

*Practical Approaches to Reconstruction and
Analysis for 3D and Dynamic 3D Computed
Tomography*

Coban, Sophia Bethany

2017

MIMS EPrint: **2017.44**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

PRACTICAL APPROACHES TO
RECONSTRUCTION AND ANALYSIS
FOR 3D AND DYNAMIC 3D
COMPUTED TOMOGRAPHY

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF SCIENCE AND ENGINEERING

2017

Sophia Bethany Coban
School of Mathematics

Contents

Abstract	4
Declaration	6
Copyright Statement	7
Acknowledgements	8
Author Contributions	10
1 Introduction	13
1.1 The X-ray CT Problem	14
1.2 Tomographic image reconstruction	16
1.2.1 Analytic reconstruction	17
1.2.2 Algebraic reconstruction	19
1.3 Format and Outline of the Thesis	23
1.3.1 Image analysis for practical CT	24
1.3.2 Compressive sensing in practical sparse tomography	26
1.3.3 Reconstructing a non-linear CT problem	27
1.3.4 Fast data acquisition and dynamic 3D reconstruction in syn- chrotron X-ray CT	28
1.4 Suggested further work	30
2 Assessing the suitability of computed tomography reconstruction meth- ods through physical quantification techniques	37
3 Predicting performance of sparsity-regularized X-ray tomography:	

experimental results using glass-bead data	64
4 Non-linearity in monochromatic transmission tomography	83
5 4-D imaging of sub-second dynamics in pore-scale processes using real-time synchrotron X-ray tomography	102
A SophiaBeads Datasets Project Documentation and Tutorials	118
B When do the iterative reconstruction methods become worth the effort?	141
C Effect of sparsity and exposure on total variation regularized X-ray tomography from few projections	146

The University of Manchester

Sophia Bethany Coban

Doctor of Philosophy

Practical approaches to reconstruction and analysis for 3D and dynamic 3D computed tomography

February 7, 2017

The problem of reconstructing an image from a set of tomographic data is not new, nor is it lacking attention. However there is still a distinct gap between the mathematicians and the experimental scientists working in the computed tomography (CT) imaging community. One of the aims in this thesis is to bridge this gap with mathematical reconstruction algorithms and analysis approaches applied to practical CT problems.

The thesis begins with an extensive analysis for assessing the suitability of reconstruction algorithms for a given problem. The paper presented examines the idea of extracting physical information from a reconstructed sample and comparing against the known sample characteristics to determine the accuracy of a reconstructed volume. Various test cases are studied, which are relevant to both mathematicians and experimental scientists. These include the variance in quality of reconstructed volume as the dose is reduced or the implementation of the level set evolution method, used as part of a simultaneous reconstruction and segmentation technique. The work shows that the assessment of physical attributes results in more accurate conclusions. Furthermore, this approach allows for further analysis into interesting questions in CT. This theme is continued throughout the thesis.

Recent results in compressive sensing (CS) gained attention in the CT community as they indicate the possibility of obtaining an accurate reconstruction of a sparse image from severely limited or reduced amount of measured data. Literature produced so far has not shown that CS directly guarantees a successful recovery in X-ray CT, and it is still unclear under which conditions a successful sparsity regularized reconstruction can be achieved. The work presented in the thesis aims to answer this question in a practical setting, and seeks to establish a direct connection between the success of sparsity regularization methods and the sparsity level of the image, which is similar to CS. Using this connection, one can determine the sufficient amount of measurements to collect from just the sparsity of an image. A link was found in a previous study using simulated data, and the work is repeated here with experimental data, where the sparsity level of the scanned object varies. The preliminary work presented here verifies the results from simulated data, showing an “almost-linear” relationship between the sparsity of the image and the sufficient amount of data for a successful sparsity regularized reconstruction.

Several unexplained artefacts are noted in the literature as the ‘partial volume’, the ‘exponential edge gradient’ or the ‘penumbra’ effect, with no clear explanation for their cause, or established techniques to remove them. The work presented in this paper shows that these artefacts are due to a non-linearity in the measured data, which comes from either the set up of the system, the scattering of rays or the dependency of linear attenuation on wavelength in the polychromatic case. However, even in monochromatic CT systems, the non-linearity effect can be detected. The paper shows that in some cases, the non-linearity effect is too large to ignore, and the reconstruction

problem should be adapted to solve a non-linear problem. We derive this non-linear problem and solve it using a numerical optimization technique for both simulated and real, γ -ray data. When compared to reconstructions obtained using the standard linear model, the non-linear reconstructed images show clear improvements in that the non-linear effect is largely eliminated.

The thesis is finished with a highlight article in the special issue of Solid Earth, named “*Pore-scale tomography & imaging - applications, techniques and recommended practice*”. The paper presents a major technical advancement in a dynamic 3D CT data acquisition, where the latest hardware and optimal data acquisition plan are applied and as a result, ultra fast 3D volume acquisition was made possible. The experiment comprised of fast, free-falling water-saline drops traveling through a pack of rock grains with varying porosities. The imaging work was enhanced by the use of iterative methods and physical quantification analysis performed. The data acquisition and imaging work is the first in the field to capture a free falling drop and the imaging work clearly shows the fluid interaction with speed, gravity and more importantly, the inter- and intra-grain fluid transfers.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

- i.** The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii.** Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii.** The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv.** Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s Policy on Presentation of Theses.

Acknowledgements

I would like to express my greatest appreciation to Prof. William R.B. Lionheart and Prof. Philip J. Withers, my research supervisors, for their patient guidance, enthusiastic encouragement and their persistent confidence in me and my abilities.

Further, I would like to thank Dr. Martin Lotz and Prof. Joost K. Batenburg, my examiners, for a wonderful discussion and their useful critiques that have finalized this thesis.

In addition I would like to offer my special thanks to Dr. Samuel A. McDonald, Dr. Katherine J. Dobson, Dr. Robert C. Atwood, Dr. Jakob S. Jørgensen, Dr. Daniil Kazantsev, and Dr. Ilker Meric, my co-authors and colleagues, for their willingness to give their time so generously, sharing their invaluable experiences, and making their support available any time.

I want to thank BP and the School of Mathematics for the financial support, and in particular, acknowledge my BP mentor Dr. Tom Lynch, BP-ICAM programme manager Sheetal Handa, my colleagues in the Inverse Problems Group, Henry Moseley X-Ray Imaging Facility and the BP-ICAM team, who have played minor but very important roles in my Ph.D. experience, and contributed greatly to my development as a researcher.

I would also like to thank my wonderful friends Bahar Arslan, Aizhan Maden, Francis Watson, Michael Crabb, Russel Miller, Zainab Alsoufi and Naeem Desai for their endless support and encouragement throughout, and Roy Lindon, my A-Level maths teacher, for making me fall in love with this beautiful subject.

Lastly, I would like to dedicate this work to my wonderful mother and my partner for their infinite patience, encouragement and unconditional love. Without their support and warm hugs, none of this would have been possible.

For Love and the Beloved.

Author Contributions

Following are the contributions done by all authors in the published papers in Chapters 2-5, and additional material given in Appendices 1-3, in the numbered order.

Assessing the suitability of a reconstruction method with physical quantification techniques

Authors: Sophia B. Coban, William R. B. Lionheart, Philip J. Withers.

The materials in this paper were produced by SBC, with contributions to and comments on the text done by WRBL and PJW.

Predicting performance of sparsity-regularized X-ray tomography: experimental results using glass-bead data

Authors: Jakob S. Jørgensen, Sophia B. Coban, William R. B. Lionheart, Samuel A. McDonald, Philip J. Withers.

Extended work from [26]. Data acquisition experiments for new sparse beads data were designed by JSJ and SBC, and collection of data done by SBC, JSJ with the guidance of SAM. Materials (reconstruction and quantification) presented in paper are produced by JSJ and SBC. Paper is written in conjunction with JSJ and SBC, with WRBL and PJW contributing with comments to the written text.

Non-linearity in monochromatic transmission tomography

Authors: William R. B. Lionheart, Bjørn Tore Hjertaker, Rachid Maad, Ilker Meric, Sophia B. Coban, Geir Anton Johansen.

Idea is proposed by WRBL. The γ -ray system was built by GAJ and BTH. Data acquisition experiments designed and performed by WRBL, RM and BTH. Monte Carlo simulations were performed by IM as per the designed experiment plan. Theoretical

work is done by WRBL with contributions from SBC. Mathematical formulations were produced by SBC and WRBL. Development and implementation of the optimization approach were done by SBC. All authors contributed to the manuscript with WRBL leading the project.

4-D imaging of sub-second dynamics in pore-scale processes using real-time synchrotron X-ray tomography

Authors: Katherine J. Dobson, Sophia B. Coban, Samuel A. McDonald, Joanna N. Walsh, Robert C. Atwood, Philip J. Withers.

Data acquisition experiments were designed by KJD, SAM and RCA. Data set acquisition was performed by KJD, SAM, SBC, JNW and RCA. Reconstruction and quantification analysis were done by SBC and KJD. Text produced by KJD and SBC with important feedback from PJW and RCA.

SophiaBeads Dataset Project Documentation and Tutorials

Author: Sophia B. Coban

Support material for SophiaBeads Dataset Project [8]. The report is produced by SBC.

When do the iterative methods become worth the effort?

Authors: Sophia B. Coban, Samuel A. McDonald, William R. B. Lionheart, Philip J. Withers.

Conference paper produced by SBC with data acquisition experiment designed by PJW and SBC, and collected by SBC and SAM. Major contributions to the text were made by PJW and WRBL.

Effect of sparsity and exposure on total variation regularized X-ray tomography from few projections

Authors: Jakob S. Jørgensen, Sophia B. Coban, William R. B. Lionheart, Philip J. Withers.

Idea for this research was proposed by JSJ (continuation from previous work). Work adapted to be applied to the SophiaBeads Dataset Project by SBC and JSJ.

Material in the conference paper is produced both by JSJ and SBC with feedback provided by WRBL and PJW.

Chapter 1

Introduction

Computed tomography (CT), an important tool particularly in medicine and material science, is a non-invasive technique that produces reliable information via a set of measurements (projections or radiographs). This information is then used to reconstruct a cross-sectional (2D) or a volumetric (3D) image in order to study the interior of a domain. The tomographic imaging problem (henceforth the CT problem) can be described as the recovery of information, and is therefore an inverse problem.

Inverse problems are often described as ill-posed problems, namely a problem that fails one or more of Hadamard's conditions for a well-posed problem [17]. These are

1. A solution to the problem must exist;
2. The solution to the problem must be unique;
3. The solution depends continuously on the data.

We consider the continuum and discrete CT problem separately. In the continuum case, we can have a unique solution given the *right* data, but the problem can still be unstable in a reasonable norm, meaning a severely ill-posed problem. The discrete case however cannot produce a unique solution due the finite set of data yet with infinite number of unknowns, but it can be stable and therefore mildly ill-conditioned. In this thesis, we focus on discrete CT problems. In addition we know that small changes in the measured data can lead to big changes in the reconstructed image, even in the case of sufficient data in circular geometry [15]. This implies that the discrete CT problem does not satisfy Hadamard's uniqueness and stability conditions (here, we

refer to the CT problem without the use of regularization). We note here that there needs to be a shift in our tradition of working with inverse problems in the way that once we establish a problem is ill-posed, we must then consider the *golden rules* of inverse problems [30] before attempting to solve the problem. The golden rules of inverse problems are as follows. Before solving any inverse problem we must

- I. first, assess the data to be measured;
- II. include any prior information about the problem, and
- III. finally, assess various methods for solving the problem with a clear solution assessment defined.

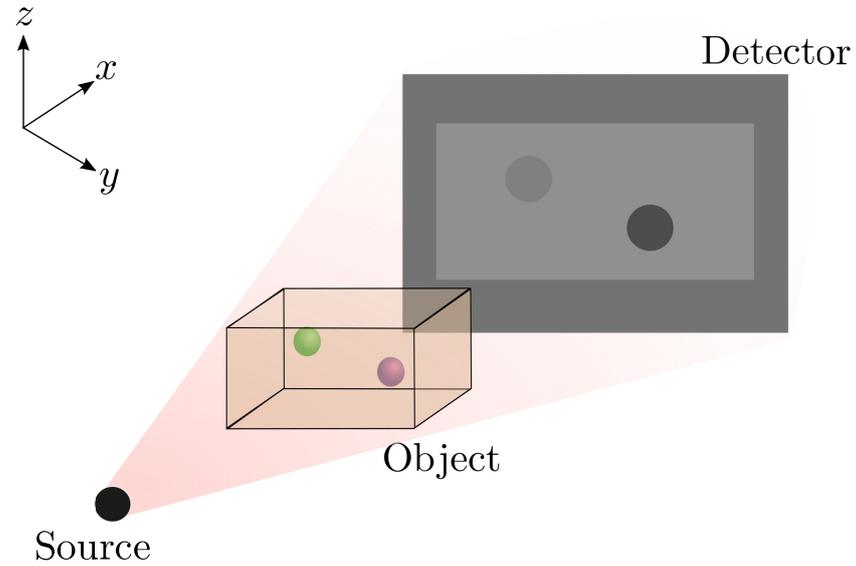
It is worth mentioning that these rules are often ignored or forgotten, while they are extremely important when solving real world problems.

Despite being an ill-posed problem, the discrete CT problem is considered to be mildly ill-conditioned, and many convex optimization and algebraic methods for well-posed problems are still viable through the use of regularization. We introduce regularization in CT later in this chapter, as well as providing a preface for each chapter in the thesis. We begin by introducing the problem itself, and common image reconstruction techniques used in practical CT today. We now continue with the underlying physics and the derivation of the CT model.

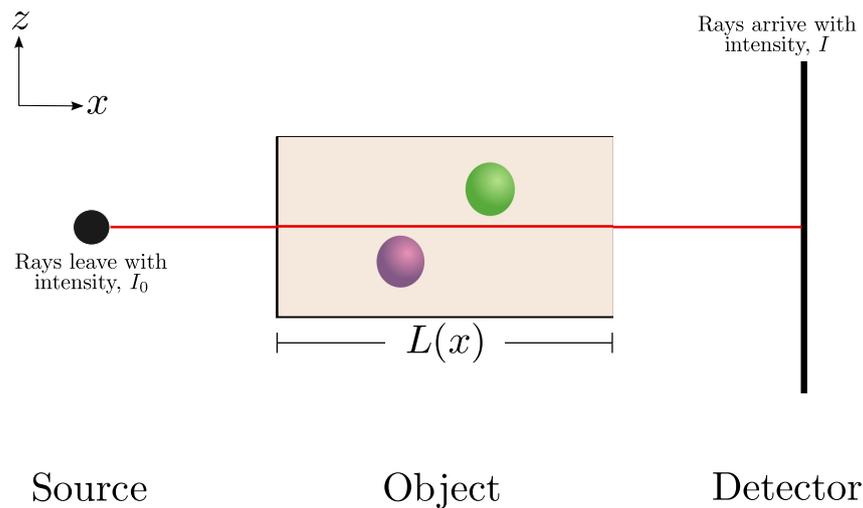
1.1 The X-ray CT Problem

Consider the simple apparatus geometry given in Figure 1.1a. In this set up, we have a point source that produces and projects X-rays onto an object with a specified energy level (total photon count or intensity). These X-rays travel through the object, and while interacting with the material, lose intensity due to scattering or absorption of photons. In technical terms, the rays get attenuated. The weakened rays reach the detector panel (consisting of many detector elements spread over a square or a rectangular grid), which records the intensity of the arriving X-ray beams. Our aim is to recover the attenuation coefficient, which is related to the material density or densities within the object in order to create a density map of the interior. During a typical scan, this process is repeated over many source positions relative to the object;

and each set of data collected at particular source position is called a projection (or radiograph in medical CT).



(a) 3D illustration of a simple apparatus set up for cone beam.



(b) 2D (xz -)view of a single ray from a point source with the intensities I_0 at the source and I at the detector.

Figure 1.1: Simple illustration of a cone beam geometry in a typical CT scanner.

The mathematical model of the CT problem is fairly straightforward. Consider the same set up, except this time we focus on a single point in a monochromatic (single energy) X-ray beam, as shown in Figure 1.1b. Let $L(x)$ be the length of the X-ray beam travelling through the object, dI be the intensity value at the point dx_L . We know that at this point dx_L , the beam has lost some intensity proportional to its initial intensity, I_0 ,

$$dI = -\mu(x)I_0 dx_L.$$

Here, $\mu(x)$ is the attenuation coefficient at location x , i.e. the rate at which the photons are absorbed or scattered. This is rearranged to get

$$\frac{dI}{I_0} = -\mu(x) dx_L.$$

Recall that this equation gives the intensity value only at one location in the X-ray beam. To account for the full beam, we have to take the sum of all the points on the line, i.e. integrate over L . This is given by

$$\int_{L(x)} \frac{dI}{I_0} = - \int_{L(x)} \mu dx_L \implies \ln \frac{I}{I_0} = - \int_{L(x)} \mu(x) dx_L,$$

which implies

$$I = I_0 e^{-\int_{L(x)} \mu(x) dx_L}. \quad (1.1)$$

In physical terms, (1.1) describes that the monochromatic beam is attenuated *exponentially* as it travels through an object. This is the result found by Lambert [29] and later independently by Beer [3]. In optics and signal processing, this is commonly referred to as the Beer-Lambert Law.

Most X-ray sources produce a polychromatic beam (beam with a range of energies), which means the attenuation coefficient depends on energy, E at beam position x ,

$$I = \int_{L(x)} I_0(E) e^{-\int_{L(x)} \mu(x,E) dx_L} dE. \quad (1.2)$$

Mathematically, the goal of X-ray CT problem is to recover the attenuation coefficient, μ , from the information at detectors, I .

1.2 Tomographic image reconstruction

The tomographic reconstruction methods used in the CT community can be separated into two groups: analytic reconstruction via direct methods, and algebraic reconstruction via iterative and statistical methods. We must note here that this does not mean

algebraic reconstruction equates iterative or statistical methods and in theory, one can use direct methods for algebraic reconstruction given the required computational power. However with the amount of data produced in real life experiments for the type of CT problems we look at, direct methods are not suitable for algebraic reconstruction. Therefore we focus on iterative and statistical methods. We briefly explain the formulation for both analytical and algebraic reconstruction, and summarize the popular direct and iterative methods.

Note that the work presented in this thesis does not include any statistical methods and therefore are not introduced in this chapter. However we refer the readers to the excellent review and analysis of statistical reconstruction methods in [14].

1.2.1 Analytic reconstruction

In his 1917 paper [37], Johann Radon invented a transform that takes a function $f(\mathbf{x})$, $\mathbf{x} = (x, y)$, defined on a plane to a function $Rf(s, \theta)$ defined on a two-dimensional space of lines, which are parameterized by distance to the origin, s and the angle θ of the normal. This is called the *classical Radon transform*. In the context of tomographic imaging, if a function f represents an unknown density in the pixel (image) domain, then the Radon transform is the projection data collected as the output of the tomographic scan, illustrated in Figure 1.2. Recall from (1.1) that at any point x in a monochromatic beam with length L , we have

$$-\log\left(\frac{I}{I_0}\right) = \int_{L(\mathbf{x})} \mu(\mathbf{x}) d\mathbf{x}_L,$$

the right-hand side of which is the definition of a Radon transform:

$$Rf(s, \theta) := \int_L f(\mathbf{x}) d\mathbf{x}.$$

Johann Radon also provided the inversion of the transform in the same paper. This work, which is the very backbone of image reconstruction, was first produced for pure mathematics and did not get any attention from the applied mathematics community. Allan Cormack reinvented the Radon transform in 1963, [11] to be used in reconstruction of a density of a body from a collection of X-ray images taken from different directions. In 1979, Cormack won the Nobel Prize in Medicine in conjunction with George Hounsfield, who independently derived a reconstruction method and built

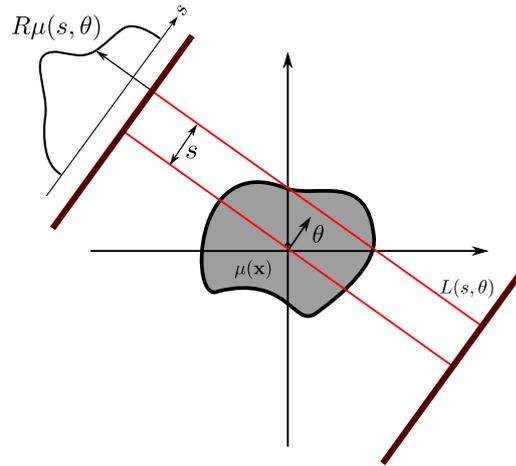


Figure 1.2: A diagram to illustrate the Radon transform. Note that a single ray in the image domain is a point in the sinogram domain.

the first medical scanner [23] a few years after Cormack’s reinvention. The derivation of the reconstruction algorithm is given in cited work by Cormack, which we briefly introduce now.

Filtered back projection

In a nutshell, the filtered back projection (FBP) is a 2D image (cross-sectional slice) reconstruction method from a set of discrete projections (line integrals). The method involves two important steps: Applying a filter Λ (often ramp filter) on the sinogram (measured data, where the filter is applied independently for each angle), $\Lambda R\mu$; and backproject using the adjoint of Radon transform, $R^*(\Lambda R\mu)$. The formula for FBP is given by

$$\mu(\mathbf{x}) = \frac{1}{4\pi} R^*(\Lambda R\mu)(\mathbf{x}). \quad (1.3)$$

We refer the interested reader to [34, 21] for the detailed derivation of the filtered back projection.

The FBP algorithm in theory should give a perfect reconstruction in the case of infinite number of one-dimensional projections of an object taken at an infinite number of angles. This is however impossible in practice and so the standard protocol applied in laboratory and synchrotron CT systems is to take as many projections as the system possibly can store in order to have a ‘sufficiently noiseless’ image (often determined by eye). This is because the inverse Radon transform can be extremely

unstable when applied to noisy data. In addition, FDK is not an exact inverse of the Radon transform, thus inviting more room for unstable reconstruction. Despite this, the FBP method and its 3D implementation for circular cone-beam, given by Feldkamp-Davis-Kress (FDK) [13], are arguably the most popular methods in practical CT, still implemented in many modern commercial CT scanners, and the resulting reconstructions are commonly taken as the ‘gold standard’.

1.2.2 Algebraic reconstruction

The problem of image reconstruction can also be considered as a set of linear equations, which then can be solved iteratively. We emphasize here again that this does not imply a set of linear equations cannot be solved directly. However due to the computational demand to store collected measurements and the set of equations for them, direct methods are not an option for algebraic reconstruction and thus we focus on iterative methods.

Consider a single ray leaving the source and arriving at the detector in Figure 1.1b, which travels in a single line. Let b_i be the log-transformed projection data (the line integral or the ray-sum of the i^{th} ray). Also consider the image space discretized into N number of pixels (voxels in 3D), $x_j, j = 1, \dots, N$. Then the algebraic expression of b_i is given by the weighted sum of the pixels the i^{th} ray has travelled through, i.e.

$$b_i = \sum_{j=1}^N a_{ij}x_j, \quad \text{where,} \quad (1.4)$$

$$a_{ij} = \begin{cases} \beta(i, j), & \text{if the ray } i \text{ intersects with the pixel } j, \\ 0, & \text{otherwise.} \end{cases}$$

Here, the value of $\beta(i, j)$ is the intersection length of ray i in each pixel j , which is calculated by tracing the ray i through the image space, as demonstrated in Figure 1.3. If we let M be the total number of rays, then (1.4) can be written in matrix form as

$$Ax = b, \quad (1.5)$$

where $b \in \mathbb{R}^{M \times 1}$ is the negative log of the normalized projection data, $x \in \mathbb{R}^{N \times 1}$ the vector of unknown pixel values and $A \in \mathbb{R}^{M \times N}$ the sparse geometry or weight matrix of the system. In practice $M > N$, meaning (1.5) is an overdetermined system with

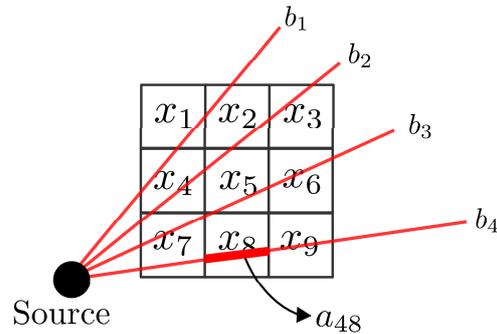


Figure 1.3: A diagram illustrating the discretization of the pixels and the projected rays in the image space.

no unique solution, and so the objective is to find an optimal set of parameters that satisfy a normal equation,

$$A^T Ax = A^T b. \quad (1.6)$$

This is done by minimizing the sum of square differences between the measured data b , and the predicted data Ax . This approach is referred to as least squares or data fitting in mathematics and statistics.

Often, M and N are too large that we are not able to store the geometry matrix A and must consider the computational cost of an algorithm when implementing new methods. Generally we are able to reduce the computational costs by performing matrix-vector calculations. Below we summarize the popular iterative methods used in practical CT.

Iterative methods

One of the oldest and most well-known iterative methods in CT is the Landweber iteration, which in its basic form is

$$x^k = x^{k-1} + \omega A^T (b - Ax^{k-1}), \quad k = 1, 2, \dots, \quad (1.7)$$

where x_k denotes the approximated solution vector at the k^{th} iteration and $0 < \omega < 2\|A^T A\|_2^{-1}$. Iterations continue until a stopping criterion is satisfied. We should note that the Landweber iteration is a special case of the gradient descent when the update is written in terms of the gradient, $x^k = x^{k-1} - \omega \nabla f(x^{k-1})$, where $\nabla f(x^k) = \|b - Ax^k\|_2^2/2$.

Another method that is similar to (1.7) is the simultaneous iterative reconstruction technique (SIRT), which has the iteration update defined as

$$x^k = x^{k-1} + C A^T D(b - Ax^{k-1}), \quad k = 1, 2, \dots,$$

where C and D are diagonal matrices with $c_{jj} = 1/\sum_i a_{ij}$ and $d_{ii} = 1/\sum_j a_{ij}$. In other words, these are the inverses of the sum of the columns (C) and the sum of the rows (D) of matrix A .

Another noteworthy method is the Kaczmarz method, better known in the CT community as the algebraic reconstruction technique (ART) [27, 21]. This method takes a ‘‘sweep’’ through the rows of A and updates the k^{th} iterated solution using

$$x^k = x^{k-1} + \lambda_k \frac{b_i - a_i^T x^k}{\|a_i\|_2^2} a_i, \quad k = 1, 2, \dots,$$

where a_i is the i^{th} row of A , b_i is the i^{th} component of b , λ_k is a relaxation parameter, and $\|a_i\|_2^2$ is precomputed.

An improved version of the ART is the simultaneous-ART (SART) [2], which is often used for limited data reconstructions. The difference between the classic and simultaneous-ART is that after all sweeps have been performed in the regular way, the iterated solution is further corrected by the average value of sweeps applied on each pixel.

The methods above (not including SART) exhibit semi-convergence when applied to ill-conditioned systems, meaning that after a certain point the methods fit the image to the noise rather than the measured data. The convergence of these methods is initially fast but becomes very slow in later iterations [28, 19]. We now introduce a type of iterative methods that are fast-converging and computationally inexpensive to implement for large systems, Krylov subspace methods.

Krylov subspace methods

For the purpose of introducing the Krylov subspace methods, let us first consider a square system with sparse geometry matrix, A . An order k Krylov subspace, $\mathcal{K}_k(A, b)$, is the linear subspace spanned by the image of b under the transformation matrix A^p , $p = 0, 1, \dots, k - 1$ (where $A^0 = I_n$),

$$\mathcal{K}_k(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{k-1}b\}. \quad (1.8)$$

In numerical linear algebra, Krylov subspace methods are a class of iterative methods derived from (1.8), and are popularly used for their convergence properties, robustness and efficiency. These methods are particularly preferred for when A is large and sparse since the product of Ab is a vector, and $A^2b = A(Ab)$ is another matrix-vector operation. This avoids filling in the zero elements in the matrix and preserves the sparsity of A . A Krylov subspace method finds an approximate solution by solving

$$x_k \in x_0 + \mathcal{K}_k(A, r_0),$$

where x_0 is the initial vector and r_0 is the initial residual vector, $r_0 = b - Ax_0$.

Perhaps the most commonly known Krylov subspace method is the conjugate gradient (CG) method [22]. The link between CG and the Krylov subspaces is when $r^k = b - Ax^k$ is orthogonal to \mathcal{K}_k , i.e. $\mathcal{K}_k \perp r^k \in \mathcal{K}_{k+1}$. CG is applied to systems that are symmetric and positive definite, which (1.5) is not. However, CG is still applicable to the CT problem by solving the normal equation $A^T Ax = A^T b$, otherwise known as the conjugate gradient least squares (CGLS) method [36]. CGLS is a popular method amongst those working in signal and image processing.

Another noteworthy but lesser known Krylov subspace method is the generalized minimum residual (GMRES) method [39]. The link between GMRES and (1.8) is that the residual r_k has the minimum norm for x_k in \mathcal{K}_k . When compared to CGLS, examples show that the residual is much smaller and GMRES requires less computational work for ill-posed problems, [4].

Image regularization techniques

When the coefficient matrix A is ill-conditioned, a slight perturbation in the collected data can lead to inaccurate results. In practice, it is often the case that b contains some levels of random noise, making the computation of an approximate solution even harder. Various approaches are available to solve this problem. The most popular and effective of those is the Tikhonov regularization [42, 19]. Tikhonov regularization transforms the problem (1.5) to

$$(A^T A + \alpha^2 L)x = A^T b,$$

or equivalently, the minimization problem $\min_{x \in \mathbb{R}^{N \times 1}} \|b - Ax\|_2$ to

$$\min_{x \in \mathbb{R}^{N \times 1}} \|b - Ax\|_2^2 + \alpha^2 \|Lx\|_2^2. \quad (1.9)$$

Here, α is the regularization parameter and L is the Tikhonov matrix. In the simplest case, the Tikhonov matrix is taken to be an identity matrix with size $N \times N$ in order to obtain results with smaller norms, i.e. penalizing the solution domain to enforce smoothness. The regularization parameter determines how severe this penalization is. In more ill-conditioned cases, the Tikhonov matrix can be chosen as a stronger lowpass operator, such as a difference operator. It is interesting to note that the solutions to (1.9) are identical to the least squares solutions of the augmented linear equation,

$$\begin{bmatrix} A \\ \alpha L \end{bmatrix} x = \begin{bmatrix} b \\ \mathbf{0} \end{bmatrix}. \quad (1.10)$$

Another popular regularization technique is the total variation model introduced in 1992 [38]. For this technique, instead of minimizing (1.9), we minimize

$$\min_{x \in \mathbb{R}^{N \times 1}} \|b - Ax\|_2^2 + \alpha^2 \|Lx\|_1. \quad (1.11)$$

Solving (1.11) allows larger jumps in the image domain, leading to a more piecewise smoothness, as opposed to an overall smoothness achieved with (1.9). This is a powerful model, however more difficult to implement than Tikhonov regularization because the function is no longer quadratic, or even differentiable [33].

Throughout the thesis, we employ Tikhonov and total variation regularization techniques. We invite the reader to explore the thesis chapters for notes on the algorithm and implementation details, and the references therein.

1.3 Format and Outline of the Thesis

The thesis consists of four research papers either published or in draft format. We introduce each paper here with the emphasis on the work produced by the student. In addition, we mention any work made public prior to the papers presented in the main body. These are included in the appendices of the thesis and act as the preliminary or supplementary work to those that are published.

The outline of the thesis is as follows. The first paper is a tool for laboratory experimentalists working in imaging to understand and examine reconstruction methods and plan in advance the data acquisition experiments. It invites experimentalists to consider multiple methods of reconstructing an image, as well as mathematicians to understand the difficulties in practical CT. In contrast, the following paper looks at exciting theoretical results in mathematics, and applies them to practical CT with a strong analysis undertone. The third paper is the application of a novel idea that could help eliminate a poorly-defined artefact in CT and possibly revolutionize the CT reconstruction for certain systems. The final paper explores the limits of latest data acquisition hardware, and with the application of an iterative reconstruction method tailored for the problem, what analysis can be extracted from the collected data. This paper in particular is an excellent evidence of multidisciplinary work that geoscientists, laboratory experimentalists and mathematicians can produce when in direct contact.

There is a distinct gap in the CT community that we believe is due to the lack of inter-disciplinary communication. This trend is due to researchers prioritizing and focusing on their specialized fields, which is not wrong, but neither does it allow for multi-disciplinary work to be carried out. One of the aims of this thesis is to bridge this gap by reporting the strengths and shortcomings of the mathematical, computational and the practical sides of the CT problem, essentially transferring information between fields within the CT community. Each work included in this thesis can be seen as an evidence of advancing the science in CT when researchers from different backgrounds work together on the same problem.

1.3.1 Image analysis for practical CT

In numerical analysis when introducing a concept, it is common to simulate an event with a pre-determined outcome, apply the concept and compare the numerical (mathematical) difference of the result to the pre-determined outcome. In image (and signal) processing, the similarities between two images are often measured by the Euclidian distance (l_2) norm or the signal-to-noise ratio. The simulations are often performed with a small, simple phantom image where some amount of Gaussian noise is included. A novel method is then tested and the resulting image is compared to the simple phantom previously established. This does not apply to practical CT, and studies with

simulated results greatly downplay many factors of data acquisition, memory management done by hardware, and the fact that a true solution is not available. This is compensated by obtaining a filtered back-projection reconstruction using a very high quality data, which is then taken as the ground truth. In cases where a high quality data is not available, this becomes a major issue. Therefore the traditional approach to analysing reconstructed images work well in simulations, but does not offer the same utility in practical CT. As argued in [43] (and many others by Prof. K. Myers), a better suited merit of establishing an accurate reconstruction is needed. Cited work by Vaishnav et al. focuses on a task-based approach more appropriate for clinical studies, whereas we aim to introduce a broader technique for image analysis and processing.

The paper included in Chapter 1 [9] introduces the idea of using specially-collected data sets (data collected using optimal and goal-specific experiment plan) where the characteristics of the sample are known. This work aims to shift the traditional look on the practical CT, and encourages readers to explore a reconstructed image beyond a grey scale. This technique of analysis also allows further look into more difficult questions in practical CT, such as “when does it become redundant to collect more data?” or “is the level of exposure or dose appropriate, or can it be reduced while the quality of extracted information remains the same?”. A short discussion is included in the paper as an example to reflect this benefit, which aims to answer the question: “when do the iterative reconstruction methods become worth the effort?”. This was also the topic of the preliminary work, which was published in the extended abstract, [10].

In order for this work to be possible, CT data had to be collected in a specific fashion. The acquisition plan used in the paper, here given in [10, Fig.1], was created so that no external factors or additional artefacts could affect the analysis. A pack of simple, smooth, Soda-Lime glass beads were chosen as the scanned object because we know the size, the shape and the texture of the sample and beads are representative of many imaging problems in material science and mathematics. The data set are released as part of the SophiaBeads Dataset Project, which aims to provide researchers with a real, raw CT data for developing, implementing or testing reconstruction algorithms. The data acquisition for the SophiaBeads Dataset Project were performed in March 2015. The data sets are made available to public via a data repository (Zenodo)

[8], along with MATLAB and C codes for users to load, pre-process and reconstruct [1]. The student has also prepared and published a technical report to act as documentation for the project codes and tutorial for applying the physical quantification techniques using an image processing suite [7]. The dataset project and preliminary work produced for quantifying reconstructions were published in the extended abstract [10]. The report and the extended abstract are included as supplementary materials in Appendices A and B.

1.3.2 Compressive sensing in practical sparse tomography

Compressive sensing is the study of recovery of a signal, image or a statistical model from fewer measurements than suggested by the Shannon–Nyquist sampling theorem [40]. The theoretical result suggests that the number of measurements needed for a sufficient recovery can be up to a logarithmic factor linear in the sparsity. The pioneering work by [5] shows that an object can be reconstructed *exactly* from incomplete set of information on the two conditions that the object we wish to recover is sparse, and the system matrix has low coherence (the column vectors of the system matrix have low similarity). This subsequently generated much interest from the CT community, with many working to recover a sparse image from few projections using total variation regularization in various algorithms [41, 16, 6], and in other imaging modalities [32, 44]. The possibility that we can reconstruct an image from a severely reduced amount of data is particularly important in medical CT where the ultimate aim is to reduce radiation dose a patient is exposed to, or in non-destructive material testing where acquisition is often limited in time (fast changes in scanned sample) or due to the particular experimental set up (limited angle availability).

Our initial work with real X-ray CT data for sparse tomography, and its results, are presented in the conference paper [26] (also included in Appendix C), which we had obtained using the SophiaBeads data sets. In the conference paper, we proposed two questions: The first question was “does TV reconstruction compensate better for high-exposure few projections data set, or low-exposure many projections data set?” This is related to a study case in the first chapter in this thesis [9, Section 5.3], except this time the focus was on sparse-regularized solutions.

The second question proposed in the conference paper was related to finding the

critical number of projections given the sparsity of the image in the gradient domain. The results for this was inconclusive, mainly due to the limited number of sparsity levels that were available through the SophiaBeads Dataset Project, of which there were two (a single-sized sample was made public through the project, and a second set of data with smaller sized bead sample were provided by the student).

The paper presented in Chapter 3 [25] is the continuation of this work where we focus on the second question proposed in the conference paper. In addition, we attempt to predict the sufficient amount of information needed to recover a sparse image with real X-ray CT data. This is an extremely important contribution for the CT imaging community, especially for scientists working on image acquisition, which ultimately means a quick analysis such as this can help maximize the output of useful CT data collected in a single experiment, which is often limited by time, dose or circumstances of in-situ experiments. The paper reports our findings, which are that there is an almost-linear relationship between the sparsity of an image and the critical number of projections needed to successfully reconstruct the image. We also performed a phase diagram analysis to deduce an agreeing number of projections, however the results were inconclusive: only a few of the predictions agreed with the critical number of projections suggested by the first analysis.

1.3.3 Reconstructing a non-linear CT problem

In many realistic settings in CT, the forward problem is a non-linear function. The effect of non-linearity can be detected in the measured data from certain systems, in particular from those with large source-detector distance that leads to a wider range of ray paths and thus wider range of linear attenuation. This non-linearity effect appears more dominantly in the polychromatic CT systems (where linear attenuation depends on a spectrum of energies), or as a result of scattering of the rays. The paper presented in Chapter 4 [31], shows that it can even appear in monochromatic CT systems where the source and detector are sufficiently far.

Very little literature is available on the non-linearity effect, with cited papers referring to it with different names such as ‘partial volume’, ‘exponential edge gradient’ or the ‘penumbra’ effect (see references in [31]). However to our knowledge, there are no established methods that can eliminate this effect. The aim of the work is to

demonstrate a systematic way for detecting non-linearity in measured data. Furthermore, we derive a non-linear problem that takes into account the non-linear nature of the measured data, and using numerical optimization methods, reconstruct and thus eliminate this effect in the reconstructed image.

We implement the trust region reflective method to reconstruct the derived non-linear model, CGLS with Tikhonov regularization and Chambolle-Pock with TV regularization to compare the non-linearity effect in linear and non-linear reconstructions. A real γ -ray data were obtained using the Bergen γ -ray tomography system, and simulated data were obtained using a ray tracing algorithm that imitated the Bergen γ -ray system geometry. The paper reports our preliminary results in the sense that no extensive analysis of more sophisticated methods, or fine-tuning of parameters were performed. Nevertheless, the preliminary results showed stark differences between the CGLS and trust-region reconstructions, with the non-linearity being pronounced in the CGLS reconstructions and completely eliminated in the trust-region reconstructions. On the other hand the Chambolle-Pock with TV regularization returned much clearer reconstructions seemingly without the non-linearity effect, however the reconstructions themselves contained inaccuracies, such as the reconstructed tube wall thickness was larger than the actual thickness.

1.3.4 Fast data acquisition and dynamic 3D reconstruction in synchrotron X-ray CT

The final chapter [12] in the thesis is a methodology paper published by Solid Earth in the special issue “*Pore-scale tomography & imaging - applications, techniques and recommended practice*”. The paper is prepared for geoscientists whose work depends heavily on data acquisition, image reconstruction and analysis techniques. The imaging challenges in this field vary from detailed reconstructions of rock decompositions in pore-scales to evolution of the solidification of volcanic lava into igneous rock in high temperatures. The published work reports the data set acquisition and our findings using the Vision Research Phantom MiroTM 310 M camera, installed at the I12 Beamline in Diamond Synchrotron, Harwell Research Facility, Oxfordshire, UK. The paper shows the impressive capabilities of the latest hardware for in-situ synchrotron X-ray

CT, when coupled with optimal data acquisition design and flexible, model-based reconstruction techniques. For this reason, the paper was chosen as one of the highlight articles of the special issue [18]. The paper also displays the use of quantifying the reconstructed volumes in order to extract relevant information (in this case, tracking the volume of inter- and intra-grain fluid transfers as the time increases, see [12, Fig. 5]).

The reconstruction, post-processing and quantification aspects involved were not emphasised in the paper due to its target audience. Therefore we will expand on these aspects here. We performed over 50 data acquisition experiments with each collected data containing between 53 to 107 frames (3D volumetric data), and each frame representing events as quick as 0.07s. These were either 360, 180 or 90 projection scans with scaled exposure time. In addition, for each experiment, a high quality data was collected (of the dry sample, prior to any wetting). An in-house FBP implementation was used during and immediately after the scans in order to determine any artefacts that could come to play in the experiments such as major sample movement, low saturation in liquid and the rocks, or a drop not being caught (the internal memory of the camera would be full before any important events happened in the acquisition).

The quantifications in [12, Section 4.1] were performed using Avizo Fire 9.1. The workflow used is very similar to the one described in [7]. Avizo was used for the visualisations for other figures, where similar workflow was reproduced.

For the images used in the paper, we obtained reconstructions from 180 projections using the total variation (TV) regularization as part of gradient projection algorithm with Barzilai-Borwein acceleration (GPBB). Details of the implementation are given in [24]. Originally this implementation took 9 hours to produce a volumetric reconstruction from a single frame. By rearranging the matrices to limit hardware memory transfer, parallelizing the process and trimming the input data where possible, the reconstruction time for each volume was reduced to between 40 minutes to an hour.

Difference reconstruction in practical CT

In addition, we experimented with difference image reconstructions by either using a high quality dry scan (initial state, if available) or the previous state. Difference image reconstruction is essentially taking the difference in projections for two separate

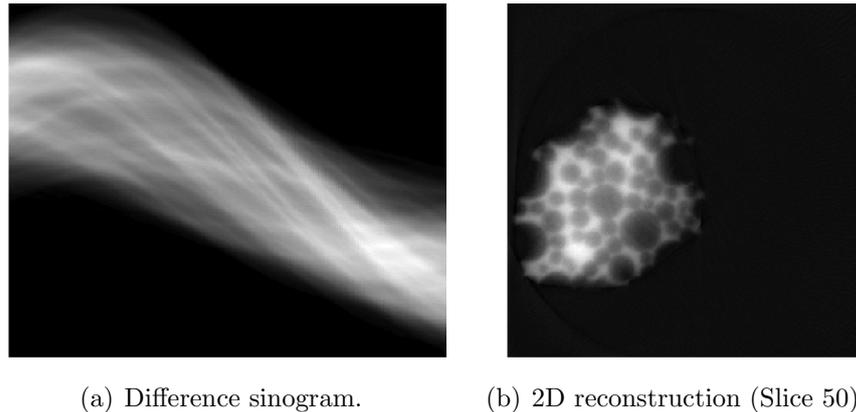


Figure 1.4: Difference reconstruction using a sample of mixed sized beads. Here the difference between the current state (arrival of first drop) and the dry-sample measurements (no drop) is reconstructed.

states, and reconstructing the resulting data. This is not a new idea but fairly novel in practice, however no practical examples are available in the CT imaging literature. We refer the readers to [20, 35] for the general idea and simulated results. For our experiments, the reconstructions initially gave poor results due to the axis shift and the odd-even reflection in collected projections (the initial or previous states did not exactly match the current state in voxels), so minor pre-reconstruction corrections had to be applied (either by shifting the indexing of voxels or applying reflection on the odd volumes to match the even ones). Reconstructions following the pre-processing are included in Figure 1.4 and Figure 1.5 where we can see the difference sinogram and slice reconstruction (Figure 1.4a and 1.4b), and evolving fluid volumes reconstructed using differences between the dry sinogram and saturated sample data for the first drip (Figure 1.5a, $t = 0.00s$), the sample at $t = 1.47s$ (Figure 1.5b) and at $t = 3.71s$ (Figure 1.5c).

1.4 Suggested further work

Note that each paper in the thesis has suggestions for further work relevant to its topic. In addition, we propose the following ideas to be studied in future.

First idea involves the entropy theory, which can be used to find an expression for the information content in a given CT data set. This expression can then be maximized to extract the most useful sections of the data.

