

ABSTRACT

Robust and Efficient Methods for Proton Computed Tomography

Paniz Karbasi, Ph.D.

Mentor: Keith Evan Schubert, Ph.D.

Proton computed tomography (pCT) is a recent promising imaging modality with the goal of generating accurate 3D maps of relative stopping power (RSP) with respect to water. Since the early developments of this imaging technique in 1970's, there have been significant improvements regarding the reconstruction of accurate RSP which makes pCT a reliable alternative to X-ray CT for planning proton therapy treatments. There are several conditions in pCT that can negatively affect the accuracy of pCT images. The goal of this dissertation is developing efficient image reconstruction methods generating accurate RSP values under both normal and critical conditions.

Robust and Efficient Methods for Proton Computed Tomography

by

Paniz Karbasi, B.S., M.S.

A Dissertation

Approved by the Department of Electrical and Computer Engineering

Kwang Y. Lee, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

Keith Evan Schubert, Ph.D., Chairperson

Reinhard Schulte, M.D.

Scott Koziol, Ph.D.

Enrique Blair, Ph.D.

Lorin S. Matthews, Ph.D.

Accepted by the Graduate School
August 2018

J. Larry Lyon, Ph.D., Dean

Copyright © 2018 by Paniz Karbasi

All rights reserved

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES	x
ACKNOWLEDGMENTS	xiv
DEDICATION	xvi
1 Introduction	1
1.1 Reconstruction Steps	2
1.2 Contributions	5
1.3 Outline	5
2 Current General and Robust Formulations	6
2.1 Current Parallel Iterative Image Reconstruction Methods	6
3 Incorporating Robustness in DROP	10
3.1 Uncertainties in MLP	11
3.2 Robust DROP	12
3.2.1 Experiments and Results of Robust DROP vs. DROP	13
3.3 Noise Removal by Perturbing the Diagonals of the Path Matrix	19
3.4 DROP with Additive Updates	21
4 Sparse Robust Estimation	23
4.1 Proof of the Compatibility of Equation 4.5 with the General Robust Equation	25
4.2 Convergence Analysis of FSAIS	28

4.3	Experiments and Results	33
4.4	Convergence Analysis of FSAIS and DROP on CTP404 Data	34
4.4.1	Removing Protons from Angle Intervals During Reconstruction of CTP404	44
4.4.2	Removing Angle Intervals During Data Read	63
4.5	Convergence Analysis of FSAIS and DROP on Pediatric Head Phantom	74
5	GPU-based Image Reconstruction	88
5.1	Distributed GPU-based Image Reconstruction without Data Transfer	89
5.2	Distributed GPU-based Image Reconstruction with Data Transfers . .	96
5.2.1	Implementation of Algorithm 2	98
6	Conclusion	103
	APPENDICES	105
	APPENDIX A Details of pCT Reconstruction Timing	106
	BIBLIOGRAPHY	107

LIST OF FIGURES

1.1	Schematic drawing of pCT scanner developed at Loma Linda University	3
2.1	Schematic drawing of DROP	8
3.1	CTP404 phantom; composition and geometry of the materials.	11
3.2	Reconstructed images of DROP and Robust DROP	15
3.3	Reconstructed images of DROP and Robust DROP, using MLP with Gaussian noise ($\eta = 1.0$)	16
3.4	Reconstructed images of DROP and Robust DROP, using MLP with Gaussian noise ($\eta = 2.0$)	17
3.5	Reconstructed image of DROP and Robust DROP while removing 120K protons	18
3.6	Vertical line segment on middle of the image	20
4.1	DROP convergence analysis on experimental CTP404	36
4.2	Experimental CTP404 pCT reconstruction by DROP	37
4.3	Convergence analysis of FSAIS on CTP404 with TVS during 1st and 2nd iterations	40
4.4	Experimental CTP404 pCT reconstruction by FSAIS with TVS during during 1st and 2nd iterations	41
4.5	Convergence analysis of FSAIS on CTP404 with TVS during 1st iteration	42
4.6	Experimental CTP404 pCT reconstruction by FSAIS with TVS during 1st iteration	43
4.7	% Discrepancy of material inserts in CTP404 while removing protons during the image reconstruction	45
4.8	DROP convergence analysis after removing protons within $0.0^\circ - 10.0^\circ$ range during reconstruction of exp. CTP404.	46

4.9	DROP convergence analysis after removing protons within a range of $0.0^\circ - 20.0^\circ$ during reconstruction of exp. CTP404.	47
4.10	DROP convergence analysis after removing protons within $0.0^\circ - 30.0^\circ$ range during reconstruction of exp. CTP404.	48
4.11	DROP convergence analysis after removing protons within $0.0^\circ - 40.0^\circ$ range during reconstruction of exp. CTP404.	49
4.12	DROP convergence analysis after removing protons within $0.0^\circ - 50.0^\circ$ range during reconstruction of exp. CTP404.	50
4.13	DROP convergence analysis after removing protons within $0.0^\circ - 60.0^\circ$ range during reconstruction of exp. CTP404.	51
4.14	Experimental CTP404 pCT reconstruction by DROP while removing protons during image reconstruction	52
4.15	FSAIS convergence analysis after removing protons within $0.0^\circ - 10.0^\circ$ range during reconstruction of exp. CTP404.	53
4.16	FSAIS convergence analysis after removing protons within $0.0^\circ - 20.0^\circ$ range during reconstruction of exp. CTP404.	54
4.17	FSAIS convergence analysis after removing protons within $0.0^\circ - 30.0^\circ$ range during reconstruction of exp. CTP404.	55
4.18	FSAIS convergence analysis after removing protons within $0.0^\circ - 40.0^\circ$ range during reconstruction of exp. CTP404.	56
4.19	FSAIS convergence analysis after removing protons within $0.0^\circ - 50.0^\circ$ range during reconstruction of exp. CTP404.	57
4.20	FSAIS convergence analysis after removing protons within $0.0^\circ - 60.0^\circ$ range during reconstruction of exp. CTP404.	58
4.21	Experimental CTP404 pCT reconstruction by FSAIS while removing protons during image reconstruction	59
4.22	% Discrepancy of Teflon and Delrin while removing protons from angle intervals with length 6° during the image reconstruction	60
4.23	% Discrepancy of Acrylic and Polystyrene while removing protons from angle intervals with length 6° during performing the iterative solver	61
4.24	% Discrepancy of LDPE and PMP while removing protons from angle intervals with length 6° during performing the iterative solver	62

4.25	Filtered sinogram while protons were removed within a range of $0.0^\circ - 30.0^\circ$	63
4.26	% Discrepancy of material inserts in CTP404 while removing protons during data read	65
4.27	Experimental CTP404 pCT reconstruction by DROP while removing protons during data read	66
4.28	Experimental CTP404 pCT reconstruction by FSAIS while removing protons during data read	67
4.29	% Discrepancy in the RSP for material inserts in CTP404 while removing protons during data read and using a median filter for FBP	68
4.30	Experimental CTP404 pCT reconstruction by DROP while removing protons during data read and using a median filter for FBP	69
4.31	Experimental CTP404 pCT reconstruction by FSAIS while removing protons during data read and using a median filter for FBP	70
4.32	% Discrepancy of Teflon and Delrin while removing protons from angle intervals with length 6° during data read	71
4.33	% Discrepancy of Acrylic and Polystyrene while removing protons from angle intervals with length 6° during data read	72
4.34	% Discrepancy of LDPE and PMP while removing protons from angle intervals with length 6° during data read	73
4.35	Representative pCT images of pediatric head reconstructed by DROP and FSAIS	76
4.36	% Discrepancy of pediatric head while removing protons during the image reconstruction	77
4.37	% Discrepancy of soft and brain tissues while removing protons from angle intervals with width 6° during data read.	79
4.38	% Discrepancy of pediatric head materials while removing protons from angle intervals with width 6° during the image reconstruction.	80
4.39	pCT images of the pediatric head reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with protons removed within a range of $0.0^\circ - 5.0^\circ$ during data read.	82

4.40	Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 5.0^\circ$ prior to generating the initial iterate w/ a median filter of radius 3 on the final iteration.	83
4.41	Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 10.0^\circ$ during data read.	84
4.42	Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 10.0^\circ$ prior to generating the initial iterate w/ a median filter of radius 3 used at the final iteration	85
4.43	Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with protons removed every 4.0° during the image reconstruction. Removing about one fourth of protons reduces the sharpness of edged in images generated by DROP, but FSAIS still generates images with sharp edges.	87
5.1	Schematic drawing of one cycle of Algorithm 2	99
5.2	Schematic drawing of left-to-right vs. top-to-bottom implementation of 2.	100
5.3	pCT reconstruction of the upper part of the exp. pediatric head . . .	102
5.4	Schematic drawing of splitting pCT tasks on 4 GPUs	102

LIST OF TABLES

2.1	Definition of frequently used mathematical terms.	9
3.1	% Discrepancy of RSP values of simulated CTP404 generated by FBP.	12
3.2	Reconstructed RSP values of pCT images in Figure 3.2 with no errors in matrix A	15
3.3	Reconstructed RSPs of DROP and Robust DROP, using MLP with Gaussian noise ($\eta = 1.0$)	16
3.4	Reconstructed RSPs of DROP and Robust DROP, using MLP with Gaussian noise ($\eta = 2.0$)	17
3.5	RSP values of reconstructed images in where 120000 protons were randomly removed.	19
3.6	Reconstructed RSP valued generated by DROP and DROP with additive updates.	22
4.1	Specifications of datasets used in this chapter.	33
4.2	Reconstructed RSP values for the materials of experimental CTP404 generated by DROP	37
4.3	RSP statistics of materials at the initial iterate of CTP404.	38
4.4	Elements of Ψ^{-1} for experimental CTP404.	38
4.5	Reconstructed RSP values for the materials of experimental CTP404 generated by FSAIS while using TVS during the first and second iterations.	41
4.6	Reconstructed RSP values for the materials of experimental CTP404 generated by FSAIS while using TVS during the first iteration.	43
4.7	DROP generated RSP values after removing protons within a range of $0.0^\circ - 10.0^\circ$ during reconstruction of exp. CTP404.	46
4.8	DROP generated RSP values after removing protons within $0.0^\circ - 20.0^\circ$ range during reconstruction of exp. CTP404.	47

4.9	DROP generated RSP values after removing protons within $0.0^\circ - 30.0^\circ$ range during reconstruction of exp. CTP404.	48
4.10	DROP generated RSP values after removing protons within $0.0^\circ - 40.0^\circ$ range during reconstruction of exp. CTP404.	49
4.11	DROP generated RSP values after removing protons within $0.0^\circ - 50.0^\circ$ range during reconstruction of exp. CTP404.	50
4.12	DROP generated RSP values after removing protons within $0.0^\circ - 60.0^\circ$ range during reconstruction of exp. CTP404.	51
4.13	FSAIS generated RSP values after removing protons within $0.0^\circ - 10.0^\circ$ range during reconstruction of exp. CTP404.	53
4.14	FSAIS generated RSP values after removing protons within $0.0^\circ - 20.0^\circ$ range during reconstruction of exp. CTP404.	54
4.15	FSAIS generated RSP values after removing protons within $0.0^\circ - 30.0^\circ$ range during reconstruction of exp. CTP404.	55
4.16	FSAIS generated RSP values after removing protons within $0.0^\circ - 40.0^\circ$ range during reconstruction of exp. CTP404.	56
4.17	FSAIS generated RSP values after removing protons within $0.0^\circ - 50.0^\circ$ range during reconstruction of exp. CTP404.	57
4.18	FSAIS generated RSP values after removing protons within $0.0^\circ - 60.0^\circ$ range during reconstruction of exp. CTP404.	58
4.19	RSP statistics of materials at the initial iterate of the experimental pediatric head phantom.	74
4.20	Elements of Ψ^{-1} for the pediatric head phantom.	75
4.21	RSP values of experimental pediatric head phantom reconstructed by DROP and FSAIS.	76
4.22	Reconstructed RSP values of pediatric head phantom by DROP and FSAIS with protons were removed within a range of $0.0^\circ - 5.0^\circ$ during data read	82
4.23	RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 5.0^\circ$ during data read w/ a median filter of radius 3 on the final iteration.	83

4.24	RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 10.0^\circ$ during data read	84
4.25	RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 10.0^\circ$ range during data read w/ a median filter of radius 3 used at the final iteration	85
4.26	RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed every 4.0° during the image reconstruction. Based on the results, both algorithms generated accurate RSP values for soft tissue, brain tissue, and Trabecular bone, while FSAIS generated a much better RSP value for Enamel.	87
5.1	Distribution of protons along z axis	90
5.2	Number of protons assigned to each GPU after data split	91
5.3	Performance comparison of pCT image reconstruction on the simulated CTP404.	92
5.4	Reconstructed RSP values of central slice of the simulated CTP404 using 1 and 2 GPUs.	93
5.5	Performance analysis of pCT image reconstruction on the experimental CTP404 using FSAIS as the iterative solver. Total reconstruction runtime included six iterations of FSAIS.	94
5.6	Reconstructed RSP values of exp. CTP404 using Algorithm 1	94
5.7	Performance analysis of pCT image reconstruction of the experimental pediatric head phantom using Algorithm 1	95
5.8	Reconstructed RSP values of the experimental pediatric head phantom by FSAIS based on Algorithm 1	96
5.9	Performance analysis of pCT image reconstruction on the experimental pediatric head phantom using Algorithm 2	99
5.10	Reconstructed RSP values of the experimental pediatric head by DROP with additive updates on two K40 GPU devices. Reconstructed RSP values of all materials except the Enamel are within a range of 1% of the predicted RSPs.	101
5.11	pCT reconstruction runtime of the entire experimental pediatric head	102

A.1	Runtime of pCT image reconstruction tasks prior to performing the iterative solver on the experimental CTP404 phantom.	106
A.2	Runtime of pCT image reconstruction tasks prior to performing the iterative solver on the experimental pediatric head phantom.	106

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my advisor, Dr. Keith Evan Schubert. I learned a lot from Dr. Schubert that can not be expressed in a few words. Not only did I learn about his incredible and unique research ideas, but I tried to learn from him “how to think” when doing research. Dr. Schubert is the advisor every single Ph.D. student would ask from God. He has always been supportive during these past four years both in my research and life. He has always given me the opportunity and confidence to pursue various research projects. Our research meetings were full of great scientific discussions, math problems to warm up the mind, and original life lessons. I can never thank my advisor enough but as the last sentence I would like to thank him for first accepting me in his research group and then giving me great advice all the way through my Ph.D. research. I would also like to thank Dr. Reinhard Schulte from Loma Linda Medical School for his valuable comments regarding my research. Dr. Schulte provided many great suggestions since the beginning of my research. It absolutely was a great honor for me to be part of the research collaboration under his supervision. I also want to thank Dr. Robert Marks, Dr. Scott Koziol, Dr. Enrique Blair, and Dr. Lorin Matthews for their valuable feedback and comments. Someone who I will never forget for her professional and emotional support and kindness is Ms. Minnie Simcik, the graduate program coordinator of the electrical and computer engineering department at Baylor. A very very big “thank you” to her. I would also like to thank my peers, Ms. Hanh Nguyen, Mr. Ritchie Cai, Mr. Blake Schultze, and Mr. Gordon Ledford for supporting me in difficult times and happy times. Hanh is a wonderful, smart, and energetic person with whom I always enjoyed talking. She has a great soul. Ritchie is a very kind and knowledgeable person, but he is also very strict. Thanks to him, I learned to

be tougher both in life and research. Blake is a talented researcher that never gave up sending super long text messages and emails including tens of hundreds of words regarding our research projects, and of course I never gave up reading those to the end. Gordon is a very nice and smart engineer, who has lots of great experiences and “words of wisdom” to share. I also acknowledge Mr. Brian Sittton for providing technical support for the use of the Kodiak Cluster at Baylor. I also want to thank Mrs. Kym Schubert, my advisor’s wonderful wife who invited my husband and me along with the other members of our research lab to their house several times during these past four years. Being far from my family had always been hard, but being with Dr. Schubert’s family was a significant relief for me. Finally, I would like to express my sincere gratitude to my wonderful husband’s parents, Mrs. Ladan Khodayar and Dr. Bijan Shakerin for their pure love. My dearest Ladan and Bijan, I love you so much. The last but not least, I sincerely thank my beloved brother, Kamyar, who encouraged me a lot and reminded me of my strengths during my difficult times. Kamyar, I am so proud of being the sister of such a sharp, kind and hard-working person.

To my best friend, my love, my dearest Bahram

To my beloved parents, Dr. Shiva Shivaei and Dr. Hamid Karbasi

To my grandparents, my eternal love, Nani and Pedar

CHAPTER ONE

Introduction

Proton computed tomography (pCT) is a recent promising imaging modality with the goal of generating accurate 3D maps of relative stopping power (RSP) with respect to water. From the early developments of this imaging technique in [1–4], to the more recent development of a preclinical pCT scanner in [5, 6], there have been many improvements regarding the reconstruction of accurate RSPs. These not only predict more reliable treatment plans, but also can be used as a validation technique prior to the treatment plan as a reliable alternative to the X-ray CT due to directly generating accurate RSP maps, exposure of tissues to smaller doses of radiation, and less uncertainty in Bragg peak location [7, 8].

One of the major concerns in therapies which use protons or other heavy charged particles, e.g., carbon ions is the range of uncertainty which is related to the changes in the tissue RSP. Even a change of a few percent in tissue RSP can result in a range of error above the desired limit of 1-2 mm, which forces the planner of proton therapy to increase margins around the target. This causes the unwanted exposure of normal tissues to high doses. Therefore, reconstruction of accurate RSP maps is one of the key goals in implementation of a pCT system. The method is based on tracking the individual protons to estimate the most likely path (MLP) [9, 10], and measuring the energy loss of each proton, and converting it to water-equivalent path-length (WEPL) [11]. Thus, one can form a linear system of equations $Ax = b$ where A is the intersection lengths of tracked protons through individual object voxels, x is the RSP maps of those voxels and b is the WEPL measurements. Solving the problem for finding the 3D maps of RSP is based on the parallel versions of an algebraic reconstruction technique (ART) [12] using projections onto convex sets. The main

advantage of parallel image reconstruction methods [13,14] is simultaneous execution over one or more graphics processing units (GPU) leveraging tens of hundreds of CUDA threads concurrently running on CUDA cores.

There are several factors that can negatively affect the results of image reconstruction causing slow convergence and/or RSP values with high discrepancy from the predicted true RSPs. These factors include: poor estimate of the initial iterate used by the iterative solver; errors in protons path estimation either errors associated with MLP calculation or the intersection length of a proton and the voxels along its path; and lack of enough proton's during a scan due to hardware issues or aiming to perform low-dose scan with a small number of protons. A standard solution to uncertain systems is to use a regularized solution,

$$x(\Psi) = (A^T A + \Psi I)^{-1} A^T b, \quad (1.1)$$

in which a diagonal matrix is added to the $A^T A$ term to make the inversion more accurate by perturbing the singular values of the matrix A . In pCT, we deal with a very large and sparse system of equations which is solved iteratively. The goal of this dissertation is developing real-time GPU based sparse compatible robust iterative solvers addressing the uncertainties in pCT and generating accurate 3D maps of RSP measurements.

1.1 Reconstruction Steps

Reconstructing accurate RSP measurements within clinically recommended time frames (i.e. under 5 minutes) is an important factor that needs to be met in developing a pCT system. During this development, there are several steps that need to be performed which can lead to performance and image quality improvements in pCT. These steps can be summarized as following [15]:

- (1) Removing outliers
- (2) Hull detection

- (3) Data cuts
- (4) Generating initial iterate
- (5) Finding the endpoints of MLP
- (6) Performing the iterative image reconstruction

Before describing the reconstructing steps in greater detail, we overview the pCT data acquisition. A prototype pCT scanner was developed at Loma Linda University consisting of a silicon-based particle tracking system and a five-stage scintillating energy detector [16, 17]. Design of the pCT data acquisition is described in [18], which is capable of measuring one to two million proton tracks per second. Each proton track includes information of the locations where a proton hits silicon-based tracking planes and the WEPL measurements of each protons are recorded. A software platform was developed to characterize the performance of the proposed pCT scanner at Loma Linda University [19], which is also illustrated in Figure 1.1.

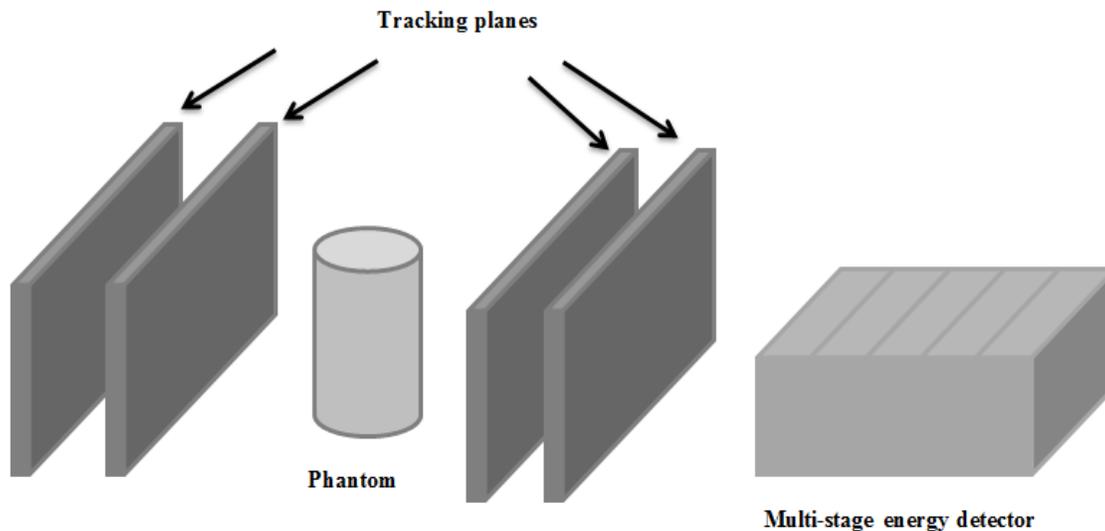


Figure 1.1: Schematic drawing of pCT scanner developed at Loma Linda University

After data acquisition, proton tracks are stored in binary file format which are stored on the solid-state drive (SSD) of the computing node. During the data

read process from SSD, a reconstruction volume is defined based on prior knowledge of the object to be reconstructed, which lets us identify and remove protons missing from the reconstruction. Next, hull detection [20,21] is performed to identify protons that missed the object, leading to identification of regions of space which do not contain any portion of the object being scanned. Hull-detection is an important step in pCT which is used in steps (4) and (5) of the pCT work-flow. A statistical analysis is performed during the hull detection to identify and remove protons with irregular relative angle or WEPL measurement due to nuclear scattering or multiple proton pile up events. The remaining protons are used to construct the sinogram representation of the data, which is then used as input to filtered back projection (FBP) as outlined in [20].

The initial iterate is generated based on a hybrid method using the convex hull and FBP image, where voxels in the FBP image that do not belong to the convex hull are removed from the initial iterate. Next, we proceed with collecting the endpoints of protons that enter and exit the image hull, which will be used in MLP calculation, and finally the iterative image reconstruction is performed on the remaining protons. In order to improve the image quality while performing the iterative solvers, one can implement the superiorization method (SM) [22] which reduces the value of a target function while seeking constraints-compatibility and has been shown to reduce noise and generate higher quality of pCT images when combined with DROP [23,24].

Implementing the reconstruction steps on CUDA enabled GPU can significantly improve the reconstruction runtime from several hours on a single CPU to a few minutes on a single GPU. Reconstruction steps (1) - (5) are computationally faster compared with step (6) which involves the MLP calculation. One of the existing solutions to developing a real-time pCT software is through distributing protons among tens of GPUs in a large cluster as described in [25–27] and solving the system using the existing parallel iterative solvers such as diagonally-relaxed orthogonal

projections [28] (DROP), but there are more efficient ways to solve the pCT problem while addressing the uncertainties in the system which will be discussed in this dissertation.

1.2 Contributions

Four ways are discussed in this dissertation to improve the reconstructed RSP accuracy under frequently occurring uncertain conditions in pCT. Three of these methods are incorporating robustness into DROP algorithm while the last proposed method is a new formula designed for accurate estimation of sparse systems such as the one we deal with in pCT.

There are also two efficient GPU-based algorithms proposed in this dissertation discussing different ways to distribute protons and the image to be reconstructed among GPUs on a single computer node with two or four GPUs.

1.3 Outline

Chapter 2 reviews the existing parallel iterative image reconstruction techniques used in pCT. Chapter 3 discusses three methods for improving the image reconstruction under critical conditions with modifying the existing iterative solver DROP. Chapter 4 introduces a novel fully-simultaneous iterative image reconstruction method and compares with the existing parallel iterative solvers, and chapter 5 discusses GPU-based algorithms for implementing the parallel iterative image reconstruction techniques.

CHAPTER TWO

Current General and Robust Formulations

Parallel image reconstruction methods are widely used in pCT due to their nature, which is compatible with single instruction multiple thread (SIMT) frameworks where projections onto convex sets (pocs) run concurrently for groups of protons. Several studies have addressed theoretical convergence and performance of these methods [14, 29, 30]. Parallel image reconstruction algorithms can be divided into three main groups of fully-simultaneous [31, 32], block-iterative (such as DROP [28] and BICAV [33]) and string-averaging techniques (such as CARP [34]) based on the different methods for performing parallel pocs. A classification of projection algorithms into block-iterative and string-averaging techniques is described in [35]. The parallel pocs can happen all at once as in fully-simultaneous methods, or be divided into several groups, which are run concurrently within a group, but sequentially between groups as in the block-iterative methods.

2.1 Current Parallel Iterative Image Reconstruction Methods

There are three major fully-simultaneous image reconstruction techniques: SART [32], Cimmino [31], and CAV [36]. Although, SART was found to have a slow convergence rate compared to Cimmino, it performed well on noisy data [32, 37]. CAV was shown to converge faster than the other fully-simultaneous methods. Convergence analysis of these methods is studied in [38] and it was shown that the fully-simultaneous methods can be written in framework of the general Landweber scheme [39–41]

$$x(k+1) = x(k) + wV^{-1}A^TW^{-1}(b - Ax(k)), \quad (2.1)$$

where w is the relaxation parameter, and W and V are two diagonal matrices of sizes $m \times m$ and $n \times n$ respectively. Different combinations of V and W give different iterative reconstruction algorithms. For example, choosing $V = I$ and $W = m\|A_i\|$ in equation 2.1 is equivalent to Cimmino’s algorithm, while to obtain the CAV algorithm based on the same equation, we need to pick W with diagonal elements $\sum_{j=1}^n s_j a_{i,j}^2$ where s_j is the number of nonzero elements in the j^{th} column of matrix A .

In [42], several of block-iterative and string-averaging techniques were compared and it was shown that string-averaging techniques such as CARP [34] achieved superior image quality in comparison to block-iterative methods, while block-iterative algorithms converged faster. OS-SART [43] achieved the fastest convergence rate among other block-iterative techniques.

When it comes to choosing an iterative reconstruction algorithm for pCT, DROP has found its place in many recent studies [28, 44–48], mainly because it has been shown to generate accurate 3D maps of RSP [15], is easily parallelized and it converges quickly. Another benefit of DROP compared to CARP is that it has less noise overall, due to generating lower standard deviations. This arises from the fact that total variation superiorization (TVS) happens in DROP for each block of projections, while in CARP, TVS is applied just once to the solution at the start of each iteration [27].

The DROP algorithm is described in the following section. First, $Ax = b$ is divided into a fixed number of blocks, where each block contains portions of the linear system of equations. Next, having the current iterate or $x(k)$, one performs the projections within the first block simultaneously, which the average of those projections is calculated and used as the point to perform the simultaneous projections onto the the linear system of equations in the second block. This way of projections continues until the last block is reached, which gives the next iterate or $x(k + 1)$.

DROP Algorithm

Initialization: $x(0) \in \mathfrak{R}^n$ is arbitrary.

Iterative Step: Given $x(k)$,

$$x(k+1) = x(k) + \lambda_k S_k \sum_{i \in I_t(k)} \frac{b_i - \langle A_i, x(k) \rangle}{\|A_i\|^2} a_i, \quad (2.2)$$

where $I = I_1 \cup I_2 \cup I_3 \cdots \cup I_M$ is a set containing M blocks, $t(k) = 1, 2, \dots, M$, and S_k is a $n \times n$ diagonal matrix where each diagonal term is $\frac{1}{s_j}$ where s_j is the number of times voxel j is intersected by path $i \in I_{t(k)}$. Figure 2.1 illustrates a high-level presentation of projection onto convex sets based on DROP.

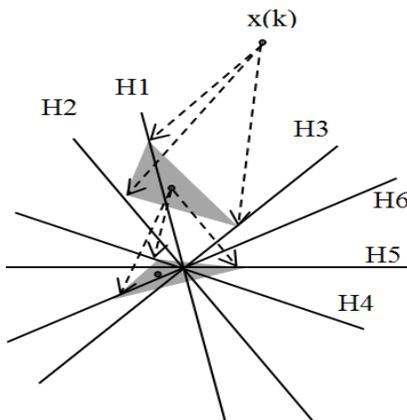


Figure 2.1: Schematic drawing of projection onto convex sets based on DROP. H_1, H_2, \dots, H_6 represent the hyper-planes, which the intersection of is the solution of the system. After splitting the data into two sets $\{H_1, H_2, H_3\}$ and $\{H_4, H_5, H_6\}$ and starting from $x(k)$, the first set of projections occurs simultaneously onto $\{H_1, H_2, H_3\}$ and the average of those will be used as the point to perform the next simultaneous projections onto $\{H_4, H_5, H_6\}$.

Because of the proven accuracy and performance of DROP in in pCT research, it is selected as the base iterative reconstruction algorithm for comparison purposes in the following chapters. Table 2.1 describes the frequently used terms in definition of algorithms such as DROP and other proposed iterative reconstruction methods in chapters three and four.

Table 2.1: Definition of frequently used mathematical terms.

Term	Definition	Chapter
A	$m \times n$ matrix containing protons path information	2, 3, 4
b	$m \times l$ vector containing the WEPL information of protons	2, 3, 4
x	$n \times 1$ vector to be reconstructed containing the RSP values	2, 3, 4
H_i	Set of hyper-planes or $\{x A_i^T x = b_i\}$, $i = 1, 2, \dots, m$	2, 4
A_i	i^{th} row of matrix A	2, 3, 4
$a_{i,j}$	An element form i^{th} row and j^{th} column of A matrix	2, 3, 4
$x(k)$	x vector at k^{th} iteration of the iterative solver	2, 3, 4
$x(k)_j$	j^{th} element of vector $x(k)$	3
λ_k	Relaxation parameter used by the iterative solver at k^{th} iteration	2, 3, 4
$I_{t(k)}$	A block of the original system of linear equations $Ax = b$	2, 3
S	$S = diag(s_j^{-1})$ of size $n \times n$	2, 3
s_j	# of times voxel j is intersected by path $i \in I_{t(k)}$	2, 3

CHAPTER THREE

Incorporating Robustness in DROP

There are certain situations in computerized tomography that might negatively affect the result of image reconstruction, such as poor selection of image reconstruction parameters (e.g. inaccurate initial iterate), lack of enough data during a scan, and overlapped projections. Lack of data during a scan could happen due to unpredictable reasons such as a hardware issue or purposely undersampled noisy data, allowing for low-dose imaging. In [49], a study on image reconstruction from a small number of projections was performed and it was shown that an algorithm based on total variation minimization can give useful results even with a small number of projections, although it is not guaranteed. In [37], a SART-type image reconstruction algorithm regularized by sparsity was proposed and shown to effectively improve the quality of reconstructed images and reduce the number of necessary projections. In [50] a multi-source SART-type image reconstruction algorithm based on the single-source SART-type method in [37] was developed to reconstruct images from overlapped projections with the goal of improving the performance of image reconstruction.

In pCT, image reconstruction based on iterative solvers can lead to inaccurate results ranging from artifacts on the image boundary to high RSP discrepancy. The underlying causes for inaccurate 3D RSP maps are mainly due to poor MLP estimations while forming the path matrix, or an imperfect initial iterate. In this chapter we discuss two main image reconstruction issues, propose hybrid methods for improving the results, and discuss the experiments on a Geant4 [51] simulated scan of the CTP404 (Sensitome) phantom which contained about 120 million proton histories with 4° increments in projection angle. The output of the image reconstruction

contains 20 slices of 200×200 where each voxel dimension is $1.0 \times 1.0 \times 2.5 \text{ mm}^3$.

Figure 3.1 shows the position of different materials inserted in CTP404 phantom.

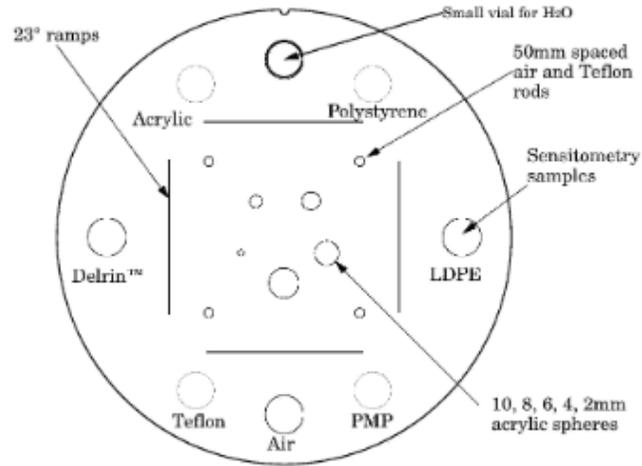


Figure 3.1: CTP404 phantom; composition and geometry of the materials.

3.1 Uncertainties in MLP

Inaccurate RSP maps with high discrepancy happen due to different reasons originating from the proton path estimations using the MLP technique and calculating the intersection length of a proton with a voxel.

In MLP, the assumption is that protons essentially traverse water due to the fact that a large portion of organs in the body are mostly water. Based on this assumption, the coefficients of the MLP are calculated and remain unchanged all the way through the proton's path estimation. Although this assumption is basically true, there are some cases in which protons traverse materials like bone, which differ significantly from the assumption made during the MLP calculation and generate inaccurate measurements of the path that a proton takes while traversing the object. In addition to the uncertainties related to MLP calculation, there are also uncertainties associated with the intersection length of a proton and a voxel which is not exact and

is calculated based on the mean chord length technique and is fixed for all the voxels along a proton’s path.

Another source of high RSP discrepancy, which mostly happens on the image boundary is caused by short proton paths (i.e. less than 30 voxels), which can only happen near the object border. Due to these short MLP paths, which also contain significant RSP discrepancy, voxels on or near the image boundary are subject to excessive updates causing visible artifacts on the boundary of the object.

3.2 Robust DROP

DROP generates accurate RSP values when $x(k)_j \cong 1$, but when $x(k)_j \gg 1$ or $x(k)_j \ll 1$, DROP tends to make the value of $x(k)_j$ closer to one. In other words, RSP values much greater than one converge slower, while RSP values much smaller than one converge faster and exceed the predicted RSPs after the first few iterations of the iterative solver. This is due to the fact that in the FBP image, (which is used as the initial iterate) the RSP value of a material like PMP, (which its RSP is much less than one), is closer to the predicted RSP of PMP, thus it converges faster. On the other hand, the RSP value of a material such as Teflon, (which is much greater than one) converges slower because the initial value of Teflon’s RSP is very far from the predicted RSP of Teflon. The % discrepancies of the RSP values generated by FBP are shown in Table 3.1.

Table 3.1: % Discrepancy of RSP values of simulated CTP404 generated by FBP. The RSP of PMP is close to the predicted RSP value of PMP, but Teflon’s RSP at the initial iterate is very far from the predicted RSP of Teflon.

Material	% Discrepancy
PMP	-1.91
LDPE	-3.46
Polystyrene	-3.70
Acrylic	-4.05
Delrin	-4.96
Teflon	-5.07

In order to prevent DROP forcing voxel values to be close to one, we perform the update of $x(k)_j$ based on the value of $(1-x(k)_j)\eta$, such that $\|E_{A_i}\| \leq \eta$, where E_{A_i} is the error in the i th row of the A matrix. Therefore, we have taken the uncertainty related to the matrix A into account. Based on these facts, we have derived a modified version of DROP that is outlined in Robust DROP¹.

Robust DROP

Initialization: $x(0) \in \mathfrak{R}^n$ is arbitrary.

Iterative Step: Given $x(k)$,

$$x(k+1) = x(k) + \lambda_k S_k \sum_{i \in I_t(k)} \sum_j \frac{b_i - \langle A_i, x(k) \rangle}{\|A_i\|^2 \pm \beta_{i,j}} a_i \quad (3.1)$$

$$\beta_{i,j} = (1 - x(k)_j)\eta \quad (3.2)$$

3.2.1 Experiments and Results of Robust DROP vs. DROP

In order to assess the outcome of Robust DROP and compare it with DROP, three experiments are performed as follows: In the first experiment, we have used the matrix A resulting from the MLP calculations, in the second experiment, we have added some Gaussian noise with zero mean and η^2 variance to each nonzero element of A to produce

$$A_{error} = A_{MLP} + \eta * N(0, 1). \quad (3.3)$$

Finally, in the third experiment, we have randomly removed 120000 of the histories from the data set and compared Robust DROP with DROP. This experiment simulates the situation that during a scan, the rate at which protons have been shot towards the object is higher than the processing rate of recording the histories, thus some of the protons are missed during the scan. This could lead to an uncertain path matrix, and therefore inaccurate RSPs during the reconstruction.

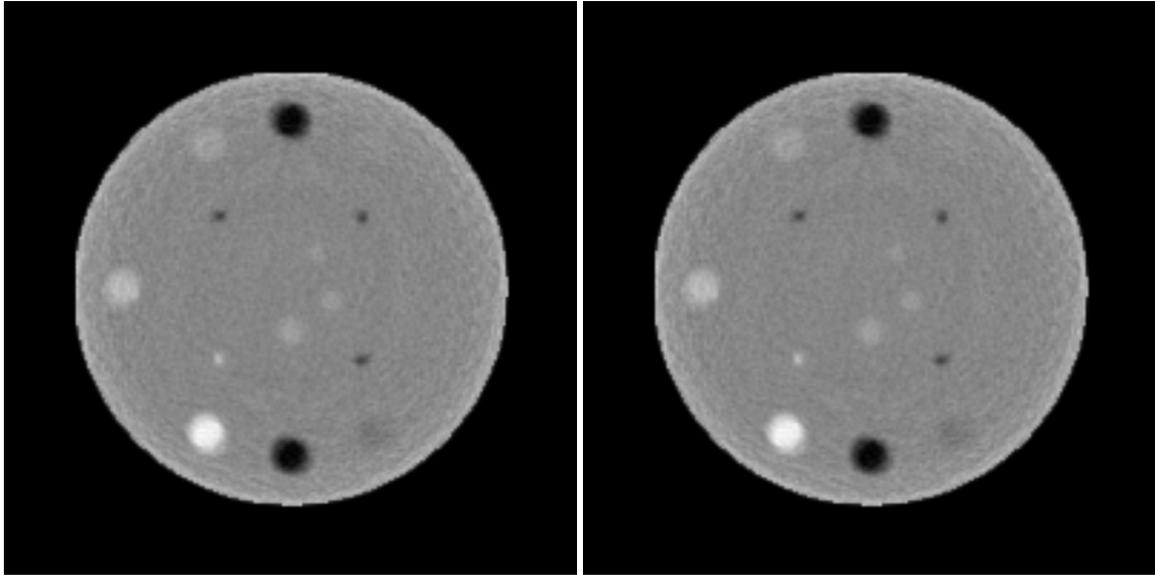
¹Some results in this chapter were published in [52]

Choosing the proper values for η depends on estimating the error in the i th row of the matrix A , which is based on the voxel size, mean chord length and number of voxels along the path. Based on the simulated CTP404 Sensitome characteristics, the value for η can range from 0.05 to 20.5.

