We will use variable splitting to express the above problem as a constrained optimization problem of the form,

$$\left(\hat{x}, \hat{y}^{(1)}, \hat{y}^{(2)}\right) = \underset{x}{\operatorname{argmin}} \left\{ \left|\left|y - f_0\left(y^{(1)}\right)\right|\right|_{\Lambda}^2 + R(x) \right\}$$
$$\text{s.t. } y^{(1)} = f_1(y^{(2)}) \text{ and } y^{(2)} = f_2(x) \tag{5.8}$$

Using ADMM, we can show that the algorithm that solves the above optimization problem is given in algorithm 10. Note that algorithm 10 also follows from algorithms 7, 8, and 9 by substituting $n = 3$ and calling the function below,

$$\left(\hat{y}^{(0)}, \hat{y}^{(1)}, \hat{y}^{(2)}, \hat{y}^{(3)}, u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}\right) \leftarrow \textsc{Invert}^{(2)} \left(\hat{y}^{(0)}, \hat{y}^{(1)}, \hat{y}^{(2)}, \hat{y}^{(3)}, u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)}\right), \tag{5.9}$$

where $y^{(0)} = y$ and $y^{(3)} = x$.

Algorithm 10 RECONSTRUCTION

while not converged **do**

  INVERSION OF FORWARD MODEL $f_2(x)$

$$\hat{x} \leftarrow \underset{x}{\operatorname{argmin}} \left\{ \frac{\mu_2}{2} \left|\left|\hat{y}^{(2)} - f_2(x) + u^{(2)}\right|\right|^2 + R(x) \right\} \tag{5.10}$$

  while not converged **do**

    INVERSION OF FORWARD MODEL $f_1\left(y^{(2)}\right)$

$$\hat{y}^{(2)} \leftarrow \underset{y^{(2)}}{\operatorname{argmin}} \left\{ \frac{\mu_1}{2} \left|\left|\hat{y}^{(1)} - f_1\left(y^{(2)}\right) + u^{(1)}\right|\right|^2 + \frac{\mu_2}{2} \left|\left|y^{(2)} - f_2(\hat{x}) + u^{(2)}\right|\right|^2 \right\} \tag{5.11}$$

    INVERSION OF FORWARD MODEL $f_0\left(y^{(1)}\right)$

$$\hat{y}^{(1)} \leftarrow \underset{y^{(1)}}{\operatorname{argmin}} \left\{ \left|\left|y - f_0\left(y^{(1)}\right)\right|\right|_\Lambda^2 + \frac{\mu_1}{2} \left|\left|y^{(1)} - f_1(\hat{y}^{(2)}) + u^{(1)}\right|\right|^2 \right\} \tag{5.12}$$

    $u^{(1)} \leftarrow u^{(1)} + \left(\hat{y}^{(1)} - f_1\left(\hat{y}^{(2)}\right)\right)$

  end while

  $u^{(2)} \leftarrow u^{(2)} + \left(\hat{y}^{(2)} - f_2(\hat{x})\right)$

end while

Note that (5.10), (5.11), and (5.12) are simpler optimization problems than the original problem in (5.7). In typical imaging systems [11], there exist efficient algorithms for solving (5.10), (5.11), and (5.12). Thus, the above algorithm represents the plug-and-play framework that solves the original reconstruction problem by iteratively solving three simpler reconstruction problems.

5.3.2 Forward Models - Composition of k Functions

The forward model for $n = k$ from equation (5.1) is given by,

$$y = \left(f_0 \circ f_1 \circ \cdots f_{k-1}\right)(x) + w \tag{5.13}$$

where $f_i(.)\ \forall\, 0 \leq i \leq k-1$ are sub-functions, $\circ$ denotes function composition, and w is the noise vector.

Then, we can formulate the reconstruction of x from y as the solution to the following optimization problem,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \left|\left| y - \left(f_0 \circ f_1 \circ \cdots f_{k-1}\right)(x) \right|\right|_{\Lambda}^2 + R(x) \right\} \tag{5.14}$$

where Λ is a diagonal matrix that models the variance of noise and $R(x)$ is a prior model for x.

We will assume that we can solve for $y^{(0)}$ (or x) from $y^{(k)}$ (or y) by calling the function below,

$$\left(\hat{y}^{(0)}, \cdots, \hat{y}^{(k)}, u^{(0)}, \cdots, u^{(k)}\right) \leftarrow \textsc{Invert}^{(k-1)}\left(\hat{y}^{(0)}, \cdots, \hat{y}^{(k)}, u^{(0)}, \cdots, u^{(k)}\right) \tag{5.15}$$

where the inputs to the function represent the initial estimates for the different parameters. Note that $y^{(0)} = y$ and $y^{(k)} = x$.