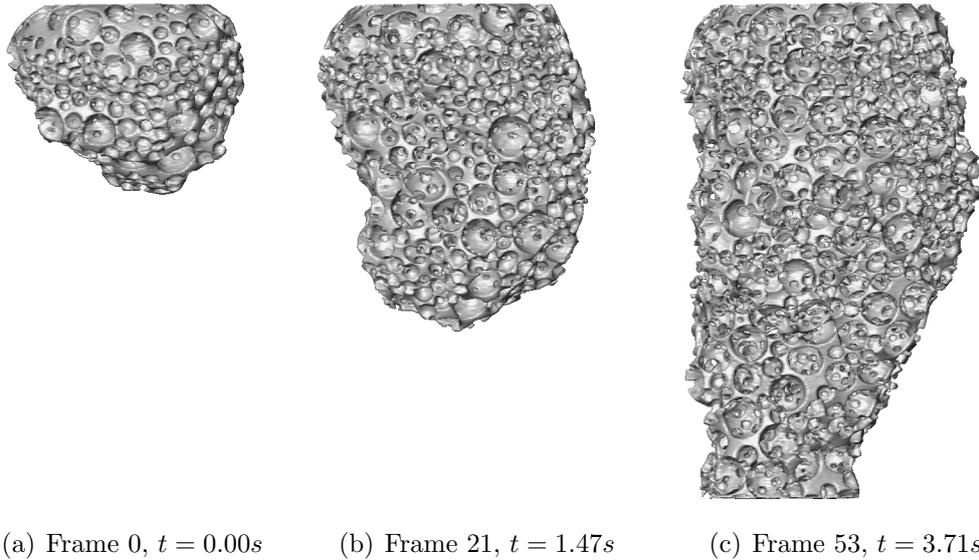


Figure 1.5: 3D renderings of difference reconstruction results where the water-saline drop is isolated. On left is the first arrival of the drop in the mixed beads pack. Second is the mid point in the flow experiment. On right is the phase where the beads pack is completely saturated and the fluid has established a path (end of experiment).

Another idea is to identify the properties of matched and unmatched forward-back projector pairs, in particular, understand how close a back projector, B must be to an exact back projector, A^T in order to obtain acceptable reconstructions but still be computationally cheap to implement. When such a back projector is implemented, we believe a Krylov subspace method, specifically GMRES, could be a suitable iterative method for solving the reconstruction problem.

Finally, we can suggest using the Low-rank and sparse matrix decomposition technique to separate a 4D reconstruction (by which we mean a series of 3D volumes changing with time) into two volumes, each representing either the foreground or the background. The next time step can then be calculated with the background volume taken away, leaving only the foreground volume to be reconstructed with the same amount of data. This would significantly reduce the computational costs of a reconstruction, and reduce the inaccuracies in the image for the next time step. This is not a new method, and the separation of foreground from background is popularly used by security software companies to track objects in a moving scene. However to our knowledge, this has never been implemented for a real 4D (3D+time) CT data.

Bibliography

- [1] SophiaBeads Datasets Project Codes – GitHub Page. <http://sophilyplum.github.io/sophiabeads-datasets/>. Accessed: 01-APR-2015.
- [2] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984.
- [3] A. Beer. Bestimmung der absorption des rothen lichts in farbigen flüssigkeiten. *Annalen der Physik*, 162(5):78–88, 1852.
- [4] D. Calvetti, B. Lewis, and L. Reichel. On the choice of subspace for iterative methods for linear discrete ill-posed problems. *Applied Mathematics and Computer Science*, 11(5):1069–1092, 2001.
- [5] E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.
- [6] K. Choi, J. Wang, L. Zhu, T. Suh, S. Boyd, and L. Xing. Compressed sensing based cone-beam computed tomography reconstruction with a first-order methoda). *Medical physics*, 37(9):5113–5125, 2010.
- [7] S. B. Coban. SophiaBeads Datasets Project Documentation and Tutorials. Technical report, School of Mathematics, University of Manchester, United Kingdom, April 2015. MIMS Technical Report.
- [8] S. B. Coban and S. A. McDonald. SophiaBeads Dataset Project, Mar. 2015.

- [9] S. B. Coban, P. J. Withers, and W. R. B. Lionheart. Assessing the suitability of computed tomography reconstruction methods through physical quantification techniques. 2016. In preparation.
- [10] S. B. Coban, P. J. Withers, W. R. B. Lionheart, and S. A. McDonald. When do iterative reconstruction methods become worth the effort? In *Proc. 13th Intl. Mtg. of Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, June 2015.
- [11] A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. *Journal of Applied Physics*, 34(9):2722–2727, 1963.
- [12] K. J. Dobson, S. B. Coban, S. A. McDonald, J. N. Walsh, R. C. Atwood, and P. J. Withers. 4-D imaging of sub-second dynamics in pore-scale processes using real-time synchrotron X-ray tomography. *Solid Earth*, 7(4):1059–1073, 2016.
- [13] L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *JOSA A*, 1(6):612–619, 1984.
- [14] J. A. Fessler. Statistical image reconstruction methods for transmission tomography. *Handbook of medical imaging*, 2:1–70, 2000.
- [15] D. V. Finch. Cone beam reconstruction with sources on a curve. *SIAM Journal on Applied Mathematics*, 45:665–673, 1985.
- [16] B. Goris, W. Van den Broek, K. J. Batenburg, H. H. Mezerji, and S. Bals. Electron tomography based on a total variation minimization reconstruction technique. *Ultramicroscopy*, 113:120–130, 2012.
- [17] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, page 49–52, 1902.
- [18] M. Halisch, H. Steeb, S. Henkel, and C. M. Krawczyk. Pore-scale tomography and imaging: applications, techniques and recommended practice. *Solid Earth*, 7(4):1141–1143, 2016.

- [19] P. C. Hansen. *Rank-deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.
- [20] P. B. Heffernan and R. A. Robb. Difference image reconstruction from a few projections for nondestructive materials inspection. *Appl. Opt.*, 24:4105–4110, 1985.
- [21] G. T. Herman. *Fundamentals of computerized tomography: image reconstruction from projections*. Springer Science & Business Media, 2009.
- [22] M. R. Hestenes and E. Stiefel. *Methods of conjugate gradients for solving linear systems*, volume 49. NBS, 1952.
- [23] G. N. Hounsfield. Computerized transverse axial scanning (tomography): Part 1. description of system. *The British Journal of Radiology*, 46(552):1016–1022, 1973.
- [24] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen. Implementation of an optimal first-order method for strongly convex total variation regularization. *BIT*, 52:329–356, 2012.
- [25] J. S. Jørgensen, S. B. Coban, W. R. B. Lionheart, S. A. McDonald, and P. J. Withers. Predicting performance of sparsity-regularized x-ray tomography: experimental results using glass-bead data. 2016. In preparation.
- [26] J. S. Jørgensen, S. B. Coban, W. R. B. Lionheart, and P. J. Withers. Effect of sparsity and exposure on total variation regularized X-ray tomography from few projections. In *Proc. 4th Intl. Mtg. on image formation in X-ray CT*, pages 279–282, 2016.
- [27] S. Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres*, 35:355–357, 1937.
- [28] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. IEEE press, 1988.

- [29] J. H. Lambert. Photometria sive de mensura et gradibus luminis, colorum et umbrae. *Augusta Vindelicorum*, 1760.
- [30] W. R. B. Lionheart. Evolution of bio-impedance (and impedance imaging): from solid foundations and in to the future, June 2016.
- [31] W. R. B. Lionheart, B. T. Hjertaker, R. Maad, I. Meric, S. B. Coban, and G. A. Johansen. Non-linearity in monochromatic transmission tomography. 2016. In preparation.
- [32] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic resonance in medicine*, 58(6):1182–1195, 2007.
- [33] J. L. Mueller and S. Siltanen. *Linear and nonlinear inverse problems with practical applications*, volume 10. Siam, 2012.
- [34] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*. Society for Industrial and Applied Mathematics, 2001.
- [35] A. Pourmorteza, H. Dang, J. H. Siewerdsen, and J. W. Stayman. Reconstruction of difference in sequential ct studies using penalized likelihood estimation. *Physics in Medicine and Biology*, 61(5):1986, 2016.
- [36] Å. Björck, T. Elfving, and Z. Strakos. Stability of conjugate gradient and Lanczos methods for linear least squares problems. *SIAM J. Matrix Anal. Appl.*, 19(3):720–736, 1998.
- [37] J. Radon. On the determination of functions from their integral values along certain manifolds. *IEEE Transactions on Medical Imaging*, 5(4):170–176, Dec 1986.
- [38] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.
- [39] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on scientific and statistical computing*, 7(3):856–869, 1986.

- [40] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.
- [41] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in medicine and biology*, 53(17):4777, 2008.
- [42] A. N. Tikhonov. *Solutions of Ill-Posed Problems*. Wiley, New York, NYC, 1977.
- [43] J. Y. Vaishnav, W. C. Jung, L. M. Popescu, R. Zeng, and K. J. Myers. Objective assessment of image quality and dose reduction in CT iterative reconstruction. *Medical Physics*, 41(7), 2014.
- [44] X. X. Zhu and R. Bamler. Tomographic sar inversion by-norm regularization—the compressive sensing approach. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10):3839–3846, 2010.

Chapter 2

Assessing the suitability of
computed tomography
reconstruction methods through
physical quantification techniques

Assessing the suitability of computed tomography reconstruction methods through physical quantification techniques

S.B. Coban*, W.R.B. Lionheart* and P.J. Withers†

School of Mathematics, School of Materials†, University of Manchester, Manchester, M13 9PL, UK*

Abstract

Conventionally the efficacy of different computed tomography reconstruction algorithms is assessed a voxel to voxel comparison between the object and the reconstructed volume. Here we quantify features of interest for the test object and use these as measures of the efficacy of the reconstructions. The technique, which we refer to as physical quantification, is used here to determine the sphericity, contacts and size of beads for a test dataset made available via the SophiaBeads Dataset Project. Using this image analysis approach a number of reconstruction methods are evaluated including FDK, CGLS, TV regularization and level set evolution methods. Our work shows that it is important to choose the optimal reconstruction strategy based on the features you want to quantify from the scan. In our case we found that the shape of the beads could be measured using TV regularisations with 8x fewer projections than the other methods.

Keywords: Computed tomography, reconstruction methods, physical quantification

1. Introduction

In computed tomography (CT) imaging there is a continuous demand for iterative reconstruction (IR) methods, adapted to specific circumstances, with the common goal of producing high quality images. For many, a high quality reconstruction means a noiseless image with clear contrast, a solution that satisfies a ‘best trade-off’ in convergence, and is close to the approximate but robust and widely-used Feldkamp-Davis-Kress (FDK) filtered back-projection algorithm [14]. The analysis techniques in image processing (signal-to-noise, contrast-to-noise) and in numerical analysis (l_2 -norm error in the residual or consecutive iterations) are commonly used to assess the quality of reconstructed images and thus the suitability of the iterative methods. However, these are poor indicators of image quality since an iteratively reconstructed solution would have spatially-dependent noise, and resolution dependent on image contrast [34]. In addition, FDK and IR methods produce different artefacts which are inaccuracies in a reconstructed image, and thus comparing these solutions with each other is not a fair comparison. As argued in [37], the assessment of image quality is challenged by these traditional techniques, and an established measure of merit for evaluating the quality of reconstruction; independent of the noise and the contrast, or its similarity to an approximate FDK solution, is necessary. An alternative, task-based method is introduced in [37] however this is designed for clinical studies, whereas our aim here is to establish a more general measure of quality for tomographic imaging.

The real aim in the CT imaging community is not to develop a reconstruction algorithm to obtain the best-looking images, but one that allows us to extract the relevant information from

a reconstruction to a desired accuracy. In this paper, we examine a strategy for extracting appropriate physical information in order to evaluate the quality of a reconstructed solution, and thus the suitability of reconstruction methods, for a given problem. The technique, which we refer to as ‘physical quantification’, is performed on reconstructions from a set of CT data, where some of the characteristics of the scanned object are known. This data, collected for the purpose of developing and testing image reconstruction techniques, was first introduced in [9]. We reintroduce the project for the sake of completeness in §2. Similar strategies are often used in medical, geophysical and non-destructive material studies to build or verify a physical model. Numerous examples are given in [35, 12, 10], which experiment with the idea of using features in a CT reconstruction to extract information relevant to the study. In the same spirit, we consider four problems to showcase the use of a physical quantification technique, and how it could differ from traditional ways of assessing the suitability and performance of a reconstruction method. These are typical problems in the imaging community, and are included to offer perspectives relevant to both mathematicians and experimental scientists working in the field. The results of these case studies are presented in §5, and their brief descriptions are as follows.

The first case study involves comparing three important types of reconstruction methods. These are the filtered back projection (FDK, an approximate direct method); the conjugate gradient least squares (CGLS, an iterative method with no explicit prior information) [26], and the total variation image reconstruction (henceforth TV Prior, an iterative method with strong prior information) [16]. The study focuses on making fair comparisons of these methods while physically quantifying the accuracy of the reconstructed results.

The following case study looks at the convergence of CGLS in the absence of prior information, and highlights how the conclusions we can draw from the traditional and our technique of assessing convergence may differ.

The third case study is an example of a more practical problem where we look at the effect of varying the signal quality (can also be thought as exposure or dose) in tomographic acquisition on the quality of reconstructed solutions. This involves studying the affect of increased noise as the amount of signal (or dose) is reduced, posing the question whether there are any advantages to higher exposure scans.

The final case study is an example of testing a novel idea against more standard methods. In this study, we explore the idea of simultaneously reconstructing and segmenting images with the means of a 2-step discrete-CGLS algorithm (DCGLS, our modification to the DART algorithm) [3]. We compare a variation of the level set evolution formulation (the distance regularized level set evolution or DRLSE) [19], here adapted as a segmentation method, with simpler techniques such as basic, Otsu’s [25] and locally adaptive [4] thresholding methods.

We finish with more discussions and further analysis of the first case study, and concluding remarks for the work presented.

2. Establishing a reference CT dataset

2.1. Description of the test data

The SophiaBeads Dataset Project [8] is a collection of cone-beam X-ray CT data sets where the number of projections are varied while the total photon count (or the total exposure time) is kept constant [9]. These data sets enable a wide range of algorithm comparisons and information content optimizations to be examined. They are also suitable for testing segmentation methods since the sample comprises just two phases (glass and air). The project

also contains scripts to read and pre-process the data prior to reconstruction, [6], which are essential in order to use the data set.

The data sets are available to download individually, with the largest one containing 2048 projections. The number of projections is reduced by a factor of 2 for the subsequent data sets, with the smallest one containing 64 projections. For the case studies in §5, we made use of all data sets available.

2.2. Scanner and sample details

The SophiaBeads data set were collected using the *320/225 kV Nikon XTEK Bay* located at the Henry Moseley X-ray Imaging Facility (HMXIF), the University of Manchester.

The apparatus consists of a cone-beam microfocus X-ray source that projects polychromatic X-rays onto a 2000×2000 pixel-length and width, 16-bit, flat detector panel. The optimal window size for the SophiaBeads reconstructions is 1564×1564 , which was determined prior to the data acquisition, see [6]. The data sets were taken using the stop/start acquisition technique to reduce angular blurring, and a 0.1mm copper filter was used throughout to eliminate beam hardening artefacts.

The sample comprised a 25mm diameter plastic tube, filled with uniform Soda-Lime Glass ($\text{SiO}_2\text{-Na}_2\text{O}$) beads size in range 2.3mm-2.6mm, with a mean value of 2.5mm. The beads in the plastic tube are packed tightly and do not move during data acquisition.

A pack of glass beads were chosen because they are representatives of many physical problems in non-destructive testing, such as a porous medium to study the flow through porous media [13, 23, 11] and granular flow of particles [36, 1].

3. Quantification Strategies

3.1. Physical assessment of the beads

Our motivation for using the SophiaBeads data sets in particular is that we know the main features of the sample: We know the sizes of the beads so we can physically quantify the quality of reconstructions and make fair conclusions. In §5 where we solve typical questions in imaging, we quantify our results using physical measures, which we define below:

- I. **Shape of the beads** – Shape3D or Shape2D: We can parameterize how close (in shape) a reconstructed bead is to a perfect sphere.
- II. **Aspect ratio of the beads** – AspectRatio3D or AspectRatio2D: The ratio of the smallest and the largest (orthogonal) width of the bead in 2D (circle) and 3D (sphere).
- III. **Sphericity of the beads**: The percentage of the reconstructed bead that fits to a perfect sphere, similar to SHAPE3D.

Most of these measures are available within the image-visualization and measuring software, Avizo Fire (versions 7 to 9) [2]. The Sphericity measure is not native to Avizo and was defined prior to the quantifications. To obtain a percentage for the sphericity, we used

$$\text{Sphericity} = 100 \frac{\pi^{1/3}(6V)^{2/3}}{S}, \quad (1)$$

where V is the volume and S is the surface area of a bead [38, 15]. Despite knowing some important properties of the beads (uniform, spherical beads; in contact but not connected),

the fact that they vary in size and proportion means that we cannot use a perfect physical measure. In order to establish a physical reference that is representative of the sample, we evaluated the mean and the standard deviation of physical quantification measures for the full 2048-projection FDK reconstruction (size: $1564 \times 1564 \times 2000$). These reference values, given in Table 1, are based on 500 reconstructed and segmented beads, thus are illustrative of the sample characteristics. For the planar (2D) quantification measures, we used the central horizontal slice (henceforth Slice 1000; size: 1564×1564) to obtain the reference values. A summary of this process is explained in Appendix C.

Quantification Measure	Idealized Ground Truth	Reference Physical Values			
		Minimum	Maximum	Mean	Standard Deviation
Shape3D	1	1.0832	1.9776	1.3274	0.00226
Shape2D	1	1.1248	2.3614	1.2143	0.00750
AspectRatio3D	1	0.7142	0.9719	0.9112	0.00172
AspectRatio2D	1	0.7948	0.9937	0.9319	0.00320
Sphericity	100	79.177	97.372	87.164	0.0132

Table 1: Listing of the idealized ground truth of perfectly spherical beads and the evaluated reference values for the given physical measures. Reference values were obtained using the full 2048-projection FDK reconstruction.

Although not used in this paper, one can also use the volume and the curvature of individual beads to determine how ‘correct’ a solution is. These were not used as quantification measures since they were used in the process of separating and filtering the reconstructed beads in Avizo, therefore are left out in order to avoid influencing the final quantification values. Instead, the independent measures listed above were chosen. For a detailed tutorial on using Avizo to segment, separate and quantify the SophiaBeads reconstructions, as well as using the released project codes, we refer the reader to [7]. We emphasize here that there are other open-source image-measuring software such as Blob3D [17], Drishti [20] or VTK [33], which can be used to apply physical measures to reconstructions.

3.2. Traditional image analysis techniques

In addition to physical measures, we also include traditional, pixel-based image analysis techniques (where applicable). These are standard measures used in mathematics to show the grey value differences between two images (the \mathcal{L}_2 error relative to the ground truth image, $\mathcal{E}_{2\text{norm}}$), or to observe changes in consecutive solutions or in estimated and measured data, at each iteration (iterative error relative to the current iterate, $\mathcal{E}_{\text{iter}}$ and the residual error relative to the measured data, \mathcal{E}_{res} , respectively). These are defined by

$$\mathcal{E}_{2\text{norm}} = \frac{\|x^* - x\|_2}{\|x^*\|_2}, \quad (2)$$

$$\mathcal{E}_{\text{iter}} = \frac{\|x_k - x_{k-1}\|_2}{\|x_k\|_2}, \quad (3)$$

$$\mathcal{E}_{\text{res}} = \frac{\|b - b_k\|_2}{\|b\|_2}, \quad (4)$$

where $\|\cdot\|_2$ denotes the (Euclidian) 2-norm, x^* denotes the ground truth image, b denotes the log of the measured data and b_k is the estimated data by the k th iterated solution x_k , defined by $b_k = Ax_k$. In cases where Equation 2 is used, the 2048-projection reconstruction (of size $1564 \times 1564 \times 500$) is taken as the ‘ground truth image’.

4. Reconstruction Algorithms

4.1. Implementation notes

The reconstruction methods mentioned in this section (except FDK) are implemented in MATLAB 2014b, with the forward and back projectors written in C [6]. Readers will notice that the CGLS is used in all subsections. Other methods used in the case studies are FDK using the in-house software, CTPro; a formulation of total variation as a regularizer with a MAP estimate (TV Prior), and our modified version of the discrete algebraic reconstruction technique (DART) with segmentation methods such as the basic thresholding, Otsu’s thresholding, locally adaptive thresholding and a variation of the level set methods (DRLSE). All formulation and algorithm implementation details are listed in Appendix A.

4.2. Presentation of results

For the 3D results, we reconstruct the central subvolume of $1564 \times 1564 \times 500$ and display the images of the central horizontal slice (Slice 1000 in the full volume, Slice 250 within the subvolume). We focus on a central region of interest (ROI) of size $440\text{px} \times 440\text{px}$ or smaller ROIs of size $100\text{px} \times 100\text{px}$ of a bead in the centre or $130\text{px} \times 130\text{px}$ near the edge (see Figure 1).

The quantification plots are obtained using the `loglog`, `semilogx` (log-linear) or `plot` (linear) functions in MATLAB, where the plotted results are the mean physical quantification values of 25 reconstructed beads, contained fully within the reconstructed volume (see Appendix C for more information).

5. Case Studies

5.1. Determining the appropriate reconstruction method via fair comparison

In this study, the aim is not to get the best results for the implemented algorithms, but rather comparable results for three important types of reconstruction methods used today: a popular approximate direct method, with an implicit prior incorporated in a filter applied to the data (FDK); an iterative method with no explicit¹ prior information (CGLS), and an iterative method with a strong prior (TV Prior). There were no optimal parameter searches done for the iterative methods. Also worth noting here that there is no data fitting error for FDK. For the CGLS reconstructions, we performed 12 iterations and for the TV Prior, 2000 iterations (chosen as the limit). These numbers were determined by how long each method took to diverge (or stall in convergence) for the 512-projection data set. This was done to achieve comparable results however, running the iterative methods until they stalled

¹An implicit prior information is provided for this method by the stopping criterion, which often is \mathcal{E}_{iter} or \mathcal{E}_{res} . However, this is replaced by a user-defined number of projections in this paper, in order to eliminate their effect on physical quantification.

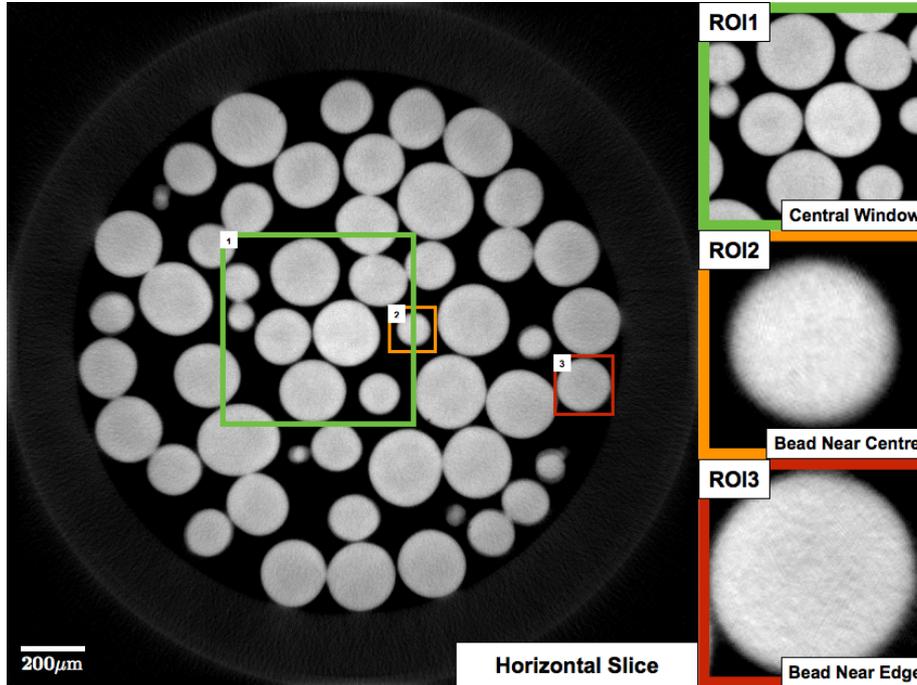


Figure 1: 2D CGLS central slice reconstruction of 512-projection dataset (4 frames), along with 3 regions of interest, namely, a central window of the sample (ROI1; of size $440\text{px} \times 400\text{px}$ or $704\mu\text{m} \times 704\mu\text{m}$), a bead near the centre (ROI2; size: $100\text{px} \times 100\text{px}$ or $160\mu\text{m} \times 160\mu\text{m}$) and a bead near the edge (ROI3; size: $130\text{px} \times 130\text{px}$ or $208\mu\text{m} \times 208\mu\text{m}$).

or diverged meant that the computation time of TV Prior increased drastically, with 2048-projection results taking approximately 4 weeks; 1024-projection results taking 2 weeks, and so on. The timings are displayed in the quantification plot, Figure 3. Other parameters for TV Prior, listed in Appendix A, were determined for 512-projection data set, and remained unchanged for the rest of the reconstructions. The central window (Figure 1, ROI1) of Slice 1000 are given below for reconstructions based on 64-, 256- and 1024-projection data sets.

Traditional image analysis methods are not suitable here as we do not have a high-quality result, with which to compare these reconstructions. Therefore we use the physical measure, Shape3D, to quantify the quality of the results. It is not surprising that the TV Prior does better than the others for all reconstructions; the TV Prior takes into account the sparsity in the gradient reconstructions, enhancing the edges of the beads. Since the gradient image of SophiaBeads reconstructions is very sparse, the TV Prior works really well. The striking observation here is how long it takes to obtain a high quality result using this method. However, this is also expected since a more sophisticated formulation requires more computation – and in the case of TV Prior, this means multiple computations of the gradient and many iterations until a level of convergence is achieved. In our case, the results for 1024- and 2048-projection data set stalled and were run until the limit of 2000 iterations was reached. The results did not show further improvements, as indicated both in the Shape3D quantification plot and the reconstructed results in Figure 2.

An important conclusion we can draw from the Shape3D quantification is that one can achieve the quality of 2048-projection FDK result with a 512-projection, or even 256-projection

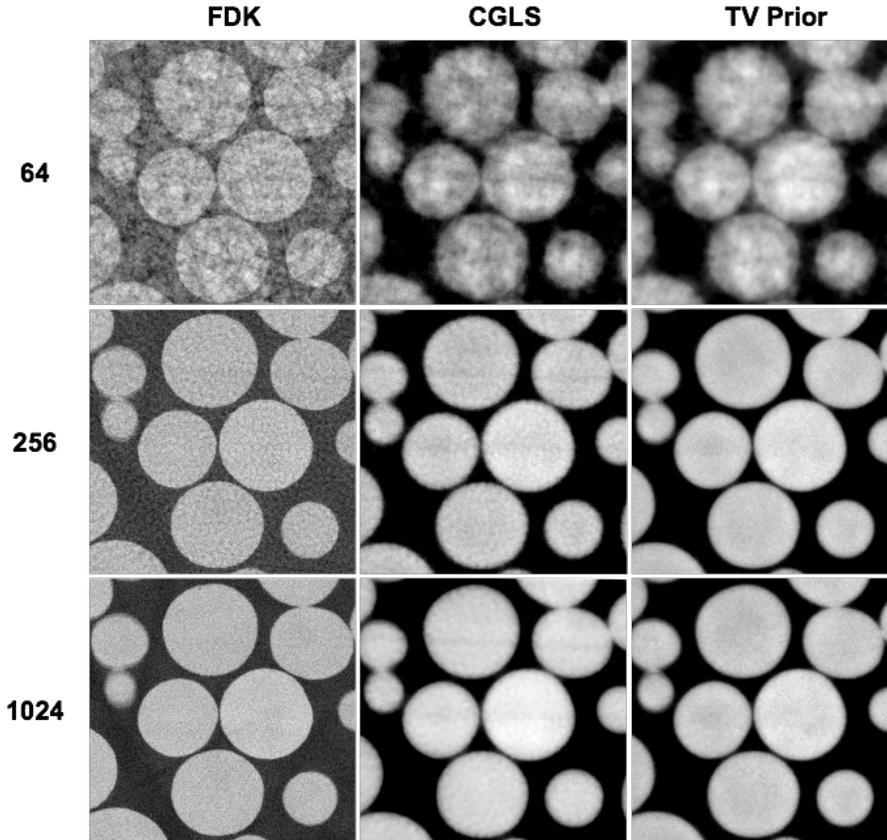


Figure 2: Central window of Slice 1000 for (top to bottom) 64-, 256- and 1024-projection data sets, reconstructed using (left to right) FDK, CGLS and TV Prior.

CGLS result. With a simple prior, the CGLS result can be improved with small increase to the run time, thus the method can be seen as a good replacement for FDK in the case of a low-dose or few-projections problem. Conversely, when we collect data with many projections (e.g. 2048-projection), we obtain reconstructions in 10 minutes using FDK, which would otherwise take days with CGLS and weeks with TV Prior.

5.2. A fresh take on the convergence study of an iterative method

In numerical analysis, a small test matrix (often with special properties or obtained via simulations) is used and the iterated solutions are compared to a true solution (a predetermined phantom) to study the behaviour of a method. Here, we have real CT data sets of varying sizes and a number of iterative methods, meaning we can study the behaviour of an iterative method in a practical setting. We choose CGLS for this study because it is a fast iterative method where one 3D iteration ($x \in \mathbb{R}^{1564 \times 1564 \times 500}$) of 64-projection data set takes (on average) 90s; 128-projection takes 180s; 256-projection data set takes 360s, and so on. Note that we exclude the 2048-projection data set in this set up due to the run time and small contributions to analysis (similar results to 1024-projection are observed).

Two questions are considered in this study: How does CGLS converge in the case of real CT data, and how would this behaviour change as the amount of data is reduced and thus becomes insufficient?

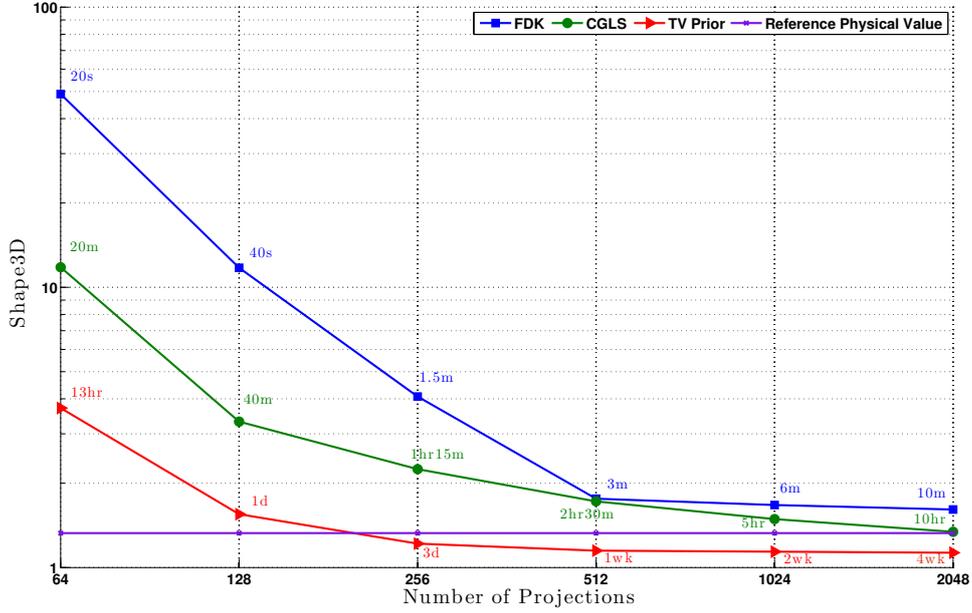


Figure 3: The results of the Shape3D analysis applied to each reconstruction. The graph also includes the computational run times for each reconstruction, annotated above each result in matching colours.

We performed 50 iterations of CGLS, saving every 5th iterated volume to be analysed using a physical measure, which is chosen to be AspectRatio3D. Also evaluated at each iteration, k , are \mathcal{E}_{iter} and \mathcal{E}_{res} (Equations 3 and 4), where $x_{k=0}$ is a zero (all black) image, taken as the starting point. The following figures show the bead near the centre (Figure 1, ROI2) reconstructions for 64-, 128- and 256-projection data sets at iterations 5 to 45 (Figure 4); the linear plots of the pixel-based error measures (Figure 5), and the linear plot of AspectRatio3D values (Figure 6).

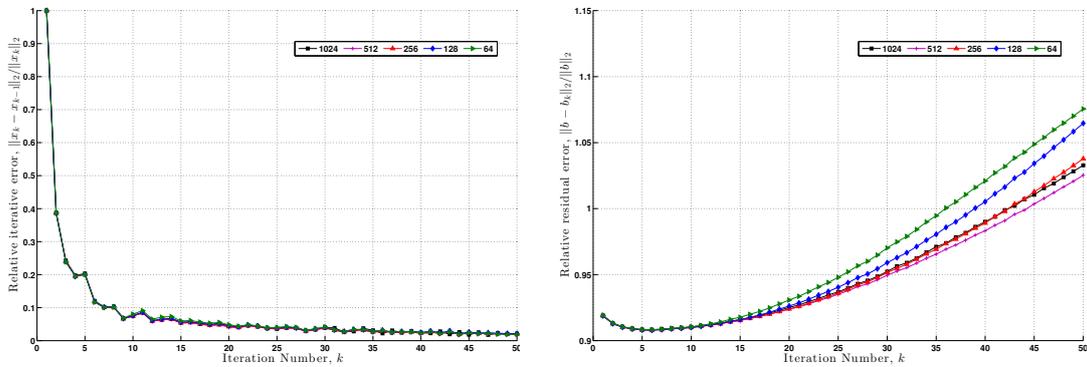


Figure 5: The relative iterative error, \mathcal{E}_{iter} (on left), and relative residual error, \mathcal{E}_{res} (on right), for the CGLS reconstructions based on 64- to 1024-projection data sets. The error norms are calculated at each iteration, k .

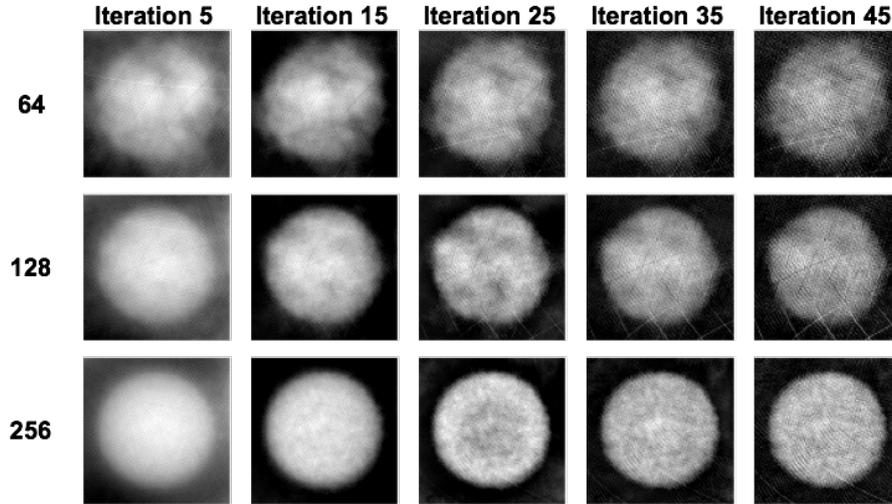


Figure 4: Reconstructions of the bead near the centre, at iteration 5 to 45 for 64-, 128- and 256-projection data sets.

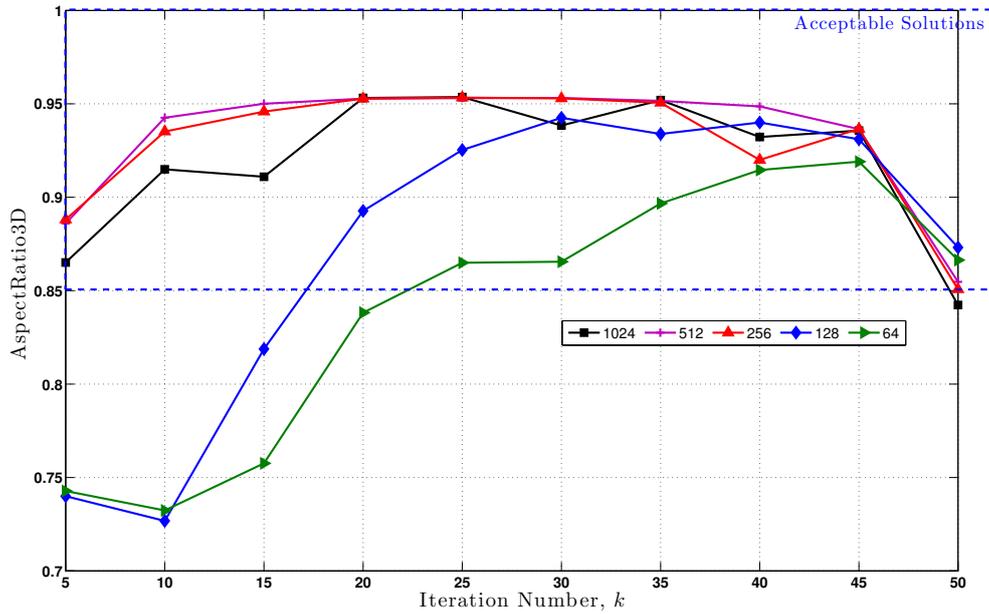


Figure 6: A linear plot of physical quantification measure, AspectRatio3D. The quantification was performed for every 5th iterated volume. The rectangle (blue dashed edges) indicate the acceptable solutions.

It is difficult to draw any conclusions from the traditional analysis techniques in Figure 5. The relative iterative error, \mathcal{E}_{iter} (on left), plot indicates that as the iteration number, k , increases, the changes between consecutive iterated solutions decrease. However, this does not mean that the method converges to a ‘better’ solution (better in the sense that the reconstructed image has less noise or the beads are accurately segmentable). This simply

means that the difference between the current and previous solution is decreasing, which could be a noisy image at iteration 50. The relative residual error plot, \mathcal{E}_{res} (on right), shows that as k increases, the difference between the estimated (b) and predicted (b_k) data decreases initially, but later increases rapidly. The minimum value of this error for all data sets is at iteration 4, suggesting that the ‘best’ approximation to a sensible solution is achieved at this point. However, upon visual inspection of the reconstructed beads in Figure 4, we see that at iteration 5, the image is still too blurry for a computer to be able to automatically segment, and in the few-projection case, deceptive even for visual perception of a human being. This shows how inaccurate standard practices can be when working with real data sets and iterative methods. Figure 6 on the other hand shows the iterated values for which the reconstructed volumes resemble more defined beads. The rectangle with blue dashed edges show the area for acceptable AspectRatio3D solutions, meaning the beads are segmentable at those locations. The graph shows that the reconstructions quickly converge to an acceptable solution but later diverge. In the case of many-projections (512-projection and above), the convergence happens very quickly and results are stable for a while, whereas in the case of insufficient data (128-projection and below), the convergence is slower and results diverge shortly after, albeit at a slower speed than 256-, 512- and 1024-projection. The interesting note is the results for the 256-projection, which are of acceptable quality as the many-projection data sets, and reach the same level of quality as the 512- and 1024-projection results in 20 iterations (3hr20m runtime).

A note we make here is that, in contrast to the error plots, the AspectRatio3D analysis agrees with the visual inspection that the solutions in the middle are acceptable solutions, and for data sets with many projections, the iterated solutions become more acceptable sooner.

5.3. The effect of varying signal in acquisition on the quality of reconstruction

In this study, we attempt to answer an interesting question from an experimental point of view: The noise is undoubtedly reduced when we collect multiple frames of a projection and average over the number of frames. However, is it worth collecting multiple frames at conservative exposure levels in few-projection problems?

The acquisition plan of SophiaBeads data sets was described in [9, Figure 1], where line B consisted of the data sets taken at decreasing signal (the conventional way, i.e. single frame) whereas line A are acquired at a constant signal (constant photon count while the number of projections is reduced, i.e. multiple frames). Line A data sets were released for the purpose of comparing algorithms, whereas line B data sets are not publicly available. However we can still obtain the line B data sets from line A by downsampling appropriately. For example, if we wanted to obtain 128-projection data set with 1 frame (line B), we would have to look up the data set with 1 frame on line A, which is 2048-projection, and downsample it by 16 (i.e. take every 16th projection). This is equivalent to the 128-projection data set taken at a single frame. In fact, we can obtain data set for any number of frames (i.e. dose) using the line A data sets: if we wanted to obtain 64-projection data set with 4 frames, we would downsample 512-projection by 8 ($32/4$). This does not convolute the analysis since the SophiaBeads projections are not taken continuously but rather the rotation pauses to take one projection at a time (stop-start scan).

The reconstructions were obtained using the same numerical set up as the previous case studies: by performing 12 CGLS iterations for each data set lying on lines A and B. The central subvolume (of size $1564 \times 1564 \times 500$) was reconstructed and Figure 7 shows the bead near the edge of the reconstructed images (Figure 1, ROI3).

We are able to use a traditional image analysis technique here and include the \mathcal{E}_{2norm} measure, where the 2048 FDK result is taken as the ‘ground truth image’ (Figure 8). For the physical quantification, we calculate the Sphericity percentage of each reconstruction. The results are given in Figure 9, which also include the Sphericity reference value, which is 87.16% (mean value from Table 1).

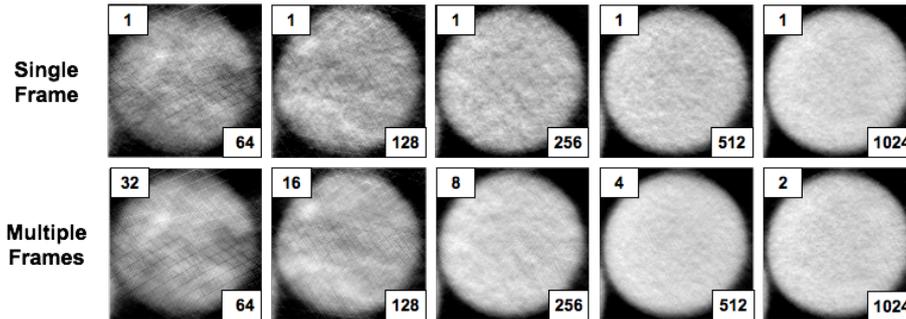


Figure 7: Reconstructions of the bead near the edge for all data set in Line B (single frame, top row) and Line A (multiple frames, bottom row).

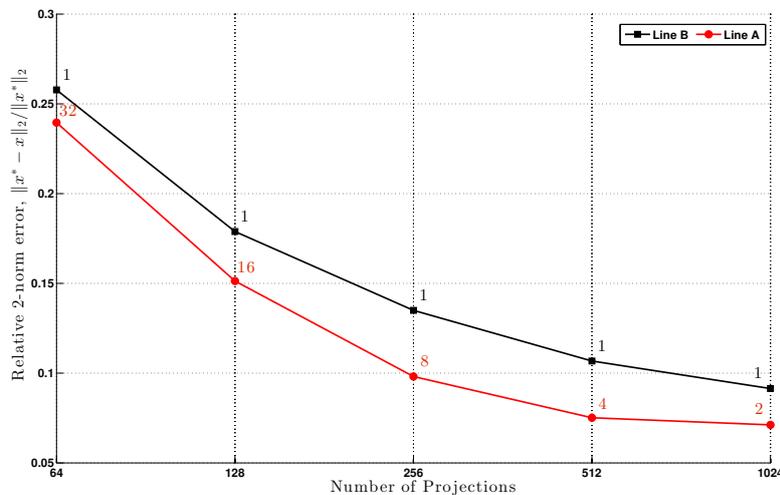


Figure 8: A log-linear plot of relative 2-norm error for all data sets. Each reconstruction is compared to the 2048-projection result obtained using the same numerical setup.

We observe that the traditional and the physical quantification measures agree, both suggesting that collecting multiple frames is giving better results (as expected due to decrease in noise) however, when closely inspected, the trade off is not as big as the intuition suggests, and collecting a single frame often gives results with comparable quality. This is observed in two locations. Firstly, compare the line A 512-projection (4 frames) and 1024-projection (2 frames) results in either graphs: Here, the total exposure is the same but the exposure-per-projection is different. However, the level of reconstruction quality at these points is almost the same, meaning 2 frames of 1024 projections does not have any benefit over collecting 4 frames of 512. Secondly, observe the difference between line B 512-projection (1 frame) and

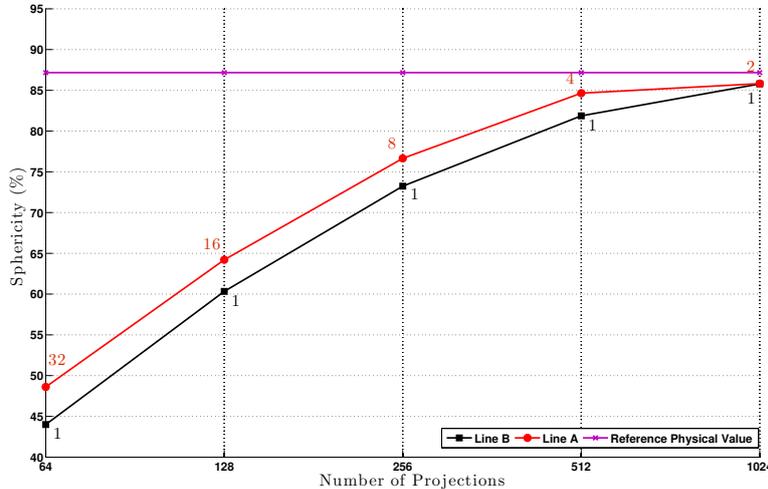


Figure 9: Sphericity of the reconstructed beads as a function of the number of projections. Reference ground truth (2048-projection FDK result) is denoted by the purple line with cross markers.

line A 256-projection (8 frames) results in the Sphericity plot. The line B solution with a single frame gives a closer quantification value to the reference value ($\approx 6\%$ difference in Sphericity) than the line A solution with 8 frames ($\approx 11\%$ difference). So despite having 4 times as much exposure-per-projection, the single frame data set still gives a better solution. This suggests that the SophiaBeads data sets were collected at too high an exposure level, reflecting a conservative tendency of CT practitioners to over-count per frame. A suggestion for future work could be looking at the limitations of iterative schemes where signal is conventionally deemed too short.