Figures 3.2a and 3.2b illustrate the results of reconstructing images with DROP and Robust DROP, respectively. Table 3.2 includes the reconstructed RSP values by DROP and Robust algorithms. Based on the results in Figures 3.2a and 3.2b, the reconstructed pCT images generated by DROP and Robust DROP do not look different visibly. Moreover, the reconstructed RSP values of different material inserts in simulated CTP404 phantom that are reported in Table 3.2, show the outcomes of these two algorithms are very close to each other.

Figures 3.3 and 3.4 illustrate the results of the reconstructed images using DROP and Robust DROP when the matrix A contains additive Gaussian noise with $\eta = 1.0$ and $\eta = 2.0$, respectively. Tables 3.3, and 3.4 show the reconstructed RSP values of simulated CTP404 using DROP and Robust DROP where the matrix A contains additive Gaussian noise with $\eta = 1.0$ and $\eta = 2.0$, respectively.

Based on the results in Figures 3.3 and 3.4, when using DROP with the noisy path matrix, we have some voxels with RSP values much less than or greater than one on the boundary of the object. Clusters of voxels that have been marked in Figures 3.3a and 3.4a are between 1 and 5 in size, and the error in RSP of these clusters is about 98%. These artifacts on the boundary of the object should be prevented because of the fact that incorrect RSP values along the boundary of the object affect the accuracy of proton therapy. Based on the images in Figures 3.3 and 3.4, and the reconstructed RSP values in Tables 3.3 and 3.4, the benefit of the robust technique mentioned in Robust DROP algorithm is clearly visible, illustrating removal of artifacts from the boundary of the object and generating much more accurate RSP values than DROP itself.



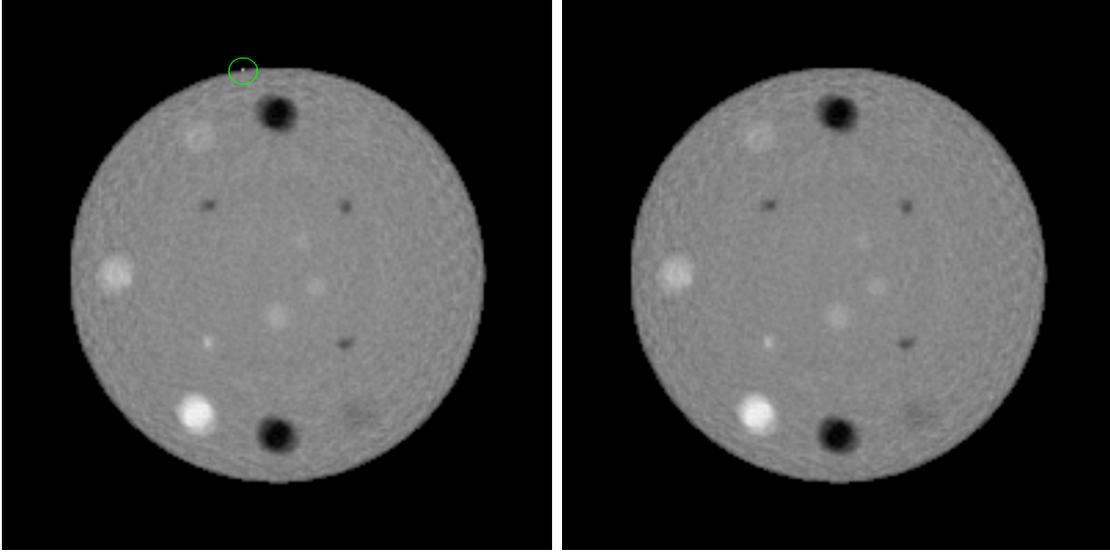
(a) DROP

(b) Robust DROP

Figure 3.2: Reconstructed image from (a) DROP, and (b) Robust DROP after six iterations, using the path matrix calculated based on the MLP. The pCT image generated by Robust DROP looks similar to the pCT image generated by DROP.

Table 3.2: Reconstructed RSP values of pCT images in Figure 3.2 with no errors in matrix A . The RSP values reconstructed by DROP and Robust DROP are very similar when there is no error in matrix A .

Material	DROP	Robust DROP	Predicted RSP
Air (bottom)	0.064	0.064	0.0013
Air (top)	0.076	0.075	0.0013
PMP	0.89	0.89	0.877
LDPE	1.01	1.01	0.997
Polystyrene	1.05	1.04	1.038
Acrylic	1.18	1.18	1.155
Delrin	1.37	1.37	1.356
Teflon	1.80	1.80	1.828



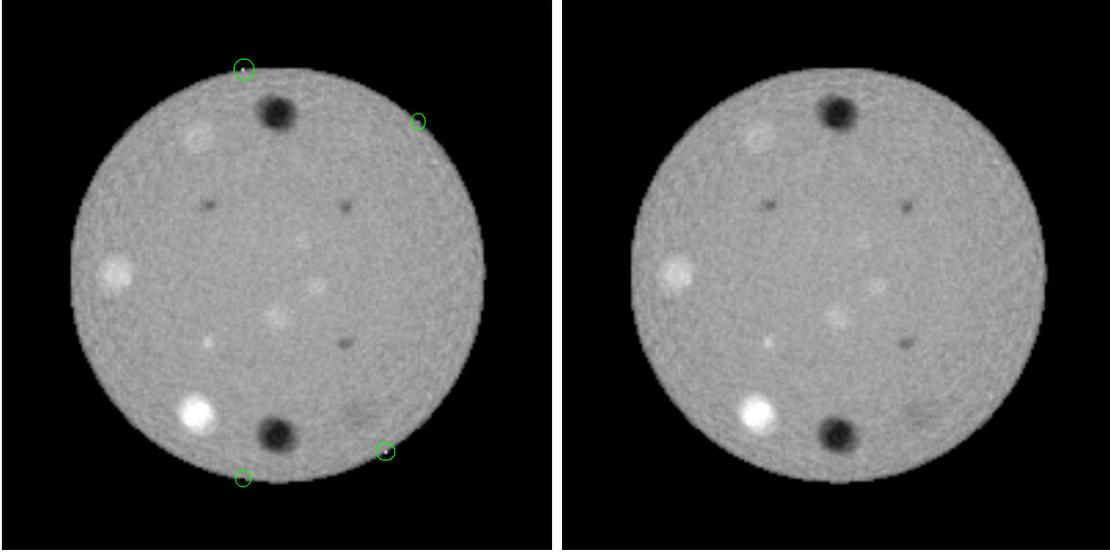
(a) DROD

(b) Robust DROD

Figure 3.3: Reconstructed image of (a) DROD, and (b) Robust DROD after 12 iterations, using the path matrix calculated based on the MLP and adding Gaussian noise ($\eta = 1.0$) to the elements of the path matrix.

Table 3.3: RSP values of reconstructed images in Figure 3.3. RSP values of all materials except Teflon generated by Robust DROD are much closer to the predicted RSP values, which indicates that Robust DROD generated more accurate results when there is additive Gaussian noise in the path matrix.

Material	DROD	Robust DROD	Predicted RSP
Air (bottom)	0.117	0.071	0.0013
Air (top)	0.119	0.074	0.0013
PMP	0.919	0.872	0.877
LDPE	1.038	0.99	0.997
Polystyrene	1.076	1.027	1.038
Acrylic	1.211	1.161	1.155
Delrin	1.39	1.34	1.356
Teflon	1.797	1.748	1.828



(a) DROD

(b) Robust DROD

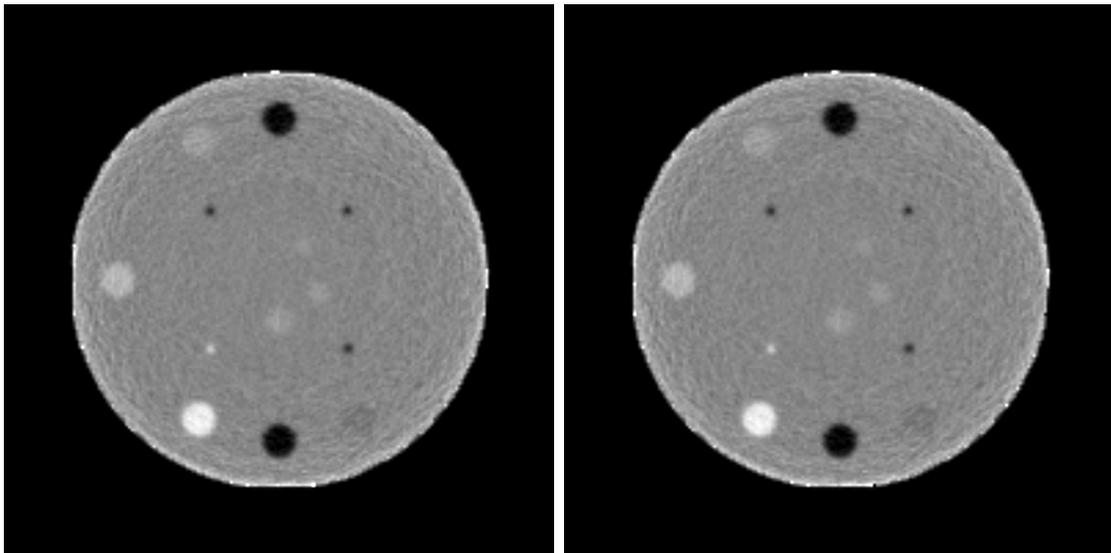
Figure 3.4: Reconstructed image of (a) DROD, and (b) Robust DROD after 12 iterations, using the path matrix calculated based on the MLP and adding Gaussian noise ($\eta = 2.0$) to the elements of the path matrix.

Table 3.4: RSP values of reconstructed images in Figure 3.4. The RSP values of all materials generated by Robust DROD are closer to the predicted RSP values, which indicates that Robust DROD generated more accurate results when there is additive Gaussian noise in the path matrix.

Material	DROD	Robust DROD	Predicted RSP
Air (bottom)	0.25	0.21	0.0013
Air (top)	0.24	0.2	0.0013
PMP	1.04	1.01	0.877
LDPE	1.17	1.12	0.997
Polystyrene	1.21	1.16	1.038
Acrylic	1.34	1.3	1.155
Delrin	1.52	1.4	1.356
Teflon	1.92	1.9	1.828

Figures 3.5(a) and 3.5(b) illustrate the results of the reconstructed images by DROP and Robust DROP, respectively, where we have randomly removed 120000 of the histories from the data set. Based on the result, in this case Robust DROP does not necessarily remove all of the artifacts on the boundary of the object, which could be due to the sudden decrease of the image voxels that get updated. When protons are removed, there are less linear equations to solve, therefore less updates to the final solution (or x), which leads to the increase of noise in the image voxels or RSP values.

Table 3.5 shows the reconstructed RSP values by DROP and Robust DROP where we have randomly removed 120000 of the histories from the data set. Based on the RSP values shown in Table 3.5, Robust DROP generates more accurate RSP values in some of the materials and is never worse than DROP in generating the RSP values for the rest of the materials.



(a) DROP

(b) Robust DROP

Figure 3.5: Reconstructed image of (a) DROP, and (b) Robust DROP after six iterations, while 120000 histories were removed randomly from the data set. Both images look similar with artifacts on the image boundaries, which occur due to the proton removals.

Table 3.5: RSP values of reconstructed images in Figure 3.5. Robust DROP generates more accurate RSP values in some of the materials and is never worse than DROP in generating the RSP values for the rest of the materials.

Material	DROP	Robust DROP	Predicted RSP
Air (bottom)	0.044	0.044	0.0013
Air (top)	0.057	0.056	0.0013
PMP	0.91	0.897	0.877
LDPE	1.004	1.001	0.997
Polystyrene	1.051	1.051	1.038
Acrylic	1.179	1.178	1.155
Delrin	1.354	1.354	1.356
Teflon	1.781	1.781	1.828

3.3 Noise Removal by Perturbing the Diagonals of the Path Matrix

When a system becomes ill-conditioned or the uncertainty is large, a robust solver is needed. In the robust estimation, one could estimate the system $Ax = b$ in the following manner:

$$x = (A^T A + \mu I)^{-1} A^T b. \quad (3.4)$$

The problem with using the above method for solving the pCT problem is that all of the rows (or columns) of the A matrix cannot be stored in memory simultaneously because of its size. Therefore, we can not compute $A^T A$ in equation 3.4 and can not perturb the diagonal. Since there is not a fast and easy way to compute $A^T A$ in equation 3.4, we tend to estimate the $A_p x = b$ with DROP while $A_p = A + \mu I$.

Figure 3.6 illustrates the vertical line segment crossing the middle of the simulated CTP404 Sensitome. Based on the results, the light grey line corresponding to the perturbed A_p with $\mu = 0.01$ is significantly improved on the image boundary compared to the dark grey line, which uses the original path matrix without perturbing

the diagonals. The reason that the perturbed path matrix generates more accurate RSP values on the image boundary is due to the fact that perturbing the diagonals of the path matrix not only modifies the eigenvalues and therefore singular values of the path matrix, but it also is equivalent to increasing the number of voxels along a proton's path within a block of DROP which causes a smaller number of the diagonal elements of matrix S resulting in smaller updates of voxels. In other words, increasing the number of voxels along a proton's path causes s_j to be increased, which is equivalent to a smaller $S = \text{diag}(s_j^{-1})$. In fact, with perturbing the matrix A , we are able to make the relaxation parameter much smaller, thus the artifacts are removed.

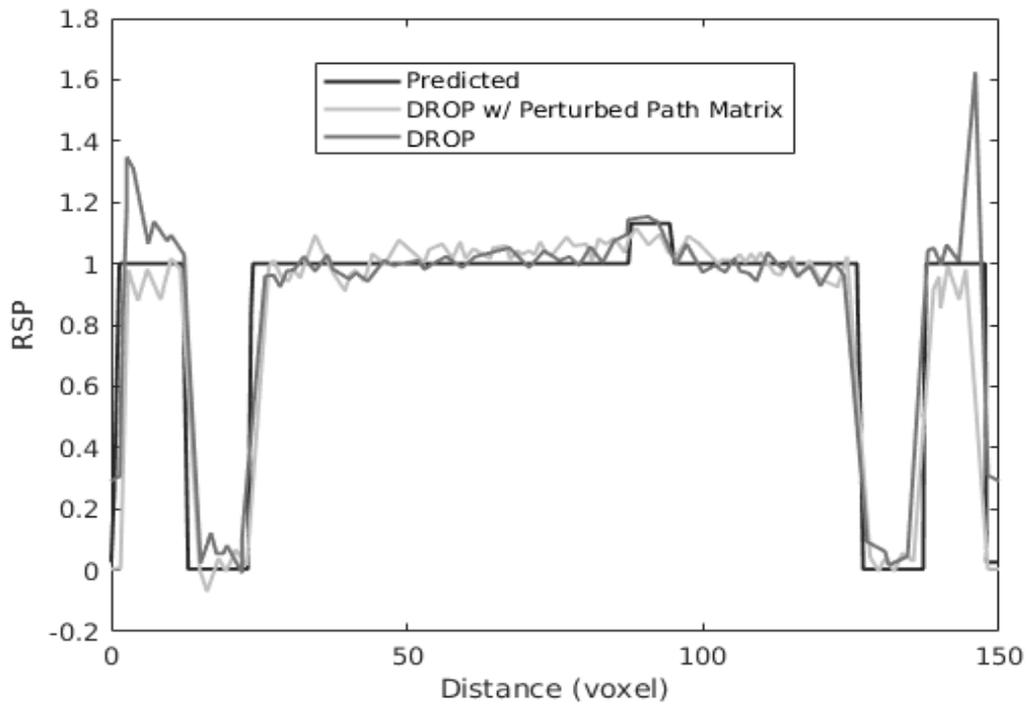


Figure 3.6: Vertical line segment on the middle of the image. Based on the results of this plot, RSP values reconstructed by DROP after perturbing the diagonal elements of path matrix is much closer to the predicted RSP values at the two ends of the plot. Using DROP without perturbing the diagonals of the path matrix causes significant spikes on the two ends of the plot which represent the clusters of voxels with high RSP values.

3.4 DROP with Additive Updates

As mentioned earlier in Chapter three, and based on our previous experiments on the simulated CTP404 Sensitome, RSP values much greater than one converge slower while the RSP values closer to one or much less than one (e.g. PMP) converge faster, which mainly happens due to the inaccurate RSP values generated by FBP.

When the initial image of the simulated CTP404 Sensitome is generated based on the FBP method, materials with RSP close to one such as PMP, LDPE, and Polystyrene, have much smaller discrepancies compared to other materials, which makes them converge faster. Based on this observation, (which is shown in Table 3.1) we can reformulate the original form of DROP as described in the following and include an additive term to address the convergence rate of different materials based on their current residue from one.

DROP with additive updates as mentioned below has the general form of DROP algorithm with the extra term $\psi(k)$, which is calculated based on the difference of the voxels from one multiplied by a user-determined constant.

DROP with Additive Updates

Initialization: $x(0) \in \mathfrak{R}^n$ is arbitrary.

Iterative Step: Given $x(k)$,

$$x(k+1) = x(k) + \lambda_k S(k) \sum_{i \in I_t(k)} \frac{b_i - \langle A_i, x(k) \rangle}{\|A_i\|^2} A_i + \psi(k) \quad (3.5)$$

$$\psi(k) = |(1 - x(k))\eta| \quad (3.6)$$

Table 3.6 compares the % discrepancy of reconstructed RSP values of different material inserts in the simulated CTP404 using eight iterations of DROP and DROP with additive updates. Based on the results in this table, the % discrepancy of RSP values of all materials is within a range of 1% of the predicted RSP values when

reconstructing with DROP with additive updates. Thus, incorporating the additive update into DROP accelerates convergence of materials such as Delrin and Teflon.

Table 3.6: Reconstructed RSP valued generated by DROP and DROP with additive updates. Based on the results, DROP with additive updates generated more accurate RSP values for all materials compared to DROP.

Material	Predicted RSP	RSP of DROP (Mean [% Discrep.])	RSP of DROP w/ additive updates (Mean [% Discrep.])
PMP	0.877	0.89 [1.48]	0.886 [0.93]
LDPE	0.997	1.01 [1.30]	0.998 [0.12]
Poly.	1.038	1.05 [1.16]	1.04 [0.15]
Acrylic	1.155	1.18 [2.54]	1.161 [0.54]
Delrin	1.356	1.37 [1.03]	1.346 [-0.74]
Teflon	1.828	1.80 [-1.53]	1.817 [-0.62]

The reason for late convergence of some materials like Delrin and Teflon is related to the image used as the initial iterate, or $x(0)$, which is discussed in more detail in the next chapter. A novel iterative fully-simultaneous image reconstruction method is presented with the focus of accelerating convergence of materials with significant discrepancy from the predicted RSP values.

CHAPTER FOUR

Sparse Robust Estimation

Over the years, a large number of methods have been developed for robust solution of mainly dense linear equations [53–56]. These methods address uncertainties, such as perturbation in the system and measurements [56], while generating a reliable solution. It has been noted that the following general matrix equation

$$x = (A^T \Phi A + \Psi)^{-1} A^T \Phi b, \quad (4.1)$$

can be used to solve any of them for the correct choice of Φ and Ψ . For instance, Total Least Squares is obtained by selecting $\Phi = I$ and $\Psi = -\sigma_{n+1} I$, with σ_{n+1} the smallest singular value of the matrix $[A \ b]$. Weighted Least Squares results from the selection of $\Phi = W^{-1}$, i.e. the inverse of the weighting matrix, and $\Psi = 0$. Ridge Regression is obtained by selecting $\Phi = I$ and $\Psi = \lambda I$, with λ the ridge parameter. Tikhonov regression is obtained by selecting $\Phi = I$ and $\Psi = \mu^2 L^T L$, with μ the regularization parameter and L the regularization matrix.

The robust solution in the form of equation 4.1 is based on calculating $A^T A$, followed by scaling and perturbing the diagonal terms of $A^T A$ and inverting the resultant matrix. These series of calculations are computationally heavy for moderate to large size dense systems. In pCT, we deal with systems of equations which are exceedingly large and sparse, solved by iterative reconstruction methods. In order to have a sparse robust solution which can be formulated as an iterative method, we start by forming the augmented system

$$\begin{bmatrix} -\Psi & A^T \\ A & \Phi^{-1} \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}. \quad (4.2)$$

The augmented system in equation 4.2 can be solved in two ways. The first way solves the second line for x then the top line for r . Alternatively, the second line of equation 4.2 can be solved for r and the top row solved for x . This yields the following equations

$$r = \Phi(b - Ax), \quad (4.3)$$

$$x = \Psi^{-1}A^T r, \quad (4.4)$$

that reduces to

$$x = \Psi^{-1}A^T\Phi(b - Ax). \quad (4.5)$$

Note that equation 4.5 is essentially of the same form as a standard iterative algorithm, with Ψ in place of λ , and Φ in the place of the per-iteration weighting. This gives a formal way for selecting values that are robust, by selecting the iterative algorithm parameters according to a robust technique. It can also be shown in next section that equation 4.5 is the same as equation 4.1, which puts certain bounds on diagonal matrices Φ and Ψ such that they must be positive definite.

Based on Equation 4.5, we can design an iterative solver considering a system of linear algebraic equations $Ax = b$, where A is an $m \times n$ real matrix. A solution point will lie in the intersection of the hyperplanes described by

$$H_i := \{x | A_i^T x = b_i\}, \quad i = 1 \cdots m. \quad (4.6)$$

Given a current approximation $x(k)$, we can write the fully simultaneous adaptive iterative solver (FSAIS) as

$$x(k+1) = x(k) + \Psi^{-1}A^T\Phi(b - Ax(k)). \quad (4.7)$$

4.1 Proof of the Compatibility of Equation 4.5 with the General Robust Equation

Consider the following augmented matrix

$$M = \begin{bmatrix} -\Psi & A^T \\ A & \Phi^{-1} \end{bmatrix}. \quad (4.8)$$

One of the nice features of matrix M is its symmetric form which enables us to apply the Cholesky factorization with the assumption that M is a Hermitian positive-definite matrix [57]. Proving the symmetric property of matrix M is a trivial task as one can easily apply the transpose operand on each block of M and the resulting matrix is the same as M . With these assumptions, the Cholesky factorization of M has the following form

$$\begin{bmatrix} -\Psi & A^T \\ A & \Phi^{-1} \end{bmatrix} = \begin{bmatrix} L_1 & 0 \\ L_2 & L_3 \end{bmatrix} \begin{bmatrix} L_1^T & L_2^T \\ 0 & L_3^T \end{bmatrix}, \quad (4.9)$$

where L_1 and L_3 are two lower triangular matrices and L_2 is a square matrix. The Cholesky factorization of the matrix M produces two sparse triangular matrices L and L^T .

Having the Cholesky factorization of matrix M , we can re-formulate the robust equation as

$$\begin{bmatrix} L_1 & 0 \\ L_2 & L_3 \end{bmatrix} \begin{bmatrix} L_1^T & L_2^T \\ 0 & L_3^T \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}. \quad (4.10)$$

In order to be able to solve this equation, we need to determine the equivalence of the L_1 , L_2 , and L_3 matrices. By multiplying the two triangular matrices obtained from the Cholesky factorization, and comparing it with the original matrix M , we end up with the following equations [58],

$$L_1 L_1^T = -\Psi, \quad (4.11)$$

$$L_1 = j\sqrt{-\Psi}, \quad (4.12)$$

$$L_2 L_1^T = A, \quad (4.13)$$

$$L_2(j\sqrt{\Psi}) = A, \quad (4.14)$$

$$L_2 = A(j\sqrt{\Psi})^{-1}, \quad (4.15)$$

$$L_2 L_2^T + L_3 L_3^T = \Phi^{-1}, \quad (4.16)$$

$$L_3 L_3^T = \Phi^{-1} - L_2 L_2^T, \quad (4.17)$$

$$L_3 L_3^T = \Phi^{-1} - A(j\sqrt{\Psi})^{-2} A^T, \quad (4.18)$$

$$L_3 L_3^T = \Phi^{-1} + A\Psi^{-1} A^T, \quad (4.19)$$

where,

$$\sqrt{\Psi} = \begin{bmatrix} \sqrt{\psi_1} & & \\ & \ddots & \\ & & \sqrt{\psi_j} \end{bmatrix}$$

Having the Cholesky factorization of the matrix M , we can solve $MX = B$ by first solving $LY = B$ for Y , and then solving $L^T X = Y$ for X . For solving the first equation $LY = B$ in the following form

$$\begin{bmatrix} L_1 & 0 \\ L_2 & L_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad (4.20)$$

we need to solve the following equations

$$L_1 y_1 = 0, \quad (4.21)$$

$$L_2 y_1 + L_3 y_2 = b, \quad (4.22)$$

which can be reduced to a single linear equation $L_3 y_2 = b$ since $y_1 = 0$. Next, we need to solve equation $L^T X = Y$ which can be written in the following form

$$\begin{bmatrix} L_1^T & L_2^T \\ 0 & L_3^T \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ y_2 \end{bmatrix}. \quad (4.23)$$

By performing the matrix multiplications on $L^T X = Y$, the following equations are achieved

$$L_1^T x + L_2^T r = 0, \quad (4.24)$$

$$L_3^T r = y_2. \quad (4.25)$$

For solving the linear systems in the above equations, first we solve equation 4.25 to obtain r , and then we can re-write equation 4.24 as $L_1^T x = y_3$ where $y_3 = -L_2^T r$, and solve for x .

In summary, we need to solve the following sets of equations to achieve a robust solution for x :

$$L_3 y_2 - b = 0, \quad (4.26)$$

$$L_1^T x + L_2^T r = 0, \quad (4.27)$$

$$L_3^T r - y_2 = 0. \quad (4.28)$$

By substituting y_2 from equation 4.28 into equation 4.26, we have the following equations

$$L_3 L_3^T r - b = 0, \quad (4.29)$$

$$(A\Psi^{-1}A^T + \Phi^{-1})r = b, \quad (4.30)$$

$$r = (A\Psi^{-1}A^T + \Phi^{-1})^{-1}b. \quad (4.31)$$

If we substitute r from equation 4.31 into $x = \Psi^{-1}A^T$ from equation 4.4, we end up with the following equation

$$x = \Psi^{-1}A^T(A\Psi^{-1}A^T + \Phi^{-1})^{-1}b. \quad (4.32)$$

Consider the following matrix identities assuming C is an invertible matrix [59, 60],

$$(A + BCD)^{-1}BC = A^{-1}(I + BCDA^{-1})^{-1}BC, \quad (4.33)$$

$$= A^{-1}B(I + CDA^{-1}B)^{-1}C, \quad (4.34)$$

$$= A^{-1}B(C^{-1} + DA^{-1}B)^{-1}. \quad (4.35)$$

Based on the above matrix identities, we can re-write equation 4.32 in the following form

$$x = \Psi^{-1}A^T(A\Psi^{-1}A^T + \Phi^{-1})^{-1}b, \quad (4.36)$$

$$= (A^T\Phi A + \Psi)^{-1}A^T\Phi b, \quad (4.37)$$

which has the form of the general robust formula in equation 4.1.

4.2 Convergence Analysis of FSAIS

In this section it will be proved that FSAIS in equation 4.7 generates sequences $x(k)$ which always converge, regardless of the initial iterate $x(0)$ and the consistency or inconsistency of the underlying system $Ax = b$. The proof is based on an convergence analysis of SART [61]. Here we form a primal optimization problem and use the gradient descent method of the dual problem to derive the convergence of FSAIS. Consider the following convex constrained optimization problem

$$\begin{cases} \text{minimize}_x & \frac{1}{2}\|x - x(0)\|_{\Psi}^2 \\ \text{subject to} & \Phi^{\frac{1}{2}}Ax = \Phi^{\frac{1}{2}}b, \end{cases} \quad (4.38)$$

where $x(0)$ is an arbitrary initial guess. We form the Lagrangian function of the primal problem by multiplying the Lagrangian coefficients or α corresponding to the equality constraints and add the result to the minimization problem. This gives the

Lagrangian function $L(x, \alpha)$ as following

$$L(x, \alpha) = \frac{1}{2} \|x - x(0)\|_{\Psi}^2 + \alpha^T \Phi^{\frac{1}{2}} (Ax - b). \quad (4.39)$$

Next, we form the dual problem by first minimizing $L(x, \alpha)$ with respect to x and then substitute the result into equation 4.39,

$$\nabla_x L(x, \alpha) = \Psi(x - x(0)) + A^T \Phi^{\frac{1}{2}} \alpha. \quad (4.40)$$

Setting $\nabla_x L(x, \alpha)$ to zero, the optimal solution is

$$x = x(0) - \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha. \quad (4.41)$$

Substituting the above optimal solution into $L(x, \alpha)$, we have

$$F(\alpha) = \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 + \alpha^T \Phi^{\frac{1}{2}} (Ax(0) - \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha - b), \quad (4.42)$$

$$= \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 + \alpha^T \Phi^{\frac{1}{2}} (Ax(0) - A\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha - b), \quad (4.43)$$

$$= \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 - \alpha^T \Phi^{\frac{1}{2}} (A\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha + b - Ax(0)), \quad (4.44)$$

$$= \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 - \alpha^T \Phi^{\frac{1}{2}} (A\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha - \alpha^T \Phi^{\frac{1}{2}} (b - Ax(0))), \quad (4.45)$$

$$= \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 - \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 - \alpha^T \Phi^{\frac{1}{2}} (b - Ax(0)), \quad (4.46)$$

$$= -\frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 - \alpha^T \Phi^{\frac{1}{2}} (b - Ax(0)). \quad (4.47)$$

Thus, we can form the dual problem as an unconstrained problem which is

$$\text{minimize}_{\alpha} F(\alpha) = \frac{1}{2} \|\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha\|_{\Psi}^2 + \alpha^T \Phi^{\frac{1}{2}} (b - Ax(0)). \quad (4.48)$$

For solving the dual problem in equation 4.48 we use the gradient descent method. First, we derive the gradient of $F(\alpha)$ in equation 4.49 and then we state the gradient descent method with unit step as shown in equation 4.50,

$$\nabla F(\alpha) = \Phi^{\frac{1}{2}} (b - Ax(0)) + \Phi^{\frac{1}{2}} A\Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha, \quad (4.49)$$

$$\alpha(k+1) - \alpha(k) = -\nabla F(\alpha(k)), \quad (4.50)$$

$$= -\Phi^{\frac{1}{2}}(b - Ax(0)) - \Phi^{\frac{1}{2}}A\Psi^{-1}A^T\Phi^{\frac{1}{2}}\alpha(k). \quad (4.51)$$

If we multiply both sides of equation 4.50 by $-\Psi^{-1}A^T\Phi^{\frac{1}{2}}$, we get a result which is in the form of the original robust solver stated in equation 4.7. The following set of equations prove this statement.

$$-\Psi^{-1}A^T\Phi^{\frac{1}{2}}\alpha(k+1) + \Psi^{-1}A^T\Phi^{\frac{1}{2}}\alpha(k) \quad (4.52)$$

$$= \Psi^{-1}A^T\Phi(b - Ax(0)) + \Psi^{-1}A^T\Phi A\Psi^{-1}A^T\Phi^{\frac{1}{2}}\alpha(k). \quad (4.53)$$

In order to simplify equation 4.53, consider equation 4.41 which lets us define $x(k) - x(0) = -\Psi^{-1}A^T\Phi^{\frac{1}{2}}\alpha(k)$. Also, based on equation 4.7 we can define $x(k+1) - x(0) = \Psi^{-1}A^T\Phi(b - Ax(0))$. Thus, equation 4.53 can be re-written in the form of the following equations

$$(x(k+1) - x(0)) - (x(k) - x(0)) \quad (4.54)$$

$$= \Psi^{-1}A^T\Phi(b - Ax(0)) - \Psi^{-1}A^T\Phi A(x(k) - x(0)), \quad (4.55)$$

$$= \Psi^{-1}A^T\Phi b - \Psi^{-1}A^T\Phi Ax(0) - \Psi^{-1}A^T\Phi Ax(k) + \Psi^{-1}A^T\Phi Ax(0), \quad (4.56)$$

$$= \Psi^{-1}A^T\Phi(b - Ax(k)). \quad (4.57)$$

The above equations show that the iterative solver shown in equation 4.7 is equivalent to the gradient descent method with unit step for the dual problem. Next, we derive the convergence of FSAIS based on the equation 4.50.

For any two vectors α_i and $\alpha_j \in \mathbf{R}^M$, we have

$$\|\nabla F(\alpha_i) - \nabla F(\alpha_j)\| = \|\Phi^{\frac{1}{2}}A\Psi^{-1}A^T\Phi^{\frac{1}{2}}(\alpha_i - \alpha_j)\|, \quad (4.58)$$

$$\leq \|\Phi^{\frac{1}{2}}A\Psi^{-1}A^T\Phi^{\frac{1}{2}}\|\|\alpha_i - \alpha_j\|, \quad (4.59)$$

$$\leq \|\alpha_i - \alpha_j\|. \quad (4.60)$$

Based on a definition of DROP algorithm and the assumptions that matrix A is sparse¹(where $a_{i,j}$ is close to one²), Φ is a $m \times m$ diagonal matrix such that $\phi_i = \frac{1}{\|A_i\|^2} < 1$, and Ψ^{-1} is a $n \times n$ diagonal matrix where $\psi_j^{-1} = \lambda_k s_j^{-1} \ll 1$ (where s_j is the number of times voxel j is intersected by path $i \in m$), thus,

$$\|\Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}}\| \leq \|\Phi^{\frac{1}{2}} A\| \|\Psi^{-1}\| \|A^T \Phi^{\frac{1}{2}}\|, \quad (4.61)$$

$$= \|\Psi^{-1}\| \|\Phi^{\frac{1}{2}} A\| \|A^T \Phi^{\frac{1}{2}}\|, \quad (4.62)$$

$$\simeq \|\Psi^{-1}\|, \quad (4.63)$$

$$\leq 1. \quad (4.64)$$

The inequality 4.60 states that $\nabla F(\alpha)$ is Lipschitz continuous (with $L = 1$) which tells us that F is upper bounded by a quadratic, therefore

$$F(\alpha_{k+1}) \leq F(\alpha_k) + \langle \alpha_{k+1} - \alpha_k, \nabla F(\alpha_k) \rangle + \frac{1}{2} \|\alpha_{k+1} - \alpha_k\|^2, \quad (4.65)$$

$$= F(\alpha_k) - \frac{1}{2} \|\alpha_{k+1} - \alpha_k\|^2, \quad (4.66)$$

$$= F(\alpha_k) - \frac{1}{2} \|\nabla F(\alpha)\|^2. \quad (4.67)$$

$F(\alpha_k)$ will decrease until α_k remains unchanged. Moreover, since $\alpha_k - \alpha_{k-1} = -\Phi^{\frac{1}{2}}(b - Ax(0)) - \Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha_{k-1}$, and thus $-\Phi^{\frac{1}{2}}(b - Ax(0)) = \alpha_k - \alpha_{k-1} + \Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha_{k-1}$, we can re-write equation 4.50 in the following form

$$\alpha_{k+1} - \alpha_k = \alpha_k - \alpha_{k-1} + \Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha_{k-1} - \Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}} \alpha_k, \quad (4.68)$$

$$= (I - \Phi^{\frac{1}{2}} A \Psi^{-1} A^T \Phi^{\frac{1}{2}})(\alpha_k - \alpha_{k-1}). \quad (4.69)$$

Similarly, we can derive a bound on $\|\alpha_k - \alpha_\star\|$ where α_\star is the optimal solution. For this, we apply the quadratic bound stated in inequality 4.65, while substituting

¹The number of non-zero elements of a row in matrix A are typically about 50 - 200.

² $a_{i,j}$ assuming a voxel size of 1 in all dimensions.

$\alpha_{k+1} = \alpha_*$ to obtain the following inequalities:

$$F(\alpha_k) \leq F(\alpha_*) + \langle \alpha_k - \alpha_*, \nabla F(\alpha_k) \rangle + \frac{1}{2} \|\alpha_k - \alpha_*\|^2, \quad (4.70)$$

$$\leq F(\alpha_*) - \|\alpha_k - \alpha_*\| \|\nabla F(\alpha_k)\| + \frac{1}{2} \|\alpha_k - \alpha_*\|^2, \quad (4.71)$$

$$\leq F(\alpha_*) - \|\alpha_k - \alpha_*\| - \|\nabla F(\alpha_k)\| + \frac{1}{2} \|\alpha_k - \alpha_*\|^2. \quad (4.72)$$

Based on the gradient descent assumption $F(\alpha_*) \leq F(\alpha_k)$, the inequality 4.73 is true,

$$-\|\alpha_k - \alpha_*\| - \|\nabla F(\alpha_k)\| + \frac{1}{2} \|\alpha_k - \alpha_*\|^2 \leq 0, \quad (4.73)$$

which leads to the following result showing the convergence of $\|\alpha_k - \alpha_*\|$:

$$\|\alpha_k - \alpha_*\| \leq 2\|\nabla F(\alpha_k)\|. \quad (4.74)$$

Moreover, we can prove the convergence of $\|\alpha_{k+1} - \alpha_k\|$ and therefore the exponential decay of $\|x(k+1) - x(k)\|_\Psi$ and $\|x(k) - x(\star)\|_\Psi$, since

$$\|x(k+1) - x(k)\|_\Psi = \|\Psi^{-\frac{1}{2}} A^T \Phi^{\frac{1}{2}} (\alpha_{k+1} - \alpha_k)\|, \quad (4.75)$$

$$= \|A^T \Phi^{\frac{1}{2}} (\alpha_{k+1} - \alpha_k)\|_{\Psi^{-1}}, \quad (4.76)$$

$$\leq \|\Phi^{\frac{1}{2}} (\alpha_{k+1} - \alpha_k)\|_\Phi, \quad (4.77)$$

$$= \|\alpha_{k+1} - \alpha_k\|. \quad (4.78)$$

Thus, we can conclude that the gradient descent is convergent. The proof used here is the same proof used for the convergence of SART in [61] which states that gradient descent is convergent while the step size is less than $\frac{2}{w}$ or $w < 2$ in equation 2.1 and $V = \Psi$ and $W = \Phi^{-1}$ are positive definite diagonal matrices. The FSAIS algorithm is outlined in equation 4.79.

FSAIS Algorithm

Initialization: $x(0) \in \mathfrak{R}^n$ is arbitrary, *Iterative Step:* Given $x(k)$,

$$x(k+1)_j = x(k)_j + \psi_j^{-1} \sum_{i=1}^m [(b_i - \langle A_i, x(k) \rangle) \phi_i] a_{i,j}. \quad (4.79)$$

4.3 Experiments and Results

For analyzing the convergence rate of FSAIS compared with the existing iterative solver DROP, we have performed three different experiments on two experimental datasets, namely CTP404 Sensitome and pediatric head phantom (PedHead). Table 4.1 shows the reconstruction parameters used for generating the pCT images of datasets used in the experiments of this chapter.

Table 4.1: Specifications of datasets used in this chapter.