5.3.3 Forward Models - Composition of $k+1$ Functions

The forward model for $n = k+1$ from equation (5.1) is given by,

$$y = \left(f_0 \circ f_1 \circ \cdots f_{k}\right)(x) + w \tag{5.16}$$

where $f_i(.)\ \forall\, 0 \leq i \leq k$ are sub-functions, $\circ$ denotes function composition, and w is the noise vector.

Then, we can formulate the reconstruction of x from y as the solution to the following optimization problem,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \left|\left| y - \left(f_0 \circ f_1 \circ \cdots f_{k}\right)(x) \right|\right|_{\Lambda}^2 + R(x) \right\} \tag{5.17}$$

Algorithm 11 RECONSTRUCTION

1: **while** not converged **do**

2: INVERSION OF $f_k(x)$

$$\hat{x} \leftarrow \underset{x}{\operatorname{argmin}} \left\{ \frac{\mu_k}{2} \left|\left| y^{(k)} - f_k(x) + u^{(k)} \right|\right|^2 + R(x) \right\}$$

3: INVERSE OF $(f_0 \circ \cdots \circ f_{k-1})\left(y^{(k)}\right)$

$$\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(k)}, u^{(1)}, \cdots, u^{(k)}\right) \leftarrow \text{INVERT}^{(k-1)}(\hat{y}^{(0)}, \cdots, \hat{y}^{(k)}, \hat{x}, u^{(0)}, \cdots, u^{(k+1)})$$

4: $u^{(k)} \leftarrow u^{(k)} + \left(\hat{y}^{(k)} - f_k(\hat{x})\right)$ ▷ $u^{(k)}$ is dual vector

5: **end while**

6: **return** $\left(\hat{y}^{(1)}, \cdots, \hat{y}^{(k)}, \hat{x}, u^{(1)} \cdots, u^{(k+1)}\right)$

where Λ is a diagonal matrix that models the variance of noise and $R(x)$ is a prior model for x.

We can then formulate the above unconstrained optimization problem in the form of a constrained optimization problem as shown below,

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \left|\left| y - (f_0 \circ \cdots \circ f_{k-1})\left(y^{(k)}\right) \right|\right|_{\Lambda}^2 + R(x) \right\} \text{ s.t. } y^{(k)} = f_k(x). \quad (5.18)$$

Thus, it follows from ADMM that we can solve the above constrained optimization problem using algorithm 11. In step (3), we estimate $\hat{y}^{(k)}$ using the following new quadratic prior model for $y^{(k)}$,

$$R\left(y^{(k)}\right) = \frac{\mu_k}{2} \left|\left| y^{(k)} - f_k(x) + u^{(k)} \right|\right|^2 \quad (5.19)$$

We can see that algorithm 11 agrees with the algorithm 7 when $n = k + 1$ and with the new prior model $R\left(y^{(k)}\right)$ defined in (5.19). Hence, we have proved the validity of our plug-and-play forward model framework using mathematical induction. However, note that we haven't proved convergence here. A sufficient condition for convergence is if the functions $f_i \; \forall \, 0 \leq i \leq n-1$ in (5.1) are all convex.

LIST OF REFERENCES

LIST OF REFERENCES

[1] K. A. Mohan, S. Venkatakrishnan, L. Drummy, J. Simmons, D. Parkinson, and C. Bouman, "Model-based iterative reconstruction for synchrotron X-ray tomography," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, pp. 6909–6913, May 2014.

[2] K. A. Mohan, S. Venkatakrishnan, J. Gibbs, E. Gulsoy, X. Xiao, M. De Graef, P. Voorhees, and C. Bouman, "4D model-based iterative reconstruction from interlaced views," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, pp. 783–787, April 2015.

[3] K. A. Mohan, S. Venkatakrishnan, J. Gibbs, E. Gulsoy, X. Xiao, M. De Graef, P. Voorhees, and C. Bouman, "TIMBIR: A method for time-space reconstruction from interlaced views," *IEEE Trans. on Computational Imaging*, vol. 1, pp. 96–111, June 2015.

[4] J.-B. Thibault, K. D. Sauer, C. A. 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.

[5] P. J. Pickhardt, M. G. Lubner, D. H. Kim, J. Tang, J. A. Ruma, A. M. del Rio, and G.-H. Chen, "Abdominal CT with model-based iterative reconstruction (MBIR): Initial results of a prospective trial comparing ultralow-dose with standard-dose imaging," *American Journal of Roentgenology*, vol. 199, no. 6, pp. 1266–1274, 2012.

[6] Y. Yamada, M. Jinzaki, Y. Tanami, E. Shiomi, H. Sugiura, T. Abe, and S. Kuribayashi, "Model-based iterative reconstruction technique for ultralow-dose computed tomography of the lung: A pilot study," *Investigative Radiolo.*, vol. 47, pp. 482–489, 2012.

