University of Antwerp

Improved X-ray CT reconstruction techniques with non-linear imaging models

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Science at the University of Antwerp to be defended by

Nathanaël Six

Promotors: Prof. dr. Jan Sijbers &
Prof. dr. Jan De Beenhouwer
Antwerp, 2024

Doctoral committee

Chairman:

Prof. dr. Dirk Lamoen	University of Antwerp

Promotors:

Prof. dr. Jan Sijbers	University of Antwerp
Prof. dr. Jan De Beenhouwer	University of Antwerp

Other members:

Prof. dr. Rudi Penne	University of Antwerp
Prof. dr. Alessandro Olivo	University College London
Dr. Marius Costin	CEA-LIST

Original title: *Improved X-ray CT reconstruction techniques with non-linear imaging models*

Dutch title: *Verbeterde reconstructietechnieken voor X-stralen CT met niet-lineaire beeldvormingsmodellen*

Keywords: *X-ray computed tomography, non-linear optimization, phase contrast imaging, discrete tomography*

Published and distributed by Nathanaël Six. The research presented in this thesis was supported by the Flemish fund for scientific research (FWO) grant: 11D8319N.

Contact information

✉ imec-Visionlab, Department of Physics
University of Antwerp
Universiteitsplein 1, 2610 Antwerp, Belgium
✍ nathanael.six@uantwerpen.be

Acknowledgments

When I applied for a PhD position back in 2017, I had no idea what I was getting myself into. Performing research has been an extremely interesting and exciting experience, but at the same time, I have lost a lot of sleep worrying about confusing results and rapidly approaching deadlines. It has been a long journey to get to this thesis, made more difficult by a disorienting global pandemic. Nevertheless, having made it to the end fills me with happiness and pride. I have met some amazing people and we have had great times together. I am grateful to everyone who has supported me over the past years, professionally or personally and I am excited for what the future will bring for us.

I would like to take this opportunity to thank all the people who helped in making this thesis a reality. First and foremost, to my supervisors, Jan Sijbers and Jan De Beenhouwer. From the moment we first met, your enthusiasm for research has been obvious and infectious. Thank you for giving me the opportunity to work on this project and for believing in me through all these years. Furthermore, to the members of my doctoral jury, thank you for being a part of this journey and for the time and effort you have spent on reviewing my work. A special thanks goes out to Alessandro Olivo, as well as Charlotte Hagen and everyone at the AXIm group at University College London, for being so kind to welcome me in London and giving me the chance to learn and grow at a different lab, as well as supporting the growing edge illumination group at Vision Lab.

Next, I would like to thank my friends and colleagues at the Vision Lab: Tim, Daniel, Jonathan, Jens, Pavel, Nicholas, Pieter-Jan and countless others. You were always there for discussions on any topic, work or otherwise, and your camaraderie has made these years so much more enjoyable.

I would also like to take a moment to sent my thanks out to my

family and my friends. First to my parents, for their unending support, patience and love throughout so many years. Next, to Oma, Opa and Noémie, for always being there for me without question. To all my close friends, I am forever grateful for your support and encouragement during these demanding times. A special mention goes out to Jeffrey, for his unwavering belief that I would finish this project. It took me a while, but I could eventually prove you right.

Of course the last and most important thank you goes out to Maíra, for being my partner in the most difficult time. Your support has carried me to the end.

Thank you all for being a part of this incredible journey.

Nathanaël Six
Antwerp, January 2024

Samenvatting

X-ray computertomografie (CT) is een krachtige en niet-invasieve techniek om de interne structuur van een object te visualiseren aan de hand van een reeks X-ray radiografieën. Reconstructiealgoritmes worden gebruikt om vanuit projectiedata een 3D-representatie van het object op te stellen. Een model van het X-ray acquisitieproces wordt gebruikt door reconstructiealgoritmes en deze algoritmes vereisen een groot aantal projecties om goed te functioneren. In bepaalde toepassingen moet het aantal projecties echter beperkt worden, om de totale afgeleverde dosis te verminderen, de acquisitietijd te verkorten, of vanwege geometrische beperkingen. Bovendien hebben de meest gebruikte algoritmes een eenvoudig lineair voorwaarts model voor X-ray attenuatie dat de werkelijke acquisitie niet nauwkeurig modelleert. Ten slotte zijn conventionele reconstructiealgoritmes in CT niet efficiënt wat betreft rekenkracht. In deze thesis zullen we verbeterde reconstructiealgoritmes voor CT ontwikkelen door nauwkeurigere niet-lineaire voorwaartse modellen en verschillende numerieke optimalisatiemethodes voor deze modellen te onderzoeken. Zowel conventionele X-ray CT als fasegevoelige X-ray CT zullen worden onderzocht. In fasegevoelige X-ray CT zijn aanvullende contrasten beschikbaar omdat de X-ray beeldvormingsopstelling gevoelig is gemaakt voor refractie-effecten van de X-ray bundel. Hieronder volgt een overzicht van de verschillende hoofdstukken van de thesis.

In Hoofdstuk 1 worden de concepten geïntroduceerd die nodig zijn voor deze thesis. Eerst worden de basisprincipes van X-ray beeldvorming en X-ray CT beschreven. Vervolgens wordt een overzicht gegeven van de effecten van undersampling en modelfouten op de reconstructie en worden gevestigde methoden om met deze effecten om te gaan beschreven. Daarna worden de basisconcepten van numerieke optimalisatie voor CT geïntroduceerd. Verschillende gradient descent- en quasi-Newtonmethoden worden beschreven, omdat ze een cruciale rol zullen spelen in

wat volgt. Ten slotte wordt een fasegevoelige X-ray beeldvormingsmodaliteit, genaamd edge illumination, geïntroduceerd en de oorsprong van het nieuwe contrast wordt beschreven. De standaard verwerkingsprocedures om de contrasten te ontwarren worden uitgelegd en de standaard reconstructietechnieken die worden gebruikt voor edge illumination CT worden getoond.

In Hoofdstuk 2 wordt een nieuw reconstructiealgoritme voorgesteld, genaamd poly-DART, om om te gaan met undersampled data van polychromatische bronnen, gebaseerd op een heuristisch discreet tomografisch reconstructiekader, DART. DART is een reconstructiekader waarin voorkennis over het aantal objectmaterialen wordt benut. In monochromatische X-ray CT is aangetoond dat DART leidt tot hoogwaardige reconstructies, zelfs vanuit een laag aantal projecties of een beperkte scankijk. De meeste X-ray bronnen zijn echter polychromatisch, wat leidt tot beam hardening-effecten die de prestatie van DART aanzienlijk verslechteren. Het voorgestelde discrete poly-DART-algoritme benut de voorkennis over de attenuatiewaarden met behulp van DART en houdt tegelijkertijd rekening met de polychromatische aard van de X-ray bron. De resultaten tonen aan dat poly-DART leidt tot een aanzienlijk verbeterde segmentatie op polychromatische data verkregen uit Monte Carlo simulaties en op experimentele data, in vergelijking met DART.

In Hoofdstuk 3 wordt een verbetering van de partitioneringsstap in DART voorgesteld. Veel op discrete tomografie gebaseerde methoden vertrouwen op een verdeel-en-heersprocedure om het volume in delen te reconstrueren, wat hun looptijd en reconstructiekwaliteit verbetert. De procedure is echter gebaseerd op statische regels, wat redundante berekening introduceert en de efficiëntie vermindert. In dit hoofdstuk wordt een update-strategiekader geïntroduceerd dat dynamische regels mogelijk maakt en meer controle biedt voor verdeel-en-heersmethoden voor discrete tomografie. Het kader wordt geïllustreerd door Tabu-DART te introduceren, dat het voorgestelde kader combineert met DART. Uit gereconstrueerde experimenten met gesimuleerde en echte data blijkt dat de benadering vergelijkbare of verbeterde reconstructiekwaliteit oplevert in vergelijking met DART, met aanzienlijk lagere rekencomplexiteit. Bovendien kan de voorgestelde strategie worden gecombineerd met het voorgestelde poly-DART-algoritme uit Hoofdstuk 2, wat de techniek verbetert.

In Hoofdstuk 4 wordt het gebruik van quasi-Newton-methoden onderzocht om een polychromatische doelfunctie te minimaliseren zonder de noodzaak om het beeld in verschillende materiaalgebieden te segmenteren. De voorgestelde doelfunctie kan ook eenvoudig worden uitgebreid met regularisatietermen in een wiskundig kader. Er wordt aangetoond

dat de quasi-Newton-methoden andere statistische of algebraïsche reconstructietechnieken kunnen overtreffen. De reconstructiekwaliteit en de projectiefout voor reconstructies van zowel Monte Carlo-gesimuleerde data als experimentele data worden onderzocht. Van de overwogen quasi-Newton-methoden blijkt Gauss-Newton-Krylov het beste te presteren. Vergeleken met een recent voorgestelde polychromatische algebraïsche reconstructietechniek bereiken quasi-Newton-methoden een lagere reconstructiefout en hebben ze een verhoogde convergentiesnelheid.

In Hoofdstuk 5 wordt een nieuwe gezamenlijke reconstructiemethode voorgesteld, die gebruikmaakt van een gecombineerd voorwaarts model om de drie contrasten die verkregen kunnen worden met edge illumination-beeldvorming simultaan te reconstrueren, zonder de noodzaak van een preprocessing stap. De gezamenlijke reconstructiebenadering vermindert reconstructietijden, omdat de preprocessing stap wordt overgeslagen en maakt een veel flexibeler acquisitieschema mogelijk, omdat er geen noodzaak is om bij elke projectiehoek een volledige illuminatiecurve te samplen. De voorgestelde doelfunctie wordt geminimaliseerd met behulp van verschillende niet-lineaire optimalisatiemethoden, en een gesplitste Barzilai-Borwein-gradiëntmethode wordt geïdentificeerd als de meest praktische kandidaat. In vergelijking met de state-of-the-art verbetert onze methode de reconstructiekwaliteit op subsampled datasets, waardoor de reconstructie van drie contrasten uit single-shot datasets mogelijk is.

In Hoofdstuk 6 besluiten we de thesis door onze algemene bevindingen uit de verschillende onderzoekshoofdstukken te beschrijven. Verder geven we een overzicht van mogelijkheden voor toekomstig onderzoek gebaseerd op het werk in deze thesis.

Summary

X-ray computed tomography (CT) is a powerful and non-invasive technique to visualise the internal structure of an object from a set of X-ray radiographs. Reconstruction algorithms are used to map projection data to a 3D volume. A model of the X-ray acquisition process is used by reconstruction algorithms and the algorithms require a large number of projections to function well. However, in certain applications, the number of projections has to be limited, to reduce total delivered dose, lower acquisition time or because of geometrical constraints. Furthermore, the most commonly used algorithms have a simple linear forward model for X-ray attenuation that does not model the real acquisition accurately. Finally, conventional reconstruction algorithms in CT are not efficient with respect to computation time. In this thesis, we will develop improved reconstruction algorithms for CT by investigating more accurate non-linear forward models, and different numerical optimisation approaches for these models. Both conventional X-ray CT and phase contrast X-ray CT will be considered. In phase contrast X-ray CT, additional contrasts are available since the X-ray imaging setup is made sensitive to refraction effects of the X-ray beam. In what follows, an overview of the different chapters of the thesis is given.

In Chapter 1, the concepts needed for this thesis are introduced. First, the basics of X-ray imaging and X-ray CT are described. Next, an overview is given of the effects undersampling and model errors have on the reconstruction and established methods to deal with these effects are examined. Then, basic concepts of numerical optimisation for CT are introduced. Different gradient descent and quasi-Newton methods are described, as they will play a crucial role in what follows. Finally, a phase sensitive X-ray imaging modality, called edge illumination, is introduced and the origin of the new contrast is described. The standard processing procedures to disentangle the contrasts are explained and the

standard reconstruction techniques used for edge illumination CT are shown.

In Chapter 2, a novel reconstruction algorithm, called poly-DART, is proposed, to deal with undersampled data from polychromatic sources, based on a heuristic discrete tomographic reconstruction framework, DART. DART is a reconstruction framework in which prior knowledge on the number of object materials is exploited. In monochromatic X-ray CT, DART has been shown to lead to high-quality reconstructions, even with a low number of projections or a limited scanning view. However, most X-ray sources are polychromatic, leading to beam hardening effects, which significantly degrade the performance of DART. The proposed discrete poly-DART algorithm exploits sparsity in the attenuation values using DART and simultaneously accounts for the polychromatic nature of the X-ray source. The results show that poly-DART leads to a vastly improved segmentation on polychromatic data obtained from Monte Carlo simulations and on experimental data, compared to DART.

In Chapter 3, an improvement to the partitioning step in DART is proposed. Many discrete tomography-based methods rely on a divide-and-conquer procedure to reconstruct the volume in parts, which improves their run-time and reconstruction quality. However, the procedure is based on static rules, which introduces redundant computation and diminishes the efficiency. In this work, an update strategy framework is introduced that allows dynamic rules and increases control for divide-and-conquer methods for discrete tomography. The framework is illustrated by introducing Tabu-DART, which combines the proposed framework with DART. Through simulated and real data reconstruction experiments, it is shown that the approach yields similar or improved reconstruction quality compared to DART, with substantially lower computational complexity. Furthermore, the proposed strategy can be combined with the proposed poly-DART algorithm from Chapter 2, improving the technique.

In Chapter 4, the use of quasi-Newton methods to minimise a polychromatic objective function is investigated, without the need for segmenting the image into different material regions. The proposed objective function can also be easily extended with regularisation terms in a mathematically sound framework. It is shown that the quasi-Newton methods can outperform other statistical or algebraic reconstruction techniques. The reconstruction quality and the projection error for reconstructions of both Monte Carlo simulated data and experimental data are investigated. From the considered quasi-Newton methods, Gauss-Newton-Krylov is found to perform best. Compared to a recently proposed polychromatic algebraic reconstruction technique, quasi-Newton solvers reach a lower reconstruction error and have an increased convergence speed.

In Chapter 5, a novel joint reconstruction method is proposed, which utilises a combined forward model to reconstruct the three contrasts obtainable with edge illumination imaging simultaneously, without the need for a retrieval step. The joint reconstruction approach reduces reconstruction times, as the retrieval step is skipped, and allows for a much more flexible acquisition scheme, as there is no need to sample a full illumination curve at each projection angle. The proposed objective function is minimised using different non-linear optimisation approaches and a split Barzilai-Borwein gradient method is identified as the most practical candidate. Compared to the state-of-the-art, our method improves reconstruction quality on subsampled datasets, allowing the reconstruction of three contrasts from single-shot datasets.

In Chapter 6, we conclude the thesis by describing our general findings from the different research chapters. Furthermore, we provide an outlook on opportunities for future research built upon the work in this thesis.

Contents

CHAPTER 1

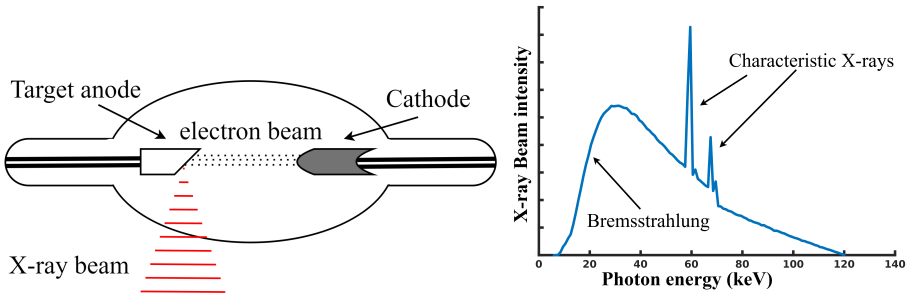
Introduction

1.1 Fundamentals of X-ray CT imaging

X-ray computed tomography (CT) is a non-destructive technique which attempts to reconstruct the internal structure of an object from a set of X-ray projections taken from different angles. This section outlines the primary principles of tomographic X-ray imaging. The basics of X-ray generation and attenuation are described, explaining how single X-ray projections are formed. Next, we explore tomographic geometries and basic reconstruction algorithms. For an in-depth exploration of these topics, the readers are referred to [1] and [2].

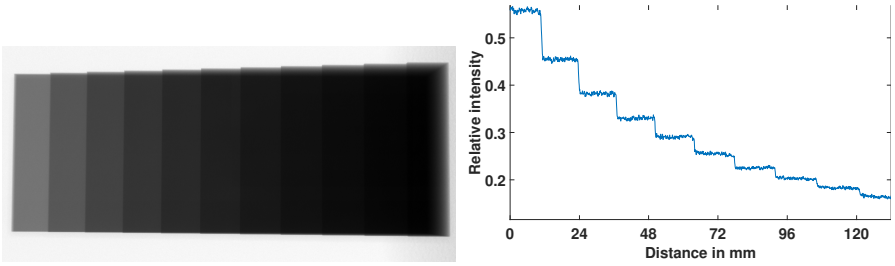
1.1.1 X-ray generation

In both medical and laboratory CT systems, X-rays are typically produced using an X-ray tube, as illustrated in Fig. 1.1a. The generation process commences by heating a cathode to around 2400 K, sufficient to overcome the binding energy of electrons, thus leading to their emission. The emitted electrons are then accelerated by a high voltage, generally in the range of 10 to 300 kV, applied between the cathode and anode, and will subsequently collide with the metal anode. Upon collision, the electrons interact with the matter of the anode and decelerate quickly and these interaction generate X-rays. An example of an X-ray spectrum, i.e. the distribution of different photon energies emitted by an X-ray tube, is shown in Fig. 1.1b. The continuous line is comprised of the bremsstrahlung and the discrete high peaks are from characteristic radiation. The physical processes generating these two types of radiation are different. Bremsstrahlung is the result of electrons slowing down and deflecting as they move through the anode material. At each interaction, the electron loses energy, and this lost energy is emitted in the form of an X-ray photon. Characteristic X-rays are generated when an incoming



(a) (b)

Figure 1.1. Schematic drawing of an X-ray tube (a) and example of an emitted X-ray spectrum (b).



(a) (b)

Figure 1.2. X-ray projection (a) with line profile (b) of an aluminium step wedge. The first step is 5 mm thick, all subsequent steps add 3 mm.

electron ejects an inner-shell electron from an atom in the anode. A higher-shell electron will move to the inner shell and in the process emit a photon. The emitted photons only have a limited number of possible energies, which are dependent on the material of the anode, and therefore show up as discrete peaks in the emitted X-ray spectrum.

1.1.2 Attenuation of X-rays by matter

X-rays are known for their high capability to penetrate matter. However, as an X-ray beam passes through an object, it loses intensity exponentially. An X-ray projection of a step phantom is shown in Fig. 1.2a. The phantom is made of aluminium and looks like a staircase, with each step adding 3 mm of additional material. It is clear from the line profile that when the steps increase in thickness linearly, the intensity loss happens exponentially.

The amount of intensity lost depends on several factors, including the distance the beam travelled through the material, the atomic com-

position of the material, and the density of the material. In the case of a homogeneous object, these last two factors can be combined into a single value, the attenuation coefficient μ of the object. For non-homogeneous objects, the attenuation value is dependent on the position in the object, so it becomes a spatial function $\mu(\boldsymbol{s})$, where $\boldsymbol{s}$ represents the position within the object. The Beer-Lambert law offers a formula that expresses how much of the X-ray beam will be absorbed when passing through a homogeneous material. Let I be the intensity of the beam after travelling through the material, l the distance travelled, and I_0 the original intensity of the beam, the Beer-Lambert law is then expressed as follows:

$$I = I_0 \mathrm{e}^{-\mu l}. \tag{1.1}$$

To describe the attenuation of a non-homogeneous object, the Beer-Lambert law can be straightforwardly extended using a line integral. For each infinitesimal length along the line, the attenuation is given by the homogeneous Beer-Lambert law in Eq. 1.1. So, to give the complete attenuation of the X-ray beam along a line, all the infinitesimal attenuations need to be summed up as follows:

$$I = I_0 \mathrm{e}^{-\int \mu(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}}, \tag{1.2}$$

with $\boldsymbol{s}$ the position vector in object space. As shown in Fig. 1.1b, an X-ray tube usually emits a wide spectrum of X-rays. In reality, photons with different energies interact with matter differently, thus the attenuation coefficient is energy dependent. Again, the attenuation model can be updated to reflect the energy dependence. Assume that the spectrum has minimal and maximal energy, ϵ_{min} and ϵ_{max} respectively. Let $I_0(\epsilon)$ be the original energy-dependent intensity of the beam and let $\mu(\boldsymbol{s}, \epsilon)$ be the energy-dependent spatial function of the attenuation coefficients in the object, then the polychromatic Beer-Lambert law along a line is given as:

$$I = \int_{\epsilon_{min}}^{\epsilon_{max}} I_0(\epsilon) \mathrm{e}^{-\int \mu(\boldsymbol{s},\epsilon)\,\mathrm{d}\boldsymbol{s}} \mathrm{d}\epsilon. \tag{1.3}$$

In practice, the monochromatic formula in Eq. 1.2 is the expression most commonly used to model X-ray attenuation. It is an approximation of Eq. 1.3 at one so-called *effective* energy in the spectrum. The monochromatic formula is used often because Eq. 1.2 can be easily linearised, and Eq. 1.3 can not. Linearising Eq. 1.2 is done by transforming it to the following equivalent equation:

$$-\ln\left(\frac{I}{I_0}\right) = \int \mu(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}. \tag{1.4}$$

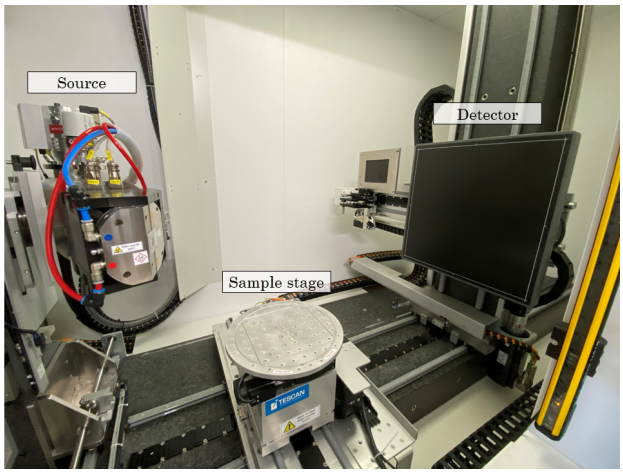


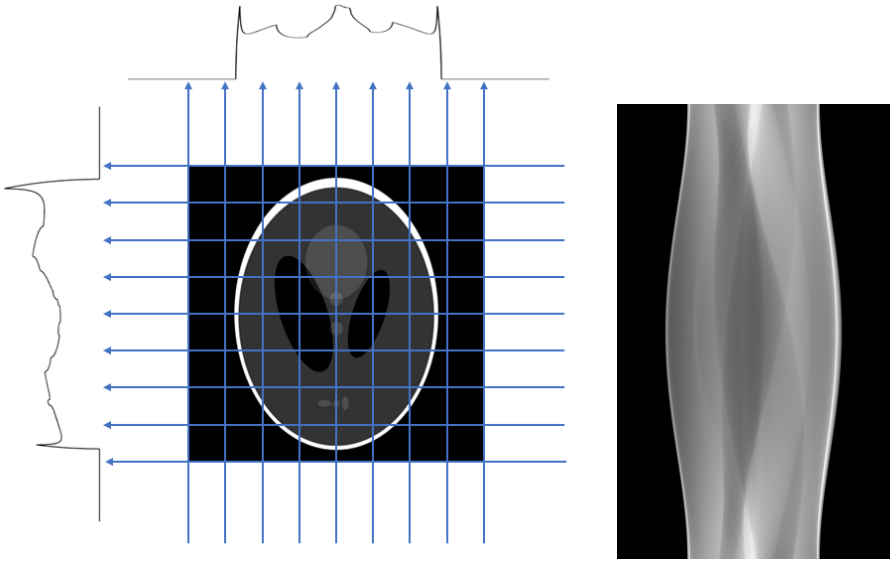
Figure 1.3. FleXCT scanner system with rotating sample stage at Vision Lab, University of Antwerp with the main parts indicated.

The division by I_0 is called flatfielding or normalising, as it transforms the measured intensity image I into a relative intensity, and the image I_0, which is the projection image without sample present, is called the flatfield. Taking the negative natural logarithm is called log-correcting. The same technique of flatfielding and log-correcting cannot be used to linearise Eq. 1.3, as there is no closed-form inverse function for general sums or integrals of exponential functions.

The next section will discuss how the inverse problem can be solved, that is, how attenuation values $\mu(\boldsymbol{s})$ can be reconstructed from measured intensities I.

1.1.3 Tomographic imaging

With X-ray imaging, it is possible to acquire 2D images of an object, called radiographs. Radiographs are projections or X-ray shadows of the object, where the detected intensity is predicted by the Beer-Lambert law. It is clear that it is impossible to infer from a single projection value which specific distribution of attenuation values along the ray led to the measured loss in intensity. There are infinitely many different solutions to the problem. A clear drawback of radiography is, therefore, that all spatial information in the projection direction is lost as the 3D object is projected onto a 2D plane. To regain the spatial information, multiple measurements are required. If the position of the X-ray source and detector relative to the object is changed, then the lines along which the attenuation values are integrated change and a new measurement of the same object is



(a) **(b)**

Figure 1.4. Figure showing (a) two orthogonal parallel beam projections of a 2D Shepp-Logan phantom and (b) the sinogram containing 1000 parallel beam projections of the Shepp-Logan phantom, with projection angles evenly spaced in $[0, 180°]$.

obtained. Therefore, multiple projections are acquired while the source and detector are rotated around the sample (or, equivalently, the sample itself rotates), leading to X-ray tomography. An example of an X-ray scanner, the FleXCT system at Vision Lab, University of Antwerp [3], is shown in Fig. 1.3. Since each measurement is of the same spatial function $\mu(\boldsymbol{s})$ of attenuation values in the object, given enough measurements, it is possible to retrieve a full 3D representation of the object, through the use of a CT reconstruction algorithm. Such algorithms depend on the projection geometry that is used in the acquisition. In 2D tomography, i.e. projections of 2D slices onto a 1D line detector, the most common geometries are parallel beam or fan beam. In a parallel beam geometry all lines from source to detector along which the projection is taken are parallel to each other - see Fig. 1.4, while in a fan beam geometry the lines all originate from a single source point and fan out towards the detector. In 3D equivalent projection geometries exist, parallel beam and cone beam. For a 3D parallel beam, the rays are detected at a plane instead of a line, thus can be seen as a stack of 2D parallel beam geometries. With a cone beam geometry the rays form a cone instead of a fan as they go from the point source to the detector plane. The central slice of a

cone beam projection is the same as that of a fan beam projection from the same point source. The next section describes how to reconstruct an object from its projections.

1.1.4 Computed tomography

Given a set of tomographic projections, the goal of computed tomography (CT) algorithms is to reconstruct the 3D object. In this section the projection geometry is first assumed to be parallel beam and the 2D case will be presented and it is noted whenever the presented algorithms would need to be adapted for other projection geometries. The problem of reconstructing a 3D volume from 2D projections in a 3D parallel beam geometry is equivalent to reconstructing a 2D slice from 1D projections in a 2D parallel beam geometry multiple times. A linear physical projection model for an X-ray image along a single ray is given by the Beer-Lambert law in Eq. 1.4. After imaging, the left-hand side of Eq. 1.4 is known, as it is the projection data, while the spatial function $\mu(\boldsymbol{s})$ is our unknown object to be reconstructed. The projection operator of the 2D object to 1D projections is mathematically described by the Radon transform of the function μ. Any line L in 2D space, can be identified by two parameters, its distance from the origin and the angle it makes with the y-axis, η and θ, respectively. All lines in a 2D space can then be written in the form:

$$L(\eta, \theta) = \{(x, y) \in \mathbb{R}^2 : \eta = x\cos(\theta) + y\sin(\theta)\}. \tag{1.5}$$

The line integral of a function $f : \mathbb{R}^2 \to \mathbb{R}$ along $L(\eta, \theta)$:

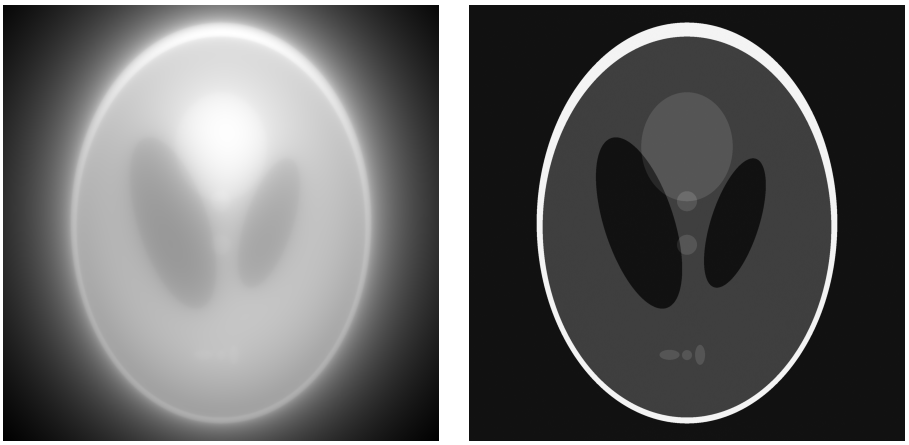
$$\begin{aligned} p(\eta, \theta) &= \int_{L(\eta,\theta)} f(\boldsymbol{s})\mathrm{d}\boldsymbol{s} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\delta(x\cos(\theta) + y\sin(\theta) - \eta)\mathrm{d}x\mathrm{d}y, \end{aligned} \tag{1.6}$$

with $\delta(0) = 1$ and $\delta(x) = 0$ for every $x \in \mathbb{R}_0$. The Radon transform of f is the function $\mathcal{R}f$ that maps each line $L(\eta, \theta)$ to the line integral of f along that line. From Eq. 1.4 it is clear that $\mathcal{R}\mu$, when μ is the function of attenuation values in space of an object, is the collection of all possible flatfielded and log-corrected X-ray projections of μ.

Analytical reconstruction techniques

The simplest reconstruction method is the basic back projection of the sinogram:

$$f(x, y) = \int_0^{\pi} p(x\cos\theta + y\sin\theta, \theta)d\theta \tag{1.7}$$



(a) **(b)**

Figure 1.5. Reconstructions of the Shepp-Logan phantom from the sinogram data in Fig. 1.4b, using (a) back projection and (b) filtered back projection.

The back projection smears the projected data back along each corresponding projection direction, resulting in a blurred reconstruction, as can be seen in Fig 1.5a. The back projection is unsuited for use as a reconstruction technique directly, however, it is an integral part of other reconstruction techniques.

Using the Fourier slice theorem [2], it can be proven that the following formula gives an exact inverse of the Radon transform:

$$f(x,y) = \int_0^{\pi} \int_{-\infty}^{\infty} P(q,\theta))|q| e^{2\pi i q(x\cos\theta + y\sin\theta)} \mathrm{d}q \mathrm{d}\theta, \tag{1.8}$$

with $P(q,\theta)$ the Fourier transform of $p(\eta,\theta)$. Applying the formula (1.8) to projection data gives back the function μ and therefore is a reconstruction algorithm, called filtered back projection (FBP). Following Eq. 1.8, the FBP algorithm consists of the following steps:

1. Calculate the Fourier transform of the measured sinogram $p(\eta,\theta)$, giving $P(q,\theta)$.
2. Multiply with a ramp-filter $|q|$, giving $|q|P(q,\theta)$
3. Calculate the inverse Fourier transform of $|q|P(q,\theta)$, resulting in a filtered sinogram.
4. Back project the filtered sinogram to obtain a reconstruction of the object.

In Fig. 1.5b the FBP reconstruction of the sinogram data from Fig. 1.4b is shown. Although FBP is theoretically the exact solution to the inverse problem, this only holds in the case where all projection data is available, i.e. the continuous projection of the function along every possible line in space is known. In practice, only a finite number of line integrals is calculated, one for each pixel on the detector, for each projection angle. The quality of the reconstruction can be severely degraded when only a small number of projections or noisy projections are available. Furthermore, the algorithm only holds for parallel beam geometries. X-ray setups following a parallel beam geometry are present at synchrotron facilities, large particle accelerators which can provide high intensity X-ray beams, however, for laboratory imaging setups, a fan-beam or cone-beam geometry is the correct model. Therefore, other analytical formulas need to be used for data from lab-based systems. However, FBP is widely used in practice as it is not a computationally intensive algorithm and therefore results in fast reconstructions. Furthermore, when a lot of high-quality data is available, FBP can also result in high-quality reconstructions.

Algebraic reconstruction techniques

There is a disconnect between the model used in analytical methods and the reality of tomographic imaging and reconstruction on a computer. The use of the Radon transform as a model for the acquisition process starts from the assumption that everything is continuous. In practice, all parts of the imaging and reconstruction framework are discrete: the detector has a finite number of pixels with a certain width and the image grid to be reconstructed on, stored in computer memory, will have a finite number of pixels. Algebraic reconstruction techniques model the problem using discretised expressions already. In the case where the linear Beer-Lambert expression Eq.,1.4 is used as the physical model, the following linear system of equations is found:

$$\boldsymbol{p} = \boldsymbol{A}\boldsymbol{x}. \tag{1.9}$$

Here, $\boldsymbol{x} \in \mathbb{R}^{n_{\text{vol}}}$ represents the lexicographically vectorised image, with $n_{\text{vol}} = n_x * n_y$, the number of voxels. Each element of $\boldsymbol{x}$ corresponds to the average linear attenuation coefficient $\mu(\boldsymbol{s})$ within the corresponding continuous region. The normalised and log-corrected projection $\boldsymbol{p}$, is a vector in $\mathbb{R}^{n_{\text{proj}}}$, with $n_{\text{proj}} = n_{\text{det}} * n_{\text{angles}}$, the product of the number of detector pixels and the number of projection angles. The projection matrix $\boldsymbol{A} \in \mathbb{R}^{n_{\text{proj}} \times n_{\text{vol}}}$ is the matrix that models the projection as a weighted sum of the elements in $\boldsymbol{x}$. The elements a_{ij} in $\boldsymbol{A}$ are the

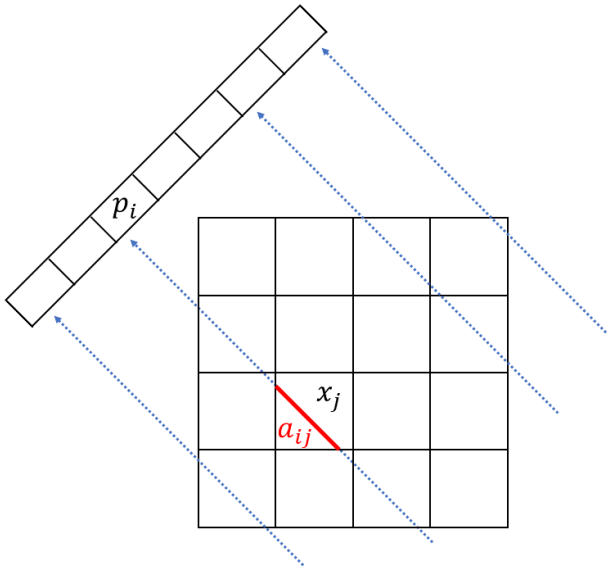


Figure 1.6. Graphical representation of the weights a_{ij} found in the projection matrix $\boldsymbol{A}$. The intersection length of the ray from the source to the detector pixel p_i with the voxel x_j is shown in red.

intersection lengths of the ray l_i with the voxel x_j. For a single parallel beam projection, a graphical representation is given in Fig. 1.6. For the projection matrices associated with a fan or cone beam geometry the concept is the same, but as the lines to each detector element have a differently angled path, the intersection lengths will be different compared to parallel beam. Multiplication with the projection matrix $\boldsymbol{A}$ can be seen as a discretised version of the Radon transform, where the function exists on a finite grid and the transform only considers a finite number of rays. Additionally, multiplication with the transposed projection matrix $\boldsymbol{A}^\top$ is a weighted distribution of the projection values on the pixels, which is a discretised version of the back projection. The goal of any reconstruction algorithm is to find a solution to Eq. 1.9. In general, the inverse of the matrix $\boldsymbol{A}$ does not exist, so the system does not have a true solution. Instead, a *best* solution $\boldsymbol{x}^*$ can be defined as the solution to a minimisation problem:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{p}\|, \tag{1.10}$$

with $\|\cdot\|$ some vector norm, typically a squared 2-norm. In case of the squared 2-norm, the best solution, then called the least-squares approx-

imation, can be found using the Moore-Penrose inverse as [4]:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{p}\|_2^2 = (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{p}. \tag{1.11}$$

However, the exact solution cannot be constructed for most real CT problems as the matrix $\boldsymbol{A}^\top \boldsymbol{A}$ is too large and ill-conditioned, as A is ill-conditioned. Algebraic reconstruction techniques will instead iteratively approximate the solution $\boldsymbol{x}^*$. Many iterative reconstruction techniques have been developed for CT, such as the Algebraic Reconstruction Technique (ART), the Simultaneous Iterative Reconstruction Technique (SIRT), and the Simultaneous Algebraic Reconstruction Technique (SART)[5, 6, 7]. However, any optimisation algorithm for linear problems can be used to solve Eq. 1.10. For example in [8] it is shown that it is possible to achieve the same reconstructions as SIRT with less computational work, with the use of Krylov optimisation techniques. In this section, SIRT [6] will be introduced , which is a Landweber algorithm to solve Eq.,1.10. More advanced algorithms for CT reconstructions will be discussed in Section 1.2. Starting with an initial guess $\boldsymbol{x}^{(0)}$, for example the 0-vector, SIRT provides new approximations for the reconstruction iteratively as follows:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \boldsymbol{C}\boldsymbol{A}^\top \boldsymbol{R}\left(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{p}\right), \tag{1.12}$$

with $\boldsymbol{C} = (c)_{ij}$ and $\boldsymbol{R} = (r)_{ij}$ diagonal matrices of inverse column and row sums, respectively.

$$c_{ii} = \left(\sum_k a_{ki}\right)^{-1}, \quad r_{ii} = \left(\sum_k a_{ik}\right)^{-1}. \tag{1.13}$$

As the operators $\boldsymbol{A}$ and $\boldsymbol{A}^\top$ can be seen as forward and back projection, the steps in the SIRT algorithm can be described as follows:

1. From the current guess $\boldsymbol{x}^{(k)}$ simulated projections $\boldsymbol{A}\boldsymbol{x}^{(k)}$ are made.
2. The simulated projections are compared to the measured data to find the *projection difference*: $\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{p}$.
3. The projection difference is then weighted with $\boldsymbol{R}$, back projected, and weighted with $\boldsymbol{C}$.
4. The resulting correction term is subtracted from $x^{(k)}$.
5. Repeat steps 1-4 until a stopping criterion is reached.

The SIRT algorithm solves the weighted least squares problem:

$$\arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{p}\|_{\boldsymbol{R}}^2, \tag{1.14}$$

and can be shown to always converge to the reconstruction with minimal error in $\boldsymbol{R}$-norm [9].

1.2 CT artefacts and specialised reconstruction techniques

In Section 1.1 basic reconstruction techniques for X-ray CT, FBP and SIRT were described. Both methods are based on the linearised monochromatic Beer-Lambert law (1.4). The model makes a number of assumptions of the imaging setup, for example, the source being a point-source, the beam being monochromatic and the object being stationary. Whenever the model does not align with reality, errors in the model will be transferred to the reconstruction as artefacts [10, 11, 1]. Additionally, for high-quality reconstructions, an adequate number of projections has to be acquired. Otherwise, the reconstruction problem is too underdetermined, and underdetermination will also result in artefacts in the reconstructed images. Artefacts obscure the interior structure, leading to a (partially) unusable scan or, worse, can lead to misinterpretation of the artefact as a feature of the real object [12]. In this section, an overview of common sources of reconstruction artefacts and their associated effect on the reconstruction will be provided. Next, algorithms designed to reduce the effects of certain artefacts are described.

1.2.1 CT reconstruction artefacts

Sources of CT reconstruction artefacts can be broadly divided into three categories:

- *Calibration errors* occur when information about the state of scanner is incorrect or unkown.
- *Undersampling* takes place when the angular sampling is insufficient.
- *Model errors*, where the linear Beer-Lambert model no longer accurately describes the real image formation process.

Each of these scenarios will be investigated separately in the following sections. Of course, in an experimental setup, multiple sources of errors

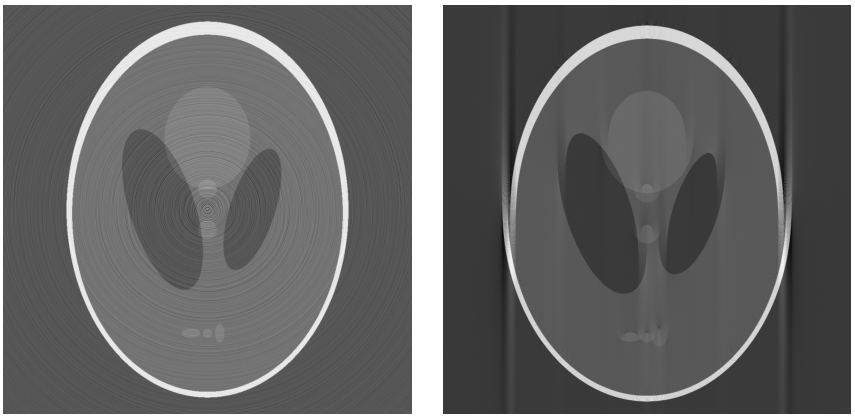
Figure 1.7. Shepp-logan phantom on 1024×1024 pixel grid.

can be simultaneously present, further degrading the quality of the reconstruction. The artefacts discussed here will be demonstrated using simulated reconstructions. From the 1024×1024 Shepp-Logan phantom, shown in Fig. 1.7, simulated projections with 1449 detector pixels will be created. Different sources of reconstruction artefacts will be present in different simulated datasets. CT reconstructions will be made using FBP from the imperfect projection data. Unless otherwise stated, 1000 simulated projections are used.