Data	Type	# Protons	Voxel dim. (xyz mm^3)	# slices
CTP404	Experimental	251×10^7	$1.0 \times 1.0 \times 2.5$	20
PedHead	Experimental	251×10^7	$1.0 \times 1.0 \times 1.0$	90

In the first experiment, FSAIS showed superior convergence rate compared to DROP for all the different material inserts in experimental CTP404 and PedHead. Results of this comparison are discussed in detail in the following sections.

For assessing the convergence behavior of FSAIS under uncertain conditions, we have performed two experiments where protons within specific angle intervals are removed. The reason for removing protons within angle intervals relates to different situations causing uncertainties in computerized tomography and/or pCT. The main reason that causes missing a sector of projection angles is that the available proton energy is too small to penetrate the object in certain directions. For example, the pelvis is shaped like a rounded box and has the hips on the side. To penetrate from the side, one would need more than 230 MeV (range of 30 cm in water) in most patients, and this is not generally available (though, PROTOM company now has a 330 MeV accelerator [62]). Thus, this leads to a sector of missing projection angles from either side.

The other possible situation is related to the data loss within specific angle intervals during a scan which is mainly related to hardware issues. The other situation related to missing protons within angle intervals is low dose computerized tomography

which is discussed in several research papers and shown to reduce mortality from lung cancer. The following experiments are performed in order to analyze FSAIS with limited data from specific angle intervals,

- (1) Removing protons within angle intervals during the image reconstruction
- (2) Removing protons within angle intervals during data read and before generating the initial iterate

Each of the above items creates an uncertain condition which can affect the results of image reconstruction in different ways. If we remove protons during data read, the FBP image which is used as the initial iterate will be affected, while if protons are removed during the image reconstruction, the assumption is that the initial iterate is safely generated in the past, but the image reconstruction with the iterative solver needs to be repeated with limited data from angle intervals due to data loss or low dose computerized tomography. Using an existing initial iterate for multiple reconstruction is a valid assumption due to the fact that there can be cases in which several scans on the same target or data is needed, thus, one can use an existing initial iterate from previous scans of the same data and perform multiple image reconstructions with the iterative solver.

4.4 Convergence Analysis of FSAIS and DROP on CTP404 Data

Image reconstructions in this chapter were executed on a single node of a compute cluster with input data read from a local solid state drive (SSD) and the majority of computations were performed in parallel on a single NVIDIA K40 GPU. The total computation time from reading of input data from SSD through the writing of reconstructed images to SSD was about 3.7 minutes for DROP and 3 minutes for FSAIS.

The image analysis program ImageJ2 1.51r [63] was used to perform quantitative analyses of reconstructed image quality of the slice number 10 of phantom. The ovular selection and measurement tool from ImageJ was used to select a 7 mm

diameter circular region of interest within the boundary of each cylindrical insert and calculate the mean and standard deviation in reconstructed RSP, with identical region selection and analyses performed for all images reconstructed with FSAIS and DROP.

Figure 4.1 along with Table 4.2 illustrate the convergence rate of DROP using the FBP image as the initial iterate and after 6 iterations while the optimal block size and relaxation parameter are 1280000 and 0.00015 respectively. Moreover, the TVS method is used during the first and second iterations of DROP and prior to performing the projections onto each block of convex sets. The criteria for choosing the optimal reconstruction parameters is based on reconstructing high quality images within the shortest amount of time with maximum number of materials with RSP discrepancy within 1% of the predicted RSPs. The reconstructed pCT image of results in Table 4.2 is shown in Figure 4.2.

Based on the results of DROP in Figure 4.1 and Table 4.2, materials such as PMP, LDPE and Polystyrene converge much faster than materials like Teflon and Delrin which have not converged after 6 iterations. The reason which causes delayed convergence of some materials like Teflon compared to accelerated convergence of other materials like PMP is due to the fact that the RSP of Teflon has a significant discrepancy from the predicted RSP in the initial iterate generated by FBP. Table 4.3 shows the discrepancy of the CTPT404 materials from the predicted RSP values at the initial iterate, which is generated by FBP. Based on the results in Table 4.3, RSP values of PMP, LDPE are much closer to their predicted RSPs than Teflon and Delrin's difference with respect to their predicted RSP values.

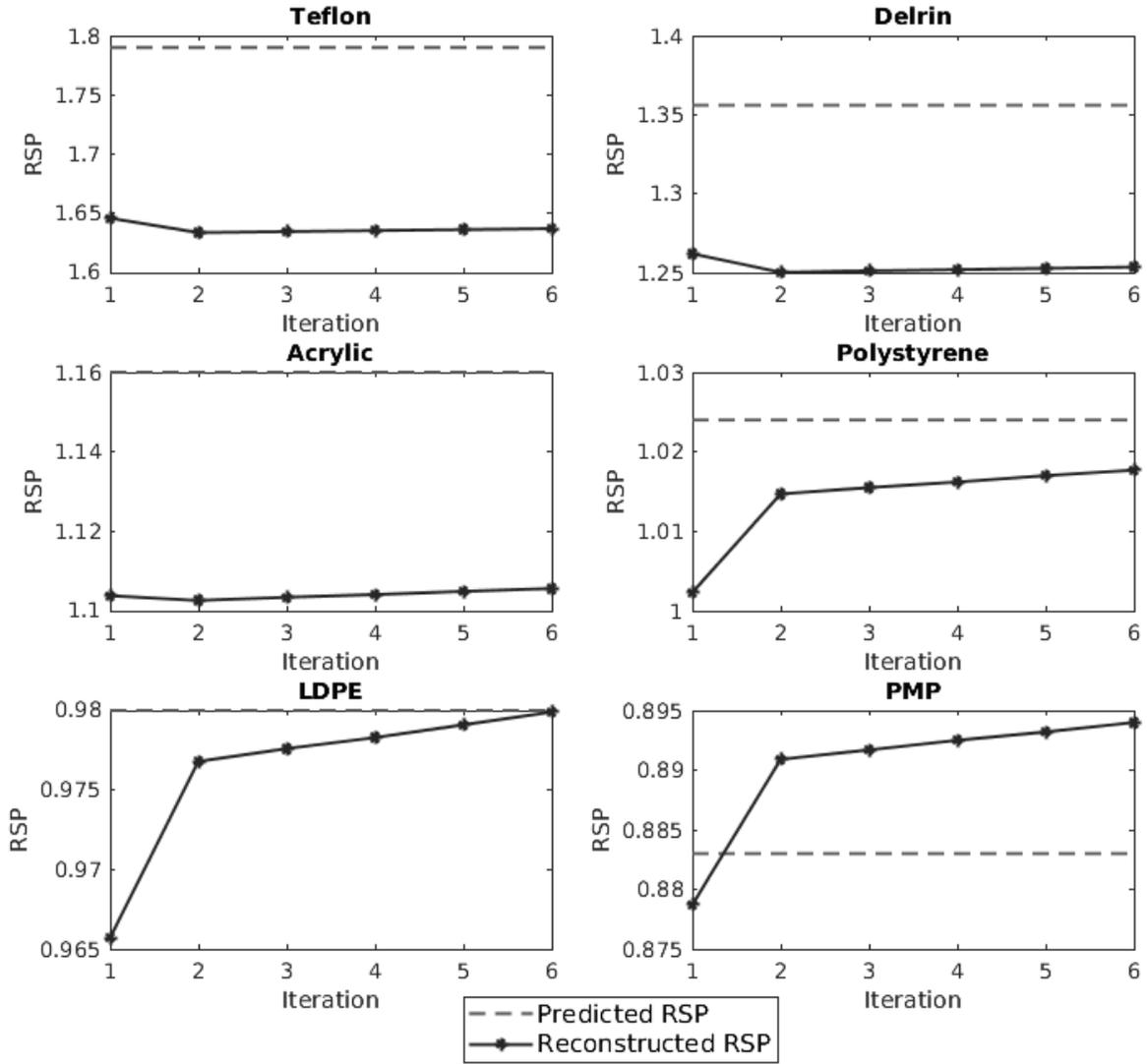


Figure 4.1: Convergence analysis of reconstructed RSP values for the materials of experimental CTP404 phantom generated by DROP with block size 1280000 and $\lambda = 0.00015$ while using TVS during the first and second iterations. % Discrepancy of PMP, LDPE and Polystyrene is less than 1% after the first iteration, while other materials do not converge even after 6 iterations of DROP. Details of the mean reconstructed RSP values are given in Table 4.2.

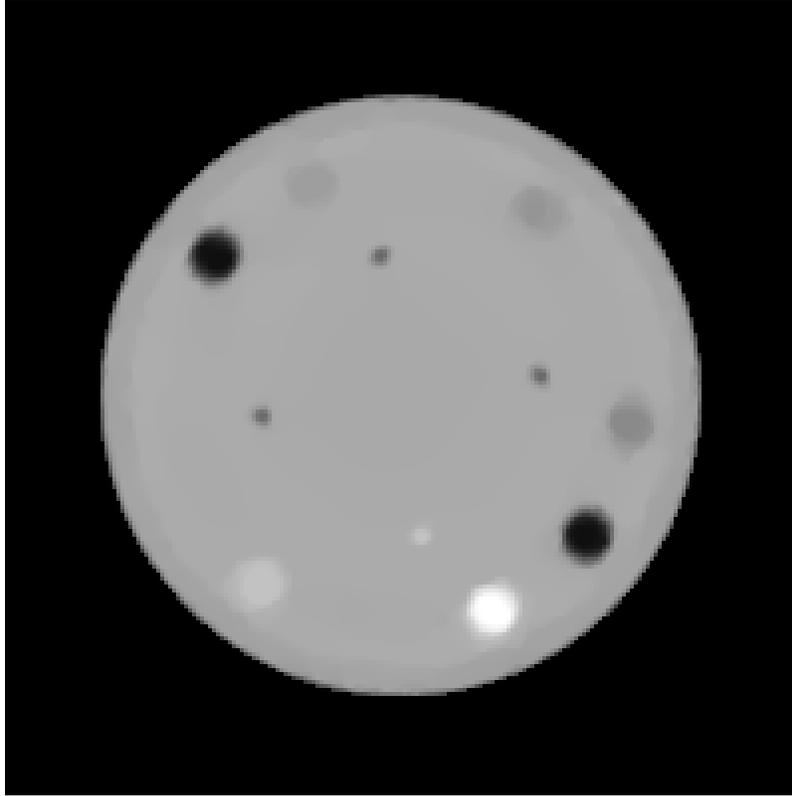


Figure 4.2: Experimental CTP404 pCT reconstruction by DROP using TVS during first and second iterations.

Table 4.2: Reconstructed RSP values for the materials of experimental CTP404 generated by DROP with block size 1280000 and $\lambda = 0.00015$ after six iterations. An important observation related to Teflon and Delrin is that DROP generated inaccurate RSP values, even worse than RSPs at the initial iterate (see Table 4.3).

Material	Reconstructed RSP	StdDev	% Discrepancy
PMP	0.894	0.008	1.25
LDPE	0.98	0.017	-0.01
Polystyrene	1.018	0.001	-0.61
Acrylic	1.106	0.006	-4.69
Delrin	1.254	0.001	-7.56
Teflon	1.637	0.016	-8.54

Table 4.3: RSP statistics of materials at the initial iterate of CTP404.

Material	Predicted	Mean	StdDev	Min.	Max.	% Discrepancy
PMP	0.883	0.8626	0.0332	0.8053	0.952	-2.31
LDPE	0.980	0.9527	0.0409	0.8655	1.0612	-2.79
Polystyrene	1.024	0.9915	0.0229	0.9405	1.0627	-3.17
Acrylic	1.160	1.1048	0.0304	1.0492	1.1763	-4.75
Delrin	1.359	1.2755	0.0214	1.2342	1.3127	-5.94
Teflon	1.79	1.6647	0.0318	1.5794	1.7207	-7.0

Knowing the RSP values of materials at the initial iterate suggests a way to pick the elements of Ψ^{-1} used in FSAIS such that RSP of materials like Teflon converge faster than materials like PMP. Based on this observation, there is a method for picking the elements of Ψ^{-1} which requires selecting elements of Ψ^{-1} for specific RSP intervals. In other words, knowing that PMP is converging faster than Teflon or Delrin, one can put some bounds on voxels about to be updated such that all voxels $x(k)_j$ belonging to certain ranges are updated while $\psi^{-1}(j)$ s for those voxels are selected based on a prior knowledge about the initial iterate. Thus, selecting smaller values of $\psi^{-1}(j)$ for voxels close to PMP's RSP while picking greater values of $\psi^{-1}(j)$ for voxels close to Teflon's RSP will help improving the convergence of all materials. Table 4.4 shows the optimal values of Ψ^{-1} for CTP404 dataset.

Table 4.4: Elements of Ψ^{-1} for experimental CTP404.

RSP range	$\psi^{-1}(j)$
0.87 - 0.91	$ 1 - x(k)_j \times 0.00000001$
0.99 - 1.03	$ 1 - x(k)_j \times 0.000085$
1.05 - 1.35	$ 1 - x(k)_j \times 0.00015$
1.13 - 1.16	$ 1 - x(k)_j \times 0.000085$
1.35 - 1.8	$ 1 - x(k)_j \times 0.000085$

Selecting RSP ranges that do not overlap is not necessarily a good way to form Ψ^{-1} because it causes pCT images do not look smooth. The RSP ranges that are shown in the first column of Table 4.4 were selected by first choosing the intervals based on Teflon and Delrin's RSP values in order to make sure those are receiving

enough updates. Then, the intervals were further created to include other materials. Moreover, if a voxel is updated within a specific range and it also belongs to a different region, it will get updated again. The starting point of a range is selected based on values close to the mean RSP values at the initial iterate or after the first iteration of algorithms. Since we use TVS during the first and second iterations of the iterative solvers, some materials are affected by that and their RPS values even get smaller than they were at the initial iterate. Thus, one needs to select RSP ranges and values of $\psi^{-1}(j)$ based on their need as the FSAIS can adapt to different conditions. As shown in Table 4.4, we use $|1 - x_j|$ multiplied by a constant. This is due to the fact that materials with RSP closer to 1.0 converge faster than materials with RSP far from 1.0 (like Teflon). This suggests we can use the difference of material RSPs to adjust the values of $\psi^{-1}(j)$ such that those with RSP values closer to 1.0 get smaller updates.

Figure 4.3 shows the reconstructed RSP values of the experimental CTP404 from the first to the last iteration of FSAIS where TVS was used during the first and second iterations. Unlike the results generated by DROP (shown earlier in Figure 4.1), Teflon, Delrin, and Acrylic get very close to the predicted RSP values in the last iteration. The reconstructed pCT image at the last iteration of FSAIS is shown in Figure 4.4. Table 4.5 shows the mean RSP and % discrepancy for material inserts in the experimental CTP404 at the last iteration of FSAIS where TVS was used during the first and second iterations. Based on the results in Table 4.5 (compared to DROP results in Table 4.2), FSAIS improves the RSP values of Teflon, Delrin, and Acrylic significantly and reduces the discrepancy in the results for those materials to less than 1%. PMP is the only material that has a discrepancy greater than 1%. This is due to the effect of TVS, which is performed during the first and second iterations of DROP and FSAIS which causes the discrepancy of PMP to grow very fast after only one iteration of the iterative solvers. Therefore, we can not lower the discrepancy of

PMP even with a small update of voxels close to the RSP value of PMP (i.e. voxels within a range of 0.87 - 0.91).

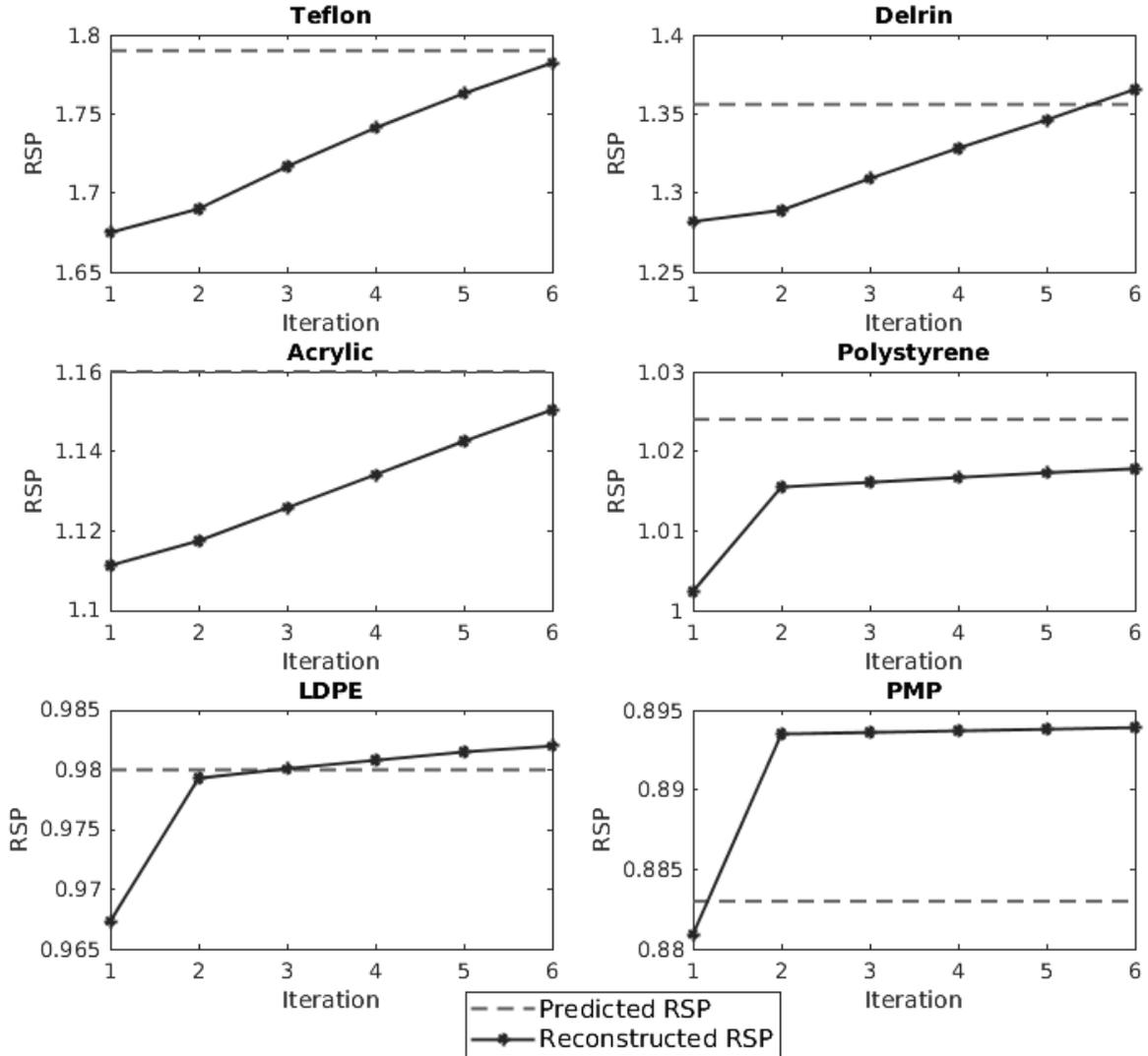


Figure 4.3: Convergence analysis of reconstructed RSP values for the materials of experimental CTP404 phantom generated by FSAIS using TVS during the first and second iterations. Details of the mean reconstructed RSP values are given in Table 4.5. Based on the results, materials such as Teflon, Delrin and Acrylic, which did not converge in the case of DROP (Figure 4.1), converge or get very close to the predicted RSPs after six iterations of FSAIS algorithm.

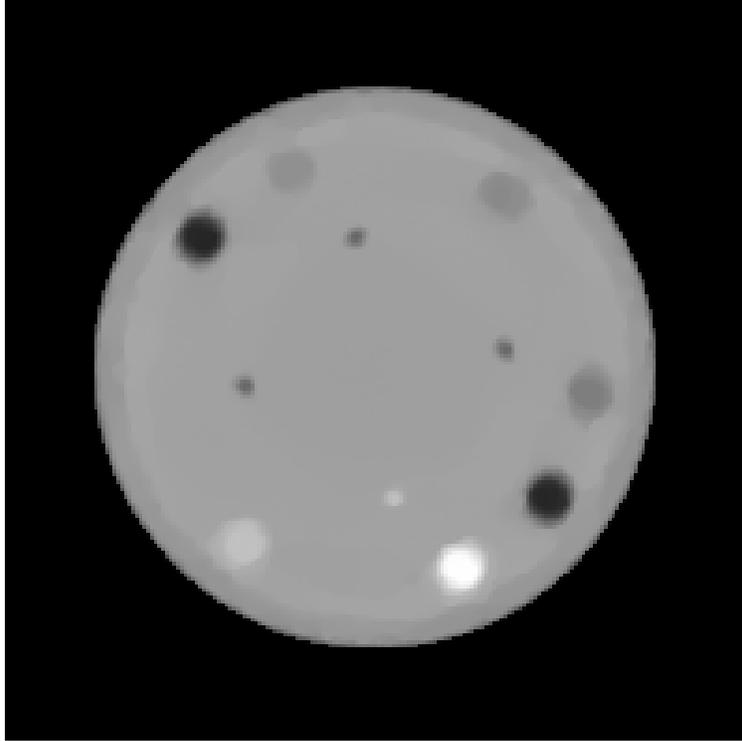


Figure 4.4: Experimental CTP404 pCT reconstruction by FSAIS using TVS during the first iteration and second iterations.

Table 4.5: Reconstructed RSP values for the materials of experimental CTP404 generated by FSAIS while using TVS during the first and second iterations. The RSP values of all materials except PMP are within a range of 1% of the predicted RSPs, which indicates FSAIS generated more accurate RSPs compared to DROP.

Material	Reconstructed RSP	StdDev	% Discrepancy
PMP	0.894	0.0072	1.24
LDPE	0.982	0.0117	0.21
Polystyrene	1.018	0.0018	-0.61
Acrylic	1.15	0.0082	-0.83
Delrin	1.366	0.0062	0.7
Teflon	1.782	0.0205	-0.43

Figure 4.5 shows the reconstructed RSP values of the experimental CTP404 from the first to the last iteration of FSAIS where TVS was used only during the first iteration. Based on the results in Figure 4.5, removing the TVS during the

second iteration improves the RSP of PMP, but has a negative effect on the RSP of Polystyrene.

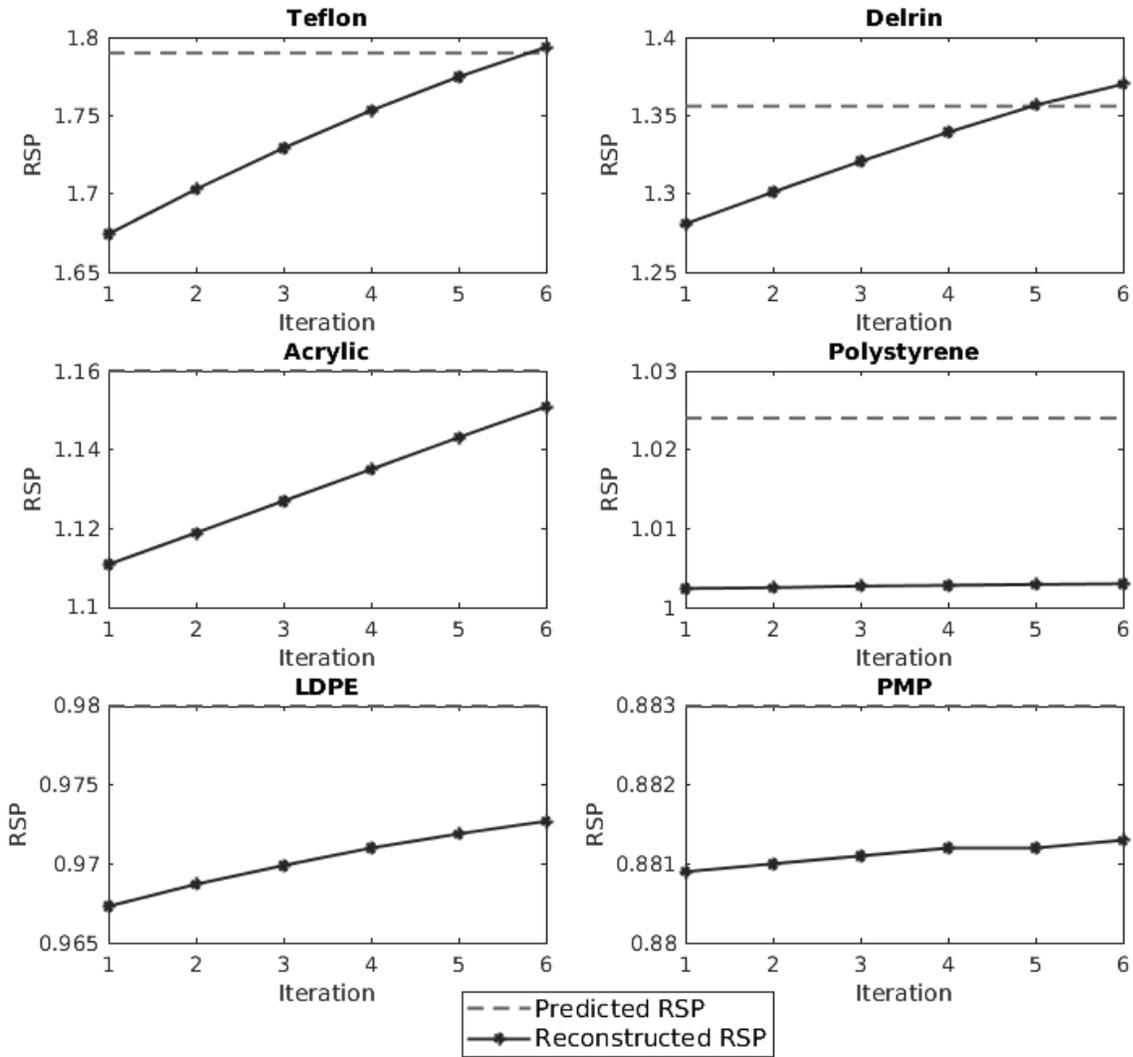


Figure 4.5: Convergence analysis of reconstructed RSP values of the experimental CTP404 phantom generated by FSAIS using TVS during the first iteration. Details of the mean reconstructed RSP values are given in Table 4.6. Based on the results, although removing the TVS during the first iteration causes the RSP value of PMP be more accurate, but it also causes a high discrepancy for the RSP of Polystyrene.

The reconstructed pCT image at the last iteration of FSAIS (where TVS was used only during the first iteration) is shown in Figure 4.6. Table 4.6 shows the mean

RSP and % discrepancy for material inserts in the experimental CTP404 at the last iteration of FSAIS (where TVS was used only during the first iteration). Results in Table 4.6 show a good reconstructed RSP value for PMP, but the pCT image of these RSP values in Figure 4.6 is not as smooth as the one in Figure 4.4, which was generated with TVS during the first and second iterations of FSAIS.

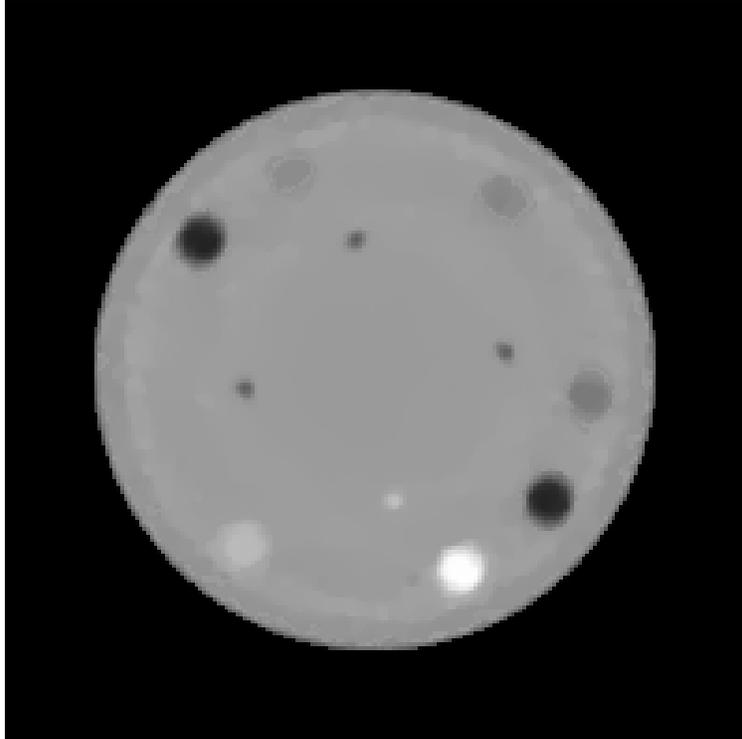


Figure 4.6: Experimental CTP404 pCT reconstruction by FSAIS using TVS during the first iteration. The image looks less smooth than the image shown in Figure 4.4.

Table 4.6: Reconstructed RSPs for the materials of experimental CTP404 generated by FSAIS while using TVS during the first iteration. The RSPs of all materials except Polystyrene and Delrin are within a range of 1% of the predicted RSPs.

Material	Reconstructed RSP	StdDev	% Discrepancy
PMP	0.881	0.0106	-0.19
LDPE	0.973	0.017	-0.74
Polystyrene	1.003	0.0045	-2.05
Acrylic	1.151	0.0089	-0.78
Delrin	1.37	0.0043	1.03
Teflon	1.794	0.0234	0.20

4.4.1 *Removing Protons from Angle Intervals During Reconstruction of CTP404*

Based on the results shown in Figure 4.7, removing protons from 0.0 to 60.0 degrees does not cause a significant change in RSP values of materials reconstructed by FSAIS and DROP in general. The % discrepancies of RSP values of Delrin, Polystyrene and LDPE reconstructed by FSAIS remain within the 1.0% range even after removing protons from 0.0 to 60.0 degree. Teflon and Acrylic's RSPs are still better than -2.0% after removing protons from 0.0 to 60.0. The reason that FSAIS does a better job in keeping the discrepancy of Polystyrene, LDPE, and PMP within 1.0% range could be due to the actual position of these materials in the CTP404 phantom which are less affected by removing protons within 0.0 to 60.0 degree. On the other hand, materials like Teflon, Delrin and Acrylic are on right side of the CTP404 phantom and can be affected more from proton removals.

Figures 4.8, 4.9, 4.10, 4.11, 4.12, 4.13 illustrate the convergence rate of the RSP of different materials reconstructed by DROP where protons were removed from different angle intervals during the image reconstruction. Figure 4.14 shows the pCT images reconstructed by DROP where protons were removed during the image reconstruction. In addition, Figures 4.15, 4.16, 4.17, 4.18, 4.19, 4.20 illustrate the convergence rate of RSP of different materials reconstructed by FSAIS. Convergence rate plots while removing protons during reconstruction have a similar pattern as the convergence plots under normal conditions which are illustrated in Figures 4.1 and 4.4. Figure 4.21 shows the pCT image of CTP404 reconstructed by FSAIS where protons were removed from different angle intervals during reconstruction. Removing protons within angle intervals during image reconstruction does not cause a blur in the reconstructed images shown in Figures 4.14 and 4.21.

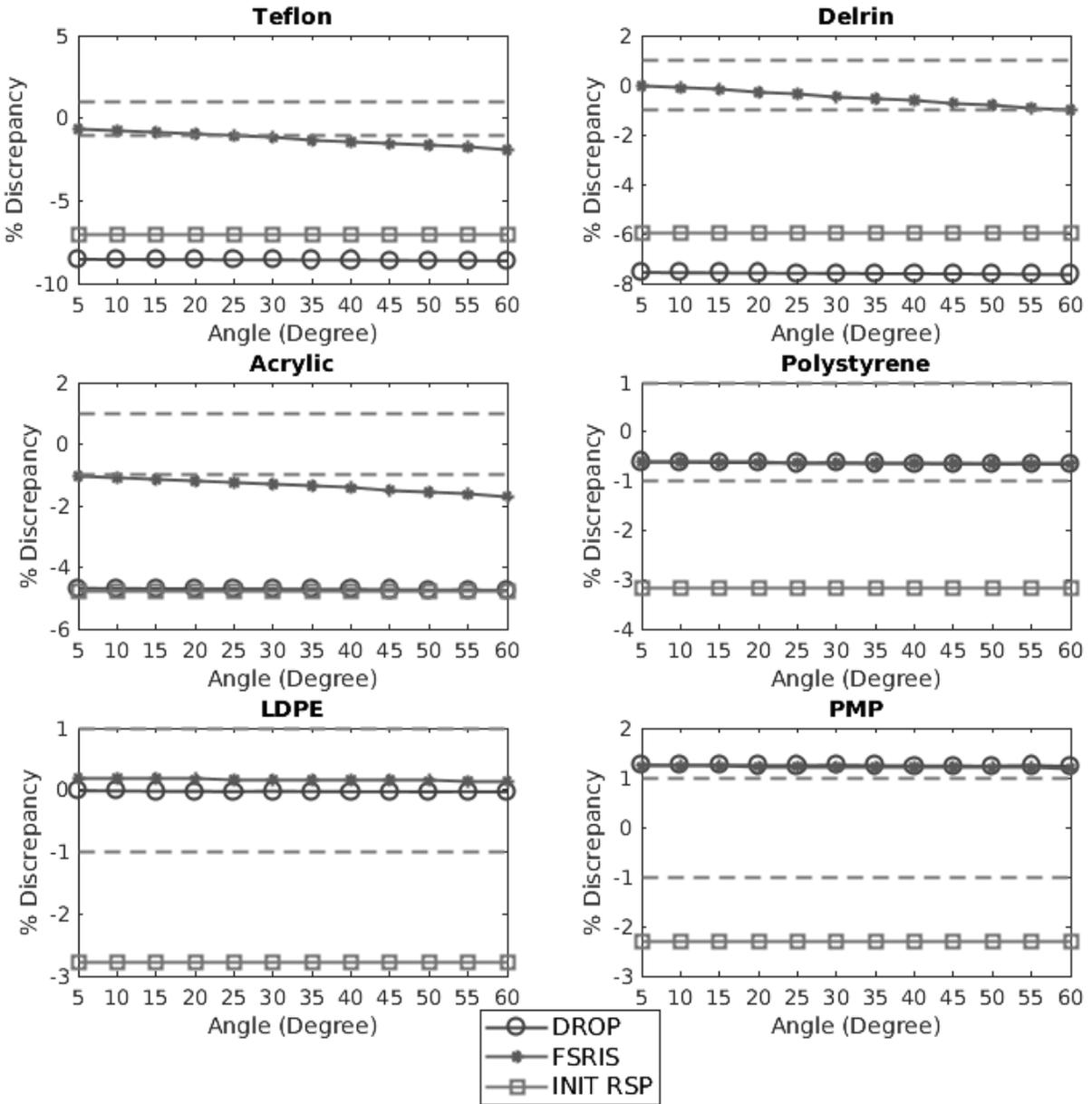


Figure 4.7: % Discrepancy of material inserts in CTP404 while removing protons during the image reconstruction. Each angle d on the horizontal axis represents the upper bound of the angle interval from which protons were removed (lower bound is zero). Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. The RSP discrepancies for all materials except Teflon and Acrylic remain within the 1% range of the predicted RSP values even after protons were removed within a range of $0.0^\circ - 60.0^\circ$. An important observation related to Teflon and Delrin is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate generated by FBP, while FSAIS is generating much more accurate results for these materials.

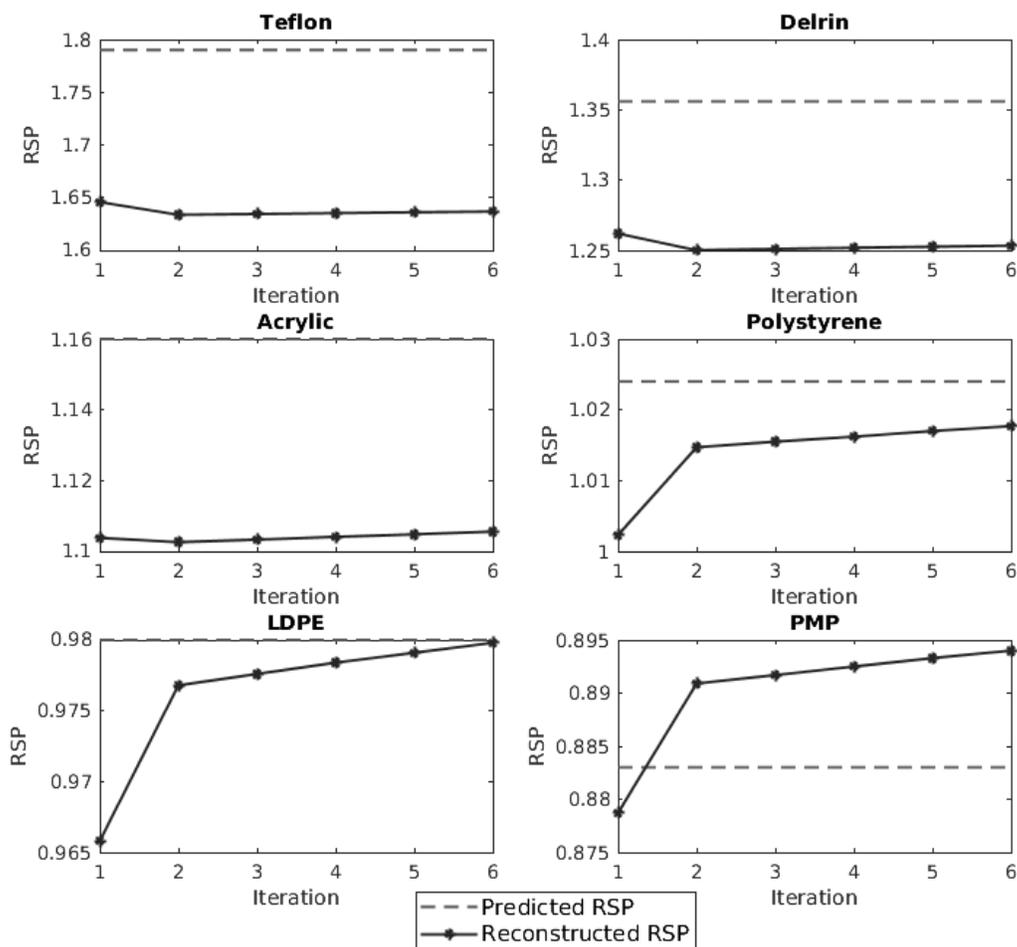


Figure 4.8: DROF convergence analysis after removing protons within a range of $0.0^\circ - 10.0^\circ$ during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.7 for the % discrepancy of materials at the 6th iteration.

Table 4.7: DROF generated RSP values after removing protons within $0.0^\circ - 10.0^\circ$ range during reconstruction of experimental CTP404.