[7] D. B. Husarik, D. Marin, E. Samei, S. Richard, B. Chen, T. A. Jaffe, M. R. Bashir, and R. C. Nelson, "Radiation dose reduction in abdominal computed tomography during the late hepatic arterial phase using a model-based iterative reconstruction algorithm: How low can we go?," *Investigative Radiolo.*, vol. 47, no. 8, pp. 468–474, 2012.

[8] H. Scheffel, P. Stolzmann, C. L. Schlett, L.-C. Engel, G. P. Major, M. Károlyi, S. Do, P. Maurovich-Horvat, and U. Hoffmann, "Coronary artery plaques: cardiac CT with model-based and adaptive-statistical iterative reconstruction technique," *European J. Radiology*, vol. 81, pp. e363–e369, 2012.

[9] Z. Yu, J. Thibault, C. Bouman, K. Sauer, and J. Hsieh, "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," *IEEE Trans. on Image Processing*, vol. 20, pp. 161 –175, Jan 2011.

[10] Z. Ruoqiao, J. B. Thibault, C. A. Bouman, K. D. Sauer, and H. Jiang, "Model-based iterative reconstruction for dual-energy X-ray CT using a joint quadratic likelihood model," *IEEE Trans. on Medical Imaging*, vol. 33, no. 1, pp. 117–134, 2014.

[11] K. A. Mohan, X. Xiao, and C. A. Bouman, "Direct model-based tomographic reconstruction of the complex refractive index," in *Proc. of IEEE Int'l Conf. on Image Proc.*, pp. 1754–1758, Sept 2016.

[12] S. V. Venkatakrishnan, L. F. Drummy, M. A. Jackson, M. De Graef, J. Simmons, and C. A. Bouman, "A model based iterative reconstruction algorithm for high angle annular dark field-scanning transmission electron microscope (HAADF-STEM) tomography," *IEEE Trans. on Image Processing*, vol. 22, no. 11, pp. 4532–4544, 2013.

[13] S. Venkatakrishnan, L. Drummy, M. Jackson, M. De Graef, J. Simmons, and C. Bouman, "Model-based iterative reconstruction for bright field electron tomography," *IEEE Trans. on Computational Imaging*, no. 99, pp. 1–1, 2014.

[14] A. Burvall, U. Lundström, P. A. C. Takman, D. H. Larsson, and H. M. Hertz, "Phase retrieval in X-ray phase-contrast imaging suitable for tomography," *Opt. Express*, vol. 19, pp. 10359–10376, May 2011.

[15] C. Phatak, M. Beleggia, and M. D. Graef, "Vector field electron tomography of magnetic materials: Theoretical development," *Ultramicroscopy*, vol. 108, no. 6, pp. 503 – 513, 2008.

[16] R. P. Yu, M. J. Morgan, and D. M. Paganin, "Lorentz-electron vector tomography using two and three orthogonal tilt series," *Phys. Rev. A*, vol. 83, p. 023813, Feb 2011.

[17] 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.

[18] K. A. Mohan, K. C. Prabhat, C. Phatak, M. D. Graef, and C. A. Bouman, "Iterative reconstruction of the magnetization and charge density using vector field electron tomography," *Microscopy and Microanalysis*, vol. 22, no. S3, pp. 1686–1687, 2016.

[19] D. Rowenhorst and P. Voorhees, "Measurement of interfacial evolution in three dimensions," *Annual Review of Materials Research*, vol. 42, no. 1, pp. 105–124, 2012.

[20] L. K. Aagesen, J. L. Fife, E. M. Lauridsen, and P. W. Voorhees, "The evolution of interfacial morphology during coarsening: A comparison between 4D experiments and phase-field simulations," *Scripta Materialia*, vol. 64, no. 5, pp. 394–397, 2011.

[21] H. A. Bale, A. Haboub, A. A. MacDowell, J. R. Nasiatka, D. Y. Parkinson, B. N. Cox, D. B. Marshall, and R. O. Ritchie, "Real-time quantitative imaging of failure events in materials under load at temperatures above 1,600° C," *Nature Materials*, vol. 12, no. 1, 2013.

[22] J. Sijbers and A. Postnov, "Reduction of ring artifacts in high resolution micro-CT reconstructions," *Physics in Medicine and Biology*, vol. 49, no. 14, p. N247, 2004.

[23] T. Donath, *Quantitative X-ray Microtomography with Synchrotron Radiation*. PhD thesis, GKSS-Forschungszentrum, 2007.

[24] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Society for Industrial and Applied Mathematics, 2001.

[25] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE Press, Nov. 2009.

[26] D. Gottleib, B. Gustafsson, and P. Forssén, "On the direct Fourier method for computer tomography," *IEEE Trans. on Medical Imaging*, vol. 19, no. 3, pp. 223–232, 2000.

[27] F. Marone and M. Stampanoni, "Regridding reconstruction algorithm for real-time tomographic imaging," *Journal of Synchrotron Radiation*, vol. 19, pp. 1029–1037, Nov 2012.