Calibration errors

The first source of error is the X-ray scanner itself. From the equation of the Beer-Lambert law 1.4 and the associated linear problem 1.9, the parameters and properties of the scanner define both I_0, the flatfield intensity, and A the projection matrix dependent on the chosen scanning geometry. Computing a reconstruction of the scanned data with I_0 or A wrongly measured will result in inaccurate reconstructions. For I_0, also known as the flatfield intensity, projections are usually taken before and/or after scanning the object, as I_0 represents the intensity measured when no attenuation occurs between source and detector. In general, I_0 is different for different detector pixels, as the detector sensitivity is not uniform. Additionally, detector pixels can malfunction over time, leading to dead pixels. When I_0 is not measured accurately, lighter and darked bands are still visible on the flat-field corrected sinogram. The bands give rise to light and dark rings in the reconstruction [13, 10, 11]. A simulated example of such ring artefacts is shown in Fig. 1.8a.

The projection matrix A is incorrectly constructed when the geometric information about the trajectory of the scanner is incorrect. Some



(a) **(b)**

Figure 1.8. Reconstructions of the Shepp-Logan phantom using FBP: (a) from projections where I_0 has unknown variations, (b) from projections where the COR is wrongly estimated by 10 pixels.

examples of geometric parameters that can be changed are the distances between source, object, and detector, the rotational step, and the centre of rotation (COR). Typically, all required parameters are reported by the scanning system. A common source of error is an incorrectly documented COR, which even for small misalignments can give severe reconstruction artefacts [14]. When the COR is misaligned, the motion of the object between projections is incorrectly described, leading to smearing and doubling artefacts throughout the reconstruction. An example of the reconstruction of the Shepp-Logan phantom with a misaligned COR of 10 pixels is shown in Fig. 1.8b.

In general, calibration errors can lead to severe artefacts in reconstructions. However, for standard acquisition setups, the effects can be minimised by adequate flatfielding and proper system maintenance.

Undersampling

Tomographic reconstruction requires a large number of projections to generate a high-quality image. For FBP, a large number of projections is required to accurately approximate the continuous problem the algorithm solves. For algebraic methods, enough samples must be taken to ensure that the minimisation problem in (1.10) is not underdetermined. When the number of projections is low or the projections do not cover the full angular range, this underdetermination will lead to reconstruction artefacts, as many local minima of the inverse reconstruction problem

(1.10) exist.

When only a limited number of projections are acquired, it is called *few-view tomography*. In Fig. 1.9 FBP reconstructions of few-view datasets are shown, with an increasing number of projections. Streaking artefacts appear in a pattern formed by backprojection. As more projections become available, the artefact pattern becomes less severe and more details in the reconstruction can be discerned. In the reconstruction from five projections, close to no information about the object can be inferred. At 10 projections, the general shape of the outside and the largest inner features can be seen. In the reconstructions from 20 and 50 projections, most of the structures can be seen, although the undersampling artefacts are still strongly present.

Another potential way undersampling can occur is when the range in which projections are taken is reduced, also called a *missing wedge*. In Fig. 1.10 reconstructions from such a dataset are shown. For both reconstructions 1000 projections are used, however, instead of sampling the full $[0, 180°]$ range, the subsets $[5°, 175°]$ and $[20°, 160°]$ are considered. In both cases, it can be observed that X-shaped streaks appear in the phantom, with the angle dependent on the missing wedge angle, lowering reconstruction quality.

In contrast to the calibration errors seen in the previous section, undersampling in X-ray CT is typically not an error but a choice. Few-view imaging allows for much faster scan times. If the reconstructions made from such a highly subsampled dataset are still usable, then acquiring fewer images is beneficial for a variety of applications, for example: fewer projections of a patient in medical imaging reduce the dose of harmful radiation [15], industrial in-line inspection only allows a few seconds per object for scanning [16] and a higher time resolution can be obtained in 4D-CT, with less scans needed [17]. The need for limited angular range tomography arises when large objects are to be scanned, such as entire cars and aeroplane wings [18]. In such cases, the source and detector are often unable to make a full rotation. Therefore, it is clear that the development of techniques that allow for accurate reconstructions from undersampled data is important to advance the field of X-ray CT imaging. Because of their practical importance, reconstructions from undersampled data will be a main focus for the research presented in this thesis, see Chapters 2, 3, and 5.

Model errors

The last type of reconstruction error to be considered are those related to deviations from the imaging model. Both FBP and SIRT assume a

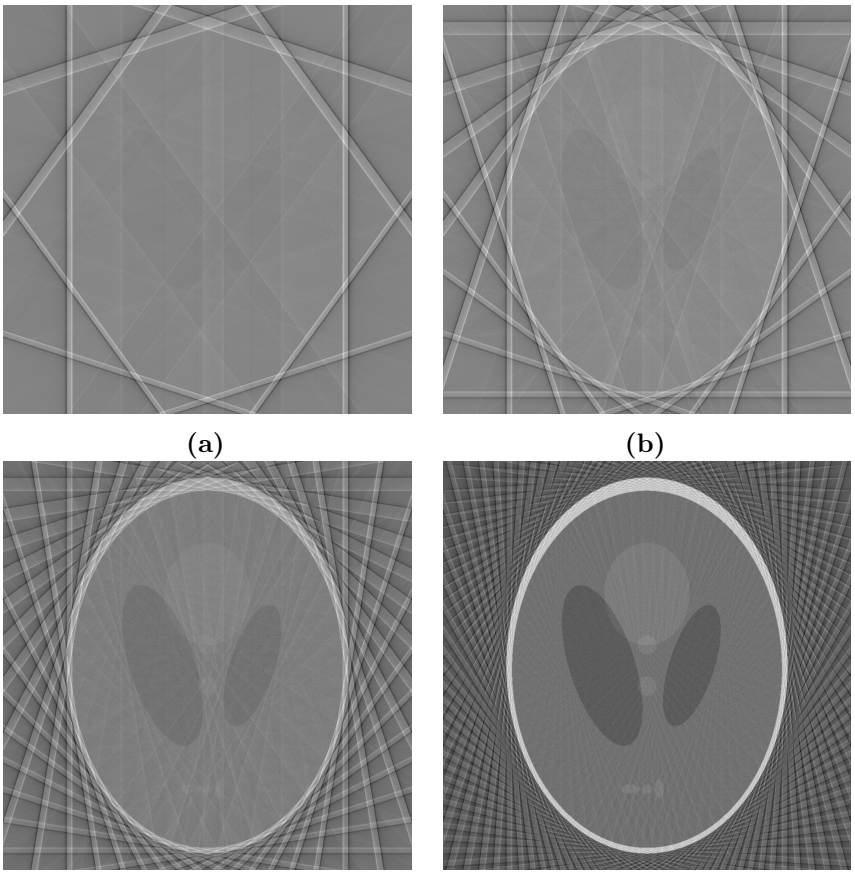
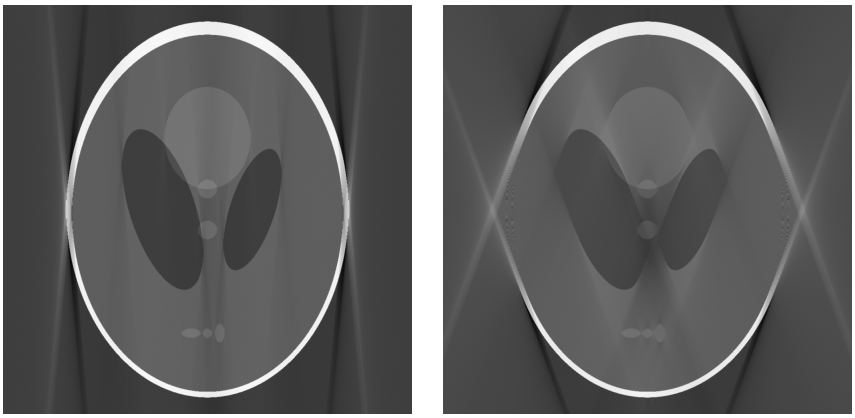


Figure 1.9. Few-view reconstructions of the Shepp-Logan phantom using FBP from (a) 5, (b) 10, (c) 20 and (d) 50 projections.

linear imaging model based on the normalised and log-corrected Beer-Lambert law (1.4) and (1.9). Reality can deviate from the assumptions of the Beer-Lambert model in a variety of ways [10, 11, 1, 19]. In this section, three common ways will be highlighted. First, movement of the object during scanning will be considered. Next, the effect of noise in the measurement will be shown. Finally, the effect of polychromatic sources is studied.

In the basic tomography models introduced in Section 1.1.4 the imaged object is assumed to be *static*, i.e. unchanging throughout the scan. Mathematically, the static nature of the object is expressed by the spatial function $\mu(\boldsymbol{s})$ being independent of time. When the object is not stationary during the acquisition, or equivalently, if the scanning setup



(a) **(b)**

Figure 1.10. Missing wedge reconstructions of the Shepp-Logan phantom using FBP from 1000 projections in (a) $[5°, 175°]$ and (b) $[20°, 160°]$.

is unstable, artefacts appear in the reconstruction. In Fig. 1.11 the FBP reconstruction of the Shepp-Logan phantom is shown, where the phantom was translated for 0.1 pixel per projection at an angle of 45°, for half of the scan in one direction, and for the second half back in the direction of the starting position. To compensate for the discrepancy between the static model and the dynamic acquisition, the large structures in the reconstructed image are deformed, while the smaller features, such as the three ellipses on the bottom, are smeared out and become unrecognisable. Movement of the object during scanning is a common problem in medical CT, where involuntary patient movement hinders clear imaging [10, 20]. Recently, 4D CT, that is, time-resolved tomography, has become an imaging modality with increasing research interest [21, 22, 23, 24, 25]. 4D CT is made possible by more accurate modelling of the dynamic imaging process.

The entire X-ray imaging chain is a stochastic process. It can be shown that the resulting process from X-ray generation to measurement by the detector is a cascaded Poisson process [1]. Therefore, noise in the acquired images is unavoidable in real measurements. However, the Beer-Lambert law does not take statistical behaviour of the X-ray beam into account, modelling the acquisition instead as a deterministic process. This discrepancy will again lead to reconstruction errors since, according to the deterministic model, any fluctuation in the projection data must be due to differences in the attenuation coefficients of the object. Reconstructions using FBP from projection data with different levels of

Figure 1.11. FBP reconstruction of the Shepp-Logan phantom with simulated movement during scanning.

simulated Poisson noise are shown in Fig. 1.12. The reconstructions from noisy data appear to be noisy themselves. The more photon counts are expected per detector pixel, the lower the effect of the noise becomes. In the case with the strongest noise, Fig. 1.12a, details of the reconstructions are fully obscured, while in the case with the most photons, Fig. 1.12d, the fluctuations in the reconstructed values are minimal. The number of detected photons can be increased by either increasing the exposure time or by increasing the source power. In this thesis, no special attention is paid to the statistical nature of the images, although noise is of course present in all experimental datasets used. A large body of literature is available on the nature of noise in CT images and algorithms attempting to diminish its adverse effects on reconstructions [26, 27, 17, 28].

The spectrum emitted by an X-ray tube is polychromatic, i.e. the tube emits photons at different energies. In reality, attenuation depends on the energy of the photons as discussed in Section 1.1.2. The energy dependence of the attenuation coefficient for different materials is shown in Fig. 1.13. From Fig. 1.13 one sees that the relationship between attenuation and energy is non-linear and different for different materials. Furthermore, the lower-energy photons are attenuated at a higher rate than the higher-energy photons. As a result, the average energy of the X-ray beam shifts upward as it travels through an attenuating object, as the low-energy photons get attenuated at a higher rate. In reconstructions utilising a monochromatic model, the energy shift leads to cupping artefacts: the attenuation is overestimated at the edge of the object and underestimated on the inside. An FBP reconstruction from simulated polychromatic data, as well as a line profile through the recon-

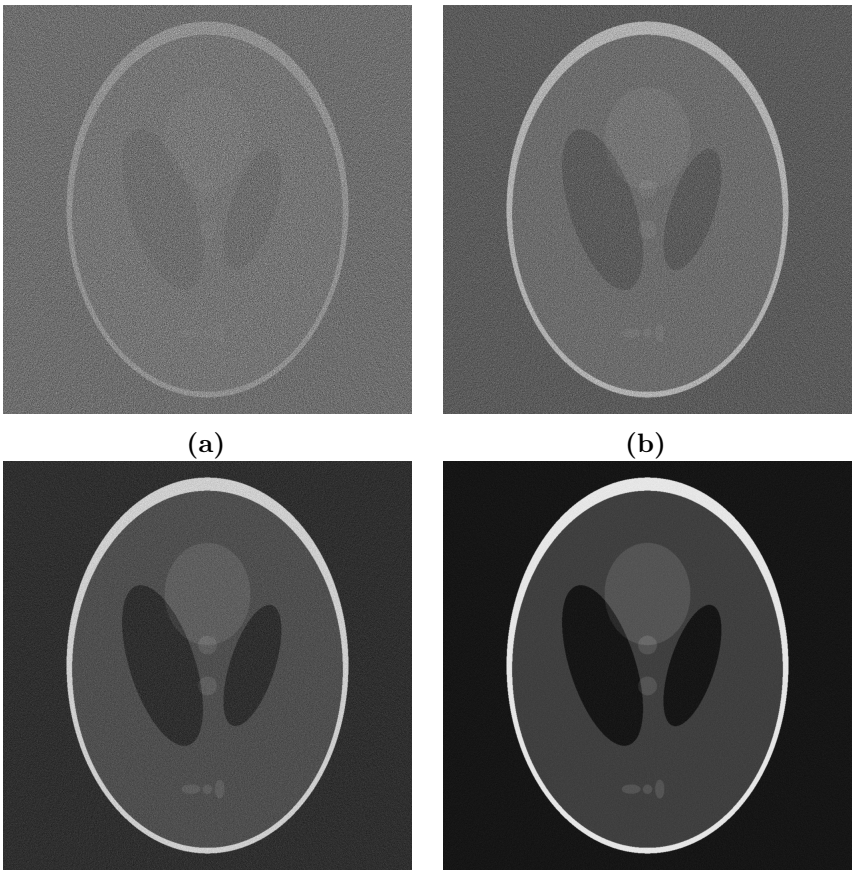


Figure 1.12. FBP reconstructions of the Shepp-Logan phantom with simulated noise corresponding to (a) 10^2, (b) 10^3 and (c) 10^4 (d) 10^5 photons per detector pixel in the flatfield.

struction of the middle column of the image, which shows clear cupping, are shown in Fig. 1.14. A side effect of cupping is that regions that are in reality homogeneous no longer are reconstructed with a single grey value, complicating segmentation procedures common in the analysis of X-ray images. Research on novel reconstruction techniques that account for polychromaticity is presented in Chapters 2, 3, and 4.

1.2.2 Advanced reconstruction techniques

In this section an overview of strategies to deal with the problems outlined in the previous section is given. It will mainly be dealing with the two sources of artefacts that are investigated in this thesis, polychromatic

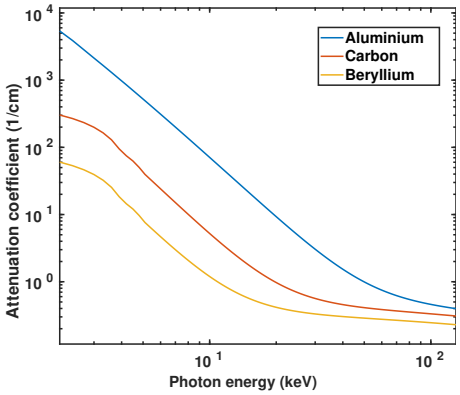
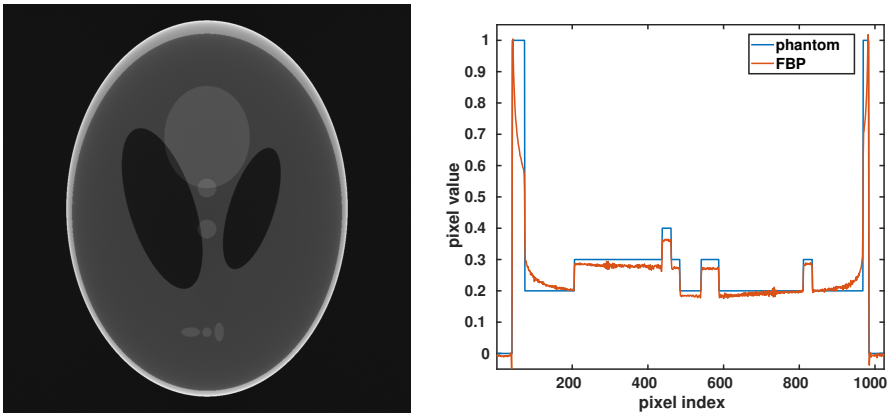


Figure 1.13. Attenuation coefficients of different materials as a function of photon energy.



(a) **(b)**

Figure 1.14. (a) FBP reconstructions of the Shepp-Logan phantom with simulated polychromatic spectrum. (b) Lineplot of the middle column of the reconstruction.

effects and undersampling.

Reconstructions from polychromatic data

There are three strategies to deal with model errors [29, 30]: hardware, image processing, and model-based approaches.

First, there are hardware solutions which attempt to change the real acquisition in such a way that it is better approximated by the chosen reconstruction model. An example of hardware solutions that reduce beam hardening is the use of a *hardware filter*, which is a small metal plate

placed directly at the source [31]. The filter removes a larger fraction of low-energy photons, *pre-hardening* the beam. Reconstructions from data acquired with a pre-hardened spectrum will show less cupping artefacts, however, the filter absorbs a number of the total photons available for imaging, so the resulting images will either be noisier or longer scan times are required.

The second option is an image preprocessing approach [32, 33]. Here, the acquired images undergo an algorithmic image processing step that attempts to correct the acquired images so that the resulting preprocessed data set fits the chosen reconstruction model. For the case of beam hardening, a preprocessing step can be seen as a function f such that:

$$f(\boldsymbol{p}_{\text{poly}}(\mu(\boldsymbol{s}))) \approx \boldsymbol{p}_{\text{mono}}(\mu(\boldsymbol{s})), \tag{1.15}$$

with $\boldsymbol{p}_{\text{poly}}(\mu(\boldsymbol{s})$ the polychromatic forward projection and $\boldsymbol{p}_{\text{mono}}(\mu(\boldsymbol{s})$ the monochromatic forward projection of an object $\mu(\boldsymbol{s})$. In other words, the function f takes the polychromatic projections as input and attempts to approximate a set of monochromatic projections from the same object. The polychromatic-to-monochromatic conversion problem is in general heavily underdetermined, therefore, some prior knowledge on the sample must be assumed. A simple version of such a preprocessing algorithm can be found by assuming that the spectrum of the source is known and that the X-ray beam only traverses two materials from source to detector: air and, for example, water. With the known spectrum and the material-dependent attenuation curve, such as the ones shown in 1.13, it is trivial to create an algorithm that transforms polychromatic measurements to equivalent line lengths through the chosen material. In practice, however, samples of only a single non-air material are rarely of interest, and the described preprocessing approach no longer provides acceptable approximations once the materials present in the sample vary too much in energy-dependent attenuation values [32]. Improvements, for example, using fitted quadratic or cubic polynomials for f, have been presented in [32]. Preprocessing techniques have the practical benefits of being generally fast and not requiring a change in the imaging setup. However, when the object to be imaged does not fulfil the assumptions of the method, the artefact can be reduced insufficiently.

The third option is to define a new model of the acquisition process and create a reconstruction algorithm based on the newly defined model. The benefit of such a *model-based reconstruction technique* is that it allows the most accurate modelling of the acquisition, therefore having the potential to generate the most faithful reconstructions. The downside is that, in general, model-based techniques require more computation time

to reconstruct an image. In the next paragraphs, an in-depth look at the construction of a model-based reconstruction algorithm for polychromatic CT will be given: the polychromatic SIRT or pSIRT technique described in [34]. First, the polychromatic algebraic forward model is introduced and then the pSIRT update step is described.

As a model for the acquisition process, the polychromatic Beer-Lambert law (1.3), described in Section 1.1.2, is used. After discretising the energy spectrum into n_e energy bins and discretising the object space, one obtains the polychromatic analog to (1.9):

$$\boldsymbol{p}(\boldsymbol{x}) = \sum_{e=1}^{n_\mathrm{e}} I_0(e) \exp(-\boldsymbol{A}\boldsymbol{x}(e)) \,, \tag{1.16}$$

where each $\boldsymbol{x}(e) \in \mathbb{R}^{n_\mathrm{vol}}$ is an object with linear attenuation coefficients at energy e, with $e \in \{1, \ldots, n_e\}$. The model in (1.16) requires a prior estimate of the energy-dependent flatfield $I_0(e)$. In (1.16) the number of unknowns is n_e times higher than in the monochromatic case, however, there is of course a strong correlation between the different values that $\boldsymbol{x}_i$ takes at the different energy levels, as it represents the same point in space and thus the same material. To reduce the number of unknowns to the same number as in the monochromatic case, basis materials are used. Each material present in the object is assumed to be one of the basis materials or is well approximated by a combination of them. The attenuation coefficients of the basis materials can be looked up for any energy bin, for example at [35]. By selecting n_m basis materials, as well as the attenuation values of the materials in all energy bins, $\mu_1(e), \ldots, \mu_{n_\mathrm{m}}(e)$, (1.16) can be rewritten as:

$$\boldsymbol{p}(\boldsymbol{x}) = \sum_{e=1}^{n_\mathrm{e}} I_0(e) \exp\left(- \sum_{m=1}^{n_\mathrm{m}} \boldsymbol{A}\boldsymbol{M}_m(\boldsymbol{x})\mu_m(e) \right), \tag{1.17}$$

with $\boldsymbol{M}_m$ triangular pointwise functions operating on the image $\boldsymbol{x}$ as follows:

$$\begin{cases} 0 & \text{if } x \notin [\mu_{m-1}(e_r), \mu_{m+1}(e_r)] \\ \frac{x-\mu_{m-1}(e_r)}{\mu_m(e_r)-\mu_{m-1}(e_r)} & \text{if } x \in [\mu_{m-1}(e_r), \mu_m(e_r)] \\ \frac{x-\mu_{m+1}(e_r)}{\mu_m(e_r)-\mu_{m+1}(e_r)} & \text{if } x \in [\mu_m(e_r), \mu_{m+1}(e_r)] \end{cases}, \tag{1.18}$$

with e_r an a priori selected reference energy. The $\boldsymbol{M}_m$ functions define a linear interpolation model assuming that every voxel is a mixture of two basis materials with neighbouring attenuation coefficients. When more basis materials are considered, the number of forward projections

increases. However, every non-zero value in $\boldsymbol{x}$ only gives rise to a non-zero value in at most two of the $\boldsymbol{M}_m(\boldsymbol{x})$, so the volumes also get more sparse as more basis materials are selected. From (1.17), one can define a normalised polychromatic projection operator polyProj(·):

$$\text{polyProj}(\boldsymbol{x}) = -\log\left(\frac{\boldsymbol{p}(\boldsymbol{x})}{\sum_{e=1}^{n_\text{e}} I_0(e)}\right). \tag{1.19}$$

Polychromatic SIRT (pSIRT) is a heuristic algorithm introduced in [34], which changes the forward model in (1.12), while keeping the backprojection of the error the same:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \boldsymbol{C}\boldsymbol{A}^\top\boldsymbol{R}\left(\text{polyProj}(\boldsymbol{x}^{(k)}) - \boldsymbol{p}\right). \tag{1.20}$$

The pSIRT technique was shown to be able to accurately remove beam hardening artefacts in clinical CT [34]. In this thesis, it will be used in Chapter 2. However, some questions about the method also naturally arise, as the backprojection of the errors was left unchanged from SIRT. Therefore in the backprojection step, the polychromaticity is no longer modelled. The method has also been shown to not converge in certain cases [36]. An investigation into improved reconstruction algorithms from the forward model in 1.17 is presented in Chapter 4.

Reconstructions from undersampled data

The simplest solution to the problem of undersampling would be to interpolate the given data so that a "full" dataset is still acquired. An example of the interpolation method is shown in Fig. 1.15, where reconstructions are made after interpolating the same datasets used for Fig. 1.9 to 1000 projections each. The reconstructions do not exhibit the typical streaking artefacts of undersampling as strongly as in Fig. 1.9, as enough projections are available. However, the reconstructions exhibit rotational blur, with the strength of the blurring dependent on the angular step between the measured projections. Smaller details in the reconstruction, such as the ellipses at the bottom of the phantom, are smeared out by the rotational blur. It is clear from these reconstructions that this approach is generally not advisable.

The interpolation solution does not address the fundamental issue with reconstructions from undersampled data, namely that the space of possible solutions is too large. Therefore, to deal with undersampled data, some sort of prior knowledge of the object is usually required. The prior knowledge will either be used to form a regularisation term, guiding the reconstruction to a solution exhibiting desired traits, or will act as

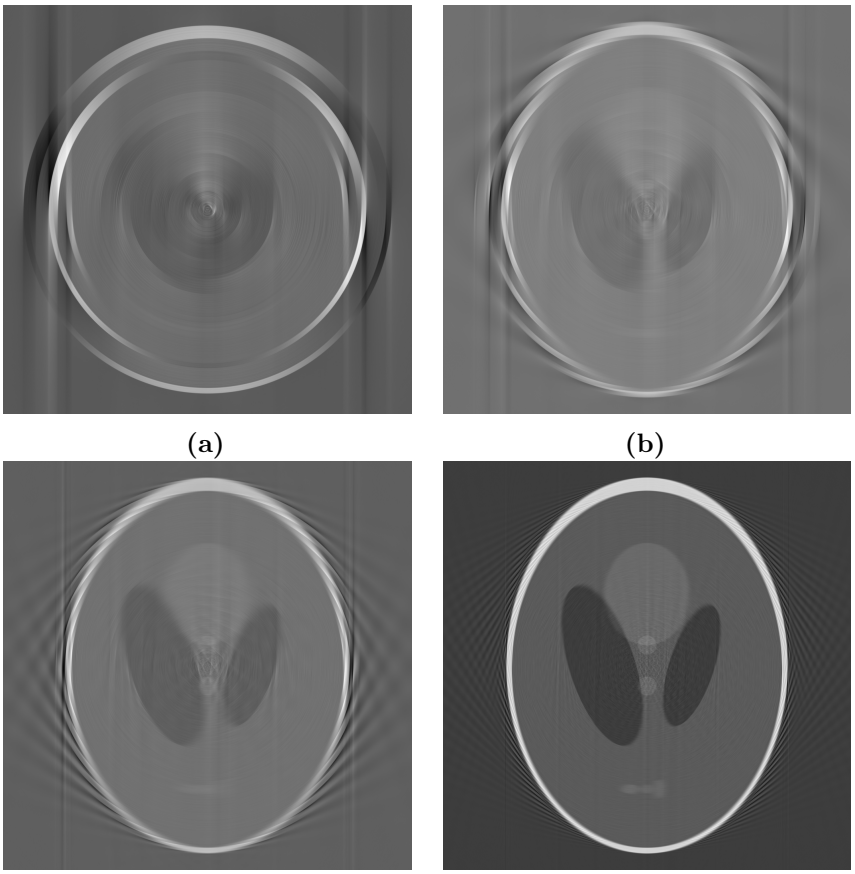


Figure 1.15. Few-view reconstructions of the Shepp-Logan phantom using FBP where the projection data is interpolated to create 1000 projections from (a) 5, (b) 10, (c) 20 and (d) 50 projections.

a constraint, reducing the space to be searched for solutions. Different types of prior information can be exploited. In [37] reconstructions from few-view projections in an industrial setting are improved by using the CAD model of the scanned piece as prior information. Many methods make use of some kind of *sparsity* of the image or a transformation of the image, that is, in some sense, the domain is restricted. If the (transformed) image is sparse in the image domain, i.e. the number of non-zero voxels is small, compressed sensing [38] techniques can be used to retrieve more accurate reconstructions. A common sparsifying transformation for X-ray CT images is the gradient, as the object itself is not sparse in the image domain, but the boundaries between different homogeneous regions are

sparse compared to the rest of the image [39, 40, 41, 42]. Another type of prior knowledge is knowledge on the values that the reconstruction can take. In industrial CT the composition of the object is known in advance, making a prior based on attenuation values a natural choice [43]. Instead of reconstructing an image $\boldsymbol{x} \in \mathbb{R}^{n_{\text{vol}}}$, the reconstruction problem is reduced to searching for the optimal $\boldsymbol{x} \in \{\rho_1, \rho_2, \ldots, \rho_{n_m}\}^{n_{\text{vol}}}$, with n_m the number of materials and ρ_i the known attenuation values of the materials. It is clear that, in theory, as the reconstruction domain is significantly reduced in size, it should be possible to reconstruct the image from a smaller number of projections. Reconstructing images with only a limited number of possible grey values is studied in Discrete Tomography (DT). For a more in-depth exploration of DT, the reader is referred to [44, 45, 46]. In this thesis, we will work with and improve on a discrete tomography technique called DART [47]. DART is a practical CT algorithm shown to work well in a variety of use-cases [48, 49, 50, 51].

DART is built upon an Algebraic Reconstruction Method (ARM), which calculates solutions to the reconstruction problem (1.9). A relaxed SIRT is commonly used as the ARM. Relaxed SIRT iteratively computes the following update step:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \lambda \boldsymbol{C}\boldsymbol{A}^\top \boldsymbol{R}\left(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{p}\right). \tag{1.21}$$

where $\lambda \in \mathbb{R}^+$ is called the relaxation parameter.

Let $\mathcal{R} = \{\rho_1 < ... < \rho_{n_m}\}$ be the set of grey values representing the different materials of which the object is composed. Then, a solution $\boldsymbol{x}$ to (1.9) is discrete if $\boldsymbol{x} \in \{\rho_1, ..., \rho_{n_m}\}^{n_{\text{vol}}}$. Given an initial SIRT reconstruction $\boldsymbol{x}^{(0)}$, the key iterative steps in the DART algorithm can be summarized as follows:

1. **Segmentation**: Let $\boldsymbol{x}^{(\ell)}$ be the output of the previous DART iterations, with $\boldsymbol{x}^{(0)}$ the output from the initial SIRT iterations. Since the grey levels in the desired reconstruction are known to be in $\mathcal{R}$, the elements of $\boldsymbol{x}^{(\ell)}$ are projected (e.g., by thresholding) onto $\mathcal{R}$. The segmented image is denoted by $\boldsymbol{s}^{(\ell)}$.

2. **Partitioning**: In the partitioning step, a divide-and-conquer procedure is initiated by labelling the image voxels into two categories: free voxels (which will be updated) and fixed voxels (which are kept fixed at their current value). If a voxel has at least one neighbouring voxel with a different grey value, it is considered a boundary voxel and is added to the set of free voxels. Otherwise, the voxel is initially considered fixed. Additionally, every non-boundary voxel

has a small but constant probability p to be included in the free set. After labelling, the reconstruction process continues on the free voxels only, while the values of the fixed voxels are not updated.

3. **Masked reconstruction**: A fixed number of SIRT iterations is then performed on the free voxels and a new image $\boldsymbol{x}^{(\ell+1)}$ is calculated by merging the updated free voxels with the fixed voxels.

4. **Smooth and repeat** After an optional smoothing is performed by convolution with a 3×3 kernel, the steps are repeated until a convergence criterion is met or a maximum number of DART iterations has been reached. In this thesis, a 3×3 median kernel M was used with weight parameter b:

$$\boldsymbol{x} = (1-b)\boldsymbol{x}^{(\ell+1)} + b(M * \boldsymbol{x}^{(\ell+1)}) \quad ,$$

where $*$ denotes the convolution operator.

Thanks to the masking step, the DART algorithm can deal with the underdetermination of the reconstruction problem. When a voxel x_i is fixed at value ρ_j, the linear system (1.9) is updated as follows:

$$\begin{bmatrix} \boldsymbol{A}_1 & \cdots & \boldsymbol{A}_{i-1} & \boldsymbol{A}_{i+1} & \cdots & \boldsymbol{A}_{n_{\text{vol}}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_{i+1} \\ \vdots \\ x_{n_{\text{vol}}} \end{bmatrix} = \boldsymbol{p} - \rho_j \boldsymbol{A}_i \qquad (1.22)$$

In the updated system the number of unknowns is reduced by one, while the number of equations remains the same. Assuming that the segmented value of x_i is correct, the DART reconstruction will provide a better estimate of the original object by solving the smaller linear system with the fixed voxels removed, compared to solving the underdetermined system directly with SIRT. The heuristic behind the partitioning into boundary and non-boundary pixels is that algebraic reconstructions tend to reconstruct large homogeneous regions well quickly, while the detailed boundary regions take significantly more iterations. The DART algorithm has a number of parameters that will influence performance and are chosen heuristically. The most important ones are the percentage of random voxels chosen to be free in the masking step and the different numbers of iterations (initial, inner and outer).

In Chapters 2 and 3, improvements to DART are presented. DART relies on accurate prior knowledge of expected grey values for reconstruction. However, as seen in Section 1.2.1, beam hardening makes homogeneous object regions non-homogeneous in the reconstruction. Polychromatic sources are unavoidable in industrial settings where the prior knowledge required for DART is readily available. An approach to deal with polychromatic data in the framework of DART is proposed and this work is presented in Chapter 2. Furthermore, the chosen percentage of random voxels is an important parameter in the DART framework. Increasing the number of fixed voxels reduces the computation per iteration. However, a lower fix percentage can make the algorithm more robust to noise and allow segmentation errors to be corrected more quickly [47]. For these reasons, the automatic optimal selection of the fix percentage would make the DART framework more practically usable. It is also clear that there is no need for the fix percentage to be static throughout the different iterations of the DART algorithm. Therefore, a novel framework to dynamically update the parameter is presented in Chapter 3. Furthermore, it is shown that the dynamic framework improves the newly proposed polychromatic DART algorithm from Chapter 2.

1.3 Non-linear optimisation for CT

In this section we introduce the basic concepts of numerical optimisation algorithms that we will use for CT. As mentioned in Section 1.2, it is not clear that either SIRT or pSIRT are optimal iterative algorithms to solve their respective CT reconstruction problems. The main drawback of iterative algebraic solvers for CT, when compared to analytical methods such as FBP, is their long computation time. Therefore, it is crucial to use efficient methods to solve optimisation problems related to CT. First, we will introduce gradient descent, a fundamental way to solve any optimisation problem. Next, a class of non-linear optimisation algorithms called quasi-Newton methods are introduced. Finally, we quickly look at the shortcomings of SIRT. For an in-depth look into numerical optimisation, we refer to [4].

1.3.1 Gradient descent

A vector-valued optimisation problem is defined as:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}), \tag{1.23}$$

with $\boldsymbol{x} \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$, the objective function. It will be assumed that f is a continuously differentiable function. The goal of an optimisation algorithm is to provide iterates $\boldsymbol{x}^{(k)}$ that converge to a (at least local) minimum of the objective function. The iterates can be written in the following form:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}, \tag{1.24}$$

with $\boldsymbol{d}^{(k)}$ the direction of the step and $\alpha^{(k)}$ the step size. In any point $\boldsymbol{x}$ the gradient of f in $\boldsymbol{x}$, $\boldsymbol{\nabla} f(\boldsymbol{x})$, is the vector pointing in the direction of steepest ascent. Therefore, as a minimisation strategy, it is natural to choose the negative gradient as the step direction, which defines the iterations of *gradient descent*:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha^{(k)}\nabla f(\boldsymbol{x}^{(k)}). \tag{1.25}$$

Gradient descent defines the direction of each step, but the step size still needs to be selected. A good choice of step size is important, as sufficient progress towards the minimum has to be made in each iteration for the algorithm to be efficient. A second optimisation problem is defined for the optimal step size in iteration $k+1$, $\alpha_*^{(k+1)}$:

$$\alpha_*^{(k+1)} = \underset{\alpha}{\operatorname{argmin}}\, f\left(\boldsymbol{x}^{(k)} + \alpha\nabla f(\boldsymbol{x}^{(k)}\right). \tag{1.26}$$

Finding $\alpha_*^{(k+1)}$ would be too expensive to solve in every iteration, thus there is a trade-off between the efficiency of the step and the time spent determining $\alpha^{(k+1)}$. The simplest option is to keep the step size constant throughout the iterations. In this case, the convergence speed of the gradient descent algorithm will be highly dependent on the choice of α. However, no computation time is needed to calculate the step size. Another option is to perform an inexact backtracking line search, that is, iteratively checking if a step size α^ℓ fulfils certain conditions, such as Wolfe, Goldstein, or Armijo [4]. If α^ℓ does not meet the conditions, $\alpha^{\ell+1} = \lambda\alpha\ell$ is tried, with $\lambda \in]0, 1[$, continuing until an acceptable step size is found. Gradient descent with line searches is intuitively clear, however, it does not perform well. Convergence to the minimum is heavily affected by ill-conditioning [52], even in the case where the optimal step size α_* is found. A special step size for gradient descent will be considered that makes the algorithm behave entirely differently from the line search version. Barzilai-Borwein steps [52], are step sizes for gradient descent, which, in iteration k, are defined as:

$$\alpha^{(k)} = \frac{\left(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right) \cdot \left(\nabla f(\boldsymbol{x}^{(k)}) - \nabla f(\boldsymbol{x}^{(k-1)})\right)}{\left(\nabla f(\boldsymbol{x}^{(k)}) - \nabla f(\boldsymbol{x}^{(k-1)})\right) \cdot \left(\nabla f(\boldsymbol{x}^{(k)}) - \nabla f(\boldsymbol{x}^{(k-1)})\right)}, \tag{1.27}$$

with an initial step size for the first iteration being chosen as either a fixed step size or found through a backtracking line search. The Barzilai-Borwein formula requires storing both the current and the previous value for both $\boldsymbol{x}$ and $\nabla f(\boldsymbol{x})$. Using Barzilai-Borwein steps leads to improved convergence speed and robustness against ill-conditioning [52]. However, convergence to a minimum is no longer guaranteed to be monotone and convergence tends to exhibit some erratic behaviour [53]. The Barzilai-Borwein formula will be revisited in Section 1.3.2.

1.3.2 Quasi-Newton methods

To overcome some of the limitations of gradient descent and similar first-order methods, quasi-Newton methods have been developed. Quasi-Newton methods approximate the Hessian matrix (the second derivative of the objective function), which can lead to significantly faster convergence rates.

Newton's method is an iterative root-finding algorithm. That is, given a differentiable function F, Newton's method iteratively approaches an x such that $F(x) = 0$. Newton's method was originally developed for 1-dimensional real-valued functions, however, it is easily extended to higher-dimensional functions. Given $F : \mathbb{R}^n \to \mathbb{R}^n$, differentiable, and its Jacobian matrix $\boldsymbol{J}_F$, the k^{th} update for Newton's method is given by [4]:

$$\boldsymbol{\Delta x}^{(k)} = -\boldsymbol{J}_F(\boldsymbol{x}^{(k)})^{-1} F(\boldsymbol{x}^{(k)})\,, \tag{1.28}$$

such that $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{\Delta x}^{(k)}$.

Newton's method can be used for optimisation problems, since every minimum of the objective function has to be a root of the gradient of the objective function. So, given the objective function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\boldsymbol{\Delta x}^{(k)} = -\boldsymbol{H}_f(\boldsymbol{x}^{(k)})^{-1} \nabla f(\boldsymbol{x}^{(k)})\,, \tag{1.29}$$

with $\boldsymbol{H}$ the Hessian matrix, gives the update step towards a critical point of f. A critical point is guaranteed to be a minimum if f is convex. For optimisation problems, the same update step can also be derived from the second-order Taylor expansion of the objective function f:

$$f(\boldsymbol{x}^{(k)} + \boldsymbol{d}) \approx f(\boldsymbol{x}^{(k)}) + \boldsymbol{d}^\top \nabla f(\boldsymbol{x}^{(k)}) + \frac{1}{2}\boldsymbol{d}^\top \boldsymbol{H}_f(\boldsymbol{x}^{(k)})\boldsymbol{d}. \tag{1.30}$$

The ideal update $\boldsymbol{d}$ should minimise the right-hand side. The minimum is found as a root of the derivative of the right-hand side with respect to $\boldsymbol{d}$, resulting again in the formula in (1.29).

The update step in (1.29) can also be seen as a search direction for (1.24). The direction given in (1.29) is called the Newton direction and can again be used in a line search method. Methods using the Newton direction tend to have faster local convergence than steepest descent methods [4].

For CT, Newton's method can be used to minimize least-squares problems, with objective function $G : \mathbb{R}^n \to \mathbb{R}$:

$$G : \boldsymbol{x} \mapsto \frac{1}{2} \|B(\boldsymbol{x}) - \boldsymbol{b}\|_2^2 \tag{1.31}$$

with $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$ and $B : \mathbb{R}^n \to \mathbb{R}^m$ some imaging model that maps $\boldsymbol{x}$ to the space of the vector $\boldsymbol{b}$.

Inversion of the Hessian $\boldsymbol{H}_G$ is infeasible for large systems of equations, as is the case for X-ray CT data. Quasi-Newton methods attempt to solve (1.29) through some, usually easier to compute, approximation of the (inverse) Hessian. Two different quasi-Newton techniques will now be considerd, Gauss-Newton-Krylov (GNK) and Broyden-Fletcher-Goldfarb-Shanno (BFGS). Furthermore, gradient descent with Barzilai-Borwein will aslo be briefly described in the context of quasi-Newton methods.