Material	Mean \pm SD	% Discrepancy
PMP	0.894 ± 0.0069	1.25
LDPE	0.9798 ± 0.0157	-0.02
Polystyrene	1.0177 ± 0.0016	-0.62
Acrylic	1.1056 ± 0.0046	-4.69
Delrin	1.2534 ± 0.0008	-7.57
Teflon	1.6369 ± 0.0155	-8.55

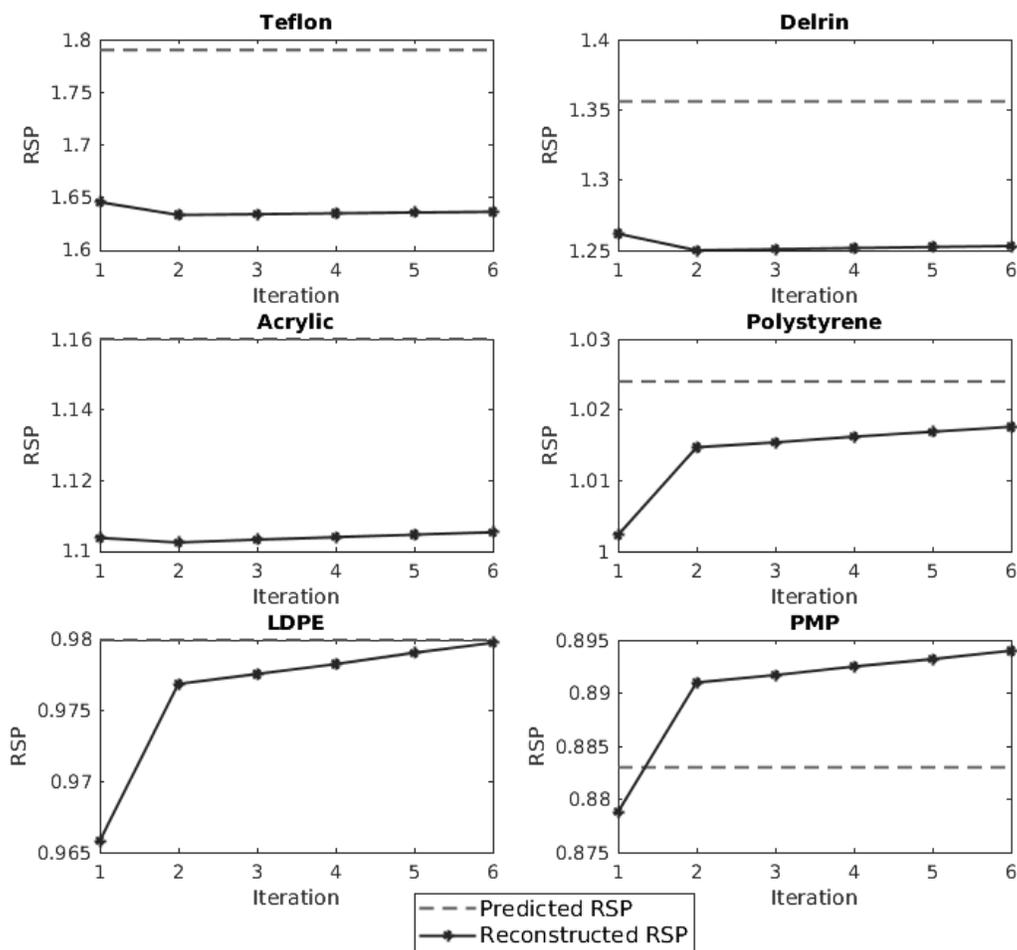


Figure 4.9: DROF convergence analysis after removing protons within $0.0^\circ - 20.0^\circ$ range during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.8 for the % discrepancy of materials at the 6th iteration.

Table 4.8: DROF generated RSP values after removing protons within a range of $0.0^\circ - 20.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.894 ± 0.0069	1.24
LDPE	0.9798 ± 0.0157	-0.02
Polystyrene	1.0176 ± 0.0016	-0.63
Acrylic	1.1054 ± 0.0046	-4.70
Delrin	1.2531 ± 0.0008	-7.59
Teflon	1.6366 ± 0.0154	-8.57

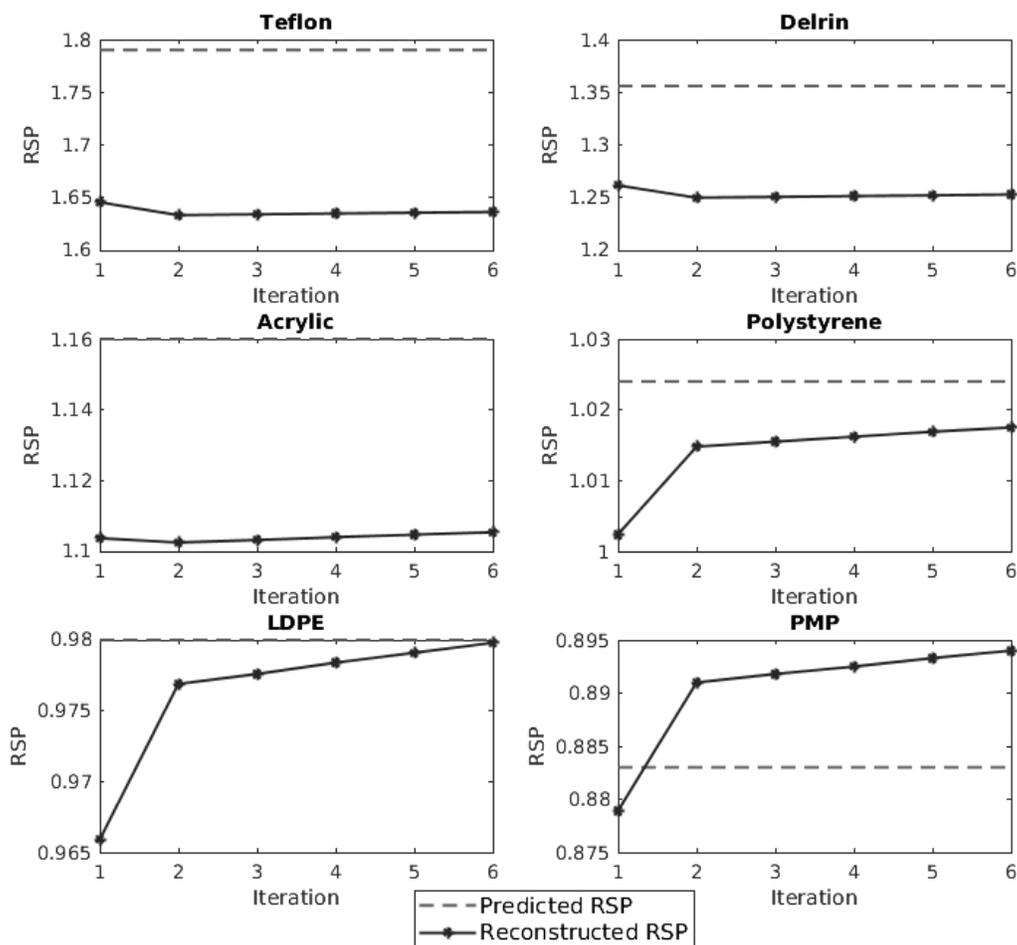


Figure 4.10: DROF convergence analysis after removing protons within a range of $0.0^\circ - 30.0^\circ$ during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.9 for the % discrepancy of materials at the 6th iteration.

Table 4.9: DROF generated RSP values after removing protons within a range of $0.0^\circ - 30.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.894 ± 0.0069	1.25
LDPE	0.9798 ± 0.0157	-0.02
Polystyrene	1.0175 ± 0.0015	-0.63
Acrylic	1.1054 ± 0.0046	-4.71
Delrin	1.253 ± 0.0008	-7.60
Teflon	1.6364 ± 0.0154	-8.58

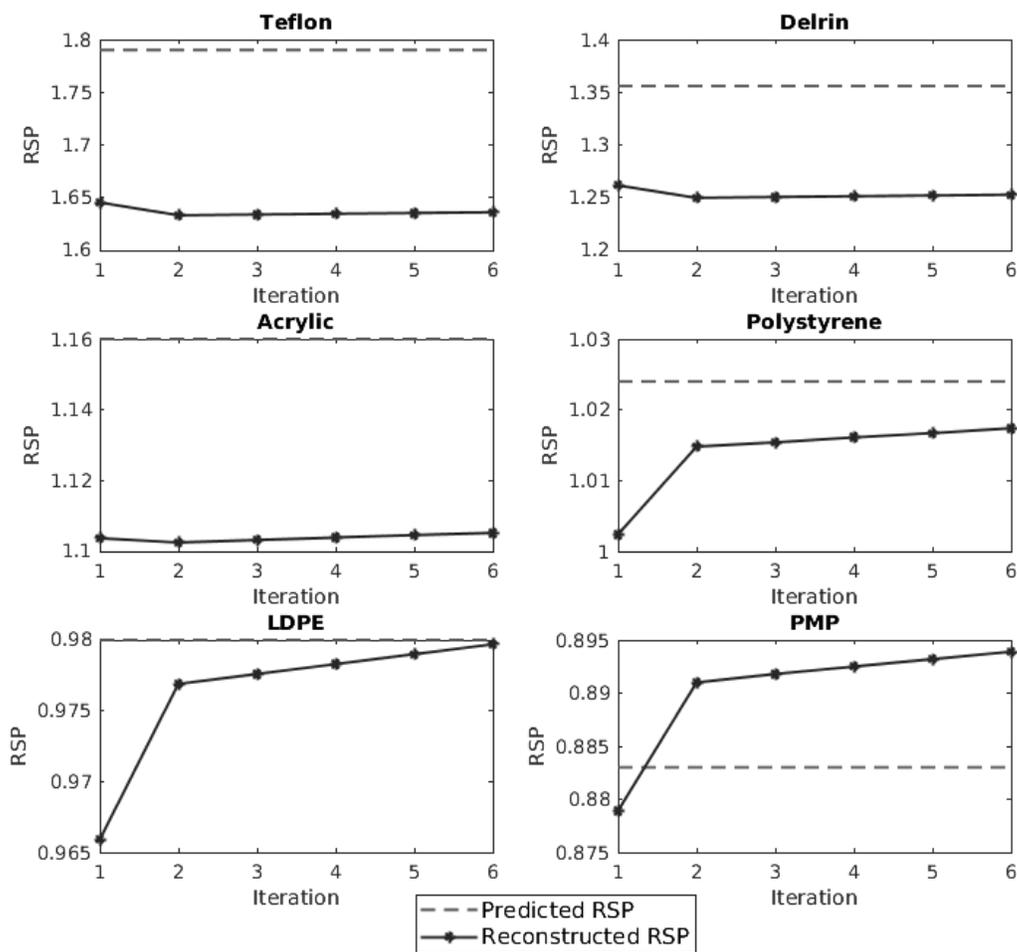


Figure 4.11: DROF convergence analysis after removing protons within a range of $0.0^\circ - 40.0^\circ$ during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.10 for the % discrepancy of materials at the 6th iteration.

Table 4.10: DROF generated RSP values after removing protons within a range of $0.0^\circ - 40.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8939 ± 0.0068	1.24
LDPE	0.9797 ± 0.0156	-0.03
Polystyrene	1.0174 ± 0.0015	-0.65
Acrylic	1.1052 ± 0.0046	-4.72
Delrin	1.2528 ± 0.0008	-7.61
Teflon	1.6361 ± 0.0154	-8.60

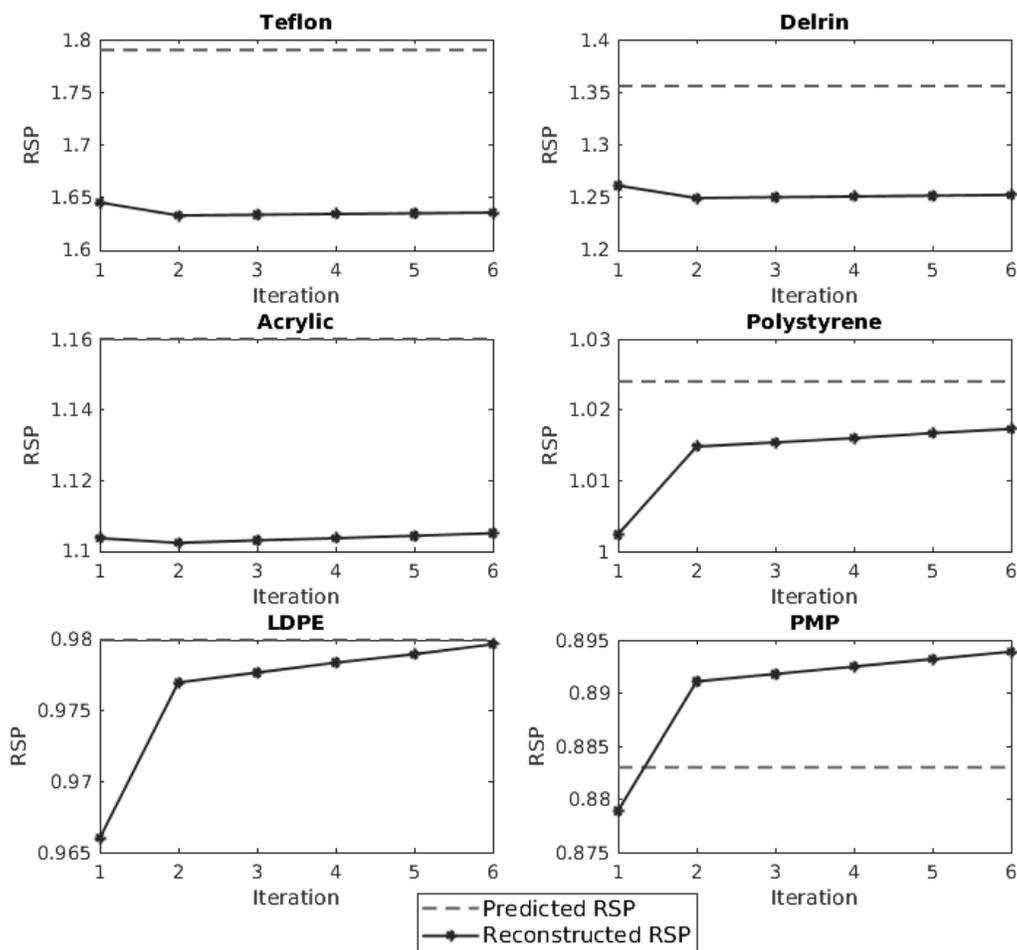


Figure 4.12: DROF convergence analysis after removing protons within a range of $0.0^\circ - 50.0^\circ$ during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.11 for the % discrepancy of materials at the 6th iteration.

Table 4.11: DROF generated RSP values after removing protons within a range of $0.0^\circ - 50.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8939 ± 0.0068	1.23
LDPE	0.9797 ± 0.0156	-0.03
Polystyrene	1.0173 ± 0.0015	-0.66
Acrylic	1.1051 ± 0.0046	-4.74
Delrin	1.2526 ± 0.0008	-7.63
Teflon	1.6358 ± 0.0154	-8.62

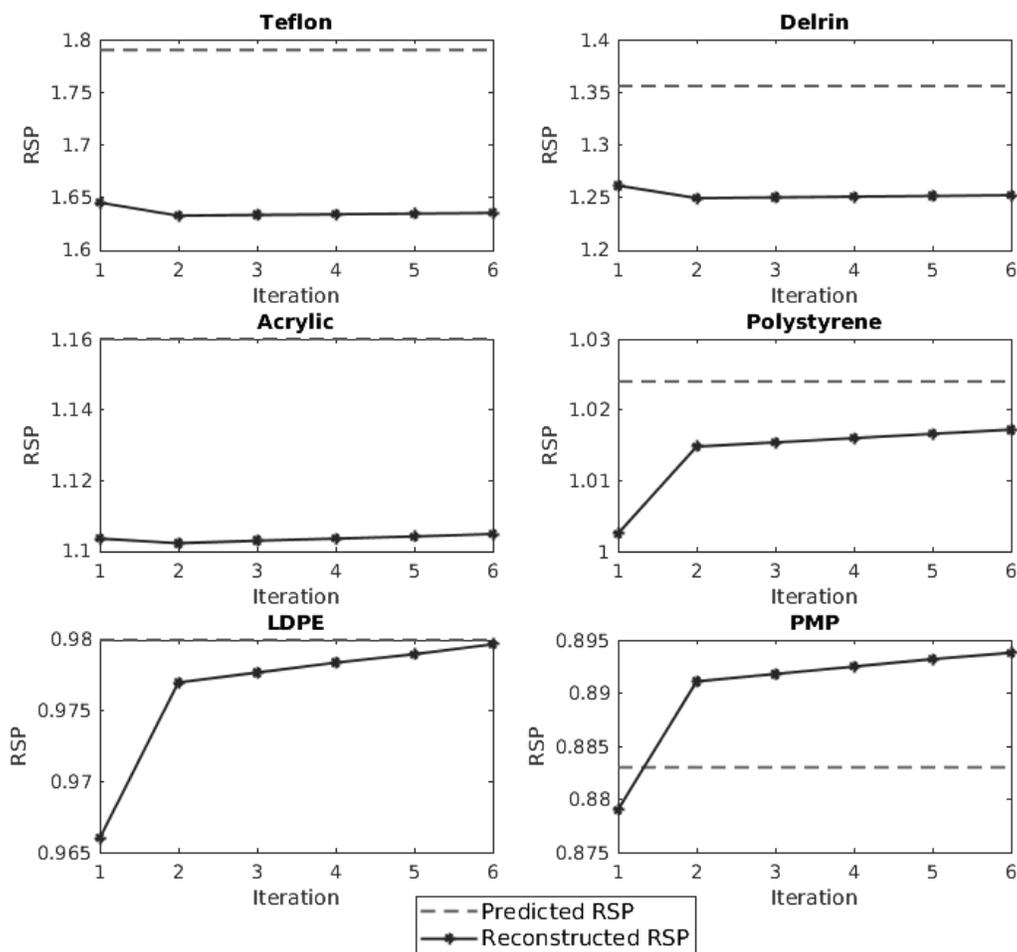
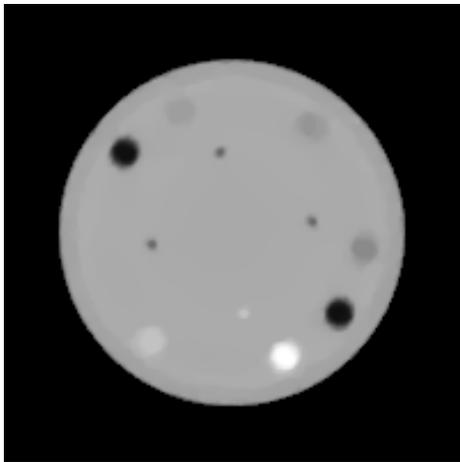


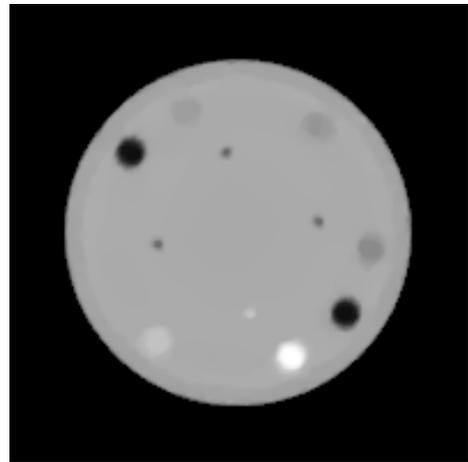
Figure 4.13: DROF convergence analysis after removing protons within a range of $0.0^\circ - 60.0^\circ$ during reconstruction of experimental CTP404. The RSP values of Teflon, Delrin and Acrylic are very far from the predicted RSP values even after six iterations. See Table 4.12 for the % discrepancy of materials at the 6th iteration.

Table 4.12: DROF generated RSP values after removing protons within a range of $0.0^\circ - 60.0^\circ$ during reconstruction of experimental CTP404.

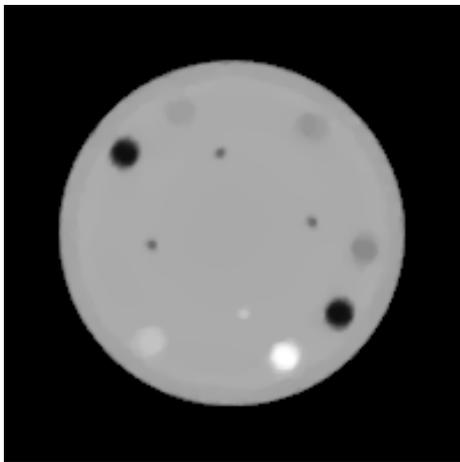
Material	Mean \pm StdDev	% Discrepancy
PMP	0.8938 ± 0.0068	1.22
LDPE	0.9797 ± 0.0156	-0.03
Polystyrene	1.0172 ± 0.0015	-0.66
Acrylic	1.1049 ± 0.0046	-4.75
Delrin	1.2523 ± 0.0007	-7.64
Teflon	1.6355 ± 0.0153	-8.63



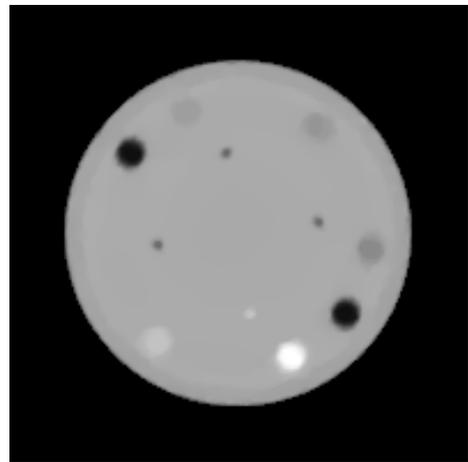
(a) $0.0^\circ - 10.0^\circ$



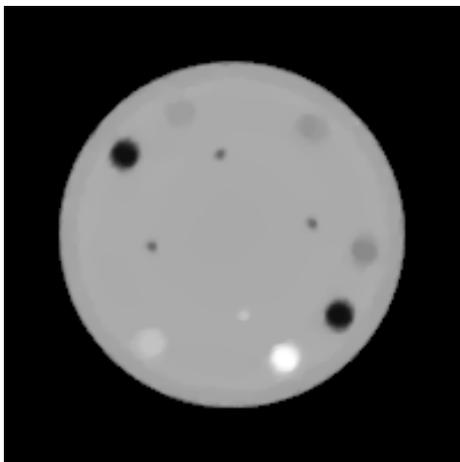
(b) $0.0^\circ - 20.0^\circ$



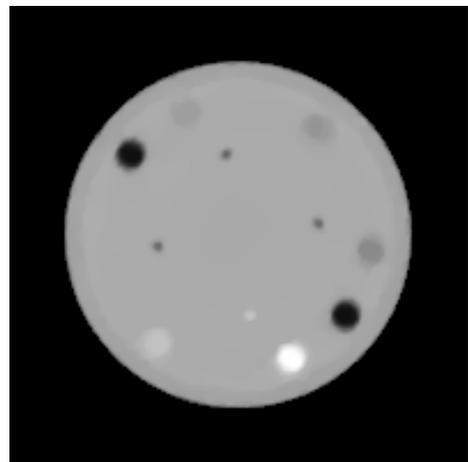
(c) $0.0^\circ - 30.0^\circ$



(d) $0.0^\circ - 40.0^\circ$



(e) $0.0^\circ - 50.0^\circ$



(f) $0.0^\circ - 60.0^\circ$

Figure 4.14: Experimental CTP404 pCT reconstruction by DROP while removing protons during image reconstruction. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the image reconstruction.

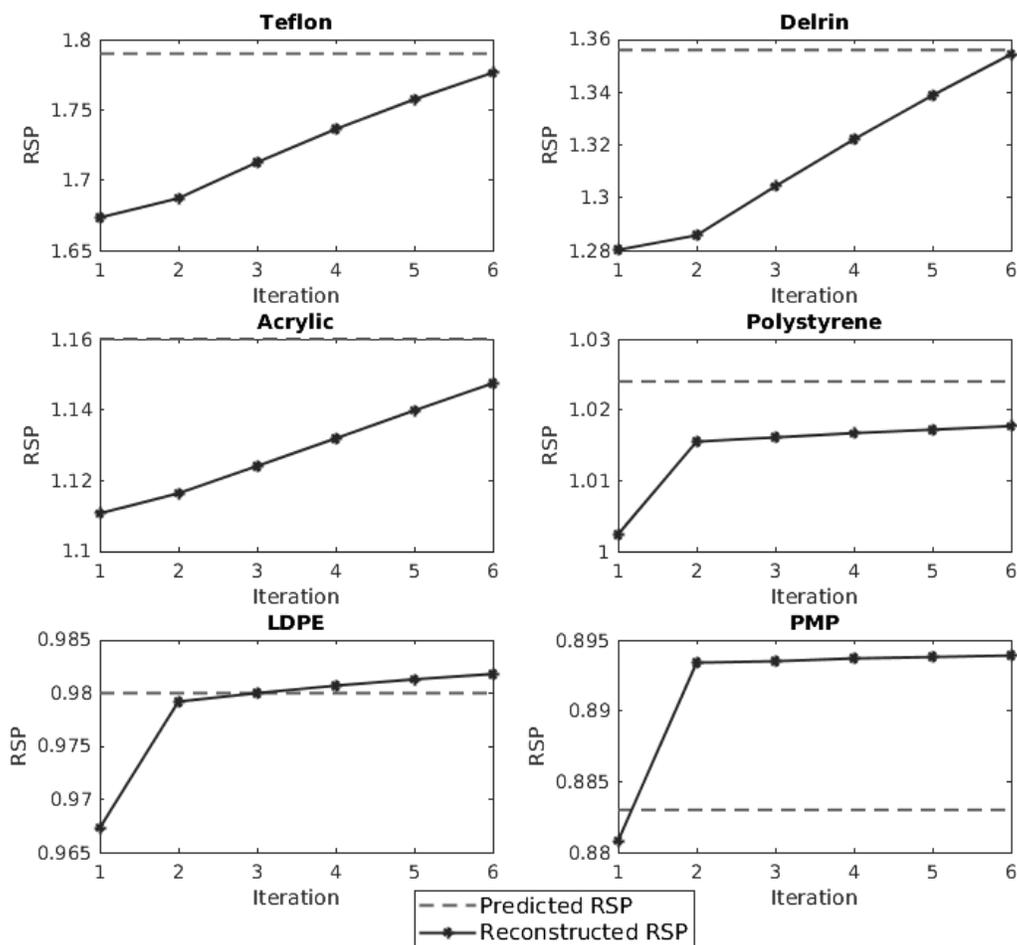


Figure 4.15: FSAIS convergence analysis after removing protons within a range of $0.0^\circ - 10.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials are very good and close to the predicted RSP values after six iterations. See Table 4.13 for the % discrepancy of materials at the 6th iteration.

Table 4.13: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 10.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8939 ± 0.0072	1.23
LDPE	0.9818 ± 0.0118	0.19
Polystyrene	1.0177 ± 0.0018	-0.61
Acrylic	1.1475 ± 0.0084	-1.08
Delrin	1.3544 ± 0.0018	-0.12
Teflon	1.7764 ± 0.0191	-0.76

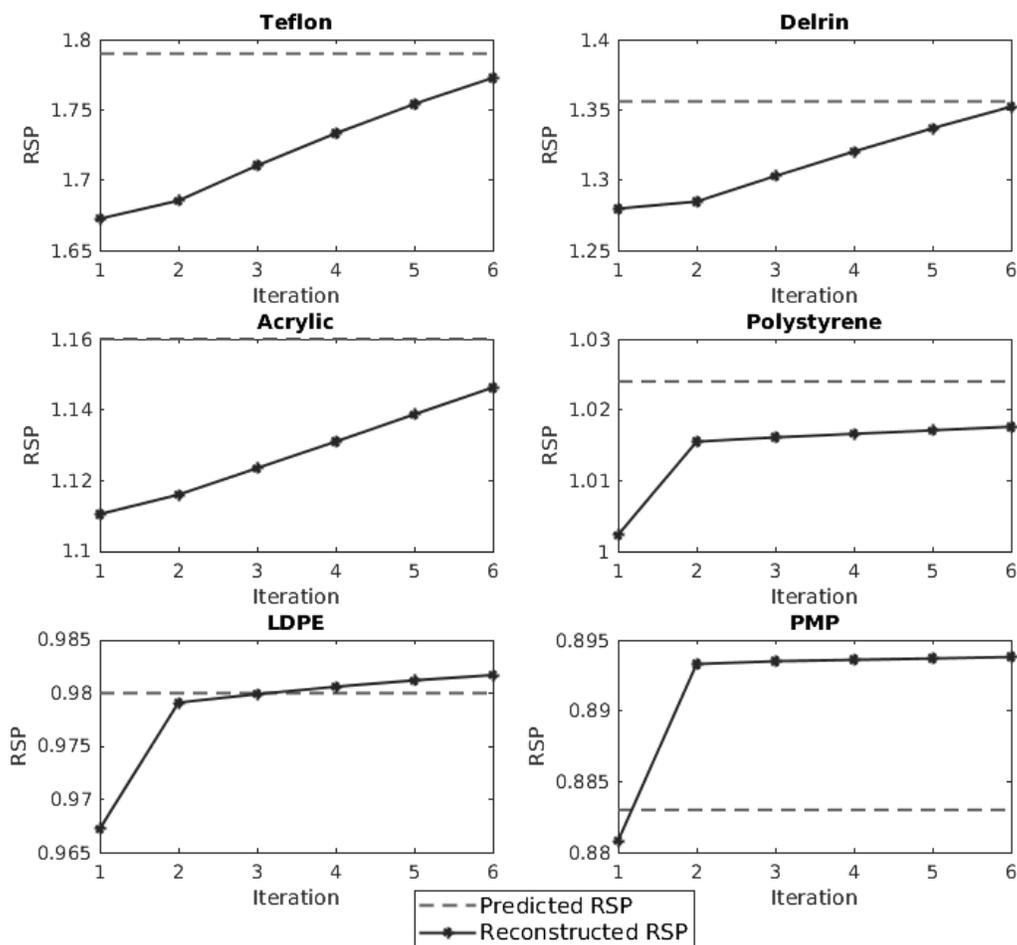


Figure 4.16: FSAIS convergence analysis after removing protons within a range of $0.0^\circ - 20.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials are very good and close to the predicted RSP values after six iterations. See Table 4.14 for the % discrepancy of materials at the 6th iteration.

Table 4.14: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 20.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8938 ± 0.0072	1.22
LDPE	0.9817 ± 0.0119	0.18
Polystyrene	1.0176 ± 0.0018	-0.62
Acrylic	1.1463 ± 0.0082	-1.18
Delrin	1.3522 ± 0.0018	-0.28
Teflon	1.7727 ± 0.019	-0.97

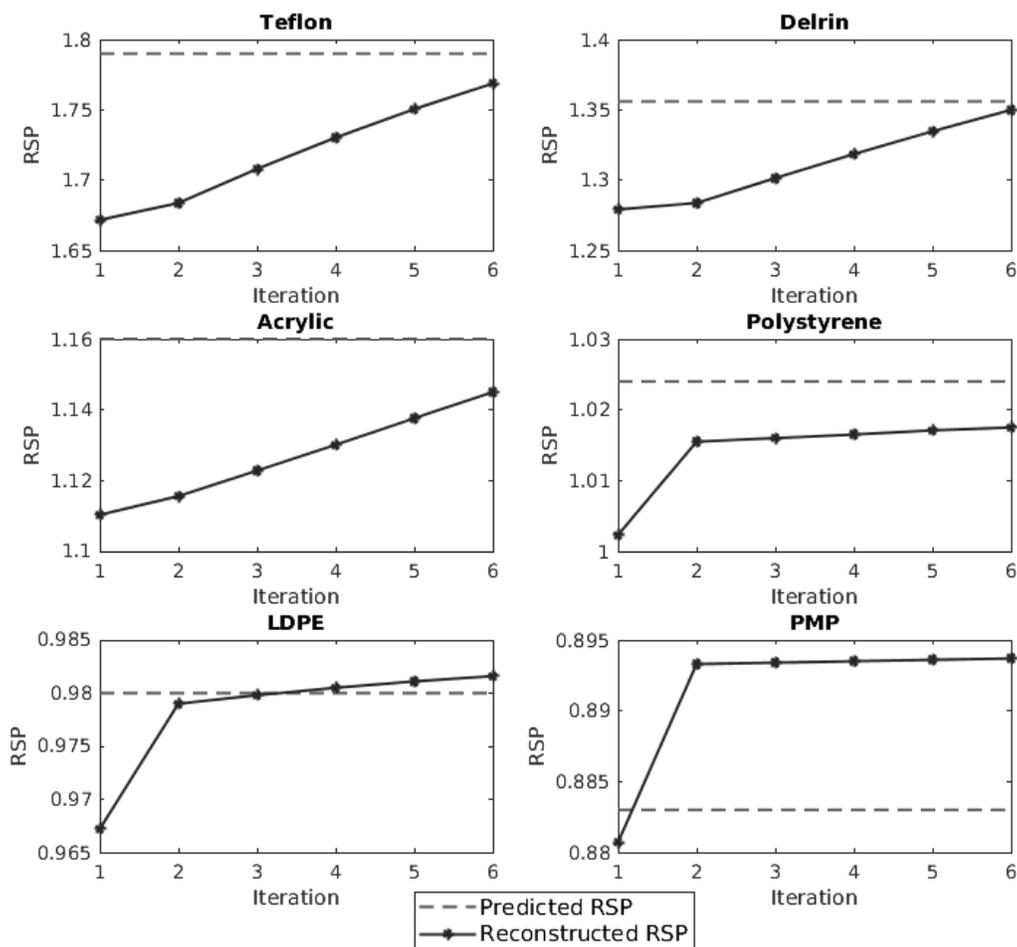


Figure 4.17: FSAIS convergence analysis after removing protons within a range of $0.0^\circ - 30.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials are very good and close to the predicted RSP values after six iterations. See Table 4.15 for the % discrepancy of materials at the 6th iteration.

Table 4.15: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 30.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8937 ± 0.0072	1.21
LDPE	0.9816 ± 0.0119	0.17
Polystyrene	1.0175 ± 0.0018	-0.63
Acrylic	1.145 ± 0.0081	-1.30
Delrin	1.3499 ± 0.0018	-0.45
Teflon	1.7687 ± 0.0187	-1.19

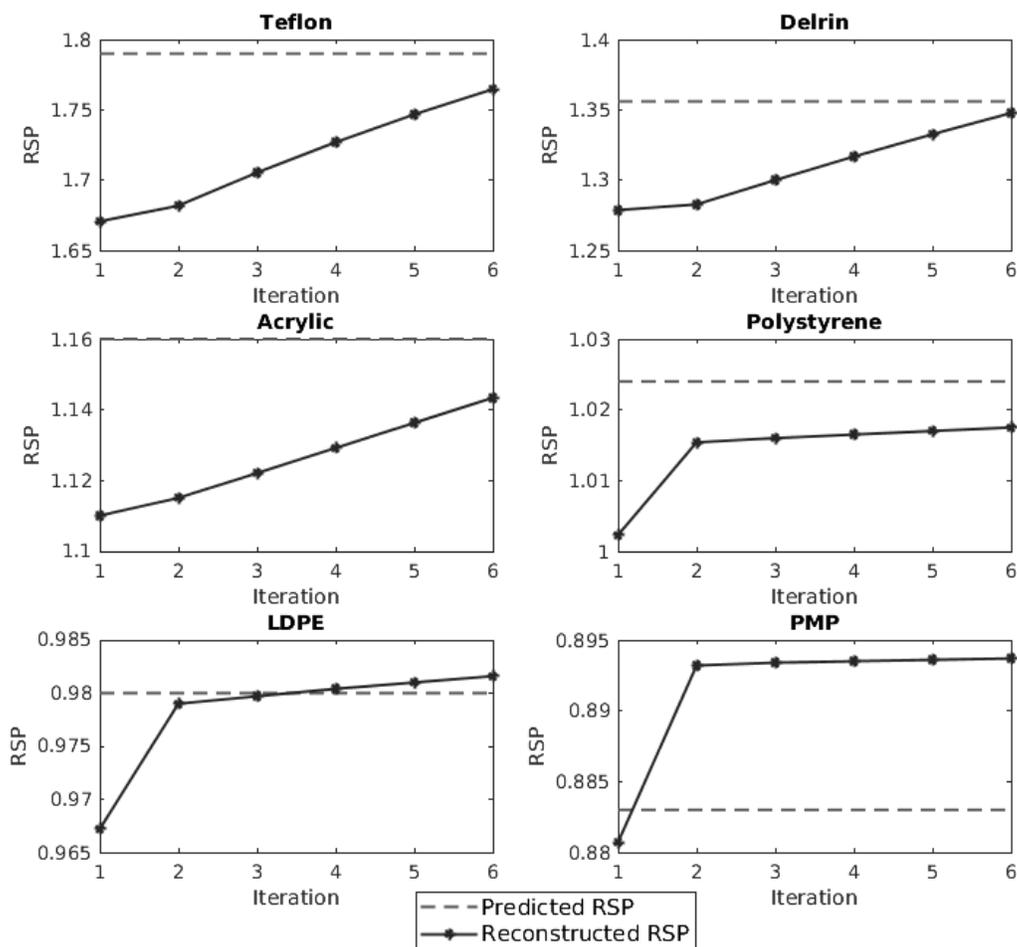


Figure 4.18: FSAIS convergence analysis values after removing protons within a range of $0.0^\circ - 40.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials are very good and close to the predicted RSP values after six iterations. See Table 4.16 for the % discrepancy of materials at the 6th iteration.

Table 4.16: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 40.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8937 ± 0.0071	1.21
LDPE	0.9816 ± 0.012	0.16
Polystyrene	1.0175 ± 0.0018	-0.64
Acrylic	1.1434 ± 0.0081	-1.43
Delrin	1.3477 ± 0.0018	-0.62
Teflon	1.7646 ± 0.0185	-1.42

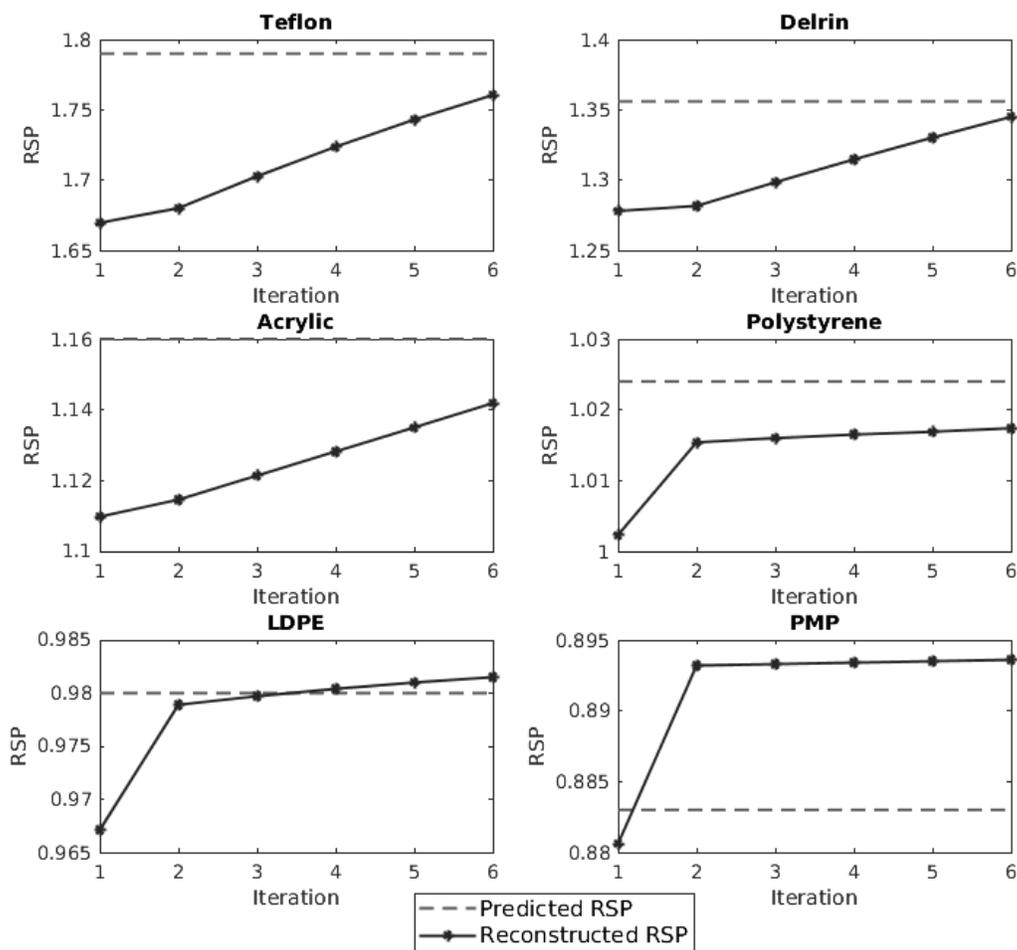


Figure 4.19: FSAIS convergence analysis after removing protons within a range of $0.0^\circ - 50.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials except Teflon and Acrylic are very close to the predicted RSP values after six iterations. See Table 4.17 for the % discrepancy of materials at the 6th iteration.