5.4. Simultaneous model-based reconstruction and segmentation – benchmarking implementations

This is an example for implementing a novel idea and benchmarking against more standard techniques to assess the accuracy and performance using a physical quantification measure. The idea implemented here is to use segmentation as part of the reconstruction algorithm. Although fairly new in practice, the idea was considered in a more theoretical sense in [21, 22]. This recently gained attention and is applied in more variations in the imaging community, namely to produce both continuous images using a Bayesian framework [32] or discrete images with a 2-step algorithm [3]. Also related is the idea of simultaneously obtaining both a reconstructed and a segmented image directly from the measured data [27], or using prior information about the phases in the sample [28].

The implementation in this study was inspired by the simple, 2-step discrete iterative reconstruction method detailed in [3], where a number of SART iterations are performed; the solution is segmented, boundary elements are determined; and SART iterations are continued for these boundary elements. Since the SophiaBeads sample only consists of uniform glass beads and air, the data sets are suitable for this study, in which we replace the SART method with CGLS (henceforth referred to as discrete-CGLS or DCGLS). This modification results in a large decrease in computational time. In [3], basic global thresholding is used, which

is a simple method that classifies pixels in an image as ‘black’ (zero) or ‘white’ (above a user-defined threshold). To benchmark this aspect of the algorithm, we replace this step with other segmentation methods. Namely, the Otsu’s thresholding algorithm, locally adaptive thresholding algorithm and the distance regularised level set evolution (DRLSE) method.

It is important to note that there were no ‘parameter fine-tuning’ tests done for these segmentation algorithms, except for DRLSE. Of all the methods implemented, DRLSE was the most unstable as the quality depended heavily on a number of parameters. As it is outside the scope of presented work, we do not go into the details of DRLSE but instead refer the reader to [19] for the derivation of the method (a summary is given in Appendix B along with the implementation details of this and other segmentation methods).

Initially 5 CGLS iterations were performed and the 5th iterated solution was used as the starting point for DCGLS. We chose a small number of initial iterations to challenge the segmentation methods, which were shown to return acceptable solutions for 256-projection data set and above (see §5.2). The DCGLS algorithm was applied to the 2D image (horizontal central slice) to reduce the computation time but the implementations are easily expendable to 3D volumes. The central windows of the reconstructions for 64- to 512-projection are given in Figure 10.

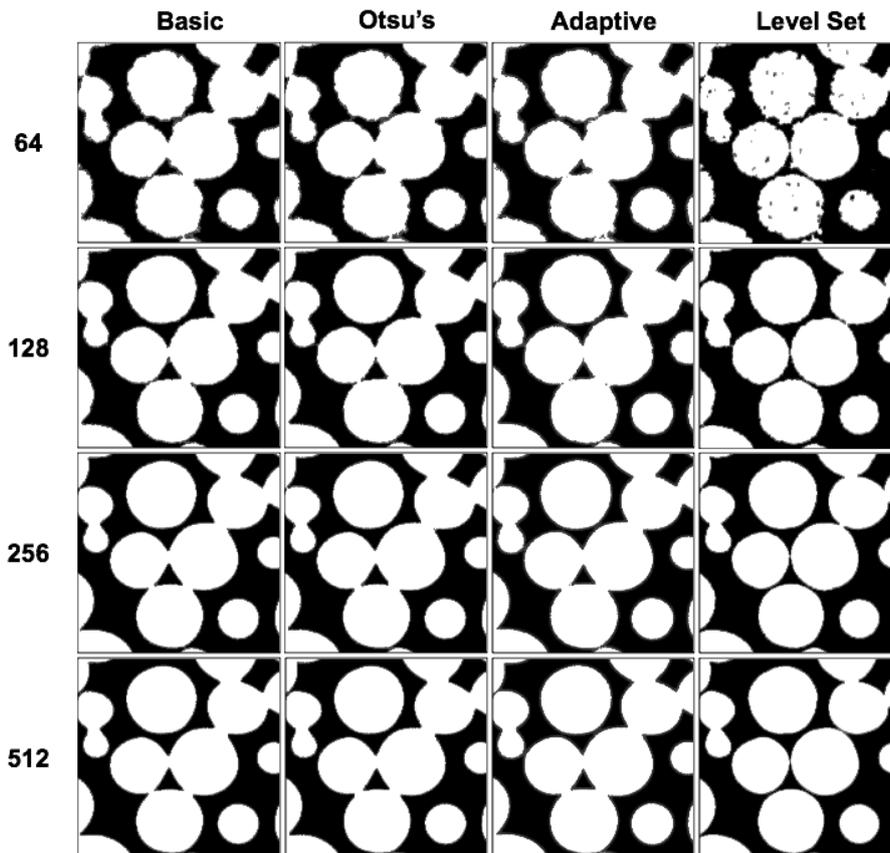


Figure 10: The central window of reconstructions for 64- to 512-projections data sets using (left to right) basic global thresholding (Basic), Otsu’s thresholding (Otsu’s), locally adaptive thresholding (Adaptive) and the distance regularised level set method (Level Set).

Traditional image analysis techniques are not feasible in this study since there is no segmented ground truth. We chose Shape2D as the physical quantification measure, which was applied to the horizontal central slice (Slice 1000). The results are given in Figure 11. The reference value is the mean Shape2D value of 2048-projection FDK reconstruction (see Table 1), denoted by a purple dashed line with cross markers.

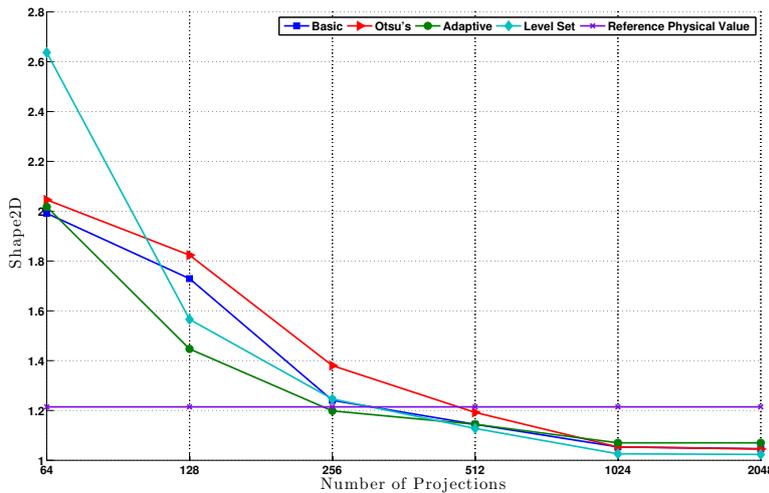


Figure 11: The log-linear plot of Shape2D quantification values for all segmentation methods for each data set, and of the full 2048-projection FDK reconstruction, included here as the reference value.

As seen in both the reconstructions and the Shape2D quantification, DCGLS with the Level Set (DRLSE) at 64-projection returns relatively poor results. This may be due to the initial reconstructed volume (CGLS, 5 iterations, recall first column in Figure 4) is not iterated long enough for the DRLSE to return a successful reconstruction. It could also be due to the DRLSE parameters being too strict for the few-projection case. Interestingly, for the same starting point, the most basic segmentation method returns better results, closely followed by the locally adaptive method. As the number of projections increase, the level set method takes over in quality, and appears to perform better than others.

Basic, locally adaptive and level set methods reach the same level of quality as the reference value at 256-projection, while Otsu's thresholding reaches this level with the 512-projection reconstruction. However, despite showing comparable Shape2D results to basic and locally adaptive methods, the level set method produces more accurate reconstructions. This conclusion is drawn from the observation that the level set segmentations show the beads as almost touching or completely separate, which is more accurate than what the others produce (beads reconstructed with other methods show merged or sintered beads for all cases, even at 2048-projection). We could of course have a stricter threshold value for the basic thresholding method, however this means more unstable results for the data sets with few projections.

6. Discussions and Further Analysis

The work presented here is applicable to various scenarios in imaging. As an example of this, as well as utilizing the SophiaBeads data sets, we tried to answer typical questions that

are of interest both to the mathematicians and the experimental scientists working in CT. In this section, we include our summary and conclusions for each problem, and further analysis into comparison of algorithms.

The first case study is a comparison of three types of reconstruction algorithms, namely a direct method (FDK), an iterative method with no explicit prior (CGLS) and an iterative method with a strong prior (TV Prior). The results were unsurprising in that the iterative methods clearly gave better quality reconstructions, where the prior information had enhanced the quality and thus gave the best result. However, the results also highlighted an aspect often ignored, which is the computational time for reconstructions. It is expected that a more sophisticated algorithm would require longer time to reconstruct an image, however, the difference in run time should question the trade-off of achieving a higher quality result. An interesting study could be the computational time or complexity of algorithms (in flops) versus the reconstruction quality for various reconstruction methods.

The following problem was studying the convergence behaviour of a method in an alternative way, and comparing this to the standard practice of convergence analysis in mathematics. Results obtained via the standard error norms did not yield conclusive results, demonstrating that the standard convergence analysis may be appropriate for simulations but they are not applicable in practical cases. Studying the physical characteristics of a reconstructed volume to determine the convergence behaviour returns more trustworthy results, which can be verified by visual inspection.

The third case study was focused on the levels of photon count (in signal or dose) and their effects on the reconstruction quality. We questioned the benefit of collecting data at constant exposure (total dose or photon count) as opposed to varying exposure (decreasing photon count). The results highlighted the conventional practice in CT, which is that even the data set with least amount of information was collected with ‘comfortable’ exposure level in order to collect the ‘best projections’. This provided unsuitable setting for this case study, which showed that collecting at high exposure did not have much benefit over collecting at lower exposure. An interesting study would be to repeat this assessment with data sets collected in more challenging settings.

The final case study was an example of benchmarking an idea, which was combining the reconstruction and segmentation steps together. The strategies we introduced in this paper helped compare various segmentation strategies used in the 2-step algorithm, DCGLS. The results highlighted the importance of choosing parameters, which determined how unstable a method can be (especially in cases with few-projection data) and the accuracy of the reconstructed solutions.

Since the SophiaBeads data sets were collected specifically for comparing reconstruction algorithms, we include further analysis into the comparison presented in §5.1.

6.1. Exploring the limits of reconstruction algorithms

The SophiaBeads data sets are collected with varying number of projections while keeping the total dose (or photon count) for each data constant. This was done to ensure fair comparison in the way that no other artefacts are introduced in the data as not to convolute the solution and influence our conclusions. In §5.1 where we compare three types of reconstruction algorithms, the resulting reconstructions were tested against a shape criterion, in this case Shape3D, to determine how close the reconstructed volumes were to a ‘perfect solution’. Another thing that could be tested here is the time it takes to obtain these reconstructions.

Figure 3 has been annotated to include the reconstruction times, but these can also be presented in a graph in a similar manner. An example is given in Figure 12. This presentation of results highlights the amount of time it takes to obtain a reconstruction, and raises important questions such as when does it make sense to implement more sophisticated methods and thus run longer reconstructions? Because it will take longer to reconstruct an image if there is more information, when does it become unnecessary to collect more information?

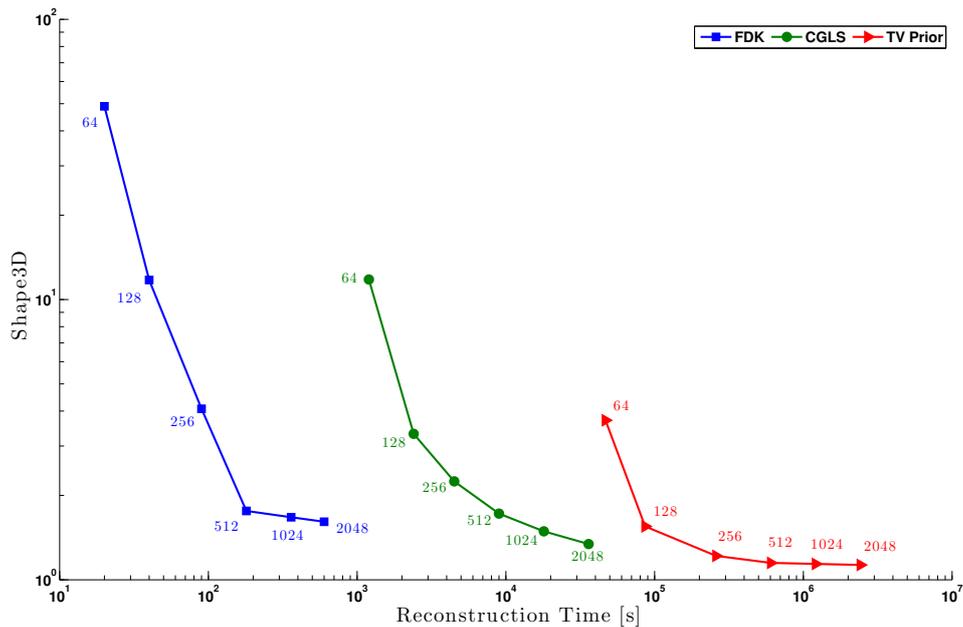


Figure 12: A log-log plot of the Shape3D results against the computational run time for the reconstruction of each SophiaBeads data set.

It is clear from Figure 12 that if we were to lower the dose, we can achieve improved results collecting 128 projections and reconstructing using TV Prior (in comparison to FDK). A similar degree of quality can be achieved with 512 projections, using CGLS, which would lead to a runtime that is approximately 10 times shorter than to obtain the TV Prior 128-projection reconstruction.

It is important to remind the reader that the CGLS reconstruction results may be improved with simple or model-based regularization, and that the TV Prior reconstruction run time can be reduced by fine-tuning the parameters to avoid stalling in convergence. We observed stalling occurred more often with larger amounts of data, but in order to ensure that the TV Prior results were comparable with one another, we have used the same parameters for all TV Prior reconstructions. Similarly with CGLS, we chose not to include regularization in order to make a comparison between an iterative method with no prior information and one with model-based prior information (CGLS vs TV Prior).

7. Conclusions

Image analysis is crucial in assessing the suitability of a reconstruction method and the accuracy of information from a reconstructed solution. Traditional techniques in image analysis (signal-to-noise ratio, l_2 norm, etc) were not adequate or able to give unbiased conclusions, and an established measure of merit was needed.

In this paper we introduced the idea of analysing a reconstructed solution using its characteristics such as shape, size and texture, referred to as physical quantification technique. We applied the measures to typical problems in CT, where we reconstructed real experimental data sets, made available via the SophiaBeads Dataset Project. This project, released for the purpose of developing and testing reconstruction methods, contains data sets taken at varying signals (or levels of dose), and the scanned sample comprises of a simple plastic tube packed with soda-lime glass beads.

We demonstrated the uses of physical quantification techniques through a number of case studies. These were chosen to be typical problems in CT in order to offer perspectives of mathematicians and experimentalists. Each study contained an example of physical characteristics being analysed, and a traditional analysis method is also included (where appropriate). This was done to highlight any possible discrepancies in conclusions drawn from both physical and traditional quantification techniques.

Presented results for each case showed that the physical quantification techniques are applicable to a wide range of problems, and offer clear results when a successful reconstruction is defined. In our case this was a full (unbroken), segmentable bead that is not connected to its neighbours.

The results also showed that the introduced techniques allowed further analysis into the comparison of methods and analysis of various levels of dose in the acquisition. With these techniques we were able to conclude that the established beads data set were collected at a conservative exposure level, which is the standard practice amongst the scientific experimentalists working in CT. A future task could be to collect a reference data set in more challenging settings, for example, when the projection view is severely limited or when the exposure-per-projection is decreased.

In addition, the results underlined the difference in computation times for a reconstruction, which was that the more sophisticated an algorithm gets, the longer the reconstruction takes. Another future task can be to limit the reconstruction time to assess the quality of results, with the goal of developing a sophisticated algorithm (correct forward model, appropriate stopping criterion, explicit prior and sensible optimization solver) that returns a reconstruction in relatively shorter amounts of time.

Acknowledgements

The authors would like to acknowledge the funding and technical support from BP through the BP International Centre for Advanced Materials (BP-ICAM), School of Mathematics and EPSRC CCPi (EP/J010456/1), which made this research possible. Funding for the Henry Moseley X-ray Imaging facility was provided by EPSRC (through grants EP/M010619, EP/K004530, EP/F007906, EP/F028431). Authors also acknowledge the guidance of Dr. Samuel A. McDonald and Dr. Julia Behnsen from Manchester X-ray Imaging Facility (MXIF).

References

- [1] T. Aste, M. Saadatfar, and T.J. Senden. Geometrical structure of disordered sphere packings. *Phys. Rev. E*, 71:061302, Jun 2005.
- [2] Avizo. *Avizo User's Guide*. FEI Visualization Sciences Group.
- [3] K.J. Batenburg and J. Sijbers. DART: A practical reconstruction algorithm for discrete tomography. *IEEE Trans Image Process*, 20(9):2542–2553, 2011.
- [4] J. Bernsen. Dynamic thresholding of gray level images. *ICPR'86: International Conference on Pattern Recognition*, SMC-9:1251–1255, 1986.
- [5] S. Beucher and F. Meyer. The morphological approach to segmentation: the watershed transformation. In *Mathematical Morphology in Image Processing*, page 433481, 1993.
- [6] S. B. Coban. SophiaBeads Dataset Project Codes, April 2015.
- [7] S.B. Coban. SophiaBeads Datasets Project Documentation and Tutorials. Technical report, School of Mathematics, University of Manchester, United Kingdom, April 2015. MIMS Technical Report.
- [8] S.B. Coban and S.A. McDonald. SophiaBeads Dataset Project, March 2015.
- [9] S.B. Coban, P.J. Withers, W.R.B. Lionheart, and S.A. McDonald. When do iterative reconstruction methods become worth the effort? In *the proceedings of The 13th International Meeting of Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, June 2015.
- [10] B. Cortet, D. Chappard, N. Boutry, P. Dubois, A. Cotten, and X. Marchandise. Relationship between computed tomographic image analysis and histomorphometry for microarchitectural characterization of human calcaneus. *Calcified Tissue International*, 75(1):23–31, 2004.
- [11] Y. Davit, G. Iltis, G. Debenest, S. Veran-Tissoires, D. Wildenschild, M. Gerino, and M. Quintard. Imaging biofilm in porous media using x-ray computed microtomography. *Journal of Microscopy*, 242(1):15–25, 2011.
- [12] A. Davour, S.J. Svrđ, P. Andersson, S. Grape, S. Holcombe, P. Jansson, and M. Troeng. Applying image analysis techniques to tomographic images of irradiated nuclear fuel assemblies. *Annals of Nuclear Energy*, 96:223 – 229, 2016.
- [13] K.J. Dobson, S.B. Coban, S.A. McDonald, J. Walsh, R. Atwood, and P.J. Withers. 4D imaging of sub-second dynamics in pore-scale processes using real time synchrotron x-ray tomography. *Solid Earth Discussions*, 2016:1–27, 2016.
- [14] L.A. Feldkamp, L.C. Davis, and J.W. Kress. Practical cone-beam algorithm. *JOSA A*, 1(6):612—619, 1984.
- [15] J.J. Friel. *Practical Guide to Image Analysis*. ASM International, 2000.

- [16] T.L. Jensen, J.H. Jørgensen, P.C. Hansen, and S.H. Jensen. Implementation of an optimal first-order method for strongly convex total variation regularization. *BIT*, 52:329–356, 2012.
- [17] R.A. Ketcham. Computational methods for quantitative analysis of three-dimensional features in geological specimens. *Geosphere*, 1:32–41, 2005.
- [18] C. Li, R. Huang, Z. Ding, C. Gatenby, D.N. Metaxas, and J.C. Gore. A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. *IEEE Trans. Image Process.*, 20(7):2007–2016, July 2011.
- [19] C. Li, C. Xu, C. Gui, and M.D. Fox. Distance regularized level set evolution and its application to image segmentation. *Image Processing, IEEE Transactions on*, 19(12):3243–3254, 2010.
- [20] Ajay Limaye. Drishti: a volume exploration and presentation tool. *Proc. SPIE 8506, Developments in X-Ray Tomography VIII*, 2012.
- [21] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1985.
- [22] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [23] J. Murison, R. Moosavi, M. Schulz, B. Schillinger, and M. Schröter. Neutron tomography as a tool to study immiscible fluids in porous media without chemical dopants. *Energy & Fuels*, 29(10):6271–6276, 2015.
- [24] S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12 – 49, 1988.
- [25] N. Otsu. A threshold selection method from gray level histograms. *IEEE Transactions on Systems Man and Cybernetics*, SMC-9:62–66, 1979.
- [26] Å. Björck, T. Elfving, and Z. Strakos. Stability of conjugate gradient and Lanczos methods for linear least squares problems. *SIAM J. Matrix Anal. Appl.*, 19(3):720–736, 1998.
- [27] R. Ramlau, E. Klann, and W. Ring. Simultaneous reconstruction and segmentation for tomography data. *PAMM*, 7(1):1050303–1050305, 2007.
- [28] M. Romanov, A.B. Dahl, Y. Dong, and P.C. Hansen. Simultaneous tomographic reconstruction and segmentation with class priors. *Inverse Problems in Science and Engineering*, 0(0):1–22, 0.
- [29] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1–4):259–268, 11 1992.
- [30] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60:259–268, 1992.

- [31] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
- [32] M.K. Schneider, P.W. Fieguth, W.C. Karl, and A.S. Willsky. Multiscale methods for the segmentation and reconstruction of signals and images. *IEEE Transactions on Image Processing*, 9(3):456–468, 2000.
- [33] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit*. Kitware, 4th edition, 2006. ISBN 978-1-930934-19-1.
- [34] J. Solomon and E. Samei. Quantum noise properties of CT images with anatomical textured backgrounds across reconstruction algorithms: FBP and SAFIRE. *Medical Physics*, 41(9), 2014.
- [35] I. Straumit, S.V. Lomov, and M. Wevers. *Analysis and Segmentation of a Three-Dimensional X-ray Computed Tomography Image of a Textile Composite*, pages 133–142. Springer International Publishing, 2014.
- [36] M. Tsukahara, S. Mitrovic, V. Gajdosik, G. Margaritondo, L. Pournin, M. Ramaioli, D. Sage, Y. Hwu, M. Unser, and Th.M. Liebling. Coupled tomography and distinct-element-method approach to exploring the granular media microstructure in a jamming hourglass. *Phys. Rev. E*, 77:061306, Jun 2008.
- [37] J.Y. Vaishnav, W.C. Jung, L.M. Popescu, R. Zeng, and K.J. Myers. Objective assessment of image quality and dose reduction in CT iterative reconstruction. *Medical Physics*, 41(7), 2014.
- [38] H. Wadell. Volume, shape, and roundness of quartz particles. *The Journal of Geology*, 43(3):250–280, 1935.

Appendix A. Iterative Reconstruction Techniques

Conjugate Gradient Least Squares

The Conjugate Gradient Least Squares (CGLS) method is the modification of the well-known Conjugate Gradient [31] where the CG method is applied to solve the least squares problem $A^T Ax = A^T b$. Here, $A \in \mathbb{R}^{m \times n}$ is the geometry matrix, $b \in \mathbb{R}^{m \times 1}$ is the measured data and $x \in \mathbb{R}^{n \times 1}$ is the reconstruction. CGLS is a popular method amongst those working in signal and image processing for its simple and computationally inexpensive implementation and fast convergence. The method is given in Algorithm 1.

Our implementation also includes a non-negativity step (negative pixel values equal to zero), applied to the final iterated solution. There is no parameter-tuning done for this implementation since the only user-defined parameter is the maximum number of iterations, k_{max} .

Algorithm 1. CONJUGATE GRADIENT LEAST SQUARES (CGLS)

Given a geometry matrix, A , a data vector b and a zero solution vector $x_0 = 0$ (a black image) as the starting point, the algorithm below gives the solution at k^{th} iteration.

```

1 Initialise the direction vector as  $d_0 = A^T b$ .
2 for  $k = 1: k_{max}$ 
3    $q_{k-1} = A d_{k-1}$ ,  $\alpha = \|d_{k-1}\|_2^2 / \|q_{k-1}\|_2^2$ 
4   Update:  $x = x + \alpha d_{k-1}$ ,  $b = b - \alpha q_{k-1}$ 
5   Reinitialise:  $q_k = A^T q_{k-1}$ ,  $\beta = \|q_k\|_2^2 / \|d_{k-1}\|_2^2$ ,  $d_k = q_k + \beta d_{k-1}$ 
6 end

```

Total Variation Regularization

The Total Variation Regularization (TV Prior) [30, 29] solves the optimization problem

$$x = \arg \min_x \frac{1}{2} \|b - Ax\|_2^2 + \alpha T_\tau(x), \quad x \geq 0. \quad (\text{A.1})$$

Here, $T_\tau(x)$ is the standard Huber-smoothed total variation, defined as

$$T_\tau(x) = \sum \Phi_\tau(\|D_j x\|_2), \quad \text{where,}$$

$$\Phi_\tau(u) = \begin{cases} |u| - \frac{1}{2}\tau & \text{if } |u| \geq \tau, \\ \frac{1}{2\tau}u^2 & \text{otherwise,} \end{cases}$$

and α is the TV regularization parameter, τ is the Huber-smoothing parameter (penalizing the non-smoothness of Equation A.1), D_j is the finite difference approximation of the image gradient at pixel j , and $\|\cdot\|_2$ is the (Euclidian) 2-norm.

The implementation uses the TVReg toolbox [16], where we used the Gradient Projection Barzilai-Borwein (GPBB) method. Using TVReg toolbox to reconstruct SophiaBeads data sets required a short script (a wrapper for the toolbox to use the forward and back-projectors provided by the SophiaBeads project), written by the authors. This script is available upon request.

For the results presented in this paper, the TV Prior iterations were run until results stalled (small changes in the gradient) or diverged for one data set (512-projection) to determine the maximum iteration number and the regularization parameter, α . These were 2000 iterations and $\alpha = 1e - 2$. The Huber-smoothing parameter τ was chosen to be sufficiently small for all runs, $\tau = 1e - 6$. The parameters remained unchanged for other reconstructions in order to obtain comparable results.

Discrete Algebraic Reconstruction Technique (DART)

Discrete Algebraic Reconstruction Technique (DART) is introduced by [3] as a discrete iterative algorithm that takes a number of SART (Simultaneous Algebraic Reconstruction Technique) sweeps as its initial solution, segment and extract the location of the boundary; iterate for those boundary pixels whilst keeping the rest of the image fixed, thus reducing the number of unknowns to the number of boundary pixels. This reduces the computational costs and the amount of uncertainty in reconstruction, making DART a suitable method for limited angle/projection experiments.

In addition, DART algorithm influences the convergence of the reconstructions by applying the threshold and thus pushing the solution towards a more educated image [3]. So coupling

this algorithm with SART (or any ART invariants) work well as SART within DART converges quicker than a regular SART run. However we note here that the convergence and the quality of the final reconstructed image depend heavily on the quality of the initial image reconstructed by SART, as well as the prior information used about the number of phases in the sample and the threshold values for separating those phases.

For our runs, we replace SART with CGLS to reduce time and computational cost, thus referring to the method as Discrete-CGLS or DCGLS throughout the paper. The initial CGLS is performed 5 times, and the 5th iterated solution is used as the starting point for the segmentation stage. The segmentation method here is the basic global thresholding algorithm, see Appendix B for details. The segmentation part was run for 3 iterations, $i_{max} = 3$ with intermediate CGLS algorithm running for 5 iterations each. DCGLS algorithm is outlined below.

Algorithm 2. DISCRETE-CGLS (DCGLS)

```

1 Run Algorithm 1 for  $k_{max} = 5$  to obtain  $x_{i=0}^k$ .
2 Declare threshold values,  $t$ .
3 for  $i = 1: i_{max}$ ,
4     Segment  $x_{i-1}$  to obtain the binary image,  $s_i$ , using the threshold values;
5     Extract boundary pixels of  $s_i$ ;  $B_i$ , and the non-boundary pixels;  $F_i$ .
6     Create image  $y_i$ :
7     if  $j \in B_i$ ,
8          $y_i(j) = x_{i-1}(j)$ ;
9     else,
10         $y_i(j) = s_i(j)$ .
11    end
12    Run Algorithm 1 with  $y_i^{(k=0)}$  as the starting point, keeping the pixels in  $F_i$  fixed.
13    Combine the final solution  $y_i^{(k=5)}$  with the segmented image,  $s_i$ ;
14    Smooth over the pixels.
15 end

```

Appendix B. Segmentation Methods

Distance Regularized Level Set Evolution Method for Segmentation

The level set methods capture dynamic changes in topology of a function (called the level set function) that evolves from a basic shape. The idea was first proposed by [24] to track propagation in fluid dynamics, though the method has been applied to many other applications. One that gained recent popularity is the image segmentation: The biggest advantage of level set methods is that they can represent contours of complex geometries and are able to track changes in the shape such as splitting or merging, making this a suitable method for image segmentation.

A drawback of this type of methods is that the level set functional (LSF) can invite numerical errors during evolution, if not reinitialized periodically or penalized within the

energy function. The version of the level set methods we implement in this paper is called the Distance Regularized Level Set Evolution (DRLSE); a variational level set formulation with a penalty on the deviation of the LSF from a signed distance function, explained in great detail in [19, 18]. A script for the implementation is available (see [18]). The level set evolution is derived from an energy functional to drive the shape of function (while penalizing it for stable evolution), and uses an edge-based active contour model for image segmentation.

Let $\phi : \Omega \rightarrow \mathbb{R}$ be the LSF on a domain Ω . The DRLSE energy functional, derived in [19] is

$$\mathcal{E}(\phi) = \frac{\delta\phi}{\delta t} = \mu\mathcal{R}_p(\phi) + \lambda\mathcal{L}_g(\phi) + \alpha\mathcal{A}_g(\phi). \quad (\text{B.1})$$

Here, $\mathcal{R}_p(\phi)$ is the level set regularization term with the potential function $p : [0, \infty) \rightarrow \mathbb{R}$; $\mathcal{L}_g(\phi)$ is the energy functional which is minimized when the zero level contour is at the object boundaries, and $\mathcal{A}_g(\phi)$ is the energy functional that computes the weighted area of the region inside the level set. The edge indicator function, g , is

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2},$$

where G_σ is a Gaussian kernel with standard deviation, σ ; I is the image in domain Ω , and $*$ is the convolution operator. The function, g , takes smaller values at object boundaries than at other locations. The parameters $\mu > 0$, $\lambda > 0$ and $\alpha \in \mathbb{R}$ are the penalty parameters acting on the energy functionals, $\mathcal{R}_p(\phi)$, $\mathcal{L}_g(\phi)$ and $\mathcal{A}_g(\phi)$, respectively. The energy functionals are defined by

$$\begin{aligned} \mathcal{R}_p(\phi) &= \int_{\Omega} p(|\nabla\phi|)dx, \\ \mathcal{L}_g(\phi) &= \int_{\Omega} g\delta(\phi)|\nabla\phi|dx, \\ \mathcal{A}_g(\phi) &= \int_{\Omega} gH(-\phi)dx, \end{aligned}$$

where δ and H are the Dirac delta and Heaviside functions, respectively.

For the results in the paper, the function p is taken as the double-well potential, defined by

$$p(u) = \begin{cases} \frac{1}{(2\pi)^2}(1 - \cos(2\pi u)) & \text{if } u \leq 1, \\ \frac{1}{2}(u - 1)^2 & \text{if } u \geq 1. \end{cases}$$

The initial zero level, ϕ_0 , is chosen as a binary step function for efficiency, which is

$$\phi_0(\mathbf{x}) = \begin{cases} -c_0 & \text{if } \mathbf{x} \in R_0, \\ c_0 & \text{otherwise,} \end{cases}$$

where $c_0 > 0$, and R_0 is a region in domain Ω . The initial zero level set contour used in this paper is a rectangle of size 300px \times 300px, shown in Figure B.1 (top left), with the evolved level set contour (top right) providing the boundaries of the segmented image for line 5 in Algorithm 2.

The quality of the segmented image with this formulation depends on the set of parameters chosen. For the results presented in §5.4, the parameters were determined by trial and error. For those runs, we let $\mu = 2$, $\lambda = 1$, $\alpha = 0.1$, $\sigma = 4$ and $c_0 = 0.1$.

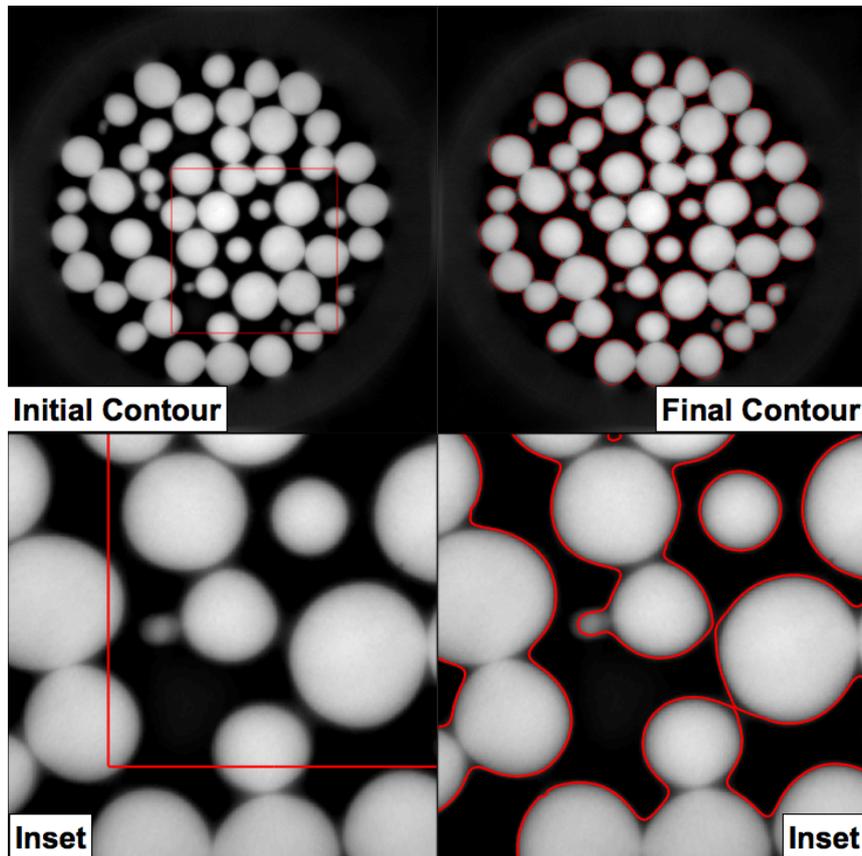


Figure B.1: The initial (top left) and final contour (top right) of the DRLS method at 2048-projection, with insets (100% zoom-in) in the bottom row. The initial contour is defined by the user.

Other Segmentation Techniques

Three more segmentation methods were used in the paper, one of which is used in the original DART algorithm. These are

- **Basic global thresholding method**

In its simplest form, this method replaces each voxel with 0 if $v_i < t$, or with the value of 1 if $v_i > t$ (binarization), where t is a user-defined thresholding value, often picked as the minimum between peaks in the histogram of the image. This method was taken as the standard in §5.4.

- **Otsu's thresholding method**

This is a more sophisticated global thresholding method, which fits each peak in a histogram to a Gaussian distribution, and (assuming there are multiple curves) finds the grey value at which the Gaussian curves intersect.

- **Locally adaptive thresholding method**

This technique is used for cases where a global threshold does not apply everywhere in the image as it fails to distinguish between the background and the foreground (i.e. when there is a continuous change in the gradient of an image). So an estimation of an optimal threshold is done in subsections of the image, and later applied to segment between the background and the sample.

The parameter-tuning was minimal for these techniques: The threshold parameter, t , in basic global thresholding was chosen as $t = 0.11$ (this is the same value used to segment in the Avizo visualisations, [7]). The locally adaptive thresholding method calculates a thresholding parameter locally, given a window size, which for these runs was chosen to be $10\text{px} \times 10\text{px}$. Otsu's thresholding automatically calculates a thresholding parameter depending on the histogram in the image, hence required no user input.

Appendix C. Quantification

A Summary of the Avizo Routine

The quantification and 3D visualization presented in this paper are done using the commercial software Avizo Fire, popularly used in materials science for interactive data and image visualization and analysis. Note that material given here is a short summary of [7], which also includes a tutorial for the project codes [6].

The project codes are programmed so after each reconstruction, the reconstructed image is saved as 32-bit single volume, which can be imported to Avizo as a 3D volume. The volume then goes through a number of operations in order for the beads to be segmented and separated using the watershed algorithm. Figure C.2 shows the outcome of stages of the Avizo workflow. The threshold value for the segmented image on left is taken as $t = 0.11$.

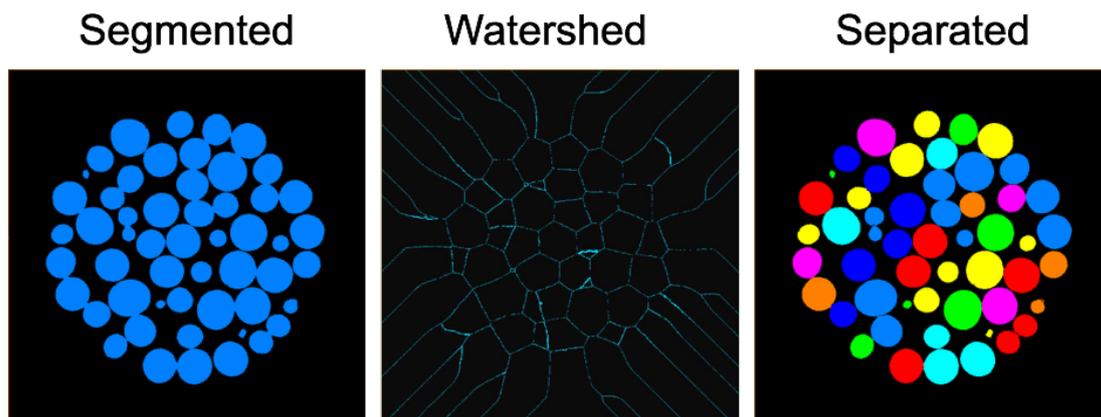


Figure C.2: The outputs of segmentation, watershed and separation of the reconstructed volume, displaying the Slice 1000 (xy-axis).

The watershed algorithm is a popular segmentation and separation tool used in image analysis. The algorithm attempts to find local minima in the image with the use of heightmaps and morphological operations, which help determine locations of markers in order to segment and separate objects in an image [5].

The Quantification and Reference Values

Since our reconstructions were performed on the central subvolume of size $1564 \times 1564 \times 500$, this meant that there were some beads not fully within the region of interest. Figure C.3 shows the reconstructed subvolume (CGLS 1024-projection, on left) where it is clear that applying the image measures on the entire subvolume would affect the analysis. To work around this, we used *Filter by Measure* module, along with the Volume3D quantification measure, to choose 25 biggest beads (in volume). By applying *Label Analysis* to these 25 beads, Avizo outputs the quantification values for the volume measures (Shape3D, AspectRatio3D, Sphericity) for each bead, along with the mean, standard deviation, minimum and maximum values of the distribution. For the planar (2D) quantification measures, we used the Area measure for filtering on the central horizontal slice to obtain a reference value.

Figure C.3 also shows a 3D rendering of the full FDK 2048-projection volume, which has 500+ beads and was used to obtain the reference values given in Table 1.

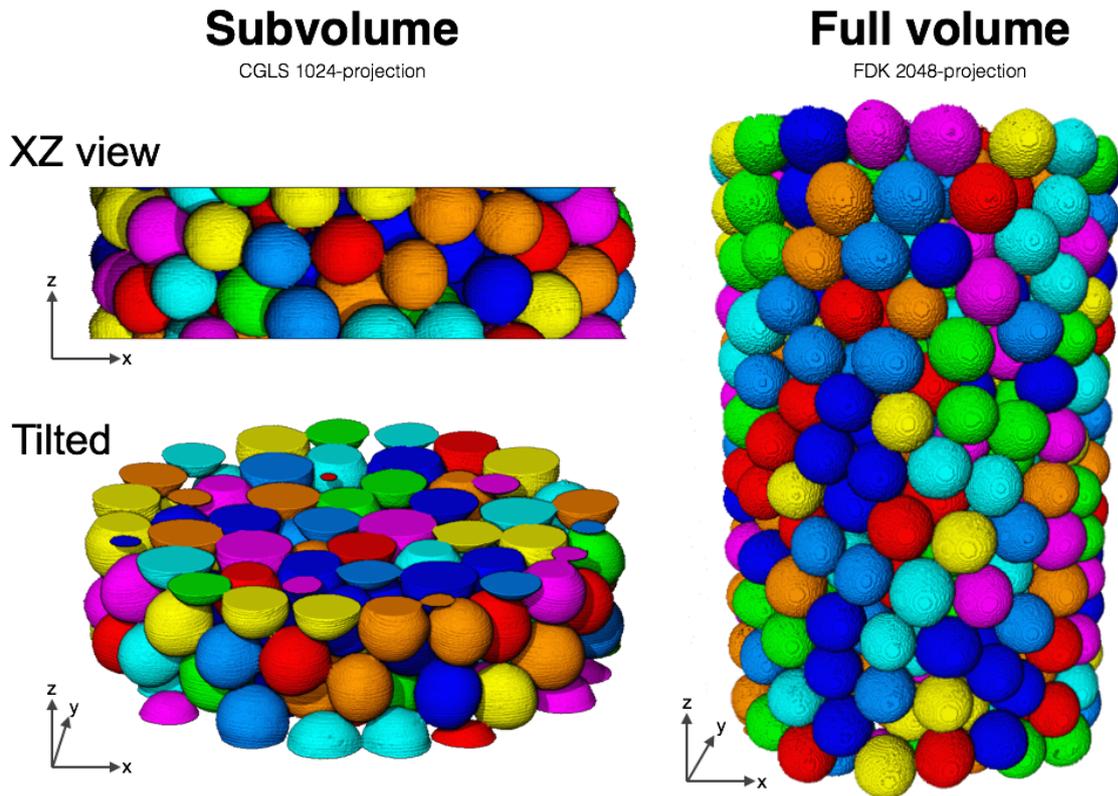


Figure C.3: The 3D renderings of the CGLS 1024-projection subvolume (centre $1564 \times 1564 \times 500$, on left with xz -axis and tilted view) and the full FDK 2048-projection volume (on right) used as a reference in the comparison studies.

Chapter 3

Predicting performance of
sparsity-regularized X-ray
tomography: experimental results
using glass-bead data

Predicting performance of sparsity-regularized X-ray tomography:
experimental results using glass-bead data

Jakob S. Jørgensen*, Sophia B. Coban†, William R. B. Lionheart†, Samuel A. McDonald‡,
Philip J. Withers‡

* *Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Kgs.
Lyngby, Denmark*

† *School of Mathematics, ‡ School of Materials, University of Manchester, Manchester, M13 9PL, UK*

Abstract

Sparsity regularization such as total-variation (TV) minimization has been demonstrated to allow accurate image reconstruction in X-ray Computed Tomography (CT) from fewer projections than standard analytical reconstruction methods. Exactly how few projections suffice and how this number may depend on the image remain poorly understood. While theoretical guarantees from Compressive Sensing (CS) connect the admissible reduction in number of projections directly to the image sparsity for certain imaging methods, this does not hold for CT, however empirical simulation work suggests a similar connection. The present work establishes for the first time for real CT data a connection between gradient sparsity and sufficient number of projections for accurate TV-regularized reconstruction.

Glass beads of 5 sizes (hence of different gradient sparsity levels) were scanned in a micro-CT scanner at increasing number of projections. The reconstruction quality was assessed as function of number of projections and gradient sparsity. The critical number of projections to obtain a satisfactory TV-regularized reconstruction was found to increase almost-linearly with the gradient sparsity. This establishes a quantitative guideline from which one may predict how few projections will be sufficient based on an expected sparsity level of a sample to be scanned. As a result, samples having a sparse gradient may benefit from an order of magnitude or more reduction in number of projections required, with clear benefits in reduced X-ray exposure or data acquisition time enabling fast time-resolved imaging.

Keywords: Computed tomography, inverse problems, sparsity, total variation, sampling

1. Introduction

X-ray Computed Tomography (CT) still predominantly uses image reconstruction methods of the analytical type, i.e., Filtered Backprojection (FBP) for parallel-beam geometry and Feldkamp-Davis-Kress algorithm for cone-beam geometry [1]. In recent years novel sparsity-regularized reconstruction methods based on an algebraic imaging model have been subject to extensive research, motivated by their ability to preserve or improve image quality from reduced data such as substantially fewer projections [2, 3, 4, 5, 6, 7]. Reconstruction from few projections is of high interest due to its potential to reduce patient X-ray exposure in medical imaging and shorter acquisition times to capture rapid material changes in non-destructive testing. The prototypical sparsity-regularized reconstruction method is Total Variation (TV)

regularization [8], which has the ability to reduce noise and subsampling artefacts while preserving sharp edges in the reconstruction. TV regularization can be interpreted as sparsity-regularization method because it encourages a sparse gradient image of the reconstruction.

Research in sparsity-regularized reconstruction has in part been motivated by new mathematical results in the field of Compressive Sensing (CS). In a nutshell, CS establishes that accurate image reconstruction is possible from a substantially reduced number of measurements. This is tied to the sparsity of the image, under certain assumptions on the imaging process such as incoherence and the restricted isometry property [9, 10]. Unfortunately the recovery guarantees of CS do not extend to cover X-ray CT [11, 12] and the success of sparsity regularization in CT therefore remains somewhat heuristic. In particular, little is known about conditions under which sparsity regularization will succeed for tomographic data. In CS the sufficient number of measurements for accurate image reconstruction using sparsity regularization methods depends directly on the the sparsity level of the image, i.e., the number of signal nonzero values or coefficients in some representation. It remains to be established whether a similar dependence holds in X-ray CT.

Previous work by one of the authors has addressed precisely this question in comprehensive studies with simulated data [13, 14, 15, 16]. Among other conclusions the results indicate that under some circumstances an almost identical relation between the sparsity level and the critical number of measurements can be observed in X-ray CT.