Gauss-Newton-Krylov

Equation (1.29) can be rewritten as the following equivalent linear system of equations:

$$\boldsymbol{H}_G(\boldsymbol{x}^{(k)})\boldsymbol{\Delta x}^{(k)} = -\nabla G(\boldsymbol{x}^{(k)}) \,, \tag{1.32}$$

which can be solved with an iterative method to obtain an approximate solution for $\boldsymbol{\Delta x}^{(k)}$. Krylov subspace methods, explained below, can be used to solve the linear system, resulting in the so-called Newton-Krylov method [54]. The Newton-Krylov approach solves the problem of having to calculate the inverse directly.

Given a set of linear equations,

$$\boldsymbol{W}\boldsymbol{x} = \boldsymbol{s}, \tag{1.33}$$

Krylov-subspace methods will iteratively minimize the residual over the r-order Krylov subspace [54]

$$\mathcal{K}_r(\boldsymbol{W}, \boldsymbol{s}) = \text{span}\{\boldsymbol{s}, \boldsymbol{W}\boldsymbol{s}, \boldsymbol{W}^2\boldsymbol{s}, \ldots, \boldsymbol{W}^{r-1}\boldsymbol{s}\} \tag{1.34}$$

in the r^{th} step. A well-known and widespread Krylov algorithm is the Conjugate Gradient method (CG)[55], which solves systems for symmetric and positive definite matrices $\boldsymbol{W}$. In this thesis, the minimal residual

method (MINRES)[56] was always used, as implemented in MATLAB, which still requires $\boldsymbol{W}$ to be symmetric, but no longer positive definite.

For least-squares problems of the form (1.31), the Hessian matrix can be approximated as [4]:

$$\boldsymbol{H}_G(\boldsymbol{x}^k) = \boldsymbol{J}_B^\top(\boldsymbol{x}^{(k)})\boldsymbol{J}_B(\boldsymbol{x}^{(k)})\,, \tag{1.35}$$

with $\boldsymbol{J}_B$ the Jacobian of the model B in our objective function. The approximation in (1.35) leads to the following system of equations:

$$\begin{aligned} &\boldsymbol{J}_B^\top\left(\boldsymbol{x}^{(k)}\right)\boldsymbol{J}_B\left(\boldsymbol{x}^{(k)}\right)\left(\boldsymbol{\Delta x}^{(k)}\right) \quad = \\ &-\boldsymbol{J}_B^\top\left(\boldsymbol{x}^{(k)}\right)\left(B(\boldsymbol{x}^{(k)}) - \boldsymbol{b}\right), \end{aligned} \tag{1.36}$$

to which an approximate solution can be constructed using MINRES.

Once again, to guarantee a sufficient decrease in the sum of squares of the residuals in each step, a backtracking line search with Armijo conditions [4] can be performed along the descent direction $\boldsymbol{\Delta x}^{(k)}$ found from (1.36). The resulting method is the Gauss-Newton-Krylov method. For the practical implementation of the method, it is important to note that there is no need to explicitly calculate and store the Jacobian or approximated Hessian matrices. In X-ray CT applications, these matrices are typically too large to store in memory. It suffices to program methods F_J and F_H, which, given a vector $\boldsymbol{y}$, directly calculate the matrix-vector products $F_J(\boldsymbol{y}) = \boldsymbol{J}\boldsymbol{y}$ and $F_H(\boldsymbol{y}) = \boldsymbol{H}\boldsymbol{y}$.

BFGS

A second quasi-Newton method to consider is BFGS. Starting from the Taylor expansion (1.30), the exact Hessian is replaced by an approximation $\boldsymbol{B}_{(k)}$:

$$f(\boldsymbol{x}^{(k)} + \boldsymbol{d}) \approx f(\boldsymbol{x}^{(k)}) + \boldsymbol{d}^\top \nabla f(\boldsymbol{x}^{(k)}) + \frac{1}{2}\boldsymbol{d}^\top \boldsymbol{B}_{(k)}\boldsymbol{d}. \tag{1.37}$$

BFGS is an algorithm that will update the Hessian approximation $\boldsymbol{B}_{(k)}$ in every iteration. The approximate Newton direction is:

$$\boldsymbol{d}^{(k)} = -\boldsymbol{B}_{(k)}^{-1}\nabla f(\boldsymbol{x}^{(k)}). \tag{1.38}$$

To make $\boldsymbol{B}_{(k)}$ a sensible approximation, it will be required to have certain properties that make it similar to the exact Hessian. Since the Hessian is always symmetric, the $\boldsymbol{B}_{(k)}$ are required to be symmetric. For the second condition, the approximation (1.37) is required to improve from one iteration to the next. Assume that $\boldsymbol{B}_k$ is already known, and have

thus made a new iterate $\boldsymbol{x}^{(k+1)} = \alpha^{(k)}\boldsymbol{d}^{(k)}$. An update to $\boldsymbol{B}_{(k)}$ needs to be defined, to construct $\boldsymbol{B}_{(k+1)}$. If the right-hand side of (1.37) is defined as $q^{(k)}(\boldsymbol{d})$, then it is trivially seen that in $\boldsymbol{d} = 0$, $q^{(k)}$ has the same value and gradient as f in the point $x^{(k)}$. Similarly, in the next iteration $q^{(k+1)}$ has the same value and gradient as f in the point $x^{(k+1)}$, regardless of what $\boldsymbol{B}_{(k+1)}$ is. The approximation $\boldsymbol{B}_{(k+1)}$ can now be required to be chosen in such a way that $q_{(k+1)}$ also matches the gradient of f in the point $\boldsymbol{x}^{(k)}$:

$$\nabla q_{(k+1)}(-\alpha^{(k)}\boldsymbol{d}^{(k)}) = \nabla f(\boldsymbol{x}^{(k+1)}) - \alpha^{(k)}\boldsymbol{B}_{(k+1)}\boldsymbol{d}^{(k)} \tag{1.39}$$

$$\nabla f(\boldsymbol{x}^{(k)}) = \nabla f(\boldsymbol{x}^{(k+1)}) - \alpha^{(k)}\boldsymbol{B}_{(k+1)}\boldsymbol{d}^{(k)}$$

$$\boldsymbol{B}_{(k+1)}(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = \nabla f(\boldsymbol{x}^{(k+1)}) - \nabla f(\boldsymbol{x}^{(k)}). \tag{1.40}$$

The expression in (1.40) is called the *secant equation*. The BFGS algorithm will require the updates to be symmetric, follow the secant equation, and be positive definite. However, updates will not be made on $\boldsymbol{B}_{(k)}$, but directly on the inverse matrices $\boldsymbol{B}^{-1}_{(k)}$. Furthermore, $\boldsymbol{B}^{-1}_{(k+1)}$ is chosen so that $\left\|\boldsymbol{B}^{-1}_{(k+1)} - \boldsymbol{B}^{-1}_{(k)}\right\|$ is small [4]. The update step for BFGS is [4]:

$$\boldsymbol{B}^{-1}_{(k+1)} = \left(\boldsymbol{I} - \frac{\boldsymbol{y}_{(k)}\boldsymbol{g}^\top_{(k)}}{\boldsymbol{g}^\top_{(k)}\boldsymbol{y}_{(k)}}\right)\boldsymbol{B}^{-1}_{(k)}\left(\boldsymbol{I} - \frac{\boldsymbol{y}_{(k)}\boldsymbol{g}^\top_{(k)}}{\boldsymbol{g}^\top_{(k)}\boldsymbol{y}_{(k)}}\right) + \frac{\boldsymbol{y}_{(k)}\boldsymbol{y}^\top_{(k)}}{\boldsymbol{g}^\top_{(k)}\boldsymbol{y}_{(k)}} \tag{1.41}$$

with $\boldsymbol{y}_{(k)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}$ and $\boldsymbol{g}_{(k)} = \nabla f(\boldsymbol{x}^{(k+1)}) - \nabla f(\boldsymbol{x}^{(k)})$.

Barzilai-Borwein

The Barzilai-Borwein step size formula of (1.27) is found by defining $\alpha^{(k)}$ as [52]:

$$\alpha^{(k+1)} = \underset{\alpha}{\operatorname{argmin}} \left\|\boldsymbol{y}_k - \alpha\boldsymbol{g}_k\right\|_2^2. \tag{1.42}$$

The Barzilai-Borwein steps provide therefore provide a scalar $\alpha^{(k+1)}$, such that the constant matrix $\frac{1}{\alpha^{(k+1)}}\boldsymbol{I}$ comes as close as possible to satisfying the secant equation. Barzilai-Borwein steps are a quasi-Newton method where in every step the optimal constant diagonal matrix approximation to the Hessian is chosen. If this matrix is defined as $\boldsymbol{B}_{(k+1)} = \frac{1}{\alpha^{(k+1)}}\boldsymbol{I}$, the gradient descent step with step size $\alpha^{(k+1)}$ can also be seen as a step with length 1 along the approximate Newton direction $-\boldsymbol{B}_{(k+1)}\nabla f(\boldsymbol{x}^{(k+1)})$. Barzilai-Borwein can therefore be seen as a quasi-Newton method, using the best constant approximation to the Hessian in each iteration, with 'best' being defined as the best at solving the secant equation.

1.3.3 SIRT as an optimisation algorithm

In Section 1.1 the SIRT algorithm was introduced. As shown in equations (1.10) and (1.14), the CT reconstruction problem is a minimisation problem of the form of (1.23) and SIRT converges to a minimum if f is chosen to be the $\boldsymbol{R}$-norm of the projection difference. It turns out that SIRT is not actually an optimal solver for the reconstruction problem. In [9], SIRT is seen as an example of Richardson iterations [57] and it is shown that the convergence of SIRT can be improved by multiplying all the update steps by a factor $\lambda \in]1, 2[$, which experimentally was shown to halve the required number of iterations. Furthermore, SIRT can also be seen as a special case of preconditioned gradient descent, with preconditioner $\boldsymbol{C}$. It suffices to calculate

$$\nabla \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{p}\|_{\boldsymbol{R}}^2 = 2\boldsymbol{A}^\top \boldsymbol{R}\left(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{p}\right), \tag{1.43}$$

to see the update step in SIRT and standard gradient descent to be the same up to the conditioner factor $\boldsymbol{C}$. SIRT uses heuristically chosen geometrical weighting in $\boldsymbol{R}$ and $\boldsymbol{C}$, however, the algorithm does not provide any proven advantages over other established minimisation techniques [58]. Therefore, algorithms derived from SIRT, such as pSIRT [34], could also be improved, at least in terms of convergence rate, by using different optimisation approaches.

Our findings on alternative optimisation of a polychromatic objective function are presented in Chapter 4. Furthermore, the methods described in this section are also applied to a different non-linear imaging model in Chapter 5

1.4 Edge illumination X-ray phase contrast imaging

In Chapter 5, edge illumination (EI) phase contrast imaging, a different X-ray imaging modality, will be studied. In this chapter, first, a brief overview of the technique is given. Next, the standard edge illumination CT reconstruction pipeline is introduced. Finally, recent developments in EI reconstructions are discussed.

1.4.1 Differential phase contrast imaging

X-ray *phase contrast* imaging is an imaging modality based on the changes in phase of the X-ray wavefront, instead of the attenuation used in conventional X-ray imaging. In the complex index of refraction of a

material:

$$n = (1 - \delta) + i\beta, \tag{1.44}$$

β is related to attenuation of the beam and δ drives the phase shifts [59]. The refractive index n is energy-dependent. For X-rays, the contrast achievable by exploiting the phase shifts can be up to 2 orders of magnitude higher than the contrast achieved by attenuation-based imaging [59]. The increased contrast available in phase contrast imaging has led to a variety of applications, from biomedical studies [60, 61, 62, 63] to non-destructive testing [64], materials science [65, 66] and security [67, 68].

Phase changes have always been present in X-ray imaging, however, the effects on the measurements in a traditional scanner setup are usually small to non-existent. A change in phase does not change the intensity of the X-rays, and detectors do not measure the phase of the wave, only the intensity. However, a change in phase results in a local change of the propagation direction of the X-rays, i.e. the X-rays refract with a certain angle dependent on the real part of the refractive index. The refraction is dependent on the first derivative of the phase shift. Given an object with spatial distribution of refractive indices $n(x, y, z) = 1 - \delta(x, y, z) + i\beta(x, y, z)$, with the z-axis the axis from source to detector, then the total attenuation μ and resulting angle of refraction along z is given by [69]:

$$\mu(x, y) = 2k \int \beta(x, y, z)\mathrm{d}z \tag{1.45}$$

$$\alpha(x, y) = \vec{\nabla}_{x,y} \int \delta(x, y, z)\mathrm{d}z, \tag{1.46}$$

with k the wave number depending only on the wavelength of the X-rays. To enable differential phase contrast imaging, an X-ray imaging setup that can measure refraction is required [59]. Since X-ray detectors measure intensity, the imaging setup needs to make refraction effects lead to intensity modulation at the detector.

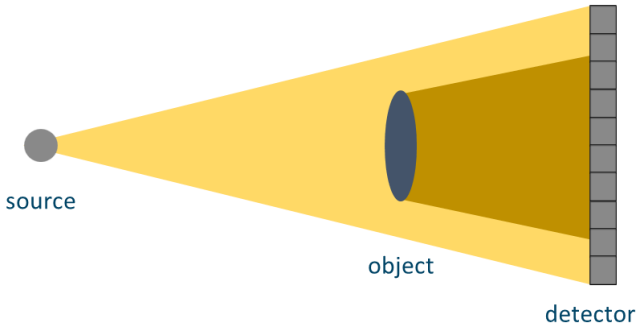
1.4.2 Edge illumination setup

Edge illumination is a differential phase contrast technique in which the X-ray beam hits the edge of the photon-sensitive part of the detector pixel. Refraction effects will then cause more or less of the beam to hit the detector pixel, resulting in the desired intensity modulation due to refraction. To create an edge illumination imaging setup in a conventional cone beam X-ray system, two gratings or *masks* are introduced [70]. The first mask, the *sample mask*, splits the cone beam into smaller *beamlets*,

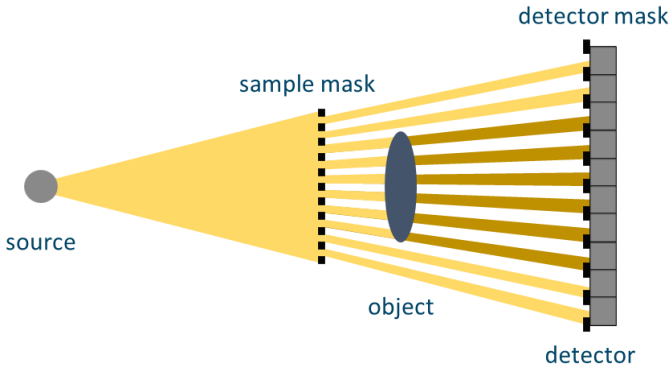
with each beamlet fully separated from the neighbouring ones. Each beamlet corresponds to a single pixel on the detector. The second mask, the *detector mask*, blocks the edges between the detector pixels, creating sensitive and insensitive regions on the detector. Each beamlet is partially blocked by the detector mask and partially hits the detector pixel. The setup is shown in Fig. 1.16. When a sample is introduced, the beamlets undergo both attenuation and refraction effects. Attenuation always leads to reduced intensity measured on the detector. Refraction effects, on the other hand, can both increase or decrease the measured intensity, depending on whether the beamlet gets refracted towards or away from the center of the detector pixel. A graphical representation of the origin of differential phase contrast in the edge illumination setup is shown in Fig. 1.17. Since the masks are constant along one axis of the detector, refraction can only be measured in the direction orthogonal to the grating bars [70]. Edge illumination is a technique that is well suited for laboratory setups, as it provides high angular sensitivity while tolerating large focal spots and polychromatic radiation [71]. In Fig. 1.18 the edge illumination setup installed at the FleXCT system at the University of Antwerp is shown [3, 72].

When an image is acquired using the setup shown in Fig. 1.16, the effects of attenuation and refractions are mixed. Therefore, to untangle the different contrast images, multiple radiographs must be acquired. By laterally translating the sample mask relative to the detector mask, the fraction of the beamlet blocked by the detector mask changes, the lateral translation is called *phase stepping*. The resulting curve obtained by measuring the intensities as a function of the mask translation is referred to as the illumination curve (IC). In Fig. 1.19 an example is shown of two illumination curves measured in the same pixel, with and without a sample present. If the masks would absorb 100% of the X-rays and the X-ray source is a perfect point source, then the IC's would have a triangular shape. In practice, however, the masks are not perfectly absorbent, and the focal spot of the X-ray source is finite. The IC is therefore assumed to approximate a Gaussian shape as a result of these imperfections. The standard edge illumination workflow is to first sample the IC by phase stepping and subsequently fitting a Gaussian curve to each of the sampled ICs.

A Gaussian curve is defined by three parameters: amplitude, mean, and variance. Comparing the change of these parameters before and after introduction of the sample gives rise to different contrasts. This comparison step is also called the *phase retrieval*. The relative change in the area under the curve gives the conventional attenuation contrast. At every phase step, the same percentage of photons is expected to be



(a)



(b)

Figure 1.16. Illustration of (a) a conventional cone beam setup, and (b) an edge illumination setup.

attenuated. The change in the mean, that is, the position of the peak, of the Gaussian curve is due to the refraction. It is clear from Fig. 1.17 that as the beam is refracted, the maximum of the IC, which is the point where the beamlet center and the pixel center align, is moved. Finally, the change in variance can also be measured, related to a contrast that has not yet been discussed, called *dark field* contrast. An increase in variance means that the beamlet got broader. The broadening is due to many refraction effects at the sub-voxel scale. Therefore, the dark-field signal is correlated to the inhomogeneity of the sample at the sub-voxel scale [59]. By defining a spatial local inhomogeneity coefficient for an object, $\sigma^2(x, y, z)$, the measured broadening S^2 at the detector can be

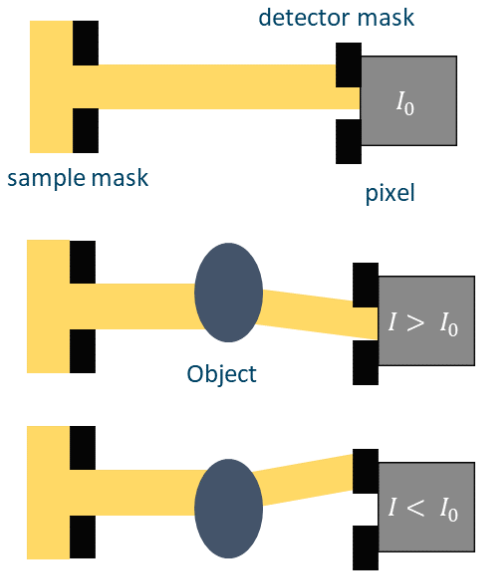
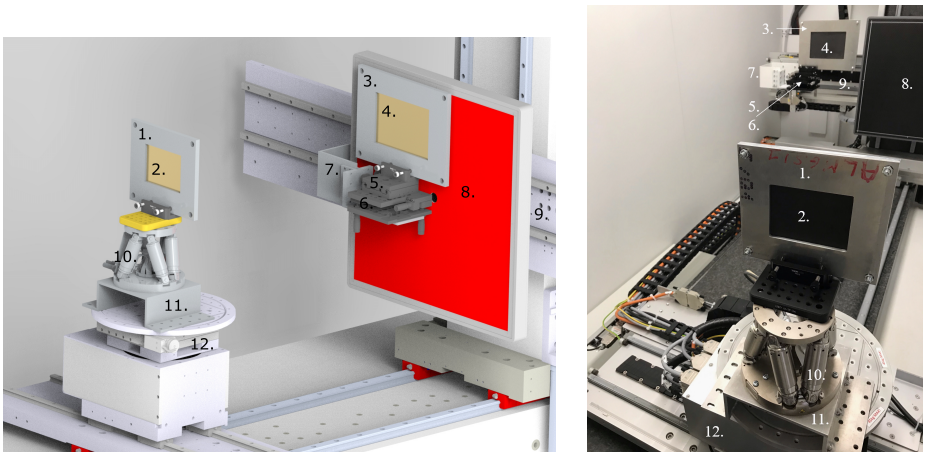


Figure 1.17. Illustration of the modulation of intensity by a refracting object without attenuation in an edge illumination set up.

defined analogously to the attenuation as [69]:

$$S^2(x, y) = \int \sigma^2(x, y, z)\mathrm{d}z. \tag{1.47}$$

To retrieve all three contrasts, attenuation, differential phase, and dark field, at least three measurements on the IC need to be acquired. With fewer images, the fitting problem becomes underdetermined in each pixel. Furthermore, since noise is present in each image, the fitting procedure will be improved by acquiring more than three samples of the IC. In Fig. 1.20 multiple phase steps of two fiber polymer blocks are shown with woven fiber bundles, held in place with plastic straws. The data was acquired at the Advanced X-Ray Imaging Group of Prof. A. Olivo, University College London. Note how in Fig. 1.20a and Fig. 1.20c the light and dark fringes are on opposite sides, as the phase steps were acquired on opposite slopes of the Gaussian IC. In Fig. 1.21 the three retrieved, flatfielded, contrasts are shown, retrieved from the phase steps in Fig. 1.20.



(a) (b)

Figure 1.18. The edge illumination set-up on the FleXCT scanner, (a) technical drawing, (b) real setup. **1.** Sample mask case **2.** Sample mask **3.** Detector mask case **4.** Detector mask **5.** Newport translation stage **6.** Newort tip/tilt platform **7.** Detector mask arm **8.** Detector **9.** Detector stage **10.** PI hexapod **11.** Sample mask bridge **12.** Sample stage. [73]

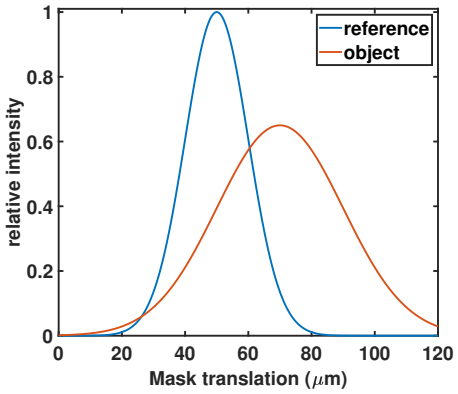
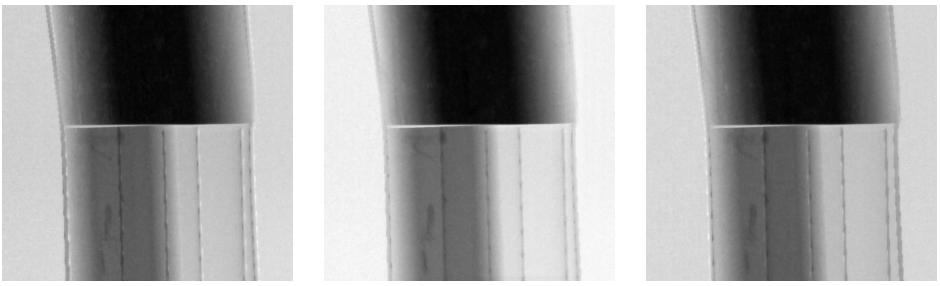


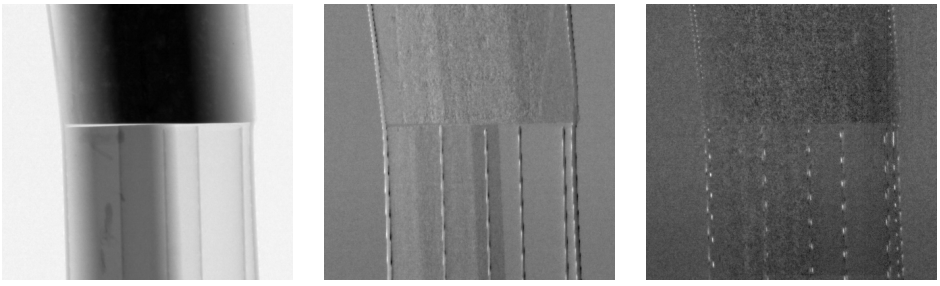
Figure 1.19. Comparison of measured ICs in one pixel with and without an object present.

There are multiple options to perform the Gaussian fitting needed for the phase retrieval. An overview was recently given in [74], where multiple iterative schemes - gradient descent, quadratic fit, and Gauss-Newton - and non-iterative phase retrievals based on moments are compared. No clear best method can be identified, as different methods provide the highest contrast-to-noise ratio for different contrasts. Furthermore, iterative methods take a considerable amount of time to run, which could



(a) −9 µm displacement **(b)** 0 µm displacement **(c)** 9 µm displacement

Figure 1.20. Different phase steps of two fiber polymer blocks with woven fiber bundles, held in place with plastic straws. (b) is the image at the top of the reference IC, (a) and (c) are measured on opposite sides of this maximum.



(a) Attenuation **(b)** Refraction **(c)** Dark field

Figure 1.21. Retrieved contrasts of two fiber polymer blocks with woven fiber bundles, held in place with plastic straws.

make the methods directly based on moments more attractive.

1.4.3 Edge illumination CT reconstruction

Each of the retrieved contrasts from images acquired with the edge illumination setup can be modelled as line integrals through an object. For attenuation and dark field, the integral equations were defined in (1.45) and (1.47), respectively. For the angle of refraction a slight change to (1.46) is required, with the measurable angle $\alpha(x, y)$ now given as [69]:

$$\alpha(x, y) = \frac{\partial}{\partial x} \int \delta(x, y, z) \mathrm{d}z, \tag{1.48}$$

since the edge illumination setup is only sensitive to refraction in one direction.

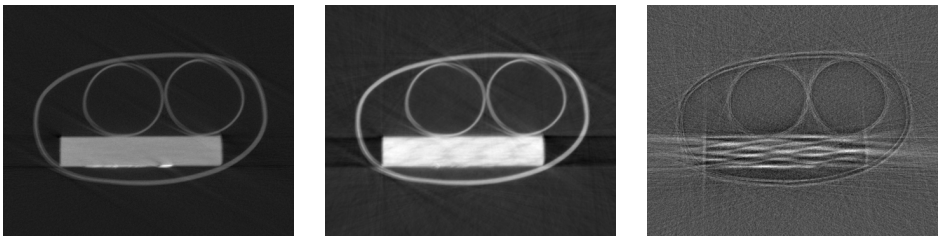
Just as in the conventional X-ray setup discussed in Section 1.1, one can again acquire edge illumination radiographs from multiple angles in an attempt to reconstruct the 3D structure of the scanned object. However, from EI data, three volumes, $\boldsymbol{x}_\mu, \boldsymbol{x}_\delta, \boldsymbol{x}_\sigma$, can be reconstructed, one for each contrast. The scan times are significantly increased compared to conventional CT, since for each projection angle, multiple phase steps must be acquired. After acquisition, a Gaussian fitting is performed to every pixel separately. The result of the fitting procedure is three separate sinograms, one for each contrast. Given $\boldsymbol{p}$ the full set of projection data acquired with EI-CT, define $\boldsymbol{p}_a, \boldsymbol{p}_b$ and $\boldsymbol{p}_c$ as the retrieved amplitude, mean position, and standard deviation, respectively. As in conventional CT, a flatfield $\boldsymbol{I}_0$ is acquired before and/or after the EI-CT acquisition. However, for EI, the flatfield scan also requires phase stepping, such that after retrieval the flatfield IC parameters $\boldsymbol{I}_{0,a}, \boldsymbol{I}_{0,b}$ and $\boldsymbol{I}_{0,c}$ can be defined. The three single contrast sinograms $\boldsymbol{p}_\mu, \boldsymbol{p}_\delta$ and $\boldsymbol{p}_\sigma$ are then defined as

$$\begin{aligned} \boldsymbol{p}_\mu &= -\ln\left(\frac{\boldsymbol{p}_a * \boldsymbol{p}_c}{\boldsymbol{I}_{0,a} * \boldsymbol{I}_{0,c}}\right) \\ \boldsymbol{p}_\delta &= \boldsymbol{p}_b - \boldsymbol{I}_{0,b} \\ \boldsymbol{p}_\sigma &= \boldsymbol{p}_c^2 - \boldsymbol{I}_{0,c}^2 \quad , \end{aligned} \tag{1.49}$$

with all operations with vectors defined pointwise.

The integral equations for the retrieved contrast described previously in (1.45) and (1.47) show that, for both attenuation and dark field contrast, reconstructions from the respective retrieved sinograms can again be made using FBP with a ramp filter. In order to reconstruct the real part of the refractive index, there are two options. The first option is to first numerically integrate the differential sinogram $\boldsymbol{p}_\delta$ along the sensitivity axis, leaving a sinogram from which $\boldsymbol{x}_\delta$ can be reconstructed using standard FBP with ramp filter. The second option is to reconstruct $\boldsymbol{x}_\delta$ from $\boldsymbol{p}_\delta$ with a filtered back projection directly, however, using a Hilbert filter [2] instead of the ramp filter. In [75] it was shown that the Hilbert filter is the theoretically correct filter to use in the case where the directional derivative of the Radon transform is considered. In Fig. 1.22 reconstructions are shown of the three contrasts made with FBP with a ramp or Hilbert filter depending on the contrast. The sinograms used for reconstruction are those of the central slice of an EI-CT dataset of the sample shown in Fig. 1.21

As an alternative to analytic methods, iterative methods can be used on each of the contrasts separately. The reconstruction problems are



(a) Attenuation **(b)** Phase **(c)** Dark field

Figure 1.22. Reconstructions from the retrieved contrast sinograms of the fiber polymer EI-CT dataset.

formulated as:

$$\begin{aligned} \boldsymbol{p}_\mu &= \boldsymbol{A}\boldsymbol{x}_\mu \\ \boldsymbol{p}_\delta &= \boldsymbol{D}\boldsymbol{A}\boldsymbol{x}_\delta \\ \boldsymbol{p}_\sigma &= \boldsymbol{A}\boldsymbol{x}_\sigma \quad , \end{aligned} \tag{1.50}$$

with $\boldsymbol{A}$ the projection matrix, $\boldsymbol{D}$ a matrix performing discretised differentiation, for example forward or central differences. For reconstructions of $\boldsymbol{x}_\mu$ and $\boldsymbol{x}_\sigma$, one could run any of the iterative reconstruction techniques discussed so far in this work. For the phase equation, a separate version of ART has been developed [76]. An effective and easy to implement candidate for an iterative reconstruction technique would be gradient descent with Barzilai-Borwein steps, as discussed in Section 1.3.

The reconstruction techniques discussed so far have all required a 2-step procedure, preprocessing the projection data such that the resulting sinograms fit well to a linear model. One could also model the entire image acquisition process and develop a single step *joint reconstruction algorithm*, reconstructing multiple contrasts at once. Two benefits of a joint approach can be identified. First, by combining the reconstruction and IC fitting steps, the fitting is, in some sense, regularised, as it has to be consistent between projections, which could add robustness. Second, the algorithm takes the projections directly as input and, therefore, no longer requires the IC for each pixel to be adequately sampled at each projection angle for the fitting. As a result, more flexibility can be achieved with acquisition schemes, such as single shot acquisitions where no phase stepping is performed per projection angle [77, 78]. Such an acquisition would be impossible to use in the 2-step pipeline, as no Gaussian can be fit to a single point. Although the acquisition can be made fully flexible, enough information about the IC's at different points still needs to be gathered to reconstruct the different contrasts. For edge illumination, a joint reconstruction algorithm reconstructing

attenuation and phase was presented in [77]. Further improvements to the method by extending the method to account for offsets in the sample mask position due to drift or vibrations are presented in [79, 80], showcasing the increased flexibility of iterative reconstruction methods. The existing joint reconstruction methods for EI do not include dark field reconstructions, which can provide additional information on sample properties [81]. Furthermore, no investigation of optimal solvers for the objective function was performed in [77]. In Chapter 5 we present a novel joint reconstruction method for EI, which enables the reconstruction of all three contrasts from single-shot datasets.

CHAPTER 2

Discrete reconstructions from polychromatic X-ray CT data

This chapter has been published as:
Six, N., De Beenhouwer, J. and Sijbers, J., 2019. poly-DART: A discrete algebraic reconstruction technique for polychromatic X-ray CT. Optics Express, 27(23), pp.33670-33682.

Parts of the published method section have been omitted because the concepts were already introduced in Chapter 1.

2.1 Introduction

X-ray computed tomography (CT) is a well-known imaging method in which the interior of an object is reconstructed from a set of X-ray radiographs. High-quality CT imaging generally requires hundreds or even thousands of radiographs acquired in a circular orbit with a large angular range. In many applications of X-ray CT, however, there is a need to lower the amount of radiographs to be acquired, either to reduce the dose [82, 83, 84] or to shorten the acquisition time [85, 86]. In time-critical processes [87, 88, 89], for example, the object may change during the scan, so reducing the number of projections could lead to improved image quality and / or higher time resolution. Industrial CT for quality control, specifically in-line inspection, again has a need for reduced scanning times and, in the case of larger parts, reconstruction from a less limited angular range [90, 82].

Unfortunately, lowering the number of projections to reduce the dose or decrease the acquisition time directly corresponds to a reduction of the sampling density in projection domain, which, if no precautions are taken,

leads to undersampling artefacts in the reconstructed CT image. Indeed, when only a few projections are available, the inverse problem of image reconstruction becomes severely underdetermined, causing streaking artefacts in the reconstructed image. By including prior knowledge about the sample, this underdetermination can be reduced to increase the quality of the reconstructed image.

One of the ways prior knowledge about the object to be imaged can be enforced, is Discrete Tomography, in which the number of possible materials in the sample is assumed to be known (and limited). For an overview of the mathematical background of discrete tomography, and several applications in medicine, we refer to [44]. The Discrete Algebraic Reconstruction Technique (DART) [47] is a practical discrete tomography algorithm which, if the sample consists of only a few materials, has been shown to generate reconstructions of much better quality than those from conventional reconstruction methods. The method has been applied in a variety of imaging domains, such as X-ray CT [48, 50, 51, 49], Electron tomography [91, 92, 93], Optical Diffraction Tomography [94], and Magnetic resonance imaging [95].

Many variations of the DART algorithm have been researched. SDART [96] and DART-ALBM [97], for example, focus on improving robustness against noise. rmwDART [98] specifically targets reduction of missing wedge artefacts. In PDART, the discreteness assumption on the whole sample is relaxed to local areas, making DART applicable to samples that are only partially discrete [93, 99]. In PDM-DART, the grey values corresponding to the different materials are estimated in an automated way [100]. ADART [101] gradually reduces the number of unknowns in each iteration, and MDART [102] follows a multi-resolution approach. EOD-DART [90], EDART [103] and TVR-DART [104] add additional prior knowledge about the sample to further improve the reconstruction quality.

All of the above DART based methods rely on the same linearised acquisition model as most classical techniques: it assumes that the log-corrected normalised projection is the sum of attenuation values along a ray. This is an adequate model for monochromatic but not polychromatic X-ray sources. As the X-ray beam travels through the sample, the low-energy photons are attenuated more easily than those with high energy. This results in the absorption along a ray from a polychromatic source being a non-linear function of sample thickness. Reconstructing such a dataset with a linear reconstruction technique will lead to well-known beam hardening (or cupping) artefacts [1]. At the same time, a single material no longer has a single defined attenuation value, as this value is energy dependent. Beam hardening artefacts and the inability

to select a clear singular attenuation value for each material, lead to inaccurate reconstructions when DART is used in combination with polychromatic projections. One of the current methods for reducing beam hardening artefacts consists of placing metal filters in front of the source, to pre-harden the beam. However, the pre-hardened spectrum is still polychromatic. Moreover, the filtering decreases the number of photons available for imaging and hence lowers the SNR of the reconstructed image.

We propose a new discrete tomographic method for polychromatic X-ray CT, which incorporates a polychromatic forward model. Such an approach has previously successfully been introduced to extend a classical algebraic reconstruction method (SART [5]) to handle polychromatic X-ray data, referred to as pSART (*polychromatic Simultaneous Algebraic Reconstruction Technique*) [34, 105]. In this paper, we propose a discrete reconstruction method that combines the best of both worlds: accounting for polychromaticity while exploiting the strength of DART to substantially reduce undersampling artefacts.

The paper starts with a brief overview of DART. Next, we explain the poly-DART algorithm and the polychromatic model that is used. Finally, different discrete reconstructions from polychromatic datasets are compared to demonstrate the improvements of the proposed method. These datasets include Monte Carlo simulated data using the GATE framework [106], as well as an experimental dataset.

2.2 Method

We propose a new algorithm based on the principles of DART [47] and pSIRT [34], aiming to exploit the benefits of both methods. Both methods were described in detail in Section 1.2.2. In the first phase of the algorithm, an initial reconstruction $\boldsymbol{x}^{(0)}$ is generated by performing a small number of pSIRT iterations. From this reconstruction and the attenuation values at the reference energy, the optimal grey values and thresholds are estimated by minimising the projection difference (i.e., the difference between the polychromatic forward projection and the projection data) in a mean squares sense. This is a polychromatic extension of the projection distance minimisation (PDM) algorithm of van Aarle et al [100]. It allows to correct for small errors in the assumed attenuation coefficients and/or in the estimated spectrum.

Next, the initial reconstruction is segmented with the estimated segmentation parameters. Based on the segmented result, the pixels are classified into boundary or non-boundary pixels. All boundary pixels and

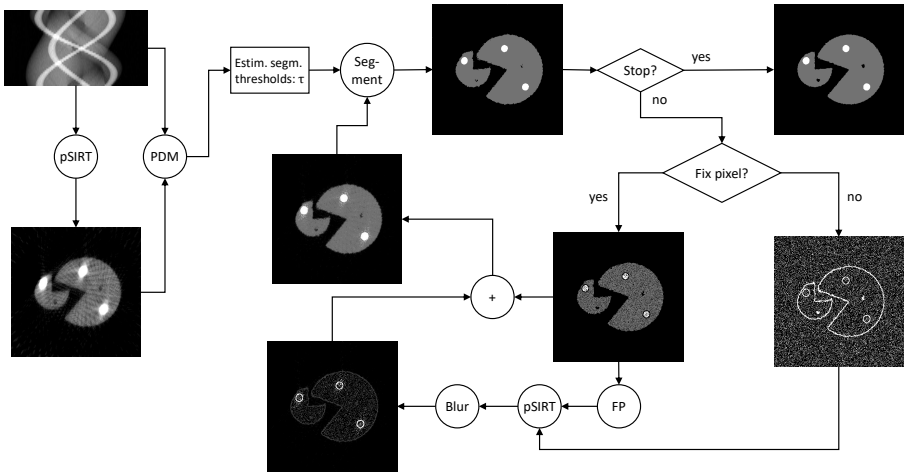


Figure 2.1. Flowchart for the poly-DART algorithm.

a set of random pixels are marked 'free' and will be updated in the next reconstruction step. The other pixels remain fixed to their segmented value. Then, the line lengths through each material for the fixed pixels are precomputed to be used in the reconstruction, as these will not change until the next segmentation step. Masked pSIRT iterations are then performed, which only update the free pixels. Furthermore, these local iterations are relaxed by a parameter λ, equal to the percentage of free pixels, as the much lower amount of free pixels (compared to a conventional update over all pixels) would otherwise lead to high fluctuations in each step of the reconstruction. Hence, the (internal) poly-DART update step becomes:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \lambda \boldsymbol{C}\boldsymbol{A}^{\top}\boldsymbol{R}\left(\text{polyProj}(\boldsymbol{x}^{(k)}) - \boldsymbol{p}\right), \tag{2.1}$$

Lastly, this reconstruction is segmented again and the previous steps are repeated until a stop criterion is met. A flowchart of poly-DART can be found in Fig. 2.1. The algorithm was implemented in Matlab, making use of the DART framework of the ASTRA toolbox [107].

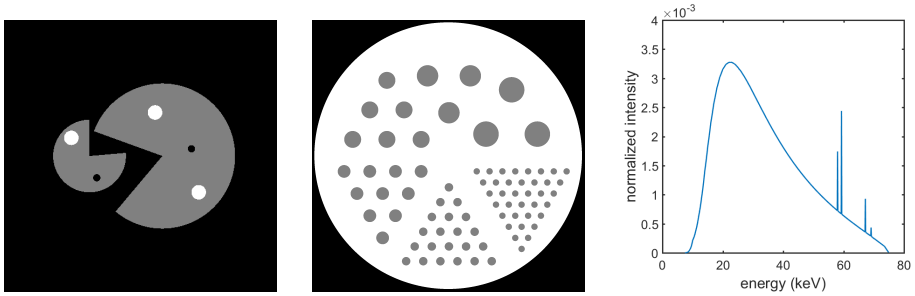
2.3 Experiments

2.3.1 Monte Carlo simulations

To validate the proposed algorithm, two simulation experiments were set up, with the phantoms shown in Figs. 2.2(a) and 2.2(b). Both phantoms consist of two materials, plexiglass and aluminium, suspended in air. The

attenuation values at different energy levels were used from the National Institute of Standards and Technology [35]. To avoid inverse crime, the phantom was analytically defined in the simulation framework, GATE [106]. For both phantoms, 300 fan-beam polychromatic projections were created using GATE. The simulated source was a 75 kVp X-ray source with a tungsten anode, of which the spectrum is shown in Fig. 2.2(c). Spekcalc was used to generate the source spectrum [108]. The projection angles were chosen in the interval $[0, 2\pi]$ using golden angle sampling, which means that the subsequent projections are $\frac{1+\sqrt{5}}{2}\pi$ radians apart from each other. The detector pixel size was 0.375 mm and the scanner set up had a magnification factor of 4. The reconstructed pixel size was 0.09375 mm. The detector was modelled as a photon counting detector with 400 pixels made of silicon, with a minimal activation energy of 5 keV. The detector was idealised in the sense that no cross-talk or blurring between pixels was present. As the projections were created using a Monte Carlo method, Poisson noise is present in the images. We simulate different noise levels by varying the number of photons from $4 * 10^5$ to $2 * 10^7$ photons for each projection, which is on average between 1000 and 50000 photons per pixel. For the experiments where varying noise levels were not taken into account, we chose the dataset with 10000 photons per detector pixel. To further simulate a real experiment, the tube spectrum (S) and detector response (D) were estimated with two step wedges. These wedges, one in plexiglass and one in aluminium, were also defined in GATE. Multiple projections were taken orthogonal to the steps and then averaged to reduce the effects of noise in the measurements. We used a maximum likelihood expectation maximisation (MLEM) algorithm to estimate the product $S(\epsilon)D(\epsilon)$ of the tube spectrum and the detector response from these measurements. The four characteristic peaks of a tungsten anode source spectrum were added manually to the initial guess for the MLEM iterations.