Table 4.17: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 50.0^\circ$ during reconstruction of experimental CTP404.

Material	Mean \pm StdDev	% Discrepancy
PMP	0.8936 ± 0.0071	1.20
LDPE	0.9815 ± 0.012	0.15
Polystyrene	1.0174 ± 0.0018	-0.65
Acrylic	1.1419 ± 0.008	-1.56
Delrin	1.345 ± 0.0017	-0.81
Teflon	1.7605 ± 0.0184	-1.65

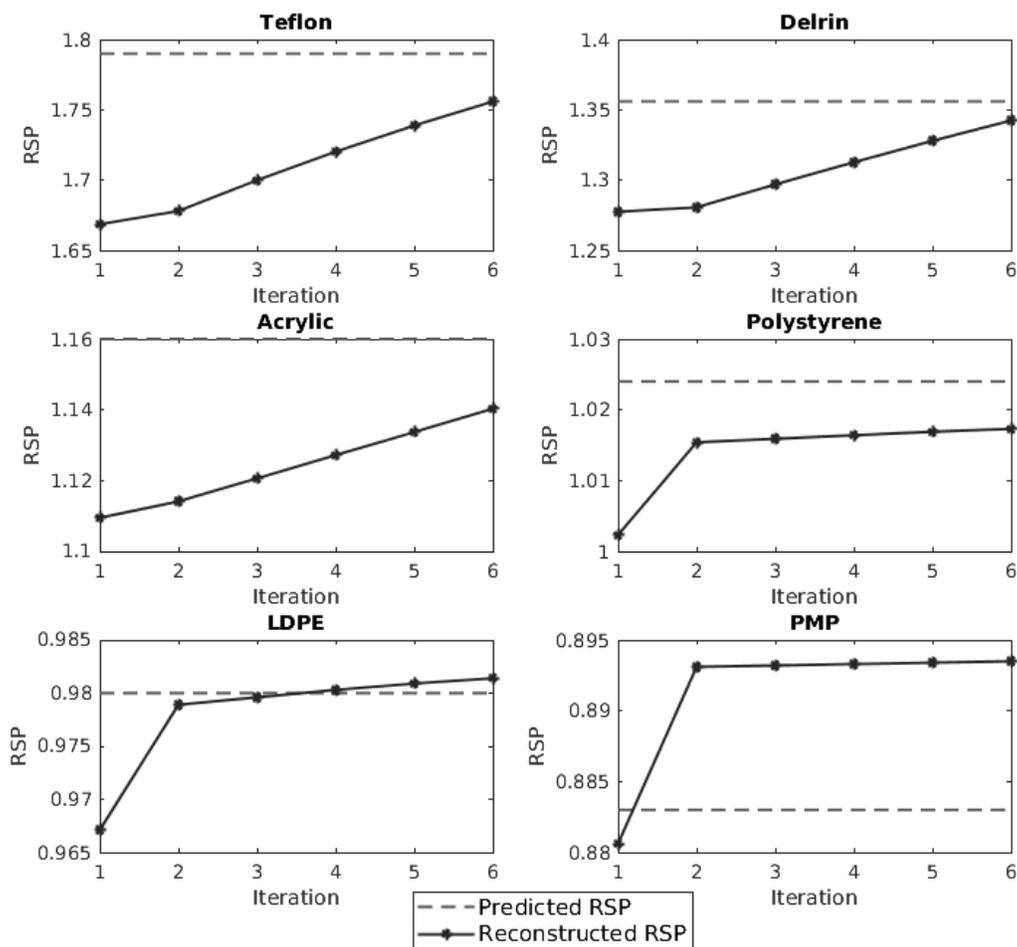
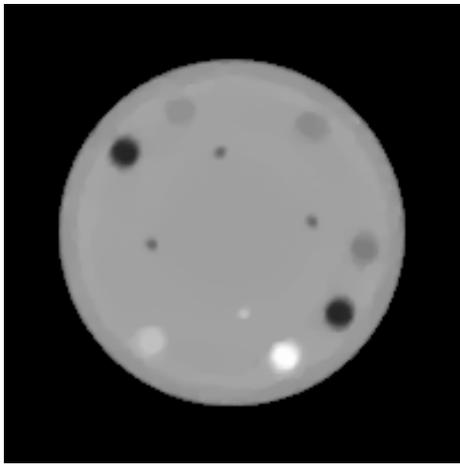


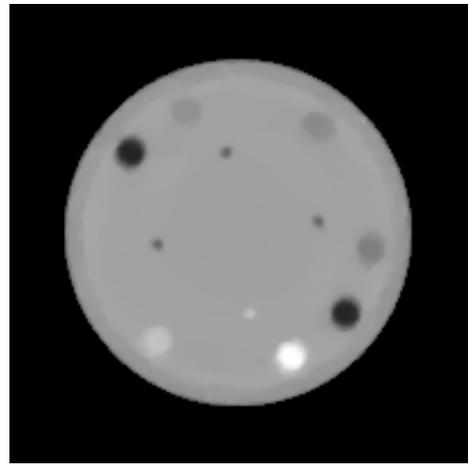
Figure 4.20: FSAIS convergence analysis after removing protons within a range of $0.0^\circ - 60.0^\circ$ during reconstruction of experimental CTP404. The RSP values of all materials except Teflon and Acrylic are very close to the predicted RSP values after six iterations. See Table 4.18 for the % discrepancy of materials at the 6th iteration.

Table 4.18: FSAIS generated RSP values after removing protons within a range of $0.0^\circ - 60.0^\circ$ during reconstruction of experimental CTP404.

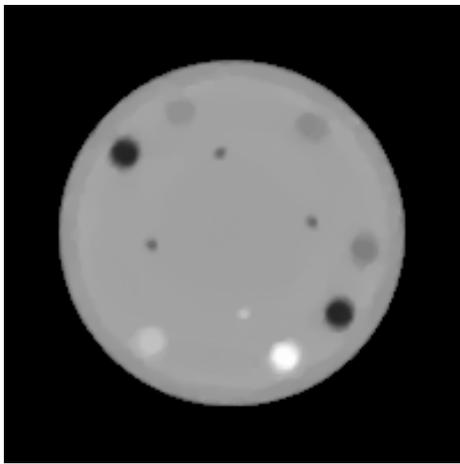
Material	Mean \pm StdDev	% Discrepancy
PMP	0.8935 ± 0.0071	1.19
LDPE	0.9814 ± 0.0121	0.14
Polystyrene	1.0173 ± 0.0018	-0.65
Acrylic	1.1403 ± 0.0078	-1.70
Delrin	1.3424 ± 0.0018	-1.00
Teflon	1.7559 ± 0.0182	-1.91



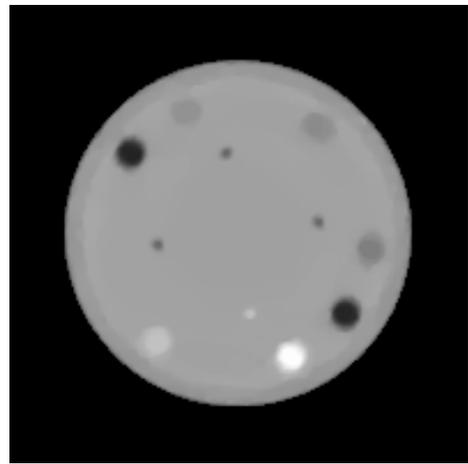
(a) $0.0^\circ - 10.0^\circ$



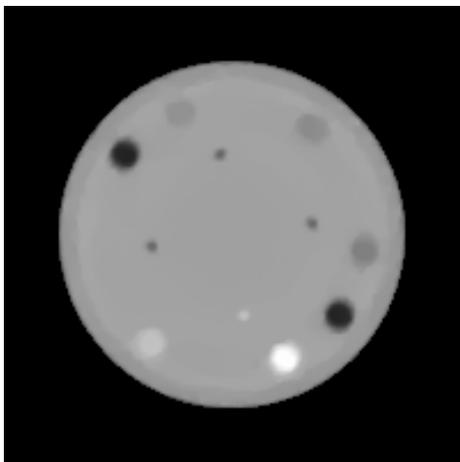
(b) $0.0^\circ - 20.0^\circ$



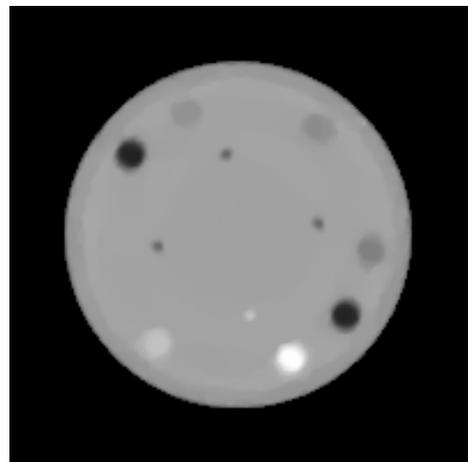
(c) $0.0^\circ - 30.0^\circ$



(d) $0.0^\circ - 40.0^\circ$



(e) $0.0^\circ - 50.0^\circ$



(f) $0.0^\circ - 60.0^\circ$

Figure 4.21: Experimental CTP404 pCT reconstruction by FSAIS while removing protons during image reconstruction. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the image reconstruction.

Figures 4.22, 4.23 and 4.24 illustrate the results of removing protons from different angle intervals with fixed length of 6° during the image reconstruction. Based on the results in these figures, the RSPs of Teflon, Delrin, and Acrylic reconstructed by DROP are even worse than the RSPs of these materials at the initial iterate.

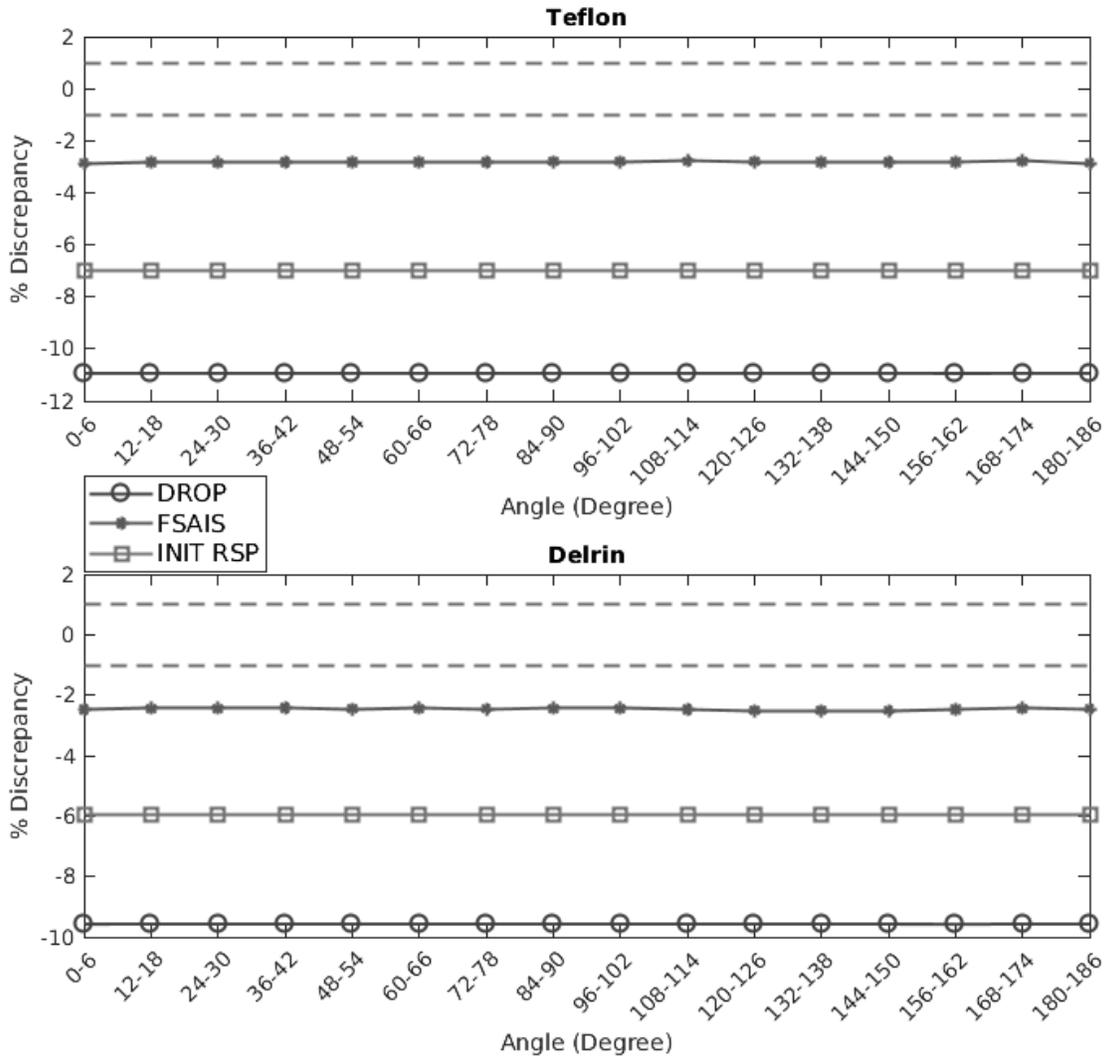


Figure 4.22: % Discrepancy of Teflon and Delrin after 6 iterations of FSAIS and DROP. Each point on the horizontal axis represents the angle interval that protons were removed from during performing the iterative solver. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. An important observation related to Teflon and Delrin is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate.

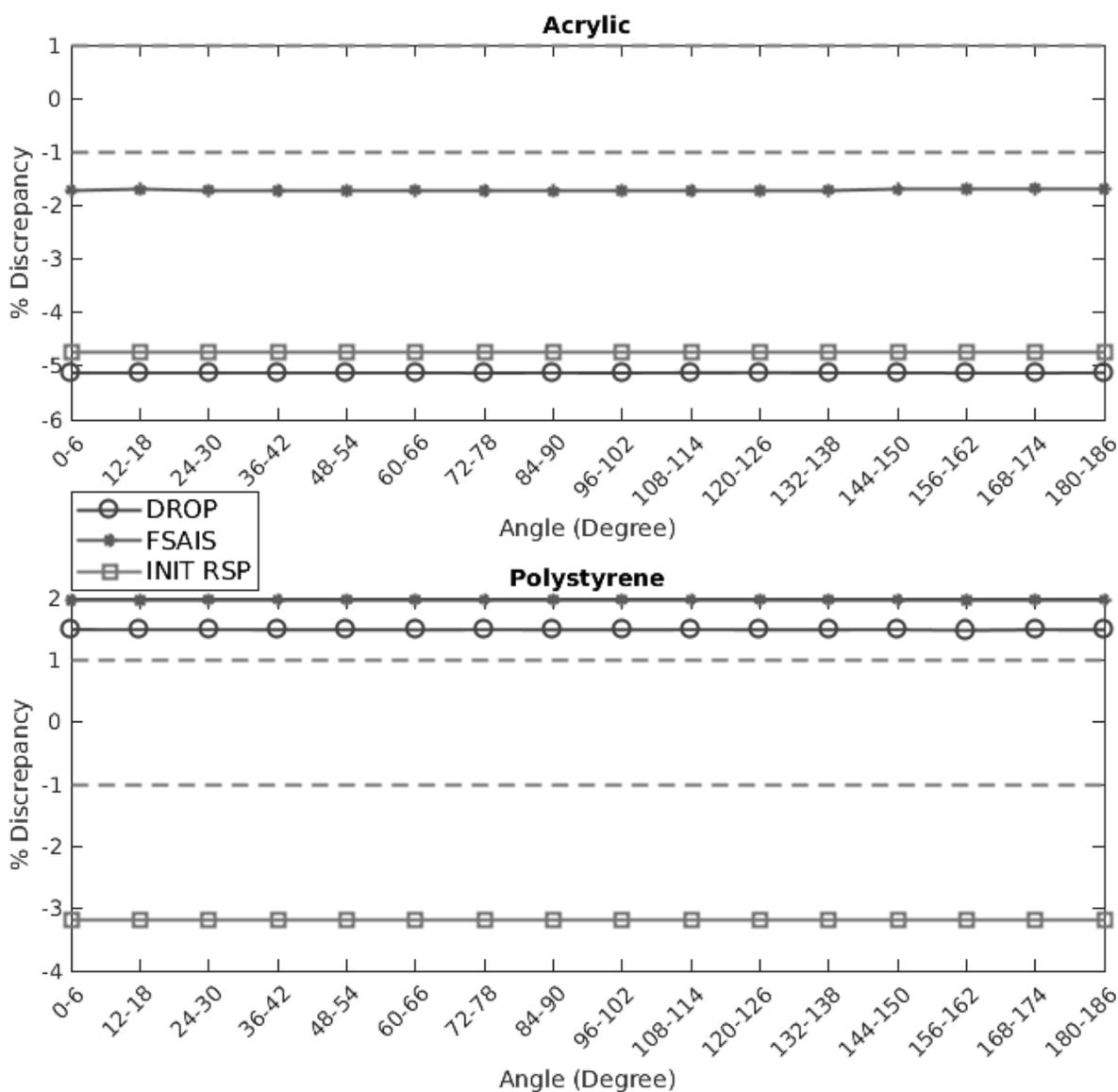


Figure 4.23: % Discrepancy of Acrylic and Polystyrene after 6 iterations of FSAIS and DROP. Each point on the horizontal axis represents the angle interval that protons were removed from during performing the iterative solver. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. An important observation related to Acrylic is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate.

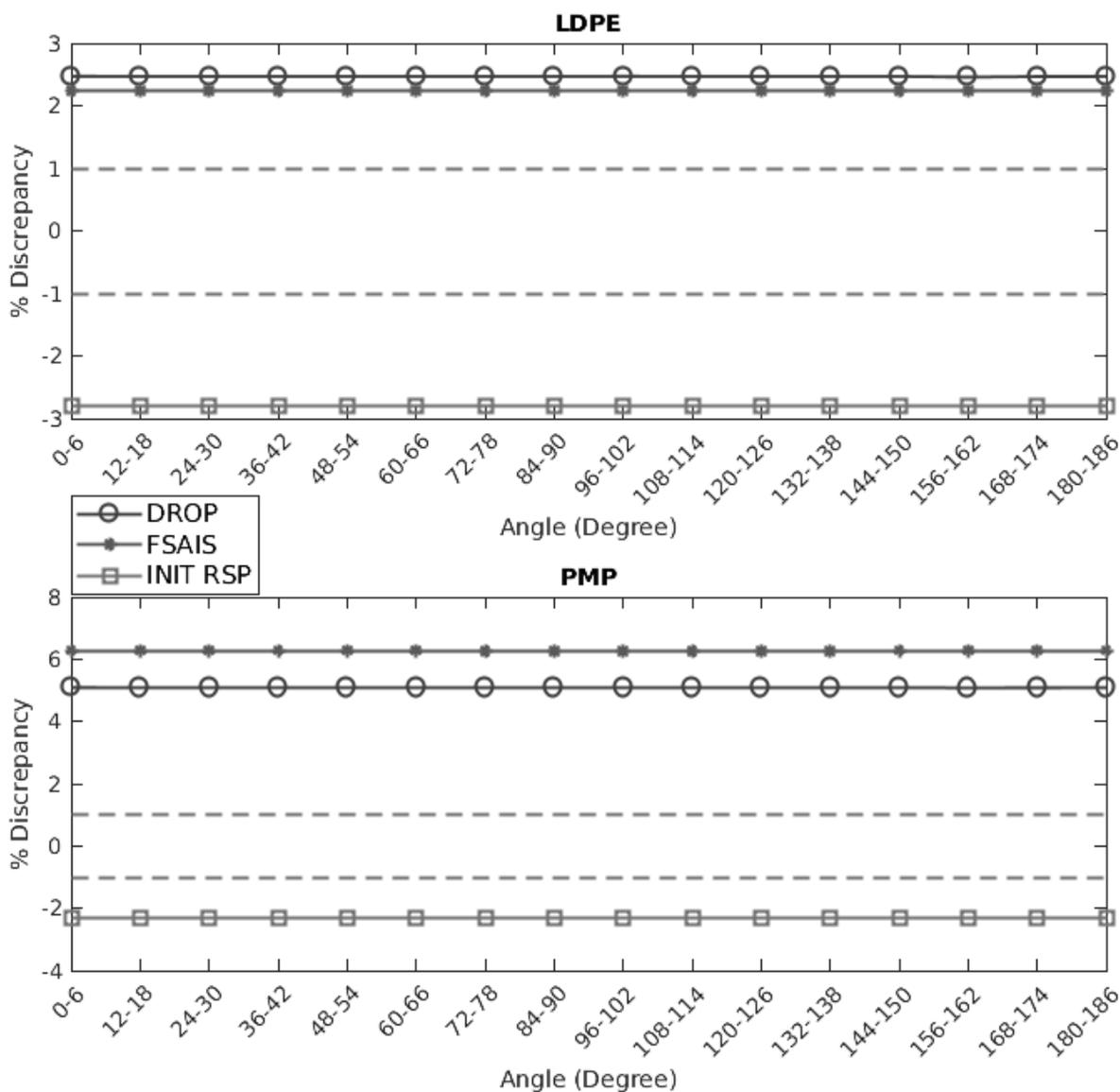


Figure 4.24: % Discrepancy of LDPE and PMP after 6 iterations of FSAIS and DROp. Each point on the horizontal axis represents the angle interval that protons were removed from during performing the iterative solver. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. Both algorithm did not generate good RSP values for LDPE and PMP.

4.4.2 Removing Angle Intervals During Data Read

Figure 4.25 shows the filtered sinogram of experimental CTP404 where protons were removed during the data read. In general, when protons are removed during data read, the FBP image is affected severely with significant noise appearing in the images. The reason that FBP images are affected by noise, arises from the fact that when protons are removed from specific angle intervals during data read, the binning process, which occurs prior to constructing the sinogram, is impacted by missing protons. The protons that are missed from specific angle intervals cause some bins are empty, which later appear in the form of lines with all zero elements in the sinogram image as shown in Figure 4.25. Figure 4.26 illustrates the result of removing protons from specific angle intervals during data read and its effect on the % discrepancy of RSP values of different materials.

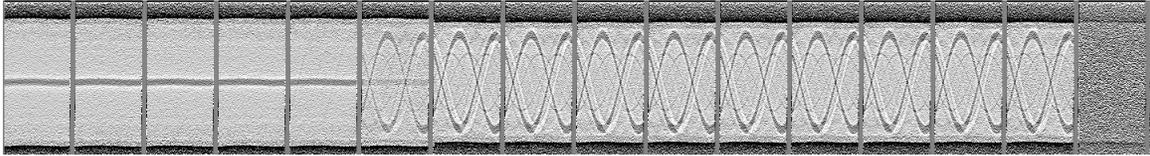


Figure 4.25: Filtered sinogram while protons were removed within a range of $0.0^\circ - 30.0^\circ$. The values of pixels in the vertical lines are zero indicate the empty bins that are generated during the binning process.

Based on the results in Figure 4.26, reconstructed RSP values of all materials except Polystyrene were less affected by the imperfect FBP images when using FSAIS. On the other hand, RSP values generated by DROP were significantly affected by imperfect FBP images. Since the reconstructed images by FSAIS and DROP used the corrupted FBP images, and therefore the final pCT images included significant noise, a median filter of radius 3 was applied to the FBP images to attenuate the effect of noise on the initial iterate. The pCT images correspond to the experiments in Figure 4.26, are shown in Figures 4.27 and 4.28 for DROP and FSAIS, respectively.

Figure 4.29 illustrates the result of removing protons from specific angle intervals during data read and its effect on the % discrepancy of RSPs in the different materials while a median filter of radius 3 was operated on the FBP image. Using the median filter caused improvements in the discrepancy of some materials using DROP. For example, for the case of Acrylic and LDPE, DROP generated more RSP values with a % discrepancy within a range of 1% where the median filter was used. On the other hand, FSAIS's results did not significantly improve using the median filter. Figures 4.30 and 4.31 show the pCT images of the experimental CPT404 reconstructed by DROP and FSAIS where protons were removed during data read and a median filter of radius 3 was operated on the FBP image.

Figures 4.32, 4.33 and 4.34 illustrate the results of removing protons from several angle intervals with fixed widths 6° . Based on the results in these figures, there are severe fluctuations in the RSPs before starting both algorithms, which is due to the fact that removing protons during data read causes significant noise in the FBP images which remains throughout the performance of the iterative solver. The general behavior of both DROP and FSAIS is similar, but FSAIS causes slightly more variations in the RSP variations because of its aggressive nature due to existence of the accelerator term $|1 - x(k)|$ compared with DROP, which performs an average of the projections within the same block of linear equations.

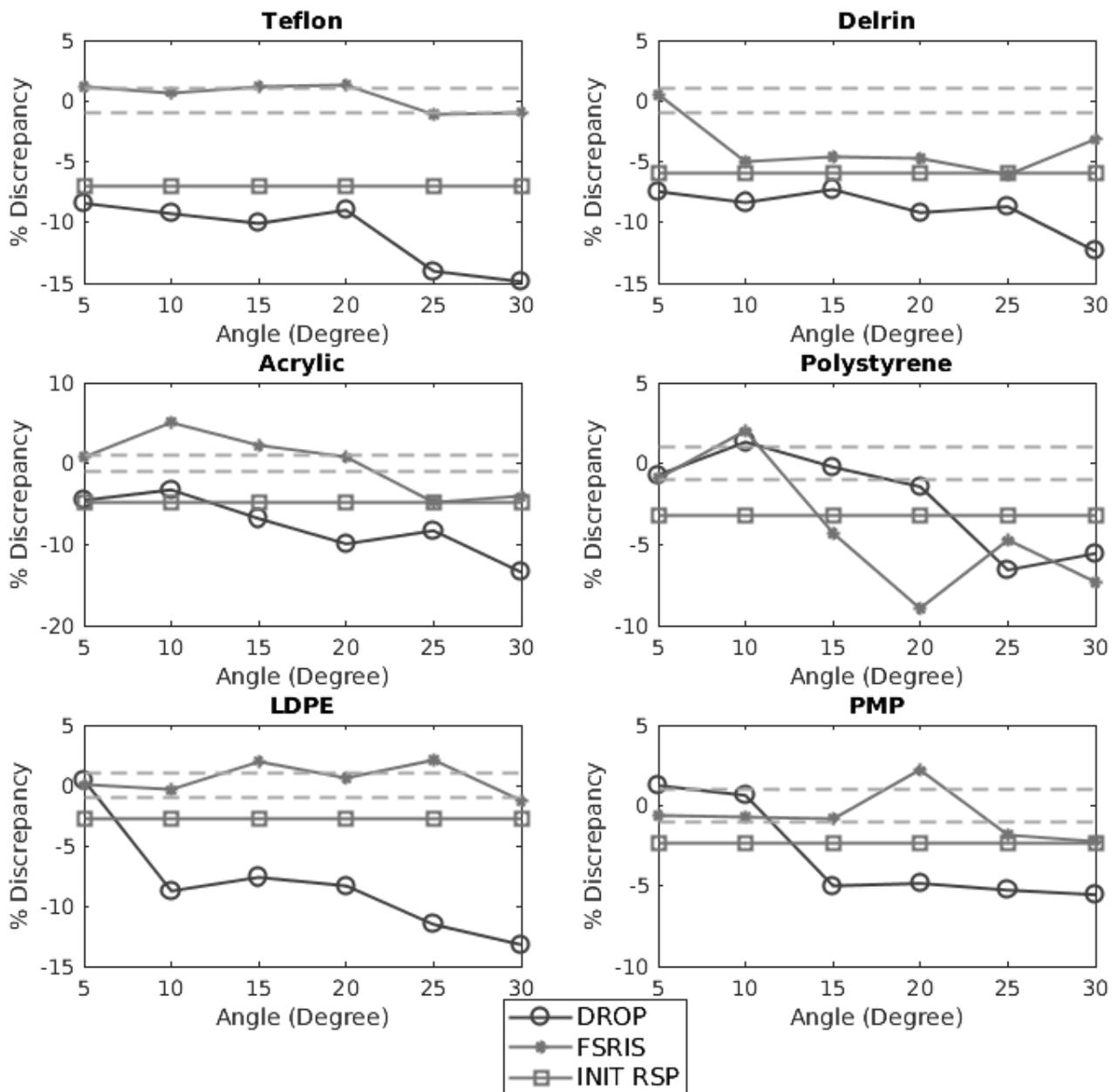
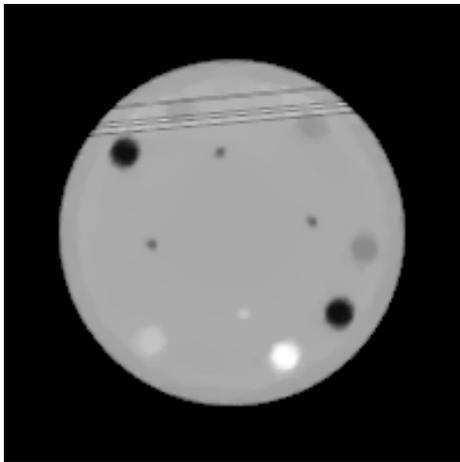
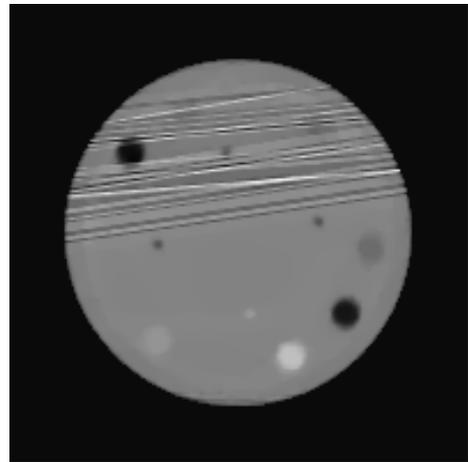


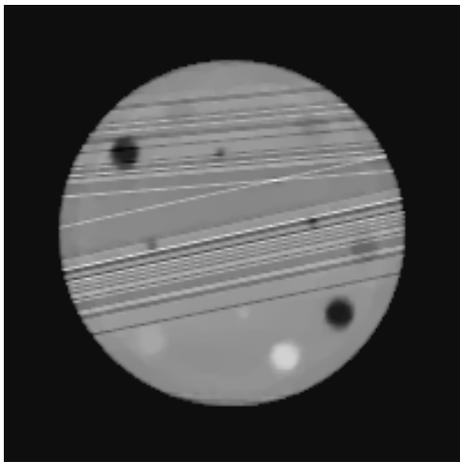
Figure 4.26: % Discrepancy in the RSP for material inserts in CTP404 while removing protons during data read. Each angle d on the horizontal axis represents the upper bound of the angle interval from which protons were removed (lower bound is zero). Teflon, Acrylic, LDPE and PMP show superior RSP accuracy for different angle intervals compared with DROP's reconstructed RSP values. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. An important observation related to Teflon, Delrin, Acrylic, LDPE, and PMP is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate.



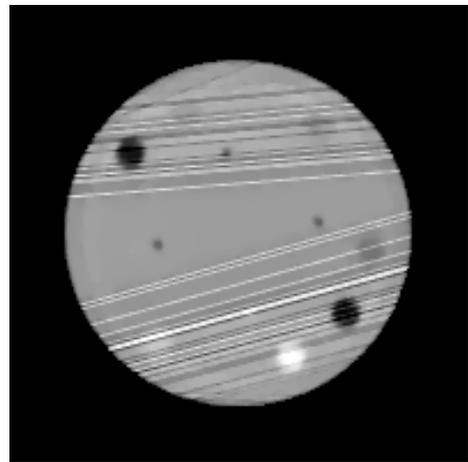
(a) $0.0^\circ - 5.0^\circ$



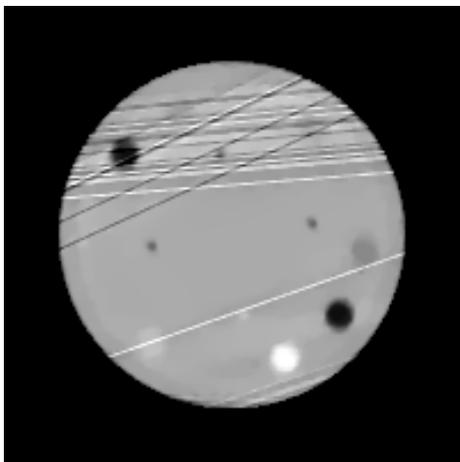
(b) $0.0^\circ - 10.0^\circ$



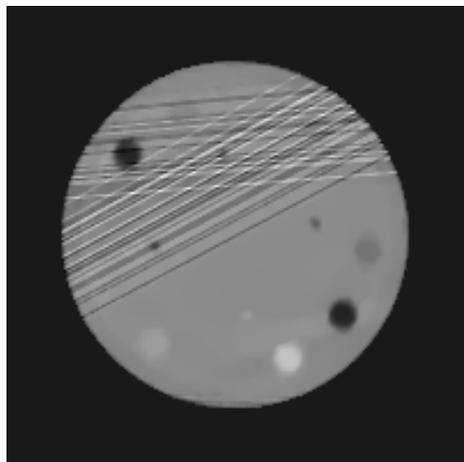
(c) $0.0^\circ - 15.0^\circ$



(d) $0.0^\circ - 20.0^\circ$

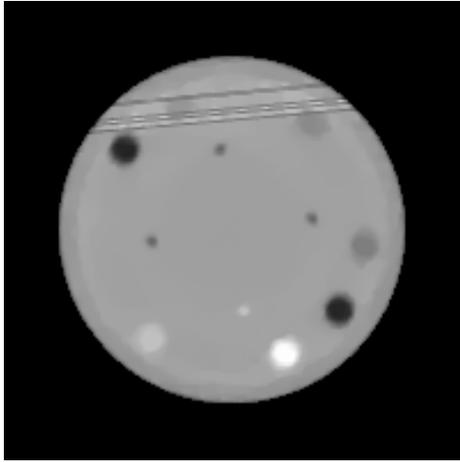


(e) $0.0^\circ - 25.0^\circ$

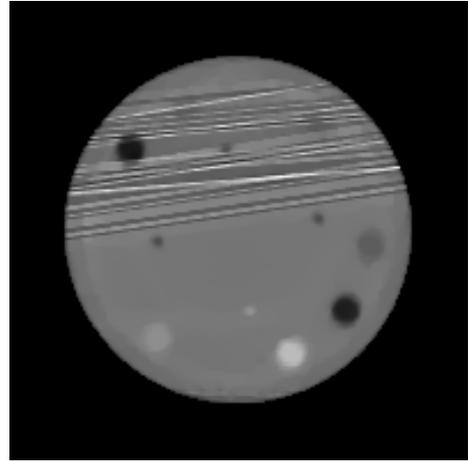


(f) $0.0^\circ - 30.0^\circ$

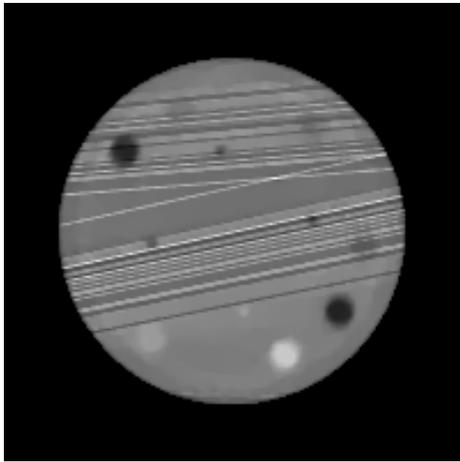
Figure 4.27: Experimental CTP404 pCT reconstruction by DROP while removing protons during data read. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the data read.



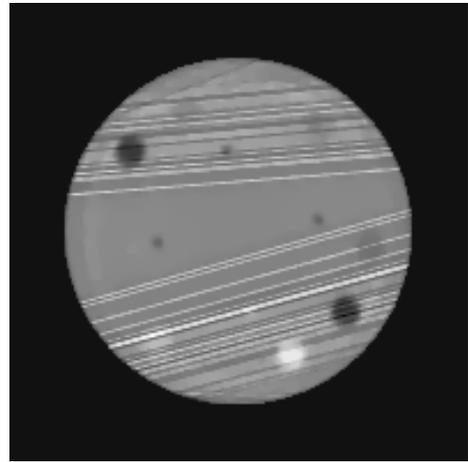
(a) $0.0^\circ - 5.0^\circ$



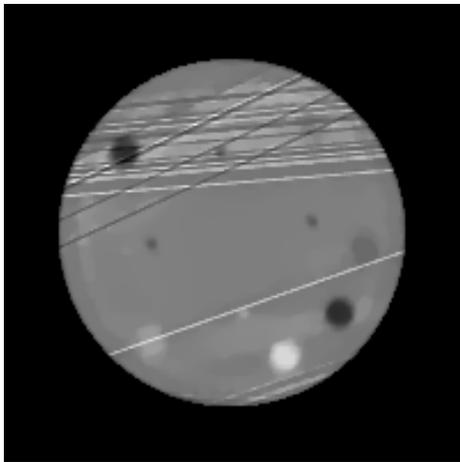
(b) $0.0^\circ - 10.0^\circ$



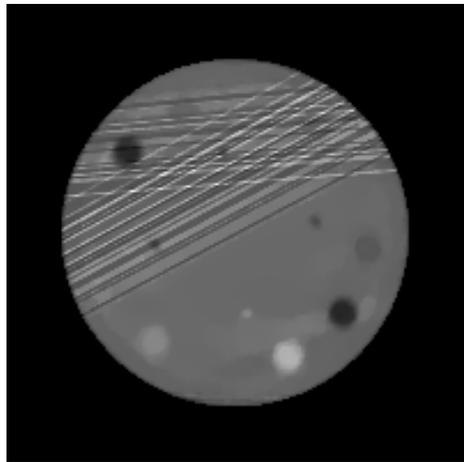
(c) $0.0^\circ - 15.0^\circ$



(d) $0.0^\circ - 20.0^\circ$



(e) $0.0^\circ - 25.0^\circ$



(f) $0.0^\circ - 30.0^\circ$

Figure 4.28: Experimental CTP404 pCT reconstruction by FSAIS while removing protons during data read. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the data read.