[28] J. Walden, "Analysis of the direct Fourier method for computer tomography," *IEEE Trans. on Medical Imaging*, vol. 19, no. 3, pp. 211–222, 2000.

[29] P. J. La Riviere and X. Pan, "Noise properties of periodic interpolation methods with implications for few-view tomography," *IEEE Trans. on Nuclear Science*, vol. 46, no. 3, pp. 639–645, 1999.

[30] M. Bertram, J. Wiegert, D. Schafer, T. Aach, and G. Rose, "Directional view interpolation for compensation of sparse angular sampling in cone-beam CT," *IEEE Trans. on Medical Imaging*, vol. 28, no. 7, pp. 1011–1022, 2009.

[31] R. R. Galigekere, K. Wiesent, and D. W. Holdsworth, "Techniques to alleviate the effects of view aliasing artifacts in computed tomography," *Medical Physics*, vol. 26, no. 6, pp. 896–904, 1999.

[32] M. Davison, "The ill-conditioned nature of the limited angle tomography problem," *SIAM Journal on Applied Mathematics*, vol. 43, no. 2, pp. 428–448, 1983.

[33] P. Rattey and A. G. Lindgren, "Sampling the 2-D Radon transform," *IEEE Trans. on Acoustics Speech and Signal Processing*, vol. 29, no. 5, pp. 994–1002, 1981.

[34] A. Faridani and L. R. Erik, "High-resolution computed tomography from efficient sampling," *Inverse Problems*, vol. 16, no. 3, p. 635, 2000.

[35] A. M. Cormack, "Sampling the Radon transform with beams of finite width," *Physics in Medicine and Biology*, vol. 23, no. 6, p. 1141, 1978.

[36] N. Willis and Y. Bresler, "Optimal scan for time-varying tomography I: Theoretical analysis and fundamental limitations," *IEEE Trans. on Image Processing*, vol. 4, no. 5, pp. 642–653, 1995.

[37] N. Willis and Y. Bresler, "Optimal scan for time-varying tomography II: Efficient design and experimental validation," *IEEE Trans. on Image Processing*, vol. 4, no. 5, pp. 654–666, 1995.

[38] J. Miao, F. Förster, and O. Levi, "Equally sloped tomography with oversampling reconstruction," *Physical Review B*, vol. 72, no. 5, p. 052103, 2005.

[39] Z. Zheng and K. Mueller, "Identifying sets of favorable projections for few-view low-dose cone-beam CT scanning," in *11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2011.

[40] Y. Kaganovsky, D. Li, A. Holmgren, H. Jeon, K. P. MacCabe, D. G. Politte, J. A. O'Sullivan, L. Carin, and D. J. Brady, "Compressed sampling strategies for tomography," *J. Opt. Soc. Am. A*, vol. 31, pp. 1369–1394, Jul 2014.

[41] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 9, p. 2791, 2009.

[42] S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli, and G.-H. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 20, p. 5653, 2008.

[43] G.-H. Chen, P. Theriault-Lauzier, J. Tang, B. Nett, S. Leng, J. Zambelli, Z. Qi, N. Bevins, A. Raval, S. Reeder, and H. Rowley, "Time-resolved interventional cardiac C-arm cone-beam CT: An application of the PICCS algorithm," *IEEE Trans. on Medical Imaging*, vol. 31, pp. 907–923, April 2012.

[44] R. Leary, Z. Saghi, P. A. Midgley, and D. J. Holland, "Compressed sensing electron tomography," *Ultramicroscopy*, vol. 131, no. 0, pp. 70 – 91, 2013.

[45] G. M. D. Godaliyadda, G. T. Buzzard, and C. A. Bouman, "A model-based framework for fast dynamic image sampling," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, pp. 1822–1826, May 2014.

[46] S. J. Kisner, E. Haneda, C. A. Bouman, S. Skatter, M. Kourinny, and S. Bedford, "Limited view angle iterative CT reconstruction," *Proc. SPIE*, vol. 8296, pp. 82960F–82960F–9, 2012.

[47] H. Wu, A. Maier, R. Fahrig, and J. Hornegger, "Spatial-temporal total variation regularization (STTVR) for 4D-CT reconstruction," in *Proceedings of the SPIE*, vol. 8313, pp. 83133J–83133J–7. 10.1117/12.911162.

[48] L. Ritschl, S. Sawall, M. Knaup, A. Hess, and M. Kachelriess, "Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior," *Physics in Medicine and Biology*, vol. 57, no. 6, p. 1517, 2012.

[49] X. Jia, Y. Lou, B. Dong, Z. Tian, and S. Jiang, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization," in *Proceedings of the 13th International Conference on Medical Image Computing and Computer-assisted Intervention: Part I*, MICCAI'10, (Berlin, Heidelberg), pp. 143–150, Springer-Verlag, 2010.