From the simulated polychromatic projection data, images were reconstructed with the following methods: pSIRT, SIRT, DART with manually optimised grey levels and poly-DART with automatically optimised grey levels. For the selection of the grey values for DART, the values with the lowest reconstruction error in the case of full angular sampling were chosen. Poly-DART was also compared to (segmented) SIRT and pSIRT. The segmentation for SIRT was performed with the same global threshold as DART. The segmentation for pSIRT was performed with the first set of estimated segmentation parameters that were obtained from the PDM step in the poly-DART algorithm. These automatically optimised grey values are only dependent on the initial pSIRT reconstruction and the



(a) Analytically defined 'Pac-Man' phantom consisting of air (black), Plexiglass (grey) and aluminium (white).

(b) Analytically defined Derenzo phantom consisting of air (black), Plexiglass (grey) and aluminium (white).

(c) Polychromatic spectrum of a 75 kVp X-ray source.

Figure 2.2. The Pac Man (a) and Derenzo (b) phantom, as well as the employed X-ray spectrum (c).

projection data, and are more accurate than using the attenuation values at the reference energy.

All poly-DART reconstructions were performed with 20% random free pixels, 50 initial and 5 inner iterations of pSIRT. The same settings were used for DART. The percentage of free pixels was chosen as advised in [47]. For the polychromatic reconstruction methods, the spectrum shown in Fig. 2.2(c) was rebinned to 70 bins. A reference energy of 55 keV was chosen emperically as optimal for pSIRT in terms of convergence speed and reliability [34]. The same reference energy was used for poly-DART. When testing the influence of the number of angles and the level of noise, 2000 (p)SIRT iterations were performed for all methods and the minimal achieved error over these iterations was then retained. That is, for both converging and diverging methods, the best reconstruction (minimal error) is retained.

2.3.2 Experimental dataset

Following the Monte Carlo simulations, a similar test was performed on an experimental dataset. The used phantom, called the Barbapapa phantom [109], consists of plexiglass with inserted aluminium rods and is shown in Fig. 2.3. A phantom with the same materials as in the simulation study was used to more easily compare the simulation results to the experimental results. A tube voltage of 130 kVp was employed. The phantom was scanned over 2400 equiangular projections in the interval $[0, 2\pi)$.

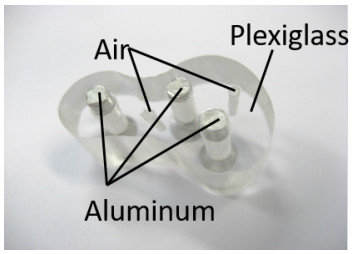
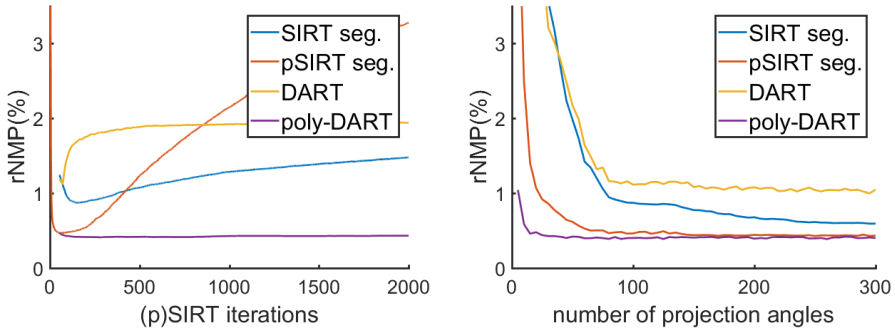


Figure 2.3. Barbapapa phantom used to generate the experimental dataset. The phantom is a smoothly shaped Plexiglass object, with three inserted aluminum rods and two air columns.

To estimate the spectrum, the same technique as in the simulation case was used. For this experiment, a PVC step wedge with 11 steps, ranging from approximately 1 mm to 18 mm in thickness, was scanned. Next, the same MLEM spectrum estimation method as in the Monte Carlo case was performed, with 130 keV as maximum energy used as a boundary condition. From the central slice of the Barbapapa dataset, images were reconstructed with the same methods: pSIRT segmented, SIRT segmented, DART, and poly-DART. All poly-DART reconstructions were performed with 10% random free pixels, 50 initial, and 10 inner iterations of pSIRT. The same settings were used for DART. A reference energy of 55 keV was chosen for pSIRT as well as for poly-DART. Since for real data no ground truth image is available, a pSIRT reconstruction from 2400 equiangular projection within $[0, 2\pi]$ was segmented using Otsu thresholding [110] and the result was regarded as the ground truth reconstruction. When testing the influence of the variable number of angles, 2000 (p)SIRT iterations were performed for all methods and the minimal achieved error over these iterations was retained. The error measure is the rNMP calculated based on the golden standard segmented reconstruction. For both the simulated and real data, all parameters (e.g. percentage of free pixels, number of inner iterations and reference energy level) were either chosen empirically or in accordance with the literature. All reconstructions were performed using the ASTRA toolbox [107].

2.4 Results

To quantify the performance of poly-DART against DART, segmented SIRT, and segmented pSIRT, we computed the *relative number of misclassified pixels* or *rNMP* [100]. The rNMP is calculated as the number of pixels in the reconstruction that have a different grey value compared



(a) Plot of rNMP vs. (p)SIRT-iterations on 100 projections.

(b) Plot of rNMP vs. number of projection angles, taken in $[0, 2\pi)$ with golden angle sampling.

Figure 2.4. Plots showing the change in error by varying (a) number of iterations and (b) number of projection angles on the simulated Pac-Man phantom.

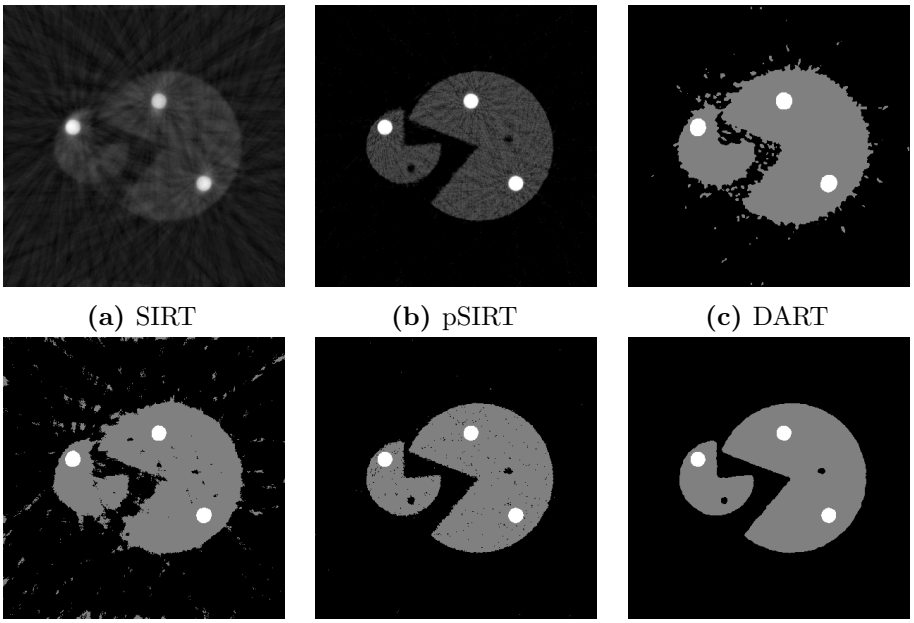
to those of the corresponding ground truth image, divided by the total number of non-background pixels in the ground truth image.

2.4.1 Monte Carlo simulations

Pac-Man phantom

In Fig. 2.4(a), the rNMP of the different methods is shown as a function of the number of (p)SIRT iterations, for 300 projection angles between 0 and 2π. The rNMP for (p)SIRT was calculated at every iteration, whereas the rNMP for (poly-)DART was calculated at every DART iteration, i.e. every 5 (p)SIRT iterations. From Fig. 2.4(a), one can observe that both DART and poly-DART converge. However, poly-DART does so to a solution with a significantly lower error, whereas DART increases the error from its start point to its converged solution. The continuous methods, SIRT and pSIRT, have a best solution after a low amount of iterations and start to diverge from that point onwards. This is a known problem for continuous methods, as they tend to overfit to the noise in the projection data.

Next, the effect of varying amounts of projection angles on the rNMP was studied. These results are shown in Fig. 2.4(b). Again, poly-DART outperforms the other methods in terms of rNMP, reaching a lower rNMP for each number of projection angles. This suggests that poly-DART benefits from both the beam hardening correction property of pSIRT and the imposed discreteness. Note that because the minimal error for



(a) SIRT **(b)** pSIRT **(c)** DART

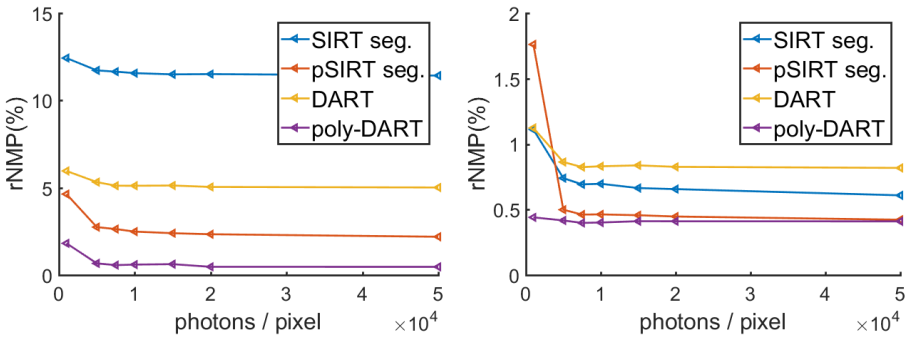
(d) SIRT segmented **(e)** pSIRT segmented **(f)** poly-DART

Figure 2.5. Comparison of the different reconstruction techniques with 40 projection angles on the simulated Pac-Man phantom.

the methods over all iterations is compared, this is the best-case scenario for SIRT and pSIRT, as they reach a minimum at some a priori unknown point. For down to about 150 projection angles, poly-DART and pSIRT have a comparable error. However, at lower amounts of projections, poly-DART outperforms pSIRT.

A comparison of the different reconstruction methods for 40 projections can be found in Fig. 2.5, where each method was run for 250 (p)SIRT iterations. From this figure, it can be observed that the poly-DART algorithm shows the least reconstruction artefacts, with the other methods showing beam hardening artefacts and streaks due to the low amount of projection angles. Furthermore, the homogeneous interior of the object is well reconstructed with poly-DART, while pSIRT shows many misclassified (black) pixels.

Finally, we tested the influence of noise on the reconstructions. Different levels of noise were simulated by varying the number of photons emitted by the source. The amount of photons ranges from, on average, 1000 to 50000 photons per detector pixel. The minimal achieved rNMP over 2000 iterations is plotted as a function of the number of photons. In Figs. 2.6(a) and 2.6(b) this rNMP is plotted for all methods for 10 and



(a) Plot of rNMP vs. photon count on 10 projection angles.

(b) Plot of rNMP vs. photon count on 100 projection angles.

Figure 2.6. Plots showing the influence of noise on the reconstruction error for the simulated Pac-Man phantom.

100 projections, respectively. From these plots, it can be observed that poly-DART consistently outperforms the other techniques, especially for low amounts of photons.

Derenzo phantom

The rNMP is not a suitable error measure for the Derenzo phantom, as every misclassified pixel contributes equally to the error. Of more interest in this phantom is to what extent the different cylinders can be discerned. Therefore, we visually show reconstructions for all methods as functions of different numbers of projections to illustrate the level of visibility. These reconstructions, all after 450 (p)SIRT iterations, are shown in Fig. 2.7. From Fig. 2.7, similar reconstruction results of the Derenzo phantom as for the Pac-Man phantom can be observed. Even when a large number of projections is available, two artefacts arise as before, beam hardening and overfitting to noise. Both SIRT and DART suffer from the beam hardening artefacts. For 50 projections and less, these beam hardening artefacts, coupled with undersampling, make it increasingly more difficult to see the cylinders at any resolution. Neither the poly-DART nor pSIRT reconstructions show significant beam hardening artefacts. Even with a large number of projections, pSIRT shows artefacts due to overfitting to noise, though the cylindrical features can still be discerned at every resolution. As the number of projections decreases, both the size and the placement of the cylinders become harder to distinguish from the pSIRT reconstruction. The poly-DART reconstructions are more accurate for each number of projections, most notably when less than 50 projections

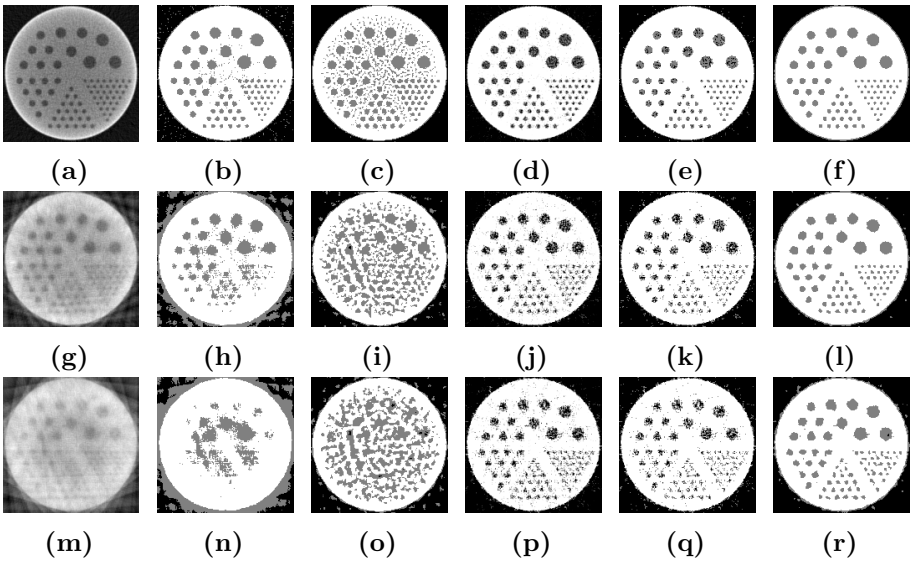


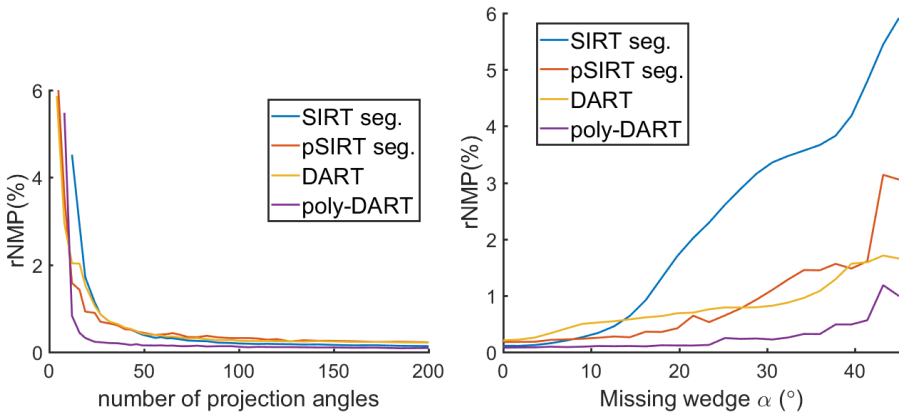
Figure 2.7. Comparison of the different reconstruction techniques on the Derenzo phantom. Techniques in the columns from left to right: SIRT, SIRT segmented, DART, pSIRT, pSIRT segmented and poly-DART. (a-f) 300 projections, (g-l) 50 projections, (m-r) 25 projections.

are present. Both position and size of the cylinders more closely match those of the phantom image in Fig. 2.2(b).

2.4.2 Experimental dataset

Based on the experimental data, two experiments were performed. First, the effect of varying amounts of projection angles on the rNMP was studied. As can be observed from Fig. 2.8(a), poly-DART outperforms the other methods in terms of misclassified pixels, reaching a lower rNMP at any number of projection angles. The plot looks similar to the one in Fig. 2.4(b), indicating that the results obtained with the Monte Carlo framework are consistent with the experimental data. Reconstructions with the different methods on the data of 40 projection angles after 450 (p)SIRT iterations are shown in Fig. 2.9.

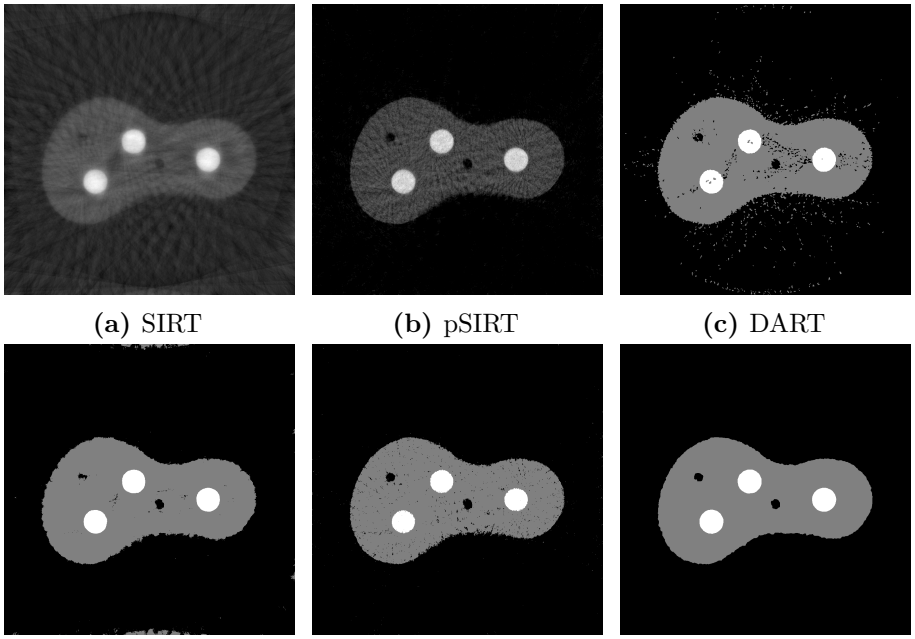
Secondly, we studied the effect of a missing wedge in the projection data, i.e. a dense sampling of the object, but over a limited range $[\alpha, \pi - \alpha] \cup [\pi + \alpha, 2\pi - \alpha], \alpha \in \mathbb{R}$. For the full sampling, 400 equiangular projections were taken from the dataset. From the previous experiment, it is clear that this is a sufficiently dense sampling for all methods to generate high-quality reconstructions. Next, a number of projections were deleted symmetrically around both 0 and π. For the created missing wedge



(a) rNMP vs. number of projection angles.

(b) rNMP vs. missing wedge.

Figure 2.8. Plots of the effect of (a) varying number of projection angles, equally spaced in $[0, 2\pi)$ and (b) a missing wedge in the projection data, on the rNMP for the experimental dataset.



(a) SIRT **(b)** pSIRT **(c)** DART

(d) SIRT segmented **(e)** pSIRT segmented **(f)** poly-DART

Figure 2.9. Comparison of the different reconstruction techniques with 40 projection angles of the experimental data.

datasets in this experiment, $\alpha \in [0, \frac{\pi}{4}]$ was used. As before, the minimal error over 2000 iterations was plotted in Fig. 2.8(b). The DART algorithm outperforms SIRT from about 15° of missing wedge and outperforms pSIRT after around 25°. The proposed poly-DART algorithm reaches the lowest reconstruction error of all the tested methods at any of missing wedge angle. This shows that the imposed prior knowledge is able to compensate for the missing data.

2.5 Conclusion

Many objects consist of a limited number of materials. Using discrete tomography, this prior knowledge can be exploited in the reconstruction of images from X-ray projection data to reduce undersampling artefacts. Current discrete tomography methods, however, do not account for polychromaticity of X-ray sources, leading to various reconstruction artefacts and limiting their applications. In this paper, poly-DART was proposed, a discrete tomography method that exploits the sparseness in the attenuation values, while taking a polychromatic projection model into account. Reconstruction experiments on both simulated and experimental data from polychromatic sources revealed that poly-DART results in substantially improved image reconstruction quality compared to DART or segmented versions of SIRT or pSIRT for polychromatic X-ray data. This allows DART, which has been successfully used in a monochromatic setting in different applications, to be extended to data acquired with polychromatic lab sources.

CHAPTER 3

Dynamic update strategy for discrete tomography methods

The work presented in this chapter was performed in collaboration with Daniel Frenkel. This chapter has been published as: Frenkel, D., **Six, N.**, De Beenhouwer, J. and Sijbers, J., 2022. Tabu-DART: a dynamic update strategy for efficient discrete algebraic reconstruction. The Visual Computer, pp.1-13.

Parts of the published method section have been omitted because the concepts were already introduced in Chapter 1.

3.1 Introduction

In X-ray Computed Tomography (XCT), the interior of an object is commonly visualised by reconstructing an image from a large number of radiographs, equiangularly acquired over 180 or 360 degrees. If scan time restrictions or geometrical constraints during scanning apply, only a small number of radiographs or a set of radiographs distributed over a limited angular range will be available, respectively. In such ill-posed limited data problems, conventional reconstruction methods, such as Filtered Back Projection (FBP) or the Simultaneous Iterative Reconstruction Technique (SIRT) [111], lead to images with severe artefacts [47] and semi-convergent behaviour [112].

Including prior knowledge about the scanned object into the reconstruction process is a well-known strategy to compensate for limited data in XCT [113, 114, 115]. A specific type of prior knowledge is exploited in Discrete Tomography (DT) [46], where the object is assumed to be composed of only a few materials. The variety of work on discrete

tomography is vast [45, 116, 117, 118], with several algorithms developed to improve robustness with respect to noise [96, 104, 119, 120], handle partially discrete images [121, 122], and polychromatic data [123, 124].

Despite their strengths, practical DT methods are computationally intensive, as they primarily rely on iterative reconstruction. To increase speed, divide-and-conquer strategies are often employed, in which only a part of the image is updated in each iteration [121, 125, 46]. Amongst the practical DT algorithms that rely on such division strategies, the Discrete Algebraic Reconstruction Technique (DART) [47] is well known for producing high-quality reconstructions of objects composed of few different materials, even in cases with a limited number of projections or projections acquired in a limited angular range [126]. DART has been successfully applied in various imaging domains [51, 91, 93, 95] and is a common benchmark method to compare new DT algorithms with [127, 128, 129, 130]. New reconstruction methods based on the DART methodology are still being introduced [123, 131, 132, 133].

Despite the benefits of DART, its computational complexity is high. One of the causes is that the update rules in DART are predetermined and hence do not change over the course of the reconstruction [134, 135]. As a result, already well-reconstructed image regions continue to be updated, leading to redundant computation. This problem has been addressed in theoretical DT, where Tabu-search theory has been combined with other DT methods such as combinatorial optimisation approaches based on Ryser's algorithm [136, 137] and with binary reconstruction based on Gibbs priors [46]. However, these approaches are infeasible for large-scale problems, due to both memory requirements and computation time. Heuristic methods such as DART have better scalability than theoretical DT methods, but still suffer from long computation times, partly caused by redundant computation.

To reduce redundant computation of DART-like methods, we propose a framework of dynamic update rules, which combines concepts from Tabu-search theory with update strategies. We introduce a probability map that adapts based on feedback received during subsequent reconstruction steps. By expressing update rules as changes to this probability map, dynamic update strategies during the reconstruction are implemented. The initialisation of this map was based on the entropy of the reconstruction, a measure used before in discrete tomography in the context of optimal projection angle selection [138] and in the non-discrete case to measure the uncertainty of the grey value [139]. As a proof-of-concept, we present such a framework for DART. Furthermore, we describe an estimation procedure for the initial state of the probability map based on image uncertainty.

3.2 Method

In Section 3.2.1, a brief overview of Tabu-search and related concepts is presented, together with potential implications of using memory structures in DART. In Section 3.2.2, the DART update step is generalised as a framework that introduces a probability map to function as a memory structure for the partitioning step (step 2) inside the algorithm. The proposed Tabu-DART algorithm is described, in which the probability map is adapted based on a dynamic set of rules and feedback received from the segmentation step. Finally, in Sections 3.2.3 and 3.2.3, the map initialisation and feedback loop are explained for Tabu-DART.

3.2.1 Memory structures and Tabu-search

Tabu-search is a variations strategy for mathematical optimisation techniques that rely on local search. The nature of local search methods makes them vulnerable to local optima. Tabu-search aids in finding the global optimum through adaptive memory structures and reaches parts of the solution space that would otherwise be left unexplored. It allows to escape from local optima and intensifies searches inside a specific region around a solution. In the next paragraph, a summary of the Tabu-search concepts is given to clarify our contribution. For a more in-depth description, we refer to [140].

There are four main factors which describe the memory structure used: recency and frequency based memory, quality, and influence. Recency-based memory stores information on recent solutions explored, and aids in preventing already visited solutions in favour of exploring worse but yet unvisited solutions. Frequency-based memory stores information on the number of times a certain attribute has appeared in recent solutions. Quality relates to the ability to differentiate between characteristics of good and bad solutions, while influence stores the impact of changes in structure of the solution. It is infeasible to store multiple solutions for large 3D volumes. Hence, recency-based memory has limited function for algebraic reconstruction with DART. The frequency of favourable attributes relating to good reconstruction can, however, be stored and exploited to improve DART. For this reason, our approach relies on frequency-based memory. The use of quality and influence metrics is limited to a feedback loop, which adapts the memory structure we propose for DART. When many reconstructions with a low error share an attribute, exploring locations in the reconstruction space where this attribute will be present increases the probability of finding a reconstruction that minimises the error. Image features, such as which pixels still change

their grey value or whether or not the boundary between different grey values has stopped evolving, are valuable attributes that can be tracked in *frequency*. In Section 3.2.3, we describe how changes with respect to such a feature can be tracked to adapt the partitioning step (step 2) in the DART algorithm and make it more efficient over time.

3.2.2 The probability map framework

In DART, the partitioning rules decide which pixels in the image are updated, and hence they have a significant impact on the quality of the resulting reconstruction. The following probability map functions as frequency based memory for the partitioning step inside DART:

$$P : \mathbb{R}^{n_{\text{vol}}} \to [0,1]^{n_{\text{vol}}}, \boldsymbol{x} \mapsto \boldsymbol{p}_{\boldsymbol{x}}. \tag{3.1}$$

Instead of one parameter p describing the probability that an interior pixel is updated in the next iteration, a probability p_{x_i} is linked to each pixel x_i, which decides whether or not to update that pixel in the next iteration. The map functions as a tracker of the frequency of change for any metric that distinguishes between pixels that are likely to be correctly classified and those that are not.

To correctly incorporate the update probability map, certain steps are different from the original DART algorithm. First, an initial state for the probability map is created after the initial SIRT reconstruction. This state is based on any available or calculated image uncertainty measure. If a region in the reconstructed volume is well-resolved, the probabilities in that region can be lowered to reduce redundancy. During each partitioning step, a random number r_i is drawn from a uniform distribution between 0 and 1, for each pixel x_i. If $r_i < p_{x_i}$, the pixel is selected for update. This samples a binary probability distribution in each pixel x_i, with probability p_{x_i} to be free. Hence, the creation of the fixed and free partitions depends entirely on the probability map. At the end of each DART iteration, a feedback loop is introduced, which updates the probability map based on the current reconstruction data. A flow chart of the Tabu-DART algorithm is shown in Fig. 3.1.

3.2.3 Probability map initialisation

An initialisation scheme is presented for the probability map to eliminate the need for the parameter p in the original DART algorithm. Initialisation is based on a generalisation of local image uncertainty as proposed by Varga et al. for binary reconstruction [141]. Each pixel can only attain a value in $\mathcal{R}$ and the probability of being equal to ρ_i is spatially dependent.

A formula for generating the probability for theoretical DT (illustrated in Fig. 3.2) is given by

$$P(x_j = \rho_i) = \frac{\#\{\boldsymbol{s} \in \{\rho_1, ..., \rho_k\}^{n_{\text{vol}}} |\ \boldsymbol{A}\boldsymbol{s} = \boldsymbol{p},\ s_j = \rho_i\}}{\#\{\boldsymbol{s} \in \{\rho_1, ..., \rho_k\}^{n_{\text{vol}}} |\ \boldsymbol{A}\boldsymbol{s} = \boldsymbol{p}\}}, \tag{3.2}$$

Hence, each pixel x_j can be linked to a probability vector $\boldsymbol{v}_{x_j} \in [0,1]^k$, where k is the number of distinct gray values in the image. The entropy, defined as

$$\mathcal{H}(x_j) = -\boldsymbol{v}_{x_j}^\top \log_k \boldsymbol{v}_{x_j}, \tag{3.3}$$

translates this vector to a single value representing uncertainty of the gray value of pixel x_j. The logarithm $\log_k$ is applied pointwise on the different components of the vector $\boldsymbol{v}_{x_j}$.

Since it is infeasible to calculate the probabilities for large images directly, we propose an extension of the approximation introduced by Varga et al. [141]. For the pixel x_j of the initial ARM reconstruction, let

$$\begin{aligned} \boldsymbol{d}_{x_j} &= \left[\frac{1}{|x_j - \rho_1|}, ..., \frac{1}{|x_j - \rho_k|}\right] \\ \boldsymbol{v}_{x_j} &= \frac{\boldsymbol{d}_{x_j}}{\|\boldsymbol{d}_{x_j}\|_1}. \end{aligned} \tag{3.4}$$

The values $\mathcal{H}(x_j)$ are used to initialize the probability map. Note that one of the denominators in (3.4) may become zero if $x_j \in \mathcal{R}$, e.g. if the condition $x_j \geq 0$ is enforced during SIRT reconstruction possibly causing x_j to be set to ρ_1. To avoid division by zero, a lower bound was selected for the denominators in $\boldsymbol{d}_{x_j}$.

Dynamic update rules

As the final part of Tabu-DART, the following set of update rules is introduced to track a stability metric based on changes between grey values for individual pixels: Define $\boldsymbol{c}_{\boldsymbol{x}}, \boldsymbol{b}_{\boldsymbol{x}} \in \mathbb{R}^n$ such that

$$\begin{aligned} c_{x_j}^{(\ell+1)} &= \begin{cases} 0, & \text{if } s_j^{(\ell+1)} = s_j^{(\ell)} \\ 1, & \text{otherwise} \end{cases} \\ b_{x_j}^{(\ell+1)} &= \begin{cases} 1, & \text{if } s_j^{(\ell+1)} \text{ is boundary} \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Then, the new probability map $\mathbf{p}_{\mathbf{x}}^{(\ell+\mathbf{1})}$ is given by

$$\boldsymbol{p}_{\boldsymbol{x}}^{(\ell+1)} = \min\left(\frac{1}{2}\boldsymbol{p}_{\boldsymbol{x}}^{(\ell)} + \boldsymbol{c}_{\boldsymbol{x}}^{(\ell+1)} + \boldsymbol{b}_{\boldsymbol{x}}^{(\ell+1)}, \mathbf{1}_{n_{\text{vol}}}\right) \tag{3.5}$$

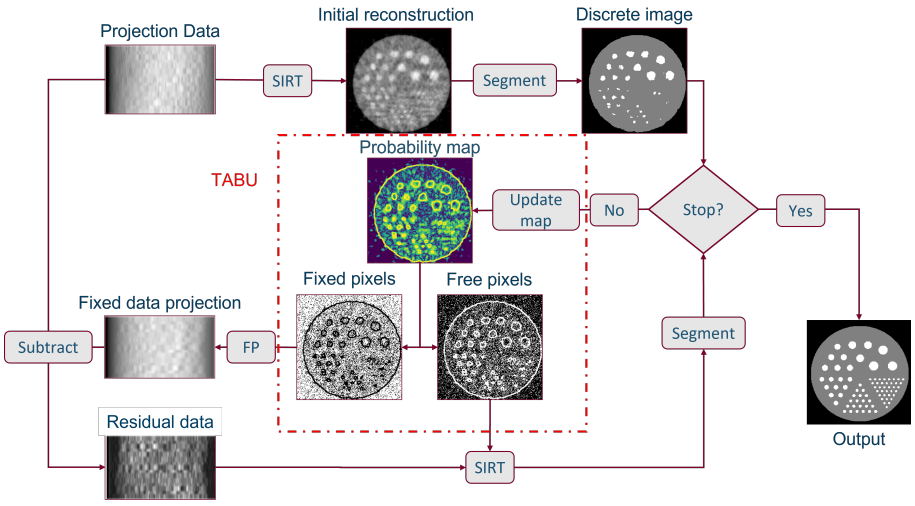
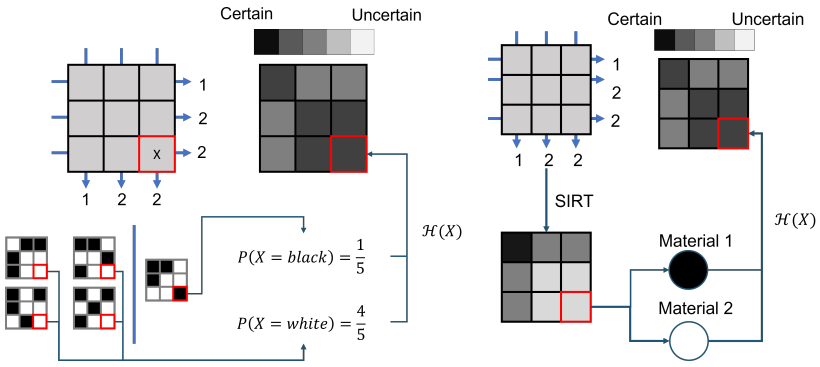


Figure 3.1. A flowchart of the Tabu-DART algorithm. After initial reconstruction and segmentation (upper part), the probability map is initialised. The map is used to label the image into free and fixed pixels (middle). The fixed pixel contribution is removed from the original projection data (left). The free pixels are then updated from the residual projection data (bottom) followed by another segmentation step. If the stopping criterion is satisfied, the output is a discrete reconstruction.



(a) Theoretical estimation of grey value uncertainty.

(b) Approximation of grey value uncertainty.

Figure 3.2. Theoretical (a) and approximate (b) uncertainty calculations in a single pixel. In (a), the exact entropy is found by counting all possible binary solutions to the projection problem. In (b), this entropy is approximated based on a SIRT solution, by using the distance between the reconstructed grey value and the a priori known grey values.

These update rules halve the probabilities of all non-boundary pixels that, when segmented, have the same grey value ρ_i as in the previous iteration. Otherwise, the probabilities are set to 1.

3.3 Experiments

Two sets of experiments were conducted. First, simulation experiments were performed to test the validity of our approach on four discrete phantoms from previous DART articles [47, 101, 134] before evaluating the accuracy of Tabu-DART on a polychromatic dataset of a plexiglass object [109]. The simulation experiments are described in 3.3.1 and the plexiglass dataset is introduced in 3.3.2.

3.3.1 Simulation experiments

Fig. 3.3 shows the phantoms that were used for the simulation experiments, which are identical to those used in previous DART publications [47, 104, 101]. The size of each phantom is 512×512 pixels. With the ASTRA toolbox [142, 107], projections were simulated following a parallel beam geometry with 512 detector values for each angle. Two cases of limited data were studied: In the first case, the acquisition range was $[0°, 180°]$ and the number of projections was varied from 2 up to 90. To maintain a uniform angular sampling distribution while studying the performance of Tabu-DART as a function of the number of projections, the latter were generated using a golden ratio angular sampling [143], which means that subsequent projections are $\frac{1+\sqrt{5}}{2}\pi$ radians apart from each other. In the second case, 90 projections were uniformly simulated, after which an increasing wedge was removed.

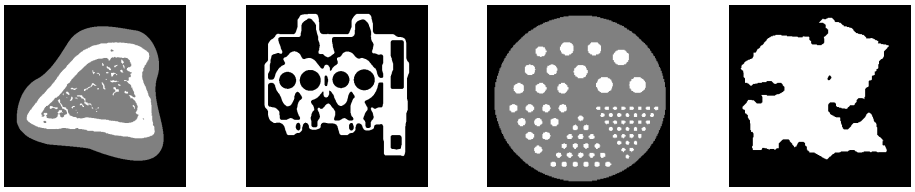
To infer whether the update probability parameter p can be avoided with our approach, we compared the Tabu-DART algorithm with the DART algorithm with a total of 12 choices for p. The best performing value for p for a specific case is denoted by *best p*. The other DART parameters were chosen according to the literature [47, 123] and have been kept constant throughout the experiment. These are shown in Table 3.1.

3.3.2 Experimental data: Barbapapa plexiglass phantom

The goal of the real data experiment is twofold. First, to provide evidence that Tabu-DART combines well with other augmentations of the original DART algorithm. Second, to study how the relaxation of the inner ARM iteration influences the overall reconstruction quality compared to DART.

Table 3.1. The values of the parameters used for DART and Tabu-DART for the basic simulation experiment.

Parameter	Value
# Initial SIRT iterations	100
# Intermediate SIRT iterations	10
The smoothing parameter	0.1
# DART iterations	100



(a) Phantom 1 **(b)** Phantom 2 **(c)** Phantom 3 **(d)** Phantom 4

Figure 3.3. The four phantoms that were used for the simulation experiments.

We reconstruct the central slice of the Barbapapa experimental dataset [109], which consists of a plexiglass block with two drilled cylindrical holes. Three aluminium rods were inserted into the block, amounting to a total of three different materials present: air, plexiglass, and aluminium. A picture of the object is shown in Fig. 3.4. A total of 2400 cone beam projections were measured over the full 360° range with a tube voltage of 130 kVp. To account for the polychromaticity of the X-ray beam, our Tabu-search framework was combined with a polychromatic version of DART, called poly-DART [123]. We refer to this polychromatic Tabu-DART algorithm as TP-DART. To this end, the polychromatic spectrum was first estimated. This was done by scanning a PVC step-wedge with steps ranging in thickness from 1 mm to 18 mm. The spectrum was then estimated using the Maximum Likelihood Expectation Maximisation algorithm as explained in [123]. A missing wedge experiment was set up, starting from 400 equiangularly distributed projections over a 360° range. Reconstructions are made from subsets of these projections with an increasingly larger missing angular wedge. These subsets consist of all projections in the range $[\alpha, 180° - \alpha] \cup [180° + \alpha, 360° - \alpha]$ with α varying from 10° to 60°. The parameters of poly-DART and TP-DART are given in Table 3.2.

The run-time parameters resulted in a total of 500 SIRT/pSIRT iterations being performed for each method. For this experiment, the

Table 3.2. Parameter values within poly-DART and TP-DART for the Barbapapa central slice reconstruction.

Parameter	Value
# Initial (p)SIRT iterations	50
# Interior (p)SIRT iterations	10
Smoothing	0.3
# DART iterations	45
Update probability	[0.05, 0.1, 0.2, 0.5]

Figure 3.4. The Barbapapa plexiglass phantom.

relaxation factor λ for each run of DART was estimated empirically as follows: At every missing wedge (10° stepsize), the projection data was reconstructed with 50 choices for λ. The relative Number of Misclassified Pixels (rNMP), i.e. the ratio between the pixels belonging to the wrong class and the total nonzero pixels, was calculated for each λ. The best λ in terms of the rNMP was kept for each choice of p and TP-DART, which yields a table for interpolation of λ for the intermediate missing wedge α. Additionally, the data was reconstructed with TP-DART, where

$$\lambda = \beta \frac{\text{Number of free pixels}}{\text{Total number of pixels}} \tag{3.6}$$

is the relaxation factor and β controls the ratio between system size and relaxation. We hypothesise that since TP-DART iteratively lowers the system size, scaling the relaxation appropriately could lead to better results. An interpolation table was also created for β. The results of scaled relaxation for TP-DART were collected separately and denoted with TP-DART scaled.

3.4 Results

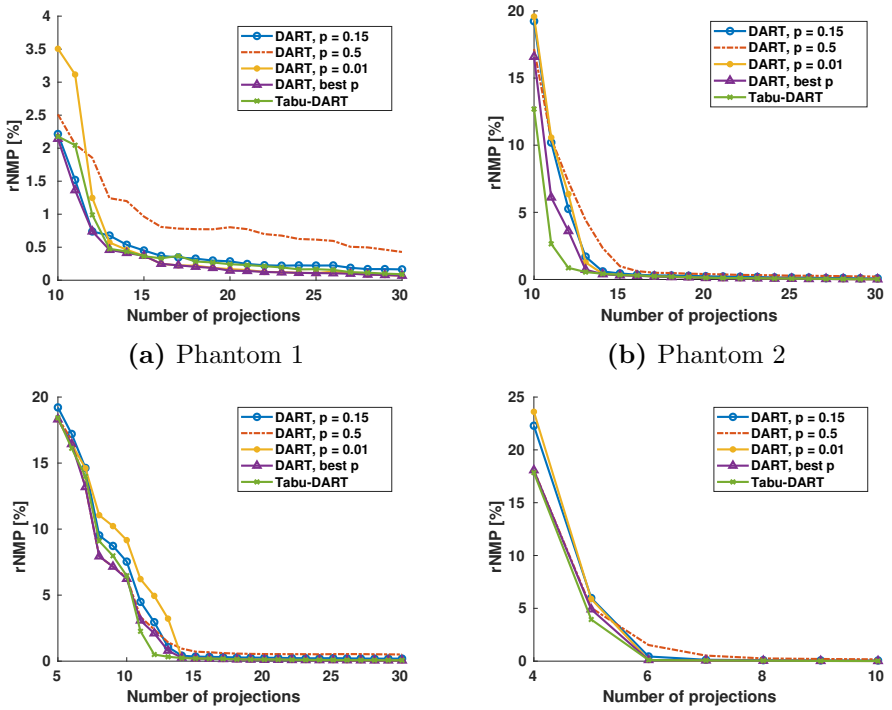
Two metrics were calculated to evaluate the performance of the algorithms in each experiment: the rNMP and a measure for the computational efficiency. The latter metric is expressed as either the total CPU time of the SIRT iterations inside one DART iteration, or as the size of the linear system. The system size is equal to the number of free pixels and expressed as a percentage.