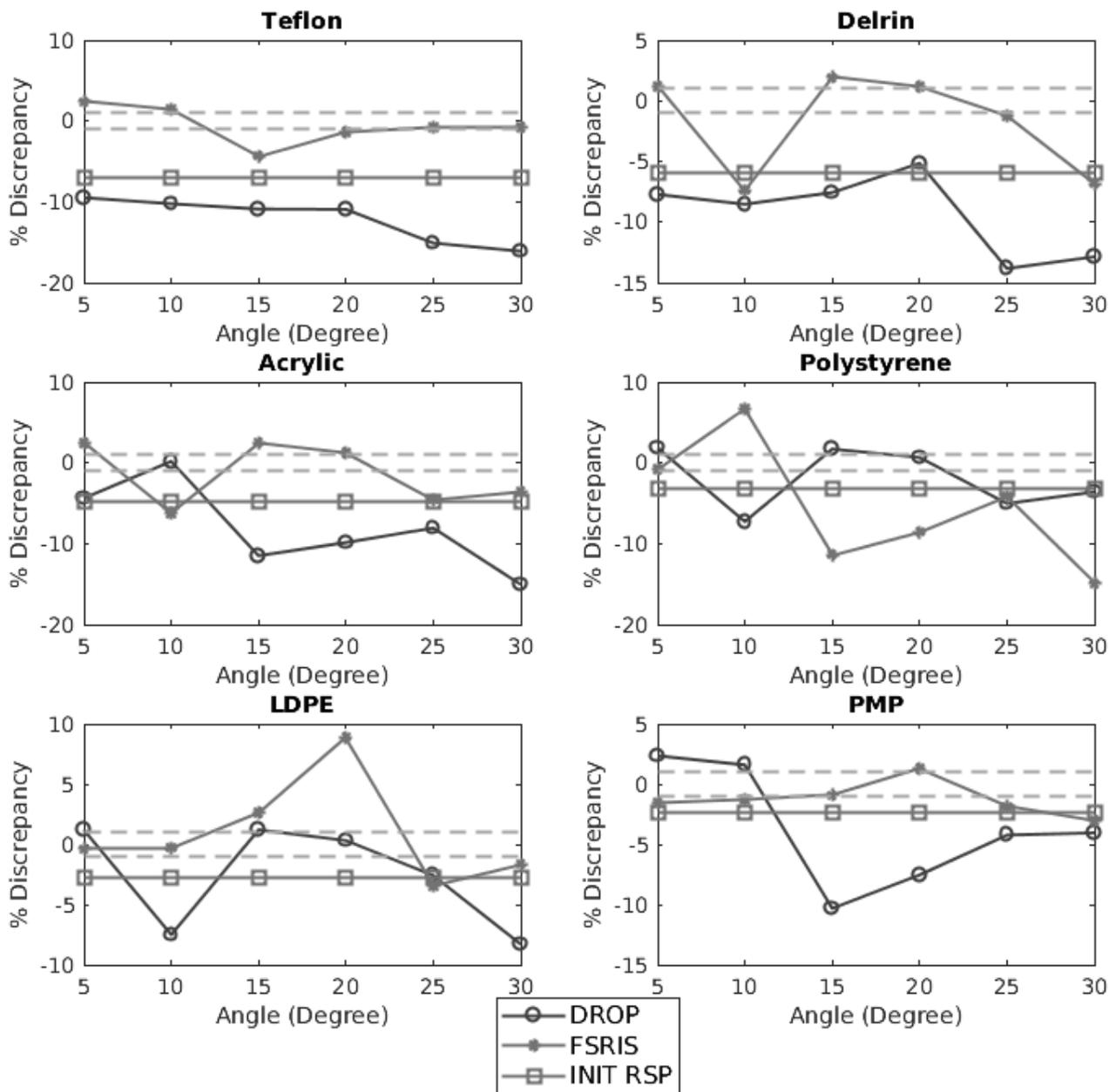
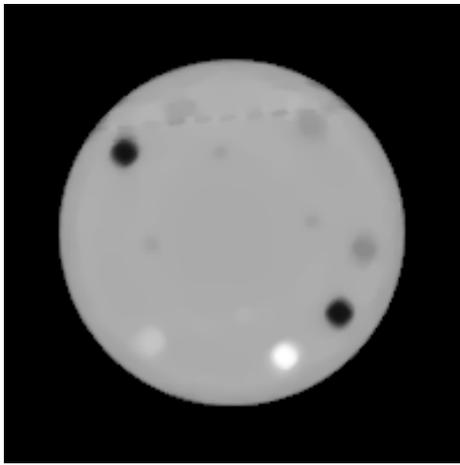
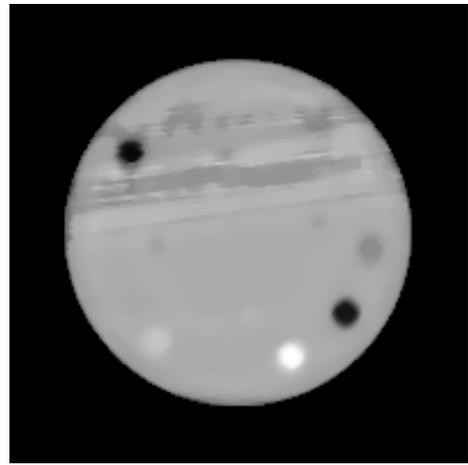


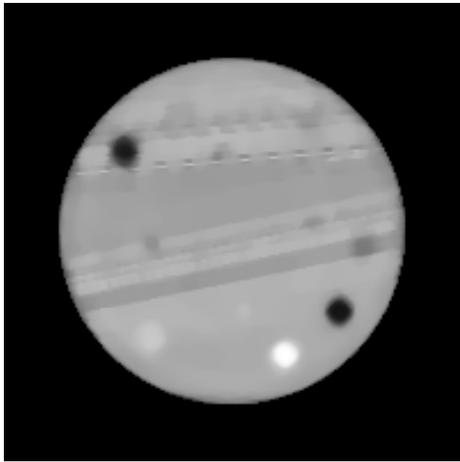
Figure 4.29: % Discrepancy of material inserts in CTP404 while removing protons during data read and using a median filter of radius 3 for FBP. Each angle d on the horizontal axis represents the upper bound of the angle interval from which protons were removed (lower bound is zero). Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. An important observation related to Teflon, Delrin, Acrylic, LDPE, and PMP is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate.



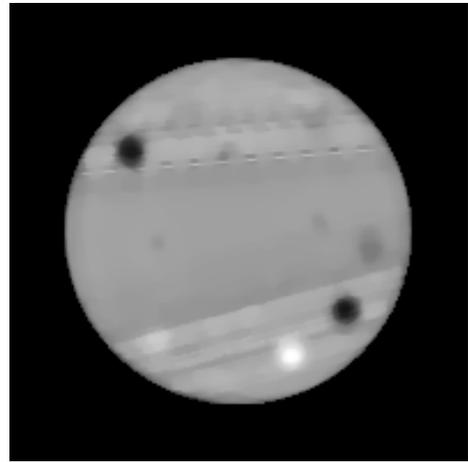
(a) $0.0^\circ - 5.0^\circ$



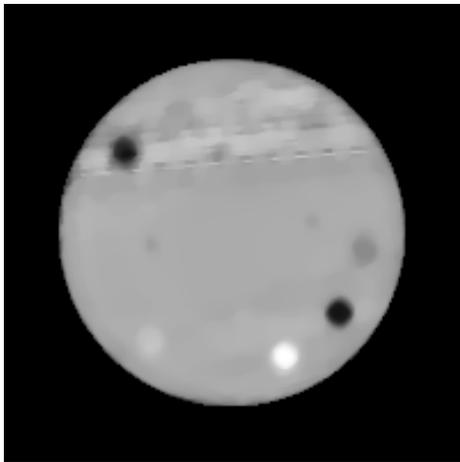
(b) $0.0^\circ - 10.0^\circ$



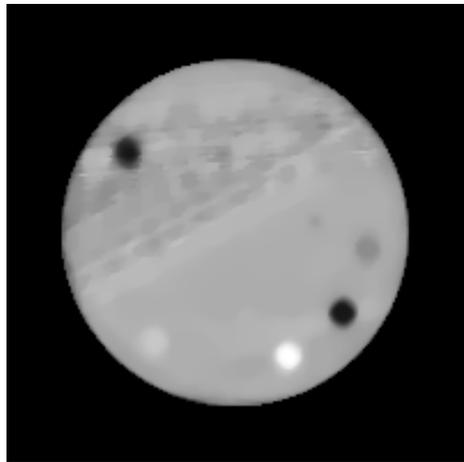
(c) $0.0^\circ - 15.0^\circ$



(d) $0.0^\circ - 20.0^\circ$

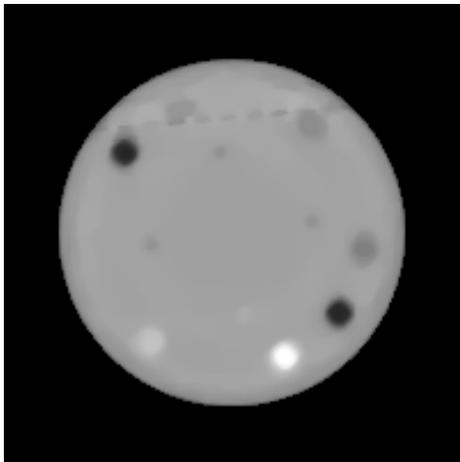


(e) $0.0^\circ - 25.0^\circ$

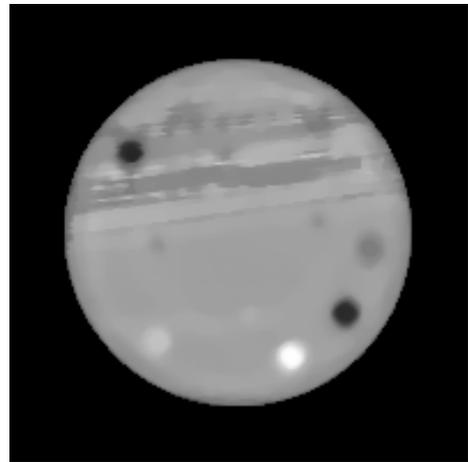


(f) $0.0^\circ - 30.0^\circ$

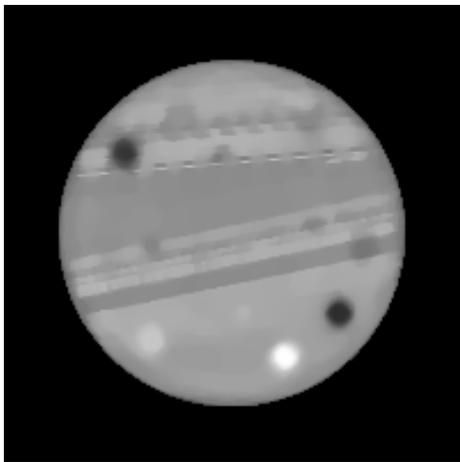
Figure 4.30: Experimental CTP404 pCT reconstruction by DROP while removing protons during data read and using a median filter of radius 3 for FBP. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the data read.



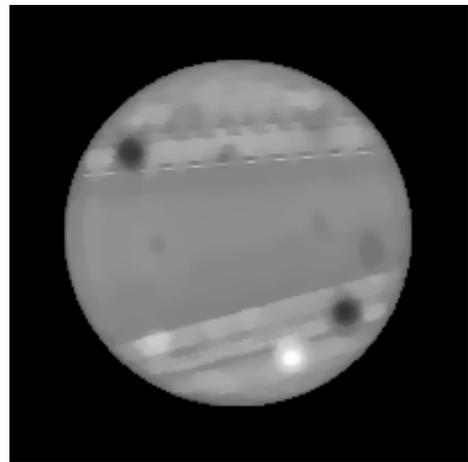
(a) $0.0^\circ - 5.0^\circ$



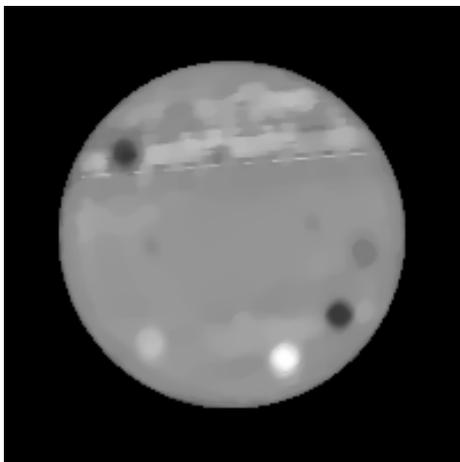
(b) $0.0^\circ - 10.0^\circ$



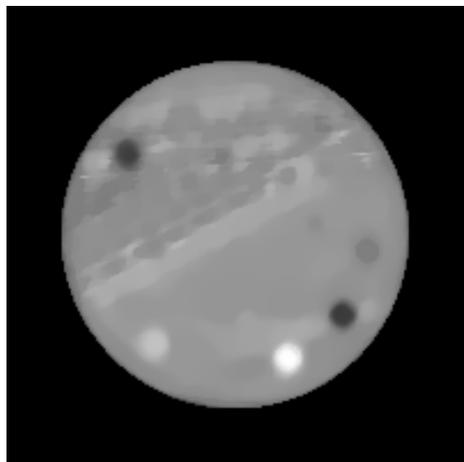
(c) $0.0^\circ - 15.0^\circ$



(d) $0.0^\circ - 20.0^\circ$



(e) $0.0^\circ - 25.0^\circ$



(f) $0.0^\circ - 30.0^\circ$

Figure 4.31: Experimental CTP404 pCT reconstruction by FSAIS while removing protons during data read and using a median filter of radius 3 for FBP. Each angle interval $0.0^\circ - d^\circ$ represents the angle interval that is removed during the data read.

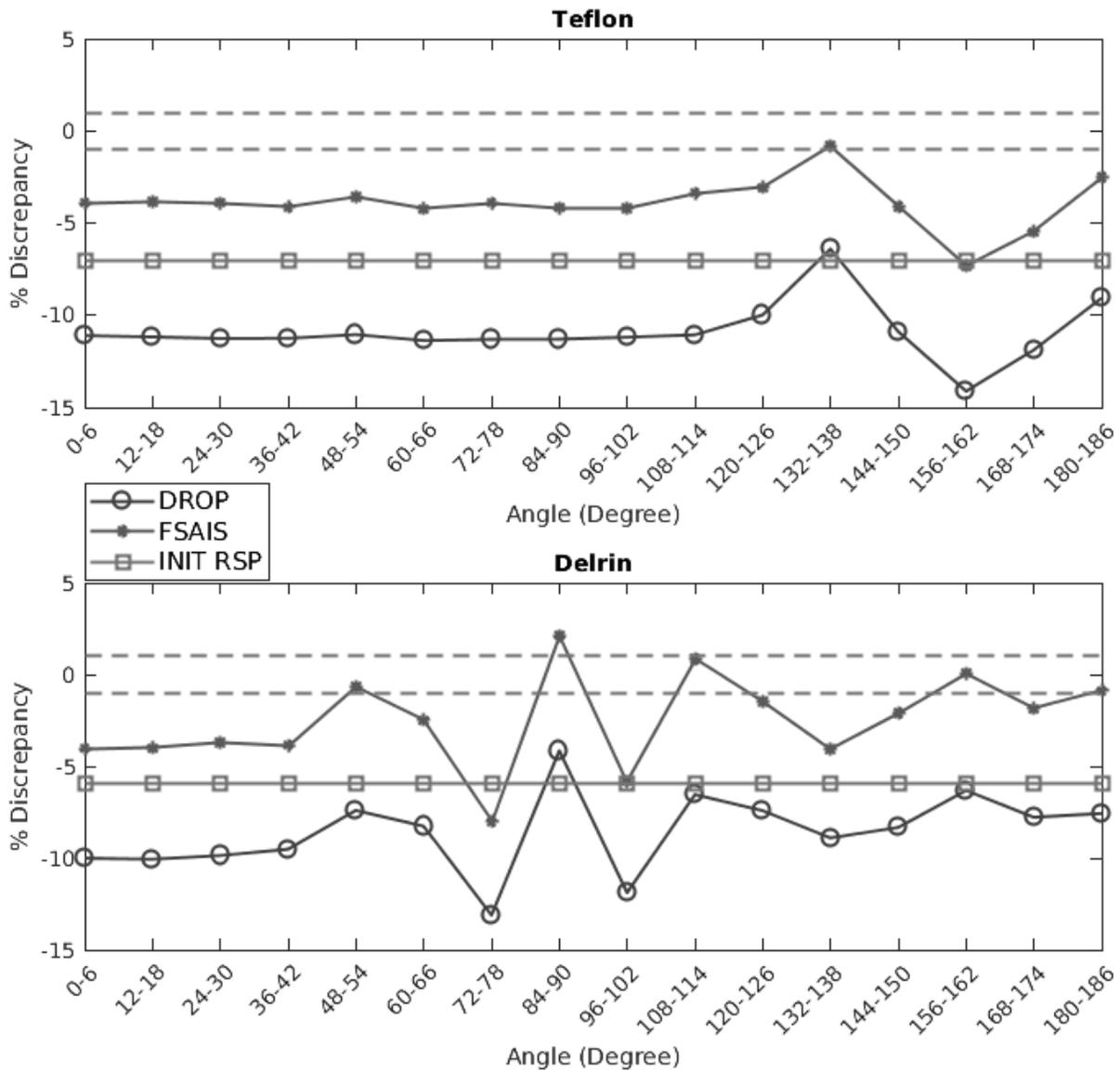


Figure 4.32: % Discrepancy in the RSP for Teflon and Delrin after six iterations of FSAIS and DROP. Each point on the horizontal axis represents the angle interval from which protons were removed during data read. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. An important observation related to Teflon and Delrin is that DROP generated very inaccurate RSP values, even worse than RSP values at the initial iterate.

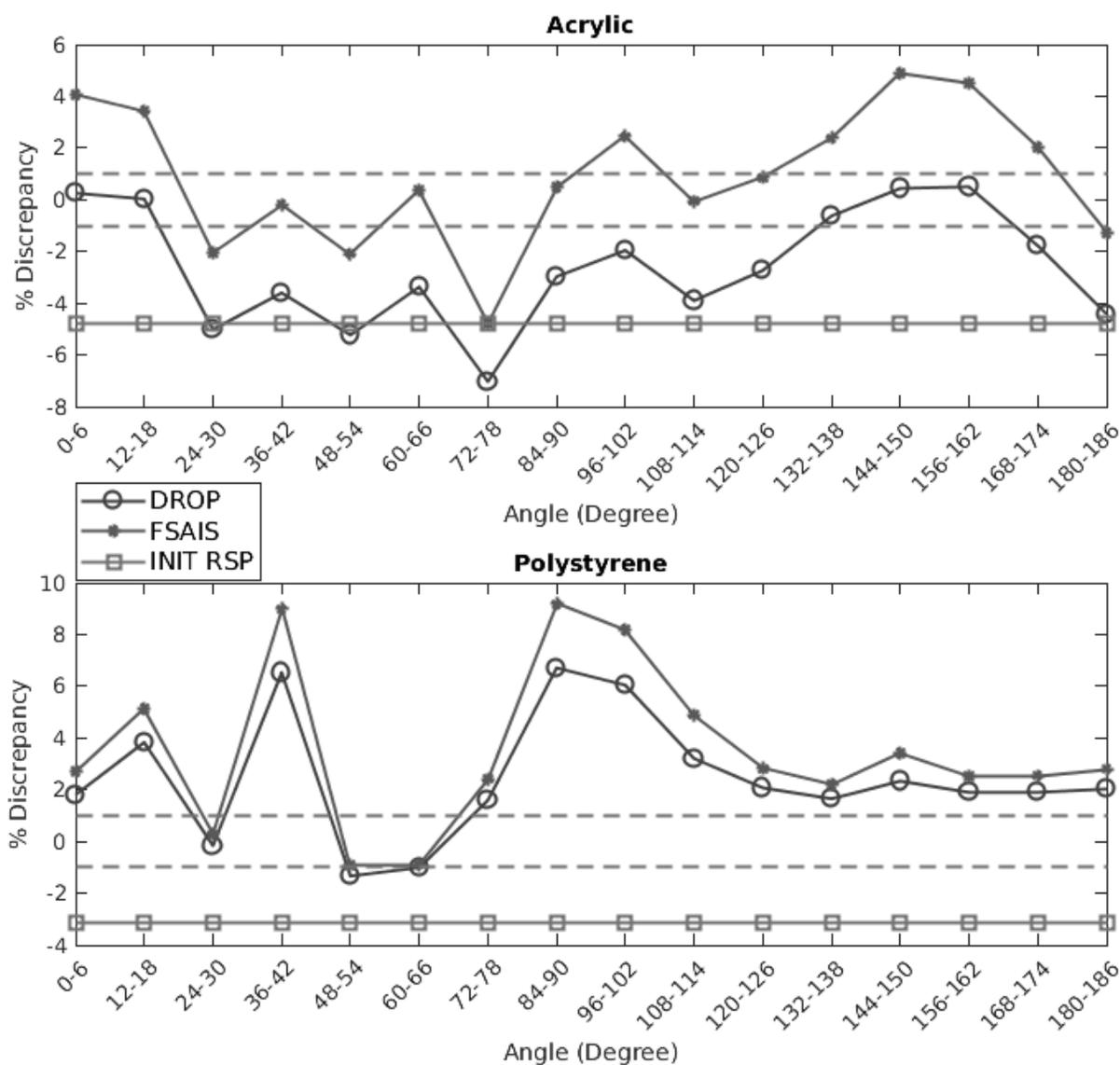


Figure 4.33: % Discrepancy in the RSP for Acrylic and Polystyrene after six iterations of FSAIS and DROP. Each point on the horizontal axis represents the angle interval from which protons were removed during data read. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. Based on the results, FSAIS does a better job of generating good RSPs for Acrylic, but for the case of Polystyrene, results of both algorithms are very similar.

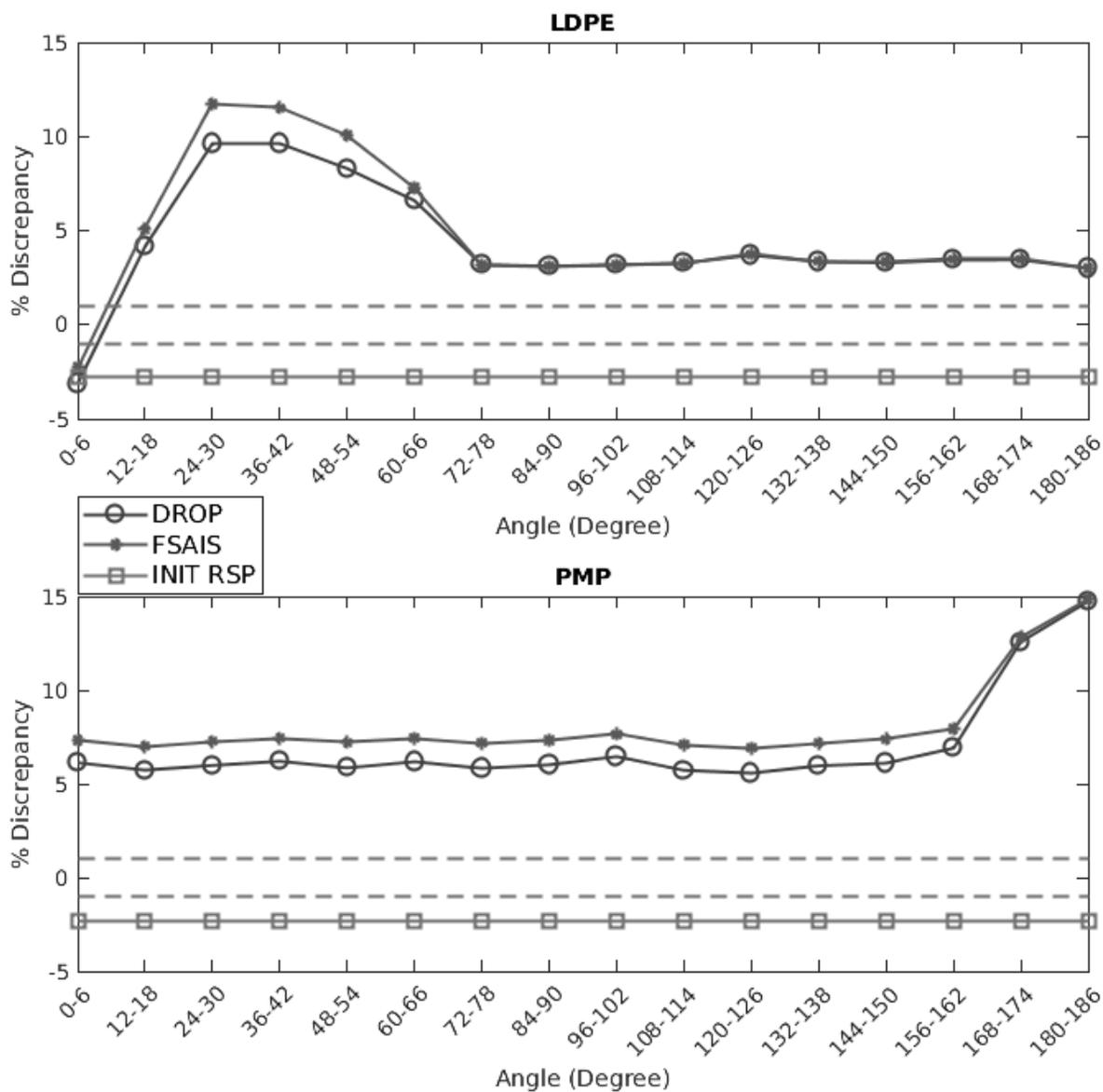


Figure 4.34: % Discrepancy in the RSP for LDPE and PMP after six iterations of FSAIS and DROP. Each point on the horizontal axis represents the angle interval form which protons were removed during data read. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. Based on the results, both algorithm generated bad RSP values for LDPE and PMP with very similar values.

4.5 Convergence Analysis of FSAIS and DROP on Pediatric Head Phantom

Figures 4.35 (a)-(f), show the slices of the pediatric head phantom reconstructed with DROP and FSAIS after six iterations. The block size and relaxation parameters used in DROP were 1280000 and 0.0015 respectively. Moreover, the TVS method is used during all six iterations of these two iterative solvers. Total reconstruction runtime of pediatric head phantom on a single K40 GPU was 9.6 and 9.1 minutes for DROP and FSAIS respectively. The image analysis program ImageJ2 was used to perform quantitative analyses of reconstructed image quality of the slice numbers 19, and 39 out of the total 90 slices overall. Tooth RSP value was selected from slice 19, while soft tissue, brain tissue, and Trabecular bone were selected from slice 39. The ovalar and polygon selection and measurement tools from ImageJ were used to select ROI and calculate the mean and standard deviation in reconstructed RSP, with identical region selection and analyses performed for all images reconstructed with FSAIS and DROP. The spinal disk in the pCT images of the pediatric head phantom is visibly indistinguishable and therefore not included in the analysis. There was another challenge related to selecting the ROI for tooth materials; Dentin, and Enamel. Since tooth Dentin and Enamel were also visibly indistinguishable, a single ROI region correspond to tooth was selected and for analysis purpose, it was assumed that the ROI could either be composed of Dentin or Enamel. Table 4.19 shows the distribution of RSP values of different materials inside the pediatric head phantom, and Table 4.20 illustrates the elements of Ψ^{-1} used by FSAIS.

Table 4.19: RSP statistics of materials at the initial iterate of the experimental pediatric head phantom.

Material	Predicted	Mean	StdDev	Min.	Max.	% Discrepancy
Soft tissue	1.037	0.979	0.0642	0.6648	1.1588	-5.59
Brain tissue	1.047	0.9778	0.055	0.8264	1.124	-6.61
Trabecular bone	1.108	1.0229	0.0582	0.886	1.1673	-7.68
Dentin	1.513	1.3841	0.0851	1.1748	1.5625	-8.52
Enamel	1.788	1.3841	0.0851	1.1748	1.5625	-22.59

Table 4.20: Elements of Ψ^{-1} for the pediatric head phantom.

RSP range	$\psi^{-1}(j)$
0.87 - 0.91	$ 1 - x(k)_j \times 0.0025$
0.96 - 1.037	$ 1 - x(k)_j \times 0.0025$
1.038 - 1.35	$ 1 - x(k)_j \times 0.00075$
1.13 - 1.16	$ 1 - x(k)_j \times 0.00075$
1.35 - 1.8	$ 1 - x(k)_j \times 0.00075$

Reconstructed RSP values of different materials of this dataset are shown in Table 4.21. Based on the results in Table 4.21, both methods generate RSP values within a range of 1% of the predicted RSP for all materials except tooth, which has a very low RSP value generated by DROP. The circular regions in Figures 4.35a and 4.35c are associated with tooth composed of dentin and enamel. Tooth Dentin and Enamel are not easily distinguishable by eye but the predicted RSP values are 1.513 and 1.788 for Dentin and Enamel respectively [19]. The reconstructed RSP of tooth generated by FSAIS reported in Table 4.21 is much closer to the predicted RSP values of dentin and enamel while the RSP value of tooth reconstructed by DROP is too low. This result is similar to the results of the CTP404 in the previous section, in which the reconstructed RSP values of Teflon and Delrin generated by DROP were too low while FSAIS generated RSP values of Teflon and Delrin within a range of 1% of the predicted RSP values of these two materials. An important observation from results in Table 4.21 is that the reconstructed RSP value of brain tissue generated by FSAIS is 0.14%, while the value generated by DROP is 0.97%. This observation shows the superior behavior of FSAIS in generating accurate results for a critical and sensitive part of the head, which needs the most accurate RSP value to avoid the range uncertainty.

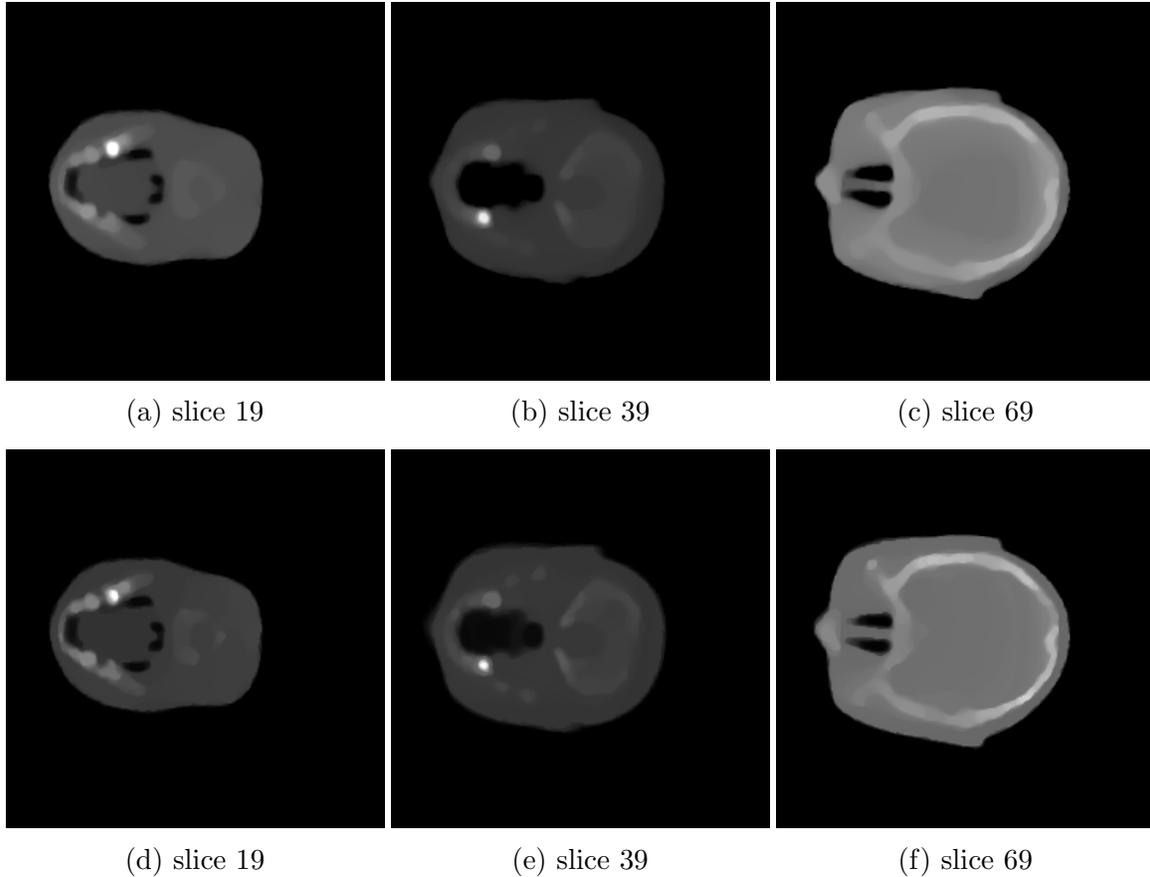


Figure 4.35: Representative pCT images of experimental pediatric head reconstructed by DROP (a)-(c) and FSAIS (d)-(f). pCT images generated by FSAIS have sharper bone areas compared with the ones generated by DROP.

Table 4.21: RSP values of experimental pediatric head phantom reconstructed by DROP and FSAIS. Both algorithms generated accurate RSP values (close to the predicted RSPs) for soft tissue, brain tissue and Trabecular bone, but FSAIS generated more accurate RSP value for Enamel.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft tis.	1.037	1.0423 \pm 0.0222	0.51	1.0432 \pm 0.0234	0.6
Brain tis.	1.047	1.0572 \pm 0.0015	0.97	1.0485 \pm 0.0025	0.14
Trab. bone	1.108	1.1077 \pm 0.0033	-0.03	1.1169 \pm 0.0117	0.8
Dentin	1.513	1.3857 \pm 0.0372	-8.41	1.7359 \pm 0.0692	14.17
Enamel	1.788	1.3857 \pm 0.0372	-22.5	1.7359 \pm 0.0692	-2.91

Figure 4.36 illustrates the result of removing protons from angle intervals $0 - d$ where d is the upper bound of each angle interval. In this experiment, protons were removed during the image reconstruction. The optimal reconstruction parameters for FSAIS were the same as reported in Table 4.20. The relaxation parameter and block sized used for DROP were 0.0015 and 1280000, respectively.

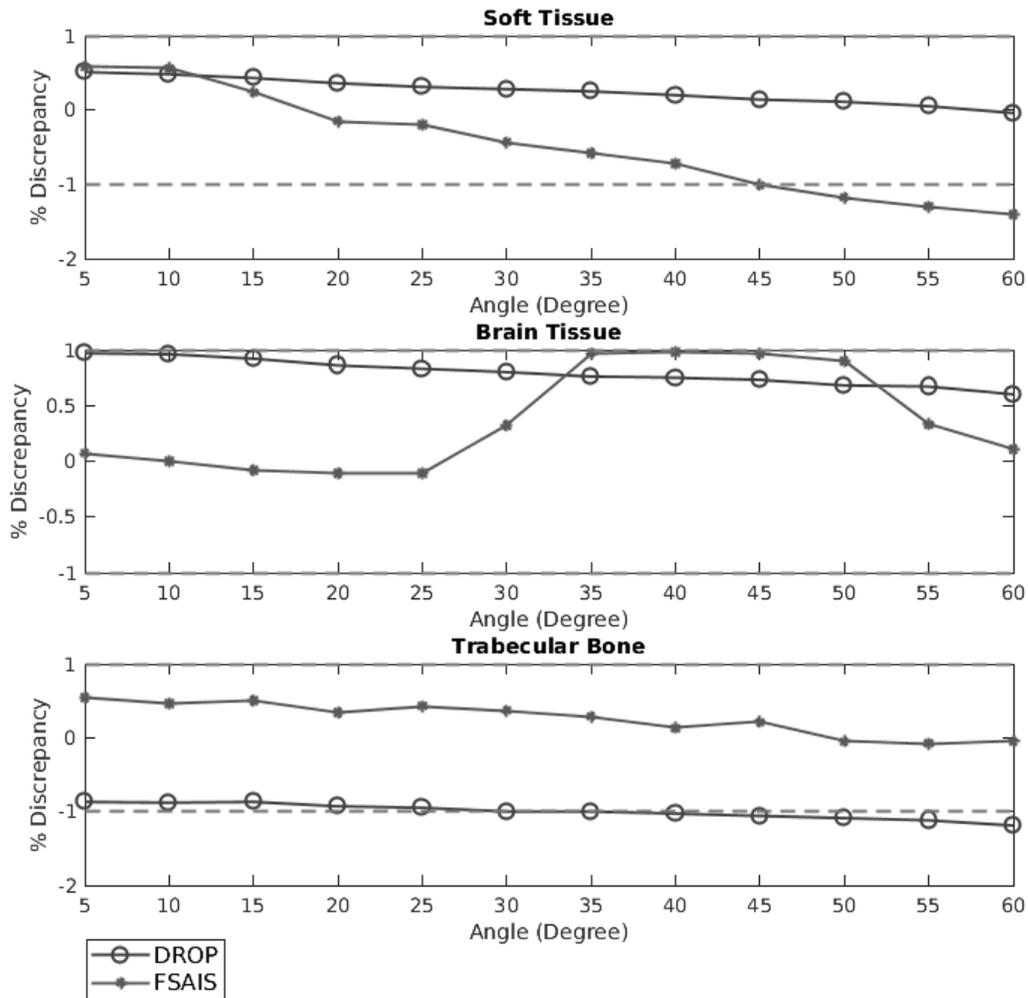


Figure 4.36: % Discrepancy in the RSP reconstructed for materials in pediatric head phantom after six iterations of FSAIS and DROP. Each number d on the horizontal axis represents the angle interval $0.0^\circ - d^\circ$ that is removed during the image reconstruction. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. Both algorithms generated accurate RSP values even where protons within a range of $0.0^\circ - 60.0^\circ$ were removed.

Based on the results in Figure 4.36, removing protons during the image reconstruction caused the % discrepancy in the RSP reconstructed for materials to move towards negative values. In the case of soft tissue, the % discrepancy changed more rapidly while using FSAIS compared with DROP. In the case of Trabecular bone, the rate of changes of % discrepancy is similar for both FSAIS and DROP, although there are more oscillations seen in FSAIS.

An important observation in Figure 4.36 is that the % discrepancy in the RSP value for brain tissue reconstructed by FSAIS is still very close to zero even after removing protons from $0.0^\circ - 25.0^\circ$. When protons were removed from $0.0^\circ - 30.0^\circ$, there was a sudden increase in the RSP discrepancy of the brain tissue generated by FSAIS, which was not expected since there were more protons that were removed (compared with the $0.0^\circ - 25.0^\circ$ case), therefore it was expected that the accuracy of RSP values decrease for all materials. The FSAIS algorithm has the accelerator term $|1 - x(k)|$, which caused the sudden change of the RSP value of brain tissue. In order to verify the effect of the accelerator term in the sudden change of the RSP value of brain tissue, we removed the accelerator term and repeated the experiment for the same angle interval $0.0^\circ - 30.0^\circ$. Based on the result of this experiment, the % discrepancy for materials in Figure 4.36 decreased to -2.54 , -2.13 , and -3.24 for soft tissue, brain tissue, and Trabecular bone, respectively. This indicates that removing the accelerator term caused the RSP discrepancy of materials decreased as expected, thus the accelerator term can cause unexpected sudden changes in the RSP values.

Figure 4.37 illustrates the results of removing protons from different angle intervals with fixed width of 6° during the data read. Based on the results and as expected, removing protons during data read causes line artifacts on the FBP images which leads to fluctuations in the RSPs during the image reconstruction by the iterative solver. Based on the results in Figure 4.37, DROP performs better than FSAIS on keeping the % discrepancy of materials closer to zero, except for the brain

tissue for which FSAIS generates significantly better results through different angle intervals. The reason FSAIS generates better results for brain tissue is related to the fact that the brain tissue is in general less affected by noise appearing in the FBP images due to its geometrical location in the pediatric head phantom.