The goal of the present work is to test whether the same result can be verified with real X-ray CT data. Specifically, we target two questions:

- Q1** Can a relation between the critical number of projections for TV-regularized reconstruction and the image sparsity level be observed for real X-ray CT data?
- Q2** If yes, can the critical number of projections be predicted from the image sparsity level using the empirical phase-transition established in previous simulation work [13]?

In the conference contribution [17] we attempted to answer **Q1** using the SophiaBeads data set [18], which is a collection of micro-CT data sets of glass beads. While appropriate for its stated purpose of comparing reconstruction algorithms in a fixed-exposure trade-off study and in principle also interesting for sparsity regularization, we found the data set to be of limited use for addressing **Q1** due to too few sparsity and sampling levels available. This prompted us to acquire a new and more extensive glass bead data specifically designed for addressing this question. The current paper presents this data set and employs it to address **Q1** and **Q2**.

This paper is organized as follows. In Sec. 2 we outline the experimental setup for acquiring the new extensive glass bead data set. Sec. 3 gives details the computational setup we employ including TV-regularized reconstruction and the optimization algorithm to solve it as well as image quality measures and determination of reference images and sparsity levels. Sec. 4 presents the results of our studies and Sec. 5 offers a discussion of results before concluding the work in Sec. 6.

2. Experimental setup

2.1. Previous work with SophiaBeads

The SophiaBeads Dataset Project [18] is a collection of cone-beam X-ray CT data sets where the number of projections is varied while the total photon count (or the total exposure

Sample Code	B1	B2	B3	B4	B5
Glass bead diameters (mm)	5	3	1	0.600 – 0.425	0.300 – 0.212
Sparsity measure 1 (%) (without segmentation)	1.29	2.21	4.58	8.34	11.60
Sparsity measure 2 (%) (with segmentation)	0.48	0.85	2.25	4.56	7.90

Table 1: Listing of sample information and the corresponding sparsity levels (percentage of nonzero pixels in 2000×2000 reconstructed images).

time) is kept constant. For each acquisition, the number of projections was reduced by a factor of 2 to preserve the equiangular projections, starting with 2048, down to 64 projections. The total exposure time was kept constant for each acquisition to reduce the experimental noise and its effects on the analysis of reconstruction techniques [19].

The sample comprised of glass beads and air, packed in a plastic container. The simplicity and the fact that any cross section through the sample will be piecewise constant makes this suitable for examining the capabilities of sparse-regularization techniques, in particular TV-regularization.

While the data set is appropriate for performing a wide range of algorithm comparisons, results presented in [17] proved to be a limitation of the data sets as we achieved unclear results for the SophiaBeads data sets with 128 and 64 projections. The study was inconclusive due to the nature of the data sets, namely the inflexibility of subsampling projections to analyse for data content higher than 64 projections but lower than 128 projections.

Another shortcoming of the SophiaBeads data sets was that there were only two distinct sample sizes available (1mm and 2.5mm diameters), meaning we were limited to two sparsity levels (5.1% and 2.2%, respectively).

Therefore, using our experiences and the acquisition plan of SophiaBeads data sets as our template, we designed and performed a new set of experiments, which we refer to as SparseBeads, specifically targeted at assessing sparsity-regularized reconstruction at reduced number of projections.

2.2. Design of SparseBeads experiments and acquisition details

The preservation of the equiangular projections and flexible subsampling motivated the use of a *highly composite number*, which refers to a positive integer that has more divisors than any smaller number. A realistic highly composite number for the purpose of our study was 2520, which was chosen to be the maximum number of projections collected over 360° .

Similar to the SophiaBeads experiments, the SparseBeads samples consisted of a simple plastic tube packed with uniform Soda-Lime glass ($\text{SiO}_2\text{-Na}_2\text{O}$) beads¹, of which there were 5 different sizes. The diameters of these bead samples, along with the assigned experimental code for each sample, are given in Table 1. Each sample was scanned at 3 separate vertical positions. The table also specifies 2 sets of estimates of the gradient sparsity levels, i.e., the fraction of nonzero values in the gradient image of each bead. The gradient sparsity

¹Lead-free Soda-Lime glass beads distributed by BioSpec Products, Inc. Address: PO Box 788, Bartlesville, OK 74005. Website: www.biospec.com

estimation procedures are described in Sec. 3.2. At this point it suffices to notice that – as intended – both sets of gradient sparsity levels increase when the bead diameter is decreased.

The SparseBeads data sets were collected using the *320/225 kV Nikon XTEK Bay*, located at the Manchester X-ray Imaging Facility (MXIF), the University of Manchester. The apparatus consists of a cone-beam microfocus X-ray source that projects polychromatic X-rays onto a 2000×2000 pixel-length and width, 16-bit, flat detector panel. All experiments were performed with 1000ms exposure. The projections were collected as part of a batch scan, where the stop-start technique was used to reduce the angular blurring. To minimize possible beam hardening artefacts, a 0.1mm copper filter was used during data acquisition.

3. Computational setup

3.1. Image quality assessment

Meaningful image quality assessment is a complex task. In principle, there is no general notion of image quality. As noted by Barrett [20] it is most meaningful to assess the quality of an image in relation to a particular task that it is intended to solve. A large body of this work addresses task-based image quality assessment in medical imaging (see [20] and references therein), where scans are generally performed with a specific purpose in mind, such as detection of certain features of interest, and the task is thus well-defined. In addition, [21] makes the same argument applied to a more general context.

The present work is not motivated by a specific practical imaging task but from trying to establish theoretical connections observed in the recent work based on simulation studies [13, 14, 15, 16, 17]. As image quality metric this previous work used the relative 2-norm error (closely related to the root-mean-square error) defined as

$$E_2(x) = \frac{\|x - x_{\text{ref}}\|_2}{\|x_{\text{ref}}\|_2},$$

which is a generic and widely used quality metric and in particular appropriate for measuring if an image exactly matches the reference image, which was the question of interest. While recognizing that exactly matching the reference image is not possible for real data we employ this metric here in order to keep the connection to the previous work. Note that only pixels within the disc-shaped field of view are included in the computation.

The relative 2-norm error may however not always capture fully if important features have been faithfully reconstructed. We therefore also employ additional quality measures which directly quantify whether certain geometric properties are preserved, as demonstrated in [21]. We do this by selecting a number of glass beads from the segmented reconstruction and do a quantification of the relevant properties using the software Avizo. We then compare with the same property measured on the same beads in the reference image and define the associated quality measure to be

$$E_Q = \frac{|M(x) - M(x_{\text{ref}})|}{M(x_{\text{ref}})},$$

where M here indicates the chosen geometry measure. In the present case of approximately spherical glass beads we will assess the following three geometric properties:

- Aspect ratio: The ratio of the shortest to longest measured diameters of a bead.

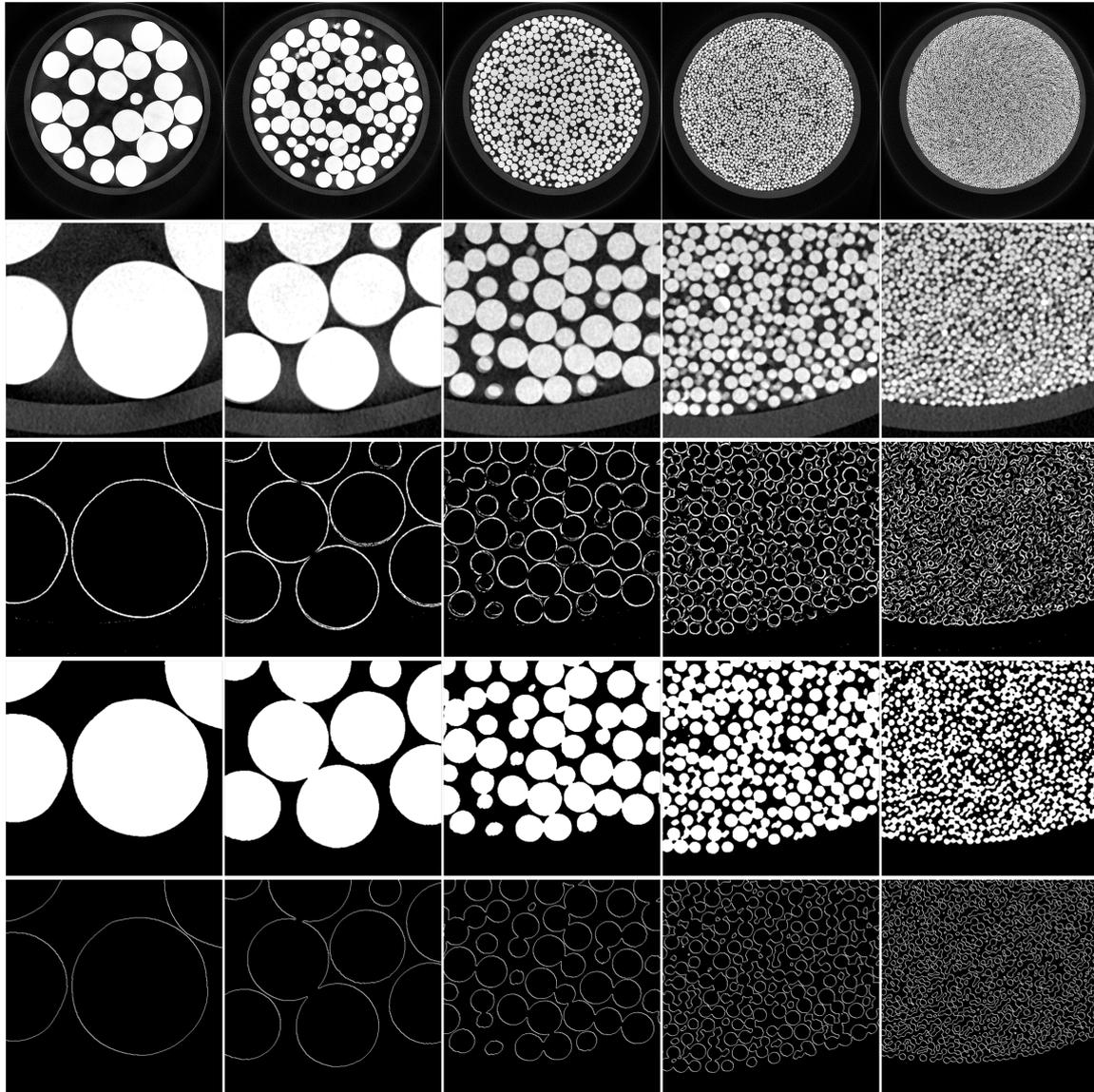


Figure 1: Reference images for layer L2 of bead sizes B1–B5 obtained as 40 CGLS iterations followed by median filtering. 1st row: Full image. 2nd row: zoom to 450×450 pixel region. 3rd row: Thresholded gradient magnitude images, threshold value 2×10^{-4} , basis for sparsity measure 1. 4th row: Manually segmented images. 5th row: Gradient magnitude of manually segmented images, basis for sparsity measure 2.

- Shape: The ratio of the surface area of the object to the volume of a sphere. This can also be thought as the cubic-root of the sphericity measure defined in [22].
- Volume: The bead volume determined from the number of pixels belonging to a bead.

The resulting quality measures will be referred to as E_A , E_S and E_V . As a qualitative supplement to the quantitative measures, visual inspection will also be employed.

3.2. Reference image and gradient sparsity level determination

A reference image is required to quantitatively assess reconstruction errors. In this study the reference image is obtained as the 40-iteration CGLS reconstruction using all 2520 projections, followed by the application of a 5-by-5 median filter. The filtering is used to reduce the noise present in the reconstruction in order to have a reference image that is more piecewise constant, as expected from a sample consisting of glass beads, plastic and air. The filter size was chosen empirically by determining the best trade-off between noise reduction and preservation of spatial resolution across all bead samples and test filter sizes 3, 5, 7, 9, 11. While the larger beads allow large filters, the smaller ones are subject to extensive oversmoothing/loss of spatial resolution with filter sizes 7 and above, prompting the use of filter size 5×5 pixels. In Fig. 1 the selected reference images are shown along with gradient images.

Two methods are used for determining the gradient sparsity level of the reference images. Measure 1 computes the gradient magnitude at each pixel using forward differencing in each direction followed by squaring, summing directions and taking the square root. Since the images are obtained as reconstructions from real data some noise is present, which yields small nonzero values that should not be counted as part of the bead edges. For this reason only values larger than an empirically determined threshold are counted, and the threshold value was set to 2.0×10^{-4} . The thresholded gradient magnitude images capture exactly the edges between beads and air in all cases. The edge between the plastic container and air is below the threshold and thus corresponding pixels are not counted towards the sparsity level. This method may have a tendency to overestimate the number of gradient nonzeros, since several pixels across any given boundary will be counted due the transition from background to bead stretching over a few pixels. In principle the transition is sharp and to better capture this Measure 2 introduces a segmentation of the reference image before forward differencing. In this way a single band of pixels will have nonzero gradient and estimated sparsity values be lower and possibly more representative of the actual beads. The sparsity level values in Table 1 are given as the percentage of nonzeros relative to the total number of pixels, 2000^2 .

3.3. Total variation regularization and algorithm

We denote the log-transformed projection data by b , the 2D fan-beam system matrix by A , a reconstructed image by x , and the number of projections by N_θ .

To determine a TV-regularized reconstruction (which can be seen as the maximum a posteriori estimate in a Bayesian formulation) of the discrete imaging model $Ax = b$ we solve the optimization problem

$$x_{\text{TV}} = \arg \min_x \frac{1}{2N_\theta} \|b - Ax\|_2^2 + \alpha T_\tau(x), \quad (1)$$

where x_{TV} is the TV-regularized solution, α is the TV regularization parameter and $T_\tau(x)$ is the standard Huber-smoothed total variation, defined as

$$T_\tau(x) = \sum_j \Phi_\tau(\|D_j x\|_2), \quad \text{where}$$

$$\Phi_\tau(u) = \begin{cases} |u| - \frac{1}{2}\tau, & \text{if } |u| \geq \tau, \\ \frac{1}{2\tau}u^2, & \text{otherwise.} \end{cases}$$

Here, τ is the Huber-smoothing parameter, D_j is the finite difference approximation of the image gradient at pixel j , and $\|\cdot\|_2$ is the (Euclidean) 2-norm.

Smoothing is used to ensure a solution by smooth optimization techniques, which are generally faster than their non-smooth counterparts. However, depending on the choice of smoothing parameter, τ , this might modify the reconstruction. In our implementation, we use a sufficiently small value of τ relative to the image values that the smoothing effects are negligible. This is discussed in more detail in next section.

The normalization by N_θ helps to compare reconstructions obtained at different N_θ by compensating the magnitude of the first term. As a result, a fixed α value yields the same balance between the two terms at different N_θ , which reduces the range of relevant α values, making it more practical to find the optimal value.

A disc shaped masking (covering the beads and the walls of the sample holder) was used to ensure a reconstruction of the relevant field of view, and to reduce the number of unknown pixel values and thus the computation time. The pixels outside the field of view were set to zero since there are no other objects but air outside the sample holder. The masking was applied by determining the pixel indices within the sample holder, and passing only those pixels to the forward and back-projector.

3.4. TV regularization parameter selection

When it comes to parameters, CGLS only requires a number for iterations to be performed, and FDK requires a choice for filter. However for TV regularization, we need to determine a number of parameters. These can be divided into two categories as also pointed out by [23]: problem-specific parameters and algorithm parameters. The problem-specific parameters are chosen to refine the mathematical solution in the optimization problem. In the present case the problem-specific parameters are the regularization parameter α and the Huber-smoothing parameter τ , as they occur in (1). The algorithm parameters specify the behaviour of the optimization algorithm used to solve the mathematically defined problem. In our case, we use the GPBB (Gradient Projection with Barzilai-Borwein acceleration) method from the TVReg [24] MATLAB[®] package and the noteworthy algorithm parameters are the maximum numbers of iterations to perform and the threshold value used in the convergence criterion; in addition a number of internal algorithm parameters, for example regarding step-size selection can also be chosen.

For problem-specific parameters, the Huber-smoothing parameter is fixed at $\tau = 10^{-5}$ in all studies. A smaller value of τ means less smoothing and thus a better approximation to the ideal non-smoothed TV-regularization problem (which corresponds to $\tau = 0$). On the other hand, convergence speed of the algorithm decays as τ is decreased [24], so in practice a trade-off must be used. To obtain a good approximation to the non-smoothed TV, it is sufficient that τ is small relative to the smallest difference between pixel values, which in the present study was of the order of 10^{-4} – 10^{-3} . Further, we empirically assessed the effect of τ by comparing reconstructions at $\tau = 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$. Large differences (much smoother images) were evident at higher τ values, while at $\tau = 10^{-5}$ extremely small differences could be observed compared to the smaller values.

The regularization parameter α balances how strongly TV regularization is enforced compared to minimizing the data fidelity term alone. Larger values of α yield a smoother image while smaller values yield images exhibit no regularization, see Fig. 2. The optimal regularization parameter α , i.e., the value that yields in some sense the best reconstructed image, depends on many non-trivial factors including noise, number of projections, and scaling of the image pixel values. In the present study the optimal parameter was selected as the one yielding the smallest relative 2-norm error of TV reconstructions across a range of α -values.

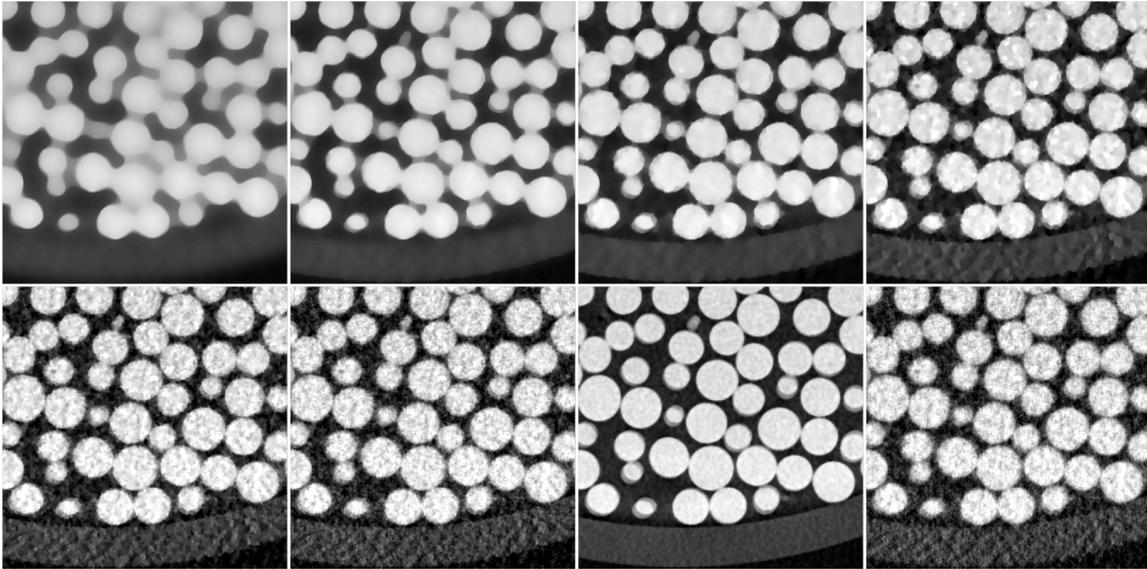


Figure 2: Effect of TV regularization parameter α on B3L2 image at 252 projections ($ds = 10$). Top, left to right: $\alpha = 3.16 \times 10^{-2}, 1.0 \times 10^{-2}, 3.16 \times 10^{-2}, 1.0 \times 10^{-3}$. Bottom: $\alpha = 3.16 \times 10^{-4}, 1.0 \times 10^{-4}$, reference image ($ds = 1$), CGLS reconstruction ($ds = 10$) for comparison.

The reconstruction corresponding to the optimal value will simply be referred to as the TV-regularized reconstruction.

Algorithm parameters were chosen to ensure an accurate solution to the TV-regularization problem, while preventing excessive computing times. In practice we observed negligible changes to the image beyond 10,000 iterations and therefore used this number as the maximum number of iterations to run. In addition we found it useful to limit unnecessary iterations to save time by terminating if the norm of the gradient of the objective function in (1) became sufficiently small. In practice we used the `epsb_rel` input of TVReg to 10^{-7} , as no noticeable changes were observed in tests beyond this value.

4. Results

4.1. Visual inspection of reconstructions at reduced numbers of projections

In total 2520 equiangular projections were acquired during each tomographic scan. This number is a highly composite number, meaning it has a large number of divisors, implying that a large number of perfectly equiangular subsets of projections can be extracted. Specifically 2520 has 48 divisors: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 15, 18, 20, 21, 24, 28, 30, 35, 36, 40, 42, 45, 56, 60, 63, 70, 72, 84, 90, 105, 120, 126, 140, 168, 180, 210, 252, 280, 315, 360, 420, 504, 630, 840, 1260, 2520. It is intuitively convenient to express the number of projections N_θ used along with the corresponding downsampling factor f_{ds} , i.e., $N_\theta = 2520/f_{ds}$, for example, 252 projections corresponds to a downsampling factor of 10, $ds = 10$.

All bead sizes B1–B5 each with three layers L1–L3 are TV-reconstructed from projection data at the downsampling factors listed above. Fig. 3 shows a representative selection of (zoomed-in) reconstructed images. Our previous simulation work showed that the undersampling factor at which accurate TV-reconstruction is admitted grows when images become more

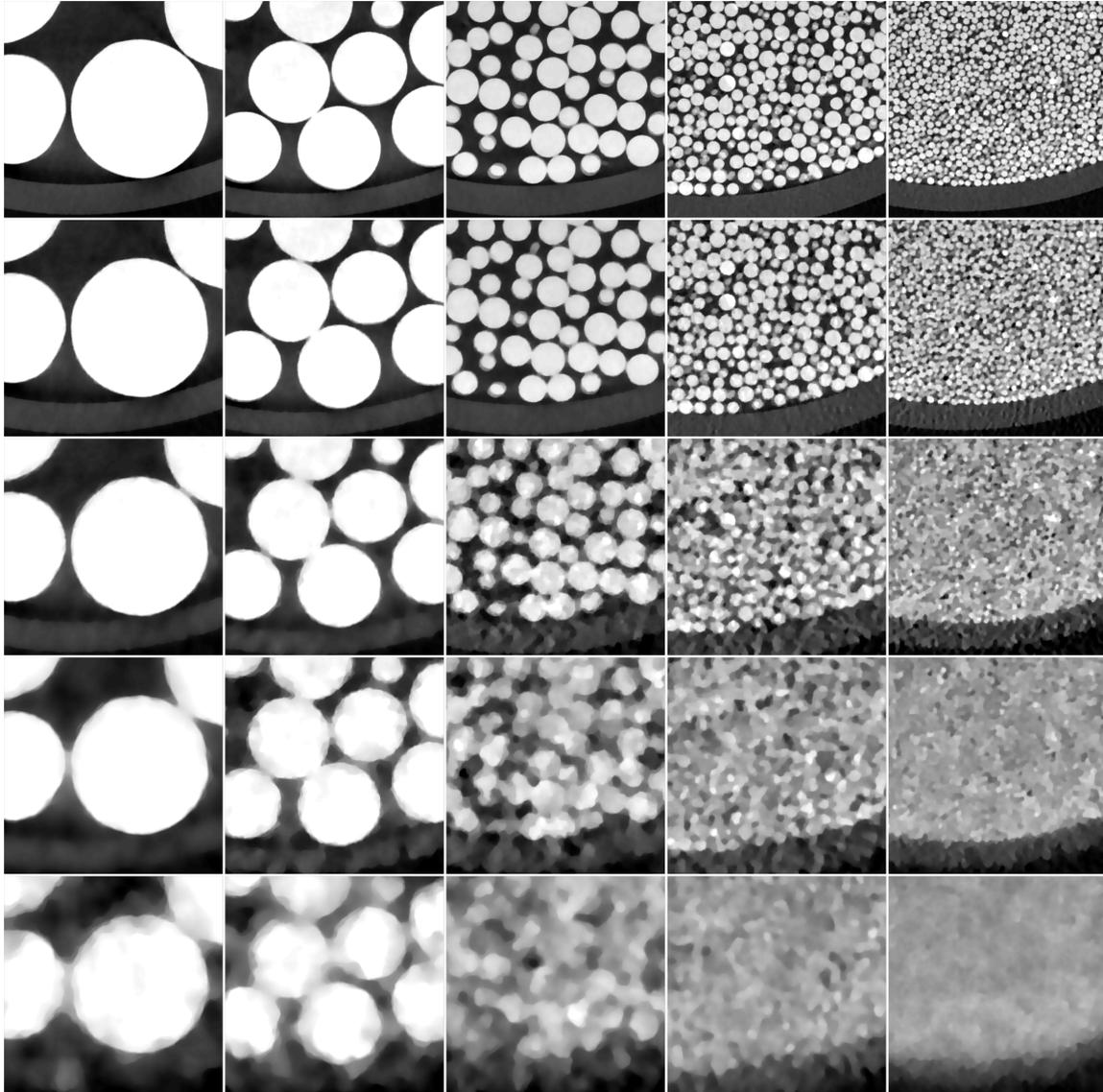


Figure 3: TV reconstructions. Left to right: B1L2 – B5L2. Top to bottom: 2520 (1), 504 (5), 126 (20), 63 (40), 30 (84) projections (downsampling factor, i.e. 2520 divided by number of projections).

sparse (have more zero pixels). Thus we expect larger beads (which have a more sparse gradient image) to allow accurate reconstruction at greater undersampling factors than smaller beads. Looking down columns of Fig. 3 this expectation appears to be confirmed visually: reconstructions of the largest bead B1L2 appear sharp and clean even at undersampling level 20–40. With decreasing bead size, we observe that smaller undersampling factors yield reasonable reconstructions, roughly along the SW-NE diagonal, with the smallest beads, B5 only admitting reconstruction at undersampling factors between 1 and 5.

4.2. Reconstruction error vs. number of projections and image sparsity

To substantiate the visually observed connection between bead size and admitted under-sampling level we compute the relative 2-norm errors at all 48 undersampling factors for all 5 bead sizes, see Fig. 4. To allow for a more detailed interpretation of the results we display the resulting error curves both in linear and double-logarithmic plots.

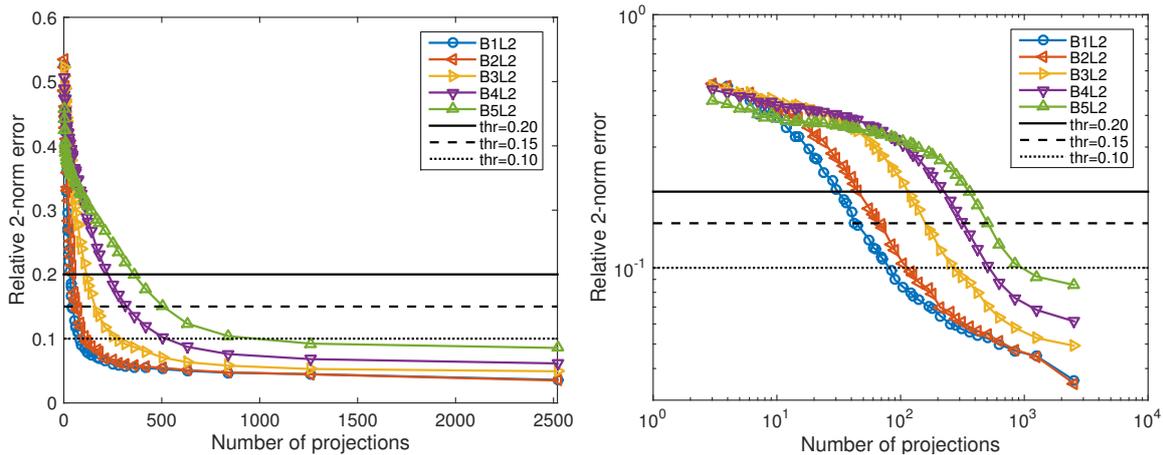


Figure 4: TV reconstruction errors vs numbers of projections for single images (layer L2) at the five bead sizes (B1–B5). Left: linear scale. Right: double-log scale.

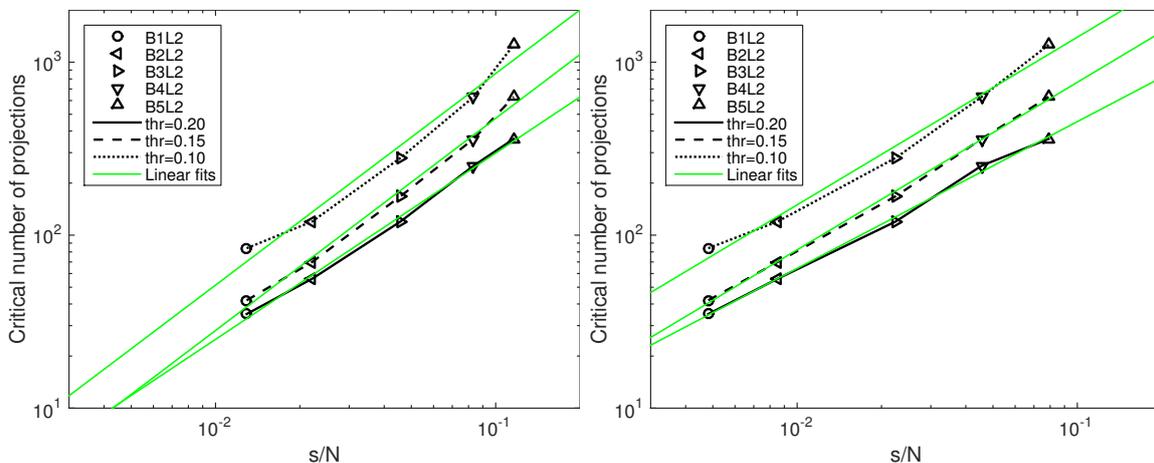


Figure 5: Relation between gradient sparsity level s/N , where s is the number of nonzero gradient values and N the total number of pixels, and the critical number of projections for a satisfactory TV-regularized reconstruction. Left: Sparsity levels determined using measure 1. Right: Using measure 2. The critical number of projections grows almost linearly with the sparsity level.

In the linear plot, different decay of reconstruction error as function of numbers of projections is observed among the bead sizes. In agreement with the visual inspection, the reconstruction error of the largest B1L2 decays rapidly to an almost constant level after a couple of hundred projections. B2L2 shows a similar decay but reconstruction errors are slightly larger. For B3L2, B4L2, B5L2 as beads become smaller, the error decays become more gradual and error values larger. This confirms the expectation that smaller beads, which are less sparse

in their gradient image, require more projections to achieve the same level of reconstruction accuracy than larger beads.

In the double-logarithmic plot a systematic decay trend is even more pronounced: All reconstruction error curves have a similar decay shape and the number of projections to reach the same reconstruction error increases systematically with decreasing bead size. Again, this supports our hypothesis of a simple link between sparsity level and critical sampling level.

To quantify the observed connection between sparsity and critical sampling level we will declare a reconstruction satisfactory if the relative error is below a given threshold. Clearly the choice of threshold is somewhat arbitrary and we thus compare three different threshold values ($E_2 = 0.20, 0.15$ and 0.10), as illustrated in Fig. 4. The smallest threshold is chosen at the point after which all error curves are settling at an almost constant level; the largest at the point where all curves have started their steep decay; and the third threshold values as the average of the two. We determine the critical number of projections for each bead size (and each threshold) as the smallest number of projection that produce a satisfactory reconstruction. The critical number of projections are plotted against the gradient sparsity levels in Fig. 5 (left and right: sparsity measures 1 and 2). The double-logarithmic plots reveal in all cases an approximate linear appearance, which is supported by best linear (in the double-logarithmic scale) models, corresponding to power functions $N_{\text{crit}} \approx b(s/N)^a$, where obtained a and b coefficients along with R^2 coefficients of determination are reported in Table 2. All R^2 values are close to 1, especially for thresholds 0.15 and 0.20, indicating that

Sparsity measure	Threshold	a	$b/10^3$	R^2
1	0.20	1.0771	3.5623	0.9968
	0.15	1.2266	7.9794	0.9938
	0.10	1.2231	14.3665	0.9758
2	0.20	0.8487	3.1985	0.9970
	0.15	0.9686	7.1175	0.9985
	0.10	0.9693	12.9925	0.9874

Table 2: Linear regression results on logarithm-transformed data, yielding relation $N_{\text{crit}} \approx b(s/N)^a$, which due to all a coefficients being close to 1 can be approximated by $N_{\text{crit}} \approx b(s/N)$.

the linear model (power function after exponentiation) is appropriate. With the a -coefficient (exponent) being close to 1 in all cases it is reasonable to approximate the relation by a direct linear relationship

$$N_{\text{crit}} \approx b \frac{s}{N}. \quad (2)$$

In other words, from the presented data we observe an approximate linear relation between the gradient image sparsity and the critical number of projections using TV-regularized reconstruction. This establishes a positive answer to our targeted question **Q1**.

Note from Table 2 that the linear coefficient b , which determines how quickly the critical number of projection grows with sparsity level, depends on the sparsity determination method. This is also observed in and can be explained by Fig. 5. The critical numbers of projections are the same in the left and right plots, only the sparsity values differ. Since sparsity method 1 yields more nonzeros than sparsity method 2, the curves are shifted to the right and are steeper, which corresponds to larger b -values.

4.3. Phase diagram analysis

Having established a positive answer to **Q1**, we now focus on **Q2**, which is to assess whether the found relation agrees with the empirical phase transition established in previous simulation work [13].

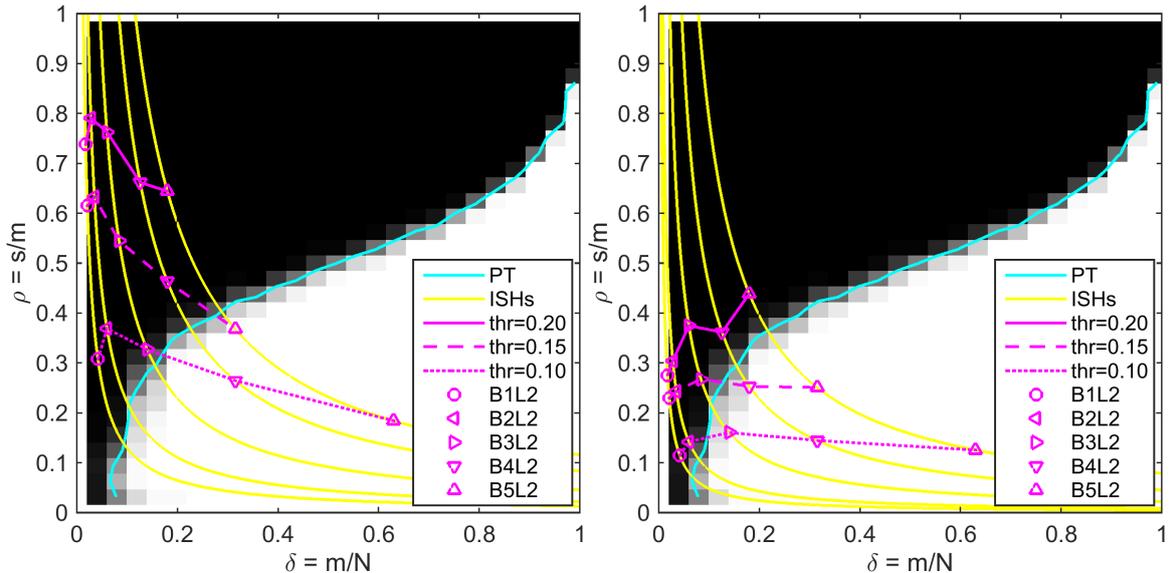


Figure 6: Empirical phase diagrams for sparsity measures 1 (left) and 2 (right). The empirical phase-transition (PT) curve sharply separates a full-recovery region (in white) from a no-recovery region (in black) and specifies the expected critical number of projections for a satisfactory reconstruction. Observed critical numbers of projections for each bead size are plotted as points based on the sparsity level and connected using lines with line styles indicating the threshold used. Auxiliary iso-sparsity hyperbolas (ISHs) shown trace out curves of constant sparsity levels of each bead size.

In [13] it was found empirically that a sharp phase transition as in CS could be observed for TV-regularized reconstruction of a class of images with sparse gradient magnitude. This phase transition occurs in the $(\delta, \rho) \in [0, 1]^2$ phase space, where $\delta = m/N$ is the number of measurements relative to the total number of pixels and $\rho = s/m$ is the sparsity level relative to the number of measurements. A combination of an image of a given sparsity and a given number of measurements acquired determine a position in the phase diagram. If below, the phase-transition curve essentially all such combinations were found to yield a satisfactory reconstruction using TV-regularization, while above implies failed reconstructions. The phase-transition curve specifies the expected critical sets of (δ, ρ) values separating the full-recovery and no-recovery regions of the (δ, ρ) phase space. The empirically determined phase-transition curve found in [13] is shown in Fig. 6.

If one knows the (gradient) sparsity s of an image with N pixels, then the phase transition curve specifies the critical number of measurements needed for TV-regularized reconstruction to produce a satisfactory reconstruction. By combining the expressions of δ and ρ we obtain

$$\rho = \left(\frac{s}{N}\right) \frac{1}{\delta},$$

from which it is seen that this image lies on a hyperbola in phase space, henceforth referred to as an iso-sparsity hyperbola (ISH). The point at which the ISH and the phase-transition curve

intersect is the critical number of measurements, m that suggests a satisfactory reconstruction at the given sparsity level. In [13] this approach predicted that 69.3 projections would suffice to recover a digitized phantom image of a walnut, which was almost perfectly correct, since the walnut was found empirically to be recovered at 68 projections.

In the present study instead of computing predicted critical numbers of projections we will compare the observed critical number of projections for all bead sizes directly with the phase-diagram predictions by plotting each one into the phase diagram. If the phase transition can predict the critical number of projections, each point should lie on or close to the phase-transition curve. This is done in Fig. 6 for both sparsity measures and all three satisfactory reconstruction thresholds. ISHs at the sparsity level of each bead size are added to visually connect each fixed sparsity and illustrate the transition across the phase diagram.

We observe that only few of the points are close to the phase-transition curves, and in particular for sparsity measure 1, most points are situated too high above the curve, which may again indicate that this method overestimates the true sparsity value. The points for sparsity measure 2 occur closer to the phase transition but do not appear to follow it at any of the chosen thresholds. From the presented data we must therefore conclude that the phase transition does not provide a useful method for predicting the critical number of projections from the (gradient) sparsity level. In other words, a negative answer to **Q2**.

Based on the almost-linear relationship between sparsity and critical sampling level (2), the lack of agreement with the phase transition should not be a surprise. We can convert (2) into the phase space variables to see what the relation corresponds to in the phase diagram. Given that each projection consists of $\sqrt{N} = 2000$ we find the critical number of measurements $m_{\text{crit}} = N_{\text{crit}}\sqrt{N}$ and the critical value of ρ to be

$$\rho_{\text{crit}} = \frac{s}{m_{\text{crit}}} = \frac{\sqrt{N}}{b}, \quad (3)$$

which is a constant independent of δ . For example for sparsity measure 2 and threshold value 0.15, this yield a constant of 0.2810. This matches quite well with the observed, almost-horizontal, dashed curve in Fig. 6, which lies at $\rho \approx 0.25$.

In other words the observed almost-linear relationship providing a positive answer to **Q1** is incompatible with the non-linear phase transition and thus results in a negative answer to **Q2**.

4.4. Bead quantification error measures

To assess whether the simple choice of the relative 2-norm error as the quality measure affects the obtained results we repeat the experiment using the quantification measures of aspect ratio, shape and volume of the selected beads. As the process of determining the measures is rather laborious we only provide values for 13 selected downsampling factors across the range of the 48 divisors of 2520. Fig. 8 shows the error measures as function of numbers of projections. While the behaviour is less defined than for the relative 2-norm error, the overall behaviour is very similar for the four largest bead sizes: from approximately constant errors the bead sizes curves decay rapidly at different numbers of projections in the same order as before until a quite steady level is reached. The errors determined for the smallest bead size B5 have been omitted as it was not possible to segment and separate enough beads to provide a meaningful error measure. A single threshold for a satisfactory reconstruction is chosen as 0.1 for all three quality measures and the critical number of projections are computed and

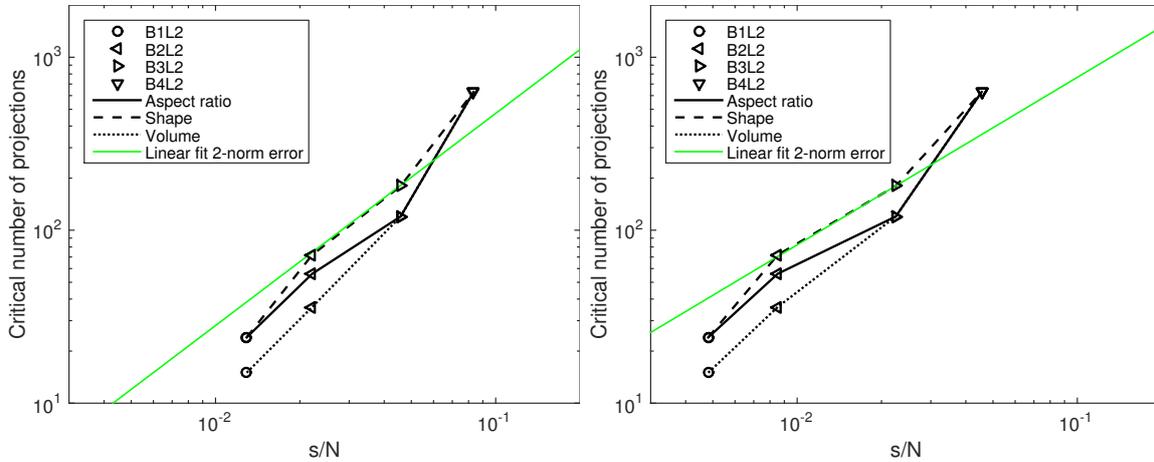


Figure 7: Critical number of projections for a satisfactory TV-regularized reconstruction as measured by the aspect ratio, shape and volume quantification errors. To help compare with the relative 2-norm error measure results in Fig. 5, the obtained linear models for threshold 0.15 are also included.

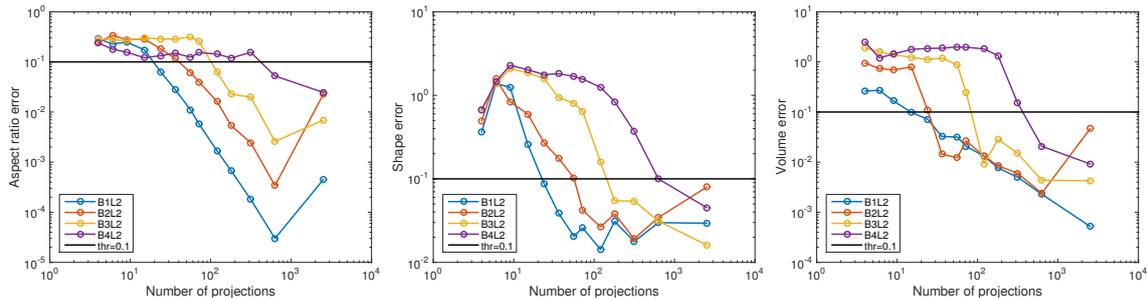


Figure 8: Aspect ratio, shape and volume quantification errors of TV-regularized reconstructions as function of numbers of projections. Compare Fig. 4.

plotted as function of the two sparsity levels in Fig. 7. In both cases a linear trend is again observed but no linear regression is performed. Instead to allow comparison with the relative 2-norm error results in Fig. 5 we display the previously obtained linear fits corresponding to the threshold of 0.15. In both cases the slope matches closely with the quantification results. We thus conclude that the quantification measures support the same positive answer to **Q1** as we obtained for the relative 2-norm error measure. This strengthens the conclusion as the quantification measures may be more representative of reconstruction quality, as they more physically reflect geometric properties of the glass beads.

For completeness we also compare the quantification measure results with the phase diagram in Fig. 9. Similar to the relative 2-norm results, the observed critical number of projections do not consistently occur close to the phase-transition curve. We conclude again that the results of the conducted study are not consistent with a positive answer to **Q2** which questioned whether there is a connection between the critical sampling levels and the sparsity level as given by the phase transition.

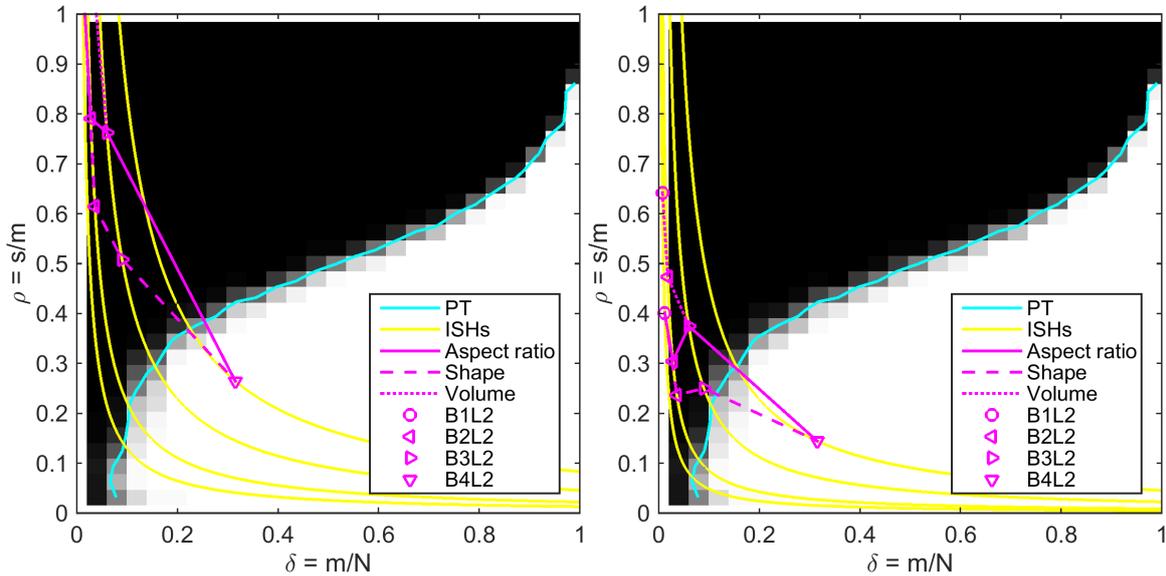


Figure 9: Comparison of observed critical number of projections by aspect ratio, shape and volume quantification errors with the phase diagram prediction of critical sampling level.