[50] Z. Tian, X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "Low-dose 4DCT reconstruction via temporal nonlocal means," *Medical Physics*, vol. 38, no. 3, pp. 1359–1365, 2011.

[51] J. Hinkle, M. Szegedi, B. Wang, B. Salter, and S. Joshi, "4D CT image reconstruction with diffeomorphic motion model," *Medical Image Analysis*, vol. 16, no. 6, pp. 1307–1316, 2012.

[52] J. W. Gibbs, K. A. Mohan, E. B. Gulsoy, A. J. Shahani, X. Xiao, C. A. Bouman, M. De Graef, and P. W. Voorhees, "The three-dimensional morphology of growing dendrites," *Nature Scientific Reports*, vol. 5, Jul 2015.

[53] R. M. Cramblitt and J. P. Allebach, "Analysis of time-sequential sampling with a spatially hexagonal lattice," *J. Opt. Soc. Am.*, vol. 73, pp. 1510–1517, Nov 1983.

[54] D. Kim, S. Ramani, and J. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE Trans. on Medical Imaging*, vol. 34, pp. 167–178, Jan 2015.

[55] K. Sauer and C. Bouman, "Bayesian Estimation of Transmission Tomograms Using Segmentation Based Optimization," *IEEE Trans. on Nuclear Science*, vol. 39, pp. 1144–1152, 1992.

[56] C. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. on Image Processing*, vol. 5, pp. 480 –492, Mar. 1996.

[57] S. Saquib, C. Bouman, and K. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," *IEEE Trans. on Image Processing*, vol. 7, pp. 1029–1044, Jul 1998.

[58] S. Oh, A. Milstein, C. Bouman, and K. Webb, "A general framework for nonlinear multigrid inversion," *IEEE Trans. on Image Processing*, vol. 14, no. 1, pp. 125–140, 2005.

[59] J. Zheng, S. Saquib, K. Sauer, and C. Bouman, "Parallelizable Bayesian tomography algorithms with rapid, guaranteed convergence," *IEEE Trans. on Image Processing*, vol. 9, no. 10, pp. 1745–1759, 2000.

[60] H. Erdogan and J. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. on Medical Imaging*, vol. 18, no. 9, pp. 801–814, 1999.

[61] C. A. Bouman, *Model Based Image Processing*. 2013.

[62] E. K. Chong and S. H. Zak, *An introduction to optimization*, vol. 76. John Wiley & Sons, 2013.

[63] M. Jacobson and J. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. on Image Processing*, vol. 16, pp. 2411–2422, Oct 2007.

[64] M. Razaviyayn, M. Hong, and Z. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[65] M. Kamasak, C. Bouman, E. Morris, and K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans. on Medical Imaging*, vol. 24, pp. 636 –650, May 2005.

[66] J. W. Cahn and J. E. Hilliard, "Free energy of a nonuniform system. I. interfacial free energy," *The Journal of Chemical Physics*, vol. 28, no. 2, pp. 258–267, 1958.

[67] P. Jin, C. Bouman, and K. Sauer, "A method for simultaneous image reconstruction and beam hardening correction," in *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–5, Oct 2013.

[68] X. Wang, A. Sabne, S. Kisner, A. Raghunathan, C. Bouman, and S. Midkiff, "High performance model based image reconstruction," in *Proceedings of the 21st ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, PPoPP '16, (New York, NY, USA), pp. 2:1–2:12, ACM, 2016.

[69] P. Tartaj, M. del Puerto Morales, S. Veintemillas-Verdaguer, T. Gonzlez-Carreo, and C. J. Serna, "The preparation of magnetic nanoparticles for applications in biomedicine," *Journal of Physics D: Applied Physics*, vol. 36, no. 13, p. R182, 2003.

[70] M. Beleggia and Y. Zhu, "Electron-optical phase shift of magnetic nanoparticles I. basic concepts," *Philosophical Magazine*, vol. 83, no. 8, pp. 1045–1057, 2003.

[71] S. J. Norton, "Unique tomographic reconstruction of vector fields using boundary data," vol. 1, pp. 406–412, Jul 1992.

[72] K. B. Winters and D. Rouseff, "Tomographic reconstruction of stratified fluid flow," *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 40, pp. 26–33, Jan 1993.

[73] K. B. Winters and D. Rouseff, "A filtered backprojection method for the tomographic reconstruction of fluid vorticity," *Inverse Problems*, vol. 6, no. 4, p. L33, 1990.

[74] H. Braun and A. Hauck, "Tomographic reconstruction of vector fields," *IEEE Trans. on Signal Processing*, vol. 39, pp. 464–471, Feb 1991.

[75] L. Desbat and A. Wernsdorfer, "Direct algebraic reconstruction and optimal sampling in vector field tomography," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1798–1808, Aug 1995.