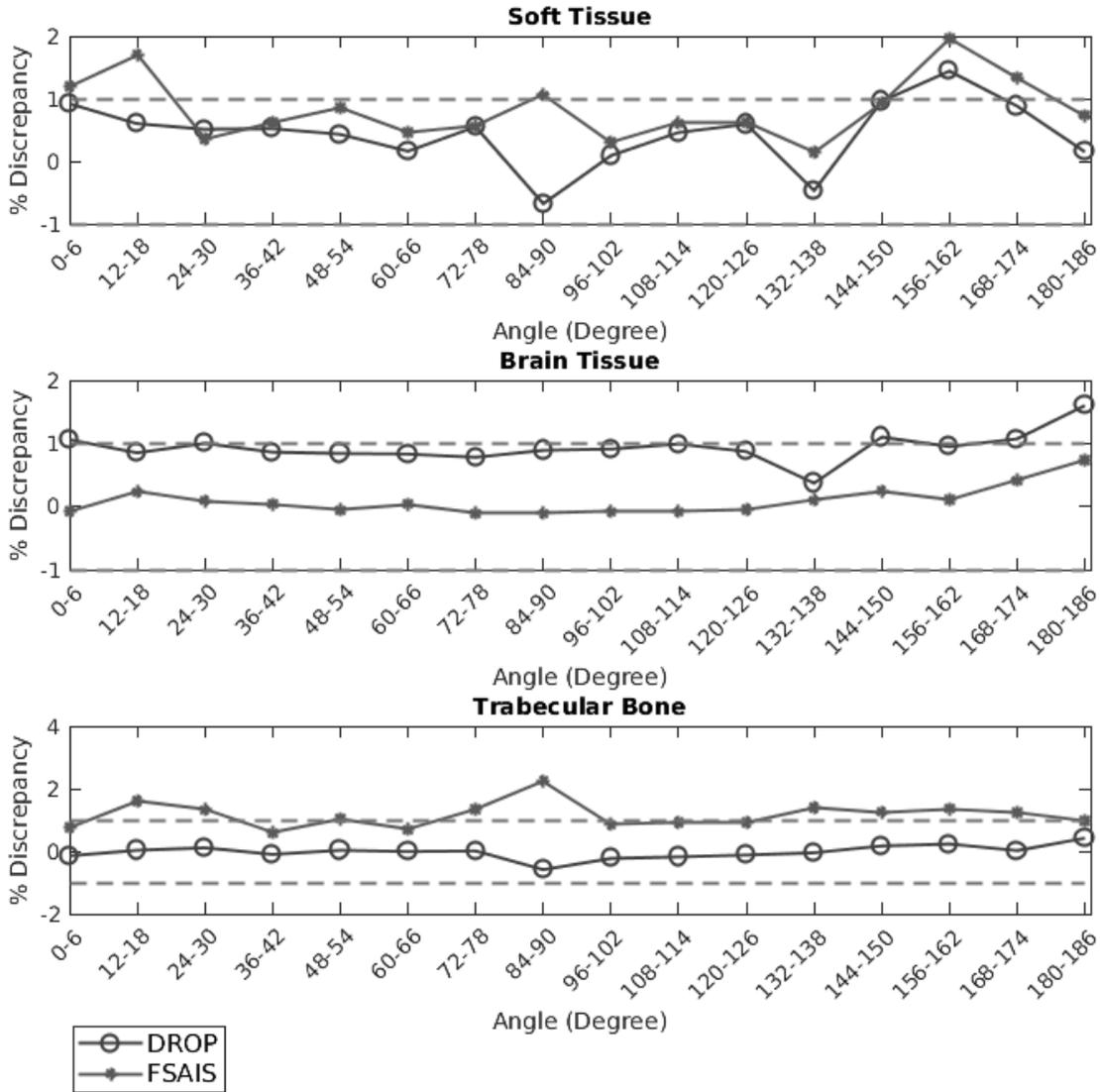


Figure 4.37: % Discrepancy of soft and brain tissues in pediatric head phantom after six iterations of FSAIS and DROF. Each point on the horizontal axis represents the angle interval from which protons were removed during data read. Dashed lines indicate the desired range of % discrepancy which is between -1 and 1. Both algorithms generated accurate RSP values for most of the angle intervals.

Figure 4.38 illustrates the results of removing protons from different angle intervals with fixed width of 6° during the image reconstruction. Based on these results, removing protons during the image reconstruction does not cause significant fluctuations in the RSPs. Results in Figure 4.38 show the % discrepancy generated by both algorithms FSAIS and DROP is consistent with the results reported in Table 4.21, which were related to image reconstruction under normal conditions (i.e. no protons were removed), because only a small number of protons were removed.

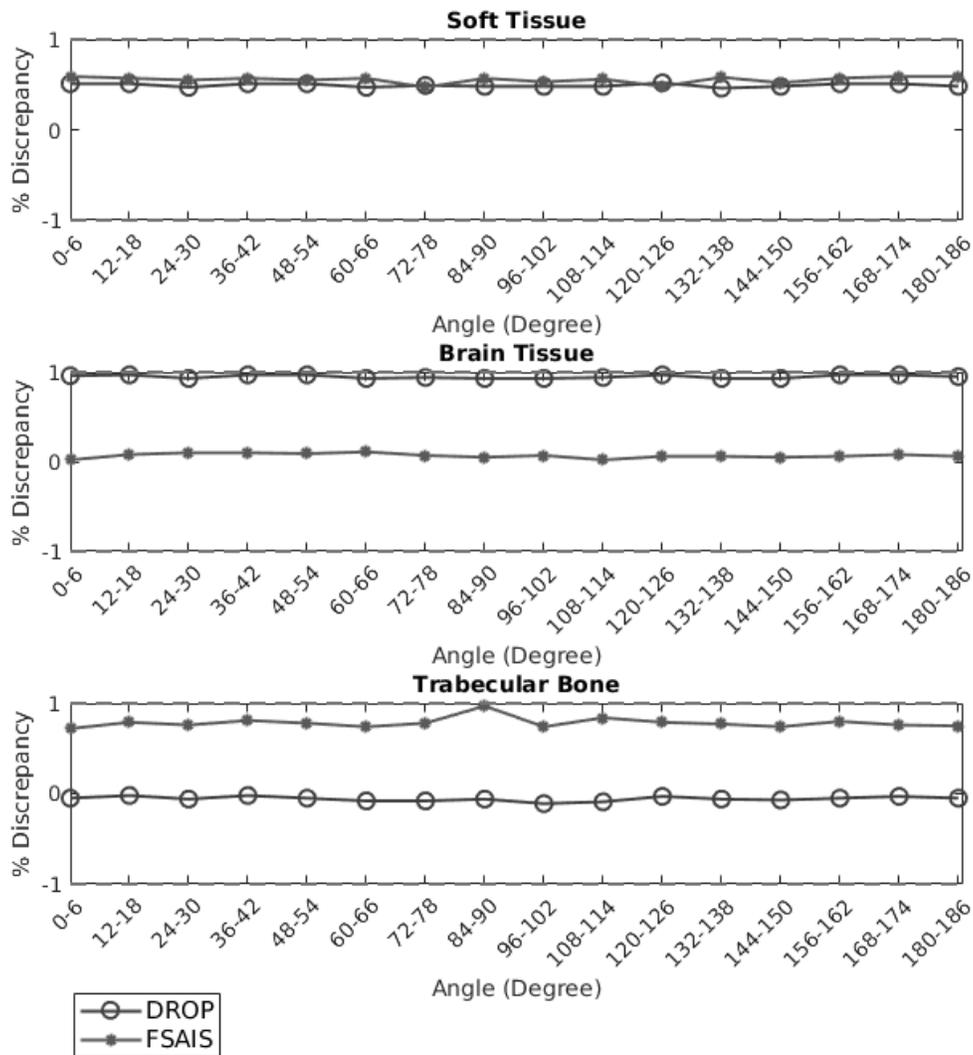


Figure 4.38: % Discrepancy of materials in the pediatric head phantom after six iterations of FSAIS and DROP. Dashed lines indicate the desired range of % discrepancy.

Figure 4.39 shows the reconstructed pCT images by DROP and FSAIS while protons within a range of $0.0^\circ - 5.0^\circ$ were removed during data read. As expected, the FBP image used as the initial iterate was affected by proton removals during data read. RSP accuracy of different materials is still good while protons were removed within a range of $0.0^\circ - 5.0^\circ$ based on the results in Table 4.22.

Figure 4.40 shows the reconstructed pCT images by DROP and FSAIS while protons within a range of $0.0^\circ - 5.0^\circ$ were removed during data read and a median filter of radius 3 was applied to the final pCT image (i.e. image generated at the 6th iteration of the iterative solver). The reason that a median filter was applied to the final pCT image rather than the FBP image is related to the fact that the median filter on the FBP image did not improve the pCT images reconstructed by FSAIS because of the reason that FSAIS tends to generate images with sharper edges and did not work well when this behavior was combined with the median filter. Based on the results in Figure 4.40, median filter caused an improvement on pCT images generated by DROP because of the fact that DROP's pCT images were smoother, while FSAIS generated pCT images with sharper edges, therefore less affected by the median filter. The reconstructed RSP values of different materials within the pediatric head phantom corresponding to the pCT images in Figure 4.40 are shown in Table 4.23.

Figure 4.41 shows the reconstructed pCT images by DROP and FSAIS while protons within a range of $0.0^\circ - 10.0^\circ$ were removed during data read. The FBP image used as the initial iterate was significantly affected by proton removals during data read. RSP accuracy of different materials is still good while protons were removed within a range of $0.0^\circ - 10.0^\circ$ based on the results in Table 4.24. Figure 4.42 shows the reconstructed pCT images by DROP and FSAIS while protons within a range of $0.0^\circ - 10.0^\circ$ were removed during data read and a median filter of radius 3 was applied to the pCT image at the 6th iteration. (see Table 4.25 for the RSPs of Figure 4.42).

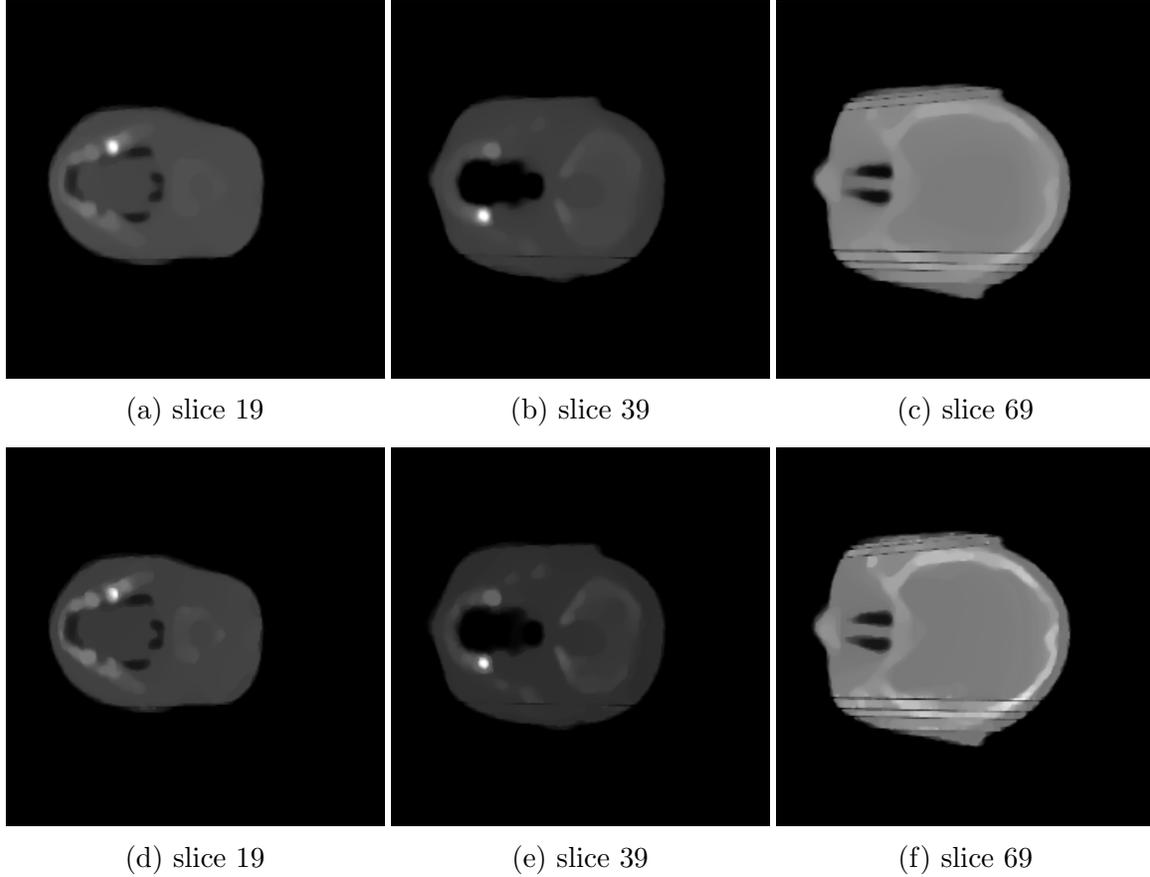


Figure 4.39: pCT images of the pediatric head reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with protons removed within a range of $0.0^\circ - 5.0^\circ$ during data read.

Table 4.22: Reconstructed RSP values of pediatric head phantom by DROP and FSAIS with protons were removed within a range of $0.0^\circ - 5.0^\circ$ during data read. Both algorithms generated accurate RSP values (close to the predicted RSPs) for soft tissue, brain tissue and Trabecular bone, but FSAIS generated more accurate RSP value for Enamel.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft ti.	1.037	1.0433 ± 0.0225	0.61	1.0412 ± 0.0229	0.41
Brain ti.	1.047	1.0572 ± 0.0012	0.97	1.0463 ± 0.0015	-0.07
Trab. bone	1.108	1.1069 ± 0.0028	-0.1	1.1198 ± 0.0132	1.06
Dentin	1.513	1.3869 ± 0.0368	-8.33	1.741 ± 0.0658	15.06
Enamel	1.788	1.3869 ± 0.0368	-22.43	1.741 ± 0.0658	-2.62

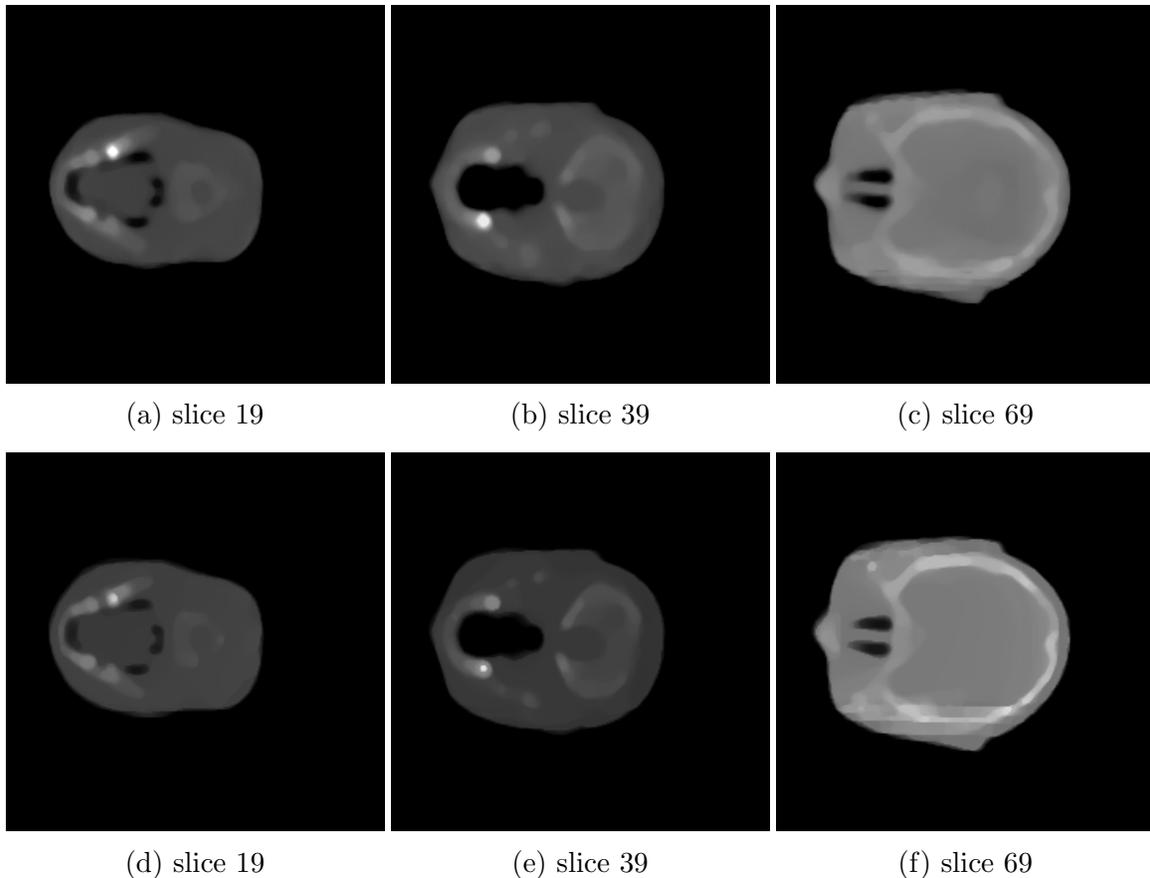


Figure 4.40: Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 5.0^\circ$ prior to generating the initial iterate. A median filter of radius 3 was used at the 6th iteration of DROP and FSAIS.

Table 4.23: RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 5.0^\circ$ during data read. A median filter of radius 3 was used at the 6th iteration of DROP and FSAIS. Using a median filter on FSAIS's result generated more accurate RSP values for soft tissue, brain tissue, Trabecular bone, and Enamel compared to DROP.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft ti.	1.037	1.0324 \pm 0.0247	-0.44	1.0411 \pm 0.0224	0.4
Brain ti.	1.047	1.0385 \pm 0.0044	-0.81	1.0463 \pm 0.0015	-0.07
Trab. bone	1.108	1.0908 \pm 0.0065	-1.55	1.1200 \pm 0.0131	1.08
Dentin	1.513	1.3881 \pm 0.0423	-8.25	1.7410 \pm 0.0658	15.06
Enamel	1.788	1.3881 \pm 0.0423	-22.36	1.7410 \pm 0.0658	-2.63

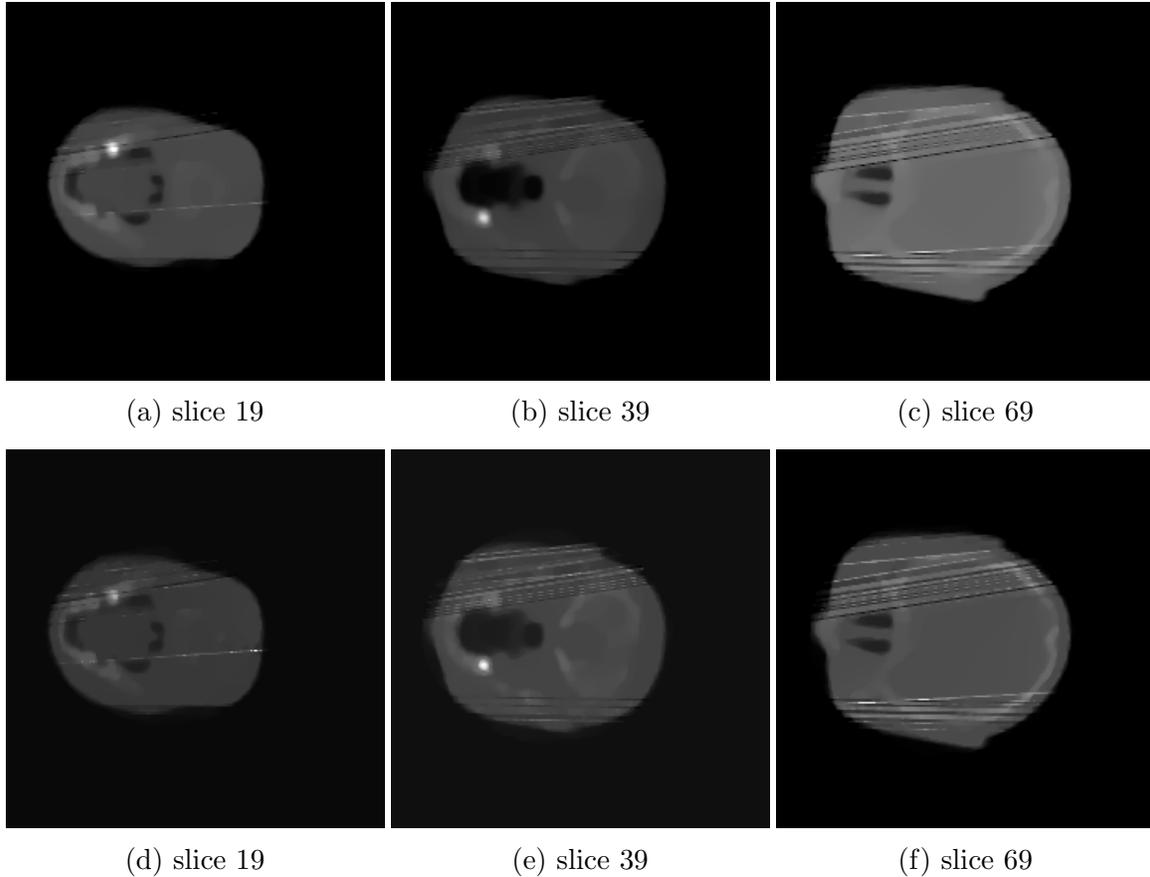


Figure 4.41: Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 10.0^\circ$ during data read.

Table 4.24: RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 10.0^\circ$ during data read. Based on the results FSAIS generated more accurate RSP values for brain tissue, while DROP generated a more accurate RSP value for soft tissue and Trabecular bone.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft ti.	1.037	1.022 ± 0.0711	-1.45	1.0552 ± 0.1499	1.76
Brain ti.	1.047	1.0589 ± 0.003	1.14	1.0509 ± 0.0039	0.37
Trab. bone	1.108	1.1121 ± 0.0029	0.37	1.1241 ± 0.0145	1.45
Dentin	1.513	1.4058 ± 0.0869	-7.09	1.7015 ± 0.1786	12.46
Enamel	1.788	1.4058 ± 0.0869	-21.37	1.7015 ± 0.1786	-4.83

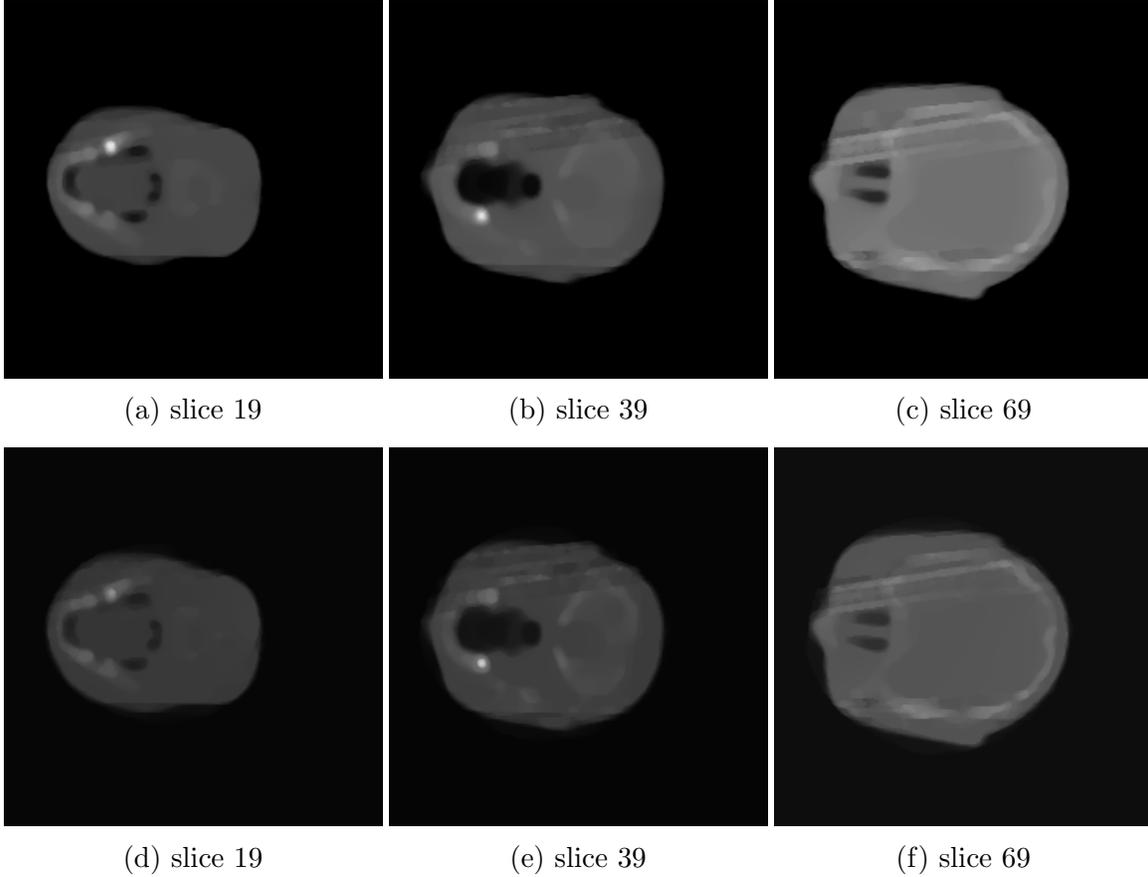


Figure 4.42: Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with histories removed within a range of $0.0^\circ - 10.0^\circ$ prior to generating the initial iterate. A median filter of radius 3 is used at the 6^{th} iteration of DROP and FSAIS.

Table 4.25: RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed within a range of $0.0^\circ - 10.0^\circ$ range during data read. A median filter of radius 3 is used at the 6^{th} iteration of DROP and FSAIS. Using a median filter on FSAIS's result generated more accurate RSP values for soft tissue, brain tissue, and Enamel compared to DROP.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft ti.	1.037	1.0233 \pm 0.0359	-1.32	1.0390 \pm 0.0285	0.19
Brain ti.	1.047	1.0590 \pm 0.0032	1.15	1.0512 \pm 0.0039	0.4
Trab. bone	1.108	1.1123 \pm 0.0030	0.39	1.1256 \pm 0.0149	1.59
Dentin	1.513	1.3978 \pm 0.052	-7.61	1.7459 \pm 0.0907	15.39
Enamel	1.788	1.3978 \pm 0.052	-21.82	1.7459 \pm 0.0907	-2.35

The final experiment in this chapter is where the protons were removed every 4.0° during the image reconstruction. The results of this experiment are shown in Figure 4.43 and Table 4.26. Removing protons every 4.0° causes one fourth of protons be removed during the image reconstruction. In the case of pediatric head phantom, there are about 100 million protons that pass the statistical cuts and will be used by the iterative solver.

When protons were removed every 4.0° , there were about 25 million protons used by the iterative solver which is significantly less data and expected to reduce the sharpness and RSP accuracy of different materials within the pediatric head phantom, but based on the results in Figure 4.43, FSAIS still generated pCT images with sharp edges especially for the case of slice 69 which the sharp edges are much better than the same slice generated by DROP.

The % discrepancy for RSP values of materials inside the pediatric head phantom in Table 4.26 shows that both algorithms FSAIS and DROP generated good RSP values for Trabecular bone, while in the case of brain tissue, FSAIS significantly outperforms DROP. The RSP value of the soft tissue generated by DROP is much better compared to FSAIS, but FSAIS still generates a very accurate RSP value for soft tissue.

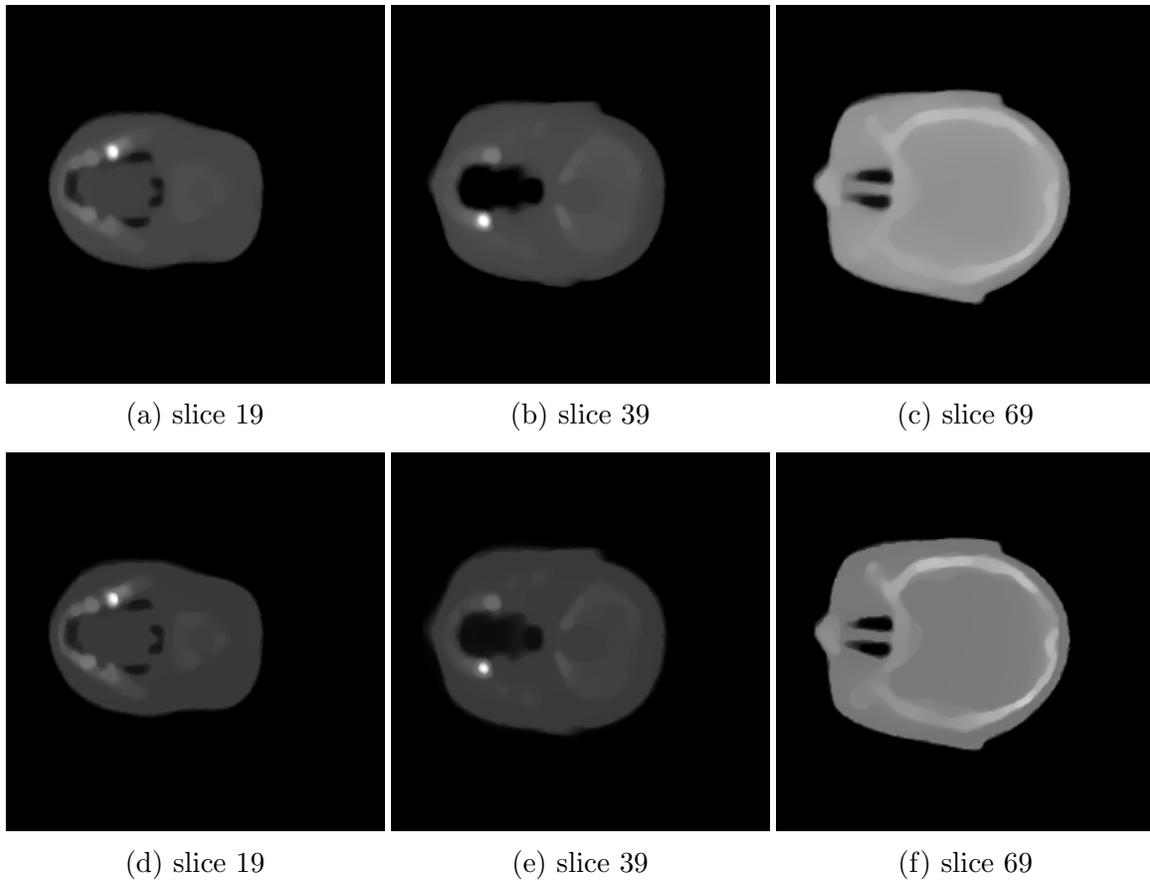


Figure 4.43: Representative experimental head phantom pCT images reconstructed by DROP (a)-(c) and FSAIS (d)-(f) with protons removed every 4.0° during the image reconstruction. Removing about one fourth of protons reduces the sharpness of edges in images generated by DROP, but FSAIS still generates images with sharp edges.

Table 4.26: RSP values reconstructed from experimental pediatric head phantom by DROP and FSAIS with protons removed every 4.0° during the image reconstruction. Based on the results, both algorithms generated accurate RSP values for soft tissue, brain tissue, and Trabecular bone, while FSAIS generated a much better RSP value for Enamel.

Material	Pred. RSP	DROP RSP (Mean \pm SD)	% Discrep.	FSAIS RSP (Mean \pm SD)	% Discrep.
Soft ti.	1.037	1.0396 \pm 0.0224	0.25	1.0271 \pm 0.0222	-0.95
Brain ti.	1.047	1.0549 \pm 0.0014	0.75	1.0486 \pm 0.0017	0.15
Trab. bone	1.108	1.1045 \pm 0.0029	-0.32	1.1119 \pm 0.0039	0.35
Dentin	1.513	1.3812 \pm 0.0392	-8.71	1.6725 \pm 0.086	10.54
Enamel	1.788	1.3812 \pm 0.0392	-22.75	1.6725 \pm 0.086	-6.54

CHAPTER FIVE

GPU-based Image Reconstruction

The pCT problem fits well within the Single Instruction Multiple Threads (SIMT) parallel programming paradigm of Graphics Processing Units (GPUs) since we treat each proton as an individual thread that can be processed in parallel. Although there is a great need for small pCT runs as a validation technique in clinics, when it comes to imaging an adult torso, because of the greatly increased problem size, the reconstruction time grows in proportion to the increased size of the reconstructed object's volume [27].

In recent years, several research papers have developed GPU-based image reconstruction techniques for pCT. In [25], based on a comparison of the reconstruction time on a single machine vs. a GPU cluster, it has been demonstrated that the reconstruction time can be reduced from 7 hours to 53 seconds for a dataset of size 131 million protons. In [26], a hybrid approach that uses both Message Passing Interface (MPI) and GPUs were implemented for performing the image reconstruction and it was demonstrated that image reconstruction in pCT can be accelerated through this approach. Using this approach on the same cluster the reconstruction runtime has been improved and reduced to 43 seconds for a similar size dataset. In a recent study in [27], it was shown (using the same approach as the first evaluation of the pCT software in [26],) that the execution time for generating accurate RSP values for a dataset of size 131 million protons is almost 30 seconds running on 60 processors (60 CPU cores and 60 GPUs).

The majority of pCT runtime from data read to writing the generated 3D maps of RSP values to SSD is spent during the parallel iterative solver mainly due to MLP computations. Because of memory limitations, the matrix A containing the

MLP calculations can not be fully stored in GPU memory even considering the sparse nature of this matrix. For the small simulated CTP404 phantom, it requires about 6 GB of memory using the sparse compressed storage formats. The memory of modern GPU devices, such as K40 or P100, is 12 to 16 GB. Considering the limited memory of the GPU devices and the size of matrix A for an experimental dataset which is at least 5 to 10 times greater than the simulated data, calculating the matrix A before performing the iterative solver requires lots of data transfers between the CPU and GPU memory which can act as barriers for parallel tasks. One good solution is performing the MLP calculations during the the iterative solver. In other words, when a block of protons are loaded into the GPU memory to be processed by the iterative solver, the MLP of those protons is calculated right before updating $x(k)$ of that block by the iterative solver. The benefit of this method for performing the MLP is that bigger blocks of protons can be loaded into the GPU memory, thus, more calculations are performed in parallel.

In this chapter, we discuss two distributed GPU-based image reconstruction algorithms with and without data transfers among the GPU devices and discuss their performance on the CTP404 and pediatric head phantoms. Details of simulated CTP404 and experimental datasets were described in chapters 3 and 4.

5.1 Distributed GPU-based Image Reconstruction without Data Transfer

In order to benefit from the Nvidia Tesla GPU devices, we have developed a fast and easy to implement reconstruction algorithm that can run on systems with at least two GPU devices with the advantage of eliminating the need of data transfers among the GPU devices. The algorithm we have designed relies on the following two facts¹:

- (1) only a small fraction of protons intersect multiple slices along the vertical axis

¹Some results in this section were published in [64]

- (2) within a slice, only a small portion of protons pass through and are used in reconstruction of that specific slice

Based on the first observation, which is reported with details in Table 5.1, we know that a proton’s path does not have a significant deviation along the z axis or vertical slices of the reconstruction volume. This enables us to consider an algorithm without having a significant concern about the GPU-to-GPU data transfers which could potentially add some considerable timing constraints.

The second observation allows us to split the protons among the available GPU devices such that each GPU reconstructs a portion of the 3D image. Each GPU only needs some fraction of the protons, and this lowers the amount of required memory and processing time per GPU leading to a faster reconstruction time overall.

Table 5.1: Percentage of protons passing through image slices along the z axis for the simulated and experimental CTP404 phantom with the 2.5 mm slice thickness.

# of slices	Simulated CTP404	Experimental CTP404
1	32	42
2	43	36
3	19	15
4	4	3
≥ 5	2	4

The general structure of the proposed multi-GPU based reconstruction technique without data transfers can be seen in Algorithm 1. The first step is the division of the reconstruction volume into several overlapping regions based on the number of available GPU devices, while each region is assigned to a unique key. The overlap is set at twice the number of slices for a 96% coverage (see Table 5.1). In the second step, each proton is associated with the image region entered into and exited from. In the third step, based on the number of protons that pass through each image region, the required memory for each proton is allocated on each GPU. Lastly (fourth step),

we perform the iterative solver on each GPU. Finally, in the overlap of two regions, where there are two reconstructions of each slice, we select the reconstruction that is either closest to the non-overlapped slices of its region or a slice from the overlapped regions based on distribution of protons. For example, in reconstructions presented in this section performing on CTP404, there are two regions, one containing slices 0 – 12 and the other containing slices 7 – 19. The overlap corresponds to slices 7 – 12, and the final image will be made of slices 0 – 9 from the first region and 10 – 19 of the second. Therefore there are cases in which one image region contains more protons due to uneven data distribution. In such a case, slices in the overlapped regions can be more accurate.

Algorithm 1 multi-GPU based reconstruction algorithm without data transfers

procedure MULTI-GPU BASED RECONSTRUCTION

- 1: Setup image regions
 - 2: Identify protons passing regions
 - 3: Allocate memory for protons per GPU
 - 4: Iteratively solve per GPU
 - 5: Select slice
-

When using the proposed reconstruction algorithm described in Algorithm 1, there are different numbers of protons passing through each image region which are reported in Table 5.2. Based on the results in Table 5.2, protons are evenly distributed between the two image regions for the simulated data, while the top image region of the experimental data encompasses 1.6 times the number of protons that pass through the bottom image region.

Table 5.2: Number of protons (millions) in each image region used by GPU devices to perform the iterative solver.

Data	Original Size	Slices 0-12	Slices 7-19
Simulated CTP404	21	14	14
Experimental CTP404	73	67	41

Execution times reported in the next sections heavily depend on the number of protons traversing the image regions and the distribution of data.

In order to test the performance of Algorithm 1, we used two different block sizes of 1280000 and 320000 for the experimental and simulated CTP404 phantom, respectively. The block size is the number of protons to be processed in parallel based on DROP or FSAIS. When it comes to the performance of the iterative solver, in general, larger block sizes decrease the runtime of a single iteration but do not necessarily generate accurate RSP values. Table 5.3a shows runtime of one iteration of DROP with additive updates on a single GPU and double GPU systems for the simulated CTP404.

Table 5.3: Performance comparison of pCT image reconstruction on the simulated CTP404.

(a) Runtime (sec.) of 1 iteration of DROP with additive updates.			(b) Total reconstruction runtime (sec.) of pCT image reconstruction.		
System	1 GPU	2 GPUs	System	1 GPU	2 GPUs
K40	13.3	7.6	K40	154.1	96.3
P100	2.3	1.5	P100	44.9	40.2

Execution times reported in Table 5.3a heavily depend on the number of protons traversing the image regions and the distribution of data. Based on the results in Table 5.3a, the runtime of the iterative solver on P100 is about 1.53 times faster after distributing data and is similar to the expected speedup of $\frac{21}{14}$ or 1.5. On the other hand, the speedup of Algorithm 1 on K40 system is about 1.75, which is greater than expected. The reason comes from the difference between the memory sizes of K40 and P100 GPU devices. In fact, the smaller memory of the K40 system in comparison to P100 leads to efficient usage of the cache on K40 leading to a faster runtime of a single iteration. The total pCT reconstruction runtime of simulated CTP404 including data read and data cleaning tasks, along with TVS iterations used at the beginning of each iteration of DROP was reported in Table 5.3b. Reconstructed RSP values of

the simulated CTP404 are reported in Table 5.4, which show accurate reconstruction results using Algorithm 1.