5. Discussion

Upon close inspection Fig. 3 shows some artefacts, mostly clearly seen for B3L2 and B4L2 and take the form of mainly smaller beads having a double appearance with two almost identical copies slightly offset. The same artefacts are present in the CGLS reconstructions in Fig. 1, which means that it is not an artefact of TV regularization but rather inherent to the data. We believe this artefact arises from bead movement during the scan, as beads were not fixed and had some freedom to move, despite best efforts to firmly pack the bead containers. We used a stop-start scan procedure to obtain projections at fixed angles rather than over a small angular range as is the case in a continuous scan mode. We believe the constant stopping and starting resulted in a periodic jerking of the container which may have caused bead movement. The movement artefacts may be removed by filling voids for example by epoxy to fix beads, however this would require ensuring that no air bubbles are trapped inside to preserve the simple object composition of just beads and background. Another option may be to subject the packed beads to sintering, which would attach all beads thus fix into a rigid structure, while preserving the spherical shape of beads apart from at interfaces and avoid issues of trapped air. A third option to avoid movement artefacts may be to use 3D-printing to produce a test phantom of varying sparsity. One may also argue that in any real data acquisition some inconsistencies will be present and the particular movement artefacts encountered in the present data set is just one realization of inconsistencies and therefore something that is simply part of carrying out a real-data study. The micro-CT scanner used to acquire the present data had been conditioned daily following local laboratory procedures to ensure the instrument was ready, notably that the X-ray source was stable during acquisition. No further calibration or optimization such as source focusing immediately prior to data collection was done; doing so may further improve the quality of data to help strengthen the achievable conclusions.

Other possible sources of errors compared to a simpler ideal synthetic data scenario include

computational aspects in the numerical solution of the TV-regularized reconstruction problem for large-scale data. We used the ASTRA Tomography Toolbox [25] for GPU-acceleration of the forward and back-projection operations, which are the most expensive steps involved. For efficiency reasons this uses only single precision and involves an unmatched forward-backprojector pair implementation, i.e., the back-projection operation is not equivalent to applying the conjugate of the forward projection operation represented by the matrix A . This may lead to inaccuracies and convergence issues [26], but the effect may be kept at bay by the additional regularity enforced by TV regularization. Furthermore a larger number of iterations may yield slightly different reconstructions, but as we observed essentially no change in the image after 10,000 iterations in a range of inspected cases, we believe the reported observations will remain unchanged. Finally, any minor effect of using the Huber-smoothed TV may be removed by using an algorithm for non-smooth optimization such as the Chambolle-Pock algorithm, but as explained in Sec. 3.4 we believe to have already reduced the smoothing effect to be negligible. In the present work we chose to use the TVReg algorithm because we have found it to work very robustly when applied to real data and interfacing to the ASTRA Tomography Toolbox across a range of parameter regimes. This is in our experience not always the case with implementations of advanced reconstruction methods.

6. Conclusion

The presented data suggest that in the range of sparsity levels considered the critical number of projections for satisfactory reconstruction by TV-regularization grows approximately linearly with the image gradient sparsity. To our knowledge this is first empirical evidence and quantification of a link between gradient image sparsity and the number of projections sufficient for accurate reconstruction by TV-regularization. Such a connection was demonstrated in previous simulation work by one of the authors and might have been anticipated from compressive sensing, but prior to the present work it was an unaddressed question whether this was also the case in practice.

We believe that under similar conditions the empirically established relation may be used as a guideline for determining the minimum number of projections to acquire for a satisfactory TV-regularized reconstruction based on the expected sparsity of the sample. For relatively sparse samples this may lead to a large reduction in scanning efforts (time, X-ray exposure, etc.) of an order of magnitude or more, as opposed to standard practice of dense angular sampling employed at present. An important question for future work is how general the observed relation is when changing the scan conditions and the type of sample scanned.

Other future work include to address limitations described in the previous section such as bead movement to design and acquire a more optimized data set. This may further corroborate the link between sparsity and critical sampling level and possibly reveal a closer relation with the phase-transition behaviour observed for simulated data.

References

- [1] X. Pan, E. Y. Sidky, M. Vannier, Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?, *Inverse Probl* 25 (2009) 123009.
- [2] E. Y. Sidky, C.-M. Kao, X. Pan, Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT, *J Xray Sci Technol* 14 (2006) 119–139.

- [3] E. Y. Sidky, X. Pan, Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization, *Phys Med Biol* 53 (2008) 4777–4807.
- [4] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, X. Pan, Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT, *Phys Med Biol* 55 (2010) 6575–6599.
- [5] G.-H. Chen, J. Tang, S. Leng, Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets, *Med. Phys.* 35 (2008) 660–663.
- [6] J. Song, Q. H. Liu, G. A. Johnson, C. T. Badea, Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT, *Med. Phys.* 34 (2007) 4476–4483.
- [7] L. Ritschl, F. Bergner, C. Fleischmann, M. Kachelrieß, Improved total variation-based CT image reconstruction applied to clinical data, *Phys Med Biol* 56 (2011) 1545–1561.
- [8] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1992) 259–268.
- [9] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans Inf Theory* 52 (2006) 489–509.
- [10] D. L. Donoho, Compressed sensing, *IEEE Trans Inf Theory* 52 (2006) 1289–1306.
- [11] S. Petra, C. Schnörr, Average case recovery analysis of tomographic compressive sensing, *Linear Algebra Appl* 441 (2014) 168–198.
- [12] S. Petra, C. Schnörr, TomoPIV meets compressed sensing, *Pure Math. Appl.* 20 (2009) 49–76.
- [13] J. S. Jørgensen, E. Y. Sidky, How little data is enough? Phase-diagram analysis of sparsity-regularized X-ray computed tomography, *Philos Trans R Soc Lond Ser A* 373 (2015) 20140387.
- [14] J. S. Jørgensen, C. Kruschel, D. A. Lorenz, Testable uniqueness conditions for empirical assessment of undersampling levels in total variation-regularized X-ray CT, *Inverse Probl. Sci. Eng.* 23 (2015) 1283–1305.
- [15] J. S. Jørgensen, E. Y. Sidky, P. C. Hansen, X. Pan, Empirical average-case relation between undersampling and sparsity in X-ray CT, *Inverse Problems and Imaging* 9 (2015) 431–446.
- [16] J. S. Jørgensen, E. Y. Sidky, X. Pan, Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in X-ray CT, *IEEE Trans Med Imaging* 32 (2013) 460–473.
- [17] J. S. Jørgensen, S. B. Coban, W. R. B. Lionheart, P. J. Withers, Effect of sparsity and exposure on total variation regularized X-ray tomography from few projections, in: *Proc. 8th Intl. Mtg. on inverse problems*, pp. 279–282.

- [18] S. B. Coban, S. A. McDonald, SophiaBeads Dataset Project, Available online from: <http://dx.doi.org/10.5281/zenodo.16474>, 2015.
- [19] S. B. Coban, P. J. Withers, W. R. B. Lionheart, S. A. McDonald, When do iterative reconstruction methods become worth the effort?, in: Proc. 13th Intl. Mtg. of Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine.
- [20] H. H. Barrett, K. J. Myers, Foundations of Image Science, John Wiley & Sons, Hoboken, NJ, 2004.
- [21] S. B. Coban, W. R. B. Lionheart, P. J. Withers, Assessing the suitability of computed tomography reconstruction methods through physical quantification techniques, (in preparation) (2016).
- [22] H. Wadell, Volume, shape, and roundness of quartz particles, *The Journal of Geology* 43 (1935) 250–280.
- [23] X. Han, J. Bian, E. L. Ritman, E. Y. Sidky, X. Pan, Optimization-based reconstruction of sparse images from few-view projections, *Phys Med Biol* 57 (2012) 5245–5273.
- [24] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, S. H. Jensen, Implementation of an optimal first-order method for strongly convex total variation regularization, *BIT* 52 (2012) 329–356.
- [25] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, J. Sijbers, The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography., *Ultramicroscopy* 157 (2015) 35–47.
- [26] G. L. Zeng, G. T. Gullberg, Unmatched projector/backprojector pairs in an iterative reconstruction algorithm, *IEEE Trans Med Imaging* 19 (2000) 548–555.

Acknowledgements

The authors are grateful to Dr. Julia Behnsen, Manchester X-ray Imaging Facility for assistance in data collection. The work by JSJ was supported by the project “High-Definition Tomography” funded by Advanced Grant No. 291405 from the European Research Council. JSJ also acknowledges the support and hospitality from School of Mathematics and School of Materials at the University of Manchester for hosting him during the time the presented work was carried out. The work by SBC was supported by the School of Mathematics (University of Manchester), EPSRC CCPi (EP/J010456/1), and BP through the BP International Centre for Advanced Materials (BP-ICAM). We acknowledge funding from EPSRC for the X-ray imaging facility under EP/F007906/1 and EP/F028431/1.

Data availability

Upon acceptance of the paper the supporting data will be made available from Zenodo.

Chapter 4

Non-linearity in monochromatic transmission tomography

Non-linearity in monochromatic transmission tomography

William R.B. Lionheart¹, Bjørn Tore Hjertaker², Rachid Maad², Ilker Meric³,
Sophia B. Coban¹, and Geir Anton Johansen⁴

¹School of Mathematics, University of Manchester

²Department of Physics and Technology, University of Bergen

³Department of Electrical Engineering, Bergen University College

⁴Faculty of Engineering and Business Administration, Bergen University College

Abstract

While it is well known that X-ray tomography using a polychromatic source is non-linear, as the linear attenuation coefficient depends on the wavelength of the X-rays, tomography using near monochromatic sources are usually assumed to be a linear inverse problem. When sources and detectors are not treated as points the measurements are the integrals of the exponentials of line integrals and hence non-linear. In this paper we show that this non-linearity can be observed in realistic situations using both experimental measurements in a γ -ray tomography system and simulations. We exhibit the Jacobian matrix of the non-linear forward problem and show that it can be solved iteratively. Applying this algorithm to experimental data we show that improved reconstructions can be obtained.

Keywords: Exponential edge effect, non-linear partial volume effect, penumbra, gamma ray tomography, x-ray tomography, non-linear reconstruction

1 Introduction

X-ray or γ -ray transmission tomography (CT) is generally taken to be a linear method in that the vector of the logarithm of the measurements on each detector, for each source, is a linear function of the linear attenuation in the region being imaged. Under this assumption the problem of recovering the linear attenuation in each pixel (or voxel) in the image reduces to one of linear algebra. The matrix of the resulting linear system will be mildly ill-conditioned (at least for fairly complete data) and standard regularized inversion methods or discrete approximations to explicit reconstruction algorithms for the continuum case work reasonably well [14]. There are several reasons for a linear model not to fit CT data. In the polychromatic case the dependence of the linear attenuation on wavelength is often the largest effect and is important in medical, dental, non-destructive testing and security screening applications. Scattering also results in non-linearity [6], as well as a forward problem that more highly coupled as material out of the direct beam path between a source and detector can affect the measurement. This has been identified as the *penumbra* effect [12, 11]. This can be counteracted by collimation of the detectors using a ‘scatter grid’ but in some systems such as the fast switched sourced RTT security screening CT system this is not possible [19, 20]. However even nearly monochromatic CT systems using for example a γ -ray or synchrotron source, and even without scatter, can exhibit non-linearity.

For simplicity we will consider a two dimensional problem (generalization to three dimensions is straightforward). Consider a fixed source and detector pair both represented by a line segment. Let $S(s)$ be a point on the source point parametrized by arc length s and $D(s)$ similarly a point on the detector. From Beer-Lambert’s law [1], flux density I measured at a point in the detector

due to a point in the source with flux density I_0 is

$$I(D(s)) = I_0(S(s')) \exp \left(- \int_{L(S(s'), D(s))} \mu \, d\ell \right), \quad (1)$$

where the integral is taken over the line $L(S(s'), D(s))$ between source and detector points, and $d\ell$ is the measure on that line. The total flux for this source-detector pair is thus

$$I = \iint I_0(S(s')) \exp \left(- \int_{L(S(s'), D(s))} \mu \, d\ell \right) \, ds' \, ds, \quad (2)$$

and we see the problem: the sum of exponentials is not the exponential of the sum. This non-linearity is clearly bigger for larger sources and detectors giving a wider range of ray paths, and also increases as the variation in μ is such that the ray paths in a source and detector pair encounter a wider range of linear attenuation. The non-linearity has been described as a “non-linear partial volume effect” [6], the volume here referring to the region defined by all the rays between a source and detector pair and the fact that part of that volume has a different attenuation from another part and the exponentials of the integrals are combined. It has also been described in the literature as “exponential edge-gradient effect” [10, 5]. For essentially two dimensional problems the phenomena has been noticed in the out of plane direction when the attenuation varies on the length scale of the slice width. According to one review paper “Hopefully, this effect on the variance is small because accounting for it seems to be challenging” [16].

When using large sources and detectors it has been noticed that the non uniform density of rays across the volume needs to be accounted for [21, 12] and that this can improve reconstruction under some circumstances even without taking account of the non-linearity.

2 Mathematical formulation

The starting point of this work is to consider a finite set of positions of sources and detectors of finite size. Although we can consider the continuous case of the Radon or X-ray transform where there are infinitely many point sources and detectors, there is no obvious candidate for a limiting continuum case with infinitely many measurements with finite sources and detectors. Our investigation will therefore concern finitely many measurements. We can only recover a finite number of parameters in the the image space so we suppose for simplicity that the linear attenuation coefficient can be represented as

$$\mu(x) = \sum_{k=1}^{K^2} \mu_k \chi_k(x), \quad (3)$$

where χ_k are the characteristic functions of some partition of the region of interest in to K^2 picture elements, such as square pixels. We will denote by \mathbf{m} the vector of non-negative coefficients μ_i . We will assume that measurements are taken on K^2 pairs of line segments $S_k D_k$ thought of as a source, detector pair.

Our data vector \mathbf{d} is given by the non-linear forward problem

$$\mathbf{d} = \mathbf{F}(\mathbf{m}), \quad (4)$$

where

$$F_k(\mathbf{m}) = \log \left(\iint I_0(S_k(s')) \exp \left(- \int_{L(S_k(s'), D_k(s))} \mu \, d\ell \right) \, ds' \, ds \right). \quad (5)$$

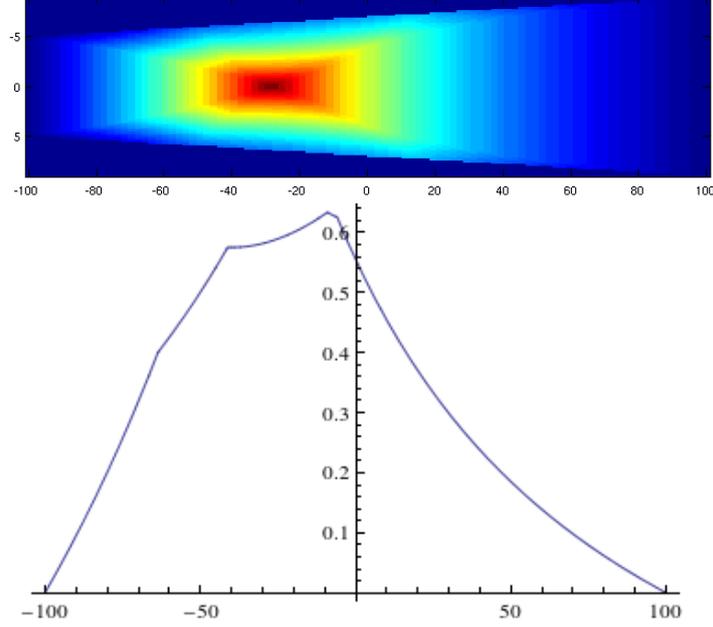


Figure 1: The normalized linearized sensitivity for a simple case where source and detector are parallel and opposite. In this case the dimensions in mm are $W_1 = 18, W_2 = 10, L = 200$. The lower plot shows a cross section for $y = 3$ illustrating that the function is continuous but not differentiable.

The Jacobian matrix, $\partial F_k / \partial \mu_l$ is

$$-\frac{1}{\exp F_k(\mathbf{m})} \iint I_0(S_k(s')) \exp\left(-\int_{L(S_k(s'), D_k(s))} \mu dl\right) \left(\int_{L(S_k(s'), D_k(s))} \chi_l\right) ds' ds. \quad (6)$$

In the simple case $I_0 = 1$, the derivative at $\mathbf{m} = 0$ is simply

$$\frac{\partial F_k(\mathbf{0})}{\partial \mu_l} = -\frac{1}{\exp F_k(\mathbf{0})} \iint \int_{L(S_k(s'), D_k(s))} \chi_l dl ds' ds, \quad (7)$$

which coincides with the usual linearized problem for CT taking account of non-zero source and detector size [21, 12].

In the limiting case, as the pixels tend to points, the integral reduces to the *point sensitivity*, which is the density of lines at a point in the image. The point sensitivity for a simple case where source and detector are parallel and opposite with source width W_1 , detector width W_2 separated by L and $I_0 = 1$ is

$$S(x, y) = \frac{1}{2W_1W_2} r_{1,2} r_{2,1}, \quad (8)$$

where

$$r_{i,j} = \left(-W_i/2, W_i/2, y - \frac{d_j(L+2x)}{2(L-2x)}, y + \frac{d_j(L+2x)}{2(L-2x)}\right)$$

and $r(a, b, c, d)$ is the length of the intersection of the intervals (a, b) and (c, d) (see Appendix A). This rather complicated formula that simply expresses the product of the lengths of the line

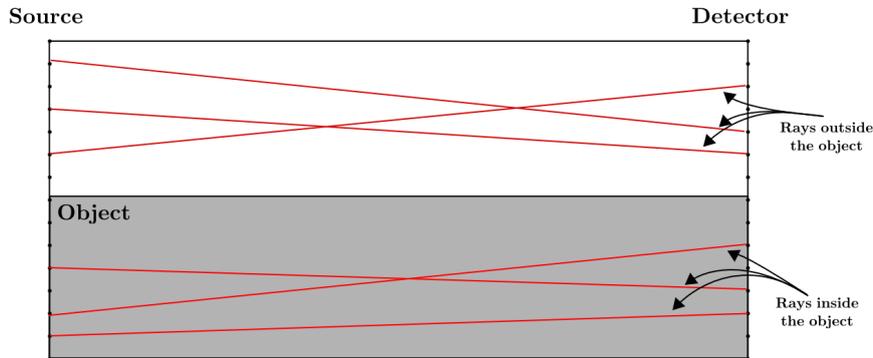


Figure 2: A simple diagram illustrating NLPVE.

segments that are subsets of the source and detector that can be connected by straight lines to the point (x, y) in question. While this is continuous the function is only piece-wise differentiable, and as can be seen from the example in fig.1 markedly inhomogeneous. Perhaps the most notable feature is that a ‘small’ object is not detected close to the source or detector or at the edges of the ‘beam’ joining the source and detector. Even for systems where the non-linearity is not especially pronounced this non-uniform sensitivity might still be sufficient to have a substantial effect on the accuracy of reconstructed images.

The non-linear effect can be illustrated nicely with a source and detector parallel and the same width. First consider an object which occupies half width of the rectangle between the source and detector. As a very crude approximation to illustrate the effect, consider only rays entirely inside the object (linear attenuation μ) or entirely missing the object (linear attenuation 0) and assume that the distance L between source and detector is sufficiently larger than the width W that we can ignore the variation in the path lengths. We have approximately the measurement as a function of μ

$$F(\mu) = \log \frac{1}{2} W (1 + \exp(-\mu L)).$$

In this case we have, by taking series expansions,

$$F(\mu) - F(0) = -\frac{\mu L}{2} + \frac{\mu^2 L^2}{8} + O(\mu L)^3,$$

which gives us a ‘rule of thumb’ that the magnitude of the non-linear partial volume effect in this case is around half the square of the linear approximation to the difference in the logarithmic data.

A simple approximation to the integral over lines joining source and detector can be achieved by choosing uniformly spaced quadrature points on each source and detector, and summing the intersection length with each pixel over all the lines joining quadrature points on each source and detector pair. Suppose there are N pairs of quadrature points for each source-detector pair. Let R be the $NK \times K$ matrix of line intersection lengths of pixels of lines joining quadrature points. Let P be the $K \times NK$ matrix that performs the (weighted) sum over the all the lines in a source detector pair. In the linear approximation to CT the matrix PR would be the matrix of the system we seek to solve. Instead, using this approximation, we solve

$$\mathbf{d} = \tilde{F}(\mathbf{m}) := \log P \exp(R\mathbf{m}), \quad (9)$$

where the log and exp applied to vectors are taken to act component-wise. The Fréchet derivative is

$$D\tilde{F}(\mathbf{m})\delta\mathbf{m} = (\text{diag} P \exp(R\mathbf{m}))^{-1} P \text{diag} (\exp(R\mathbf{m})) R \delta\mathbf{m}. \quad (10)$$

3 Reconstruction algorithms

Let p be the number of projections, r be the number of rays per projection, $M = pr$, and K be the length (and width) of the image domain in pixels. For both linear and non-linear reconstructions, the image \mathbf{m} is reconstructed over a square grid of $K \times K$. To approximate the integrals over the source and detector we choose N_S and N_D quadrature points on the source and detector respectively and $N = N_D N_S$. We denote by R the geometry matrix of size $NM \times K^2$. Each row of R is obtained using Jacob's ray tracing algorithm [18] to find the intersection length of a point between a quadrature point on the source and detector. The matrix P is the $M \times NM$ matrix that simply takes a weighted sum of rows corresponding to each ray within a given source-detector pair. We denote by \mathbf{d} the (vectorized) measured data with size $M \times 1$. We found by numerical experiment that for our geometry using numbers of quadrature points larger than $N_S = N_D = 5$ did not change the results to three decimal places.

For the linear model, we need to solve $PR\mathbf{m} = \mathbf{d}$, although it will be overdetermined if we choose $K^2 < M$ so we require a least squares solution

$$\arg \min_m \|PR\mathbf{m} - \mathbf{d}\|_2^2$$

using a straightforward implementation of conjugate gradient method adapted to least squares problems (henceforth CGLS) [17]. The CGLS method is essentially the conjugate gradient method applied to solve the normal equation $A^T A x = A^T b$. CGLS is a popular method amongst those working in signal and image processing for its simple and computationally inexpensive implementation and fast convergence. For our reconstructions, the CGLS iterations were performed until a tolerance value, ϵ , was reached or a maximum number of iterations, i_{\max} , was performed. The tolerance value is defined as the 2-norm of the residual vector calculated at each iteration. The algorithm was also applied to the generalized Tikhonov regularization problem [8]

$$\arg \min_m \|PR\mathbf{m} - \mathbf{d}\|_2^2 + \alpha^2 \|L\mathbf{m}\|_2^2,$$

where L is the regularization matrix and $0 < \alpha$ is the regularization factor. For our runs for both numerically simulated and real γ -ray data reconstructions, the regularization matrix L was chosen to be the identity matrix, and $\alpha = 0.01$. Another obvious choice of L would be a finite difference gradient which would penalize less smooth images. In that case a larger K could be used as the regularized system, formed by appending L beneath PR , would still be over determined.

In addition we include total variation regularized (TV) reconstructions to compare with the CGLS and non-linear optimization methods. The method employed here is based on the primal-dual algorithm outlined by Chambolle and Pock in [3], which solves the following optimization problem

$$\arg \min_m \frac{1}{2} \|PR\mathbf{m} - \mathbf{b}\|_2^2 + \alpha \|D\mathbf{m}\|_{1,2}.$$

Here, the mixed $l_1 - l_2$ norm $\|\cdot\|_{1,2}$ is defined as

$$\|D\mathbf{m}\|_{1,2} = \sum_{i=1}^{K^2} \|D_i\mathbf{m}\|_2,$$

where $D_i\mathbf{m}$ is the forward-difference approximation of the gradient at voxel i . The implementation of the primal-dual algorithm also requires four additional parameters that define the primal and dual step-sizes, $\rho, \tau > 0$, and upper and lower bounds for the image domain. The lower bound is known in our case since we cannot have a negative attenuation coefficient, meaning we have three parameters and the regularization factor α to fine-tune for best reconstructions. Despite the required number of parameters and longer runtime due to increased demand for computational power, TV is expected to give more accurate images where objects have distinct boundaries and homogeneous interiors. For this reason TV regularization is often used as a deblurring or denoising method in image processing.

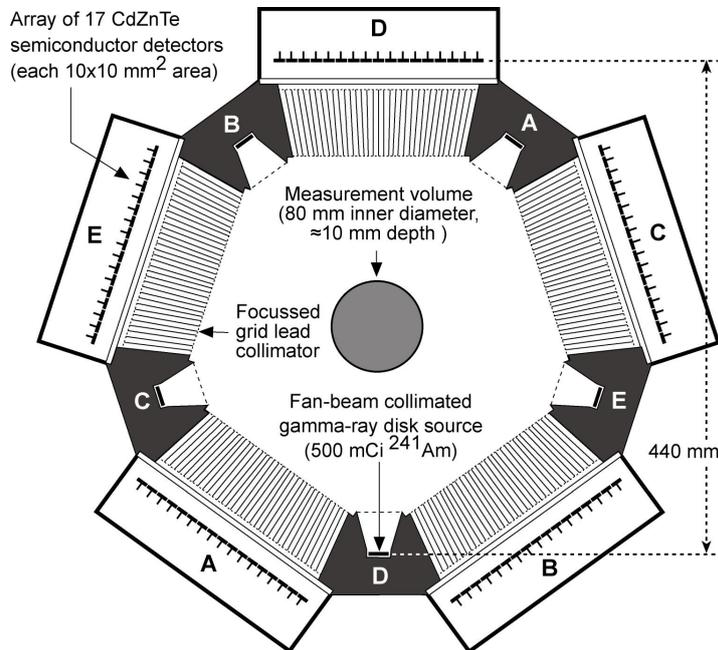


Figure 3: Cross section of Bergen γ -tomography system. The lead collimator plates at the midline of the detector elements are shown but they were removed in the detector array used for our experiments leaving only the plates between each element.

The non-linear reconstructions are obtained using the trust region method [2] with the reflective transform [4] applied at each iteration. The trust region methods minimize a quadratic model function, $Q(p)$ given by

$$\min_p Q(p) = f_k + J_k^T p + \frac{1}{2} p^T H_k p, \quad \text{s.t. } \|p\| \leq \Delta_k,$$

where $f_k = f(p_k)$ is the objective function, $J_k = \nabla f(p_k)$ is the Jacobian, $H_k = \nabla^2 f(p_k)$ is the Hessian matrix, and Δ_k is the radius of the trust region, over which the function $Q(p)$ is minimized. In our implementation, we use $H_k = J_k^T J_k$ (trust region Newton's method). Δ_k is calculated at each iteration, which is based on the agreement between the quadratic function Q and the objective function f at the previous iteration. See cited literature and [15] for implementation details. We chose to implement a trust region method as these methods are more robust than line search methods for solving ill-conditioned systems.

4 Description of apparatus

The Bergen γ -ray tomography system described in [9] was used for this experimental study, see fig.3. The system was designed for fast imaging of oil, gas and water cross sectional distributions within a pipe. The system is made up of a total of 5 ^{241}Am radioisotopes with an activity of 500 mCi and 5 detector modules each containing a total of 17 CdZnTe detectors. The relatively large source (18mm diameter) and detector size (10mm across) separated by 440mm was large enough to expect an observable non-linearity. The ^{241}Am source with a 59.5 keV principal emission peak is a good approximation to monochromatic radiation.

The CdZnTe detector elements are $10 \times 10 \text{ mm}^2$ in cross section and 2mm thick, chosen to effectively stop any low-energy γ -ray from being transmitted through an element. Lead plates are fitted between each element to reduce the detection of scattered photons: a 'scatter grid'. In its

normal operation there are similar lead plates bisecting each element but these were removed as we wanted to measure the effect of all the ray paths between source and detector element. These collimation plates would be expected to reduce the non-linearity at the expense of also reducing the counts measured. Each detector array is positioned opposite a source. The source aperture is designed to collimate the radiation in a fan beam that coincides with the 80mm diameter circular aperture where the phantoms are placed. This aperture is itself surrounded by a tube of Perspex 5mm thick. The system has five nominally identical source and detector array pairs distributed equiangularly. The counting threshold of each detector is set to 48 keV, effectively screening out the lower emission peaks of the source as well as minimizing the effects of scattered radiation, e.g. those photons that undergo Compton scattering and lose a certain fraction of their initial energy. Before inserting a phantom a background reading was obtained, averaging the counts on the detector array in use until the ratios of the number of counts in each element to the total counts was constant to three decimal places. Each measurement of the phantom used the total number of counts over the same time. In practice we averaged over 30s for each measurement.

5 Monte Carlo simulation

It is a well-known fact that scattered radiation affects the accuracy of reconstructed images. This is also the case in the Bergen γ -ray tomography system with fixed source and detector modules as illustrated in fig.3. The question that arises then is the following: What is the contribution from scattered radiation to the overall non-linearity as compared to radiation directly transmitted through the phantom and onto the detector module? This question cannot be answered easily by experiments as it is not an easy task to separate scattered (i.e. collided) and uncollided components. Monte Carlo (MC) simulations, on the other hand, provide a means to numerically track individual photons in the geometry of the γ -ray tomography system, and provide the possibility of identifying these as collided and uncollided photons. Thus, in order to study the relative importance of the non-linear effects as compared to scattered radiation, a simple Monte Carlo (MC) model of the Bergen γ -ray tomography system has been implemented using the general purpose MC code, MCNP6.1 [7].

Here, the first task was to benchmark the MC modeling against experimental data. Since the focus was on determining the overall non-linear effects in the pertinent γ -ray tomography system, a detailed implementation of entire system geometry was deemed unnecessary. This is because the non-linear effects stem from the partial coverage of an object spanned by the rays between a source and the corresponding detector. A portion of these rays may undergo different amounts of attenuation. Thus, the non-linear effects are not related to the finer details of the measurement geometry with the exception of collimation grid. Therefore, only a single detector module, containing a total of 17 CdZnTe semiconductor detectors, was implemented. In addition, the lead collimator blades placed in front of each detector module were implemented. In order to replicate the experimental conditions, every second collimator, the ones in the middle of each element, was removed from the geometry. Thus, the final geometry implemented in MCNP6.1 contains 17 lead collimator blades. As the main purpose of the MC simulations has been the separation of collided and uncollided fluxes incident upon the detectors, no other details such as the detector and source housing were implemented. In accordance with the geometry of the experimental setup the source-to-detector distance was set to 440 mm. The source was defined as a disk source with a diameter of 18 mm emitting 59.5 keV photons whereas the pipe inner and outer diameters were set to 80 mm and 90mm, respectively. In table 1, a list of materials used in the MC simulations is provided.

At the energy of interest, the CdZnTe detectors exhibit nearly 100.0% detection efficiency. Thus, secondary electron transport could be safely neglected. The secondary electron transport was turned off in the entire problem geometry and the simulations were run as a “photon-only” problem using detailed photon transport physics. Furthermore, the transport cut-off energy for photons was set to 48 keV in the entire geometry. This is in line with the experiments as the counting threshold of all detectors is set to 48 keV. For all 17 detectors in the geometry, so-called

Table 1: Listing of materials used in the MC simulations.

Component	Elements	Weight fractions	Density [g/cm ³]
Pipewall (PMMA)	H, C, O	0.08, 0.6, 0.32	1.19
Phantom(polypropylene)	H, C	0.14, 0.86	0.92
Air	N, O, Ar, C	0.755, 0.232, 0.012, 0.001	0.0012
Detectors(CdZnTe)	Cd, Zn, Te	0.4, 0.1, 0.5	5.78
Collimators	Pb	1.0	11.35

$f4$ tallies giving the flux averaged over a cell were defined. For the purpose of benchmarking the MC model, simulations were run with the pipe filled with air and polypropylene (see table 1) representing empty and full pipe measurements, respectively.

Moreover, the above mentioned $f4$ tallies were divided into collided and uncollided contributions using the so-called uncollided secondaries card in conjunction with special tally treatment [7]. The purpose was to identify the number of collisions source particles suffer prior to reaching the cells defining the detectors and also, to identify secondary particles such as characteristic X-rays that may reach the detectors without suffering collisions. The uncollided secondaries card thus was used to label these secondaries as collided particles. In this way, it was possible to separate the uncollided source particles from the collided in addition to any secondaries that may reach the detectors. Everything other than uncollided source particles were treated as collided.

Using the above mentioned settings, the effects of scattered radiation could be studied and compared to the case of “no scatter”. The results obtained are given in Sec.7.1.

6 Phantoms and experimental protocol

The biggest observable non-linear effect is expected when the rays between a source-detector pair encounter the largest contrast in the line integral of the linear attenuation. One simple way in which this happens is a plane interface between air and a more attenuating substance with the plane aligned with the source detector pair but only partly in the path. Such a phantom was already available and had been used with the Bergen system. It consist of half of a solid, 80mm diameter cylinder of polypropylene truncated by an axial plane through its mid-line. This was designed to emulate a pipe half full of oil and locating such an interface accurately within the intended use of the system. In each of the five detector arrays the 9th of the 17 detector elements diametrically opposite the centre of a source and we thus expect this to show the largest non-linearity. For comparison a full solid cylinder of the same polypropylene was used. The protocol used was to collect a set of data with the half-cylinder phantom aligned with one source and the opposite detector element 9. The phantom was then turned through half a turn and the measurement repeated. Finally the same measurement was repeated with the solid cylinder phantom. To avoid statistical fluctuations due to the arrival time statistics of the photons data was collected over a 30s period in three blocks of 10s.

Denote the array of normalized logarithmic counts for a each detector element and a given source for the half-cylinder in the two positions by L_1 and L_2 and the full cylinder by L_{12} . The normalization is taken with all phantoms removed and compensates for variations in the sensitivity of the detector elements. If the logarithmic counts were linear in the linear attenuation we would satisfy the superposition principle $L_1 + L_2 = L_{12}$. We thus consider $L_1 + L_2 - L_{12}$ as a measure of the non-linearity.

The second phantom used consists of an 80mm diameter polypropylene cylinder with cylindrical holes of diameter 26mm bored so that the axis of the cylinder and the hole are parallel and separated by 20mm. See fig.4b. Two solid plugs machined to be a tight fit in the hole were made from the same batch of polypropylene. A solid cylinder would have served the same purpose as the two hole cylinder with both plugs fitted, but experiments showed a difference between the

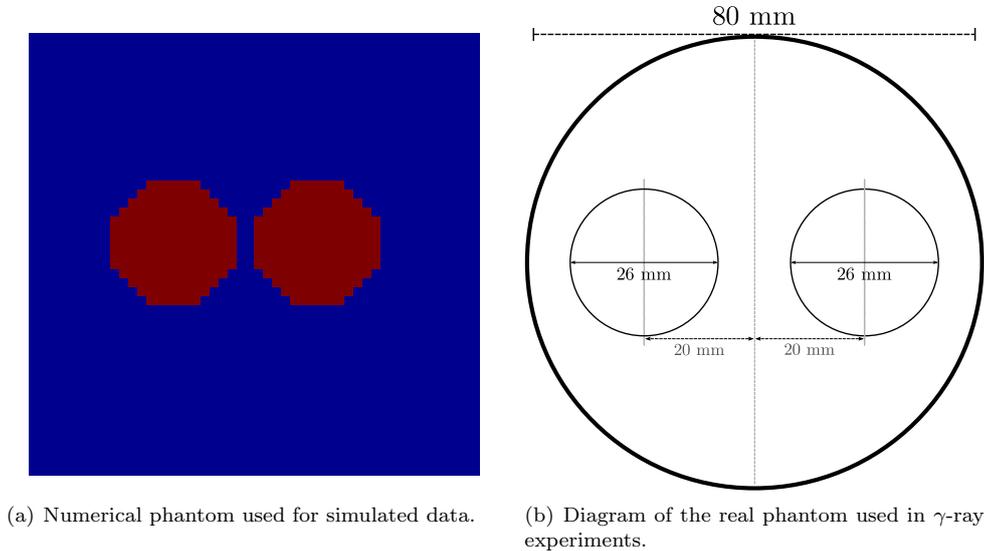


Figure 4: Phantom images used in the simulations (left) and the diagram of the phantom object used for collecting real data (right).

linear attenuation factor of our solid cylinder and the polypropylene available to construct the new phantom. The design and dimensions of this phantom were typical of those used to test the Bergen combined γ -ray and capacitance dual sensor system. Of course we could explore non-linearity using just the plugs suspended in air, but as the apparatus is arranged for a vertical imaging plane it would be more difficult to ensure the plugs were positioned accurately. Let L_0 denote the logarithmic counts for the two hole phantom and L_1 and L_2 the logarithmic counts with one of the plugs in a hole and L_{12} for both plugs. If the system was linear we would expect $L_1 - L_0$ to be equivalent to the normalized counts for just one plug in air. So superposition would give $L_1 - L_0 + L_2 - L_0 = L_{12} - L_0$ so we have $L_1 + L_2 - L_{12} - L_0$ as a measure of the non-linearity.

We also collected data suitable for comparing linear and non-linear reconstruction from multiple projection. The Bergen system is intended to be used for fast imaging using data from only five projections, collected simultaneously from five sources. To remove the limited data effect of only five projections and to eliminate variation caused by differences in the sources and detector arrays we collected imaging data by rotating the half cylinder and two hole phantom in angular increments to collect a fully sampled fan-beam projection data set. Again, the measurement time for each projection was 30s and thus, the statistical fluctuations in the data were negligible.

Finally our initial numerical experiments used a numerical phantom, see fig.4a. This consisted of a square pixel grid with a background of $\mu = 1$ with circular contrasting objects with $\mu = 2$. The dimensions of the square were taken so that the diagonal was contained in the 80mm region of interest. To avoid ‘inverse crimes’ the simulated data was generated using a grid with $2K \times 2K$ pixels Gaussian pseudo random noise was added to the data.

7 Results

7.1 Results of MC simulations

In the following, the results of MC simulations will be presented. Firstly, the MC implementation of the measurement geometry was benchmarked with experimental data. All simulations were run for a total of 10^8 primary photon histories, ensuring a relative statistical error of less than 2% for all tallies in the problem, i.e. collided and uncollided fluxes incident on the detectors for empty pipe, full pipe and phantom measurements. Each simulation was run on a six-core Intel Xeon 2.4

GHz CPU and took about 25 minutes to complete.

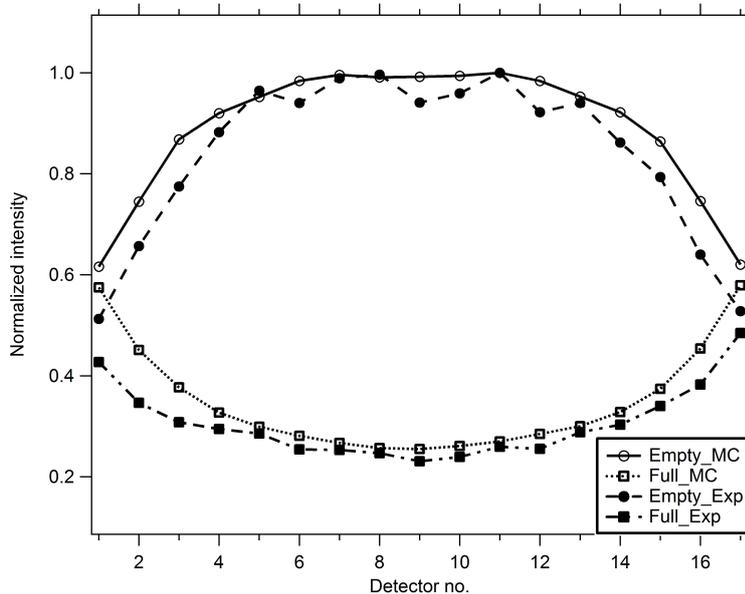


Figure 5: The normalized intensities recorded by each detector for empty and full pipes. The open legends show the results of MC simulations whereas the full legends show the experimental results. The normalization of the intensities is performed by dividing the intensities in each detector by the largest intensity obtained in the empty pipe measurements.

As can be seen in fig.5, the results of MC simulations showing the total flux of photons incident on the detectors agree reasonably well with the experimental data, especially for the central detectors in both full and empty pipe measurements. The MC simulations and experiments does, however, show a divergent behaviour for the side detectors. For the side detectors, as fig.5 reveals, the MC simulations predict higher intensities. This may be explained by considering the following factors; finer details of the measurement geometry are ignored in the MC modeling, only nominal values of the pipe dimensions are considered and positioning of the lead collimators. MC simulations do confirm this, as e.g., using slightly larger pipe diameter improves the agreement between MC simulations and experiments. As given in the following, altering these dimensions does not change the results significantly with respect to the non-linear partial volume effects vs. scattered radiation. Thus, for the purposes of this work, it was concluded that the MC simulations, given the results shown in fig.5, gave a reasonably well approximation of the experimental apparatus.

The non-linear effect was studied using a half-circular phantom simulating a pipe half oil filled and half empty (i.e., air). As given in Sec.7.2, the experiments show a significant non-linear effect at the interface between phantom and air whereas the largest non-linearity is seen for the centre-most detector, i.e. detector 9, which spans this interface. It would therefore be interesting to see whether this effect would be equally pronounced when there is no scattered radiation. The MC predicted photon flux incident on the detectors were separated into collided and uncollided components. As mentioned earlier, collided component includes contributions from both scattered primary photons as well as any secondary photons from atomic relaxations. The non-linearity was calculated for the uncollided flux as outlined in Sec.6.

As expected, and as shown in fig.6, non-linearity turns out to be largest for the centre-most detector, i.e. detector 9, in this detector module. More importantly, the benchmarked MC results show that the non-linearity will be substantial, even when scattered radiation is entirely eliminated.

The above observation partially confirms the statement that the non-linear partial volume effects are present even when the scattered radiation is eliminated entirely and that the non-

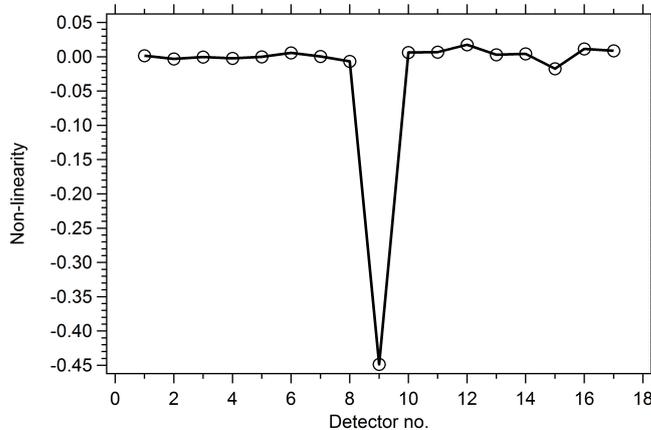


Figure 6: The MC generated non-linearity plotted against the detector number in a given detector module obtained for the half-circular phantom using only the uncollided photon flux. The results show that the non-linearity will be substantial (for the center-most detector, i.e. detector 9) even in case of no scattered photons.

linearity in the transmitted radiation is the dominant cause of errors in reconstructed images. A direct comparison based on the use of logarithms of collided and uncollided intensities is, however, difficult as the normalization of intensities is not straightforward. In addition, the overall intensity of scattered radiation is, in the given geometry, an order of magnitude less than the uncollided intensity. An indirect approach is therefore based upon reconstructing the phantom images using uncollided and total fluxes. Utilizing the reconstructed pixel densities, the gas volume fraction (GVF) can then be calculated for these two cases. The half-circular phantom represents a case with 50% GVF. The error in the calculated GVF can then be used as an indicator of the magnitude of the non-linear effects stemming from scattered radiation as well as that of transmitted radiation.

As only one source-detector module pair was implemented, full and empty pipe simulations were repeated five times using different random number seeds to obtain all five projections. The same procedure was utilized for the simulation of the phantom whereas the phantom was rotated at intervals of 72° . The tallied photon fluxes were separated as total (i.e., sum of collided and uncollided fluxes) and uncollided. The images of the phantom were reconstructed using the MC simulated data through a simultaneous multiplicative algebraic reconstruction technique (SMART) [13]. The GVF was calculated as an average weighted sum of the reconstructed pixel densities, $\alpha = 1/A \sum \alpha_i A_i$ where A is the pipe cross-section area, α_i are the reconstructed pixel densities and A_i are the pixel areas.

In fig.7, the MC simulated, normalized intensities obtained for the two cases, i.e. total and uncollided fluxes, are shown along with the rotation of the half-circular phantom. It should be kept in mind that, to accelerate the simulations, a photon transport cut-off energy of 48 keV was applied in all of the simulations. Thus the results reflect the intensities obtained for photons arriving at the detectors with energy greater than 48 keV. As mentioned earlier, this is in line with the experiments as a counting threshold of 48 keV is applied to each detector in the γ -ray tomography system.