3.4.1 Simulation results

First, DART with different values of p was compared to Tabu-DART in terms of rNMP, for both the few-view and the missing wedge case. This experiment has been repeated ten times with different seeds. Fig. 3.5 and Fig. 3.6 show the mean rNMP for each choice of parameter p and Tabu-DART for increasing number of projections and increasing missing wedge, respectively. For phantoms 2 and 3 in the few-view case, Tabu-DART performs noticeably better than the other three DART algorithms in terms of rNMP in the case with varying angles, when the number of projections is very limited. For the other two phantoms, Tabu-DART remains competitive towards DART with the best performing value of p.

The missing wedge experiment (Fig. 3.6) yields shows that Tabu-DART performs comparable to the best choice for p, especially when the missing wedge is high. Three specific missing wedges (small, medium, and large) were selected for each phantom for an in-depth study, and for those, the experiment was repeated 50 times with different seeds for the random number generator. Fig. 3.7 shows the box plot data of the rNMP for DART and Tabu-DART for the small and medium wedge choices. A lower rNMP and lower variance is observed for Tabu-DART compared to DART. Both DART and Tabu-DART start from a different initial map, and this map is constant per algorithm in each of the 50 seeded repeats. Our approach consistently feeds back data and dynamically changes the set of pixels to be updated, while DART has no feedback loop. This leads to the higher variance on the rNMP for DART, as the free pixel selection is largely influenced by random chance. The difference in visual quality between DART with the worst and best performing value for p is shown in Fig. 3.8. The contrast between the best and worst choice is evident, which emphasises the importance of a good parameter choice for these experiments. Tabu-DART, on the other hand, yields a superior visual quality without relying on the p parameter.

Fig. 3.9 shows the average CPU time of 10 SIRT iterations in seconds for varying angles. As very little of the background is selected in the case



(a) Phantom 1 **(b)** Phantom 2

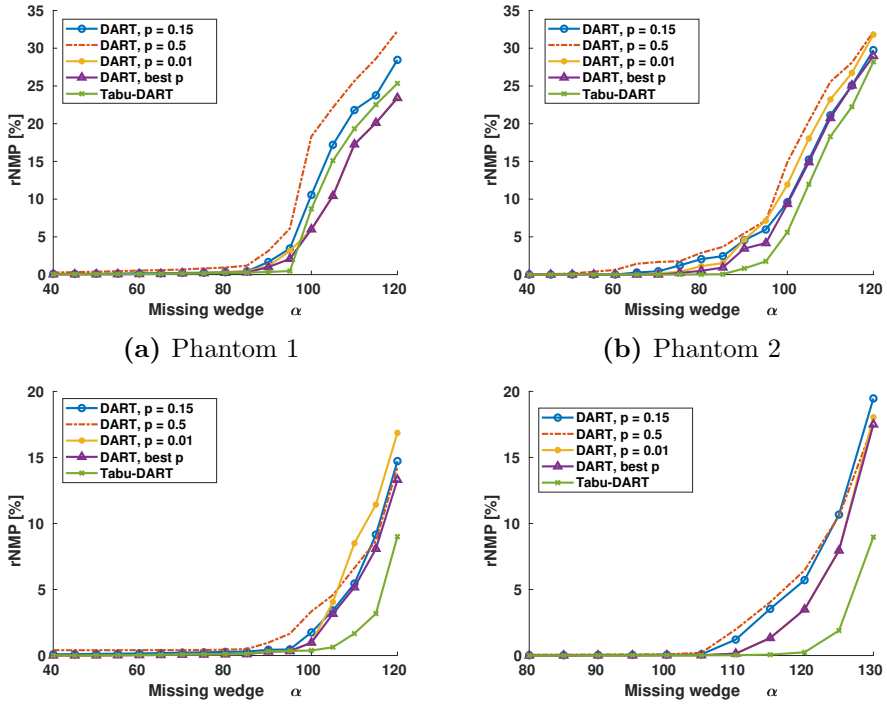
(c) Phantom 3 **(d)** Phantom 4

Figure 3.5. The rNMP for Tabu-DART and DART with four different choices of p as a function of the number of projections.

of $p = 0.01$, it is not surprising that this choice for p leads to the fastest algorithm. However, our approach is comparable in speed. This is due to the feedback procedure, which iteratively removes already stable regions from the reconstruction. Hence, the average size of the linear system decreases, yielding the observed low computation time together with a high reconstruction quality in terms of the rNMP. We conclude that our approach outperforms the DART algorithm both in rNMP and visual quality for different types of noiseless scenarios. A final remark is that earlier simulation studies [47, 134] show that lower values of p lead to a lower rNMP. In practice, once noisy data is involved, the higher values of p tend to yield a lower rNMP. We show evidence for this claim in the next experiment.

3.4.2 Barbapapa plexiglass phantom

Fig. 3.10a shows the rNMP of all methods tested for the different choices of λ. Fig. 3.10b shows the rNMP of TP-DART for varying scaling factor



(a) Phantom 1 **(b)** Phantom 2

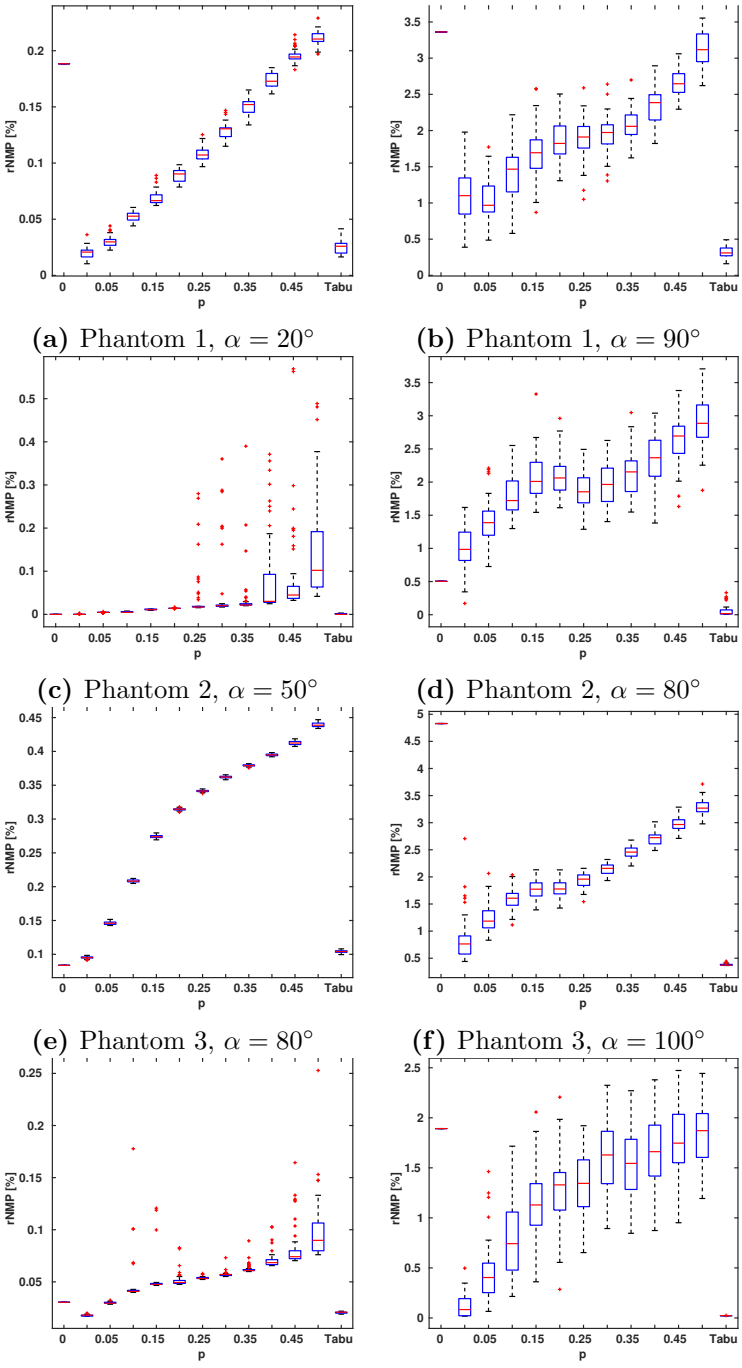
(c) Phantom 3 **(d)** Phantom 4

Figure 3.6. The rNMP for Tabu-DART and DART with four different choices of p for increasing missing wedge.

Table 3.3. The relaxation values λ and scaling factors β which comprise the interpolation table for the Barbapapa plexiglass reconstruction.

Missing wedge α	10	20	30	40	50	60
λ for $p = 0.05$	0.1	0.1	0.12	0.12	0.92	1.0
λ for $p = 0.1$	0.16	0.18	0.2	0.2	0.76	1.0
λ for $p = 0.2$	0.24	0.34	0.32	0.3	0.94	1.0
λ for $p = 0.5$	1.0	0.96	0.66	1.0	1.0	1.0
λ for TP-DART	0.1	0.14	0.16	0.18	0.66	0.96
β for TP-DART	2.0	2.5	2.5	2.5	5.0	10

β. It can be observed that DART with lower p-values has a better defined minimum compared to high values of p. The same occurs for TP-DART and TP-DART scaled. The common trait that low p-values and TP-DART share is the lower number of freed pixels. Hence, we reason that this sensitivity to the relaxation factor is related to the system size. A lower system size implies an increased sensitivity of each pixel to noisy



(a) Phantom 1, $\alpha = 20°$ **(b)** Phantom 1, $\alpha = 90°$

(c) Phantom 2, $\alpha = 50°$ **(d)** Phantom 2, $\alpha = 80°$

(e) Phantom 3, $\alpha = 80°$ **(f)** Phantom 3, $\alpha = 100°$

(g) Phantom 4, $\alpha = 100°$ **(h)** Phantom 4, $\alpha = 110°$

Figure 3.7. Box plots of the rNMP for different missing wedge α.

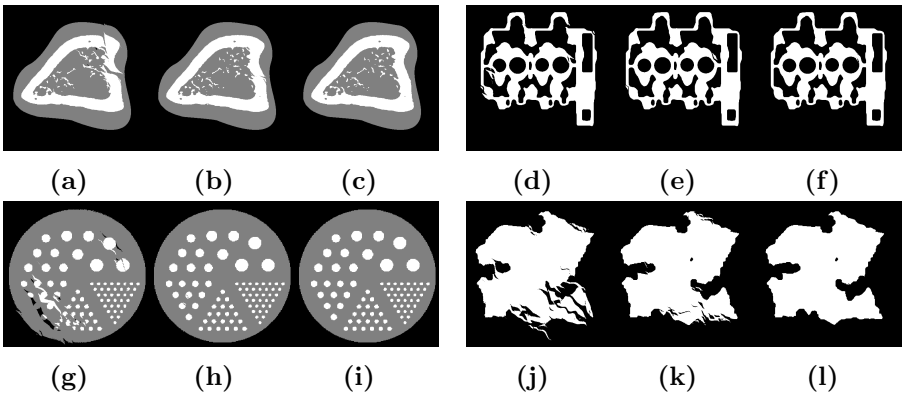
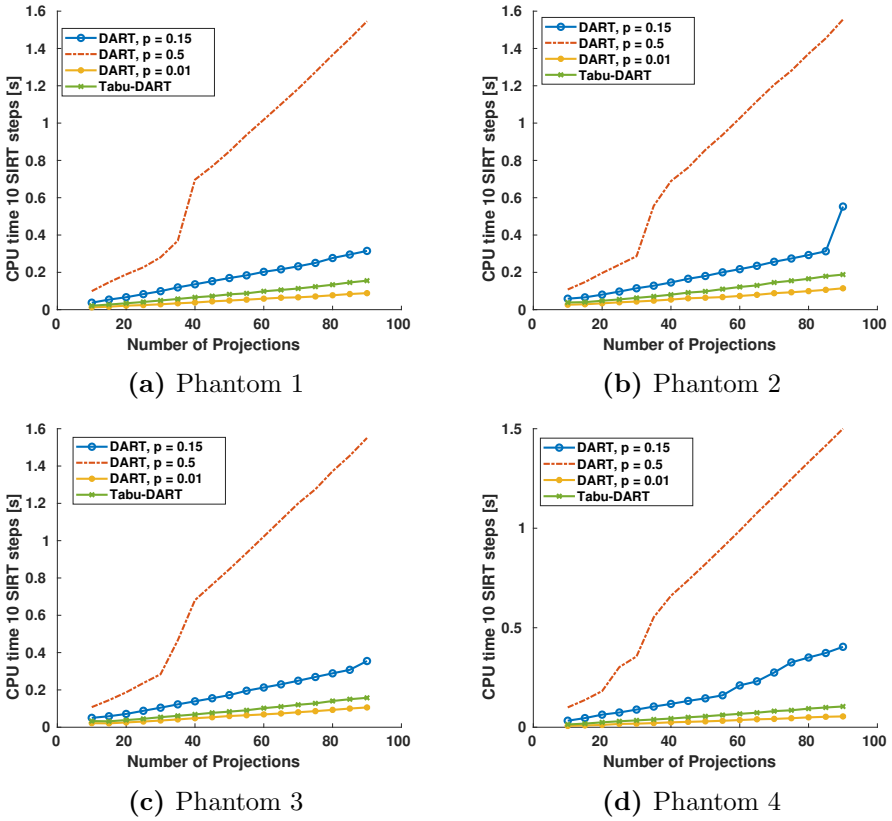


Figure 3.8. Missing wedge reconstructions of the simulation phantoms 1 (a-c), 2 (d-f), 3 (g-i), and 4 (j-l) for the worst and best performing p (left and center) compared to the reconstruction made with Tabu-DART (right). The missing wedge corresponds to a choice shown in Fig. 3.7: for phantom 1-4, α is $90°, 80°, 100°$ and $110°$, respectively.

data, due to increased convergence speed. Relaxation is necessary to counteract semi-converging behaviour. However, over-relaxation lowers convergence speed. The higher choices of p are innately more resistant to semi-convergent behaviour, and hence the impact of relaxation is lower since their convergence rate is slower. When additional projection data is removed, the reconstruction error increases due to lack of data instead of noise. Therefore, less relaxation is necessary, which results in a higher choice for λ in (1.21). The entries in Table 3.3 support this since the best performing values for λ and β increase as the missing wedge increases. This also means that smaller choices of p benefit more from relaxation. The final argument is that the best λ for poly-DART with $p = 0.5$ is almost exclusively $\lambda = 1.0$. In summary: The lower the choice of p for poly-DART, the more important the selection of the correct relaxation factor becomes. Furthermore, optimal λ selection for TP-DART is similar to the optimal choice for poly-DART with a small p.

The reason for only introducing a scaling factor β for the relaxation in TP-DART is because poly-DART relies on the same update rules as DART, which on average frees $100p$ percent of the pixels plus the boundary. The change in system size for poly-DART iterations is negligible compared to TP-DART, which makes the scaled relaxation with a scaling factor identical to the relaxation with a different fixed λ.

Fig. 3.11a shows the rNMP of the reconstructed images for varying missing wedge. All methods have very similar rNMP when the missing wedge is low, which was also the case for the simulation experiments. The



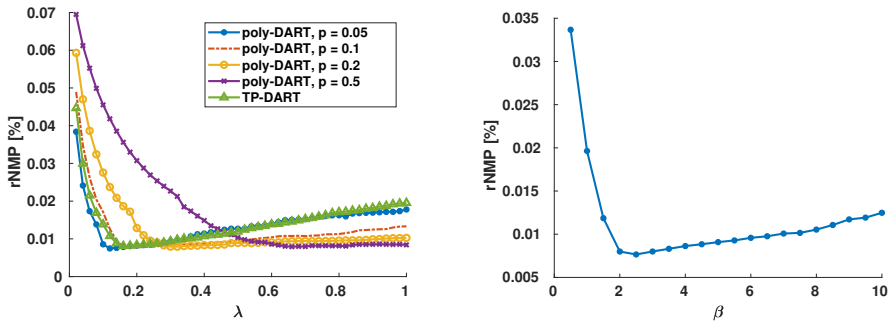
(a) Phantom 1 **(b)** Phantom 2

(c) Phantom 3 **(d)** Phantom 4

Figure 3.9. The average CPU time for 10 SIRT iterations when varying the angles for each phantom. We observe that Tabu-DART has a low average CPU time due to the probability map feedback procedure.

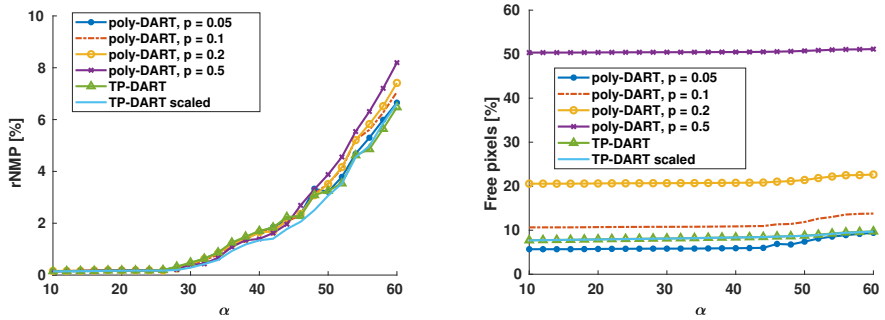
choice of p has negligible effect if there is sufficient data to reconstruct the object. For large values of α, TP-DART shows a consistently lower rNMP than poly-DART. Overall, TP-DART scaled has the lowest rNMP for each value of α with a system size of the same order as poly-DART with $p = 0.05$ (Fig. 3.11b).

Despite real-world projection data, the lower choices for p yield a lower rNMP for this object. For the Barbapapa phantom, the optimal relaxation parameter λ, with respect to the rNMP, was chosen, (cfr. Table 3.3). This implies that there exists a cut-off where relaxation stops benefiting the DART algorithm. Our experiments provide evidence that this cut-off depends on both the amount of projection data and the choice p. Table 3.3 shows a large jump for λ once $\alpha \geq 50°$. It is also from this point on that $p = 0.05$ outperforms higher choices of p.



(a) (b)

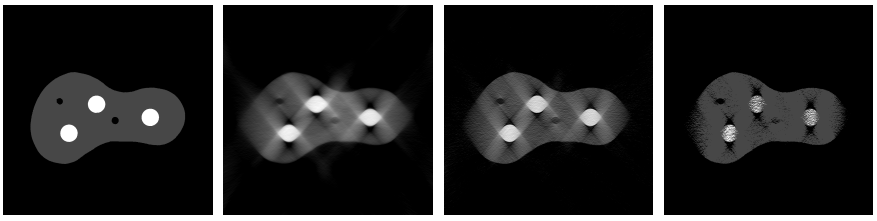
Figure 3.10. (a) The rNMP for varying relaxation λ for the Barbapapa plexiglass phantom. (b) The rNMP for Tabu-DART with changing relaxation with scaling factor β.



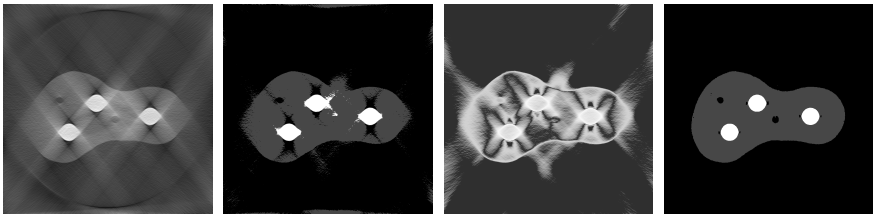
(a) (b)

Figure 3.11. The rNMP (a) and the percentage of free pixels (b) as a function of missing wedge α for the Barbapapa plexiglass phantom.

Two conclusions can be drawn from the results. The first is that estimating λ based solely on system size will yield poor results if the available projection data is insufficient. Second, relaxation based on system size with a scaling factor β dependent on the amount of data available is indispensable towards the proper functioning of TP-DART for experimental projection data. Even in the case of polychromatic data, our approach based on Tabu-search showed favourable results with respect to DART. The reconstructed image for a missing wedge of 40° is shown in Fig. 3.12. Due to the large missing wedge, the pSIRT and SIRT reconstructions show large streak artefacts, which drastically influence the quality of the segmentation (Fig. 3.12b up to Fig. 3.12f). The initial probability map used in TP-DART captures these artefacts



(a) Reference reconstruction **(b)** Initial pSIRT **(c)** Final pSIRT **(d)** Final pSIRT segmented



(e) Final SIRT **(f)** Final SIRT segmented **(g)** Initial probability map **(h)** TP-DART scaled reconstruction

Figure 3.12. pSIRT, SIRT and TP-DART reconstructions for the missing wedge experiment with $\alpha = 40°$. TP-DART yields highly increased visual quality. The final number of (p)SIRT iterations for each method was 500.

(Fig. 3.12g), but the final TP-DART output contains no missing wedge artefacts. In fact, the TP-DART reconstruction is very similar to a reference reconstruction created with pSIRT for the entire dataset of 2400 projections (Fig. 3.12a). This implies that the feedback structure of the algorithm is able to correct errors created during the initial SIRT reconstruction.

When considering visual quality of the reconstructions, no clear best method emerged.

3.5 Discussion and outlook

In summary, the proposed probability map plays the role of frequency-based memory and aids in choosing more optimal regions for further reconstruction. It is able to retain which regions are already stable and uses this information to completely remove them from the reconstruction problem, increasing the efficiency and speed of the DART iterations over time. The initialisation procedure suggested above eliminates the need for the update probability parameter p.

Table 3.4. FLOP counts for operations occurring during a SIRT iteration.

Operation	FLOP count
$\boldsymbol{x} - \boldsymbol{y}$	n
$\boldsymbol{x}^\top \boldsymbol{y}$	$2n - 1$
$\boldsymbol{A}\boldsymbol{x}$ (sparse)	$\mathcal{O}(sn)$
$\boldsymbol{d}\boldsymbol{x}$ (diagonal D)	n

3.5.1 Complexity analysis based on floating point operations

The experiments show that our update strategy reduces the system size on which the SIRT algorithm is run. A good measure for iterative algorithms is the number of FLOating Point operations (FLOPs) needed to perform an iteration. In this section, each addition, subtraction, multiplication, and division is counted as one FLOP. A theoretical speedup can be measured by counting FLOPs for the SIRT algorithm. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be s-sparse, i.e. $|\boldsymbol{A}| = s$. Let $\boldsymbol{d} \in \mathbb{R}^{n \times n}$ be any diagonal matrix. The FLOP counts for different matrix/vector operations present in SIRT are summarised in Table 3.4 [144]. In practice, the entries of $\boldsymbol{A}$ are calculated on the fly, resulting in additional overhead depending on the number of pixels in the image and the number of non-zero entries in $\boldsymbol{A}$. The complexity of multiplying a vector by $\boldsymbol{A}$ is hence $\mathcal{O}(sn)$. The only operation in SIRT which is not yet accounted for, is the creation of the $\boldsymbol{R}, \boldsymbol{C}$ matrices, which are diagonal matrices that have the inverses of the row and column sums on their diagonal, respectively. To create a matrix of the form:

$$\boldsymbol{R} = \begin{bmatrix} \frac{1}{\sum_{j=1}^{m} a_{1j}} & & & \\ & \frac{1}{\sum_{j=1}^{m} a_{2j}} & & \\ & & \ddots & \\ & & & \frac{1}{\sum_{j=1}^{m} a_{mj}}, \end{bmatrix} \tag{3.7}$$

the entry $\frac{1}{\sum_{j=1}^{m} a_{ij}}$ requires $n - 1$ additions and 1 division, which totals n FLOPs when counting divisions as one operation. This is repeated for each row of the m-by-n matrix which means m times n FLOPs for a total of mn. The SIRT update step can be decomposed into a sequence

of matrix-vector multiplications with costs:

$$\begin{aligned} \boldsymbol{x}_1 &= \boldsymbol{A}\boldsymbol{x}^{(k)} && \text{cost: } \mathcal{O}(sn) \\ \boldsymbol{x}_2 &= \boldsymbol{p} - \boldsymbol{x}_1 && \text{cost: } m \\ \boldsymbol{x}_3 &= \boldsymbol{R}\boldsymbol{x}_2 && \text{cost: } m \\ \boldsymbol{x}_4 &= \boldsymbol{A}^\top \boldsymbol{x}_3 && \text{cost: } \mathcal{O}(sn) \\ \boldsymbol{x}_5 &= \boldsymbol{C}\boldsymbol{x}_4 && \text{cost: } n \\ \boldsymbol{x}_6 &= \boldsymbol{x}^{(k+1)} + \boldsymbol{x}_5 && \text{cost: } n \end{aligned} \tag{3.8}$$

Since the creation of matrices $\boldsymbol{C}$ and $\boldsymbol{R}$ only happens once per sequence of SIRT updates, the cost is omitted in further complexity calculation. From (3.8) it is trivial to find that the total complexity in terms of FLOPs is of $\mathcal{O}(sn+m)$. However, the s non-zero entries of $\boldsymbol{A}$ are spread equally across each column since each ray i that passes through a pixel j yields a non-zero value a_{ij}. If instead $n-k$ pixels are removed from reconstruction in the masking step, a total of $n-k$ columns is removed from $\boldsymbol{A}$. This leads to a linear decrease in the number of remaining entries $s_k < s$ and hence the new complexity becomes $\mathcal{O}(s_k k + m)$. A study on DART performed earlier [101] pointed out that if enough pixels are set fixed, certain rays only pass through vacuum and fixed pixels. These zero rays lead to zero rows in the matrix $\boldsymbol{A}$. Hence, the number of non-zero detector readings is lowered to a value $m_k < m$. The final complexity of masked SIRT becomes $\mathcal{O}(s_k k + m_k)$ which is certainly higher than a linear reduction considering $m \ll n$ in typical discrete tomography applications.

3.5.2 Memory requirements

The gain in computational efficiency comes at the price of storage memory. To run the Tabu-DART algorithm we presented, two additional image-sized matrices need to be stored. The first one is for the probability map. This cost cannot be avoided, since the entire purpose of the map is to serve as a memory for the algorithm. The second matrix is required to store the segmentation from the previous DART iteration. This allows to track changes in grey values between two iterations. Since grey value classes of the previous iteration can be represented by integer numbers representing the class, the memory demand can be reduced by working with short bit integers at the cost of extra processing.

3.5.3 Outlook

The criteria for dynamic update rules are not limited to image stability. Our approach can utilise metrics such as the Reconstructed Residual Error [145], image stability [134], or image uncertainty [141]. Algorithms such as MDART [102] and ADART [101] can be easily represented with a probability map, illustrating that our proposed technique is, in fact, a generalisation of the original DART approach to a dynamic framework. The development of additional dynamic update strategies based on image uncertainty is a point of reference for future work.

3.6 Conclusion

A generic framework based on Tabu-search was proposed to aid divide-and-conquer strategies for algebraic discrete tomography methods. Our framework relies on a probability map that functions as a memory structure, which can be adapted through feedback obtained during the run-time of the algorithm. This concept was applied to DART, for which we introduced new dynamic update rules and a stronger initialisation phase based on local image uncertainty. The method was subjected to a simulation study using different discrete phantoms and an experimental polychromatic dataset of a plexiglass block with aluminium rods. The experiments provided evidence of increased visual imaging quality as well as lower rNMP rates and lower average computation time compared to the original DART algorithm. The generic nature of our approach makes it ideal to be combined with other discrete algebraic methods that rely on divide-and-conquer strategies.

CHAPTER 4

Quasi-Newton methods for polychromatic X-ray CT

This chapter has been published as:
Six, N., Renders, J., Sijbers, J. and De Beenhouwer, J., 2021. Gauss-Newton-Krylov for reconstruction of polychromatic X-ray CT images. IEEE Transactions on Computational Imaging, 7, pp.1304-1313.

Parts of the published method section have been omitted because the concepts were already introduced in Chapter 1.

4.1 Introduction

Computed tomography (CT) is a non-destructive imaging technique to reconstruct the interior of an object from multiple X-ray radiographs. Classical reconstruction methods, such as the filtered back projection (FBP) and the simultaneous iterative reconstruction technique (SIRT) [2], rely on a linearised acquisition model, assuming that the log-corrected normalised projection is the sum of attenuation values along a ray. This is an adequate model for monochromatic X-ray beams, for example those employed at synchrotron facilities. However, most lab X-ray sources are polychromatic with a broad spectrum. As the X-ray beam travels through the sample, the low-energy photons are attenuated more easily than those with high energy, causing the effective energy of the beam to shift upward as it passes through more material, which is known as beam hardening. As a result, the absorption along a ray from a polychromatic source is a non-linear function of the sample thickness. Reconstructing such a dataset with a method based on a monochromatic inversion formula will lead to artefacts in the reconstruction. These artefacts are usually in the form of cupping, when attenuation values toward the middle of an object are undervalued and the values at the edges are overvalued, as well as

dark streaks between different regions of higher attenuating materials separated by lower attenuating materials [1].

In an attempt to remove or reduce beam hardening artefacts, numerous techniques have been developed, both in hardware and software. Common practice on the hardware side is placing thin metal filters in front of the source to pre-harden the beam [31]. However, the pre-hardened spectrum is still polychromatic. Moreover, filtering decreases the number of photons available for imaging, and hence lowers the signal-to-noise ratio of the reconstructed image. Another option is to use a dual-energy X-ray setup, where an object is scanned using two different X-ray spectra, with distinct effective energies[146]. The attenuation coefficients can then be reconstructed with specialised algorithms. These algorithms are application specific, but have been successfully used for bone mineral density measurements [147], reduction of polychromatic artefacts from dense inclusions in non-organic material [148] and to more accurately measure body composition [149].

Algorithmically, beam hardening artefacts can be reduced by linearising the data, i.e. by mapping the polychromatic projections to monochromatic projections [33, 32]. Alternatively, the inherent polychromaticity of the beam can be modelled in the reconstruction algorithm itself. One class of algorithms that follow this approach is the statistical beam hardening reduction methods [150, 151, 27], where polychromaticity is modelled and the reconstruction problem is solved through a maximum likelihood algorithm. A simpler algorithm was described in [34], where the authors propose a polychromatic variant of SIRT. The proposed polychromatic model in [34] does not require segmentation, unless a contrast agent is used, and models each voxel as a mixture of two neighbouring basis materials. The reconstruction problem is solved heuristically by substituting polychromatic forward model for the linear projection in SIRT but leaves the other parts of the update step unchanged. The forward model has thus been changed, but the backprojection, which returns these errors to reconstruction space, is still the same linear model.

We propose using a quasi-Newton method to solve the polychromatic reconstruction problem arising from the model proposed in [34]. We have considered Gauss-Newton-Krylov (GNK)[54], limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)[152, 153] and Gradient Descent with Barzilai-Borwein steps (GD-BB) [52]. The backprojection in SIRT can be viewed as the transpose of the Jacobian of the system of equations, through which SIRT is seen to be a gradient descent method with a particular choice of step size. The proposed method uses the transposed Jacobian of the polychromatic model, which can be interpreted as the

polychromatic backprojection. Quasi-Newton methods are used to minimise an objective function consisting of the polychromatic projection error term and optional regularisation terms. The proof of principle of this idea was recently demonstrated on simulated data at ICIP 2020 [154]. In this work, we give a broader overview of the theoretical aspects, as well as more reconstruction experiments, including reconstructions from an experimental X-ray CT dataset. To highlight the practical applicability of the method, both the simulated and the experimental dataset have noisy projections, and the spectral information of the source was estimated using projections of a step wedge phantom. Another approach, only considering L-BFGS [152, 153], was proposed in [155]. In [155] a different forward model for the polychromatic image formation was employed, the log-corrected version of the model used in IMPACT [150], which uses less exact material information than the one proposed in [34]. Furthermore, only perfect synthetic data was considered in [155], without noise, and with perfect spectral and material information. We will present results of our method on realistically simulated Monte Carlo data as well as real data.

4.2 Method

4.2.1 Jacobian of the polychromatic model

We wish to minimise the polychromatic projection error term, with optional regularisation term $R(\boldsymbol{x})$:

$$\mathrm{Obj}(\boldsymbol{x}) = \frac{1}{2}\left\|\mathrm{polyProj}(\boldsymbol{x}) - \boldsymbol{b}\right\|_2^2 + \lambda R(\boldsymbol{x}) \tag{4.1}$$

with $\boldsymbol{b} \in \mathbb{R}^{n_{\mathrm{proj}}}$ the log-corrected measured polychromatic projection data, polyProj the projection operator defined in (1.19), and $\lambda \in \mathbb{R}$ a weighting term. To calculate the gradient of (4.1) or the Hessian approximation of (1.35), the Jacobian matrix of $\mathrm{polyProj}(\boldsymbol{x})$ is needed. In this section, this Jacobian will be described. All parts of the polychromatic forward model in (1.17) are differentiable, except for the triangular functions $\boldsymbol{M}_m(\boldsymbol{x})$. We can create differentiable variants of these functions by convolving them with a continuous function. We chose to use a triangular function T_ε defined as follows:

$$T_\varepsilon(x) = \begin{cases} 0 & \text{if } x \notin [-\varepsilon, \varepsilon] \\ \frac{x+\varepsilon}{\varepsilon} & \text{if } x \in [-\varepsilon, 0] \\ \frac{\varepsilon-x}{\varepsilon} & \text{if } x \in [0, \varepsilon] \end{cases}. \tag{4.2}$$

The convolutions $T_\varepsilon * \boldsymbol{M}_m$, which we will call $\boldsymbol{TM}_m^\varepsilon$, are differentiable with derivative

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\boldsymbol{TM}_m^\varepsilon\right) = \left(\frac{\mathrm{d}}{\mathrm{d}x}T_\varepsilon\right) * \boldsymbol{M}_m. \tag{4.3}$$

It is furthermore trivial that $\lim_{\varepsilon\to 0} \boldsymbol{TM}_m^\varepsilon = \boldsymbol{M}_m$. The $\boldsymbol{TM}_m^\varepsilon$ functions will be substituted for the $\boldsymbol{M}_m$ in the objective function (4.1). The parameter ε was empirically set at an optimal value of 10^{-4} for all experiments in this paper. With this substitution, one arrives at the following gradient of the objective function:

$$\boldsymbol{\nabla}\,\mathrm{Obj}(\boldsymbol{x}) = \boldsymbol{J}_{\mathrm{poly}}^\top(\boldsymbol{x})\left(\mathrm{polyProj}(\boldsymbol{x}) - \boldsymbol{b}\right) + \lambda\boldsymbol{\nabla}R(\boldsymbol{x}), \tag{4.4}$$

with $\boldsymbol{J}_{\mathrm{poly}}(\boldsymbol{x})$ the Jacobian matrix of the polyProj function in $\boldsymbol{x}$. This Jacobian matrix can be seen to be equal to the following:

$$\boldsymbol{J}_{\mathrm{poly}}(\boldsymbol{x}) = \boldsymbol{D}_3 \sum_{e=1}^{n_\mathrm{e}} I(e)\boldsymbol{D}_1 \sum_{m=1}^{n_\mathrm{m}} \boldsymbol{A}\boldsymbol{D}_2\mu_m(e) \tag{4.5}$$

where

$$\begin{aligned}
\boldsymbol{D}_1 &= \mathrm{diag}\left(\exp\left(-\sum_{m=1}^{n_\mathrm{m}} \boldsymbol{A}\boldsymbol{TM}_m^\varepsilon(\boldsymbol{x})\mu_m(e)\right)\right)\\
\boldsymbol{D}_2 &= \mathrm{diag}\left(\boldsymbol{TM}_m^{\varepsilon\,\prime}(\boldsymbol{x})\right),\\
\boldsymbol{D}_3 &= \mathrm{diag}\left(\exp(-\mathrm{polyProj}(\boldsymbol{x}))^{-1}\right)
\end{aligned}$$

with diag operating on k-dimensional vectors and turning them into $k\times k$-diagonal matrices. Thus $\boldsymbol{D}_1, \boldsymbol{D}_3 \in \mathbb{R}^{n_\mathrm{proj}\times n_\mathrm{proj}}$ and $\boldsymbol{D}_2 \in \mathbb{R}^{n_\mathrm{vol}\times n_\mathrm{vol}}$. As mentioned earlier, for implementation only a method handling multiplication with a vector needs to be made. In such an implementation, multiplication with $\boldsymbol{D}_1, \boldsymbol{D}_2$ or $\boldsymbol{D}_3$ is done as a pointwise multiplication of vectors instead. Multiplication with the Jacobian or the transposed Jacobian requires, respectively, n_m forward or backward projections. This gives each application with the Jacobian a complexity of $\mathcal{O}(n_\mathrm{m} * \max(n_\mathrm{vol}, n_\mathrm{proj}))$.

4.2.2 Regularization

Any regularisation term $R(\boldsymbol{x})$ can be added to the model, as long as its gradient and Hessian can be calculated. In this paper, we have used a differentiable version of anisotropic total variation (TV) regularisation as an example in some of the experiments. The anisotropic TV of a 2D image $\boldsymbol{x}$ is defined as:

$$\begin{aligned}
TV_a(\boldsymbol{x}) = \sum_i \sum_j &\left|\boldsymbol{x}(i+1,j) - \boldsymbol{x}(i,j)\right| \\
&+ \left|\boldsymbol{x}(i,j+1) - \boldsymbol{x}(i,j)\right|.
\end{aligned} \tag{4.6}$$

There exists a matrix $\boldsymbol{D}$ such that, if $\boldsymbol{x}$ is now considered in vectorized form, $TV_a(\boldsymbol{x}) = \sum_i (\sqrt{(\boldsymbol{D}\boldsymbol{x})^2})_i$. Anisotropic TV is not a differentiable function, so we will use the following differentiable approximation, for $\beta \in \mathbb{R}, \beta > 0$:

$$TV_a(\boldsymbol{x}) \approx \sum_i (\sqrt{(\boldsymbol{D}\boldsymbol{x})^2 + \beta})_i. \tag{4.7}$$

We will denote this approximation as $TV(\boldsymbol{x})$. The gradient and Hessian of this regularisation term are:

$$\begin{aligned} \boldsymbol{\nabla} TV(\boldsymbol{x}) &= \boldsymbol{D}^\top \left(\frac{\boldsymbol{D}\boldsymbol{x}}{TV(\boldsymbol{x})} \right) \\ \boldsymbol{H}_{TV}(\boldsymbol{x}) &= \boldsymbol{D}^\top \mathrm{diag} \left(\frac{\beta}{TV(\boldsymbol{x})^{\frac{3}{2}}} \right) \boldsymbol{D}. \end{aligned} \tag{4.8}$$

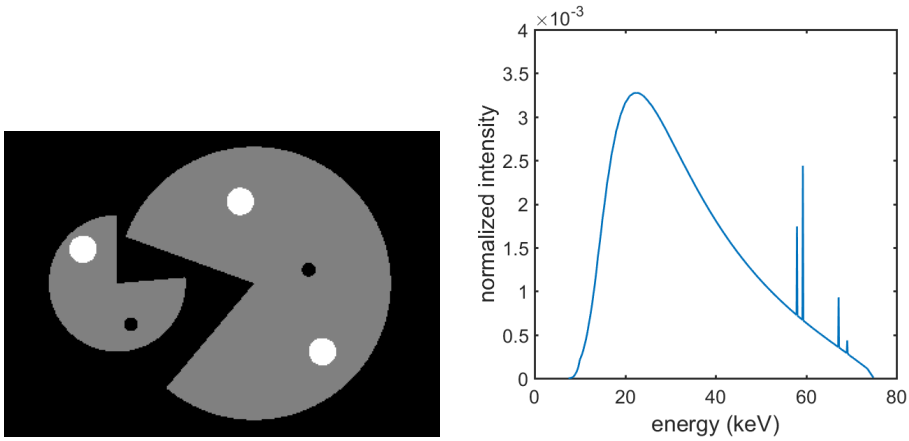
All methods were implemented in MATLAB, making use of the ASTRA toolbox to handle forward and backward projections on the GPU [142, 107]. The implementation of the methods, as well as the Monte Carlo dataset described in Section 4.3.1, are available on GitHub and archived in Zenodo[156].

4.3 Experiments

To validate and compare the different reconstruction approaches, a Monte Carlo simulation and a real data experiment were set up.

4.3.1 Monte Carlo data

A Monte Carlo simulation was performed with the Pie phantom shown in Fig. 4.1(a). The phantom consists of two materials, plexiglass and aluminium, suspended in air. The attenuation values at different energy levels were obtained from the National Institute of Standards and Technology [35]. To avoid the inverse crime, the phantom was analytically defined in the simulation framework, GATE [106]. Polychromatic fan-beam projections were simulated with GATE. A 75 kVp X-ray point source with a tungsten anode was simulated, of which the spectrum is shown in Fig. 4.1(b). Spekcalc [108] was used to generate the source spectrum. The projection angles, 300 in total, were chosen using golden angle sampling, which means that subsequent projections are $\frac{1+\sqrt{5}}{2}\pi$ radians apart. The detector pixel size was 375 μm and the scanner setup had a magnification factor of 4. The reconstructed pixel size was 93.75 μm. The detector was modelled as an idealised photon counting line detector



(a) Analytically defined 'Pie' phantom consisting of air (black), plexiglass (grey) and aluminum (white).