Table 5.4: Reconstructed RSP values of central slice of the simulated CTP404 using one and two GPUs. For the case of two GPUs, the slice belonged to the top half of the image.

Material	1 GPU (% discrepancy)	2 GPUs (% discrepancy)
PMP	0.886 (0.33)	0.887 (0.45)
LDPE	0.986 (0.61)	0.988 (0.82)
Polystyrene	1.032 (0.78)	1.033 (0.87)
Acrylic	1.163 (0.25)	1.162 (0.17)
Delrin	1.349 (-0.73)	1.347 (-0.88)
Teflon	1.7895 (-0.03)	1.786 (-0.22)

Table 5.5 shows the performance of pCT image reconstruction on the experimental CP404 phantom using FSAIS as the iterative solver. Based on the timing results in Table 5.5a, using two K40 GPU devices not only does not cause performance improvement, but it is slightly slower than running on one GPU. In general, based on the distribution results in Table 5.2, after splitting data among GPUs, about 91% of protons belong to the top half of the image running on one of the two available GPUs. Thus, the expected speedup is around 1.08. In reality and in an actual implementation of Algorithm 1, after one iteration of the iterative solver, two GPU devices need to be synced before starting the next iteration which causes some overhead per iteration. Moreover, at the end of each iteration, there are two consecutive GPU to CPU transfers followed by two consecutive data writes to SSD instead of one transfer and data write using single GPU. Due to these facts, reconstruction runtime using Algorithm 1 and based on uneven distribution of protons does not outperform the runtime of pCT reconstruction on one GPU.

The total pCT reconstruction runtime of experimental CTP404 including data read and data cleaning tasks, along with TVS performing during the first and second

iterations of FSAIS is reported in Table 5.5b. Also, the total runtime on two GPU devices includes the runtime of distributing the data among GPU devices which was about 20 seconds. RSP values of the reconstructed experimental CTP404 are reported in Table 5.6 for the case of one and two GPU devices. Running the same experiments on P100 GPU devices show no difference between the total computation time, but the runtime of one iteration of FSAIS is slightly faster on one GPU due to uneven distribution of protons as discussed earlier.

Table 5.5: Performance analysis of pCT image reconstruction on the experimental CTP404 using FSAIS as the iterative solver. Total reconstruction runtime included six iterations of FSAIS.

(a) Runtime (sec.) of 1 iteration of FSAIS. (b) Total reconstruction runtime (min.).

System	1 GPU	2 GPUs	System	1 GPU	2 GPUs
K40	8.7	9.5	K40	3.0	3.4
P100	1.7	2.1	P100	1.6	1.6

Table 5.6: Reconstructed RSP values of slice number 10 from the top half of the experimental CTP404 pCT image using two GPU devices. See Table 5.5b for total reconstruction runtime of the RSP values. The reconstructed RSP values of all materials except PMP are within a range of 1% of the predicted RSPs.

Material	1 GPU (% discrepancy)	2 GPU (% discrepancy)
PMP	1.24	1.42
LDPE	0.21	0.24
Polystyrene	-0.61	-0.55
Acrylic	-0.83	-0.9
Delrin	0.70	0.24
Teflon	-0.43	0.05

Table 5.7a shows the pCT reconstruction time of experimental pediatric head phantom using FSAIS and running on K40 and P100 GPU devices. The slice thickness of reconstructed image of pediatric head is 1 mm and 96% of protons pass through no more than six vertical slices. This observation is similar to the distribution of

protons along vertical slices of CTP404 in Table 5.2. The number of protons processed in parallel by FSAIS were 1280000. There are about 100 million protons prior to performing the iterative solver, which after splitting the image regions into half and adding the six extra slices to each image region, results in 56 and 63 million protons associated with the top and bottom image regions, respectively. The total pCT computation time for the experimental pediatric head phantom in Table 5.7b includes data read and data cleaning tasks, along with performance of TVS during every iteration of FSAIS. The runtime for distributing the data among the two GPU devices was about 25 seconds. Results in Table 5.7a show speedups by factors of 2.3 and 2 from running pCT reconstruction on K40 and P100 GPU devices, respectively. Both of these speedups are greater than the expected speedup $\frac{100}{63}$, due to memory efficiency. The speedup of the Algorithm 1 on the K40 system is greater than the speedup on P100, which as discussed earlier in this section is due to the smaller memory of the K40 system in comparison to P100, resulting in the efficient usage of the cache on K40 allowing for a faster runtime of a single iteration. Reconstructed RSP values of Algorithm 1 are reported in Table 5.8.

Table 5.7: Performance analysis of pCT image reconstruction of the experimental pediatric head phantom using Algorithm 1 and FSAIS as the iterative solver. Total reconstruction runtime included six iterations of FSAIS.

(a) Runtime (sec.) of 1 iteration of FSAIS. (b) Total reconstruction runtime (min.).

System	1 GPU	2 GPUs	System	1 GPU	2 GPUs
K40	60.6	26.5	K40	9.1	6.1
P100	16.8	8.2	P100	3.5	3.1

Results in Table 5.8 show good accuracy of reconstructed RSP values for all materials. Spinal disk’s RSP is slightly lower than -1% , which can be due to removing the protons that traverse more than six vertical slices based on step 2 of Algorithm 1.

Table 5.8: Reconstructed RSP values of the experimental pediatric head phantom by FSAIS using two GPU devices. See Table 5.7b for total reconstruction runtime of the RSP values. The reconstructed RSP values of soft tissue, brain tissue and Trabecular bone are within a range of 1% of the predicted RSPs.

Material	Mean RSP	% Discrepancy
Soft tissue	1.0428	0.56
Brain tissue	1.0512	0.4
Trabecular bone	1.1053	-0.24
Dentin	1.6874	11.52
Enamel	1.6874	-5.63

5.2 Distributed GPU-based Image Reconstruction with Data Transfers

As discussed in the previous section, designing a parallel distributed image reconstruction model for pCT requires consideration of how the reconstruction volume and protons are split among the available GPU devices. The main problem of dividing the reconstruction volume among several GPU devices is that after we split the data and assign each section to a GPU, voxels in the slices close to the split section don't get enough updates which causes low estimates for the RSP values and can delay the convergence of the iterative solver. The reason for this comes from the fact that in order to accurately estimate the proton paths based on the MLP, we need to use the coordinates of entry and exit points of protons with respect to the object. Therefore, we collect only the protons that both enter and exit the image hull. Thus, if a slice is close enough to the split section of the image, it may include proton paths that enter and exit a slice belonging to one GPU, and traversing a slice or slices that belong to another GPU. With this assumption, slices that include a part of the MLP without the entry and exit points, will not get any updates. It is trivial to mention that all the other slices with reasonable distance from the split section get enough updates because the entry and exit points of protons and the voxels along the MLP all belong to the slices on the same GPU.

The proposed parallel GPU-based iterative reconstruction algorithm in this section is based on proton path deviations along the vertical slices of the reconstruction volume based on the observations in Table 5.1. Knowing that the proton path deviations are insignificant along the vertical slices, we have proposed a distributed image reconstruction method with data transfers which is outlined in Algorithm 2.

Before starting the image updates based on Algorithm 2, data is distributed among two GPU devices in the following way (steps 1 – 8 of Algorithm 2, Figure 5.1):

- img_0 : slices^{2,3} 0 to $\frac{NUM_SLICES}{2} - slices_2_xfer - 1 \rightarrow GPU_A$
- img_1 : slices $\frac{NUM_SLICES}{2} - slices_2_xfer$ to $\frac{NUM_SLICES}{2} - 1 \rightarrow GPU_A$
- img_2 : slices $\frac{NUM_SLICES}{2}$ to $\frac{NUM_SLICES}{2} + slices_2_xfer - 1 \rightarrow GPU_B$
- img_3 : $\frac{NUM_SLICES}{2} + slices_2_xfer$ to $NUM_SLICES - 1 \rightarrow GPU_B$

After distributing the problem among the two GPU devices on a compute node, one identifies the protons which enter and exit each image section and allocate the memory for them on the associated GPU devices (steps 10 – 18 of Algorithm 2). At this point, one performs the image updates using the iterative image reconstruction techniques such as DROP or FSAIS. After performing the iterative solver, the image regions close to the split section of the original image are transferred from one GPU to another (steps 20 – 30 of Algorithm 2), and this process is repeated based on a user-determined integer NUM_CYCLES representing the number of cycles this process needs to be performed. A high level description of Algorithm 2 is also illustrated in Figure 5.1.

²Assuming slice indexing starts from 0

³ $slices_2_xfer$ is the number of vertical slices that needs to be transfers between the GPUs, based on a knowledge of proton path deviations along the vertical axis.

Algorithm 2 Pseudocode definition of the distributed GPU based reconstruction algorithm with data transfers

procedure DISTRIBUTED IMAGE RECONSTRUCTION WITH DATA TRANSFERS

```

1: set  $img_0 = x_0[0 : r \times c \times (\frac{s}{2} - slices\_2\_xfer - 1)]$ 
2: set  $img_1 = x_0[r \times c \times (\frac{s}{2} - slices\_2\_xfer) : r \times c \times (\frac{s}{2} - 1)]$ 
3: set  $img_2 = x_0[r \times c \times \frac{s}{2} : r \times c \times (\frac{s}{2} + slices\_2\_xfer - 1)]$ 
4: set  $img_3 = x_0[r \times c \times (\frac{s}{2} + slices\_2\_xfer) : r \times c \times (s - 1)]$ 
5: set  $hull_0 = hull[0 : r \times c \times (\frac{s}{2} - slices\_2\_xfer - 1)]$ 
6: set  $hull_1 = hull[r \times c \times (\frac{s}{2} - slices\_2\_xfer) : r \times c \times (\frac{s}{2} - 1)]$ 
7: set  $hull_2 = hull[r \times c \times \frac{s}{2} : r \times c \times (\frac{s}{2} + slices\_2\_xfer - 1)]$ 
8: set  $hull_3 = hull[r \times c \times (\frac{s}{2} + slices\_2\_xfer) : r \times c \times (s - 1)]$ 
9:  $h \leftarrow 0$ 
10: while  $h < NUM\_PROTONS$ 
11:   if ENT\_AND\_EXT\_HULL( $h, hull_0, hull_1$ )
12:     ALLOCATE\_DEVICE\_MEMORY( $gpu_A, h$ )
13:   elseIf ENT\_AND\_EXT\_HULL( $h, hull_2, hull_3$ )
14:     ALLOCATE\_DEVICE\_MEMORY( $gpu_B, h$ )
15:   else
16:     FREE\_HOST\_MEMORY( $h$ )
17:    $h \leftarrow h + 1$ 
18: endWhile
19:  $i \leftarrow 0$ 
20: while  $i < NUM\_CYCLES$ 
21:   PERFORM\_ITERATIVE\_SOLVER( $img_0, img_1, gpu_A$ )
22:   PERFORM\_ITERATIVE\_SOLVER( $img_2, img_3, gpu_B$ )
23:   copy  $img_2$  from  $gpu_B$  to  $gpu_A$ 
24:   PERFORM\_ITERATIVE\_SOLVER( $img_0, img_1, img_2, gpu_A$ )
25:   PERFORM\_ITERATIVE\_SOLVER( $img_3, gpu_B$ )
26:   copy  $img_1, img_2$  from  $gpu_A$  to  $gpu_B$ 
27:   PERFORM\_ITERATIVE\_SOLVER( $img_1, img_2, img_3, gpu_B$ )
28:   PERFORM\_ITERATIVE\_SOLVER( $img_1, gpu_A$ )
29:    $i \leftarrow i + 1$ 
30: endWhile

```

5.2.1 Implementation of Algorithm 2

Instead of updating the images from top to bottom, there is an alternative way to implement Algorithm 2, which is updating the image with the same blocks of protons three times and then moving to the next block of protons. For instance, if we divide the proton data into 10 blocks such that each block updates specific portions of the image as illustrated in Figure 5.2, one performs the updates from left

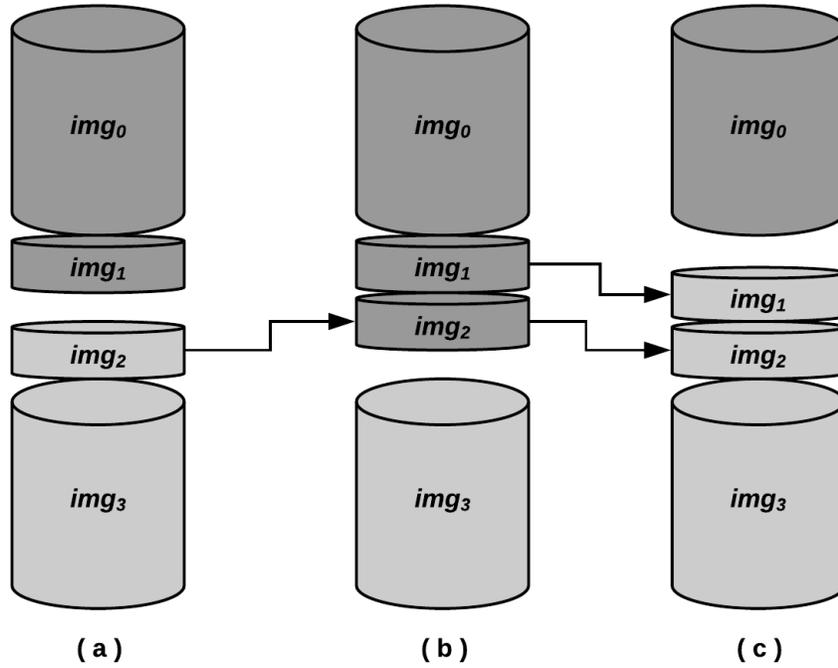


Figure 5.1: Schematic drawing of one cycle of Algorithm 2.

to right which is equivalent to three updates on block #1 before starting the updates of the second block, which is in contrast to updating images with consecutive blocks. Traversing the blocks based on the left to right approach leads to a performance improvement of the reconstruction time of data. This implementation performs projections onto blocks and updates the image after each such projection. Thus a block-iterative method such as DROP with additive updates is more compatible with this implementation compared with fully simultaneous methods which require updating the image after performing projections onto all rows of matrix A . The computation time of Algorithm 2 on pediatric head phantom is reported in Table 5.9.

Table 5.9: Runtime of the pediatric head image reconstruction using Algorithm 2 with two K40 GPUs. DROP w/ additive updates was used as the iterative method.

Task	Runtime (sec.)
1 cycle of Algorithm 2	67.5
Total reconstruction runtime	204.5

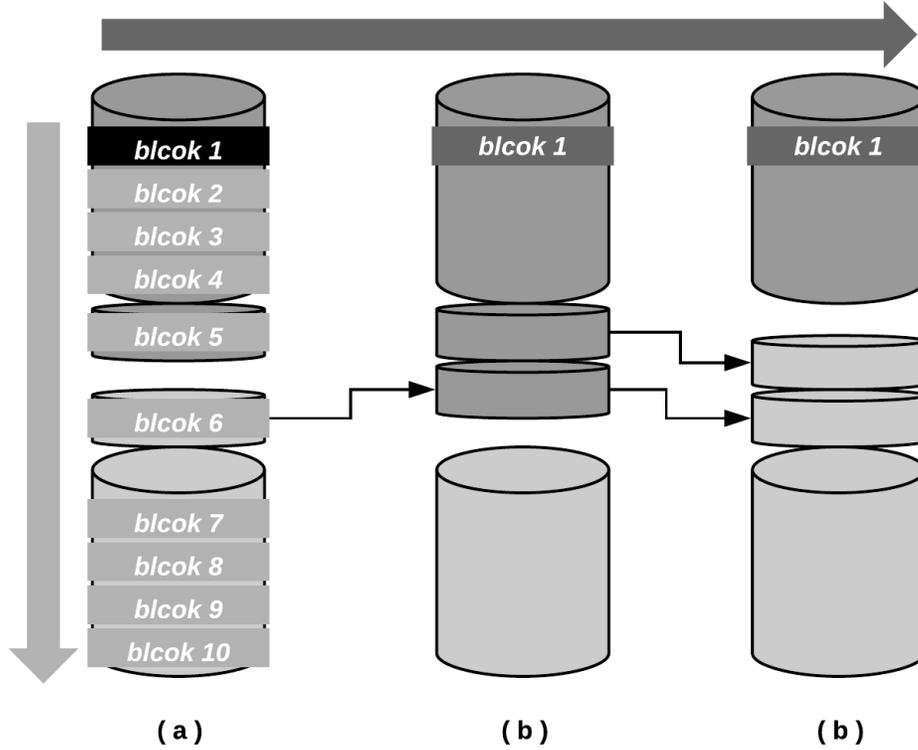


Figure 5.2: Illustration of left-to-right vs. top-to-bottom implementation of Algorithm 2. Instead of updating the images from top to bottom, there is an alternative way to implement Algorithm 2, which is updating the image with the same blocks of protons three times and then moving to the next block of protons. For instance, if we divide the proton data into 10 blocks such that each block updates specific portions of the image, one performs the updates from left to right which is equivalent to three updates on block #1 before starting the updates of the second block, which is in contrast to updating images with consecutive blocks.

Based on the results shown in Table 5.9, the total runtime of pediatric head image reconstruction is about 3.41 minutes after one cycle of Algorithm 2 and the results in Table 5.10 show accurate reconstruction of RSP values of the pediatric head phantom after one cycle of Algorithm 2.

Data transfers between two GPU devices is implemented using CUDA's `cudaMemcpyPeerAsync` method which uses the unified virtual addressing (UVA). UVA enables one address space for all CPU and GPU memories which copies the bytes between two GPU devices along the interconnect (PCIe) path with no staging through CPU memory. The P100 GPU devices used in timing analysis reported earlier in

Table 5.10: Reconstructed RSP values of the experimental pediatric head by DROP with additive updates on two K40 GPU devices. Reconstructed RSP values of all materials except the Enamel are within a range of 1% of the predicted RSPs.

Material	Mean RSP	% Discrepancy
Soft tissue	1.0329	-0.4
Brain tissue	1.0437	-0.32
Trabecular bone	1.1189	0.98
Dentin	1.5161	0.20
Enamel	1.5161	-15.20

this chapter did not support the peer-to-peer memory transfers, thus Algorithm 2 was only implemented on two K40 GPUs which supported the peer-to-peer memory transfers.

The experimental pediatric head phantom is composed of two scans, the first of which corresponds to the infimum or the lower part of the head, and the second scan corresponds to the supremum or upper part of the head. All the experiments related to the pediatric head phantom in chapter four and five were related to the lower part of the pediatric head phantom. The total number of protons included in the two scans of the experimental pediatric head phantom is about 502 million. A reconstructed pCT image from the upper part of the experimental pediatric head phantom is represented in Figure 5.3.

In order to have an assessment of the total reconstruction runtime on the entire head dataset composed of 502 million protons, different tasks of pCT were assigned to a different number of GPU devices as illustrated in Figure 5.4. Data read in Figure 5.4 includes some of the preprocessing tasks such as removing outliers, binning and hull detection. Pre-recon step in Figure 5.4, which runs on two K40 GPU devices, includes tasks such as statistical analysis and cuts, generating the FBP image, and obtaining the endpoints of protons used in MLP calculations. The reconstruction step runs on four K40 GPU devices such that the lower part of data from the first scan runs on a

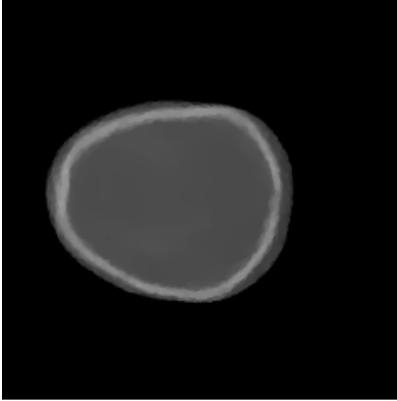


Figure 5.3: pCT reconstruction of the upper part of the experimental pediatric head.

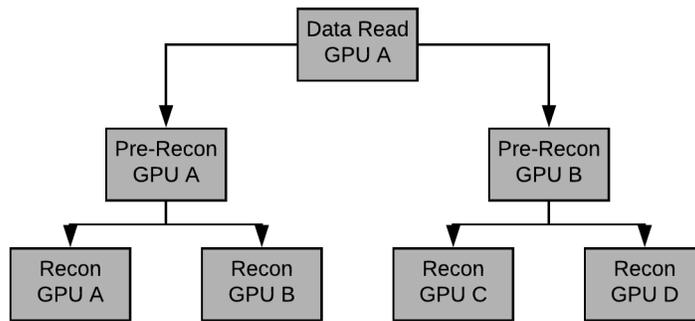


Figure 5.4: Schematic drawing of splitting pCT tasks on a node with four K40 GPU devices for reconstructing the entire experimental pediatric head dataset.

pair of GPU devices and the upper part of data from the second scan runs on another pair of GPU devices. Since each of the scans are isolated from each other and run on two GPU devices, either Algorithm 1 or 2 can be used in practice. Table 5.11 shows the runtime of different pCT tasks on the entire experimental pediatric head dataset using Algorithm 2 with total runtime of about 7.5 minutes.

Table 5.11: pCT reconstruction runtime of the entire experimental pediatric head.

Task	Runtime (sec.)
Data read and Preprocessing	194.5
Pre-reconstruction	175.7
Reconstruction	75.4

CHAPTER SIX

Conclusion

Proton computed tomography (pCT) is a promising imaging modality and a potential reliable alternative to other imaging techniques such as X-CT used in cancer treatment planning. There are two key goals in developing a reliable pCT system for a clinical setup: (1) accurate reconstruction of 3D maps of relative stopping power (RSP), and (2) implementing computationally fast image reconstruction methods.

There are some conditions in pCT that can have negative effects on the RSP values of reconstructed images. One such condition can be missing protons during a scan due to thickness of the body (e.g. pelvis) or hardware issues, or purposely removing protons for low-dose image reconstruction with the goal of potentially less damage to normal body tissues. Another issue which frequently occurs in pCT is the delayed convergence of materials with RSP values much greater than one, which is mainly due to the inaccurate RSP generated by FBP image which is used as the initial iterate in pCT.

In this dissertation, three variations of one of the existing widely used iterative solvers in pCT were studied under uncertain conditions and shown to have superior accuracy of the generated RSP values for certain cases such as having inaccurate initial iterate, and existence of noise in proton path estimations.

Moreover, a novel sparse compatible robust iterative solver algorithm was designed and proven to converge both in theory and experiment while generating RSP values of different materials within a range of 1% their corresponding predicted RSP values.

Generating accurate RSP values under normal and critical conditions is one of the key goals in pCT, but being able to reconstruct RSP values within clinically

recommended time frames is another key goal in pCT which needs to be met. In order to leverage the modern GPU devices on a single computing node, two methods for distributing the pCT tasks among multi-GPU systems were discussed in chapter five of this dissertation. The performance of these GPU based reconstruction methods using two or four GPUs showed good results close to the large GPU clusters consisting of hundred GPU devices.

Ion CT methods [65] such as pCT, carbon ion CT, and helium ion CT [66] as emerging imaging modalities all need to generate accurate RSP values using real-time image reconstruction algorithms. All the image reconstruction techniques discussed in this dissertation are not only limited to pCT, but can be used in other ion CT modalities for generating reliable results within the shortest amount of time.

APPENDICES

APPENDIX A

Details of pCT Reconstruction Timing

Computation time of different pCT image reconstruction tasks on K40 and P100 systems are reported in Tables A.1 and A.2. Data read is an iterative process where one data chunk at a time is read from SSD and processed such that protons that do not traverse the reconstruction volume are removed. During data read, hull detection is executed and the binning process is performed to group protons based on their relative angle and WEPL information. The information from the binning process is later used in one of the preprocessing steps to remove the protons that do not belong to a specific range of the standard deviation. Also, during preprocessing, protons that pass the statistical cuts are used to generate the FBP image. The final step before the image reconstruction is identifying and removing protons that do not enter or exit the hull, and those which meet this condition will later be used by the iterative solver.

Table A.1: Runtime of pCT image reconstruction tasks prior to performing the iterative solver on the experimental CTP404 phantom.

Task	K40 time (sec.)	P100 time (sec.)
Data read and preprocessing	80	53
Obtaining the endpoints for MLP	40	10

Table A.2: Runtime of pCT image reconstruction tasks prior to performing the iterative solver on the experimental pediatric head phantom.

Task	K40 time (sec.)	P100 time (sec.)
Data read and preprocessing	125	93
Obtaining the endpoints for MLP	60	19

BIBLIOGRAPHY

- [1] AM Cormack and AM Koehler. Quantitative proton tomography: preliminary experiments. *Physics in Medicine & Biology*, 21(4):560, 1976.
- [2] KM Hanson, JN Bradbury, TM Cannon, RL Hutson, DB Laubacher, R Macek, MA Paciotti, and CA Taylor. The application of protons to computed tomography. *Journal of Computer Assisted Tomography*, 2(5):671–674, 1978.
- [3] KM Hanson, JN Bradbury, TM Cannon, RL Hutson, DB Laubacher, RJ Macek, MA Paciotti, and CA Taylor. Computed tomography using proton energy loss. *Physics in Medicine & Biology*, 26(6):965, 1981.
- [4] KM Hanson, JN Bradbury, RA Koeppel, RJ Macek, DR Machen, R Morgado, MA Paciotti, SA Sandford, and VW Steward. Proton computed tomography of human specimens. *Physics in Medicine & Biology*, 27(1):25, 1982.
- [5] RP Johnson, VA Bashkurov, G Coutrakon, V Giacometti, P Karbasi, NT Karonis, CE Ordoñez, M Pankuch, HF-W Sadrozinski, KE Schubert, et al. Results from a prototype proton-ct head scanner. *Physics Procedia*, 90:209–214, 2017.
- [6] Vladimir A Bashkurov, Robert P Johnson, Hartmut F-W Sadrozinski, and Reinhard W Schulte. Development of proton computed tomography detectors for applications in hadron therapy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 809:120–129, 2016.
- [7] Tia Plautz, Vladimir Bashkurov, V Feng, F Hurley, RP Johnson, C Leary, S Macafee, A Plumb, V Rykalin, HF-W Sadrozinski, et al. 200 mev proton radiography studies with a hand phantom using a prototype proton ct scanner. *IEEE transactions on medical imaging*, 33(4):875–881, 2014.
- [8] Martina Bucciantonio and Fabio Sauli. Proton computed tomography. *Modern Physics Letters A*, 30(17):1540024, 2015.
- [9] R. Schulte, S. Penfold, J. Tafas, and K.E. Schubert. A maximum likelihood proton path formalism for application in proton computed tomography. *Med. Phys.*, 35:4849–4856, November 2008.
- [10] Charles-Antoine Collins Fekete, Paul Doolan, Marta F Dias, Luc Beaulieu, and Joao Seco. Developing a phenomenological model of the proton trajectory within a heterogeneous medium required for proton imaging. *Physics in Medicine & Biology*, 60(13):5071, 2015.

- [11] R.F. Hurley, R.W. Schulte, V.A. Bashkirov, A.J. Wroe, A. Ghebremedhin, H.F.-W. Sadrozinski, V. Rykalin, G. Coutrakon, P. Koss, and B. Patyal. Water-equivalent path length calibration of a prototype proton ct scanner. *Med. Phys.*, 39:2438–2446, 2012.
- [12] Lev M Bregman. The method of successive projection for finding a common point of convex sets. *Sov. Math. Dok.*, 162(3):688–692, 1965.
- [13] Yair Censor, Scott N Penfold, and Reinhard W Schulte. Systems and methodologies for proton computed tomography, December 8 2015. US Patent 9,207,193.
- [14] Scott Penfold and Yair Censor. Techniques in iterative proton ct image reconstruction. *Sensing and Imaging*, 16(1):19, 2015.
- [15] Blake Schultze, Paniz Karbasi, Valentina Giacomelli, Tia Plautz, Keith E Schübert, and Reinhard W Schulte. Reconstructing highly accurate relative stopping powers in proton computed tomography. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2015 IEEE*, pages 1–3. IEEE, 2015.
- [16] Reinhard Schulte, Vladimir Bashkirov, Tianfang Li, Zhengrong Liang, Klaus Mueller, Jason Heimann, Leah R Johnson, Brian Keeney, HF-W Sadrozinski, Abraham Seiden, et al. Conceptual design of a proton computed tomography system for applications in proton radiation therapy. *IEEE Transactions on Nuclear Science*, 51(3):866–872, 2004.
- [17] RW Schulte, V Bashkirov, R Johnson, HF-W Sadrozinski, and KE Schubert. Overview of the llumc/ucsc/csusb phase 2 proton ct project. *Transactions of the American Nuclear Society*, 106:59, 2012.
- [18] Robert P Johnson, Joel DeWitt, Scott Macafee, Hartmut F-W Sadrozinski, David Steinberg, and Andriy Zatserklyaniy. A fast tracker data acquisition system for pct. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, pages 2529–2531. IEEE, 2012.
- [19] Valentina Giacometti, Vladimir A Bashkirov, Pierluigi Piersimoni, Susanna Guatelli, Tia E Plautz, Hartmut F-W Sadrozinski, Robert P Johnson, Andriy Zatserklyaniy, Thomas Tessonnier, Katia Parodi, et al. Software platform for simulation of a prototype proton ct scanner. *Medical physics*, 44(3):1002–1016, 2017.
- [20] B. Schultze, M. Witt, Y. Censor, K.E. Schubert, and R. Schulte. Performance of hull-detection algorithms for proton computed tomography reconstruction. In S. Reich and A.J. Zaslavski, editors, *Infinite Products of Operators and Their Applications*, volume 636 of *Contemporary Mathematics*, pages 211–224. American Mathematical Society, 2015.

- [21] Blake Schultze, Micah Witt, Keith E Schubert, Robert F Hurley, Vladimir Bashkirov, Reinhard W Schulte, and Ernesto Gomez. Space carving and filtered back-projection as preconditioners for proton computed tomography reconstruction. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, pages 4335–4340. IEEE, 2012.
- [22] Gabor T Herman, Edgar Garduño, Ran Davidi, and Yair Censor. Superiorization: An optimization heuristic for medical physics. *Medical physics*, 39(9):5532–5546, 2012.
- [23] SN Penfold, Reinhard W Schulte, Yair Censor, and Anatoly B Rosenfeld. Total variation superiorization schemes in proton computed tomography image reconstruction. *Medical physics*, 37(11):5887–5895, 2010.
- [24] Blake Schultze, Yair Censor, Paniz Karbasi, Keith E Schubert, and Reinhard W Schulte. An improved method of total variation superiorization applied to reconstruction in proton computed tomography. *arXiv preprint arXiv:1803.01112*, 2018.
- [25] Kirk L Duffin, Nicholas T Karonis, Caesar E Ordonez, Michael E Papka, George Coutrakon, Bela Erdelyi, Eric C Olson, and Thomas D Uram. An analysis of a distributed gpu implementation of proton computed tomographic (pct) reconstruction. In *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion.*, pages 166–175. IEEE, 2012.
- [26] Nicholas T Karonis, Kirk L Duffin, Caesar E Ordoñez, Bela Erdelyi, Thomas D Uram, Eric C Olson, George Coutrakon, and Michael E Papka. Distributed and hardware accelerated computing for clinical medical imaging using proton computed tomography (pct). *Journal of Parallel and Distributed Computing*, 73(12):1605–1612, 2013.
- [27] Caesar E Ordoñez, Nicholas Karonis, Kirk Duffin, George Coutrakon, Reinhard Schulte, Robert Johnson, and Mark Pankuch. A real-time image reconstruction system for particle treatment planning using proton computed tomography (pct). *Physics Procedia*, 90:193–199, 2017.
- [28] Yair Censor, Tommy Elfving, Gabor T Herman, and Touraj Nikazad. On diagonally relaxed orthogonal projection methods. *SIAM Journal on Scientific Computing*, 30(1):473–504, 2008.
- [29] Scott Penfold, Reinhard W Schulte, Yair Censor, Vladimir Bashkirov, Scott McAllister, Keith Schubert, and Anatoly B Rosenfeld. Block-iterative and string-averaging projection algorithms in proton computed tomography image reconstruction. 2010.
- [30] Scott Nicholas Penfold. Image reconstruction and monte carlo simulations in the development of proton computed tomography for applications in proton radiation therapy. 2010.

- [31] Gianfranco Cimmino. Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari. *La Ricerca Scientifica (Roma)*, 1:326–333, 1938.
- [32] Anders H Andersen and Avinash C Kak. Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984.
- [33] Yair Censor, Dan Gordon, and Rachel Gordon. Bicav: A block-iterative parallel algorithm for sparse systems with pixel-related weighting. *IEEE Transactions on Medical Imaging*, 20(10):1050–1060, 2001.
- [34] Dan Gordon and Rachel Gordon. Component-averaged row projections: A robust, block-parallel scheme for sparse linear systems. *SIAM Journal on Scientific Computing*, 27(3):1092–1117, 2005.
- [35] Yair Censor, Stavros Andrea Zenios, et al. *Parallel optimization: Theory, algorithms, and applications*. Oxford University Press on Demand, 1997.
- [36] Yair Censor, Dan Gordon, and Rachel Gordon. Component averaging: An efficient iterative parallel algorithm for large and sparse unstructured problems. *Parallel computing*, 27(6):777–808, 2001.
- [37] Hengyong Yu and Ge Wang. Sart-type image reconstruction from a limited number of projections with the sparsity constraint. *Journal of Biomedical Imaging*, 2010:3, 2010.
- [38] Ming Jiang and Ge Wang. Convergence studies on iterative algorithms for image reconstruction. *IEEE Transactions on Medical Imaging*, 22(5):569–579, 2003.
- [39] Louis Landweber. An iteration formula for fredholm integral equations of the first kind. *American journal of mathematics*, 73(3):615–624, 1951.
- [40] Andreas Kirsch. *An introduction to the mathematical theory of inverse problems*, volume 120. Springer Science & Business Media, 2011.
- [41] Per Christian Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, volume 4. Siam, 2005.
- [42] S. N. Penfold, R. W. Schulte, Y. Censor, V. Bashkurov, S. Macallister, K. E. Schubert, and A. B. Rosenfeld. Block-iterative and string-averaging projection algorithms in proton computed tomography image reconstruction. In Y. Censor, M. Jiang, and G. Wang, editors, *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, Madison, WI, USA, 2010. The Huangguoshu International Interdisciplinary Conference, Medical Physics Publishing.
- [43] Ge Wang and Ming Jiang. Ordered-subset simultaneous algebraic reconstruction techniques (os-sart). *Journal of X-ray Science and Technology*, 12(3):169–177, 2004.

- [44] R Davidi, RW Schulte, Y Censor, and L Xing. Fast superiorization using a dual perturbation scheme for proton computed tomography. *Transactions of the American Nuclear Society*, 106(73p-76):2013, 2012.
- [45] Robert P Johnson. Review of medical radiography and tomography with proton beams. *Reports on Progress in Physics*, 81(1):016701, 2017.
- [46] C Civinini, D Bonanno, M Brianzi, M Carpinelli, GAP Cirrone, G Cuttone, D Lo Presti, G Maccioni, S Pallotta, N Randazzo, et al. Proton computed tomography: iterative image reconstruction and dose evaluation. *Journal of Instrumentation*, 12(01):C01034, 2017.
- [47] Jiahua Zhu and Scott Penfold. Total variation superiorization in dual-energy ct reconstruction for proton therapy treatment planning. *Inverse Problems*, 33(4):044013, 2017.
- [48] M Bruzzi, C Civinini, M Scaringella, D Bonanno, M Brianzi, M Carpinelli, GAP Cirrone, G Cuttone, D Lo Presti, G Maccioni, et al. Proton computed tomography images with algebraic reconstruction. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 845:652–655, 2017.
- [49] Gabor T Herman and Ran Davidi. Image reconstruction from a small number of projections. *Inverse problems*, 24(4):045011, 2008.
- [50] Hengyong Yu, Changguo Ji, and Ge Wang. Sart-type image reconstruction from overlapped projections. *Journal of Biomedical Imaging*, 2011:1, 2011.
- [51] Sea Agostinelli, John Allison, K al Amako, Jo Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G 2 Barrand, et al. Geant4a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [52] Paniz Karbasi, Blake Schultze, Valentina Giacometti, Tia Plautz, Keith E Schubert, Reinhard W Schulte, and Vladimir A Bashkirov. Incorporating robustness in diagonally-relaxed orthogonal projections method for proton computed tomography. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2015 IEEE*, pages 1–4. IEEE, 2015.
- [53] Gene H Golub, Per Christian Hansen, and Dianne P O’Leary. Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1):185–194, 1999.
- [54] Gene H Golub and Charles F Van Loan. An analysis of the total least squares problem. *SIAM journal on numerical analysis*, 17(6):883–893, 1980.
- [55] Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

- [56] K. E. Schubert. *A New Look at Robust Estimation and Identification*. PhD thesis, University of California, Santa Barbara, 2003.
- [57] Lloyd N Trefethen and David Bau III. *Numerical linear algebra*, volume 50. Siam, 1997.
- [58] Gilbert W Stewart. *Matrix perturbation theory*. 1990.
- [59] Harold V Henderson and Shayle R Searle. On deriving the inverse of a sum of matrices. *Siam Review*, 23(1):53–60, 1981.
- [60] Daniel J Tylavsky and Guy RL Sohie. Generalization of the matrix inversion lemma. *Proceedings of the IEEE*, 74(7):1050–1052, 1986.
- [61] Ming Yan. Convergence analysis of sart: optimization and statistics. *International Journal of Computer Mathematics*, 90(1):30–47, 2013.
- [62] Jacobus M Schippers and Antony J Lomax. Emerging technologies in proton therapy. *Acta Oncologica*, 50(6):838–850, 2011.
- [63] Curtis T Rueden, Johannes Schindelin, Mark C Hiner, Barry E DeZonia, Alison E Walter, Ellen T Arena, and Kevin W Eliceiri. Imagej2: Imagej for the next generation of scientific image data. *BMC bioinformatics*, 18(1):529, 2017.
- [64] Paniz Karbasi, Ritchie Cai, Blake Schultze, Hanh Nguyen, Jones Reed, Patrick Hall, Valentina Giacometti, Vladimir Bashkirov, Robert Johnson, Nick Karonis, et al. A highly accelerated parallel multi-gpu based reconstruction algorithm for generating accurate relative stopping powers. *arXiv preprint arXiv:1802.01070*, 2018.
- [65] David C Hansen, Niels Bassler, Thomas Sangild Sørensen, and Joao Seco. The image quality of ion computed tomography at clinical imaging dose levels. *Medical physics*, 41(11), 2014.
- [66] Lennart Volz, Charles-Antoine Collins-Fekete, Pierluigi Piersimoni, Robert P Johnson, Vladimir Bashkirov, Reinhard Schulte, and Joao Seco. Stopping power accuracy and achievable spatial resolution of helium ion imaging using a prototype particle ct detector system. *Current Directions in Biomedical Engineering*, 3(2):401–404.