The images reconstructed using the data plotted in fig.7 are given in fig.8. The reconstructed images were obtained for a total of 20 iterations on a 32×32 square grid. Using the reconstructed pixel densities without scatter rejection, a GVF of 36.7% was obtained whereas that obtained using pixel densities with scatter rejection, i.e. the uncollided flux, was calculated to be 37.5% giving an absolute deviation of 26.6% and 25% from the true GVF, respectively. These results support the statement that scattered radiation only has a marginal effect as compared to the non-linear effects in the transmitted intensity.

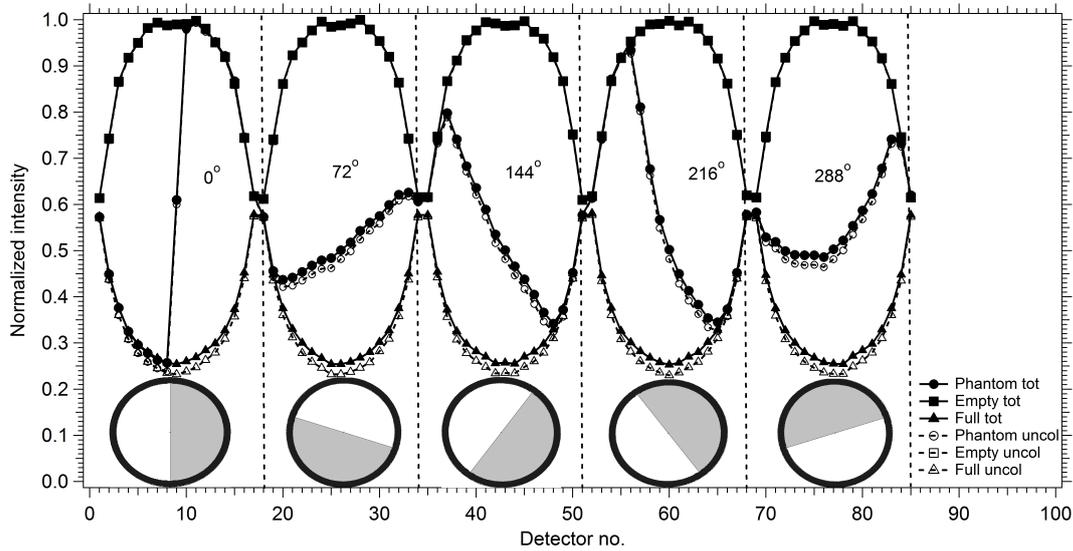


Figure 7: The MC generated, normalized intensities shown for both total (i.e., sum of collided and uncollided) and uncollided fluxes. Open legends and dashed lines show the intensities for the uncollided fluxes whereas the full legends and solid lines show those of total fluxes.

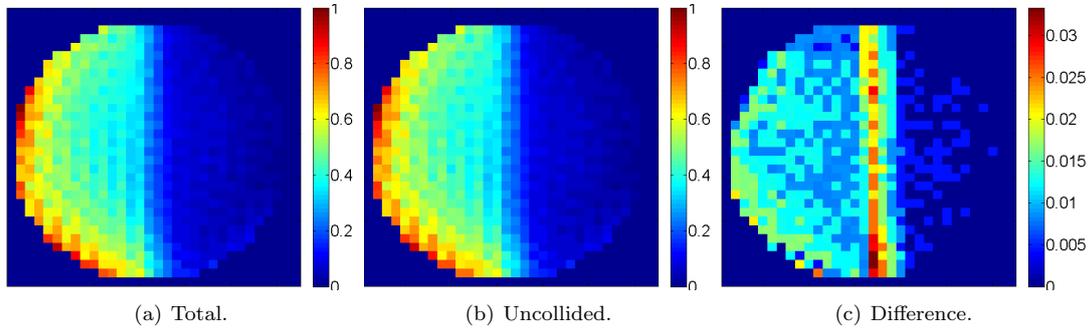


Figure 8: Reconstructed images of the half-circular phantom using MC generated projection data (a) without scatter rejection, (b) with scatter rejection and (c) the absolute difference between the two images. Images are reconstructed over a 32×32 square grid. The colour legend shows the GVF in each pixel.

It should, however, be emphasized that effects of scattered radiation will be dependent on the measurement geometry, more specifically on the source-detector separation as well as the collimation grid used to reduce the effects of scattered radiation. The geometry of the Bergen γ -ray tomography system is optimized to minimize the flux of scattered radiation incident on the detectors. The source-detector module separation is large compared to the pipe dimensions, the detectors are heavily collimated using lead blade collimators and a counting threshold of 48 keV is applied in order to further eliminate the influence of scattered radiation. These are factors that contribute to the fact that scattered radiation intensity in the given geometry is greatly suppressed. Thus, a new set of MC simulations were run where all of the lead collimators were removed and the source-detector module separation was reduced to half the initial value of 440 mm. This increases the flux of scattered radiation incident on the detectors by about a factor of 5. Here, the corresponding intensity profiles and reconstructed images will not be shown explicitly as these are essentially similar to the ones obtained for the original system shown in fig.7 and

fig.8. However, it is noteworthy that removing the collimators and reducing the source-detector module separation produces a GVF of 36.9% when the total fluxes are used. On the other hand, the calculated GVF equals 39.5% for the uncollided fluxes where the contribution of scattered radiation is simply rejected. This gives, respectively, an absolute deviation of 26.2% and 21% from the true GVF, representing a non-negligible change in the calculated GVF.

7.2 Results of experiments

Our initial experiment on the half circular phantom was to determine if there was a significant non-linear effect. The non-linearity was calculated as described in Sec.6 and the results are given in fig.9.

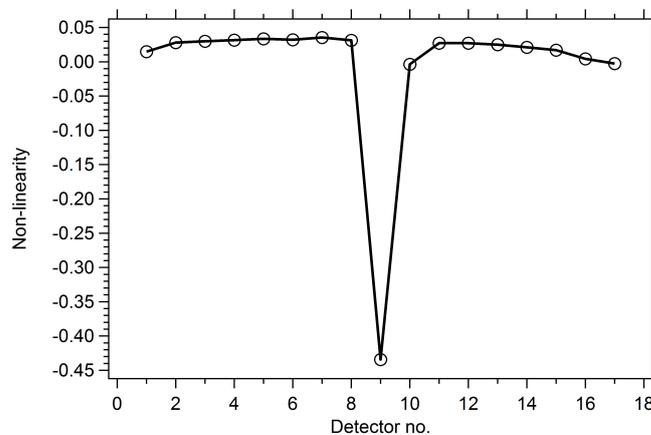


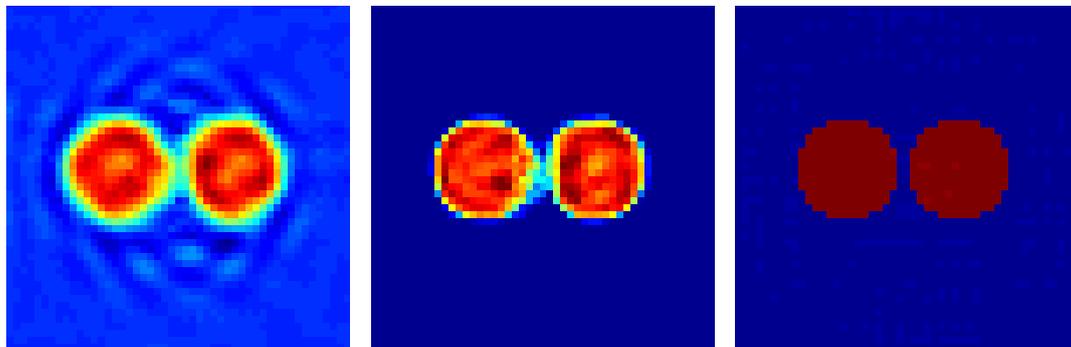
Figure 9: The experimental non-linearity plotted against the detector number in a given detector module obtained for the half-circular phantom. The results show that the non-linearity will be highest for the centre-most detector, i.e. detector 9, that spans the interface between air and plastic.

The figure show a graph of non-linearity in the logarithmic data from this experiment. As expected the non-linearity is highest for the source detector pair that spans the interface between air and plastic. As the phantom simulates a half full pipe of oil it was interesting that the non-linearity is significant for this important case.

We then tested the algorithms described in Sec.3 on the data from the numerical phantom with 2° rotations, giving 180 projections, and real γ -ray data with the Bergen γ -tomography system, also with 2° rotations. Fig.10 shows the reconstructions obtained via CGLS with Tikhonov regularization, primal-dual with TV regularization and the trust region reflective methods. For the CGLS runs, we chose the tolerance value $\epsilon = 1e-4$, and maximum number of iterations to perform to be $i_{\max} = 10^4$ for both numerically simulated and real γ -ray data reconstructions. The same tolerance value was also used in the trust region reflective method. In addition, TV parameters were fine-tuned separately for simulated and real data experiments: $\rho, \tau = 1.5$, $\alpha = 1e-5$, upper bound chosen as 1 for simulated; $\rho, \tau = 1.9$, $\alpha = 1e-3$, upper bound 5 for real γ -ray data reconstructions. The simulated geometry for these runs was set up to imitate the Bergen γ -tomography system, namely with 17 rays projected onto a flat array of detectors from a single source. The data was generated using the non-linear forward model applied to the phantom image in fig.4a, and 10% white Gaussian noise was added to the simulated γ -ray data. Of course this level of noise would represent a much shorter averaging time, in practice closer to how the apparatus is typically used for in-situ measurements. Except that of course in its standard use only five projections would be taken simultaneously.

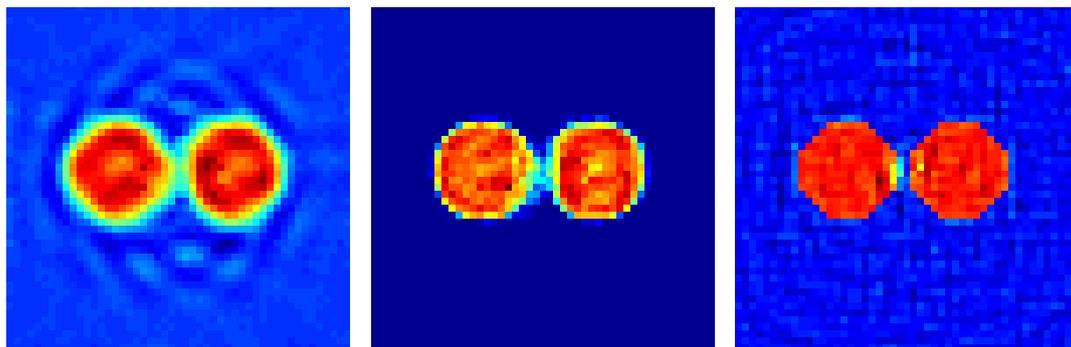
The numerical experiment was repeated with simulated data collected over 60 equiangular

projections. The results are shown in fig.11. The same reconstruction parameters were used as the previous case with 180 projection angles.



(a) CGLS with Tikhonov regularization. (b) Primal-dual with TV regularization. (c) Trust region reflective method.

Figure 10: Simulation results for linear and non-linear reconstructions. Simulated data is generated with 17 rays per projection, 180 projections. Images are reconstructed over a 50×50 square grid.



(a) CGLS with Tikhonov regularization. (b) Primal-dual with TV regularization. (c) Trust region reflective method.

Figure 11: Simulations are repeated with fewer angles (60 projections).

Finally we present reconstructions from the experimental γ -ray data collected with the Bergen γ -tomography system using the linear and non-linear models from the two hole phantom, illustrated in fig.4b. The results are shown in fig.12. In all reconstructions the perspex ring is clearly shown and the holes are visible. The CGLS method with Tikhonov regularization produces the poorest image (in quality) while primal-dual with TV and the trust region reflective reconstructions are very close in accuracy. However it should be noted that the over-smoothing with TV regularization is clear within the holes and the perspex ring, and further, over-estimating the thickness of the ring in comparison to CGLS with Tikhonov and trust region reflective methods. This is also demonstrated in the simulated results where the two hexagonal shapes are over-smoothed into more circular discs.

8 Discussion and Conclusions

Computed tomography problem is generally taken as a linear problem in the sense that the logarithm of the collected data on each detector, from each source, is a linear function of the attenuation

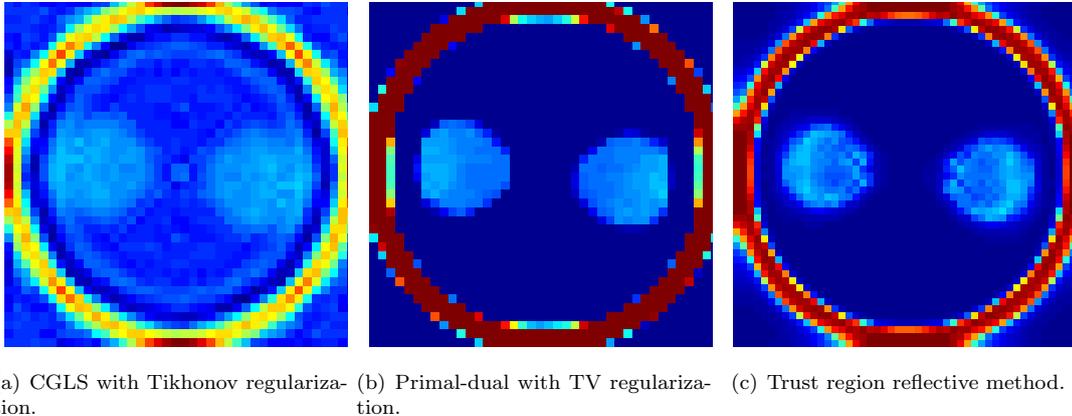


Figure 12: Real data reconstructions using the linear and non-linear models. Provided data contains 17 rays per projection, 180 projections. Images are reconstructed over a 50×50 square grid.

coefficient. However, as we demonstrated by considering the Beer-Lambert law, the sum of exponential does not imply the exponential of the sum. This effect, (rarely) noted in the literature as the “exponential edge-gradient” or “non-linear partial volume” effect, is magnified for systems with large sources and detectors producing a wider range of ray paths.

In this paper, we demonstrate that this non-linearity is clearly present in a monochromatic tomography system. For what we believe to be the first time in literature, we perform a non-linear reconstruction taking account of this effect using regularized non-linear optimization methods.

For the reconstructed results presented in this paper, we used the trust region reflective method to numerically optimize the derived non-linear model. The results were compared with the reconstructions obtained with a popular reconstruction method, CGLS. The stark differences in the CGLS and the trust region reflective reconstructions highlight the dominance of the non-linearity effect around the edges of the objects in the phantoms. This effect is also highlighted in simulated primal-dual TV regularized reconstructions where two hexagonal discs are close to one another.

These initial reconstruction results show that for sufficiently large sources and detectors taking account of the non-linearity can produce dramatic improvements. However there is still much work to be done. We have solved a non-linear problem where the linearization about the background case is well explored and clearly invertible under the right conditions on the pixel grid. Detailed analysis is required to show that our regularized non-linear optimization problem does not suffer from local minima. Also we used a standard optimization algorithm and a more customized algorithm for this problem is likely to be more efficient. We chose the simplest regularization that imposes a smoothness condition on the image. We also implemented a more sophisticated algorithm for objects with distinct boundaries and relatively homogeneous interiors, for which total variation regularization gave better images, as did the imposition of strict upper and lower bounds using constrained optimization methods.

In the specific case of the Bergen γ -ray tomography system, the lead collimators were expected to reduce counts registered due to scattering so we are confident that the non-linearity we observed is largely due to the mechanism we describe. The results of MC simulations of the system confirm this statement. The MC simulations predict that the intensity of the scattered radiation is about an order of magnitude less than the transmitted radiation intensities for the given system. This is not surprising as the system has been designed and optimized to minimize the effects of scattered radiation. The reconstructed images of the MC simulated half circular phantom show also that the error in the calculated GVF based on the reconstructed pixel densities has the same order of magnitude even for scatter-free data. Most importantly, for the half circular phantom, the non-linearity in the center-most detector is substantial even in case of full scatter rejection,

indicating the predominance of non-linear partial volume effects. Through MC simulations, it was also demonstrated that, by removing the collimation grid entirely and reducing source-detector separation, scattered radiation has a more pronounced effect on the reconstruction accuracy. However, the non-linear reconstruction on experimental data as well as absolute deviations from the true GVF calculated for the MC generated total and scatter-free data reveal that it would be necessary to take into account also the non-linearity for a substantial improvement in the reconstruction accuracy. For polychromatic tomography systems, MC simulations would also be able to predict how much of the non-linearity in a polychromatic system is due to the size of the source and detectors as opposed to the variation of attenuation with energy (beam hardening). In X-ray systems the source would be expected to be much less uniform. In [11] a Gaussian was used to model the source intensity of the focal spot in a micro-focus laboratory CT system. Such variation of I_0 is easily incorporated in our algorithm if it is known.

Acknowledgements

WRBL is partly funded by The Royal Society Wolfson Research Merit award, WRBL and SBC would like to thank EPSRC for support from grants EP/M022498/1, EP/I01912X/1 and EP/M010619/1 and the EU for COST Action MP1207 EXTREMA. The design and development of the University of Bergen γ -ray tomography was funded by the Research Council of Norway, Norsk Hydro ASA (now merged into Statoil ASA) and the University of Bergen. Recent hardware upgrade development work has been funded by GCE Subsea (grant# 301602). IM would like to acknowledge a former student, IVM Moreira for access to reconstruction software used for reconstructing the MC simulated data.

References

- [1] A. Beer. Bestimmung der absorption des rothen lichts in farbigen flssigkeiten. *Annalen der Physik*, 162(5):78–88, 1852.
- [2] M. A. Branch, T. F. Coleman, and Y. Li. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing*, 21(1):1–23, 1999.
- [3] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.
- [4] T. F. Coleman and Y. Li. On the convergence of interior-reflective newton methods for nonlinear minimization subject to bounds. *Mathematical Programming*, 67(1):189–224, 1994.
- [5] B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens. Metal streak artifacts in x-ray computed tomography: a simulation study. *Nuclear Science, IEEE Transactions on*, 46(3):691–696, 1999.
- [6] G. H. Glover and N. J. Pelc. Nonlinear partial volume artifacts in x-ray computed tomography. *Medical physics*, 7(3):238–248, 1980.
- [7] T. Goorley et al. Initial MCNP6 release overview. *Nuclear Technology*, 180:298–315, 2012.
- [8] P. C. Hansen. *Rank-deficient and Discrete Ill-posed Problems: Numerical Aspects of Linear Inversion*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.
- [9] G. A. Johansen, T. Frøystein, B. T. Hjertaker, and Ø. Olsen. A dual sensor flow imaging tomographic system. *Measurement Science and Technology*, 7(3):297, 1996.
- [10] P. M. Joseph and R. D. Spital. The exponential edge-gradient effect in x-ray computed tomography. *Physics in medicine and biology*, 26(3):473, 1981.

- [11] A. Kueh, J. M. Warnett, G. J. Gibbons, J. Brettschneider, T. E. Nichols, M. A. Williams, and W. S. Kendall. Modelling the penumbra in computed tomography. *Journal of X-Ray Science and Technology*, 24(4):583–597, 2016.
- [12] H. Li, A. Kingston, G. Myers, B. Recur, and A. Sheppard. 3D x-ray source deblurring in high cone-angle micro-ct. *IEEE Transactions on Nuclear Science*, 62(5):2075–2084, Oct 2015.
- [13] R. Maad, B. T. Hjertaker, and G. A. Johansen. Experimental analysis of high-speed gamma-ray tomography performance. *Measurement Science and Technology*, 19:085502, 2008.
- [14] F. Natterer. *The Mathematics of Computerized Tomography*, volume 32. Siam, 1986.
- [15] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer, New York, NY, 2. ed. edition, 2006.
- [16] J. Nuyts, B. De Man, J. A. Fessler, W. Zbijewski, and F. J. Beekman. Modelling the physics in the iterative reconstruction for transmission computed tomography. *Physics in medicine and biology*, 58(12):R63, 2013.
- [17] Å. Björck, T. Elfving, and Z. Strakos. Stability of conjugate gradient and Lanczos methods for linear least squares problems. *SIAM J. Matrix Anal. Appl.*, 19(3):720–736, 1998.
- [18] E. Sundermann, F. Jacobs, M. Christiaens, B. De Sutter, and I. Lemahieu. A fast algorithm to calculate the exact radiological path through a pixel or voxel space. 1998.
- [19] W. M. Thompson, W. R. B. Lionheart, E. J. Morton, M. Cunningham, and R. D. Luggar. High speed imaging of dynamic processes with a switched source x-ray ct system. *Measurement Science and Technology*, 26(5), 2015.
- [20] N. Wadson, E. Morton, and W. R. B. Lionheart. Scatter in an uncollimated x-ray ct machine based on a geant4 monte carlo simulation. In *SPIE Medical Imaging*, pages 76223E–76223E. International Society for Optics and Photonics, 2010.
- [21] H. Yu and G. Wang. Finite detector based projection model for high spatial resolution. *Journal of X-ray Science and Technology*, 20(2):229–238, 2012.

A Point sensitivity

Consider a two dimensional example where (see figure 13) The source is a line segment $S'S$ and the detector is a parallel line segment DD' placed opposite the source at a distance L . Let X be the point under consideration. Consider the lines XS and XS' from the ends of the source meeting the detector line DD' at Q and Q' which may or may not be within the detector. Similarly extend lines DX and $D'X$ from either end of the detector to meet the source line SS' at P and P' . The only rays passing through X that result in a measurement are those between points in the intersection of line segments $DD' \cap QQ'$ and $SS' \cap PP'$ so the point sensitivity is proportional to the product of the lengths of those intervals.

The length of the intersection of the intervals (a, b) and (c, d) is

$$r(a, b, c, d) = \max\{0, \min\{b - a, b - c, d - c, d - a\}\}$$

as can be seen from consideration of all the possible cases. Introducing coordinates with the origin in the centre of the region between source and detector and $X = (x, y)$ we have $S^\pm = (L/2, \pm W_S/2)$ $D^\pm = (L/2, \pm W_D/2)$

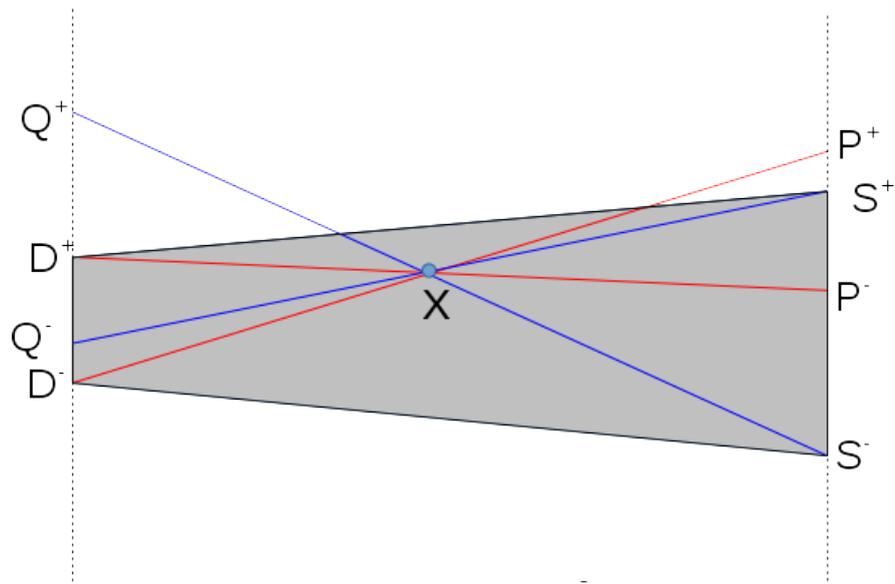


Figure 13: Simple two-dimensional case of point sensitivity.

Chapter 5

4-D imaging of sub-second
dynamics in pore-scale processes
using real-time synchrotron X-ray
tomography

Solid Earth, 7, 1059–1073, 2016
 www.solid-earth.net/7/1059/2016/
 doi:10.5194/se-7-1059-2016
 © Author(s) 2016. CC Attribution 3.0 License.



4-D imaging of sub-second dynamics in pore-scale processes using real-time synchrotron X-ray tomography

Katherine J. Dobson^{1,a}, Sophia B. Coban², Samuel A. McDonald³, Joanna N. Walsh³, Robert C. Atwood⁴, and Philip J. Withers³

¹Ludwig-Maximilians-Universität München, Department für Geo-und Umweltwissenschaften, 80333 München, Germany

²School of Mathematics, University of Manchester, Manchester, M13 9PL, UK

³School of Materials, University of Manchester, Manchester, M13 9PL, UK

⁴Diamond Light Source, Harwell Innovation Campus, Didcot, OX11 0DE, UK

^anow at: Durham University, Department of Earth Sciences, South Road, Durham, DH1 3LE, UK

Correspondence to: Katherine J. Dobson (katherine.dobson@durham.ac.uk)

Received: 1 March 2016 – Published in Solid Earth Discuss.: 21 March 2016

Revised: 1 June 2016 – Accepted: 7 June 2016 – Published: 15 July 2016

Abstract. A variable volume flow cell has been integrated with state-of-the-art ultra-high-speed synchrotron X-ray tomography imaging. The combination allows the first real-time (sub-second) capture of dynamic pore (micron)-scale fluid transport processes in 4-D (3-D + time). With 3-D data volumes acquired at up to 20 Hz, we perform in situ experiments that capture high-frequency pore-scale dynamics in 5–25 mm diameter samples with voxel (3-D equivalent of a pixel) resolutions of 2.5 to 3.8 μm . The data are free from motion artefacts and can be spatially registered or collected in the same orientation, making them suitable for detailed quantitative analysis of the dynamic fluid distribution pathways and processes. The methods presented here are capable of capturing a wide range of high-frequency nonequilibrium pore-scale processes including wetting, dilution, mixing, and reaction phenomena, without sacrificing significant spatial resolution. As well as fast streaming (continuous acquisition) at 20 Hz, they also allow larger-scale and longer-term experimental runs to be sampled intermittently at lower frequency (time-lapse imaging), benefiting from fast image acquisition rates to prevent motion blur in highly dynamic systems. This marks a major technical breakthrough for quantification of high-frequency pore-scale processes: processes that are critical for developing and validating more accurate multiscale flow models through spatially and temporally heterogeneous pore networks.

1 Introduction

Porosity, permeability, and flow in geological systems are all highly dynamic. Changes in the confining conditions (e.g. pressure, temperature), flow volume, fluid chemistry/viscosity, or suspension composition can drive mass transport (through processes such as precipitation, dissolution, deposition, or erosion) and change the connectivity and tortuosity of the pore network. The evolving porosity and permeability then cause further changes in both micro- and macro-scale flow. These mass transport processes and the passage of fluid–fluid and fluid–rock reaction fronts through heterogeneous geological systems occur at the pore scale. Developing a detailed appreciation of macroscopic processes as diverse as pollutant transport, hydrocarbon recovery, CO₂ sequestration, storage of nuclear waste, aquifer management, nutrient bio-accessibility, building stone preservation, and hydrothermal deposit formation requires a quantitative understanding of the multiscale effect of pore-scale processes.

Synchrotron imaging for dynamic geoscience applications

Laboratory and synchrotron X-ray computed micro tomography (usually XCT and sXCT respectively) are ideal approaches for imaging these processes as they allow in situ observation in a high-resolution, non-destructive way. The general principles of X-ray tomography, reconstruction, and data processing as related to geoscience applications are covered

in a series of recent reviews (Bultreys et al., 2016; Cnudde and Boone, 2013; Hess et al., 2011; Maire and Withers, 2014; Wildenschild and Sheppard, 2013). For the specific application of understanding pore-scale processes, XCT and sXCT are now becoming widely used for both qualitative and quantitative imaging of complex natural pore networks and the distribution of liquid(s) within them (Al-Raoush et al., 2011; Al-Raoush and Willson, 2005; Berg et al., 2013; Bhreasail et al., 2012; Boone et al., 2014; Bultreys et al., 2015a; Cnudde and Boone, 2013; Dewanckele et al., 2012; Geraud et al., 2003; Herring et al., 2013; Iglauer et al., 2011; Katuwal et al., 2015; Ma et al., 2016; Naveed et al., 2013b; Olafuyi et al., 2010; Sakellariou et al., 2003; Sok et al., 2010; Wildenschild et al., 2002; Wildenschild and Sheppard, 2013). The data are also being used as a basis of, and validation, for numerical simulations (Al-Raoush and Papadopoulos, 2010; Alhashmi et al., 2015; Bultreys et al., 2015b, 2016; Degruyter et al., 2010; Menke et al., 2015; Naveed et al., 2013a; Raeini et al., 2014, 2015; Fourie et al., 2007).

Standard operation is to collect a set of 2-D “projections” or “radiographs” at constant angular spacing (hereafter angular density) while the sample is rotated through 180° or 360°. Standard 3-D tomographic image datasets require seconds (synchrotron), minutes (synchrotron and laboratory), or even hours (laboratory) to acquire (Maire and Withers, 2014). Until recently this has limited 3-D experimental investigations of dynamic processes because the critical pore-scale processes occur over much shorter durations. Many key fluid–rock and fluid–fluid interactions therefore remain poorly constrained. Furthermore, while numerical simulations can now consider multiple mass and thermal transport processes simultaneously and can incorporate realistic pore geometries (see Bultreys et al., 2016, for in depth review), many generally still lack validation from experimental data from natural systems where observations are needed with a range of temporal and spatial resolutions.

In recent years, careful alignment of 3-D datasets collected at fixed time points over moderate duration experiments has enabled “time-integrated” tomography as a tool for quantification of fluid dynamics and porosity evolution (Andrew et al., 2014, 2015; Armstrong et al., 2014a; Berg et al., 2013; Blunt et al., 2013; Herring et al., 2013, 2014; Lin et al., 2016; Menke et al., 2015; Wildenschild and Sheppard, 2013). The increased image acquisition rates now available at third-generation synchrotron facilities have driven development of faster “continuous” imaging. In this mode, sample rotation is not stopped for each projection; instead data are acquired over a narrow arc (typically 0.1–0.5°). For most geological materials, collection of the projection data can be just a few seconds (Andrew et al., 2015; Berg et al., 2013; Bhreasail et al., 2012; Pistone et al., 2013, 2015; Youssef et al., 2014). However, 2-D imaging (the same projection acquisition rates but without rotation) is still needed to observe processes occurring on a timescales of milliseconds to a few seconds, such as individual Haynes jumps and the sub-

sequent relaxation dynamics (Armstrong et al., 2014b; Berg et al., 2013). Two-dimensional imaging cannot capture pore and pore–fluid interface morphology, pore throat orientation and size, or the location of the events within the fluid volume.

Advances in temporal resolution through camera and beamline technology have been followed by highly parallelised iterative reconstruction methods which require far fewer projections yet still achieve adequate reconstructed image quality (Batenburg and Sijbers, 2011; Brabant et al., 2014; Kaestner et al., 2015; Kazantsev et al., 2015a, b; Van Eyndhoven et al., 2015). However, direct in situ observation of sub-second events and processes, as well as those that cause rapid change over longer durations, still remain challenging. Pushing time-integrated tomography towards true long duration 4-D imaging would therefore have high impact on our understanding of pore-scale processes.

Here we present a state-of-the-art 4-D (3-D + real time) imaging methodology that enables visualisation and quantitative assessment of dynamic pore-scale processes in real-time (3-D acquisition rates of up to 20 Hz) over variable experimental durations. We present two key methodology developments: sub-second real-time imaging of fluid transport up to an order of magnitude faster than previously presented and the extension of that imaging protocol to allow slower or variable flow dynamics to be quantitatively assessed over longer time periods. We illustrate the potential of these developments for providing insight into pore-scale processes using three case studies showing preliminary data from experiments that quantify initial wetting, the evolution from dynamic to steady state flow, and the evolution of the reaction front during chemical dilution or fluid mixing.

2 Experimental set-up

Disaggregated sandstone gravels (1–2 or 2–4 mm size fraction) were loaded into a reusable cylindrical gravity-fed flow cell and mounted on the JEEP i12 beamline (Diamond Light Source, Harwell, UK) (Fig. 1). Full beam specification can be found in Drakopoulos et al. (2015). The cell diameter for this system can be varied (up to 50 mm diameter) according to experimental need and X-ray transmission through the sample. Perspex sample chambers of 6 and 25 mm internal diameter were used here. A circular piece of fine mesh was placed at the bottom of the cell above the outlet feed to prevent fine material being washed through to the outflow reservoir (maximum volume 125 mL in the configuration used). Particles were added to a minimum depth of 15 mm to ensure no effects from basal flow were observed in the imaging field of view. Particle diameters were sufficiently small that the influence of the cell geometry on particle packing was minimal. The random nature of the packing structure was checked before each experiment through the collection of a single tomography dataset (Fig. 2). In the well-cemented host sandstones (prior to disaggregation) grain sizes ranged from 200

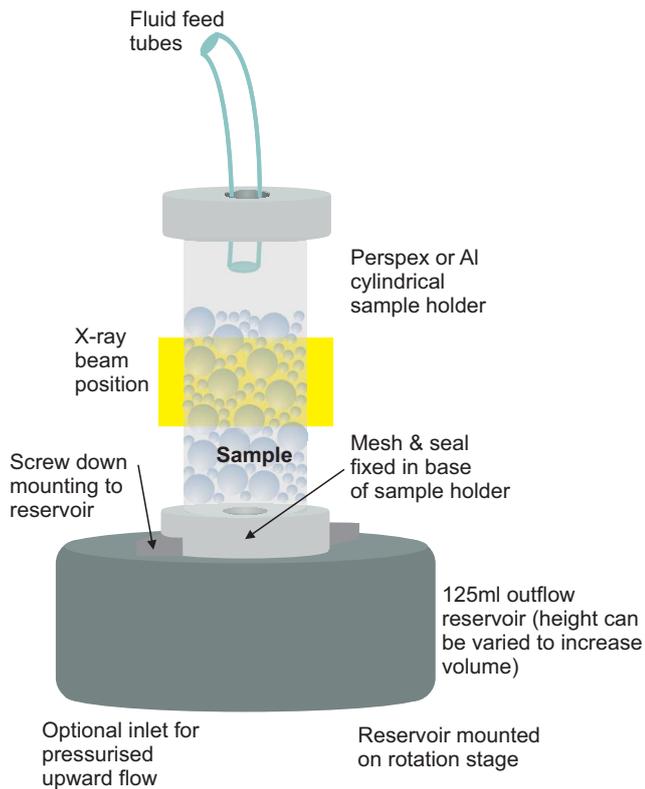


Figure 1. A schematic showing the simple peristaltic pump drip-fed gravity-driven flow cell. The system was designed to be modifiable to incorporate laminar and multiphase flow and to be suitable for use with confining pressure and variable temperatures. All key dimension (sample cell diameter, height, and material; reservoir volume; feed) can be adjusted to fit experimental parameters and beamline imaging conditions.

to 500 μm , with pore diameters of 10–150 μm . Mercury and helium porosimetry and water absorption measurements on the disaggregated samples were 13–18, 15–23, and $\sim 19\%$ respectively, and the measured permeabilities ranged from 250 to 600 md. In the configuration used here, fluid was supplied at a constant flow rate by a 12-channel peristaltic pump operated from the beamline control room (outside the X-ray enclosure or “hutch”) so that flow rates and compositions could be adjusted in real time, i.e. while 3-D data are being acquired. The inlet tube was attached centrally to the top of the cell using a slip ring to ensure a fixed drip point (approx. 5 mm above the centre of the rotating sample). Flow rates were varied from 40 to 200 $\mu\text{L min}^{-1}$, which corresponded to a drip every 2 and 10 s. The majority of experimental runs were performed using a 6 M KI solution to ensure high contrast between air, fluid, and rock particles. Pilot scans showed un-doped H_2O and some oils could be observed under the hard X-rays of the JEEP white beam (50–150 kV), although low contrast would make post-processing challenging.

The flow cell was designed to be easily modified and can accommodate base fed laminar flows, simultaneous flow

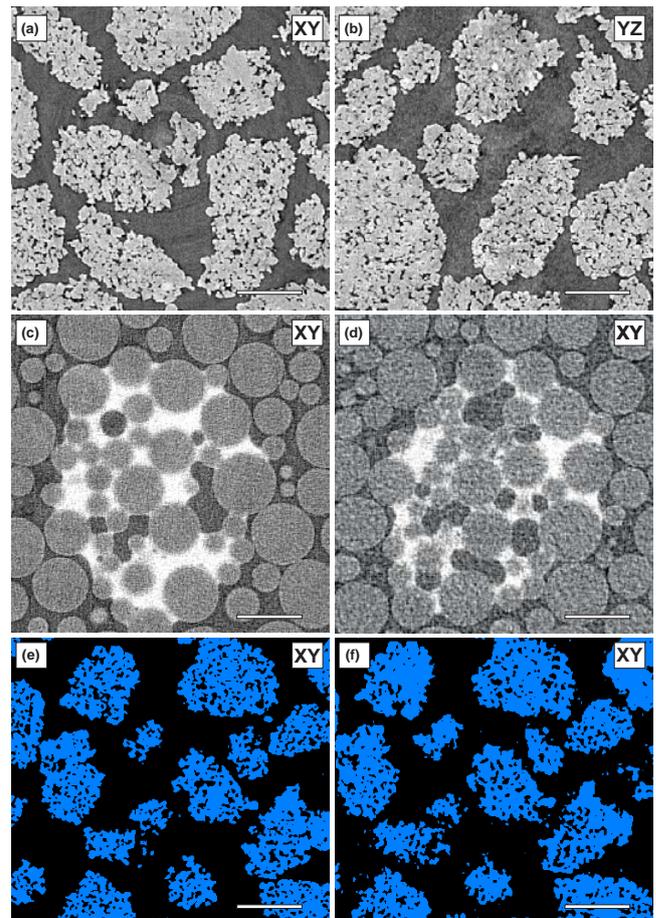


Figure 2. Two-dimensional reconstructed slices through a typical gravel sample. High-resolution (1800 projection data) perpendicular (a; XY) and parallel to flow direction (b; YZ) reconstructed using filtered back projection. (c) A 2-D slice (XY) through a simple spherical bead pack containing a small volume for KI solution reconstructed from 360 projections (2°) using filtered back projection; (d) the same 2-D slice reconstructed from 180 projections (1°) using filtered back projection; (e, f) comparison of the segmentation of the rock possible from the 1800 and 180 projection data. All scale bars are 200 μm .

of multiple fluids (variable composition or temperature), or multiphase suspensions. Thin-walled (1 mm) aluminium sample chambers can be used for both variable thermal and pressure experiments. The cell can be loaded with disaggregated material (e.g. soil, sand, gravel, glass beads) or with more coherent cored soil/rock samples of any diameter. However, when designing ultra-high-speed experiments or processes that cause subtle changes in attenuation/composition, some higher density or large volume solid samples will increase the minimum exposure time and may also cause artefacts that reduce the ability to resolve features of interest.

3 Ultra-fast image acquisition

The maximum rotation speed of the JEEP tomography stage is 10 Hz (continuous rotation). As synchrotron tomography datasets are acquired over 180°, 3-D volumes can be acquired at up to 20 Hz. The actual operational rotation speed for the given experiment will be controlled by the exposure time for each projection and the total number of projections collected per 3-D dataset. Exposure time is controlled by X-ray flux and sample density. Imaging at the JEEP beamline can be performed using a polychromatic (white, higher flux) or monochromatic X-ray beam (selectable energy range 53–150 kV). Images were acquired using a Vision Research Phantom Miro™ 310 M camera (up to 5000 images per second). In monochromatic beam mode, the exposure time for the large diameter cell was 200 μ s (3.8 μ m pixel resolution), which dropped to 90 μ s for the smaller diameter system (2.5 μ m pixel resolution). The system can be in continuous rotation at this speed, in one direction, for as long as required. Beam hardening can occur under white beam in response to the energy dependence of the attenuation, but none was seen in our test data. While white beam would allow shorter exposure times (between 30 and 50 % reduction in the test scans), there would be lower contrast between the low density phases and more X-ray scattering and associated noise in the images. Imaging in monochromatic beam gave increased contrast between air and water in a three phase system, without requiring long exposures. The diameter of the samples used here exceeds the field of view of the highest magnification module available at JEEP (1.25 μ m pixel resolution), but region-of-interest images could be obtained with a 200 μ s exposure. Qualitative analysis of the images at different rotation speeds showed no visible dependence of liquid distribution on acquisition (rotation) speed, suggesting the centrifugal effect on this sample and fluid was negligible.

In most “traditional” tomography experiments, over 1000 projections are collected for each 3-D volume. We collected high-resolution data (1800 projections, 0.2 and 0.5 s per 3-D data acquisition) for each dry sample prior to flow initiation to characterise the inter- and intra-grain porosity of each sample (Fig. 2a, b). This dataset can also be used as a structural prior for iterative reconstruction methods (e.g.; Kazantsev et al., 2015a, b; Van Eyndhoven et al., 2015). Reducing the angular density of the projections will shorten overall scan time. The reconstructed image quality does reduce as the projection density drops, but a significant increase in temporal resolution can be achieved. To achieve the maximum 20 Hz 3-D image acquisition rate, it is necessary to reduce the number of projections to 250 (200 μ s exposure) and 550 (90 μ s exposure) per scan.

Under-sampling (using fewer projections) to this degree does cause image quality to degrade, making differences between phases of similar attenuation difficult to observe and reducing the precision with which phase boundaries and volumes can be defined. When features of interest are relatively

small, or have complex geometries, then under-sampling can prevent quantitative analysis. However, here the contrast between the saline solution, rock, and air is high, and the primary target is the sub-second distribution of the fluid phase. Using the routine filtered back-projection reconstruction algorithms, testing with as few as 360 (Fig. 2c) and 180 projections (Fig. 2d) gave adequate images for basic quantification (Fig. 2e). Comparison of the two under-sampled datasets shows little increase in quality and little overall change between the 360 and 180 projection scans. For the smaller cells, it is possible to collect 360 projections at maximum image acquisition frequency.

The Miro camera has fixed on-board data storage, which allowed a maximum of approximately 20 000 projections to be collected for the field of view needed in this experiment (in this set-up, the number of projections is dependent on the field of view needed). The under-sampling allowed between 53 and 141 3-D datasets to be acquired in a single experimental run (2.5–7 s of acquisition at maximum frequency). Capturing a specific event with this short a duration of collection is challenging, especially as the high rotation speeds prevent observation of small volume changes in the live view radiographs. During collection the camera is armed and then triggered. Arming the camera starts projection acquisition, but those data are automatically overwritten until the trigger signal is received. The user can define the trigger to mark (a) the last image (all existing projections on the card are saved), (b) the first image (projections are collected until the card volume is full), or (c) an event (a user-defined number of projections before and after the trigger is saved). The latter option allows a small amount of buffer time for the user to capture a specific event, such as a drip entering the sample or a flow volume/composition change. Automatic triggering mechanisms such as light gates can be incorporated although in this experiment a webcam was used to observe the delivery of each drip.

To clarify the real-time nature of the data, the following terminology is used in all subsequent discussion: one or more scans are collected during each experiment, and each scan contains a number of 3-D “frames” which were acquired at a known speed and with a given frequency (defined by the time between the first projections of sequential frames). Data are offloaded from the camera between scans (download takes between 5 and 7 min). Several scans can be collected during an experiment. To highlight specific processes and features, most figures presented below show only 2-D slices (either perpendicular or parallel to flow) or a sub-volume of the reconstructed dataset. The post-processing and image segmentation are the same for all examples and were performed in Avizo™ and MATLAB. The methods are outlined in Appendix A.

Case study: observing sub-second dynamics using ultra-fast imaging

Wetting processes and the formation of rivulet networks are poorly understood in heterogeneous media. The processes operate on short timescales and are strongly influenced by sample morphology. Non-invasive 4-D tomography is ideal for capturing the dynamics of these behaviours. Figure 3 shows the evolution of the fluid distribution in a sub-volume of the sample immediately below the drip point. The data were collected under continuous acquisition mode, with a 3-D frame acquisition frequency of 0.07 s (system operating at 7 Hz rotation speed).

The first drip of KI solution (shown in blue) enters the dry sandstone bed (rendered in translucent grey) between frames F0 and F1 (Fig. 3). Initially ($t = 0.07$ s) the KI solution fills the inter-grain pore space to the front of the sub-volume (A) and then immediately begins lateral transport into the intra-grain porosity in the neighbouring grains (Fig. 3b). By the second frame ($t = 0.14$ s) the inter-grain volume has largely emptied, and the majority of the fluid is now within the grains. Intra-grain transport has also occurred with capillary-driven wetting of grains not in direct contact with inter-grain fluid (e.g. C), and drainage of previously saturated intra-grain pores can also be observed from other viewing angles. Between F1 and F2 most grains in contact with the inter-grain fluid have become partially or fully saturated. The remaining intra-grain fluid is generally limited to a few small pores and surface films (D). The figure only shows the relaxation after the first drip. As the experiment continues (F3 to F5), the fluid moves away from the initial location, with grain saturation increasing by radial and vertical redistribution, with oblique upward propagation of the saturation front in some instances (E). Most grains show an expansion and coalescence of grain-surface fluid films after initial intra-grain saturation (F), although grain surface wetting does occur with no, or prior to, intra-grain saturation (G). Capillary forces continue to redistribute the intra-grain pore fluid over the remainder of the experiment, with saturation of many grains decreasing as the system moves towards equilibrium. Grains at both the upper and lower extent of the wetted region (H) are almost completely saturated and then drained again within 1.54 s.

Vertical slices taken through the same sub-volume of the sample in the second scan of the experiment show the arrival of the third drip (Fig. 4). In this experiment the fluid feed was stopped during data download. When the pump was restarted, the first drip was used as the trigger for the scan acquisition. The initial frames of the second scan were before the arrival of the drip and act as a wet reference. Differences in the fluid volume across the download period result from the fall of a drip at the very end of the experiment, immediately after fluid feed was stopped. Figure 4 captures the major changes in the fluid distribution in the inter-grain porosity (blue). At this time, the intra-grain fluid distribution (grey)

has become largely stable, although a few grains still show some variability (Fig. 4b).