(b) Polychromatic spectrum of a 75 kVp X-ray source.

Figure 4.1. Ground truth data employed in GATE for the simulation experiments: the Pie phantom (a) and the X-ray spectrum (b).

with 400 pixels made of silicon, with a minimal activation energy of 5 keV. Charge sharing between the pixels was not simulated. As the projections were created using a Monte Carlo method, the generated data is Poisson distributed. For each projection 8×10^6 photons were simulated, which is on average 20k photons per pixel. Both absorption and scattering of the photons were simulated. To further simulate a real experiment, the spectrum $I_0(\epsilon)$ was estimated with two simulated step wedges, one of plexiglass and one of aluminium. Multiple projections were taken orthogonal to the steps and then averaged to reduce the effects of noise in the simulation measurements. A maximum likelihood expectation maximisation (MLEM) algorithm [157] was used to estimate the spectrum $I_0(\epsilon)$ on 70 energy bins. The four characteristic peaks of a tungsten anode source spectrum were added to the initial guess for the MLEM iterations, based on prior knowledge.

From the simulated polychromatic projection data, images were reconstructed with the following methods: pSIRT, gradient descent with Barzilai-Borwein steps (GD-BB), Gauss-Newton-Krylov (GNK), limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS), nonlinear conjugate gradients (NCG) and the polychromatic maximum likelihood maximisation technique IMPACT. All methods used the same four basis materials: vacuum, plexiglass, aluminium, and iron. The same polychromatic projection error term was used for GD-BB, GNK, L-BFGS, NCG,

and pSIRT. IMPACT employs a different polychromatic model and, as such, mimics a different objective function. Next, the effect of the chosen number of inner iterations for GNK is examined. Then, we investigated the effect of a different choice of basis materials. In this experiment, we used a basis of vacuum, water, and iron. Lastly, the methods directly compatible with regularisation, GD-BB, NCG, L-BFGS, and GNK, were compared when reconstructing the same data with TVmin regularisation. The regularisation term had empirically chosen parameters $\lambda = 1$ and $\beta = 10^{-7} \times \mu_{\max}(e_r)$, with $\mu_{\max}$ the attenuation coefficients of the most attenuating basis material. A reference energy of 55 keV was chosen as optimal for pSIRT in terms of convergence speed and reliability [34]. The same reference energy was used for the other techniques. Unless otherwise specified, L-BFGS was run storing the two last image and gradient updates, and the Gauss-Newton-Krylov methods used 5 Krylov iterations in each Newton iteration.

4.3.2 Experimental data

Following the Monte Carlo simulations, a similar test was performed on the central slice of an experimental cone beam dataset. The Barbapapa phantom [109], consists of plexiglass with inserted aluminium rods and is shown in Fig. 4.2. A phantom with the same materials as in the simulation study was used to more easily compare the simulation results to the experimental results. The phantom was scanned over 300 equiangular projections in the interval $[0, 2\pi)$. The detector rows have 968 pixels, each pixel has a size of 149.6 μm. A magnification of 4.28 was achieved, resulting in a voxel size of 34.95 μm. A tube voltage of 130 kVp was employed.

To estimate the spectrum, the same technique as in the simulation case was used. For this experiment, a PVC step wedge with 11 steps, ranging from approximately 1 mm to 18 mm in thickness, was scanned. Next, the same MLEM spectrum estimation method as in the Monte Carlo case was performed, with 130 keV as maximum energy constraint. From the central slice of the dataset, images were reconstructed with the same methods: pSIRT, GD-BB, GNK, L-BFGS, NCG and IMPACT. All methods used the same four basis materials: vacuum, plexiglass, aluminium, and iron.

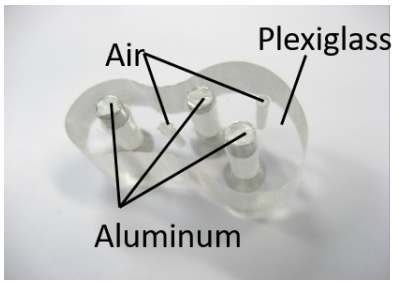
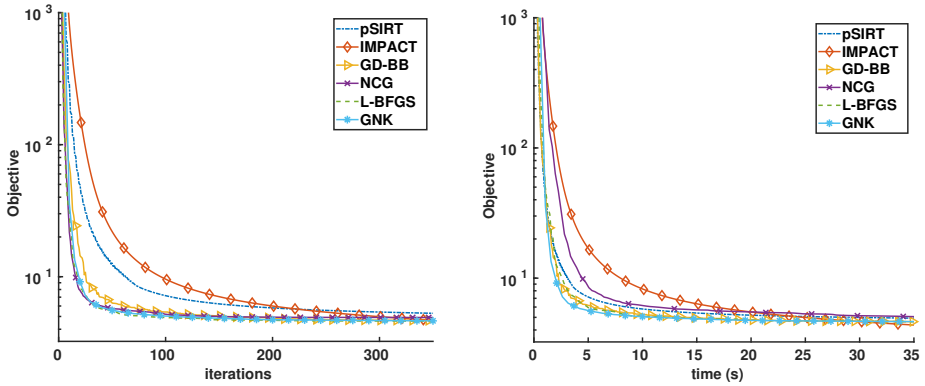


Figure 4.2. Barbapapa phantom used to generate the experimental dataset. The phantom is a smoothly shaped plexiglass object, with three inserted aluminum rods and two air columns.



(a) Objective function value against iterations.

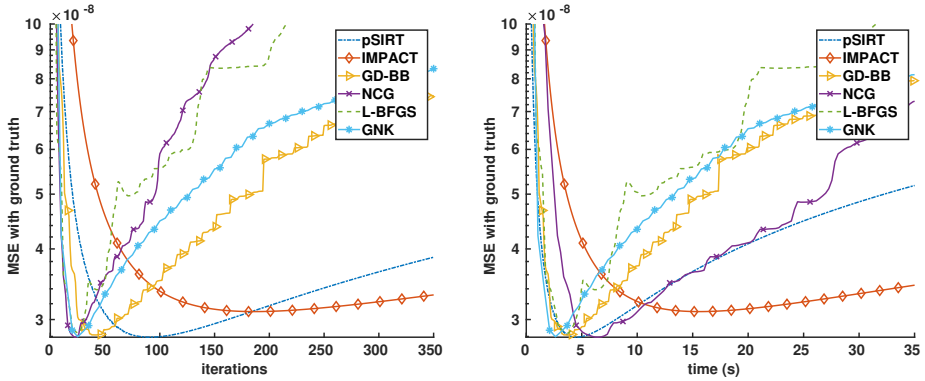
(b) Objective function value against computation time.

Figure 4.3. Plots of the objective function for the simulated data for all considered methods against number of (inner) iterations and computation time. Note that IMPACT uses a different model and thus minimises a different objective function.

4.4 Results

4.4.1 Monte Carlo data

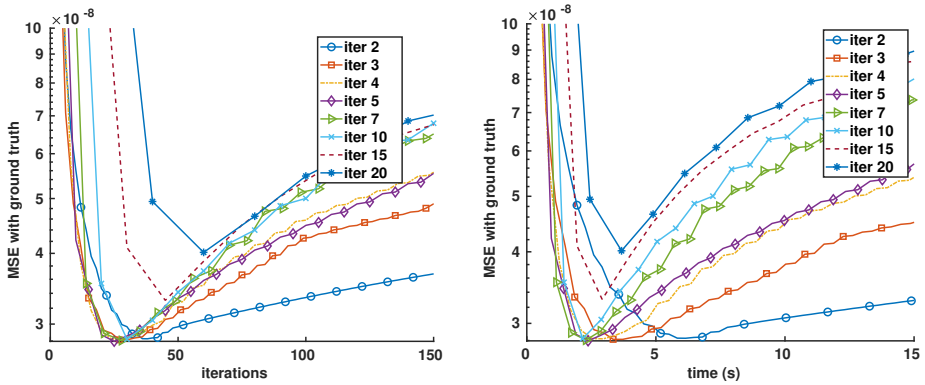
To quantitatively compare the results of the different methods, we computed the objective function at each iteration. Note that pSIRT does not have convergence theorems guaranteeing it minimises any error, so it does not have any true objective function, instead we used the first term of (4.1), which is the same objective function as the other non-regularized methods use, with the exception of IMPACT. IMPACT uses a different polychromatic model and as such has a different objective function, this



(a) MSE with the ground truth against iterations.

(b) MSE with the ground truth against computation time.

Figure 4.4. Plots of MSE with the ground truth for the simulated data for all considered methods against number of (inner) iterations and computation time.

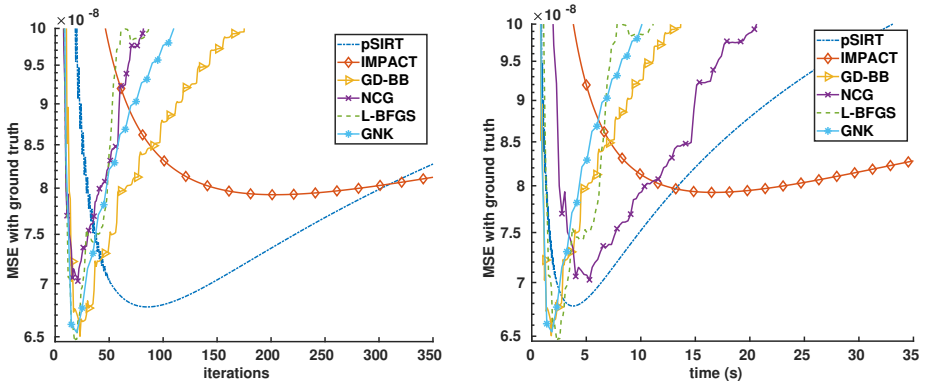


(a) MSE with the ground truth against iterations.

(b) MSE with the ground truth against computation time.

Figure 4.5. Plots of MSE with the ground truth for the simulated data for differing numbers of inner iterations against number of iterations and computation time.

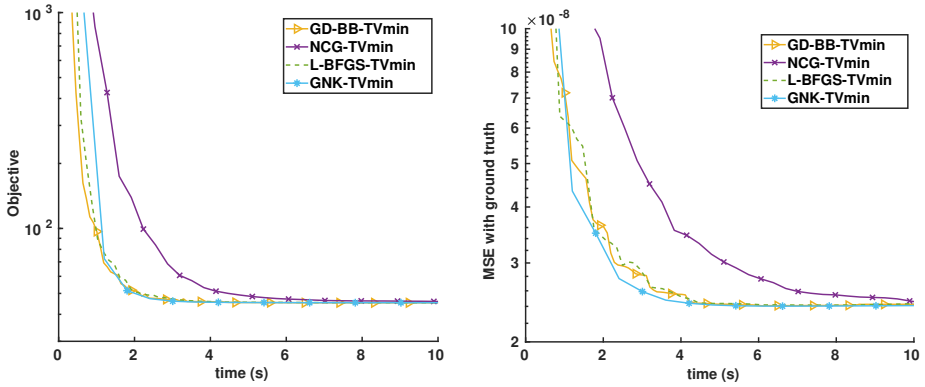
makes it difficult to compare IMPACT with the other methods with this metric. However, this is a comparable measure for all methods on how fast they converge to a solution of the system. The error was then plotted as a function of the number of iterations as well as computation time. For Gauss-Newton-Krylov, since it works on an inner-outer iteration scheme, the number of performed inner iterations was chosen for a fair comparison. Convergence in a more mathematical sense can be seen in the plots



(a) MSE with the ground truth against iterations.

(b) MSE with the ground truth against computation time.

Figure 4.6. Plots of MSE with the ground truth for the simulated data for an alternative material basis, consisting of vacuum, water and iron, against number of iterations and computation time.



(a) Objective function value against computation time.

(b) MSE with the ground truth against computation time.

Figure 4.7. Plots of different error measures for the simulated data for regularized methods.

comparing error to number of iterations, however, as computation time is a more important factor for real world applications, it is our main focus. The methods were all timed in MATLAB for the plots as functions of computation time. The full runtime for 1000 iterations was timed, and the errors, calculated after each iteration, were plotted assuming that each iteration takes an equal amount of time. Only the multiplication with the projection matrix was performed on GPU with the ASTRA toolbox [142, 107]. Furthermore, since this is synthetic data, the error with

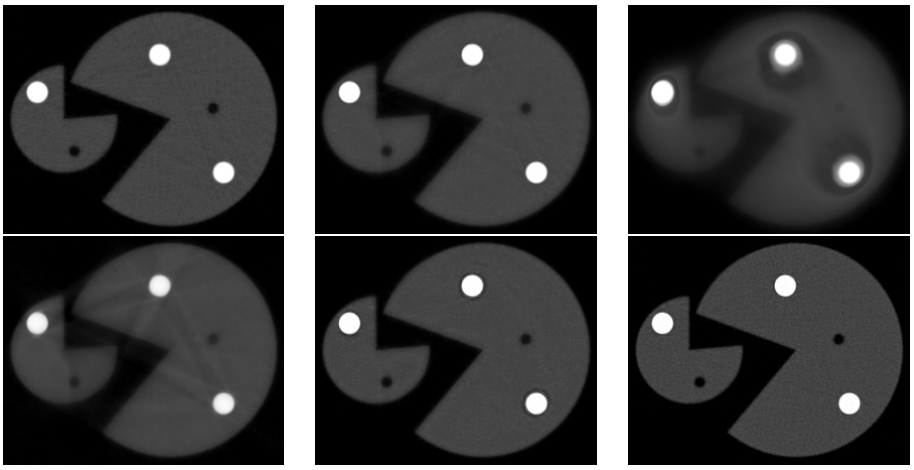


Figure 4.8. Comparison of the different reconstruction techniques on simulated data, after 3 seconds of reconstruction time. From top left to bottom right: pSIRT, GD-BB, NCG, IMPACT, L-BFGS, GNK.

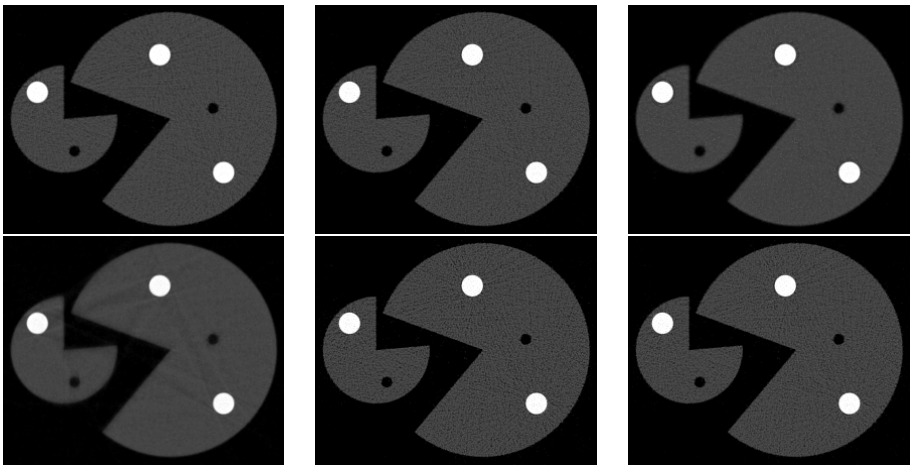


Figure 4.9. Comparison of the different reconstruction techniques on simulated data, after 10 seconds of reconstruction time. From top left to bottom right: pSIRT, GD-BB, NCG, IMPACT, L-BFGS, GNK.

respect to the ground truth was calculated. The ground truth was defined analytically in GATE, so it was recreated from the analytical definitions on the same 400×400 grid as the reconstruction for these calculations. These error measures are shown in Fig. 4.3 and Fig. 4.4. Based on the plots against iteration in Fig. 4.3 it is clear that IMPACT converges the slowest, followed by pSIRT, GD-BB and finally NCG, L-BFGS and GNK all tied as fastest. However, in the plot against computation time

it becomes clear that on a time scale NCG performs poorly. This is easily explained, as it is the only method that requires strong Wolfe conditions for the line search. Furthermore, GD-BB performs comparable to L-BFGS on a time scale, while GNK slightly outperforms the other methods. Lastly, we note that IMPACT reaches the lowest objective from all methods, but it is a different objective function, which makes it difficult to compare to the other methods. In Fig. 4.4 the mean squared error against the ground truth, a measure for the reconstruction quality, is shown. All methods show semi-convergence, which is to be expected due to the presence of noise in the measurements. We also note that IMPACT has the highest minimum in MSE with the ground truth. All other reconstruction techniques reach a similar minimal error with the ground truth. On a time scale, this minimum is reached first by GNK, then GD-BB and L-BFGS, then by pSIRT and finally by NCG. The reconstructions are compared in Fig. 4.8 and Fig. 4.9, after letting each reconstruction technique run for 3 and 10 seconds, respectively. The reconstructions show very little beam hardening artefacts, as expected. Only IMPACT still has some darker streaks visible after 10 seconds. In Fig. 4.8 it is clear that GNK has reached a good reconstruction the fastest. The low time reconstruction of L-BFGS shows some peculiar dark ring artefacts around the metal cylinders, however, these artefacts are removed after a longer runtime. After 10 seconds, all methods reach similar reconstructions, with L-BFGS and GNK showing the strongest overfitting artefacts due to the noise in the projection data.

Next, the influence of the chosen number of inner iterations in GNK was examined. When few iterations of the Krylov method are run, the solution is too far removed from the Newton step, too many iterations and the resulting step can be overfitted. In Fig. 4.5 the ground truth MSE error is compared for different choices of number of inner iterations. All choices between 3 and 10 lead to a similar minimum in error. From comparison of the plots one sees that for these choices the minimum is achieved after a similar amount of inner iterations, however, on a time plot the choices drift more apart. This is because in GNK the line search step only happens once per outer iteration. We conclude that the choice of number of inner iterations is important, but the method is not overly sensitive to it.

Then, the influence of a different choice of basis materials was investigated. In the other experiments, we use a basis that is as close as possible to the actual materials in the phantom. In this experiment, we chose vacuum, water, and iron as basis. This means that the plastic and aluminium parts need to be estimated as a combination of water and iron. The ground truth MSE of the reconstruction with this alternative basis

is shown in Fig. 4.6. The plot shows that the minimally achieved error is higher with this new basis, which is to be expected as the real materials are now represented as a combination of water and iron. Interesting to note is that, in contrast to the results with the exact basis, NCG and pSIRT no longer reach the same minimum as the quasi-Newton methods. We conclude that the quasi-Newton methods still perform well when less prior material knowledge is included.

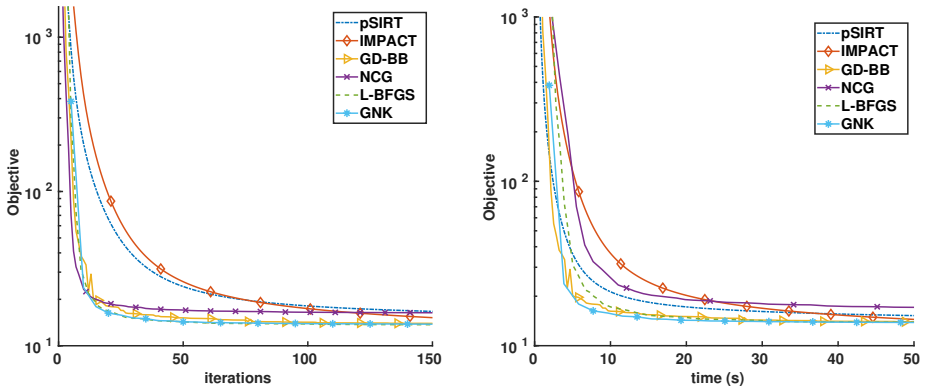
Lastly, the effect of regularisation with TV was investigated. The results are shown in Fig. 4.7. The regularised versions converge slower than the unregularized methods. This was to be expected, as they are also minimising a more complex objective function. In the comparison to the ground truth, the regularised methods reach a lower MSE than the unregularized methods and show vastly reduced semi-convergence. As with the unregularized reconstructions, NCG is a slower method, while GNK, GD-BB and L-BFGS are comparable. However, all methods converge to the same error.

4.4.2 Experimental data

A similar reconstruction experiment was run on the experimental dataset. However, since there is no ground truth image, we can only compare the objective and the reconstructed images. The changes in objective function of the different methods are shown in Fig. 4.10. From these plots, one can observe that again pSIRT, GD-BB, L-BFGS and GNK seem to converge to a comparable objective. NCG converges to a higher objective and IMPACT converges to a lower error, however, we note again that IMPACT uses a different objective function. In terms of convergence compared to iteration, very comparable results to the simulated data are seen. On a time scale, there is more difference between GD-BB, L-BFGS and GNK compared to the simulated data, with GNK and GD-BB clearly outperforming L-BFGS and GNK slightly outperforming GD-BB. Reconstructions after 10 seconds are shown in Fig. 4.11. The comparison between the methods at this point is more similar to the reconstructions of the simulated dataset after 3 seconds. This is not surprising, as the experimental dataset has over six times more unknowns than the simulated dataset.

4.5 Conclusion

Beam hardening artefacts reduce the quality of X-ray CT images. By including a polychromatic model in the reconstruction framework, these



(a) Objective function values against iterations.

(b) Objective function values against computation time.

Figure 4.10. Plots of the objective function values for the experimental data for all considered methods against number of (inner) iterations and computation time. Note that IMPACT uses a different model and thus minimizes a different objective function.

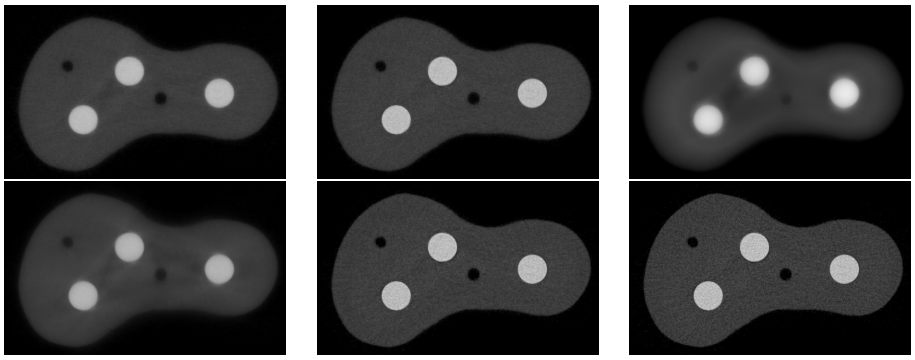


Figure 4.11. Comparison of the different reconstruction techniques on experimental data, after 10 seconds of reconstruction time. From top left to bottom right: pSIRT, GD-BB, NCG, IMPACT, L-BFGS, GNK.

effects can be mitigated. In this work quasi-Newton solvers are considered to minimise the polychromatic objective function arising from a recently proposed polychromatic model in [34]. This approach is mathematically justified and facilitates regularisation. To allow the use of quasi-Newton solvers, the Jacobian of the polychromatic model was first derived and presented. Multiple methods were compared: pSIRT, gradient descent with Barzilai-Borwein steps, limited-memory Broyden–Fletcher–Goldfarb–Shanno, Gauss-Newton-Krylov, non-linear conjugate gradients and a statistical reconstruction method IMPACT. Res-

ults on Monte-Carlo simulated data and on experimental data were shown. In all cases, GD-BB, L-BFGS and GNK were the best performing methods when taking reconstruction time into account, with GNK being the best. The results show that quasi-Newton methods in general and GNK in particular should be considered as a practical alternative to other beam hardening correction algorithms, algebraic or statistical.

CHAPTER 5

Joint multi-contrast CT for edge illumination X-ray imaging

This chapter has been published as:
Six, N., Renders, J., De Beenhouwer, J. and Sijbers, J., 2024. Joint multi-contrast CT for edge illumination X-ray phase contrast imaging using split Barzilai-Borwein steps. Optics Express, 32(2), pp.1135-1150.

5.1 Introduction

X-ray phase contrast imaging (XPCI) offers complementary contrasts to conventional X-ray absorption: differential phase or refraction contrast, which relates to the refraction of the X-ray beam, and a scatter or dark-field contrast, which relates to scattering caused by sub-voxel micro structures [158]. Compared to absorption contrast, these contrasts have demonstrated the ability to generate higher contrast in soft tissue and enable the visualization of sub-voxel structures, such as those found in composite materials [59]. The applications of XPCI have expanded beyond biomedical studies [60, 61, 62, 63] to non-destructive testing [64], materials science [65, 66], and security [67, 68].

Among the emerging XPCI techniques, edge illumination (EI) is particularly suitable for use with lab-based X-ray sources that have a large focal spot and a polychromatic spectrum [69, 59]. In this technique, the phase information can be retrieved by displacing two strongly absorbing masks relative to each other, with the object positioned in between. Conventionally, multiple images at a single view angle are acquired, one for each mask displacement. In every detector element, in every projection, this results in a curve, called the illumination curve. Subsequently, a

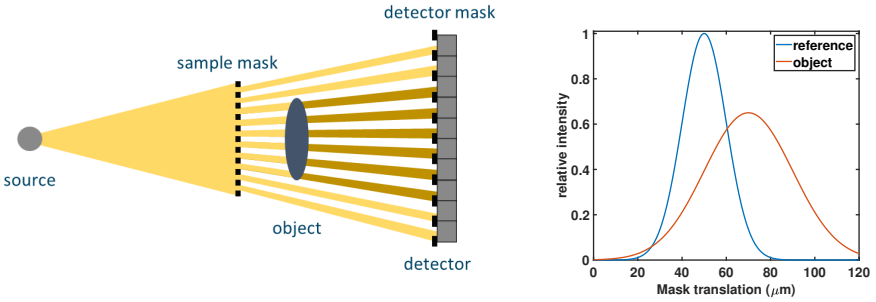
Gaussian curve is fitted to the illumination curves of each pixel in each projection, both before and after the introduction of a sample. From the two fitted Gaussians, three complementary contrasts are retrieved: absorption, refraction, and scatter. This step is called *phase retrieval*. Finally, for each of those contrasts, an image is individually reconstructed using filtered backprojection (FBP) or iterative reconstruction techniques.

Unfortunately, EI XPCI requires adequate sampling of the illumination curve at every detector pixel, resulting in long scan times, particularly in computed tomography (CT) imaging. Furthermore, in CT imaging, millions to billions of Gaussian curve fits need to be performed for the contrast retrieval, resulting in a lengthy preprocessing step. Additionally, there is no direct information sharing between the contrasts during reconstruction, limiting their potential synergies.

An alternative to the conventional two-step-approach, in which phase contrast images are reconstructed after phase retrieval, is a one-step-approach, in which phase contrast images are directly reconstructed from the measured X-ray data. In this approach, different contrasts are reconstructed simultaneously, without a separate fitting and retrieval step. In grating based Talbot-Lau interferometry (GBI), statistical approaches to one-step joint reconstruction have been proposed [159, 160], where a likelihood function is minimized. These approaches were also shown to help reduce scan times as greater freedom in the scanning geometry is allowed [78]. Unfortunately, these techniques cannot be directly applied to EI setups due to the different nature of the measurements between GBI and EI. Both the hardware components as well as the way X-ray refraction affects the measurements, differ. In GBI the resulting changes in interference patterns are compared, while in EI the deflection of the center of the beamlet is considered. Consequently, the forward models developed for GBI are not applicable to EI data.

Recently, an EI reconstruction method was presented in [77] that reconstructs refraction and absorption contrasts separately in a joint reconstruction using line searches. This method achieves a similar flexibility in acquisition setup for EI as reported in [78] for GBI. However, this method does not include scatter reconstructions, which can provide additional information on sample properties [81]. Moreover, no investigation into different suitable solvers for the objective function was performed in [77].

In this work, we propose a novel joint reconstruction method using a full EI forward model without retrieval, enabling the simultaneous reconstruction of all three contrasts. The proposed approach reduces reconstruction times, as the retrieval step is skipped and allows more flexible acquisition schemes. We demonstrate the effectiveness of the



(a) EI imaging set up **(b)** EI measurements

Figure 5.1. (a) overview of an EI setup and (b) flatfield and measured IC.

proposed method through reconstructions of both simulated and experimental datasets, showing results on single-shot datasets which cannot be reconstructed with the two-step method.

5.2 Method

5.2.1 Edge illumination CT forward model

In EI imaging, the refraction of the X-rays in the object is quantified by incorporating two gratings, called the sample and detector mask, into the traditional X-ray setup, as depicted in Fig. 5.1a. The X-ray beam is split into smaller beamlets by the sample mask, with each beamlet corresponding to a pixel on the detector. These beamlets are partially blocked by the detector mask. By performing a phase stepping procedure, which involves lateral translation of the sample mask relative to the detector mask, the fraction of the beamlet blocked by the detector mask changes. The resulting curve obtained by measuring the intensities as a function of the mask translation is referred to as the illumination curve (IC), shown in Fig. 5.1b. The IC is assumed to approximate a Gaussian shape. After the introduction of a sample, three contrasts can be retrieved by measuring the changes in the IC, in a step called *phase retrieval*: absorption, refraction, and scatter. The change of area under the IC is a measure of the absorption of the sample, the shift of the peak of the IC measures the refraction of the beamlet by the sample, and the broadening of the IC corresponds to the scatter contrast.

We represent the vectors corresponding to absorption, refraction, and scatter reconstructions as $\boldsymbol{x}_\mu$, $\boldsymbol{x}_\delta$, and $\boldsymbol{x}_\sigma$, respectively, where each vector

has dimensions $n \times 1$. To combine these vectors into a single vector, we define $\boldsymbol{x} \in \mathbb{R}^{3n}$. We assume a Gaussian model for the IC, such that the flatfield intensity $\boldsymbol{p}_0$, as function of the phase step ξ, can be modeled as:

$$\boldsymbol{p}_0(\xi) = a_0 \exp\left(-\frac{(\xi - b_0)^2}{2c_0^2}\right), \tag{5.1}$$

with a_0, b_0, c_0 the amplitude, mean and standard deviation, respectively. These parameters can be found by fitting a Gaussian to the measured flatfield data. In this paper, only a single fit is performed, namely on the mean illumination curve obtained by spatial averaging the intensities of each (phase stepped) flat field. That is, only three parameters are fitted in the preprocessing step.

We now consider models that describe the effects of the introduction of the object $\boldsymbol{x}$ on the beamlets. The absorption of the object, with absorption coefficients $\boldsymbol{x}_\mu$, is modeled by the Beer-Lambert law in X-ray CT as:

$$\boldsymbol{I}(\boldsymbol{x}_\mu) = \boldsymbol{I}_0 \exp\left(-\boldsymbol{A}\boldsymbol{x}_\mu\right), \tag{5.2}$$

with $\boldsymbol{I}$ the measured intensity, $\boldsymbol{I}_0$ the flatfield intensity, and $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ the discretized X-ray CT projector. In the case of EI, the area under the IC is considered to decrease caused by absorption. Therefore, we employ the same equation as (5.2), but where now $\boldsymbol{I}$ and $\boldsymbol{I}_0$ represent the integral of the measured and flatfield IC, respectively. Refraction of the object, which is based on the real part of the refraction index (δ), causes the mean of the measured IC to shift as follows:

$$\boldsymbol{s}(\boldsymbol{x}_\delta) = \boldsymbol{D}\boldsymbol{A}\boldsymbol{x}_\delta, \tag{5.3}$$

with $\boldsymbol{D} \in \mathbb{R}^{m \times m}$ a discrete differential operator. In what follows, $\boldsymbol{D}$ is the forward differences operator, which has been shown to be a good choice for X-ray phase contrast applications in [161]. Lastly, the dark-field contrast models scattering in the object due to the presence of sub-voxel structures. This shows as broadening of the IC, which we model this as the convolution of the IC with a normal distribution $\mathcal{N}(0, \boldsymbol{A}\boldsymbol{x}_\sigma)$. As such, we obtain the following joint forward projection model, where all operations with vectors are pointwise, except for the matrix-vector products with $\boldsymbol{A}$ and $\boldsymbol{D}$:

$$\boldsymbol{p}(\boldsymbol{x}, \xi) = \exp\left(-\boldsymbol{A}\boldsymbol{x}_\mu\right) \frac{a_0 c_0}{\sqrt{c_0^2 + \boldsymbol{A}\boldsymbol{x}_\sigma}} \exp\left(-\frac{(\xi - b_0 - \boldsymbol{D}\boldsymbol{A}\boldsymbol{x}_\delta)^2}{2\left(c_0^2 + \boldsymbol{A}\boldsymbol{x}_\sigma\right)}\right). \tag{5.4}$$

Analogously to the construction of $\boldsymbol{x}$, we can still represent $\boldsymbol{p}$ as a vector when multiple phase steps ξ_i per projection angle are considered, by

concatenating the different $\boldsymbol{p}(\boldsymbol{x}, \xi_i)$ to create a single vector, which we denote as $\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\xi})$.

5.2.2 Jacobian of the model

Given the forward model in (5.4) and some measured data $\boldsymbol{b}$, the reconstruction problem can be seen as a non-linear minimization problem of the following least-squares objective function:

$$\boldsymbol{f}(\boldsymbol{x}) = \frac{1}{2} \left|\left|\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\xi}) - \boldsymbol{b}\right|\right|_2^2, \tag{5.5}$$

with the reconstruction $\tilde{\boldsymbol{x}}$ being defined as:

$$\tilde{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}). \tag{5.6}$$

Non-linear optimization problems of the form of (5.6) can be solved in a variety of ways [4]. In this work, versions of gradient descent and quasi-Newton methods were considered. A key advantage of the proposed approach is that, unlike the conventional two-step method of reconstructing after phase retrieval, solving (5.6) does not necessitate fully sampled illumination curves for every pixel in each projection. This property was demonstrated in [77] for absorption and refraction contrast.

For the solvers considered in this work, the Jacobian of $\boldsymbol{p}(\boldsymbol{x}, \boldsymbol{\xi})$ with respect to $\boldsymbol{x}$ is required. We will denote this Jacobian of $\boldsymbol{p}$ with respect to $\boldsymbol{x}$ in a point $\boldsymbol{x_0}$ as $\boldsymbol{J_p^x}(\boldsymbol{x_0})$. Given this Jacobian, the gradient of $\boldsymbol{f}$ in (5.5) can be written as [4]:

$$\nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}_0} = \boldsymbol{J_p^x}(\boldsymbol{x}_0)^\top (\boldsymbol{p}(\boldsymbol{x_0}, \boldsymbol{\xi}) - \boldsymbol{b}). \tag{5.7}$$

Furthermore, for quasi-Newton methods such as Gauss-Newton and Levenberg-Marquardt [162], the required Hessian matrix of a least-squares problem can be approximated using the Jacobian matrix as:

$$\boldsymbol{H} \approx \boldsymbol{J}^\top \boldsymbol{J}. \tag{5.8}$$

The Jacobian $\boldsymbol{J_p^x}$ can be seen as a block matrix consisting of the Jacobians with respect to the separate contrasts considered in the model:

$$\boldsymbol{J_p^x} = \begin{bmatrix} \boldsymbol{J_p^{x_\mu}} & \boldsymbol{J_p^{x_\delta}} & \boldsymbol{J_p^{x_\sigma}} \end{bmatrix} \tag{5.9}$$

In the case of a single phase step per projection angle, one can calculate the Jacobian of (5.4) with respect to the different contrasts as follows:

$$\begin{aligned} \boldsymbol{J}_{\boldsymbol{p}}^{\boldsymbol{x}_\mu}(\boldsymbol{x}_0) &= -\mathrm{diag}\left(\boldsymbol{p}(\boldsymbol{x}_0,\xi)\right)\boldsymbol{A} \\ \boldsymbol{J}_{\boldsymbol{p}}^{\boldsymbol{x}_\delta}(\boldsymbol{x}_0) &= \mathrm{diag}\left(\frac{\boldsymbol{R}(\boldsymbol{x}_0,\xi)}{\boldsymbol{S}(\boldsymbol{x}_0)}\,\boldsymbol{p}(\boldsymbol{x}_0,\xi)\right)\boldsymbol{D}\boldsymbol{A} \\ \boldsymbol{J}_{\boldsymbol{p}}^{\boldsymbol{x}_\sigma}(\boldsymbol{x}_0) &= \mathrm{diag}\left(\frac{1}{2\boldsymbol{S}(\boldsymbol{x}_0)}\left(\frac{\boldsymbol{R}(\boldsymbol{x}_0,\xi)^2}{\boldsymbol{S}(\boldsymbol{x}_0)}-1\right)\boldsymbol{p}(\boldsymbol{x}_0,\xi)\right)\boldsymbol{A}, \end{aligned} \tag{5.10}$$

where

$$\begin{aligned} \boldsymbol{R}(\boldsymbol{x}_0,\xi) &= \xi - b_0 - \boldsymbol{D}\boldsymbol{A}\boldsymbol{x}_{0,\delta}, \\ \boldsymbol{S}(\boldsymbol{x}_0) &= c_0^2 + \boldsymbol{A}\boldsymbol{x}_{0,\sigma}. \end{aligned} \tag{5.11}$$

with all operations between vectors being executed pointwise, multiplications between matrices being standard matrix-multiplication and diag the operator transforming a vector into a diagonal matrix. If multiple phase steps per projection angle would be acquired, the resulting Jacobian is the sum of the Jacobians of the form described in (5.10) at each of the phase steps. This Jacobian matrix is typically too large to store in memory in real CT use cases, so instead functions are implemented that directly calculate the result of matrix multiplication with this Jacobian, based on the formulas in (5.10). All multiplications with the system matrix $\boldsymbol{A}$ or its transpose are performed on the GPU using the ASTRA toolbox [107].

5.2.3 Split gradient descent

Gradient descent is a widespread iterative method for solving optimization problems of the form of (5.6), where the next iteration is found from the previous as [4]:

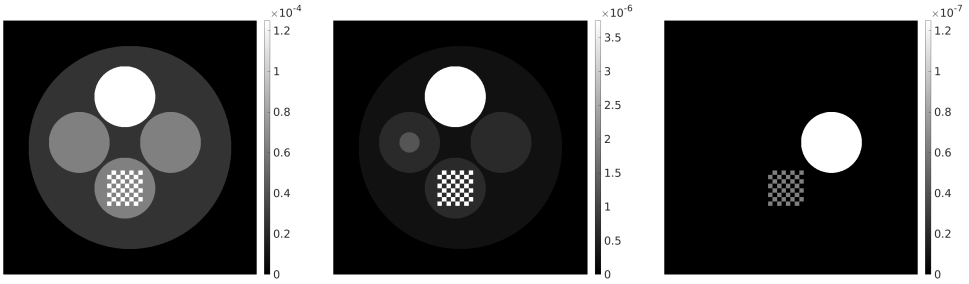
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha^{(k)} \nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k)}}\,. \tag{5.12}$$

Analogous to the Jacobian, the gradient of (5.5) with respect to $\boldsymbol{x}$ can be divided into partial gradients:

$$\nabla_{\boldsymbol{x}} \boldsymbol{f} = \begin{bmatrix} \nabla_{\boldsymbol{x}_\mu} \boldsymbol{f}^\top & \nabla_{\boldsymbol{x}_\delta} \boldsymbol{f}^\top & \nabla_{\boldsymbol{x}_\sigma} \boldsymbol{f}^\top \end{bmatrix}^\top, \tag{5.13}$$

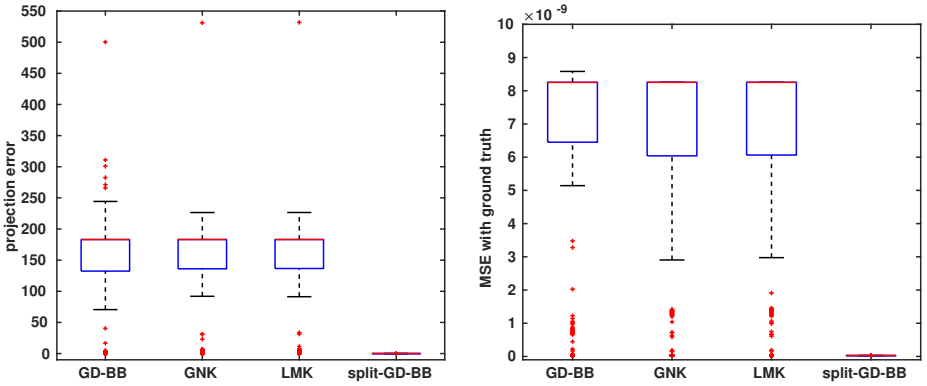
For finding a suitable value of $\alpha^{(k)}$ in (5.12) multiple options exist [4]. In this paper, we used the Barzilai-Borwein scheme [52] to calculate the step size in iteration k:

$$\alpha^{(k)} = \frac{\left(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\right) \cdot \left(\nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k)}} - \nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k-1)}}\right)}{\left(\nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k)}} - \nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k-1)}}\right) \cdot \left(\nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k)}} - \nabla_{\boldsymbol{x}}\, \boldsymbol{f}|_{\boldsymbol{x}^{(k-1)}}\right)}. \tag{5.14}$$



(a) Absorption **(b)** Refraction **(c)** Scatter

Figure 5.2. Different contrast channels, (a) absorption, (b) refraction and (c) scatter, of the ground truth phantom used for the simulation experiments.



(a) Boxplot of projection error **(b)** Boxplot of MSE with respect to the ground truth

Figure 5.3. Box plots showing the variation, by using different relative scaling of the contrasts, in minimally achieved projection error (a) and MSE with respect to the ground truth (b) for different optimizers of the EI objective function.

In addition to using this method to find a single step size $\alpha^{(k)}$, we also considered a split gradient descent method where three separate step sizes $\alpha_1^{(k)}, \alpha_2^{(k)}, \alpha_3^{(k)}$ are calculated for the partial gradients shown in (5.13). This allows the step size of the three contrasts to vary independently, with the hypothesis being that this would alleviate the need to find scaling parameters between the different contrasts.