[76] S. Lade, D. Paganin, and M. Morgan, "3-D vector tomography of doppler-transformed fields by filtered-backprojection," *Optics Communications*, vol. 253, no. 46, pp. 382 – 391, 2005.

[77] S. Lade, D. Paganin, and M. Morgan, "Electron tomography of electromagnetic fields, potentials and sources," *Optics Communications*, vol. 253, no. 46, pp. 392 – 400, 2005.

[78] V. Stolojan, R. Dunin-Borkowski, M. Weyland, and P. Midgley, "Three-dimensional magnetic fields of nanoscale elements determined by electron-holographic tomography," in *CONFERENCE SERIES-INSTITUTE OF PHYSICS*, vol. 168, pp. 243–246, Philadelphia; Institute of Physics; 1999, 2001.

[79] E. Völkl, L. F. Allard, and D. C. Joy, *Introduction to Electron Holography*. Springer US, 1999.

[80] E. Humphrey, C. Phatak, A. Petford-Long, and M. D. Graef, "Separation of electrostatic and magnetic phase shifts using a modified transport-of-intensity equation," *Ultramicroscopy*, vol. 139, pp. 5 – 12, 2014.

[81] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1, pp. 185 – 203, 1981.

[82] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert, "Highly accurate optic flow computation with theoretically justified warping," *International Journal of Computer Vision*, vol. 67, no. 2, pp. 141–158, 2006.

[83] C. Phatak and D. Gursoy, "Iterative reconstruction of magnetic induction using lorentz transmission electron tomography," *Ultramicroscopy*, vol. 150, pp. 54 – 64, 2015.

[84] K. M. Holt, "Angular regularization of vector-valued signals," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, pp. 1105–1108, May 2011.

[85] D. Keren and A. Gotlib, "Denoising color images using regularization and "correlation terms"," *J. Vis. Comun. Image Represent.*, vol. 9, pp. 352–365, Dec. 1998.

[86] M. Arigovindan, M. Suhling, C. Jansen, P. Hunziker, and M. Unser, "Full motion and flow field recovery from echo doppler data," *IEEE Trans. on Medical Imaging*, vol. 26, pp. 31–45, Jan 2007.

[87] P. D. Tafti and M. Unser, "On regularized reconstruction of vector fields," *IEEE Trans. on Image Processing*, vol. 20, pp. 3163–3178, Nov 2011.

[88] X. Wang, K. A. Mohan, S. J. Kisner, C. Bouman, and S. Midkiff, "Fast voxel line update for time-space image reconstruction," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, pp. 1209–1213, March 2016.

[89] E. Humphrey and M. D. Graef, "On the computation of the magnetic phase shift for magnetic nano-particles of arbitrary shape using a spherical projection model," *Ultramicroscopy*, vol. 129, pp. 36 – 41, 2013.

[90] M. Beleggia, Y. Zhu, S. Tandon, and M. D. Graef, "Electron-optical phase shift of magnetic nanoparticles II. Polyhedral particles," *Philosophical Magazine*, vol. 83, no. 9, pp. 1143–1161, 2003.

[91] M. D. Graef, *Introduction to conventional transmission electron microscopy*. Cambridge University Press, 2003.

[92] B. S. He, H. Yang, and S. L. Wang, "Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 337–356, 2000.

[93] S. L. Wang and L. Z. Liao, "Decomposition method with a variable parameter for a class of monotone variational inequality problems," *Journal of Optimization Theory and Applications*, vol. 109, no. 2, pp. 415–429, 2001.

[94] A. V. Bronnikov, "Reconstruction formulas in phase-contrast tomography," *Optics Communications*, vol. 171, no. 46, pp. 239 – 244, 1999.

[95] T. E. Gureyev, T. J. Davis, A. Pogany, S. C. Mayo, and S. W. Wilkins, "Optical phase retrieval by use of first Born and Rytov-type approximations," *Appl. Opt.*, vol. 43, pp. 2418–2430, Apr 2004.

[96] D. Paganin, S. C. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins, "Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object," *Journal of Microscopy*, vol. 206, no. 1, pp. 33–40, 2002.

[97] X. Wu, H. Liu, and A. Yan, "X-ray phase-attenuation duality and phase retrieval," *Opt. Lett.*, vol. 30, pp. 379–381, Feb 2005.

[98] M. Beltran, D. Paganin, K. Uesugi, and M. Kitchen, "2D and 3D X-ray phase retrieval of multi-material objects using a single defocus distance," *Opt. Express*, vol. 18, pp. 6423–6436, Mar 2010.

[99] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Nonlinear phase retrieval using projection operator and iterative wavelet thresholding," *IEEE Signal Processing Letters*, vol. 19, pp. 579–582, Sept 2012.