Passage of the drip causes transient fluid bridging (Fig. 4a). Most bridging connections are maintained throughout the drip cycle, but some transient connections are formed with every drip, and others form only occasionally. Temporary isolation of fluid (enhanced by capillary-driven draw in of the surface film during equilibration as observed at A; Fig. 4) could control chemical reaction rates in reactive systems. Upon reconnection the relatively small volume of this fluid would be rapidly diluted, but rates of dissolution would be locally reduced by limited fluid supply. A later part of the experiment is shown in Movie 1 (see Supplement).

Although the focus of this contribution is on method development, we include a limited selection of the quantitative analysis performed on these samples, as an illustration of the adequacy of the data quality. Quantitative analysis of inter- and intra-grain fluid volume shows the strong fluctuation over the course of two drips (Fig. 5) and the general instability of the fluid volume as it is redistributed. After the first drip, the saturation of the intra-grain porosity becomes more gradual, with significant increases on the arrival of subsequent drips. The inter-grain distribution equilibrates more slowly.

4 Ultra-high-speed imaging for longer-duration experiments

The experiments presented above allow capture of the high-speed dynamics but have very short overall durations (< 10 s). Capturing longer duration processes with the same 3-D frame acquisition rates requires lower frame acquisition frequencies. The Miro camera allows frames to be acquired either continuously (as above) or at user-defined intervals (hereafter called gapped acquisition). Gapped acquisition decreases frame acquisition frequency, while maintaining frame acquisition speeds and the resultant data therefore benefit from a lack of motion induced blurring in the reconstructed data.

Although any time or angular gap can be defined, in these experiments they were limited to odd multiples of 180° . This ensures that the initial projection of every frame was collected at the same angle. Registration of the data in this manner saves significant time aligning datasets during post-processing. A gapped acquisition with a spacing of 1 (G1) collects projections over the 0 – 180° sector of every rotation (one full rotation between acquisition start points); a gapped acquisition with spacing of 5 (G5) acquires projections from 0 to 180° on every third rotation (three full rotations between acquisition start points).

Although not used in this experiment, the camera memory can also be partitioned, allowing for a fixed and constant number of projections/frames in each partition. This could enable more efficient imaging across different tempo-

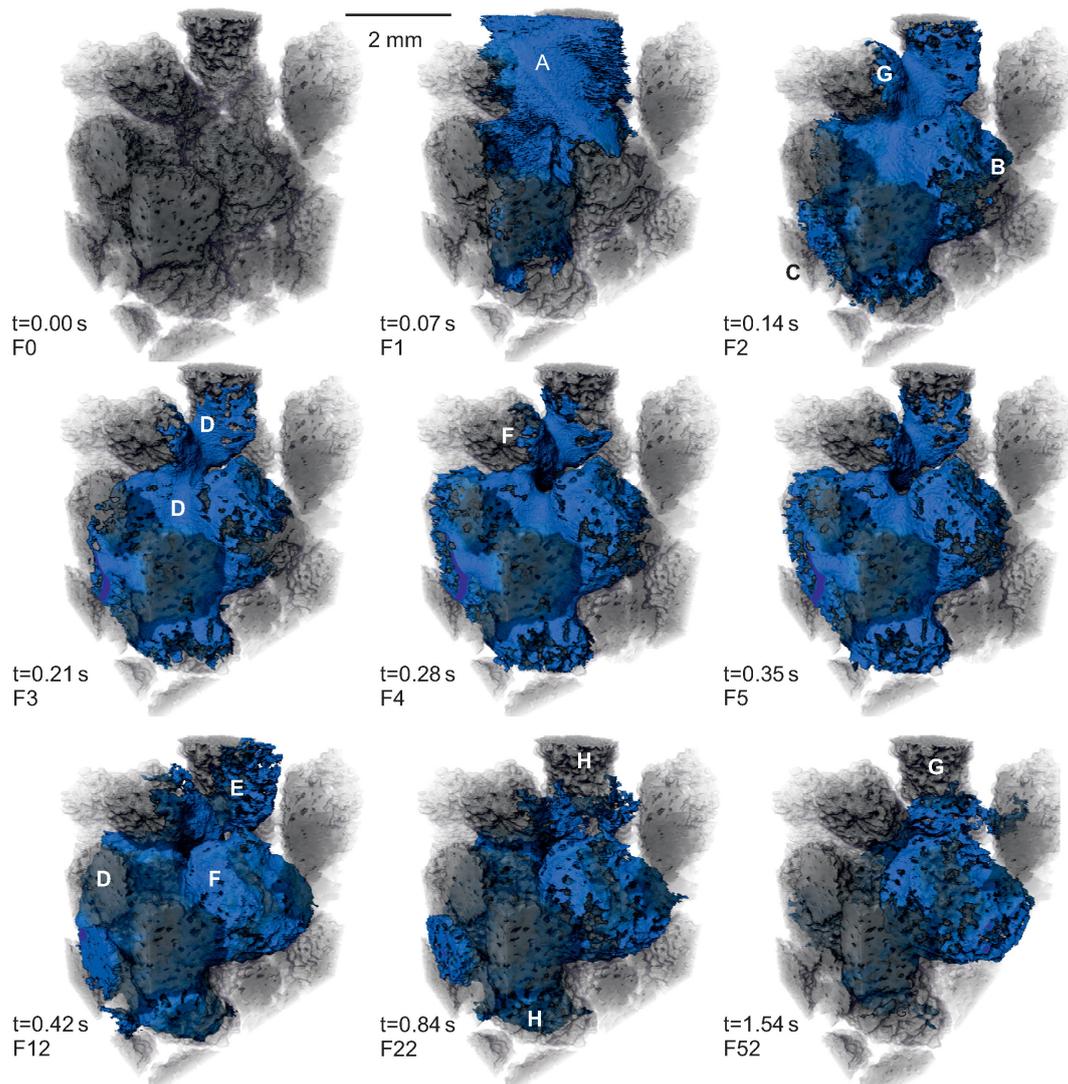


Figure 3. Time series showing the wetting of the sandstone gravel under the first drip. Three-dimensional frame acquisition frequency = $0.07/3\text{-D}$ volume (14 Hz). The first frame shows the dry sample with the sandstone grains rendered in semi-transparent grey. From the second frame the KI solution is shown in blue. Note the rapid changes in both the inter- and intra-grain fluid distribution (see text for discussion). The rendered volume is a small sub-volume of the larger dataset and shows a volume approximately $3.5\text{ mm} \times 3.5\text{ mm} \times 2\text{ mm}$. Data are from sample LH_9, scan 1.

ral scales in one scan or the inclusion of much longer delays between high speed frame acquisitions.

4.1 Case study: capturing multiscale dynamics in pore-scale processes

To show the suitability of this method for capturing the variability of pore-scale flow dynamics, wetting experiments were scanned repeatedly at different gap lengths (continuous through to G25). Between 53 and 106 frames were collected in each scan. High-frequency data acquisition (continuous or G1) was used to capture the first few drips, with longer acquisition intervals (G3–G25) used to capture the slower

processes as the flow network evolved towards steady state through time (Fig. 6). In some experimental runs, variability in the high-speed processes through time was also assessed by repeating continuous G1 or G3 acquisitions at intervals among G5 or G7 acquisitions.

The larger fluid volume in these images for longer durations experiments makes visualisation of the 3-D fluid distribution challenging to show on the page (as in Fig. 3), and so a semi-transparent render of a distance map is used to show the variability in the local fluid volume (Fig. 6). Areas of inter-grain fluid appear as red regions, while the majority of the intra-grain fluid is shown in blue and green. In all scans at all frame acquisition frequencies, we see significant vari-

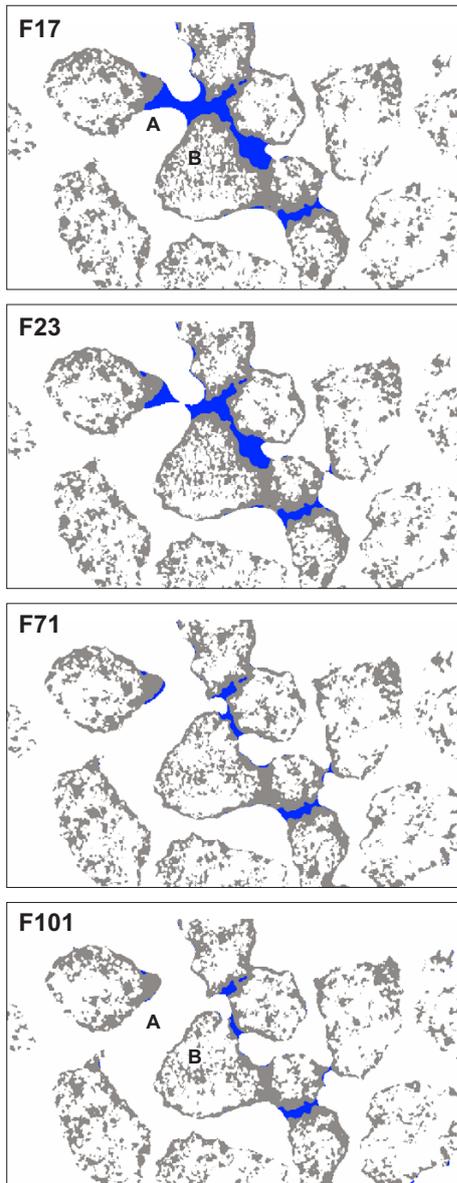


Figure 4. A single vertical 2-D slice through the second frame. The frames shown cover the passage of the through the sample. The inter-grain fluid is shown in blue and the intra-grain fluid in grey. Wet rocks are grey; KI solution outside the rocks is blue; air and dry rocks are in the background. Field of view is $3.5 \text{ mm} \times 2 \text{ mm}$. Data are from sample LH_9, scan 3.

ability in the distribution of the KI solution (Fig. 6) over the 35–60 min experiments.

From the same data it is possible to track the local changes in saturation at the grain scale. Well-established network analysis and morphological quantification algorithms can be applied to extract details of saturation on the pore scale (Fig. 7) or on the geometry of individual phase interfaces and contact angles (Andrew et al., 2014, 2015). For under-sampled data, higher levels of image noise in filtered back-

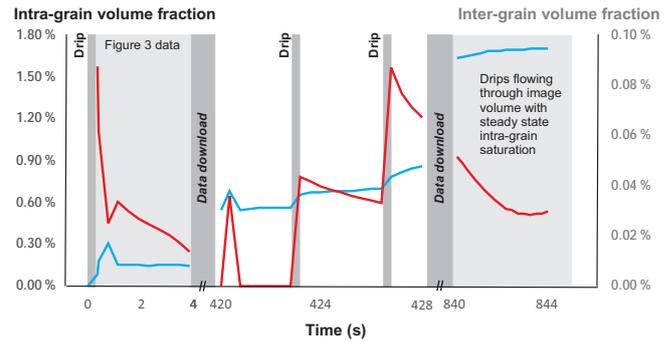


Figure 5. The change in fluid volume within the inter- (red) and intra-grain (blue) porosity over the experiment showing the arrival and subsequent equilibration of two drips. After the first drip the system equilibrates to very low inter-grain saturation. Data are from sample LH_9, scans 1 to 3.

projection reconstructions may limit quantitative surface curvature and contact angle analysis, but optimised iterative reconstruction methods may give more suitable data (see Sect. 5 below).

4.2 Case Study: in situ observation of mixing dynamics

One key application for ultra-high-speed in situ imaging is to track the interaction between multiple phases: capturing both spatial and temporal heterogeneity of replacement, dilution, and chemical reactions. After 30 min under a 6M saline feed the system was flushed with H_2O . Scanning was performed at G1 (0.07 s scan acquisition time, 7 Hz acquisition frequency); the scan started on arrival of the first H_2O drip. Other experimental runs tested replacing H_2O with KI solution and replacing oil with KI solution. The need for ultra-high-speed imaging to capture the complex and spatially variable mixing and dilution processes can be seen in Fig. 8. Dilution of the uppermost part of the local inter-grain network is instantaneous (Fig. 8a), but percolation of the dilution/replacement front through the image volume takes approximately 5 s (Fig. 8a–f). After this time there is little major change in the local greyscale values (Fig. 8g, h), implying a shift in mixing mechanisms. The 2-D slices through the same sub-volume perpendicular to the flow direction (Fig. 8) show the dilution is extremely heterogeneous and can occur on a range of timescales in neighbouring regions. The central pore volume (orange box, F30) is diluted over 10–15 scans, whereas a neighbouring region (connected in 3-D) is replaced in under five frames (yellow box F35). Some grains show variability in the KI concentration within the porosity, implying that, although the bulk porosity of this sandstone is approx. 20 % and the connectivity is high, individual grains have discrete pore networks within them (blue boxes F45). Using slower data acquisition methods would not capture the dynamics of replacement.

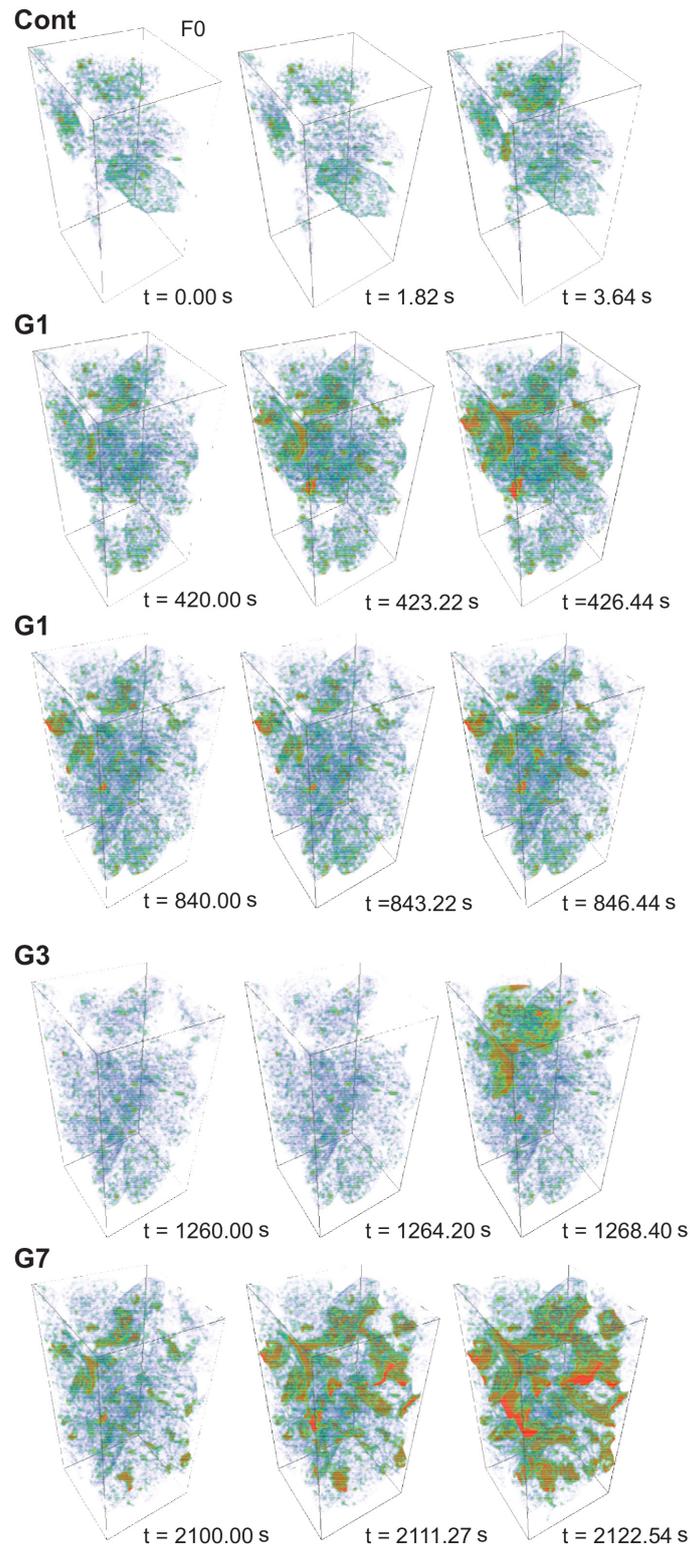


Figure 6. Three-dimensional renders of a gapped experimental run. For each scan series the first, middle, and last frame of each acquisition period is shown. Only the fluid volume has been rendered, with the colour representing the distance to the fluid interface. Cool colours represent small fluid volumes; warm colours represent fluid more than $20\ \mu\text{m}$ from the interface. The field of view is $1.5\ \text{mm} \times 1.5\ \text{mm} \times 3\ \text{mm}$. Data from sample LD_1, scans 1 to 5.

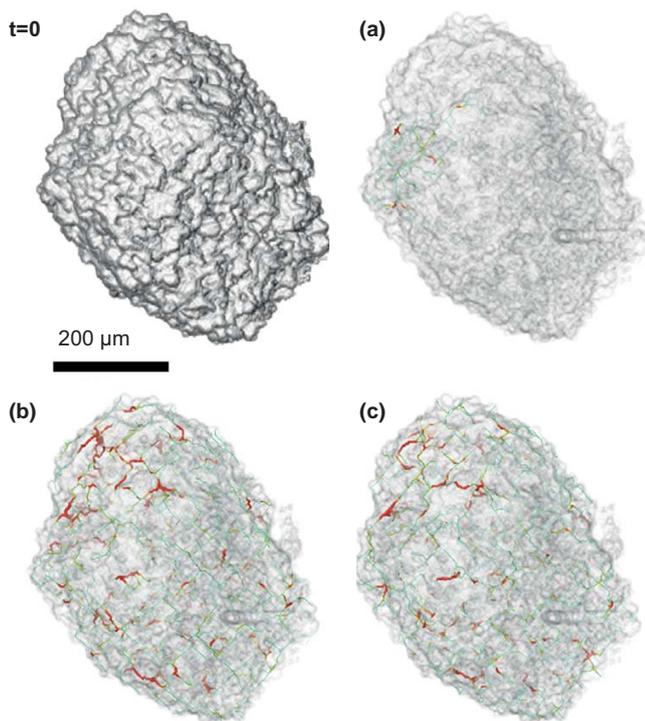


Figure 7. A 3-D render of a single grain showing the progression of saturation. The images show the evolution from first contact with the fluid to a steady saturated state over three scans. The colours and the size of tubes representing the pore network correspond to the local volume of the fluid within the pores: thick red tubes represent the higher local volume; thin blue tubes represent lower local saturation. Data are from sample LD_1, showing the dry scan ($t = 0$) and the final frames of (a) scan 1 (continuous), (b) scan 2 (G1), and (c) scan 3 (G1).

5 Moving beyond the state of the art

The ultra-high-speed imaging presented here allows capture of wetting and transport behaviour at the moderate-to-high spatial resolution necessary for quantitative understanding of the processes involved. The 3-D frame acquisition speed and frequency are substantially faster than has been previously achieved. Events such as Hayne jumps (millisecond timescale) and the subsequent relaxation (a few seconds) have been observed in 2-D imaging (acquisition of projections without rotation) (Armstrong et al., 2014b). The method presented here can provide data on the 3-D nature of those changes, at comparable timescales. Such quantitative data can be used to develop or validate larger-scale flow models and to better incorporate pore-scale processes.

The data presented here are reconstructed using standard filtered back-projection reconstruction algorithms. However, these methods are not optimal for under-sampled data. Recent advances in iterative reconstruction algorithms use information about the different phases contained within the data to enhance the quality of the reconstruction, with great

effect (Fig. 9). Visual comparison of this type of method (Fig. 9b, d) with reconstructions using filtered back projection (Fig. 9a, c) shows the improvement in phase resolution that can be achieved on fewer projection data. This example applies an iterative optimisation solver, and along total variation penalisation is incorporated as a regularisation step (Rudin et al., 1992; Little and Jones, 2010) that takes the sample porosity into account (i.e. uses the dataset themselves as prior information), with refinement at each iteration. The iterative data have fewer line artefacts and lower noise and permit accurate phase segmentation from fewer than 45 projections.

Iterative methods have significant potential to develop imaging dynamic pore-scale processes further still. Of specific interest are those methods that use high-quality dry scan data as a structural prior (Kazantsev et al., 2014, 2015a, b; Van Eyndhoven et al., 2015) and those that divide the data volume into static and dynamic regions using the static information from sequential scans to better define the reconstruction volume (Kazantsev et al., 2014, 2015a, b; Van Eyndhoven et al., 2015). Both allow reconstruction of better-quality images from under-sampled data (data with fewer projections) and also allow data of comparable or better quality to those displayed here to be produced from datasets with as few as 18 projections (Van Eyndhoven et al., 2015). These methods can improve the quality of the quantitative data, especially on the key phase interfaces; however, the key advance is likely to come from working with still higher degrees of under-sampling. However, iterative methods are computationally expensive and can take an order of magnitude longer to reconstruct. The methods applied here took between 15 (Fig. 9d) and 45 (Fig. 9b) min per volume, depending on the number of projections in being used. As such, their use in processing every frame of long-duration high-frame-rate data for near-real-time assessment remains unrealistic at present. This may change as computational efficiency increases, but we envisage that iterative methods are more practically applied retrospectively to achieve higher spatial resolution over short periods of critical interest.

At the 3-D volume acquisition rates achieved here, the projection exposure times and mechanical limit of the rotation stage means that imaging faster is not currently possible for the current samples. For higher-density materials, or higher resolution cameras where longer exposure times mean slower overall volume acquisition, iterative methods may improve temporal resolution. However, on the JEEP experimental setup, the main advantage of iterative methods is the ability to acquire fewer projections per scan and so collect more scans per experiment without sacrificing image quality. This could enable experimental run times to be extended by up to an order of magnitude.

The methods presented here provide a workflow for visualisation and quantification of sub-second dynamics in porous media (Fig. 10). We used the disaggregated sample to capture two spatial scales of fluid-pore interaction, but

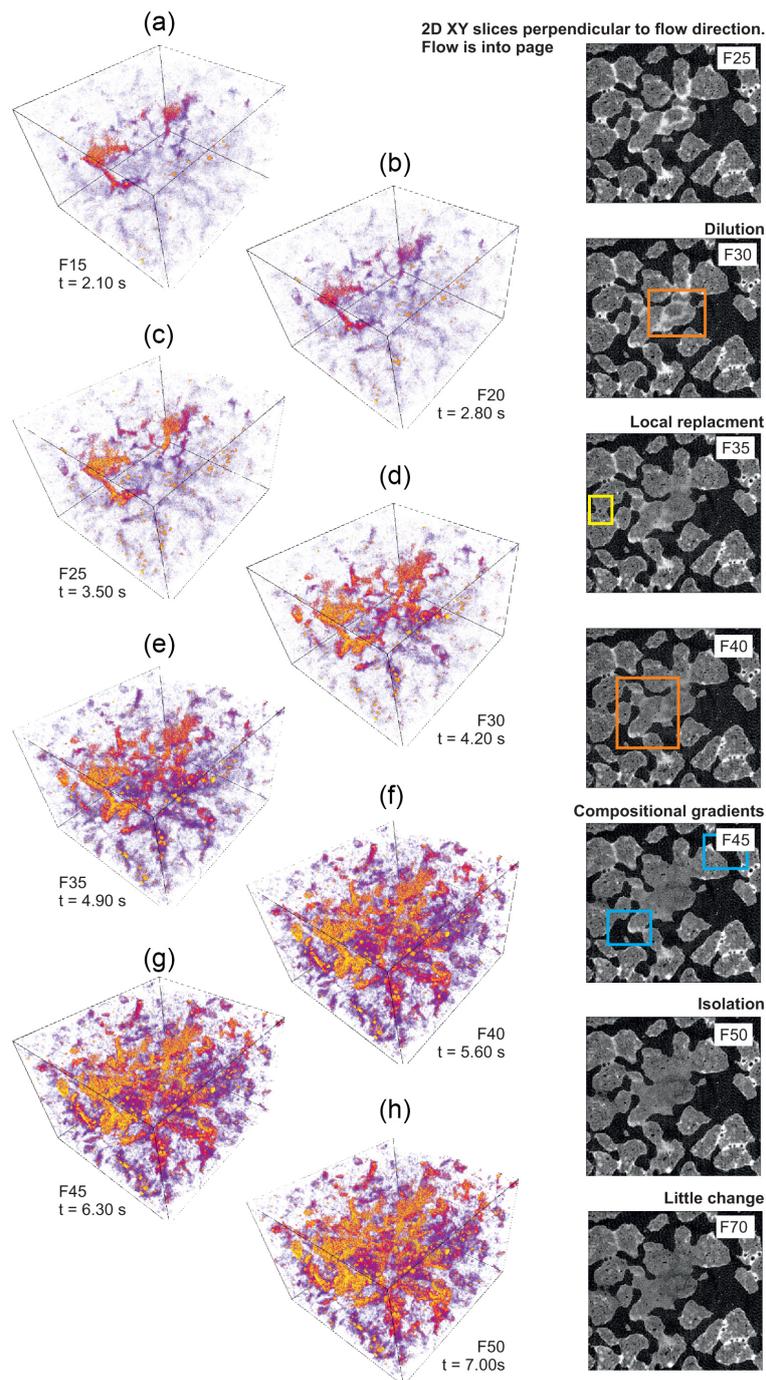


Figure 8. In situ observation of dilution. The 6M saline solution feed is replaced by H_2O at $t = 0$. (a–h) Three-dimensional renders of the differential density (greyscale) between the initial saline fluid volume and the saline + water mixture at the times shown. The frame acquisition rate was 0.07 s and acquisition frequency was 7 Hz (g). Data are from sample LD_3, scan 1. Yellow represents complete replacement with H_2O . Cooler colours represent volumes that are undergoing dilution. Volumes remaining at 6M KI are not shown, as they have undergone no change.

the overall experimental set-up can equally be used for solid core samples at the same spatial and temporal resolution. This experiment was performed at below maximum resolution, and imaging finer structures (grains or pores) at resolu-

tions of $1.2\ \mu\text{m}$ is possible. For higher-resolution or higher-density samples slightly slower projection acquisition times may be needed, and in extreme circumstances this will reduce overall frame acquisition frequency. The ability of the

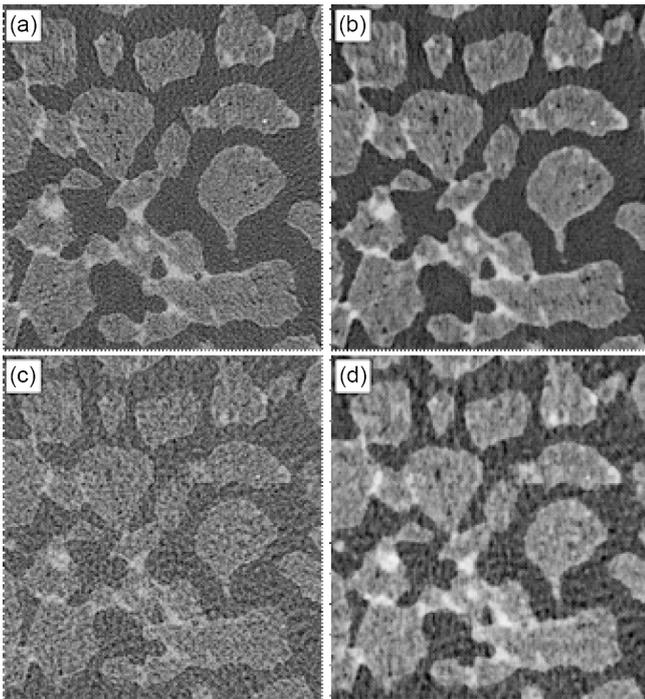


Figure 9. Two-dimensional reconstructed slices through a gravel sample using filtered back projection (**a**, **c**) and an iterative optimisation method with the total variation regularisation incorporated as prior information (**b**, **d**). Reconstructions are obtained using 180 projections (**a**, **b**) and 90 projections (**c**, **d**). The cropped area shown is approximately $3.5 \text{ mm} \times 3.5 \text{ mm}$. Note the increased definition of both the inter- and intra-grain porosity that can be achieved for a given number of projections. Data are from sample LD_3, first frame, scan 1. The optimised iterative methods took approximately 45 min (**b**, 180 projections) and 15 min (**d**, 90 projections) per 3-D frame, making them a viable element of the post-processing workflow for key datasets.

experimental set-up to acquire large numbers of projections at constant angular density means that even higher temporal resolution could be achieved. Reconstruction of volumes from datasets with small rotational offsets (typically about $10\text{--}20^\circ$, i.e. scan 1 reconstruction from 0 to 180° , scan 2 reconstruction from 10 to 190° , etc.) could increase temporal resolution by over an order of magnitude (Van Eynhoven et al., 2015), especially when coupled to iterative methods that can reduce image noise and enhance phase boundary definition in the reconstructed data.

Complex multiphase, spatially heterogeneous micro- and pore-scale processes control many key macro-scale geological system responses. However, the method is not limited to the simple drip-fed experiments on geological materials presented here and can as be easily applied to high-speed processes operating in any material (metals, ceramics, plastics, biological materials, foodstuffs, etc.) suitable for synchrotron X-ray tomography at JEEP/i12. It can also be applied in other dynamic systems for samples and materials

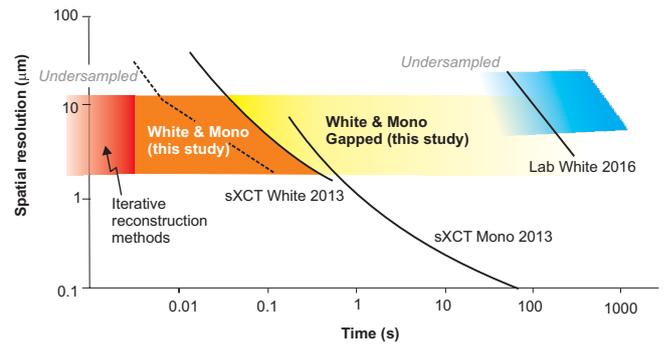


Figure 10. A schematic showing the current state-of-the-art imaging capabilities for synchrotron and laboratory XCT scanning (modified after Bultreys et al., 2016; Maire and Withers, 2014).

undergoing modification through treatment at extreme temperature (hot or cold) or experiencing deformation (tension, compression, cyclic fatigue, fracture, or shear). The development of the real-time 4-D imaging methods we have presented means that quantitative understanding of all of these processes is now possible, in porous media and beyond.

6 Data availability

The projection data and the reconstructed volumes of these experiments are available from the authors.

Appendix A: Post-processing and image analysis

Visualisation and quantification was performed using Avizo™ and MATLAB. The greyscale volumes were cropped to the sample boundaries to reduce all subsequent processing times. When processing continuous acquisition data the frames were first aligned to allow direct comparison and quantitative analysis. For the gapped scans, all data are acquired over the same angle and no additional alignment is required. The italicised steps in the following refer to built-in algorithms in Avizo™, and the reader is directed to the associated documentation for further details of the implementation.

The edge-preserving Bilateral Filter was applied to the greyscale-filtered back-projection reconstructions to reduce image noise. Water, air, and rock phases were segmented using the Interactive Thresholding tool, and the resultant binary data volumes were used in subsequent analysis. The inter- and intra-grain porosity was defined by creating two separate binary masks: one by segmentation of the wet and dry grains and one by segmentation of the fluid (inter- and intra-granular). Logical combination of these allows calculation of the inter- and intra-granular fluid volumes. Fluid volumes were calculated using in-built image measures (Volume3-D).

Additional quantitative analysis of the intra-grain porosity was performed on single grains isolated using the MATLAB command `imfill` followed by Watershed segmentation. The pore network within the grain was displayed using the Centerline Tree module, which applies the TEASAR algorithm (Sato et al., 2000) defining local Euclidian distance to the nearest object boundary. Fluid thickness was calculated by applying a Chamfer 3-D distance map to the fluid label volume to calculate the distance to the nearest surface. The exchange of KI for H₂O was tracked by calculating a differential image between the initial scan and each subsequent frame.

The Supplement related to this article is available online at doi:10.5194/se-7-1059-2016-supplement.

Author contributions. The study was designed by Katherine J. Dobson and Sam A. McDonald; reconstruction and visualisation was led by Sophia B. Coban and Katherine J. Dobson with assistance from Samuel A. McDonald. Robert C. Atwood led camera integration at Diamond Light Source. Katherine J. Dobson, Sophia B. Coban, Sam A. McDonald, Joanna N. Walsh, and Robert C. Atwood performed the experiments at i12/JEEP. The lead author was Katherine J. Dobson, with all other authors contributing to the manuscript.

Acknowledgements. Katherine J. Dobson was supported in this work by EVOKES ERC247076 and NERC NE/M018678/1; Sophia B. Coban is supported by EPSRC EP/J010456/1, the School of Mathematics, and the BP International Centre for Advanced Materials. Philip J. Withers wishes to acknowledge the BP International Centre for Advanced Materials for funding under ICAM 03. Samuel A. McDonald is grateful to Zeiss for funding his fellowship. We thank Diamond Light Source for access to beamline I12 (beam time award EE10500-1).

Edited by: M. Halisch

References

- Alhashmi, Z., Blunt, M. J., and Bijeljic, B.: Predictions of dynamic changes in reaction rates as a consequence of incomplete mixing using pore scale reactive transport modeling on images of porous media, *J. Contam. Hydrol.*, 179, 171–181, 2015.
- Al-Raoush, R. and Papadopoulos, A.: Representative elementary volume analysis of porous media using X-ray computed tomography, *Powder Technol.*, 200, 69–77, 2010.
- Al-Raoush, R. I. and Willson, C. S.: Extraction of physically realistic pore network properties from three-dimensional synchrotron X-ray microtomography images of unconsolidated porous media systems, *J. Hydrol.*, 300, 44–64, 2005.
- Al-Raoush, R., Gordon, C., Robins, S., and Richardson, J.: Characterization of immiscible non-wetting fluids in porous media systems using synchrotron tomography, *Abstracts of Papers of the American Chemical Society*, 241, 1, 2011.
- Andrew, M., Bijeljic, B., and Blunt, M. J.: Pore-by-pore capillary pressure measurements using X-ray microtomography at reservoir conditions: Curvature, snap-off, and remobilization of residual CO₂, *Water Resour. Res.*, 50, 8760–8774, 2014.
- Andrew, M., Menke, H., Blunt, M. J., and Bijeljic, B.: The Imaging of Dynamic Multiphase Fluid Flow Using Synchrotron-Based X-ray Microtomography at Reservoir Conditions, *Transport Porous Med.*, 110, 1–24, 2015.
- Armstrong, R. T., Georgiadis, A., Ott, H., Klemin, D., and Berg, S.: Critical capillary number: desaturation studied with fast X-ray computed microtomography, *Geophys. Res. Lett.*, 41, 55–60, 2014a.
- Armstrong, R. T., Ott, H., Georgiadis, A., Rucker, M., Schwing, A., and Berg, S.: Subsecond pore-scale displacement processes and relaxation dynamics in multiphase flow, *Water Resour. Res.*, 50, 9162–9176, 2014b.
- Batenburg, K. J. and Sijbers, J.: DART: A Practical Reconstruction Algorithm for Discrete Tomography, *IEEE T. Image Process.*, 20, 2542–2553, 2011.
- Berg, S., Ott, H., Klapp, S.A., Schwing, A., Neiteler, R., Brussee, N., Makurat, A., Leu, L., Enzmann, F., Schwarz, J.-O., Kersten, M., Irvine, S., and Stampanoni, M.: Real-time 3-D imaging of Haines jumps in porous media flow, *P. Natl. Acad. Sci.*, 110, 3755–3759, 2013.
- Bhreasail, Á. N., Lee, P. D., O’Sullivan, C., Fenton, C. H., Hamilton, R., Rockett, P., and Connolley, T.: In-Situ Observation of Cracks in Frozen Soil using Synchrotron Tomography, *Permafrost Periglac.*, 23, 170–176, 2012.
- Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., Paluszny, A., and Pentland, C.: Pore-scale imaging and modelling, *Adv. Water Resour.*, 51, 197–216, 2013.
- Boone, M. A., De Kock, T., Bultreys, T., De Schutter, G., Vontobel, P., Van Hoorebeke, L., and Cnudde, V.: 3-D mapping of water in oolitic limestone at atmospheric and vacuum saturation using X-ray micro-CT differential imaging, *Mater. Charact.*, 97, 150–160, 2014.
- Brabant, L., Dierick, M., Pauwels, E., Boone, M. N., and Van Hoorebeke, L.: EDART, a discrete algebraic reconstructing technique for experimental data obtained with high resolution computed tomography, *J. X-Ray Sci. Technol.*, 22, 47–61, 2014.
- Bultreys, T., Boone, M. A., Boone, M. N., De Schryver, T., Masschaele, B., Van Hoorebeke, L., and Cnudde, V.: Fast laboratory based micro-computed tomography for pore-scale research: Illustrative experiments and perspectives on the future, *Adv. Water Resour.*, in press, 2015a.
- Bultreys, T., Van Hoorebeke, L., and Cnudde, V.: Multi-scale, micro-computed tomography-based pore network models to simulate drainage in heterogeneous rocks, *Adv. Water Resour.*, 78, 36–49, 2015b.
- Bultreys, T., De Boever, W., and Cnudde, V.: Imaging and image based fluid transport modeling at the pore scale in geological materials: A practical introduction to the current state-of-the-art, *Earth-Sci. Rev.*, 155, 93–128, 2016.
- Cnudde, V. and Boone, M. N.: High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications, *Earth-Sci. Rev.*, 123, 1–17, 2013.
- Degruyter, W., Burgisser, A., Bachmann, O., and Malaspinas, O.: Synchrotron X-ray microtomography and lattice Boltzmann simulations of gas flow through volcanic pumices, *Geosphere*, 6, 470–481, 2010.
- Dewanckele, J., De Kock, T., Boone, M. A., Cnudde, V., Brabant, L., Boone, M. N., Fronteau, G., Van Hoorebeke, L., and Jacobs, P.: 4-D imaging and quantification of pore structure modifications inside natural building stones by means of high resolution X-ray CT, *Sci. Total Environ.*, 416, 436–448, 2012.
- Drakopoulos, M., Connolley, T., Reinhard, C., Atwood, R., Magdysyuk, O., Vo, N., Hart, M., Connor, L., Humphreys, B., Howell, G., Davies, S., Hill, T., Wilkin, G., Pedersen, U., Foster, A., De Maio, N., Basham, M., Yuan, F., and Wanelik, K.: I12: the Joint Engineering, Environment and Processing (JEEP) beamline

- at Diamond Light Source, *J. Synchrotron Radiat.*, 22, 828–838, 2015.
- Fourie, W., Said, R., Young, P., and Barnes, D. L.: The Simulation of Pore Scale Fluid Flow with Real World Geometries Obtained from X-Ray Computed Tomography, *Proceedings of the Boston COMSOL Conference*, 2007.
- Geraud, Y., Surma, F., and Mazerolle, F.: Porosity and fluid flow characterization of granite by capillary wetting using X-ray computed tomography. Geological Society, London, Special Publications, 215, 95–105, 2003.
- Herring, A. L., Harper, E. J., Andersson, L., Sheppard, A., Bay, B. K., and Wildenschild, D.: Effect of fluid topology on residual nonwetting phase trapping: Implications for geologic CO₂ sequestration, *Adv. Water Resour.*, 62, 47–58, 2013.
- Herring, A. L., Andersson, L., Newell, D. L., Carey, J. W., and Wildenschild, D.: Pore-scale observations of supercritical CO₂ drainage in Bentheimer sandstone by synchrotron X-ray imaging, *Int. J. Greenh. Gas Con.*, 25, 93–101, 2014.
- Hess, K. U., Flaws, A., Muehlbauer, M. J., Schillinger, B., Franz, A., Schulz, M., Calzada, E., Dingwell, D. B., and Bente, K.: Advances in high-resolution neutron computed tomography: Adapted to the earth sciences, *Geosphere*, 7, 1294–1302, 2011.
- Iglauer, S., Paluszny, A., Pentland, C. H., and Blunt, M. J.: Residual CO₂ imaged with X-ray micro-tomography, *Geophys. Res. Lett.*, 38, L21403, doi:10.1029/2011GL049680, 2011.
- Kaestner, A. P., Trtik, P., Zarebandkouki, M., Kazantsev, D., Snehot, M., Dobson, K. J., and Lehmann, E. H.: Recent developments in neutron imaging with applications for porous media research, *Solid Earth Discuss.*, 7, 3481–3510, doi:10.5194/sed-7-3481-2015, 2015.
- Katuwal, S., Norgaard, T., Moldrup, P., Lamande, M., Wildenschild, D., and de Jonge, L. W.: Linking air and water transport in intact soils to macropore characteristics inferred from X-ray computed tomography, *Geoderma*, 237, 9–20, 2015.
- Kazantsev, D., S., O., Hutton, B. F., Dobson, K. J., Kaestner, A. P., Lionheart, W. R. B., J., W. P., Lee, P. D., and Arridge, S. R.: A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction, *Inverse Probl.*, 30, 065004, doi:10.1088/0266-5611/30/6/06500, 2014.
- Kazantsev, D., Thompson, W. M., Van Eyndhoven, G., Dobson, K., Kaestner, A.P., Lionheart, W., Withers, P. J., and Lee, P. D.: 4-D-CT reconstruction with unified spatial-temporal patch-based regularization, *Inverse Probl. Imag.*, 9, 447–467, 2015a.
- Kazantsev, D., Van Eyndhoven, G., Lionheart, W., Withers, P., Dobson, K., McDonald, S., Atwood, R., and Lee, P.: Employing temporal self-similarity across the entire time domain in computed tomography reconstruction, *Philos. T. R. Soc. A*, 373, 20140389, doi:10.1098/rsta.2014.0389, 2015b.
- Lin, Q., Barker, D. J., Dobson, K. J., Lee, P. D., and Neethling, S. J.: Modelling particle scale leach kinetics based on X-ray computed micro-tomography images, *Hydrometallurgy*, 162, 25–36, 2016.
- Little, M. A. and Jones, N. S.: Sparse Bayesian step-filtering for high-throughput analysis of molecular machine dynamics, *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2010 Proceedings Dallas, Texas, USA*, 4162–4165, 2010.
- Ma, L., Taylor, K. G., Lee, P. D., Dobson, K. J., Dowey, P. J., and Courtois, L.: Novel 3-D centimetre-to nano-scale quantification of an organic-rich mudstone: The Carboniferous Bowland Shale, Northern England, *Mar. Petrol. Geol.*, 72, 193–205, 2016.
- Maire, E. and Withers, P. J.: Quantitative X-ray tomography, *Int. Mater. Rev.*, 59, 1–43, 2014.
- Menke, H. P., Bijeljic, B., Andrew, M. G., and Blunt, M. J.: Dynamic Three-Dimensional Pore-Scale Imaging of Reaction in a Carbonate at Reservoir Conditions, *Environ. Sci. Technol.*, 49, 4407–4414, 2015.
- Naveed, M., Hamamoto, S., Kawamoto, K., Sakaki, T., Takahashi, M., Komatsu, T., Moldrup, P., Lamande, M., Wildenschild, D., Prodanovic, M., and de Jonge, L. W.: Correlating Gas Transport Parameters and X-Ray Computed Tomography Measurements in Porous Media, *Soil Sci.*, 178, 60–68, 2013a.
- Naveed, M., Moldrup, P., Arthur, E., Wildenschild, D., Eden, M., Lamand, M., Vogel, H. J., and de Jonge, L. W.: Revealing Soil Structure and Functional Macroporosity along a Clay Gradient Using X-ray Computed Tomography, *Soil Sci. Soc. Am. J.*, 77, 403–411, 2013b.
- Olafuyi, O. A., Sheppard, A. P., Arns, C. H., Sok, R. M., Cinar, Y., Knackstedt, M. A., and Pinczewski, W. V.: Experimental Verification of Effect of Size on Drainage Capillary Pressure Computed from Digitized Tomographic Images, *Int. J. Eng. Res. Africa*, 1, 1–10, 2010.
- Pistone, M., Caricchi, L., Ulmer, P., Reusser, E., and Ardia, P.: Rheology of volatile-bearing crystal mushes: Mobilization vs. viscous death, *Chem. Geol.*, 345, 16–39, 2013.
- Pistone, M., Arzilli, F., Dobson, K. J., Cordonnier, B., Reusser, E., Ulmer, P., Marone, F., Whittington, A. G., Mancini, L., Fife, J. L., and Blundy, J. D.: Gas-driven filter pressing in magmas: Insights into in-situ melt segregation from crystal mushes, *Geology*, 43, 699–702, 2015.
- Raeini, A. Q., Blunt, M. J., and Bijeljic, B.: Direct simulations of two-phase flow on micro-CT images of porous media and up-scaling of pore-scale forces, *Adv. Water Resour.*, 74, 116–126, 2014.
- Raeini, A. Q., Bijeljic, B., and Blunt, M. J.: Modelling capillary trapping using finite-volume simulation of two-phase flow directly on micro-CT images, *Adv. Water Resour.*, 83, 102–110, 2015.
- Rudin, L. I., Osher, S., and Fatemi, E.: Nonlinear total variation based noise removal algorithms, *Physica D*, 60, 259–268, 1992.
- Sakellariou, A., Sawkins, T. J., Senden, T. J., Arns, C. H., Limaye, A., Sheppard, A. P., Sok, R. M., Knackstedt, M. A., Pinczewski, W. V., Berge, L. I., and Øren, P.-E.: Micro-CT Facility For Imaging Reservoir Rocks At Pore Scales, *SEG Technical Program Expanded Abstracts*, 22, 1664–1667, 2003.
- Sato, M., Bitter, I., Bender, M. A., Kaufman, A. E., and Nakajima, M.: TEASAR: tree-structure extraction algorithm for accurate and robust skeletons, *Proceedings of the Eighth Pacific Conference on Computer Graphics and Applications*, 281–449, 2000.
- Sok, R. M., Varslot, T., Ghous, A., Latham, S., Sheppard, A. P., and Knackstedt, M. A.: Pore Scale Characterization of Carbonates at Multiple Scales: Integration of Micro-CT, BSEM, FIBSEM, *PetroPhysics*, 51, 379–387, 2010.
- Van Eyndhoven, G., Batenburg, K. J., Kazantsev, D., Van Nieuwenhove, V., Lee, P. D., Dobson, K. J., and Sijbers, J.: An Iterative CT Reconstruction Algorithm for Fast Fluid Flow Imaging, *Image Processing, IEEE Trans.*, 24, 4446–4458, 2015.