Algorithm	Time per iteration	Total time
Gaussian fitting & phase retrieval	0.0142 s	2560 s
Reconstruction after retrieval	0.0595 s	119 s
Joint reconstruction	0.111 s	221 s

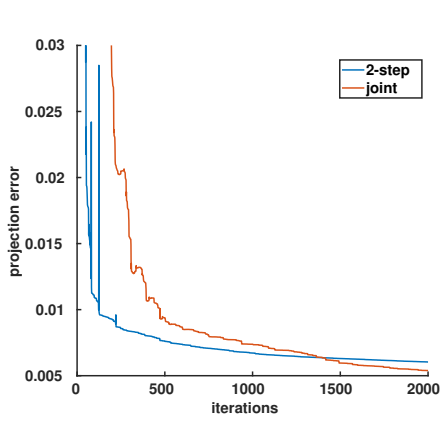
Table 5.1. Table of total times and time per iteration of the phase retrieval, reconstruction after phase retrieval and joint reconstruction on the full dataset with five phase steps. Time per iteration for the phase retrieval algorithm is the time per pixel. Reconstruction algorithms were run for 2000 iterations, phase retrieval was performed on $500 \times 360 = 180000$ pixels.

5.3 Results & discussion

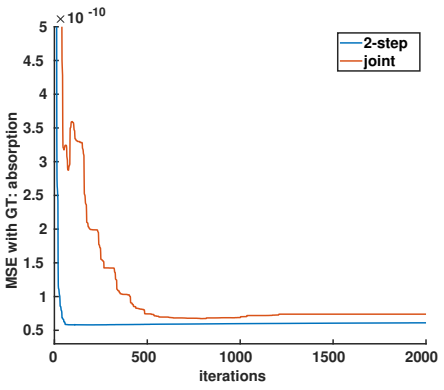
To assess the effectiveness of the proposed joint reconstruction method, we conducted evaluations on both simulated and experimental datasets. For the simulated dataset, we employed multiple optimizers and assessed their performance based on two metrics: projection error, which is also the value of the objective function, and the mean squared error (MSE) with respect to the ground truth. To ensure a fair comparison, gradient descent with Barzilai-Borwein steps was implemented for use with each of the separate contrasts after phase retrieval, using the same forward operators, $\boldsymbol{A}$ and $\boldsymbol{DA}$, as the joint reconstruction model.

5.3.1 Simulated data

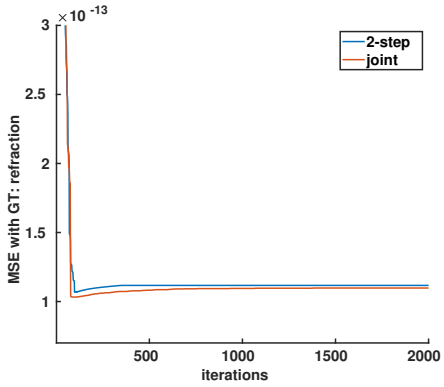
A simulation experiment was set up to select an appropriate optimization algorithm and to compare the joint reconstructions to the conventional two step workflow of reconstruction after phase retrieval. A phantom was defined on a 2000×2000 pixel grid, with a pixel size of 60 µm. The different ground truth contrast channels of the phantom are shown in Fig. 5.2. A total of 360 fan beam projection angles were considered, evenly distributed over a 360° range. For each projection angle, five projections were obtained with varying simulated sample mask displacements at $-13.5, -9, 0, 9$ and 13.5 µm, resulting in a total of 1800 fan beam projections The magnification factor was 1.25. The projections were created using the forward model in (5.4). To avoid using the same model in both simulation and reconstruction, the projections were computed on a 2000×2000 pixel grid, but reconstructed on a 500×500 pixel grid. It should be noted that, for the purpose of comparing with the conventional workflow, the model in (5.4) simulates Gaussians, which benefits the



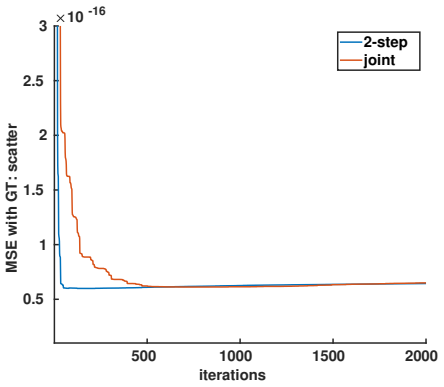
(a) Projection error.



(b) MSE with respect to ground truth absorption.



(c) MSE with respect to ground truth refraction.



(d) MSE with respect to ground truth scatter.

Figure 5.4. Plots of reconstruction quality as a function of the number of iterations for the simulated dataset, comparing the two-step method with GD-BB optimization and the joint reconstruction method with split-GD-BB optimization. Two error measures are shown: (a) projection error, measuring how well the simulated projection of the reconstruction fits with the measured data and (b-d) MSE with respect to the ground truth for each contrast channel separately.

Gaussian fitting process.

In the first experiment, we compared the performance of different optimizers for the joint reconstruction approach, on the full dataset. The same reconstructions were performed with gradient descent with Barzilai-Borwein steps (GD-BB) [52], split gradient descent with Barzilai-Borwein steps (split-GD-BB), Gauss-Newton-Krylov (GNK) [4] and Levenberg-

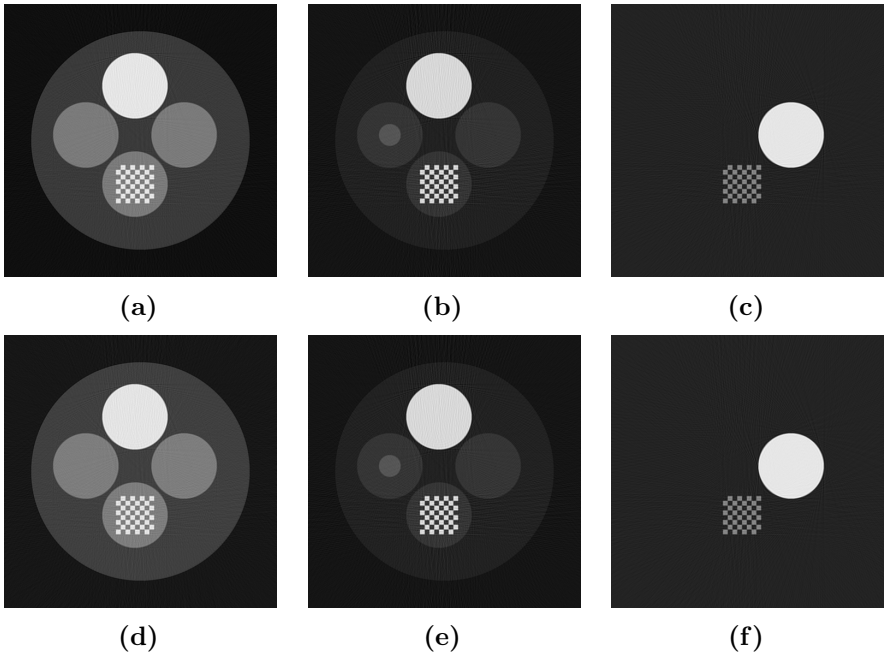


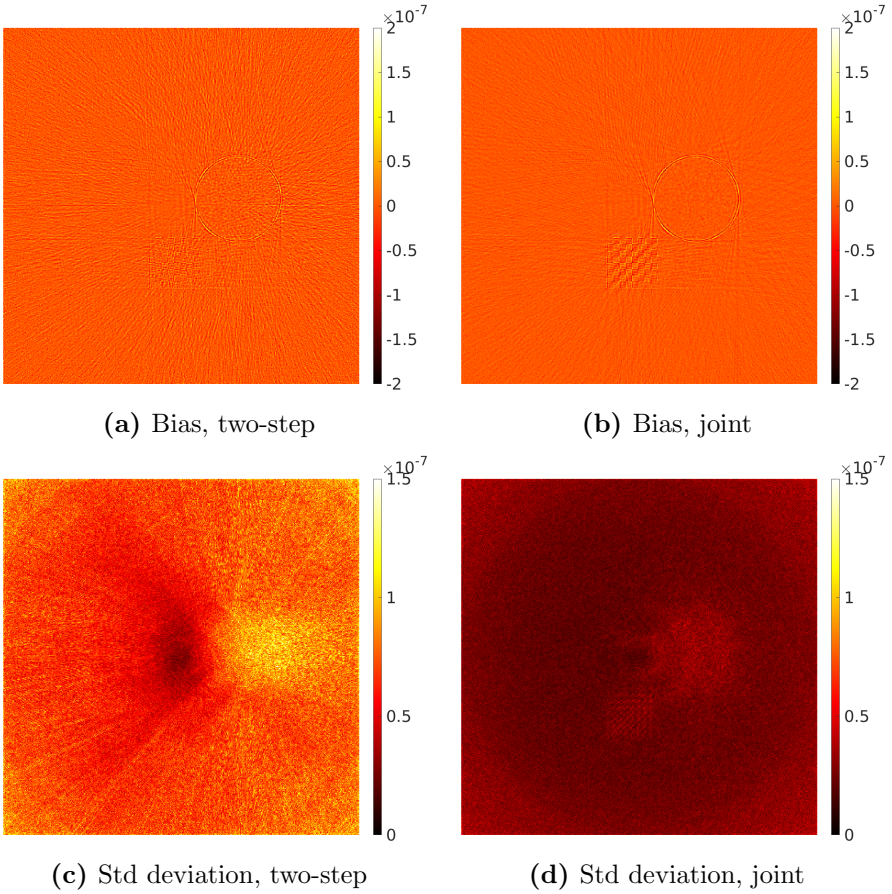
Figure 5.5. Reconstructions of the simulated dataset with 5 phase steps per projection. (a-c) two-step reconstruction method with GD-BB and (d-f) joint reconstruction method with split-GD-BB. Different contrast per column, left to right: absorption, refraction, scatter.

Marquardt-Krylov (LMK) [162]. Each contrast channel in the joint model can be scaled individually, which corresponds to changing the unit of the reconstructed values. These different potential scaling factors can have an impact on convergence speed and reconstruction quality. Therefore, the three different contrasts were each multiplied with scaling factors in $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$, resulting in $6^3 = 216$ different reconstructions for each method. The GD-BB and split-GD-BB were run for 500 iterations, GNK and LMK were run for 50 outer iterations with 10 inner Krylov iterations in each outer iteration. Two measures of the error are considered: projection error and mean squared error (MSE) with respect to the ground truth. For the projection error the reconstructed images are projected again using (5.4) and MSE with the measured data is calculated. For the MSE with respect to the ground truth, the MSE of the reconstruction with respect to each ground truth contrast channel is calculated separately. The box plots in Fig. 5.3 illustrate the minimum achieved projection error and ground truth MSE. It can be

observed that the reconstruction quality varies significantly depending on the scaling factors employed, except in the case of split-GD-BB. This can be attributed to the fact that split-GD-BB computes different step sizes for each contrast, making it less sensitive to relative scaling. Notably, the minimum achieved error in an outlier for GD-BB, GNK, and LMK, different for each one, is slightly lower compared to the split-GD-BB error. However, split-GD-BB offers the advantage of not requiring the determination of optimal parameters. The determination of scaling factors severely limits the practical application of the other optimizers for real reconstruction problems. Even when found, the GD-BB, GNK and LMK methods with optimal parameters do not output noticeably improved reconstructions over split-GD-BB without optimal parameters. Therefore, in the following experiments, only split-GD-BB was used as the optimizer for the proposed the joint reconstruction method.

Next, the reconstruction quality and convergence speed was compared between the joint reconstruction method with split-GD-BB optimizer and the two-step method with GD-BB optimizer. In both cases three step sizes were calculated with the BB formula for a single update of each contrast. Of both methods, 2000 iterations were performed and timed. For the two-step method, Gaussian curves were first fitted to each of the $360 \times 500 = 180.000$ ICs, using MATLAB's built-in LSQCURVEFIT function with trust region approach. The plots showing different errors, projection error and ground truth MSE, as function of number of iterations are shown in Fig. 5.4. The reconstructions after 2000 iterations are shown in Fig. 5.5. The total time and time per iteration are shown in Table. 5.1. From these figures, it is clear that the differences on the full dataset between the two reconstruction methods are minimal. The ground truth phantom is faithfully reconstructed, with the exception of some light streaking artefacts, for both methods. Comparing the difference images with the ground truth (not shown here), reveals that there are some additional weak artefacts at the interfaces between the materials, again of the same magnitude for both methods, which is likely due to the downsampling of the projection data. The joint reconstruction method reaches a lower projection error although differences in ground truth MSE are negligible. As both methods converge to the same reconstructions, the proposed joint reconstruction method has the benefit of providing this reconstruction faster, as it does not require a lengthy Gaussian fitting preprocessing step.

Additionally, the stability of the reconstructed values when noise is present in the projection data was investigated. The above experiment was repeated after Poisson noise was added to the projection data, for a flatfield intensity of 100000 photon counts per detector element. Twenty



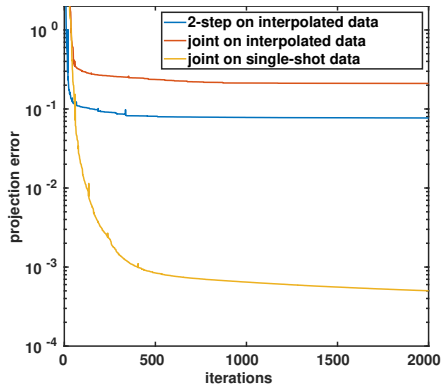
(a) Bias, two-step **(b)** Bias, joint

(c) Std deviation, two-step **(d)** Std deviation, joint

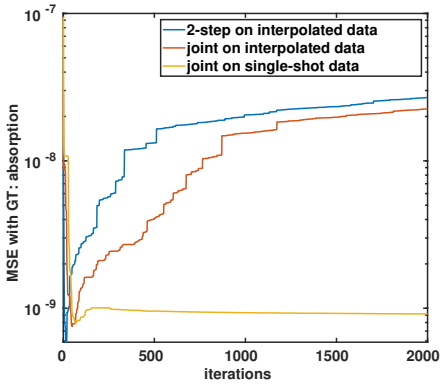
Figure 5.6. Bias (a-b) and standard deviation (c-d) maps of the scatter contrast for the simulated dataset from twenty repeated noise experiments.

independent noise realisations were considered and bias and standard deviation maps for both methods and for each contrast were calculated. For absorption and refraction both methods performed virtually identically. The bias and standard deviation maps for scatter are shown in Fig. 5.6. It can be seen that the standard deviation of the proposed method for scatter is significantly lower than that of the conventional two-step method. A possible explanation is that errors due to noise made in the IC fitting procedure propagate through, leading to less predictable reconstructions for the two-step method.

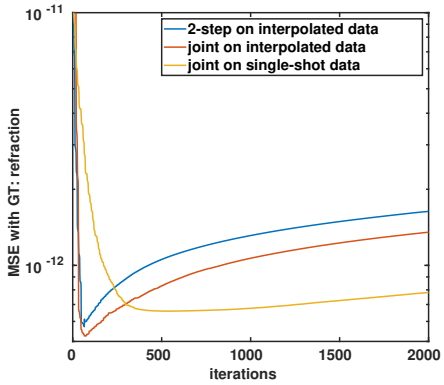
Lastly, the method was tested as a single-shot reconstruction algorithm. For each projection angle, a single projection at one mask displacement was acquired (i.e., no phase stepping at any projection



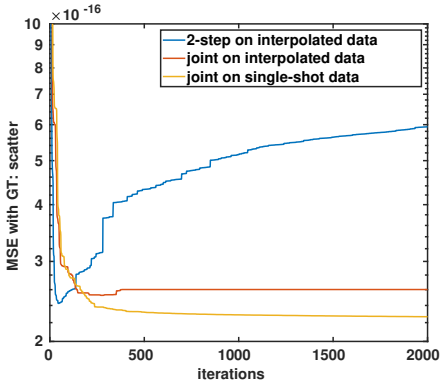
(a) Projection error.



(b) MSE with respect to ground truth absorption.



(c) MSE with respect to ground truth refraction.



(d) MSE with respect to ground truth scatter.

Figure 5.7. Plots of reconstruction quality as function of number of iterations for the simulated dataset with one phase step per projection, comparing the two-step method with GD-BB optimization on the interpolated data and the joint reconstruction method with split-GD-BB optimization on both the interpolated data and the single shot data: (a) projection error (b-d) MSE with respect to the ground truth, for each contrast channel separately.

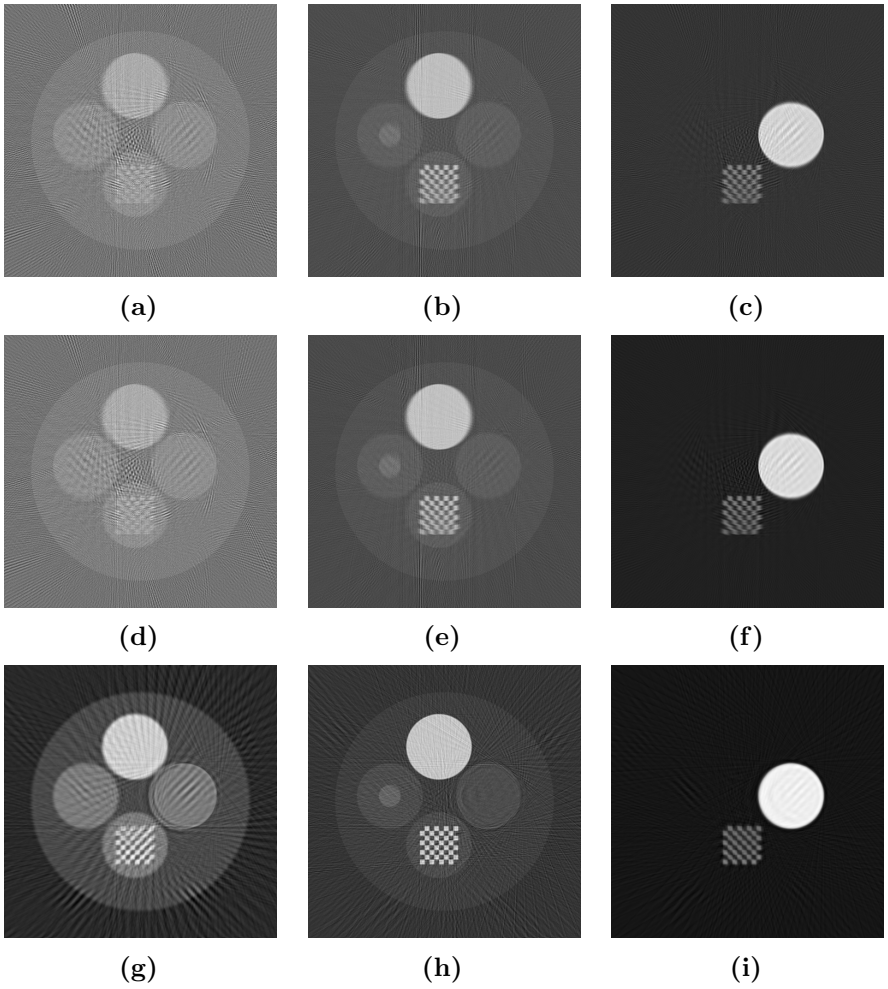


Figure 5.8. Reconstructions of the simulated single-shot data with 360 projections: (a-c) two-step method on the interpolated single-shot data, (d-f) joint reconstruction split-GD-BB on interpolated single-shot data, (g-i) joint reconstruction split-GD-BB on single-shot data. Different contrast per column, left to right: absorption, refraction, scatter.

angle). When moving to the next projection angle, the sample mask was stepped only once, to the next displacement in the full range of phase steps: $-13.5 - 9, 0, 9$ and $13.5\,\mu\text{m}$, returning to $-13.5\,\mu\text{m}$ at the end of the list. This sampling method is similar to the cycloidal CT presented in [163], where the sample undergoes a lateral displacement following each rotational step. The simulated dataset acquired this way for 360

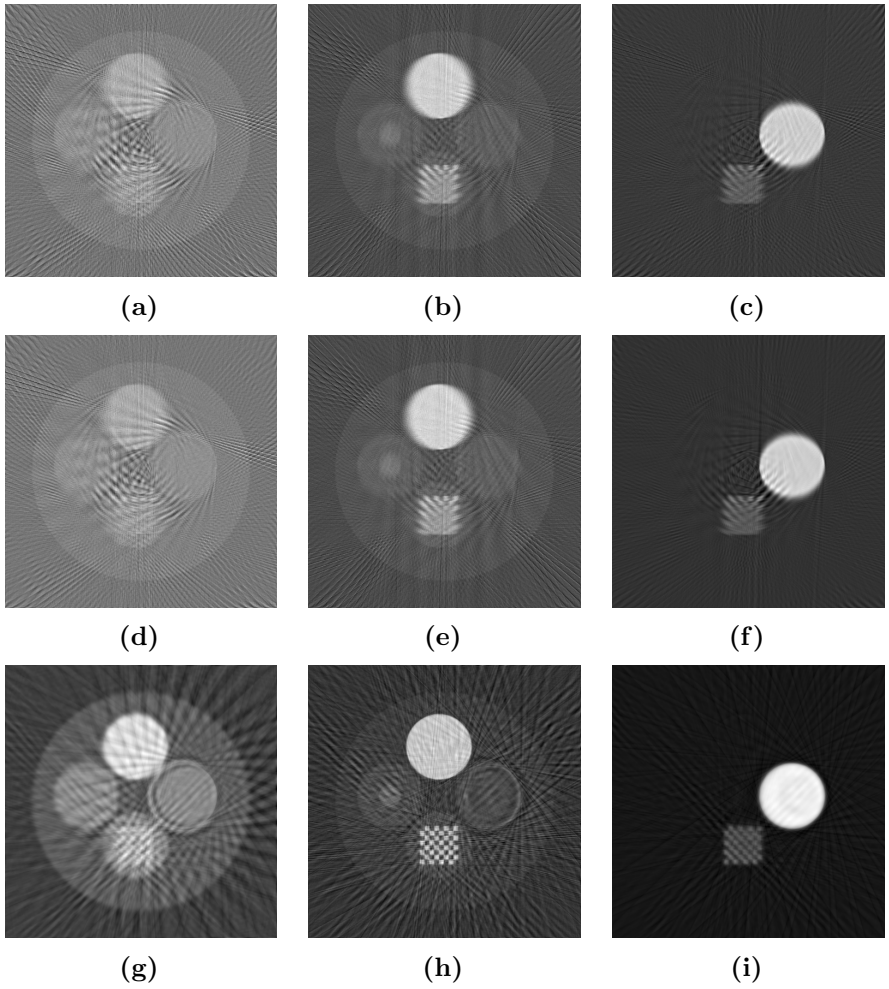


Figure 5.9. Reconstructions of the simulated single-shot data with 180 projections: (a-c) two-step method on the interpolated single-shot data, (d-f) joint reconstruction split-GD-BB on interpolated single-shot data, (g-i) joint reconstruction split-GD-BB on single-shot data. Different contrast per column, left to right: absorption, refraction, scatter.

uniformly distributed projections is a subset of the simulated dataset described before, with only 20% of the data available.

For the joint reconstruction method, the above described projection data can be reconstructed directly. However, for the conventional two-step method, this is not possible, as no Gaussian fitting can be performed to a single point in each projection pixel. To allow for a comparison

with the two-step method from the same dataset, the available data was interpolated to recreate 5 phase steps per projection. The data available at each mask displacement was used to fill in the missing projection angles at that displacement, using cubic interpolation. The two-step method was subsequently performed on this interpolated dataset. The joint reconstruction method was performed on both the original single-shot data and the interpolated dataset. The same single-shot methodology was also followed on a subset of the simulated data where only half of projection angles are taken.

The error measurements of this experiment are shown in Fig. 5.7 and reconstructions after 2000 iterations for each method are shown in Fig. 5.8. From the plots in Fig. 5.7, it can be observed that the proposed method on the non-interpolated data converges to a much lower projection error than the two-step method on the interpolated dataset, while the proposed method on the interpolated data converges to a slightly higher projection error. Furthermore, the different ground truth MSE plots show that both methods exhibits semi-convergent behaviour when performed on the interpolated data. However, the joint reconstructions on the single-shot data without interpolation do not show this semi-convergent behaviour. Although the methods on the interpolated data have a low MSE with respect to the ground truth after a low amount of iterations, the reconstruction here is extremely blurry and still has a high projection error. From the qualitative comparison in Fig. 5.8, it can be seen that both methods suffer from artefacts due to the subsampling. However, the reconstructions on interpolated data are unable to reconstruct the fine details, as they are lost in rotational blurring from the interpolation, which is especially visible in the checkerboard part of the phantom. In the reconstructions, a slight improvement of the reconstruction quality of the scatter contrast can be observed for the joint reconstruction on interpolated data, compared to the two-step method. The proposed joint reconstruction method on the data without interpolation still reconstructs most of the interior faithfully. When the number of single-shot projections is halved, it can be observed again in Fig. 5.9 that the quality of the reconstructions with the proposed method degrades much less severely than for the ones with the two-step method on interpolated data. This is to be expected, as the lower the number of single-shot projections, the more error is introduced by the interpolation method. In summary, the proposed method provides more accurate reconstructions without the need of a preprocessing step consisting of interpolation, Gaussian fitting and phase retrieval.

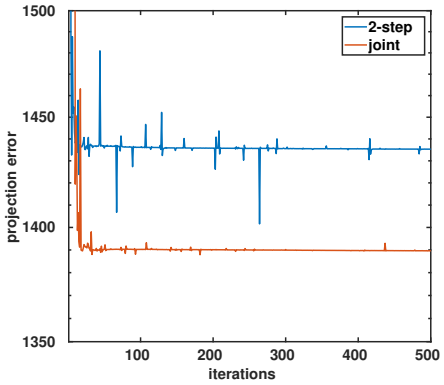


Figure 5.10. Plot of projection error as function of iterations for the experimental dataset, comparing the two-step and joint reconstruction methods.

5.3.2 Experimental dataset: fiber polymer

To further evaluate the performance of the joint reconstruction method and to confirm the conclusions from the simulation results of section 5.3.1, reconstructions from an experimental dataset were compared. The data consists of the central slice of a cone beam edge illumination dataset, acquired at the Advanced X-Ray Imaging Group, University College London. The imaged object was a fiber polymer block with fiber bundles woven in two orthogonal directions, held in place by plastic straws. Over a 360° range, 6250 projections were acquired, from 1250 uniformly spaced angles, with 5 phase steps per projection angle. From this fully sampled dataset, two reconstruction experiments were set up. First, reconstructions were made from all available data using both the conventional two-step method as well as the proposed joint reconstruction method. As no ground truth was available for this dataset, only the projection error was computed for this experiment. Next, subsets of the data were taken to simulate a single-shot dataset of the same object. Both the full range with 1250 and a subsampled range with 500 projections were considered. For this experiment, the MSE with respect to the reconstruction found in the first experiment is calculated, considering this as the best guess for the ground truth object. As undersampling artefacts in the background would be the main contributor to this MSE, the MSE with respect to the ground truth is calculated only in the region where the object is present. For the two-step method, reconstructions were again computed using GD-BB on the retrieved sinograms. As the simulation results in section 5.3.1 suggest that split-GD-BB performs

well and has less parameters that need to be tuned, only this optimizer was used for the joint reconstruction method for the experimental data. For both methods the optimizer was run for 500 iterations.

In Fig. 5.10, the evolution of the projection error over iterations is shown for the fully sampled dataset. The joint method reaches a lower projection error compared to the two-step method. The reconstructions made with the joint reconstruction method are therefore more faithful to the projection data. Erratic behaviour can be observed in the projection error curves, however, this is known behaviour of gradient descent when using Barzilai-Borwein steps [53]. In Fig. 5.11 the reconstructions after 500 iterations are shown. The reconstructions are again visually very similar on the full dataset, which is in line with the simulation results.

Next, the results of the single-shot experiment were compared. In Fig. 5.12, the evolution of the MSE with respect to the 'ground truth' reconstruction is shown and in Fig. 5.13the reconstructions are qualitatively compared. The joint method reaches a lower MSE with respect to the ground truth reconstruction compared to the two-step method on interpolated data in almost all cases. There is one exception: in refraction contrast with all 1250 projections, the measured error for the two-step method is slightly lower than the joint reconstruction method. In general, this is again in line with the simulation results, showing that the joint reconstruction method is able to compute accurate reconstructions from single-shot datasets. In Fig. 5.13, the reconstructions can be compared. The reconstructions with the joint method on the single-shot data are sharper than the two-step reconstructions on the interpolated data. The reconstructions of the interpolated data tend to have less background artefacts, as everything gets smoothed out slightly due to the rotational blurring. It is important to note that this blurring effect is not due to the two-step method, but inherent to the interpolation. Therefore, if the blurred reconstruction would qualitatively be preferred, the joint reconstruction method can also be employed on the interpolated dataset. This would result in similar reconstructions, without the need for a preprocessing step, as in the full data case studied before.

Future work building upon the proposed work can be envisioned in multiple fronts. To use the greater freedom in acquisition schemes, one could investigate ideal sampling schemes with varying rotational, dithering [163] and phase steps, as well as varying exposure times, using the proposed method. In modeling, more accurate physics models can be investigated, extending the forward model to include for example effects of polychromatic sources and directionality of scatter. Other choices for the differential operator for refraction contrast can also be considered. Furthermore, regularisation terms for the different contrast channels, as

well as joint regularisation using data from multiple contrast channels at once could be considered. Additionally, the forward model proposed here can also be used as the model for iterative statistical reconstruction methods, by adjusting the objective function according to a chosen noise model.

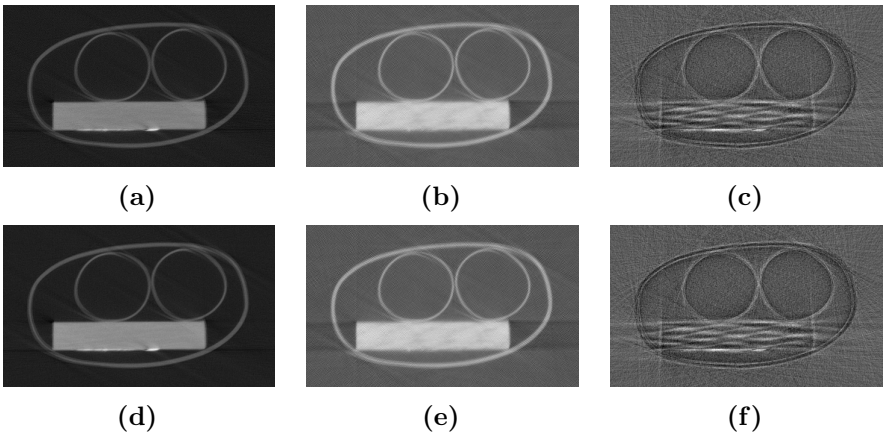
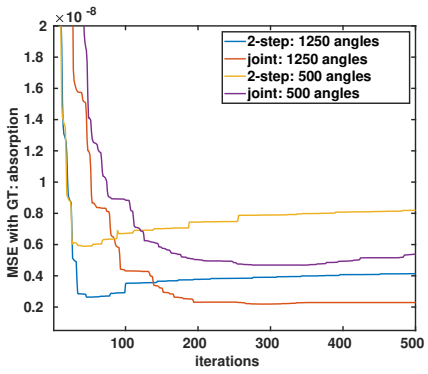


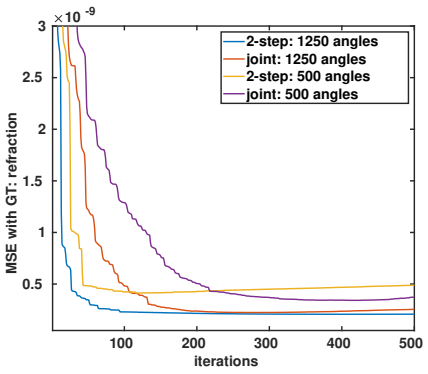
Figure 5.11. Cropped reconstructions of the fiber polymer block. (a-c) two-step method of the full dataset, (d-f) joint reconstruction method on the full dataset. Columns from left to right: absorption, refraction, and scatter contrasts.

5.4 Conclusion

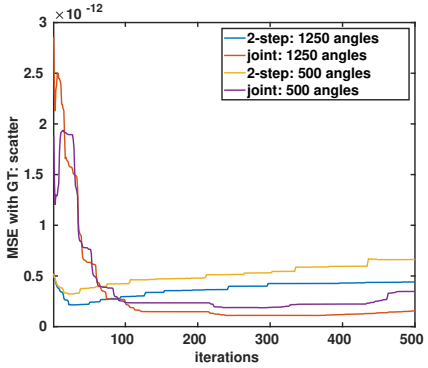
Conventional EI-based phase contrast reconstruction employs a two-step procedure: the fitting of a Gaussian function to the IC of each detector pixel to retrieve absorption, refraction, and scatter contrast, followed by the reconstruction of each of these contrasts. Significant drawbacks of this approach are the lengthy preprocessing of the data required before reconstruction and the necessity for complete sampling of each IC at every projection angle. In our paper, we introduced a joint reconstruction method capable of simultaneous reconstruction of all three contrasts: absorption, refraction, and scatter. We furthermore investigated different optimizer choices and concluded that gradient descent with a split Barzilai-Borwein scheme performed the best in terms of convergence speed and reconstruction quality with little scaling of the objective function needed. We showed that the proposed method reconstructs equally reliably as the state-of-the-art two-step procedure



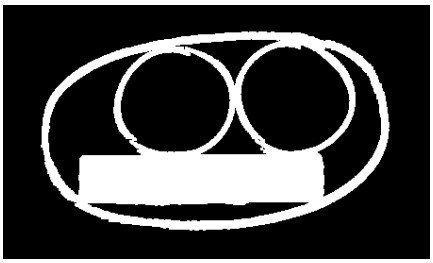
(a) MSE with respect to ground truth absorption.



(b) MSE with respect to ground truth refraction.



(c) MSE with respect to ground truth scatter.



(d) Mask for MSE.

Figure 5.12. (a-c) Plots of reconstruction quality as function of number of iterations for the experimental dataset with one phase step per projection, comparing the two-step method and the joint reconstruction method MSE with respect to the ground truth, for each contrast channel separately. The results on both 1250 and 500 projection angles are compared. (d) ROI where the MSE is calculated.

when using the same datasets, but without the need of performing Gaussian fitting and phase retrieval. Furthermore, the method allows for more freedom in acquisition schemes, as we showed that images can also be reconstructed from single-shot datasets that cannot be reconstructed with the two-step procedure without additional interpolation.

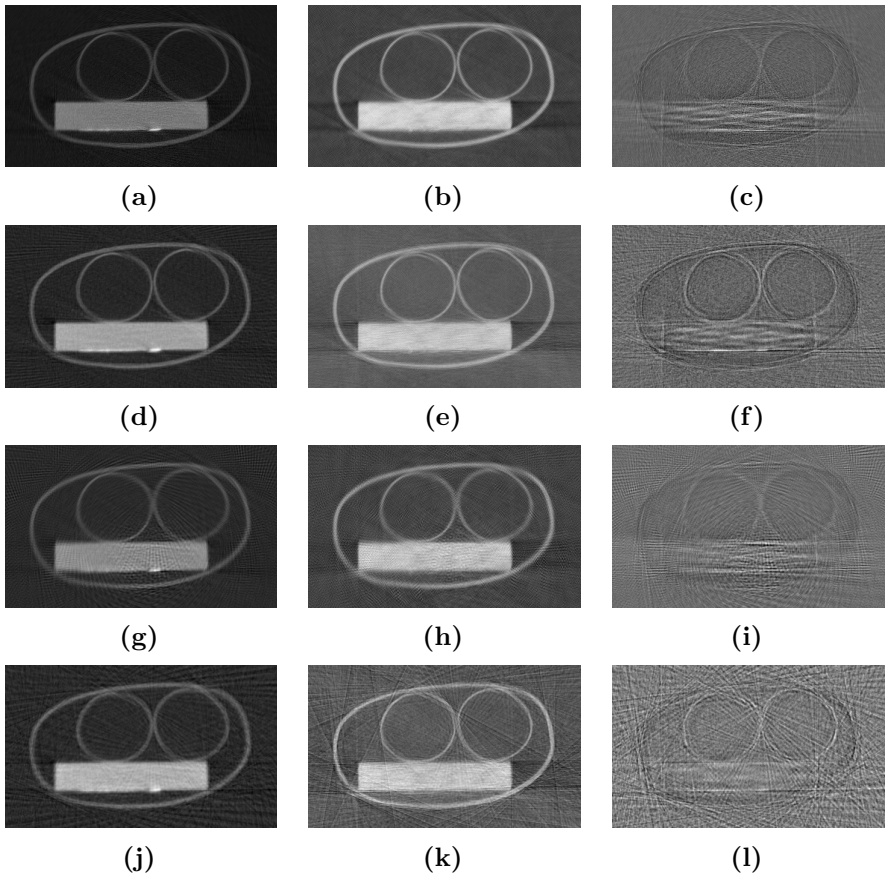


Figure 5.13. Cropped reconstructions of the fiber polymer block. (a-c) two-step method on 1250 single-shot projections, (d-f) joint reconstruction method on 1250 single-shot projections, (g-i) two-step method on 500 single-shot projections, (j-l) joint reconstruction method on 500 single-shot projections. Columns from left to right: absorption, refraction, and scatter contrasts.

CHAPTER 6

Conclusions and future perspectives

X-ray CT is a non-invasive imaging modality in which projections of an object are used to reconstruct its internal structure. The algorithms most commonly used to perform the reconstruction step, such as FBP and SIRT, can suffer from a wide range of artefacts. Reconstruction artefacts appear due to either undersampling, that is, lacking sufficient projections to solve the reconstruction problem, or due to a mismatch between the real process of image formation and the forward model used in the reconstruction algorithm. In this thesis we present multiple algorithms aiming to improve reconstruction quality in situations where the algorithms based on the standard Beer-Lambert model do not provide adequate results. In this thesis, we have attempted to incorporate more accurate physics models in the reconstruction algorithms to improve performance.

In Chapters 2 and 3, the combined problem of projections from polychromatic sources and undersampling is investigated. Starting from an existing heuristic discrete tomographic method, DART, that works on undersampled monochromatic data, we proposed a variation, called poly-DART, where the polychromaticity of the source and the energy dependence of the attenuation were included in the reconstruction method. Next, we proposed an improvement to the heuristic partitioning step underlying all DART-based methods, updating the partitioning rules throughout the algorithm. The resulting Tabu-DART framework is applied on the proposed poly-DART algorithm, further improving performance. Experiments on simulated and experimental data show that the included polychromatic forward model drastically improves the performance of discrete tomography on datasets from polychromatic sources.

In Section 1.3 we see that although SIRT is widely used and many

new algorithms are adapted from it, it is not a particularly efficient optimisation algorithm. Therefore, in Chapters 4 and 5 we investigated the use of different optimisation techniques for non-linear CT problems. In Chapter 4 different optimisation techniques are compared to reconstruct polychromatic data, with the same forward model as used in Chapter 2. The investigation shows that quasi-Newton methods outperform existing heuristic solvers, in particular a proposed polychromatic SIRT algorithm. At the same time, gradient-based methods are much more easily extended to adapt to any additional modelling or regularisation. Finally, a similar approach was used in Chapter 5, where we examine a different X-ray CT imaging modality: edge illumination phase contrast imaging. Here, we combined the three physical sources of contrast in one differentiable model and applied different quasi-Newton and gradient descent methods to the combined least-squares problem. The resulting optimisation algorithm has slightly improved performance compared to the state-of-the-art two-step reconstruction approach, while allowing much higher freedom in acquisition.

The work presented here shows that by using more accurate X-ray acquisition models, reconstructions can be achieved with little artefacts. In general, there is a trade-off between the accuracy of the modelling and the time required for reconstruction. We have shown that by employing quasi-Newton techniques the convergence speed can be improved, however, reconstructions will still take longer than with a simple SIRT algorithm.

Multiple ways forward from the presented research can be identified. Both for the conventional and EI CT models extensions of the presented model have to be considered, modelling, for example, source and detector effects, scatter, motion, etc. However, one must be vigilant that the reconstruction model does not end up being so complex that reconstruction time becomes unacceptable. Furthermore, although many quasi-Newton methods were compared, it is likely that there exist other numerical optimisation methods that have the potential to perform better on the presented models.

The work presented on polychromatic data required a priori known weights for the spectrum. The weights represent a multiplication of the emitted source spectrum and the spectral response of the detector. The spectrum estimation methods can be improved, and for better modelling, the source spectrum and the detector response should separately be estimated. Since the energy distribution of the X-ray beam changes as it passes through a sample, the spectral weight as used in the methods now does not fully capture the real effects. Given a separate source spectrum and spectral detector response, the outgoing spectrum along each ray

could easily be modelled and then multiplied with the detector response, resulting in more accurate simulated polychromatic images.

We saw that by changing models in the ARM step of DART, the discrete tomography method could be used on data that did not follow a linear acquisition model. Furthermore, in edge illumination we saw undersampling issues with the single-shot method. It would be interesting to attempt to perform discrete tomography on edge illumination datasets. Particularly interesting is, given the same number of total projections, whether it is ideal to take more phase steps with fewer projection angles or fewer phase steps with more projection angles. A challenge for applying a DART-like algorithm on EI data would be the lack of clear estimations for the expected grey value in the dark field signal, as there are still unexplained effects in the contrast, for example a flattening of the expected straight curve at higher sample thickness as shown in [164]. The undersampling issue can also be addressed with compressed sensing techniques, such as adding a total variation regularising term to the method presented in Chapter 5.