[100] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Non-linear iterative phase retrieval based on frechet derivative," *Opt. Express*, vol. 19, pp. 22809–22819, Nov 2011.

[101] V. Davidoiu, B. Sixou, M. Langer, and F. Peyrin, "Nonlinear approaches for the single-distance phase retrieval problem involving regularizations with sparsity constraints," *Appl. Opt.*, vol. 52, pp. 3977–3986, Jun 2013.

[102] S. C. Mayo, P. R. Miller, S. W. Wilkins, T. J. Davis, D. Gao, T. E. Gureyev, D. Paganin, D. J. Parry, A. Pogany, and A. W. Stevenson, "Quantitative X-ray projection microscopy: phase-contrast and multi-spectral imaging," *Journal of Microscopy*, vol. 207, no. 2, pp. 79–96, 2002.

[103] J. Baruchel, J.-Y. Buffière, E. Maire, P. Merle, and G. Peix, *X-Ray Tomography in Material Science*. Hermes Science Publications, 2000.

[104] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate discrete distributions*, vol. 444. John Wiley & Sons, 2005.

[105] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[106] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *IEEE Global Conf. on Signal and Info. Proc.*, pp. 945–948, Dec 2013.

[107] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Trans. on Computational Imaging*, vol. PP, no. 99, pp. 1–1, 2016.

APPENDICES

A. THE OFFSET ERROR CONSTRAINT MATRIX, H

In this section, we briefly discuss our choice of the H matrix. We choose the matrix H such that we enforce a zero constraint on the weighted average of the offset errors, d_i, over overlapping rectangular patches. Since the number of constraints is much less than the number of parameters, d_i, the matrix H will have many more columns than rows. Furthermore, all the rows of H sum to the same value and all the columns sum to the same value.

The offset error constraint $Hd = 0$ is better expressed as a constraint on the two dimensional form of the offset error, $\tilde{d}_{i,j}$. Let $\tilde{d}_{i,j}$ be the offset error corresponding to the i^{th} row and j^{th} column of the detector. We then impose a zero constraint on the weighted average of the offset error, $\tilde{d}_{i,j}$, over overlapping[1] rectangular patches as shown in Fig. A.1. The patches are such that they overlap half-way along both the u-axis and v-axis and cover the entire detector plane.

Let $\tilde{h}_N(i)$ be a triangular window of the form

$$\tilde{h}_N(i) = \begin{cases} i & 1 \leq i \leq N \\ 2N - i + 1 & N + 1 \leq i \leq 2N \\ 0 & \text{otherwise.} \end{cases} \tag{A.1}$$

Then, the function $h(i,j) = \tilde{h}_P(i)\tilde{h}_Q(j)$ is used to appropriately weight the offset error terms, $\tilde{d}_{i,j}$, over the rectangular patches. Let $h^{(k,l)}(i,j)$ be the weighting function for the k^{th} patch along the u-axis and l^{th} patch along the v-axis obtained by appropriately shifting $h(i,j)$. Then, the constraint corresponding to the (k,l) patch is given by,

$$\sum_{i=0}^{M_r} \sum_{j=0}^{M_c} h^{(k,l)}(i,j)\tilde{d}_{i,j} = 0. \tag{A.2}$$

[1]If the patch extends outside the support of the detector plane along any dimension, then we wrap it around from the start of that dimension.

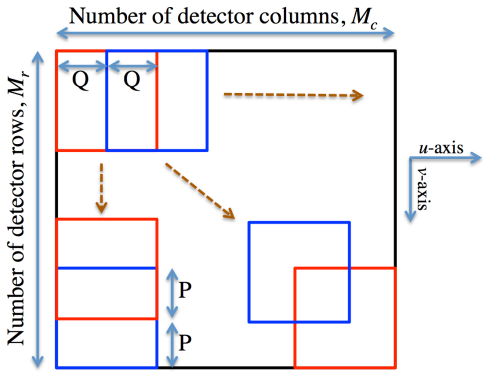


Fig. A.1. Diagram showing the overlapping rectangular patches over which the weighted sum of the offset error is zero. The patches are such that they overlap half-way along both the u-axis and v-axis and cover the entire detector plane.

where M_r and M_c are the total number of rows and columns of the detector respectively. In our application, we choose P such that it is closest to $\sqrt{M_r}$ and is a factor of M_r (similarly, we choose Q depending on M_c). Thus, the number of constraints is approximately equal to $\sqrt{M_r M_c}$. Each constraint indexed by (k, l) in (A.2) corresponds to one of the rows of H. The offset error d is the vector form of $\tilde{d}_{i,j}$ and the elements of H correspond to the weights, $h^{(k,l)}(i, j)$, used in (A.2). Thus, the matrix H has a size of $\sqrt{M_r M_c} \times M_r M_c$.