An interesting combination of the presented work would be to include polychromatic modelling of the edge illumination set up in the reconstruction algorithm, as polychromaticity has different effects on all contrasts. Some advances in modelling polychromaticity have been made in [165]. The equations in [165] show that the polychromatic effects described at the level of the retrieved contrast after Gaussian fitting are very complex, due to the mixing of multiple effects, suggesting that using an extension of the joint reconstruction method proposed in this work could be more successful than methods after phase retrieval. Furthermore, if we compare to the polychromatic model used in Chapter 4, a reconstruction method that accurately employs a similar model for EI can only be a joint reconstruction method, as the dependency can only be accurately described at the material level and not at the projection level. Particularly challenging would be the dependence of the linear dark field coefficients on energy, as experimentally verified values do not yet exist. In such a model the transmission through the grating bars, resulting in an offset in the projections, could also easily be modelled more accurately. As the X-rays first travelled through the grating bars, their spectrum is a heavily filtered version of the source spectrum.

Appendix A: Scientific contributions

Journal articles

1. **Six, N.**, De Beenhouwer, J. and Sijbers, J., 2024. Joint multi-contrast CT for edge illumination X-ray phase contrast imaging using split Barzilai-Borwein steps. Optics Express, 32(2), pp.1135-1150.

2. Vanthienen, P.J., Sanctorum, J., Huyge, B., **Six, N.**, Sijbers, J. and De Beenhouwer, J., 2023. Grating designs for cone beam edge illumination X-ray phase contrast imaging: a simulation study. Optics Express, 31(17), pp.28051-28064.

3. Frenkel, D., **Six, N.**, De Beenhouwer, J. and Sijbers, J., 2022. Tabu-DART: a dynamic update strategy for efficient discrete algebraic reconstruction. The Visual Computer, pp.1-13.

4. De Samber, B., Renders, J., Elberfeld, T., Maris, Y., Sanctorum, J., **Six, N.**, Liang, Z., De Beenhouwer, J. and Sijbers, J., 2021. FleXCT: a flexible X-ray CT scanner with 10 degrees of freedom. Optics Express, 29(3), pp.3438-3457.

5. **Six, N.**, Renders, J., Sijbers, J. and De Beenhouwer, J., 2021. Gauss-Newton-Krylov for reconstruction of polychromatic X-ray CT images. IEEE Transactions on Computational Imaging, 7, pp.1304-1313.

6. Nazemi, E., **Six, N.**, Iuso, D., De Samber, B., Sijbers, J. and De Beenhouwer, J., 2021. Monte-Carlo-based estimation of the X-ray energy spectrum for CT artifact reduction. Applied Sciences, 11(7), p.3145.

7. **Six, N.**, De Beenhouwer, J. and Sijbers, J., 2019. poly-DART: A discrete algebraic reconstruction technique for polychromatic X-ray CT. Optics Express, 27(23), pp.33670-33682.

Conference proceedings

1. Huyge, B., Vanthienen, P.J., **Six, N.**, Sijbers, J. and De Beenhouwer, J., 2023. Adapting an XCT-scanner to enable edge illumination x-ray phase contrast imaging. In Online e-journal of nondestructive testing. (Vol. 28, No. 3, pp. 1-7)

2. **Six, N.**, Renders, J., De Beenhouwer, J. and Sijbers, J., 2022, October. Joint reconstruction of attenuation, refraction and dark field x-ray phase contrasts using split Barzilai-Borwein steps. In Developments in X-Ray Tomography XIV (Vol. 12242, pp. 113-118). SPIE.

3. Sanctorum, J., **Six, N.**, Sijbers, J. and De Beenhouwer, J., 2022, October. Augmenting a conventional x-ray scanner with edge illumination-based phase contrast imaging: how to design the gratings. In Developments in X-Ray Tomography XIV (Vol. 12242). SPIE.

4. Vanthienen, P.J., Sanctorum, J., Huyge, B., **Six, N.**, Sijbers, J. and De Beenhouwer, J., 2022, October. Alternative grating designs for cone-beam edge illumination x-ray phase contrast imaging. In Developments in X-Ray Tomography XIV (Vol. 12242). SPIE.

5. Huyge, B., Sanctorum, J.,**Six, N.**, De Beenhouwer, J. and Sijbers, J., 2021, April. Analysis of flat fields in edge illumination phase contrast imaging. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (pp. 1310-1313). IEEE.

6. Iuso, D., Nazemi, E., **Six, N.**, De Samber, B., De Beenhouwer, J. and Sijbers, J., 2021, September. CAD-Based scatter compensation for polychromatic reconstruction of additive manufactured parts. In 2021 IEEE International Conference on Image Processing (ICIP) (pp. 2948-2952). IEEE.

7. **Six, N.**, Renders, J., Sijbers, J. and De Beenhouwer, J., 2020, October. Newton-Krylov methods for polychromatic X-ray CT. In 2020 IEEE International Conference on Image Processing (ICIP) (pp. 3045-3049). IEEE.

8. **Six, N.**, De Beenhouwer, J., Van Nieuwenhove, V., Vanroose, W. and Sijbers, J., 2018, November. Joint reconstruction and flat-field estimation using support estimation. In 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC) (pp. 1-4). IEEE.

9. **Six, N.**, De Beenhouwer, J. and Sijbers, J., 2018. pDART: Discrete algebraic reconstruction using a polychromatic forward model. In Fifth International Conference on Image Formation in X-Ray Computed Tomography (pp. 1-4).

Bibliography

[1] T. M. Buzug. *Computed tomography: from photon statistics to modern cone-beam CT*. Springer Science & Business Media, 2008.

[2] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. SIAM, 2001.

[3] B. De Samber, J. Renders, T. Elberfeld, Y. Maris, J. Sanctorum, N. Six, Z. Liang, J. De Beenhouwer and J. Sijbers. FleXCT: a flexible X-ray CT scanner with 10 degrees of freedom. *Optics Express*, 29(3):3438–3457, 2021.

[4] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Science & Business Media, 2006.

[5] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6:81–94, 1984.

[6] P. Gilbert. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117, 1972.

[7] R. Gordon, R. Bender and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of theoretical Biology*, 29(3):471–481, 1970.

[8] A. Van der Sluis and H. Van der Vorst. SIRT-and CG-type methods for the iterative solution of sparse linear least-squares problems. *Linear Algebra and its Applications*, 130:257–303, 1990.

[9] J. Gregor and T. Benson. Computational analysis and improvement of SIRT. *IEEE transactions on medical imaging*, 27(7):918–924, 2008.

[10] J. F. Barrett and N. Keat. Artifacts in CT: recognition and avoidance. *Radiographics*, 24(6):1679–1691, 2004.

[11] F. E. Boas, D. Fleischmann et al. CT artifacts: causes and reduction techniques. *Imaging Med*, 4(2):229–240, 2012.

[12] A.-M. Du Plessis, S. Theron and S. Andronikou. The effects of misinterpretation of an artefact on multidetector row CT scans in children. *Pediatric radiology*, 39:137–141, 2009.

[13] H. O. Aggrawal, M. S. Andersen, S. D. Rose and E. Y. Sidky. A convex reconstruction model for X-ray tomographic imaging with uncertain flat-fields. *IEEE transactions on computational imaging*, 4(1):17–31, 2017.

[14] C.-C. Cheng, Y.-T. Ching, P.-H. Ko and Y. Hwu. Correction of center of rotation and projection angle in synchrotron X-ray computed tomography. *Scientific reports*, 8(1):9884, 2018.

[15] D. Evangelista, E. Morotti and E. L. Piccolomini. RISING: A new framework for model-based few-view CT image reconstruction with deep learning. *Computerized Medical Imaging and Graphics*, 103:102156, 2023.

[16] L. Butzhammer and T. Hausotte. Effect of iterative sparse-view CT reconstruction with task-specific projection angles on dimensional measurements. In *9th Conference on Industrial Computed Tomography, Padova, Italy (iCT2019)*, 2019.

[17] X. Jia, Y. Lou, B. Dong, Z. Tian and S. Jiang. 4D computed tomography reconstruction from few-projection data via temporal non-local regularization. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010: 13th International Conference, Beijing, China, September 20-24, 2010, Proceedings, Part I 13*, pages 143–150. Springer, 2010.

[18] J. Kastner, B. Plank and C. Heinzl. Advanced x-ray computed tomography methods: High resolution CT, phase contrast CT, quantitative CT and 4DCT. *Digital Industrial Radiology and Computed Tomography (DIR 2015), Ghent, Belgium*, 2015.

[19] A. Lange, A. Kupsch, B. R. Müller and M. P. Hentschel. Edge artefacts of radiographic images by X-ray refraction. In *Proceedings 18th World Conference on Non-Destructive Testing, Durban*, 2012.

[20] A. Hertanto, Q. Zhang, Y.-C. Hu, O. Dzyubak, A. Rimner and G. S. Mageras. Reduction of irregular breathing artifacts in respiration-correlated CT images using a respiratory motion model. *Medical physics*, 39(6Part1):3070–3079, 2012.

[21] A. Nguyen, J. Renders, J. Soete, M. Wevers, J. Sijbers and J. De Beenhouwer. An accelerated motion-compensated iterative reconstruction technique for dynamic computed tomography. In *Developments in X-Ray Tomography XIV*, volume 12242. SPIE, 2022.

[22] J. Renders, J. De Beenhouwer and J. Sijbers. Mesh-based reconstruction of dynamic foam images using X-ray CT. In *2021 International Conference on 3D Vision (3DV)*, pages 1312–1320. IEEE, 2021.

[23] J. Renders, J. Sijbers and J. De Beenhouwer. Adjoint image warping using multivariate splines with application to four-dimensional computed tomography. *Medical Physics*, 48(10):6362–6374, 2021.

[24] V. Van Nieuwenhove, J. De Beenhouwer, J. Vlassenbroeck, M. Brennan and J. Sijbers. MoVIT: a tomographic reconstruction framework for 4D-CT. *Optics Express*, 25(16):19236–19250, 2017.

[25] G. Zang, R. Idoughi, R. Tao, G. Lubineau, P. Wonka and W. Heidrich. Warp-and-project tomography for rapidly deforming objects. *ACM Transactions on Graphics (TOG)*, 38(4):1–13, 2019.

[26] I. A. Elbakri and J. A. Fessler. Statistical image reconstruction for polyenergetic X-ray computed tomography. *IEEE transactions on medical imaging*, 21(2):89–99, 2002.

[27] I. A. Elbakri and J. A. Fessler. Segmentation-free statistical image reconstruction for polyenergetic X-ray computed tomography with experimental validation. *Physics in Medicine & Biology*, 48(15):2453, 2003.

[28] A. C. Silva, H. J. Lawder, A. Hara, J. Kujak and W. Pavlicek. Innovations in CT dose reduction strategy: application of the adaptive statistical iterative reconstruction algorithm. *American Journal of Roentgenology*, 194(1):191–199, 2010.

[29] N. Six, J. Renders, J. De Beenhouwer and J. Sijbers. Joint reconstruction of attenuation, refraction and dark field x-ray phase contrasts using split barzilai-borwein steps. In *Proceedings of SPIE*, volume 12242, 122420O, 2022.

[30] L. Zhu, N. R. Bennett and R. Fahrig. Scatter correction method for x-ray CT using primary modulation: theory and preliminary results. *IEEE transactions on medical imaging*, 25(12):1573–1587, 2006.

[31] J. A. Meganck, K. M. Kozloff, M. M. Thornton, S. M. Broski and S. A. Goldstein. Beam hardening artifacts in micro-computed tomography scanning can be reduced by X-ray beam filtration and the resulting images can be used to accurately measure bmd. *Bone*, 45(6):1104–1116, 2009.

[32] G. T. Herman. Correction for beam hardening in computed tomography. *Physics in Medicine & Biology*, 24(1):81, 1979.

[33] M. Kachelrieß, K. Sourbelle and W. A. Kalender. Empirical cupping correction: a first-order raw data precorrection for cone-beam computed tomography. *Medical Physics*, 33(5):1269–1274, 2006.

[34] Y. Lin and E. Samei. An efficient polyenergetic SART (pSART) reconstruction algorithm for quantitative myocardial CT perfusion. *Medical Physics*, 41(2), 2014.

[35] J. H. Hubbell and S. M. Seltzer. Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z = 1 to 92 and 48 additional substances of dosimetric interest. Technical report, National Institute of Standards and Technology-PL, Gaithersburg, MD, United States, 1995.

[36] T. Humphries. Convergence analysis of a polyenergetic SART algorithm. *Medical Physics*, 42(7):4007–4014, 2015.

[37] A. Ziabari, S. Venkatakrishnan, Z. Snow, A. Lisovich, M. Sprayberry, P. Brackman, C. Frederick, P. Bhattad, S. Graham, P. Bingham et al. Enabling rapid X-ray CT characterisation for additive manufacturing using CAD models and deep learning-based reconstruction. *npj Computational Materials*, 9(1):91, 2023.

[38] D. L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[39] J. F. Abascal, M. Abella, A. Sisniega, J. J. Vaquero and M. Desco. Investigation of different sparsity transforms for the PICCS algorithm in small-animal respiratory gated CT. *PLOS one*, 10(4):e0120140, 2015.

[40] R. Chartrand, E. Y. Sidky and X. Pan. Nonconvex compressive sensing for X-ray CT: an algorithm comparison. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 665–669. IEEE, 2013.

[41] G.-H. Chen, J. Tang and S. Leng. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical physics*, 35(2):660–663, 2008.

[42] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd and L. Xing. Compressed sensing based cone-beam computed tomography reconstruction with a first-order method. *Medical physics*, 37(9):5113–5125, 2010.

[43] B. Schillinger. Proposed combination of CAD data and discrete tomography for the detection of coking and lubricants in turbine blades or engines. *Electronic Notes in Discrete Mathematics*, 20:493–499, 2005.

[44] G. T. Herman and A. Kuba. Discrete tomography in medical imaging. *Proceedings of the IEEE*, 91(10):1612–1626, 2003.

[45] G. T. Herman and A. Kuba. *Advances in discrete tomography and its applications*. Springer Science & Business Media, Berlin, 2008.

[46] G. T. Herman and A. Kuba. *Discrete tomography: Foundations, algorithms, and applications*. Springer Science & Business Media, Berlin, 2012.

[47] K. J. Batenburg and J. Sijbers. DART: a practical reconstruction algorithm for discrete tomography. *IEEE Trans. Image Process.*, 20(9):2542–2553, 2011.

[48] K. J. Batenburg, J. Sijbers, H. F. Poulsen and E. Knudsen. DART: a robust algorithm for fast reconstruction of three-dimensional grain maps. *Journal of Applied Crystallography*, 43(6):1464–1473, 2010.

[49] J. H. Ketola, S. S. Karhula, M. A. J. Finnilä, R. K. Korhonen, W. Herzog, S. Siltanen, M. T. Nieminen and S. Saarakkala. Iterative and discrete reconstruction in the evaluation of the rabbit model of osteoarthritis. *Sci. Rep.*, 8(1):12051, 2018.

[50] G. Schena, M. Piller and M. Zanin. Discrete X-ray tomographic reconstruction for fast mineral liberation spectrum retrieval. *International Journal of Mineral Processing*, 145:1–6, 2015.

[51] E. Van de Casteele, E. Perilli, W. van Aarle, K. J. Reynolds and J. Sijbers. Discrete tomography in an in vivo small animal bone study. *J. Bone Miner. Metab.*, 36(1):40–53, 2018.

[52] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA journal of numerical analysis*, 8(1):141–148, 1988.

[53] E. G. Birgin, J. M. Martínez and M. Raydan. Spectral projected gradient methods: review and perspectives. *Journal of Statistical Software*, 60:1–21, 2014.

[54] P. N. Brown and Y. Saad. Hybrid Krylov methods for nonlinear systems of equations. *SIAM Journal on Scientific and Statistical Computing*, 11(3):450–481, 1990.

[55] R. Fletcher and C. M. Reeves. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, 1964.

[56] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 12(4):617–629, 1975.

[57] Y. Saad and H. A. Van Der Vorst. Iterative solution of linear systems in the 20th century. *Journal of Computational and Applied Mathematics*, 123(1-2):1–33, 2000.

[58] J. Gregor and J. A. Fessler. Comparison of SIRT and SQS for regularized weighted least squares image reconstruction. *IEEE Transactions on Computational Imaging*, 1(1):44–55, 2015.

[59] A. Olivo. Edge-illumination X-ray phase-contrast imaging. *Journal of Physics: Condensed Matter*, 33(36):363002, 2021.

[60] A. Momose, T. Takeda, Y. Itai and K. Hirano. Phase-contrast X-ray computed tomography for observing biological soft tissues. *Nature medicine*, 2(4):473–475, 1996.

[61] F. Arfelli, M. Assante, V. Bonvicini, A. Bravin, G. Cantatore, E. Castelli, L. Dalla Palma, M. Di Michiel, R. Longo, A. Olivo, S. Pani, D. Pontoni, P. Poropat, M. Prest, A. Rashevsky, G. Tromba, A. Vacchi, E. Vallazza and F. Zanconati. Low-dose phase contrast X-ray medical imaging. *Physics in Medicine & Biology*, 43(10):2845, 1998.

[62] M. Marenzana, C. K. Hagen, P. D. N. Borges, M. Endrizzi, M. B. Szafraniec, T. L. Vincent, L. Rigon, F. Arfelli, R.-H. Menk and A. Olivo. Synchrotron-and laboratory-based X-ray phase-contrast imaging for imaging mouse articular cartilage in the absence of radiopaque contrast agents. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2010):20130127, 2014.

[63] L. Massimi, T. Suaris, C. K. Hagen, M. Endrizzi, P. R. T. Munro, G. Havariyoun, P. M. S. Hawker, B. Smit, A. Astolfo, O. J. Larkin, R. M. Waltham, Z. Shah, S. W. Duffy, R. L. Nelan, A. Peel, J. L. Jones, I. G. Haig, D. Bate and A. Olivo. Volumetric High-Resolution X-Ray Phase-Contrast Virtual Histology of Breast Specimens With a Compact Laboratory System. *IEEE Transactions on Medical Imaging*, 41(5):1188–1195, 2022.

[64] D. Shoukroun, L. Massimi, M. Endrizzi, D. Bate, P. Fromme and A. Olivo. Edge illumination X-ray phase contrast imaging for impact damage detection in CFRP. *Materials Today Communications*, 31(October 2021):103279, 2022.

[65] Z. Hu, P. Thomas, A. Snigirev, I. Snigireva, A. Souvorov, P. Smith, G. Ross and S. Teat. Phase-mapping of periodically domain-inverted LiNbO3 with coherent X-rays. *Nature*, 392(6677):690–693, 1998.

[66] V. Revol, I. Jerjen, C. Kottler, P. Schütz, R. Kaufmann, T. Lüthi, U. Sennhauser, U. Straumann and C. Urban. Sub-pixel porosity revealed by X-ray scatter dark field imaging. *Journal of Applied Physics*, 110(4):044912, 2011.

[67] A. Olivo, D. Chana and R. Speller. A preliminary investigation of the potential of phase contrast X-ray imaging in the field of homeland security. *Journal of Physics D: Applied Physics*, 41(22):225503, 2008.

[68] A. Olivo, K. Ignatyev, P. Munro and R. Speller. Design and realization of a coded-aperture based X-ray phase contrast imaging for homeland security applications. *Nuclear Instruments and Methods in Physics Research A*, 610(2):604–614, 2009.

[69] A. Zamir, C. Hagen, P. C. Diemoz, M. Endrizzi, F. Vittoria, Y. Chen, M. A. Anastasio and A. Olivo. Recent advances in edge illumination X-ray phase-contrast tomography. *Journal of Medical Imaging*, 4(04):1, 2017.

[70] A. Olivo and R. Speller. Modelling of a novel x-ray phase contrast imaging technique based on coded apertures. *Physics in Medicine & Biology*, 52(22):6555, 2007.

[71] P. Diemoz, C. Hagen, M. Endrizzi and A. Olivo. Sensitivity of laboratory based implementations of edge illumination x-ray phase-contrast imaging. *Applied Physics Letters*, 103(24), 2013.

[72] J. Sanctorum, N. Six, J. Sijbers and J. De Beenhouwer. Augmenting a conventional x-ray scanner with edge illumination-based phase contrast imaging: how to design the gratings. In *Developments in X-Ray Tomography XIV*, volume 12242. SPIE, 2022.

[73] B. Huyge, P.-J. Vanthienen, N. Six, J. Sijbers and J. De Beenhouwer. Adapting an XCT-scanner to enable edge illumination x-ray phase contrast imaging. In *Online e-journal of nondestructive testing*, volume 28(3), pages 1–7, 2023.

[74] N. Djurabekova, I. Buchanan, A. Doherty, M. Endrizzi1, J. Enwald1, P. Munro, A. Olivo, S. Savvidis and S. Arridge. Comparison of different Gaussian fitting and phase contrast retrieval methods. Research Square preprint `https://doi.org/10.21203/rs.3.rs-2190429/v1`.

[75] Z.-F. Huang, K.-J. Kang, Z. Li, P.-P. Zhu, Q.-X. Yuan, W.-X. Huang, J.-Y. Wang, D. Zhang and A.-M. Yu. Direct computed tomographic reconstruction for directional-derivative projections of computed tomography of diffraction enhanced imaging. *Applied Physics Letters*, 89(4), 2006.

[76] J. Fu, S. Schleede, R. Tan, L. Chen, M. Bech, K. Achterhold, M. Gifford, R. Loewen, R. Ruth and F. Pfeiffer. An algebraic iterative reconstruction technique for differential X-ray phase-contrast computed tomography. *Zeitschrift für medizinische Physik*, 23(3):186–193, 2013.

[77] Y. Chen, H. Guan, C. K. Hagen, A. Olivo and M. A. Anastasio. Single-shot edge illumination X-ray phase-contrast tomography enabled by joint image reconstruction. *Optics Letters*, 42(3):619–622, 2017.

[78] M. von Teuffenbach, T. Koehler, A. Fehringer, M. Viermetz, B. Brendel, J. Herzen, R. Proksa, E. J. Rummeny, F. Pfeiffer and P. B. Noël. Grating-based phase-contrast and dark-field computed tomography: a single-shot method. *Scientific reports*, 7(1):1–8, 2017.

[79] P. Modregger, J. Meganck, C. K. Hagen, L. Massimi, A. Olivo and M. Endrizzi. Improved iterative tomographic reconstruction for x-ray imaging with edge-illumination. *Physics in Medicine & Biology*, 64(20):205008, 2019.

[80] P. Modregger, T. Korzec, J. Meganck, L. Massimi, A. Olivo and M. Endrizzi. Accelerated iterative tomographic reconstruction with x-ray edge illumination. *arXiv preprint arXiv:2202.08627*, 2022.

[81] A. Doherty, S. Savidis, A. Astolfo, L. Massimi, N. Djurabekova, C. N. Leon, M. F. Gerli, F. Iacoviello, P. Shearing, D. Norman, M. Williams, A. Olivo and M. Endrizzi. X-ray dark-field tomography using edge-illumination. In *Developments in X-Ray Tomography XIV*, volume 12242. SPIE, 2022.

[82] S. Abbas, J. Min and S. Cho. Super-sparsely view-sampled cone-beam CT by incorporating prior data. *J. Xray Science and Technology*, 21(1):71–83, 2013.

[83] H. Zhang, D. Zeng, H. Zhang, J. Wang, Z. Liang and J. Ma. Applications of nonlocal means algorithm in low-dose X-ray CT image processing and reconstruction: A review. *Med. Phys.*, 44(3):1168–1185, 2017.

[84] J. F. P. J. Abascal, M. Abella, C. Mory, N. Ducros, C. de Molina, E. Marinetto, F. Peyrin and M. Desco. Sparse reconstruction methods in X-ray CT. *Proc. SPIE 10391*, 1039112, 2017.

[85] N. Mitroglou, M. Lorenzi, M. Santini and M. Gavaises. Application of X-ray micro-computed tomography on high-speed cavitating diesel fuel flows. *Exp. Fluids*, 57(11):175, 2016.

[86] Z. Purisha, S. S. Karhula, J. H. Ketola, J. Rimpelainen, M. T. Nieminen, S. Saarakkala, H. Kroger and S. Siltanen. An Automatic Regularization Method: An Application for 3-D X-Ray Micro-CT Reconstruction Using Sparse Data. *IEEE Trans. Med. Imaging*, 38(2):417–425, 2019.

[87] M. Guilizzoni, M. Santini, M. Lorenzi, V. Knisel and S. Fest-Santini. Micro computed tomography and CFD simulation of drop deposition on gas diffusion layers. *J. Phys. Conf. Ser.*, 547(1):012028, 2014.

[88] M. Santini and M. Guilizzoni. 3D X-ray micro computed tomography on multiphase drop interfaces: from biomimetic to functional applications. *Colloid Interface Sci. Commun.*, 1:14–17, 2014.

[89] M. Santini, M. Guilizzoni, S. Fest-Santini and M. Lorenzi. A novel technique for investigation of complete and partial anisotropic wetting on structured surface by X-ray microtomography. *Rev. of Sci. Instrum.*, 86(2):023708, 2015.

[90] L. F. A. Pereira, E. Janssens, G. D. C. Cavalcanti, I. R. Tsang, M. Van Dael, P. Verboven, B. Nicolai and J. Sijbers. Inline discrete tomography system: application to agricultural product inspection. *Comput. Electron. Agric.*, 138:117–126, 2017.

[91] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado and G. Van Tendeloo. 3D imaging of nanomaterials by discrete tomography. *Ultramicroscopy*, 109(6):730–740, 2009.

[92] S. Bals, K. J. Batenburg, D. Liang, O. Lebedev, G. Van Tendeloo, A. Aerts, J. A. Martens and C. E. Kirschhock. Quantitative three-dimensional modeling of zeotile through discrete electron tomography. *Journal of the American Chemical Society*, 131(13):4769–4773, 2009.

[93] T. Roelandts, K. J. Batenburg, E. Biermans, C. Kuebel, S. Bals and J. Sijbers. Accurate segmentation of dense nanoparticles by partially discrete electron tomography. *Ultramicroscopy*, 114:96–105, 2012.

[94] M. Lee, S. Shin and Y. Park. Reconstructions of refractive index tomograms via a discrete algebraic reconstruction technique. *Optics Express*, 25(22):27415–27430, 2017.

[95] H. Segers, W. J. Palenstijn, K. J. Batenburg and J. Sijbers. Discrete Tomography in MRI: a Simulation Study. *Fundam. Inform.*, 125(3-4):223–237, 2013.

[96] F. Bleichrodt, F. Tabak and K. J. Batenburg. SDART: an algorithm for discrete tomography from noisy projections. *Computer Vision and Image Understanding*, 129:63–74, 2014.

[97] F. Yang, D. Zhang, K. Huang, Z. Gao and Y. Yang. Incomplete projection reconstruction of computed tomography based on the modified discrete algebraic reconstruction technique. *Measurement Science and Technology*, 29(2):025405, 2018.

[98] J. Liu, Z. Liang, Y. Guan, W. Wei, H. Bai, L. Chen, G. Liu and Y. Tian. A modified discrete tomography for improving the reconstruction of unknown multi-gray-level material in the missing wedge situation. *Journal of Synchrotron Radiation*, 25(6), 2018.

[99] R. Pua, M. Park, S. Wi and S. Cho. A pseudo-discrete algebraic reconstruction technique (PDART) prior image-based suppression of high density artifacts in computed tomography. *Nucl. Instrum. Methods Phys. Res. A*, 840:42–50, 2016.

[100] W. van Aarle, K. J. Batenburg and J. Sijbers. Automatic parameter estimation for the discrete algebraic reconstruction technique (DART). *IEEE Trans. Image Process.*, 21(11):4608–4621, 2012.

[101] F. J. Maestre-Deusto, G. Scavello, J. Pizarro and P. L. Galindo. ADART: an adaptive algebraic reconstruction algorithm for discrete tomography. *IEEE Trans. Image Process.*, 20(8):2146–2152, 2011.

[102] A. Dabravolski, K. J. Batenburg and J. Sijbers. A multiresolution approach to discrete tomography using DART. *PLoS One*, 9(9):1–10, 2014.

[103] L. Brabant, M. Dierick, E. Pauwels, M. Boone and L. Van Hoorebeke. EDART, a discrete algebraic reconstructing technique for experimental data obtained with high resolution computed tomography. *Journal of X-ray Science and Technology*, 22(1):47–61, 2014.

[104] X. Zhuge, W. J. Palenstijn and K. J. Batenburg. TVR-DART: a more robust algorithm for discrete tomography from limited projection data with automated gray value estimation. *IEEE Trans. Image Process.*, 25(1):455–468, 2016.

[105] T. Humphries, J. Winn and A. Faridani. Superiorized algorithm for reconstruction of CT images from sparse-view and limited-angle polyenergetic data. *Phys. Med. Biol.*, 62(16):6762–6783, 2017.

[106] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assié, D. Autret, S. Avner, R. Barbier, M. Bardiès, P. M. Bloomfield, D. Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A. F. Chatziioannou, Y. Choi, Y. H. Chung, C. Comtat, D. Donnarieix, L. Ferrer, S. J. Glick, C. J. Groiselle, D. Guez, P. F. Honore, S. Kerhoas-Cavata, A. S. Kirov, V. Kohli, M. Koole, M. Krieguer, D. J. van der Laan, F. Lamare, G. Largeron, C. Lartizien, D. Lazaro, M. C. Maas, L. Maigne, F. Mayet, F. Melot, C. Merheb, E. Pennacchio, J. Perez, U. Pietrzyk, F. R. Rannou, M. Rey, D. R. Schaart, C. R. Schmidtlein, L. Simon, T. Y. Song, J. M. Vieira, D. Visvikis, R. Van de Walle, E. Wieërs and C. Morel. GATE: a simulation

toolkit for PET and SPECT. *Phys. Med. Biol.*, 49(19):4543–4561, 2004.

[107] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg and J. Sijbers. Fast and flexible X-ray tomography using the ASTRA toolbox. *Optics Express*, 24(22):25129–25147, 2016.

[108] G. G. Poludniowski, G. Landry, F. DeBlois, P. M. Evans and F. Verhaegen. SpekCalc: a program to calculate photon spectra from tungsten anode X-ray tubes. *Phys. Med. Biol.*, 54(19):N433, 2009.

[109] G. Van Gompel, K. Van Slambrouck, M. Defrise, K. J. Batenburg, J. de Mey, J. Sijbers and J. Nuyts. Iterative correction of beam hardening artifacts in CT. *Med. Phys.*, 38(1):36–49, 2011.

[110] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man. Cybern.*, 9(1):62–66, 1979.

[111] P. C. Hansen, J. Jørgensen and W. R. Lionheart. *Computed Tomography: Algorithms, Insight, and Just Enough Theory.* SIAM, Philadelphia, 2021.

[112] B. S. van Lith, P. C. Hansen and M. E. Hochstenbach. A twin error gauge for Kaczmarz's iterations. *SIAM Journal on Scientific Computing*, 43(5):S173–S199, 2021.

[113] A. Perelli, M. Lexa, A. Can and M. E. Davies. Compressive Computed Tomography Reconstruction through Denoising Approximate Message Passing. *SIAM Journal on Imaging Sciences*, 13(4):1860–1897, 2020.

[114] C. Wang, M. Tao, J. G. Nagy and Y. Lou. Limited-Angle CT Reconstruction via the L_1/L_2 Minimization. *SIAM Journal on Imaging Sciences*, 14(2):749–777, 2021.

[115] T. Lukić and P. Balázs. Limited-view binary tomography reconstruction assisted by shape centroid. *The Visual Computer*, 38(2):695–705, 2022.

[116] P. Gritzmann, S. De Vries and M. Wiegelmann. Approximating binary images from discrete X-rays. *SIAM Journal on Optimization*, 11(2):522–546, 2000.

[117] A. Stolk and K. J. Batenburg. An algebraic framework for discrete tomography: revealing the structure of dependencies. *SIAM Journal on Discrete Mathematics*, 24(3):1056–1079, 2010.

[118] P. Balázs. A decomposition technique for reconstructing discrete sets from four projections. *Image and Vision Computing*, 25(10):1609–1619, 2007.

[119] T. Capricelli and P. Combettes. A convex programming algorithm for noisy discrete tomography. In *Advances in discrete tomography and its applications*, pages 207–226. Springer, Berlin, 2007.

[120] M. Zisler, J. H. Kappes, C. Schnörr, S. Petra and C. Schnörr. Non-binary discrete tomography by continuous non-convex optimization. *IEEE Transactions on Computational Imaging*, 2(3):335–347, 2016.

[121] T. Sanders. Discrete iterative partial segmentation technique (DIPS) for tomographic reconstruction. *IEEE Transactions on Computational Imaging*, 2(1):71–82, 2016.

[122] A. Kadu, T. van Leeuwen and K. J. Batenburg. A parametric level-set method for partially discrete tomography. In *International Conference on Discrete Geometry for Computer Imagery*, pages 122–134, 2017.

[123] N. Six, J. De Beenhouwer and J. Sijbers. Poly-dart: a discrete algebraic reconstruction technique for polychromatic x-ray ct. *Optics Express*, 27(23):33670–33682, 2019.

[124] M. Zeegers, F. Lucka and K. J. Batenburg. A multi-channel DART algorithm. In *International Workshop on Combinatorial Image Analysis*, pages 164–178, 2018.

[125] K. Batenburg. Network flow algorithms for discrete tomography. In *Advances in discrete tomography and its applications*, pages 175–205. Springer, Berlin, 2007.

[126] B. Goris, T. Roelandts, K. Batenburg, H. H. Mezerji and S. Bals. Advanced reconstruction algorithms for electron tomography: from comparison to combination. *Ultramicroscopy*, 127:40–47, 2013.

[127] A. Tuysuzoglu, W. C. Karl, I. Stojanovic, D. Castañòn and M. S. Ünlü. Graph-cut based discrete-valued image reconstruction. *IEEE Transactions on Image Processing*, 24(5):1614–1627, 2015.

[128] Y. Guo, R. Aveyard and B. Rieger. A multichannel cross-modal fusion framework for electron tomography. *IEEE Transactions on Image Processing*, 28(9):4206–4218, 2019.

[129] Y. Zhao, J. Xu, H. Li and P. Zhang. Edge information diffusion-based reconstruction for cone beam computed laminography. *IEEE Transactions on Image Processing*, 27(9):4663–4675, 2018.

[130] Z. Wei, B. Liu, B. Dong and L. Wei. A Joint Reconstruction and Segmentation Method for Limited-Angle X-Ray Tomography. *IEEE Access*, 6:7780–7791, 2018.

[131] F. Yang, D. Zhang, K. Huang, Z. Gao and Y. Yang. Incomplete projection reconstruction of computed tomography based on the modified discrete algebraic reconstruction technique. *Measurement Science and Technology*, 29(2):025405, 2018.

[132] J. Liu, Z. Liang, Y. Guan, W. Wei, H. Bai, L. Chen, G. Liu and Y. Tian. A modified discrete tomography for improving the reconstruction of unknown multi-gray-level material in the missing wedge situation. *Journal of Synchrotron Radiation*, 25(6):1847–1859, 2018.

[133] M. AnanthaLakshmi, G. Yamuna and A. SanjeeviKumar. A novel method of 3d image reconstruction using aco-based tvr-dart. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, 12(5):1–10, 2021.

[134] D. Frenkel, J. Beenhouwer and J. Sijbers. An adaptive probability map for the Discrete Algebraic Reconstruction Technique. In *10th Conference on Industrial Computed Tomography (iCT),(iCT 2020) Wels, Austria*, 2020.

[135] D. Frenkel, J. De Beenhouwer and J. Sijbers. Tabu-DART: a dynamic update strategy for the Discrete Algebraic Reconstruction Technique based on tabu-search. In *Proceedings of the 16th Virtual International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine, 19-23 July, 2021, Leuven, Belgium*, pages 173–177, 2021.

[136] P. Miklós. Discrete tomographic reconstruction of binary matrices using Tabu search and classic Ryser algorithm. In *2011 IEEE 9th International Symposium on Intelligent Systems and Informatics*, pages 387–390, 2011.

[137] P. Miklós. Tabu search reconstruction of hv-convex binary contours using classic Ryser algorithm and smart switching. In *2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI)*, pages 341–344, 2011.

[138] G. Lékó, S. Domány and P. Balázs. Uncertainty based adaptive projection selection strategy for binary tomographic reconstruction. In *International Conference on Computer Analysis of Images and Patterns*, pages 74–84, 2019.

[139] L. G. Varga, G. Lékó and P. Balázs. Grayscale uncertainty and errors of tomographic reconstructions based on projection geometries and projection sets. *The Visual Computer*:1–13, 2022.

[140] F. Glover and M. Laguna. Tabu search: effective strategies for hard problems in analytics and computational science. *Handbook of Combinatorial Optimization*, 21:3261–3362, 2013.

[141] L. G. Varga, L. G. Nyúl, A. Nagy and P. Balázs. Local and global uncertainty in binary tomographic reconstruction. *Computer Vision and Image Understanding*, 129:52–62, 2014.

[142] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg and J. Sijbers. The ASTRA Toolbox: a platform for advanced algorithm development in electron tomography. *Ultramicroscopy*, 157:35–47, 2015.

[143] T. Kohler. A projection access scheme for iterative reconstruction based on the golden section. In *IEEE Symposium Conference Record Nuclear Science 2004.* Volume 6, pages 3961–3965, 2004.

[144] R. Hunger. *Floating point operations in matrix-vector calculus.* Munich University of Technology, Inst. for Circuit Theory and Signal, Munich, 2005.

[145] T. Roelandts, K. J. Batenburg, A. J. den Dekker and J. Sijbers. The reconstructed residual error: A novel segmentation evaluation measure for reconstructed images in tomography. *Computer Vision and Image Understanding*, 126:28–37, 2014.

[146] R. E. Alvarez and A. Macovski. Energy-selective reconstructions in X-ray computerised tomography. *Physics in Medicine & Biology*, 21(5):733, 1976.

[147] W. A. Kalender, D. Felsenberg, H. K. Genant, M. Fischer, J. Dequeker and J. Reeve. The european spine phantom—a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. *European Journal of Radiology*, 20(2):83–92, 1995.

[148] K. Remeysen and R. Swennen. Beam hardening artifact reduction in microfocus computed tomography for improved quantitative coal characterization. *International Journal of Coal Geology*, 67(1-2):101–111, 2006.

[149] R. B. Mazess, H. S. Barden, J. P. Bisek and J. Hanson. Dual-energy X-ray absorptiometry for total-body and regional bone-mineral and soft-tissue composition. *The American Journal of Clinical Nutrition*, 51(6):1106–1112, 1990.

[150] B. De Man, J. Nuyts, P. Dupont, G. Marchal and P. Suetens. An iterative maximum-likelihood polychromatic algorithm for ct. *IEEE transactions on medical imaging*, 20(10):999–1008, 2001.

[151] S. Srivastava and J. A. Fessler. Simplified statistical image reconstruction for polyenergetic X-ray CT. In *IEEE Nuclear Science Symposium Conference Record, 2005*, volume 3, 5–pp. IEEE, 2005.

[152] R. H. Byrd, P. Lu, J. Nocedal and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995.

[153] R. Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2013.

[154] N. Six, J. Renders, J. Sijbers and J. De Beenhouwer. Newton-Krylov Methods for Polychromatic X-ray CT. In *2020 IEEE International Conference on Image Processing (ICIP)*, pages 3045–3049. IEEE, 2020.

[155] T. Humphries and A. Faidani. Segmentation-free quasi-newton method for polyenergetic CT reconstruction. In *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–5. IEEE, 2014.

[156] N. Six. GNK for polyCT. *Zenodo*, 2021. URL: `https://doi.org/10.5281/zenodo.5721273`.

[157] E. Y. Sidky, L. Yu, X. Pan, Y. Zou and M. Vannier. A robust method of X-ray source spectrum estimation from transmission measurements: demonstrated on computer simulated, scatter-free transmission data. *Journal of Applied Physics*, 97(12):124701, 2005.

[158] M. Endrizzi. X-ray phase-contrast imaging. *Nuclear Instruments & Methods in Physics Research A*, 878:88–98, 2018.

[159] A. Ritter, F. Bayer, J. Durst, K. Gödel, W. Haas, T. Michel, J. Rieger, T. Weber, L. Wucherer and G. Anton. Simultaneous maximum-likelihood reconstruction for x-ray grating based phase-contrast tomography avoiding intermediate phase retrieval. *arXiv preprint arXiv:1307.7912*, 2013.

[160] B. Brendel, M. von Teuffenbach, P. B. Noël, F. Pfeiffer and T. Koehler. Penalized maximum likelihood reconstruction for x-ray differential phase-contrast tomography. *Medical physics*, 43(1):188–194, 2016.

[161] N. Schenkels, J. Sijbers, W. van Aarle and W. Vanroose. A generalized bidiagonal-tikhonov method applied to differential phase contrast tomography. *arXiv preprint arXiv:1510.03233*, 2015.

[162] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.

[163] C. K. Hagen, F. A. Vittoria, O. R. i Morgó, M. Endrizzi and A. Olivo. Cycloidal computed tomography. *Physical Review Applied*, 14(1):014069, 2020.

[164] A. Doherty, S. Savvidis, C. Navarrete-León, A. Olivo, M. Endrizzi and M. F. Gerli. Edge-Illumination X-Ray Dark-Field Tomography. *Physical Review Applied*, 19(5):054042, 2023.

[165] F. A. Vittoria, G. K. Kallon, D. Basta, P. C. Diemoz, I. K. Robinson, A. Olivo and M. Endrizzi. Beam tracking approach for single–shot retrieval of absorption, refraction, and dark–field signals with laboratory x–ray sources. *Applied Physics Letters*, 106(22), 2015.