B. DISCRETE IMPLEMENTATION OF GREEN'S FUNCTION CONVOLUTION

A discrete approximation to the Green's function in (3.2), (3.3), and (3.4) is obtained by substituting $(u, v, k) = (i\Delta, j\Delta, k\Delta)$ where Δ is the sampling width (voxel width). However, the sampled value at $(i = 0, j = 0, k = 0)$ is undefined. We can approximate the value at $(0, 0, 0)$ by integrating (3.2), (3.3), and (3.4) over a voxel cube of width Δ centered at $(0, 0, 0)$. Thus, an approximation to the discrete value at $(0, 0, 0)$ for $h_C^{(u)}(u, v, w)$ is given by,

$$
\begin{aligned}
&\int_{w=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \int_{v=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \int_{u=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} h_C^{(u)}(u, v, w) du\, dv\, dw \\
&= \int_{w=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \int_{v=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \left[\int_{u=-\frac{\Delta}{2}}^{0} \frac{u}{\left|u^2+v^2+w^2\right|^{3/2}} \right. \\
&\qquad\qquad \left. + \int_{u=0}^{\frac{\Delta}{2}} \frac{u}{\left|u^2+v^2+w^2\right|^{3/2}} \right] du\, dv\, dw \\
&= \int_{w=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \int_{v=-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \left[\int_{u=-\frac{\Delta}{2}}^{0} \frac{u}{\left|u^2+v^2+w^2\right|^{3/2}} \right. \\
&\qquad\qquad \left. - \int_{u=-\frac{\Delta}{2}}^{0} \frac{u}{\left|u^2+v^2+w^2\right|^{3/2}} \right] du\, dv\, dw \\
&= 0
\end{aligned}
$$

Similarly, we can show that the corresponding integrals for $h_C^{(v)}(u,v,w)$ and $h_C^{(w)}(u,v,w)$ evaluate to zero. Thus, a discrete form of the Green's function in (3.2), (3.3), and (3.4) is given by,

$$\tilde{h}_D^{(u)}[i,j,k] = W[i,j,k]\frac{i}{|i^2+j^2+k^2|^{3/2}\Delta^2} \tag{B.1}$$

$$\tilde{h}_D^{(v)}[i,j,k] = W[i,j,k]\frac{j}{|i^2+j^2+k^2|^{3/2}\Delta^2} \tag{B.2}$$

$$\tilde{h}_D^{(w)}[i,j,k] = W[i,j,k]\frac{k}{|i^2+j^2+k^2|^{3/2}\Delta^2} \tag{B.3}$$

where $\tilde{h}_D^{(u)}[0,0,0] = \tilde{h}_D^{(v)}[0,0,0] = \tilde{h}_D^{(w)}[0,0,0] = 0$ and $W[i,j,k] = W_u[i]W_v[j]W_w[k]$ is a 3D window function where $W_u[i]$, $W_v[j]$, and $W_w[k]$ are 1D Hamming windows. The window function is necessary to prevent aliasing artifacts in the reconstruction. Since discrete convolution is implemented using discrete summation rather than continuous domain integration, the differential terms du, dv, and dw in (3.1) is accounted for by the sample width, Δ. Thus, Δ is absorbed into the expressions $\tilde{h}_D^{(u)}[i,j,k]$, $\tilde{h}_D^{(v)}[i,j,k]$, and $\tilde{h}_D^{(w)}[i,j,k]$. Thus, the modified point spread functions that can be used to compute the magnetic vector potential using discrete convolution is given by,

$$h_D^{(u)}[i,j,k] = W[i,j,k]\frac{i\Delta}{|i^2+j^2+k^2|^{3/2}} \tag{B.4}$$

$$h_D^{(v)}[i,j,k] = W[i,j,k]\frac{j\Delta}{|i^2+j^2+k^2|^{3/2}} \tag{B.5}$$

$$h_D^{(w)}[i,j,k] = W[i,j,k]\frac{k\Delta}{|i^2+j^2+k^2|^{3/2}} \tag{B.6}$$

VITA

VITA

K. Aditya Mohan is a PhD candidate at the department of Electrical and Computer Engineering, Purdue University, IN, USA. He received his B.Tech. degree in Electronics and Communication Engineering from National Institute of Technology Karnataka, Surathkal, India, in 2010 and M.S. degree in Electrical and Computer Engineering from Purdue University, West Lafayette, IN, in 2014. From 2010 to 2011, he worked as an ASIC design engineer at Nvidia, Bengaluru, India.

His research interests are in statistical signal processing, inverse problems, computational imaging, computed tomography, and machine learning. He won the student best paper award at the IEEE International Conference on Image Processing (ICIP) 2016. He was one of the finalists for the Werner Meyer-Ilse Memorial Award at the X-Ray Microscopy (XRM) Conference 2016. He was also an invited speaker at the Microscopy and Microanalysis (M&M) Conference 2016 and the ALS User Meeting 2016.