K. J. Dobson et al.: Real-time 4-D imaging of sub-second pore-scale process dynamics**1073**

Wildenschild, D. and Sheppard, A. P.: X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems, *Adv. Water Resour.*, 51, 217–246, 2013.

Wildenschild, D., Hopmans, J. W., Vaz, C. M. P., Rivers, M. L., Rikard, D., and Christensen, B. S. B.: Using X-ray computed tomography in hydrology: systems, resolutions, and limitations, *J. Hydrol.*, 267, 285–297, 2002.

Youssef, S., Oughanem, R., Rosenberg, E., Maire, E., and Mokso, R.: 4-D Imaging of Fluid Flow Dynamic in Natural Porous Media by Ultra-fast X-ray Microtomography, *International Congress on 3-D Materials Science 2014*, 2014.

Appendix A

SophiaBeads Datasets Project Documentation and Tutorials

SophiaBeads Datasets Project Documentation and Tutorials

Sophia Bethany Coban
School of Mathematics, University of Manchester
e-mail: sophia.coban@manchester.ac.uk
website: www.maths.manchester.ac.uk/~scoban

April 1, 2015

The SophiaBeads Datasets [4] are real microCT datasets, acquired specifically for implementing, testing and comparing iterative reconstruction algorithms. The main motivations for the SophiaBeads Datasets Project are providing real datasets for researchers, and introducing a framework for designing experiments and choosing appropriate reconstruction methods via fair comparisons. This aspect of our work is studied in great detail in [5, 6]. The reason we use the SophiaBeads Datasets is because we know what the reconstructions should look like: We know the insides of the sample and its characteristics, so we can quantify the reconstructions and find out how close we are to an ‘exact solution’. The details of the sample and the experiment plan are listed in the next section.

As part of this project, we have also released SophiaBeads Datasets project codes [3]. Our aim with this report is to provide the reader with enough information to work with these codes so the reader can reconstruct the datasets. Additionally, we include a tutorial for quantifying the reconstructions so the readers are able to reproduce our results presented in [5, 6]. In §3, we document the project codes and explain the main script (`sophiaBeads.m`), followed by documentation for the quantification stage using Avizo. §5.1 and §5.2 are the detailed, step-by-step tutorials for the project codes and the Avizo work, where we reproduce our results for SOPHIABEADS_512_AVERAGED as an example. For accessing the contents of this project (datasets and the source codes) and guidance on referencing, we refer the reader to §2. It is also recommended to read the additional notes and the licenses under which the project is distributed, which are explained in §6.

1 Sample Information and Experiment Plan

For the SophiaBeads experiments, we have set up a basic sample with specific measurements so we are able to quantify the reconstructions. The sample is a plastic tube with a diameter of 25mm, filled with uniform soda-lime glass ($\text{SiO}_2 - \text{Na}_2\text{O}$) beads of diameters 2.5 mm (with standard deviation 0.1 mm). The sample has been scanned in the same conditions for each dataset, where the number of projections is halved after each batch scan (starting from 2048, down to 64 projections, see Table 1 for more information). This follows the set up explained

in [5, 6]. This design allows us to understand the effects of fewer projections, and develop algorithms that deliver quality results when the information content is low (e.g. patient scans with lower dose or rapid data acquisitions for 3D+time experiments).

Number of Projections	Number of Frames (scans per batch)	Acquisition Time (per frame)	Size of Dataset
2048	1	24 mins and 10 secs	15.7 GB
1024	2	12 mins and 5 secs	7.8 GB
512	4	6 mins and 3 secs	3.9 GB
256	8	3 mins and 2 secs	1.9 GB
128	16	1 min and 30 secs	953.6 MB
64	32	45 secs	474.9 MB

Table 1: The experiment plan and information about the SophiaBeads Datasets.

The datasets are acquired using the Nikon Custom Bay X-ray CT machine located in the Manchester X-ray Imaging Facility. The reader can visit [8] to find out more about this equipment.

2 Download and Referencing Guide

The SophiaBeads Datasets are accessible via Zenodo, which is an open digital repository aimed at preserving and sharing academic/scientific results (these mainly include datasets or software used in a particular study, or supporting texts for a thesis). The reader can download each dataset as a zipped folder from [4]. When extracting the zipped folders, make a note of the path the folder is extracted to (this becomes important later).

Zenodo provides exportation of various citation formats. To cite the datasets, the reader may export the BibTeX record by clicking on the link on the right hand side of the SophiaBeads Datasets Zenodo page. It is not necessary to refer to each dataset separately; citing (using the DOI) for any of the datasets *once* is sufficient.

The SophiaBeads Datasets project codes are released on GitHub [9] and published on Zenodo [3]. The project codes can be downloaded by visiting either the GitHub repository page (download as zip or tar.gz files), or the Zenodo page (available only as a zip file). By default, Zenodo takes an archive of the SophiaBeads GitHub repository every time there is a new release. However, we recommend direct downloads from the GitHub repository for the most recent changes applied in-between major releases. The reader will still need to export the BibTeX record from the Zenodo page (to get the complete citation including the correct DOI).

The reader will need to download the source project codes in order to work with SophiaBeads Datasets. However, we ask that the reader cites both the datasets and the project codes *separately*, using the relevant DOI.

3 SophiaBeads Datasets Project Codes

As we mentioned earlier, to be able to work with the SophiaBeads Datasets, we have also released project codes for the pre-reconstruction stage. In this section, we introduce these codes, and give detailed descriptions of each script or function¹. The main script in particular, named `sophiaBeads.m`, contains the relevant functions for all the stages up to (and excluding) the quantification. A visual description of this script is given in the figure below, followed by a list of all scripts and functions, and their definitions.

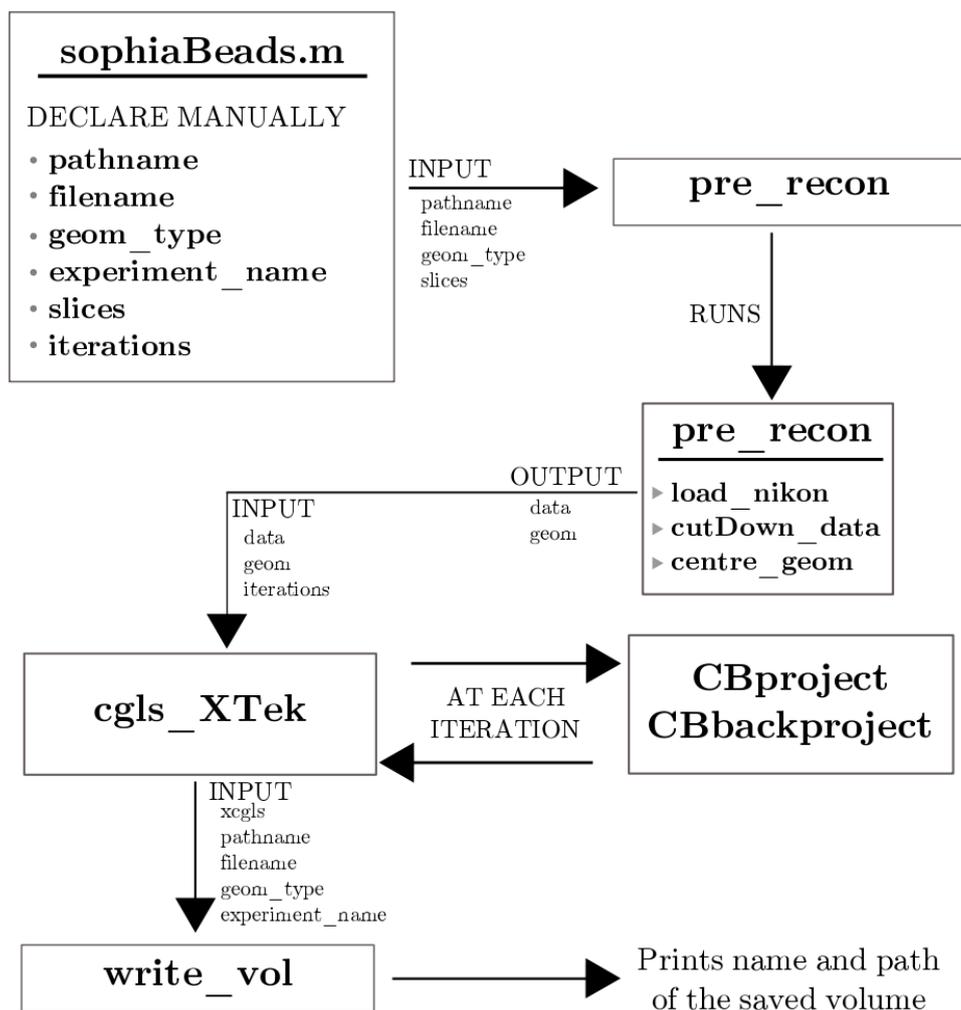


Figure 1: This is the process tree for the script `sophiaBeads.m`.

¹In this report, we assume the reader has some experience using MATLAB, including knowing the difference between a regular script and a function. Nevertheless, we adapt a certain behaviour in the report to distinguish between the two: If the name of a script is mentioned, the name will always be followed by the `.m` extension, e.g. `scriptname.m`; whereas a function will just be `functionname`.

Script/Function	Description
• Main	
<code>sophiaBeads.m</code>	A script with all the commands to carry out the reconstructions. We will be going through this script step by step in the tutorial, in §5.1.
• Pre-reconstruction	
<code>pre_recon</code>	This function is a collection of commands that return the outputs <code>data</code> (the dataset stored in MATLAB) and <code>geom</code> (the cone beam geometry parameters required for forward and back projectors).
<code>load_nikon</code>	Loads the data taken by Nikon XTek CT machine.
<code>cutDown_data</code>	Cuts the dataset down from $2000 \times 2000 \times 2000$ to $1564 \times 1564 \times 2000$. This is the size of the images during the data acquisition for all the SophiaBeads experiments.
<code>centre_geom</code> and <code>find_centre.m</code>	Functions to find the centre of rotation and apply this as a correction to <code>data</code> and <code>geom</code> before the reconstruction stage. See [7] for the correction algorithm implemented.
• C codes	
<code>CBproject.c.c</code> <code>CBbackproject.c.c</code> <code>project_single.c</code> <code>backproject_single.c</code> <code>jacobs_rays.h</code>	These are essentially the forward and back projectors provided by our colleagues W. Thompson [12] and D. Szotten [11]. These codes adapt Jacob's ray tracing algorithm as explained in [10].
• Interface for the mex files	
<code>CBproject</code> , <code>CBbackproject</code>	Interfaces written to connect a reconstruction script with the forward and back projectors.
• mex files	
<code>CBproject.c.mexa64</code> <code>CBbackproject.c.mexa64</code> <code>CBproject.c.mexmaci64</code> <code>CBbackproject.c.mexmaci64</code> <code>CBproject.c.mexw64</code> <code>CBbackproject.c.mexw64</code>	These are the outputs of <code>setup.m</code> , and are used for calculating Ax (forward projector) or A^Tb (back projector).

• Templates

<code>setup.m</code>	This script outputs the mex files for forward and back projectors. This and the next function are provided to us by our colleague N. Wadson [13].
<code>cglstXTek</code>	An example reconstruction script included as a template for the reader. The CGLS algorithm is given in [2].

• Reading/Writing

<code>read_vol.m</code>	Writing and reading reconstructed volumes. We include <code>write_tiff.m</code> as an extra file, which we do not use in the tutorial.
<code>write_vol.m</code>	
<code>write_tiff.m</code>	

The reader is reminded that using the `doc` or the `help` command in MATLAB for a particular script/function will output more details. For example:

```
>> help write_tiff
```

will output

```
write_tiff
```

```
Function to write the reconstructed volume as a set of tiff images.
Each tiff image is a slice in the z-direction.
```

```
INPUT:
```

```
vol: Reconstructed volume.
pathname: Name of the folder where the volume is to be stored.
filename: Name of file to store the volume as.
experiment_name: Name of the experiment for reconstructing this volume.
                 This is to help distinguish between volumes, avoids
                 overwriting. NOTE: This can be an empty string.
voxels: Size of the volume.
type: 'uint8' or 'uint16'. NOTE: This has to be a string.
```

```
DEFAULT VALUES:
```

```
experiment_name = ''; (empty string)
voxels = [1564 1564 2000]; (2000 slices)
type = 'uint16';
```

```
OUTPUT:
```

```
foldername: Name of folder the tiff files are saved in.
```

Copyright (c) 2015 Sophia Bethany Coban
Code is available via the SophiaBeads Datasets project.
University of Manchester.

This is useful for seeing the inputs required by the function and the default values set in the codes.

As mentioned earlier, we have collected all the commands in one script for an easy and quick run (`sophiaBeads.m`) but the readers are encouraged to experiment by running these functions individually.

4 The Quantification Stage

The quantification of the reconstructed results was done using a commercial software, named Avizo. This application is developed by FEI Visualization Sciences Group, and is popularly used in materials science for interactive data visualization and analysis. Avizo is convenient for working on the SophiaBeads Datasets as it includes a list of measures we can apply to quantify our results. Here, we give a list of actions to be used in the tutorial and their descriptions; and a list of measures that can be used to quantify our results. The results presented in [5] are obtained using Avizo Fire 8, but the tutorial is valid for Avizo Fire 7 and Avizo 9. The workflow we used is outlined in Figure 2. In the quantification tutorial, we assume the reader has no experience using Avizo.

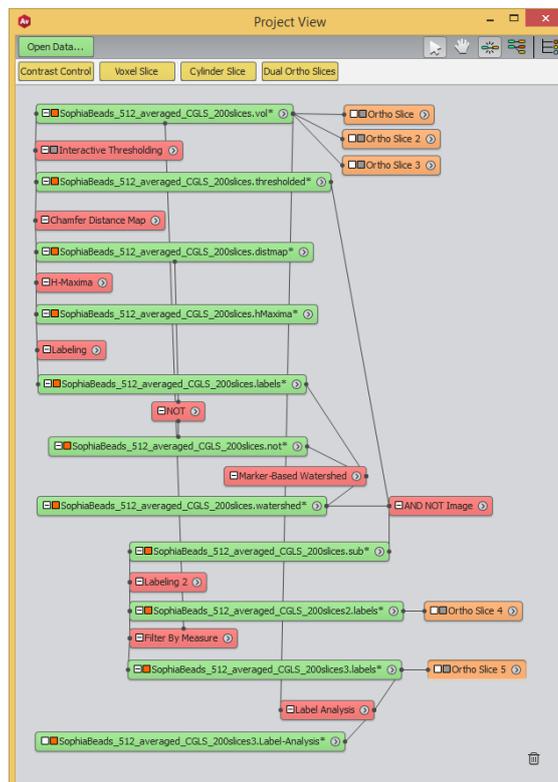


Figure 2: This is the Avizo workflow used in the quantification tutorial. This particular version is Avizo 9. However, the tutorial in §5.2 is valid for Avizo Fire 7 and 8.

The quantification tutorial is essentially the Watershed Tutorial given in [1, pages 271–293],

modified specifically for the SophiaBeads Datasets. We note here that we do not apply any noise reduction or ‘filling in the holes’ during the segmentation stages, as this would influence the analysis of the reconstructed volume. Below is the list of actions we apply, and their descriptions.

Action	Description
Interactive Thresholding	Creates a binary image using a defined thresholding range of intensity values. The thresholding range we used for all the SophiaBeads reconstructed volumes is (0.11,0.35). This was enough to separate the exterior and the beads, even for SOPHIABEADS_64_AVERAGED.
Chamfer Distance Map	Distance maps are applied to binary images (<code>.thresholded</code>). They output grey level images where each intensity value represents the minimal distance from the object boundary, in voxels. We use the Chamfer Distance Map, which is a discrete map. This cuts down the computation time whilst giving reliable results for this type of sample.
H-Maxima	Takes the distance map (<code>.distmap</code>) as an input, calculates the regional maxima, and outputs the most inner regions within the objects. This is necessary for labeling individual objects in the image.
Labeling	This is used to index all of the disconnected/segmented objects within the image. We repeat this action twice during the quantification stage: First time is to distinguish between the exterior and the objects of interest (our reconstructed volume). Second is to label the separated objects (as individual beads).
NOT	Inverts a grey level image (in our case, the input is <code>.distmap</code>).
Marker-Based Watershed	Watershed is the main tool used in the quantification tutorial. The algorithm automatically separates the beads, which are then labeled as individual objects. We use the <i>Marker-Based Watershed</i> algorithm in the tutorials.
AND NOT	This takes two inputs: <code>.thresholded</code> and <code>.watershed</code> , and outputs an image where the separation lines are subtracted from the binary image.
Filter by Measure	Measures individual objects and ranks them according to the chosen measure (see below for the list of measures we believe are suitable for the datasets).
Generate Surface	Generates a 3D surface from a label image (the output for this is <code>.surf</code>).

Surface View	Allows us to visualize the <code>.surf</code> output on the screen.
Ortho Slice	Views the current image as a slice (the horizontal, xy , slice by default, but the user can change this to view vertical slices, xz or yz).
Volumized Rendering	Renders a 3D visualization of the selected volume.

Below is the list of available measures in Avizo Fire, which we can use to quantify the quality of our reconstructed results.

Measure	Description	SophiaBeads Expected Result
SHAPE_VA3D	Outputs a shape factor depending on how spherical the object is. For a perfect sphere, the answer is 1.	SophiaBeads are not all perfect spheres (there is a fraction of the beads that are egg-shaped, which would influence the outcome) so we accept a range between 1 to 1.5 as the “perfect” spherical solution.
EQDIAMETER	Inspects the diametres of the spherical objects.	This should output the diametres of the beads, which (in the perfect case) would be 2.5 mm.
VOLUME3D	Calculates the spherical volume the objects; uses $4\pi r^3/3$ to calculate the volume of a sphere.	We expect the answer for a “perfect” reconstruction of a sphere to be 8.18 mm^3 (or $8.18 \times 10^9 \mu\text{m}^3$), with a standard deviation of 1.02 mm (or $1.02 \times 10^9 \mu\text{m}^3$).
AREA3D	Outputs the surface area of a 3D object.	The surface area of a “perfect bead” in SophiaBeads experiments should be 19.63 mm^2 (or $1.9 \times 10^7 \mu\text{m}^2$), with a standard deviation of 3.27 mm (or $0.33 \times 10^7 \mu\text{m}$).
SYMMETRY	Outputs a factor on how symmetric the objects are.	This should be 1 as we expect the reconstructed beads to be symmetric (egg-shaped beads would also influence this analysis but not by a huge factor, similar to SHAPE_VA3D).
ECCENTRICITY	Similar to the measure above, outputs a factor on how eccentric the shapes of the objects are.	A very low number would mean the shape is not at all eccentric, which is what we would expect for the beads.

There are more ideas listed for quantifying the SophiaBeads datasets in [5, 6], which are not available as measures in Avizo. The readers are welcomed to use these but are also invited to consider new ideas for quantifying own reconstructions.

5 Reconstruction and Quantification Tutorials

This section aids the readers through the process of reproducing our results, as presented in [5], and more importantly helps the reader understand the exact steps taken in our codes. In this section, we will be repeating our reconstruction experiments for the SOPHIABEADS_512_AVERAGED dataset, and present the messages, results, and the Avizo work involved.

Previously, we have mentioned a main script, where all the relevant commands are used to prepare the dataset for reconstructing and then saving in a format appropriate for Avizo (see Figure 1 for the process tree of `sophiaBeads.m`). We will now go through each command, but also print the outputs of `sophiaBeads.m`.

5.1 Reconstruction Using the Project Codes

The scripts are written in a way that everything up to the Avizo stage are automated. However, there are some variables the user will have to manually enter in `sophiaBeads.m` before running the script. These are:

- `pathname`: Declare the path to folder where the dataset is saved.
- `filename`: Declare the name of the dataset. For these runs, this is declared as `SOPHIABEADS_512_AVERAGED`.
- `geom_type`: Declare whether to perform a 2D or a 3D reconstruction. Please note that the variable is case-sensitive, and will not work if, e.g., the variable is declared as `geom_type = '2d'`.
- `experiment_name`: This is for naming purposes only. The reconstructed volume will take the name `<filename>_<experiment_name>.vol`. This variable will help distinguish between the reconstruction experiments, and avoid overwriting a previous result.
- `slices`: This variable is only valid if `geom_type = '3D'`, and will be ignored otherwise. This is to allow the reader to work on a certain number of slices instead of reconstructing the full volume (2000 slices), essentially cutting down the computation time. For our results and quantification in [5], 200 slices were adequate.
- `iterations`: This is an optional variable, and only used in the code we have written for CGLS. When implementing own reconstruction method, reader is free to replace/remove this variable as they see fit.

For this particular run, the variables above are declared as the following:

```
pathname = '/media/SophiaBeads_Datasets/SophiaBeads_512_averaged/';
filename = 'SophiaBeads_512_averaged';
```

```

geom_type = '3D';
experiment_name = 'CGLS_200slices';
slices = 200;
iterations = 12;

```

Running `sophiaBeads.m` will trigger the process outlined in Figure 1. We will now go through this by explaining the commands step by step. The first on our list is `setup`. Note that this is commented out and is only intended to be a template for those wishing to test projector/back projector techniques. If this line is uncommented, the program will output the following:

```

setup;

>> Creating the folder mex/...
>> Building with 'gcc'.
>> MEX completed successfully.
>> Building with 'gcc'.
>> MEX completed successfully.

```

Take caution as this will overwrite the existing projector and back projector mex files. Please read §6 for more information.

Pre-Reconstruction

Next is a function that runs scripts only for the pre-reconstruction phase. This is where the data gets loaded onto MATLAB, the cone beam geometry is set up, and relevant corrections are applied. Below is the list of functions in `pre_recon.m`, followed by how the function is called within `sophiaBeads.m` and their outputs in the command window. Please refer to §3 for the descriptions of these functions.

```

[data,geom] = pre_recon(pathname, filename, geom.type, slices);

>> Loading the SophiaBeads dataset (3D)...
>> Dataset is cut down to 1564 x 1564 x 200...
>> Applying centre of rotation correction...
>> Pre-reconstruction stage is complete!

```

Reconstruction and Saving

We are now ready to reconstruct the SophiaBeads dataset. To encourage readers to implement their own techniques, we include a short CGLS script in the project codes release, named `cglis_XTek`. The readers are free to use this code to obtain the SophiaBeads dataset result in these tutorials, or as a template to implement own techniques. The CGLS algorithm is implemented as described in [2], and run for 12 iterations (this was picked by trial and error, as explained in [5]). The program command and the outputs are

```

xcglis = cglis_XTek(data, geom, iterations);

```

```
>> Reconstructing the SophiaBeads dataset (3D)...
>> Iteration 1 -- Elapsed time is ---- seconds.
>> ...
>> Reconstruction is complete!
```

The code then plots the reconstructed image (for 3D, this is just the central slice).

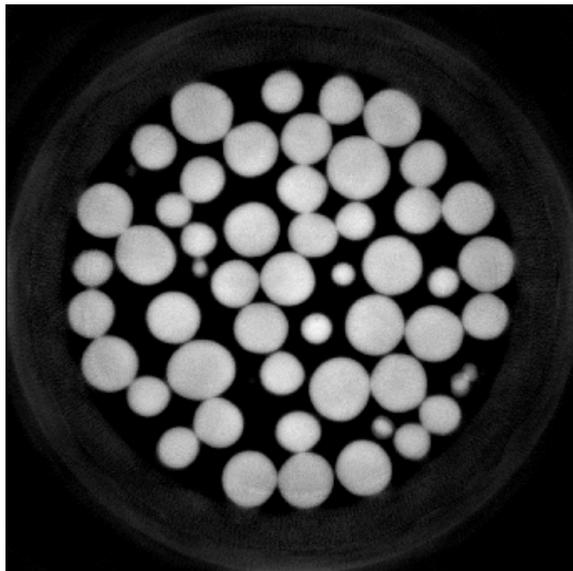


Figure 3: This is the central slice of the reconstructed volume.

The SOPHIABEADS_512_AVERAGED dataset is now reconstructed. The volume is saved as single floats, using the script `write_vol.m`:

```
volname = write_vol(xcglsl, pathname, filename, experiment_name, 'single');
```

which outputs

```
>> The reconstructed volume is written in folder
    /media/SophiaBeads_Datasets/SophiaBeads_512_averaged/.
>> The volume is saved as SophiaBeads_512_averaged_CGLS_200slices.vol.
```

The reconstructed result for this dataset is now ready to be exported to Avizo for the quantification stage.

5.2 Quantification using Avizo

For this section of the tutorials, we have divided the content into three sections:

- I. *Loading the Dataset,*
- II. *Image Segmentation and Separation,* and
- III. *Quantification and Saving Results.*

I. Loading the Dataset

Start the Avizo Fire application and, from the top of the page, select **File** → **Open Data** → **Navigate** to the folder where `write_vol.m` has saved the dataset. The window in Figure 4(a) will appear. Select “*Read complete volume into memory*”, and press OK. The reader will then be prompted to pick the file format. Pick “*Raw Data*” from the list (as highlighted in Figure 4(b)) and press OK.

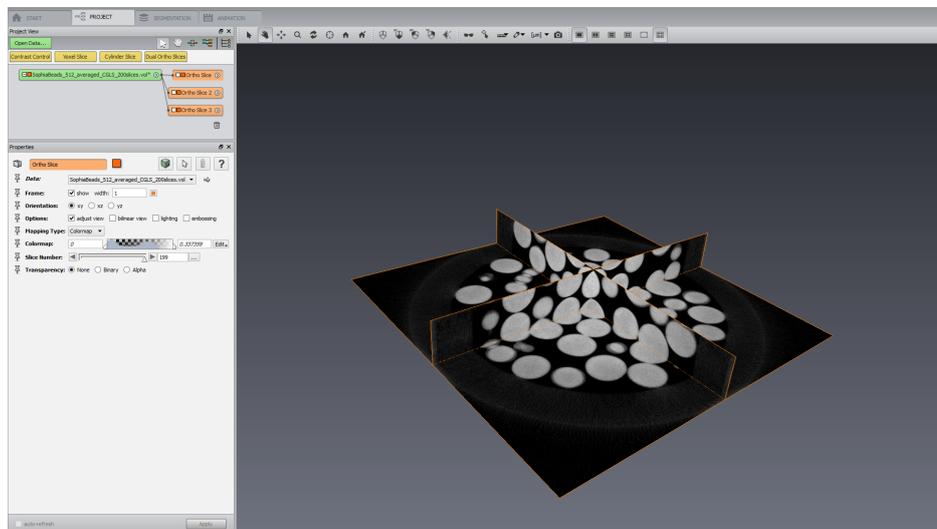
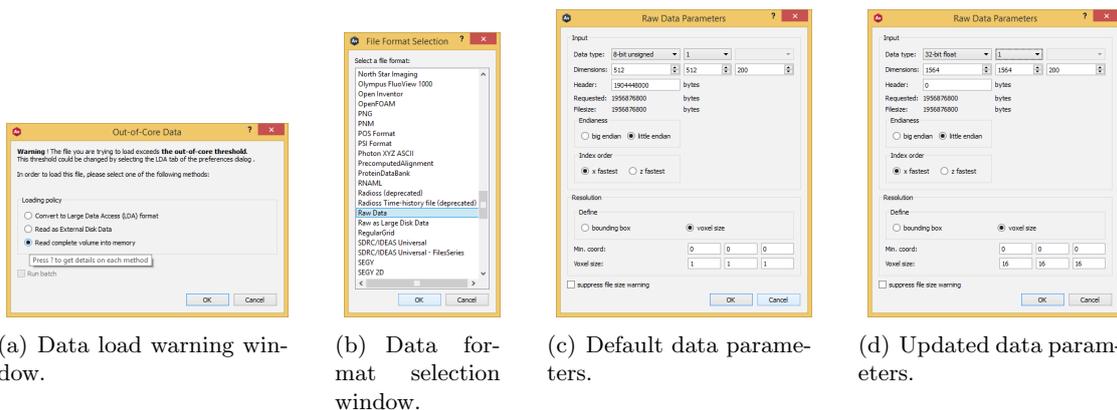


Figure 4: Data loading in Avizo.

The reader will then be asked to enter a set of dataset parameters (see Figure 4(c) for the default values, and Figure 4(d) for the changes applied for the SophiaBeads Datasets). Apply the changes below:

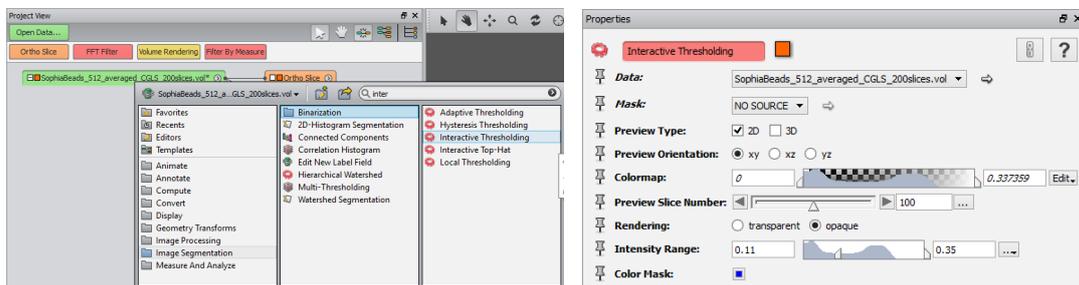
- **Data type:** Select 32-bit float from the drop-down menu.
- **Dimensions:** This should be the size of the reconstructed volume, which in this case is $1564 \times 1564 \times 200$.

- **Voxel size:** This should be as defined in the `.xtekt` files, which (for all SophiaBeads Datasets) is $[16, 16, 16]\mu\text{m}$ (this can be left as $[1, 1, 1]$ but in doing so, the results obtained during the analysis stage would have to be scaled).

Note that after these changes are applied, the **Header** should read 0 (zero). After loading the dataset, one can attach **Ortho Slices** to view the horizontal (default orientation) and vertical slices (choose xz and yz as the *Orientation* option in the **Properties** window below the **Project View**, on the left hand side of the application). The current setup should now match the one given in Figure 4(e).

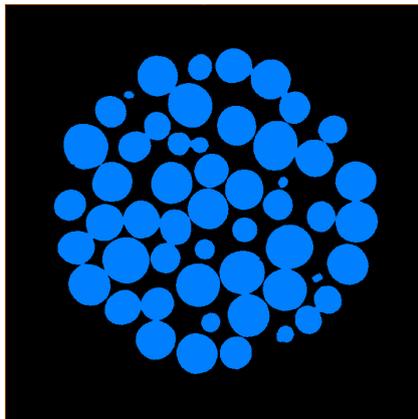
II. Image Segmentation and Separation

The first step is to get the binary image (this is to separate the objects from the background). For this, we apply the action *Interactive Thresholding*: Right click on the dataset (`.vol`) → **Image Segmentation** → **Binarization** → **Interactive Thresholding** → **Create**. See Figure 5(a) as a reference. A red (action) box will appear below `.vol` with options to modify. The only modification needed here is the *Intensity Range*, which should be from 0.11 to 0.35 (in fact, anything above 0.35 will not effect the analysis results). **This range must be the same for all the reconstructed results of SophiaBeads Datasets.** Figures below are of the option window and an ortho slice attached to the output of this action.



(a) Choosing *Interactive Thresholding*.

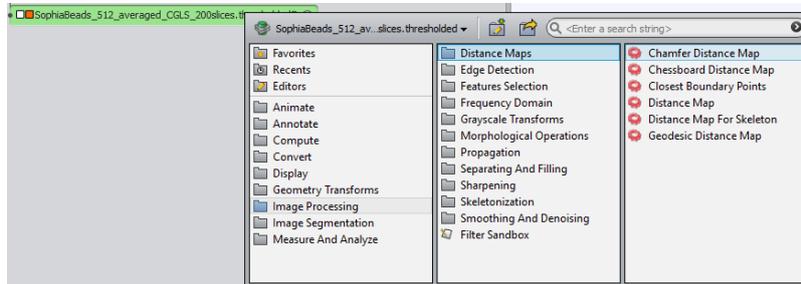
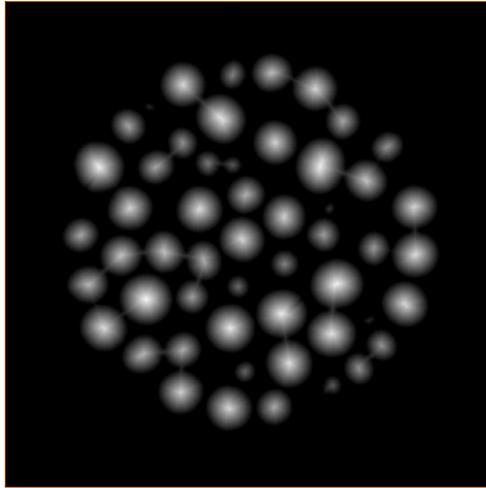
(b) Options.



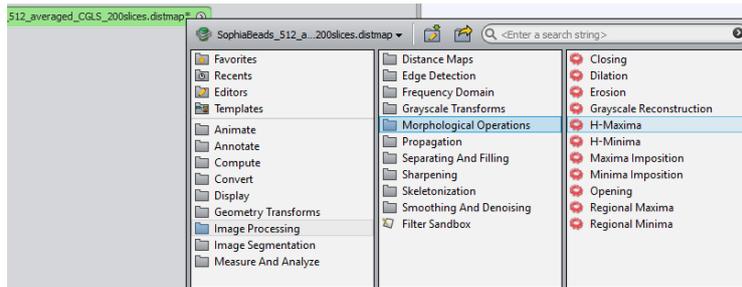
(c) Resulting ortho slice of the binary image (blue = objects of interest, black = exterior).

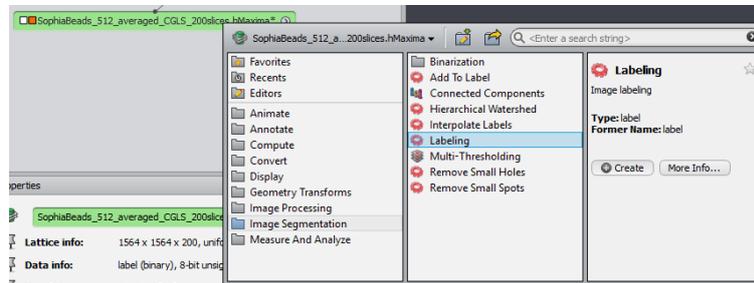
Figure 5: *Interactive Thresholding* steps for a binary image of the volume.

Next, we apply the *Chamfer Distance Map* to the binary image. To do this, right click on `.thresholded` → **Image Processing** → **Distance Maps** → **Chamfer Distance Map** → **Create** (Figure 6(a)). Note that in the **Properties** window, the reader must choose *3D* for the *Interpretation* option and select **Apply**. Attaching an ortho slice to the output gives Figure 6(b). This is followed by the *H-Maxima* action: Right click on the Distance Map output, `.distmap` → **Image Processing** → **Morphological** → **H-Maxima** → **Create**. No changes needed to make in the properties, so click **Apply**. It might take a few minutes to complete the iterations. Finally, we need to attach *Labeling* to the *H-Maxima* output. As before, right click on `.hmaxima` → **Image Segmentation** → **Labeling** → **Create** and then **Apply**.

(a) Choosing *Chamfer Distance Map*.

(b) Distance Map Output.

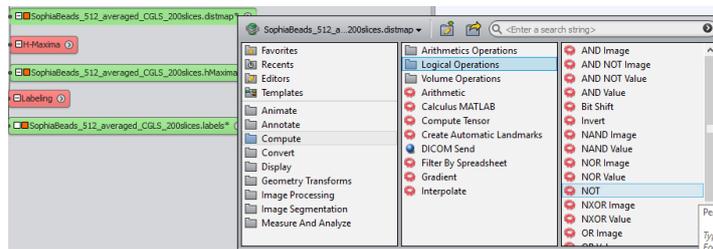
(c) *H-Maxima*.



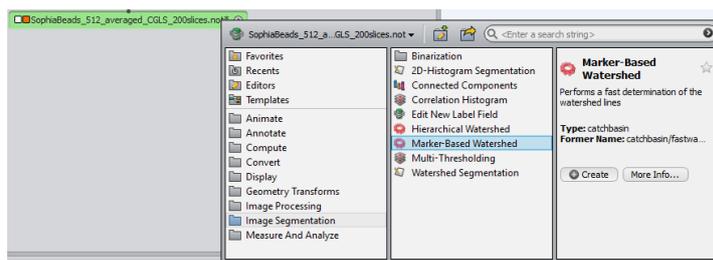
(d) Labeling.

Figure 6: Steps for detecting individual objects.

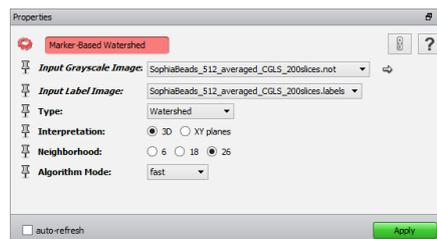
Now, we must go back to the distance map and invert it: Right click `.distmap` → **Compute** → **Logical Operations** → **NOT** → **Create**. Nothing to change in the **Properties** window so click **Apply**. We are now ready to apply the Watershed algorithm on the inverted image: Right click on the output `.not` → **Image Segmentation** → **Marker-Based Watershed** → **Create**. On the **Properties** window, for the *Input Label Image*, select the `.labels` output. See Figure 7(c) for the **Properties** window changes, and Figures 7(d) and 7(e) for the results of `.not` and `.watershed`.



(a) Choosing the NOT option.



(b) Choosing the Watershed algorithm.

(c) Watershed options: Choose *Input Label Image*.

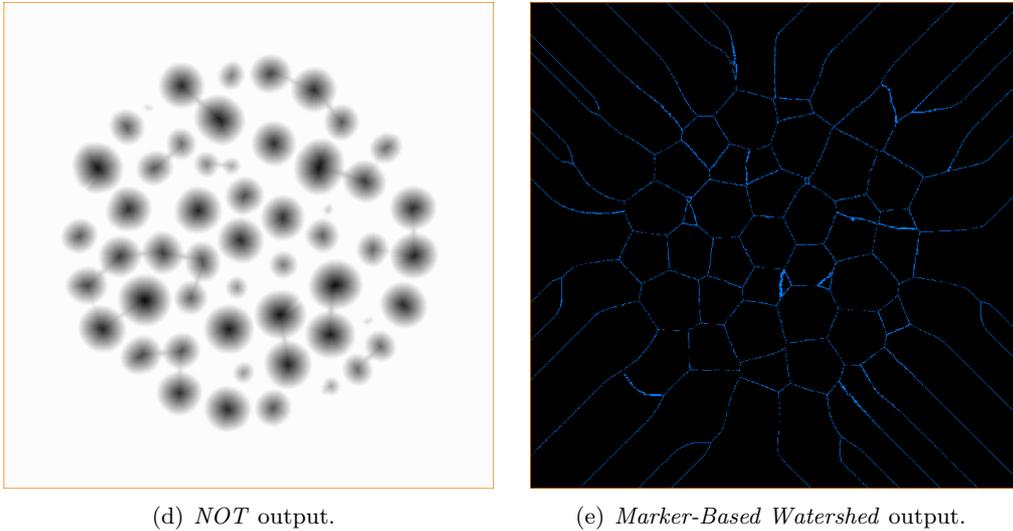
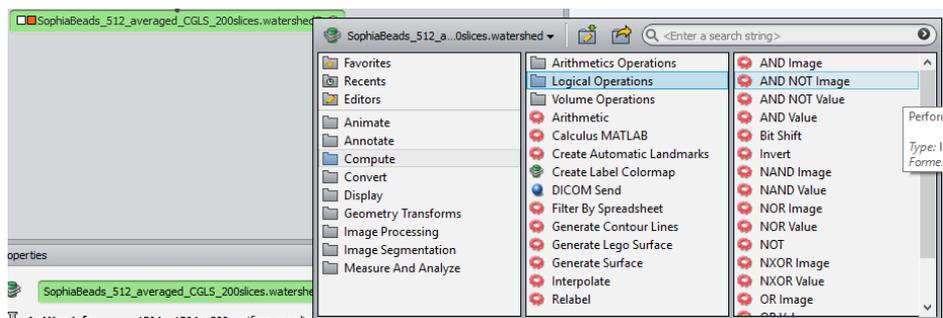
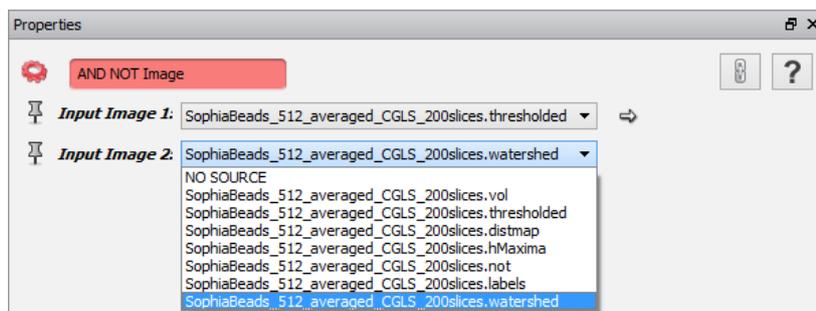


Figure 7: Steps for the Watershed algorithm application.

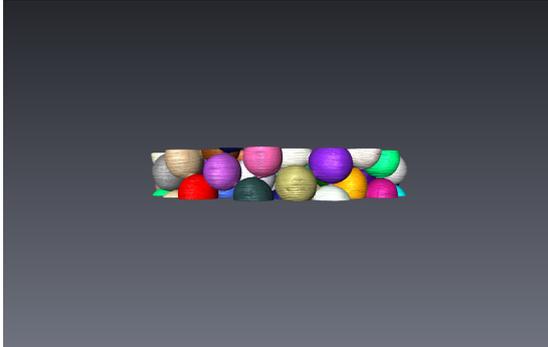
Finally, subtract the separation lines (`.watershed`) from the labeled binary image (`.thresholded`) by right-clicking on `.thresholded` → **Compute** → **Logical Operations** → **AND NOT Image** → **Create**. On the **Properties** window, choose `.watershed` output from the drop-down menu for the *Input Image 2* (see Figure 8(b)). Clicking **Apply** produces the output `.sub`. Repeat the steps for the *Labeling* process to index the separated objects. We can now quantify these objects (reconstructions of the individual beads) using the measures listed in §4.



(a) Choosing the *AND NOT Image*.



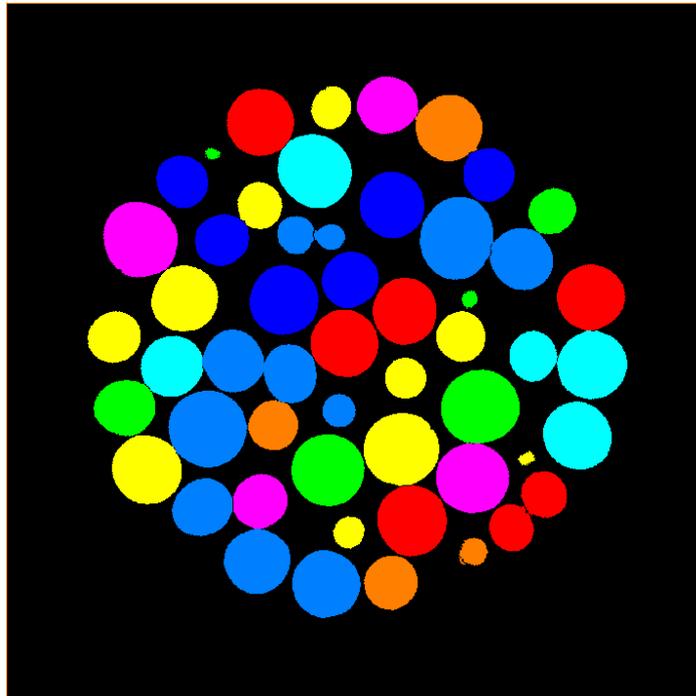
(b) *AND NOT Image* option.



(c) Visualization using *Surface View* of *Labeling 2* (vertical view).



(d) Visualization using *Surface View* of *Labeling 2* (tilted view).



(e) *Labeling 2*; Separated objects.

Figure 8: Final steps for separating the objects.

III. Quantification and Saving Results

We now have our reconstructed beads separated and individually labeled. The second label lists that there are 4329 objects (this is visible in Figure 9(a)). However, this does not mean that there are exactly 4329 beads: this number takes into account any objects of any size within the image. These include any beads that are not fully in the 200 slices window, which may or may not have been separated from other beads (see Figures 8(c) and 8(d) for the *Surface View* of the reconstructed volume). To quantify our reconstruction appropriately, we need to filter out the objects that are too small to be included in the analysis. This is done by *Filter by Measure*: Right click on the second `.labels` → **Measure and Analyze** →

Individual Measures \rightarrow Filter by Measure \rightarrow Create. We believe the most appropriate measure for filtering out is the VOLUME3D option, which sorts the objects by the descending order of their volume. See Figure 9(b) for the modified properties. Note that we pick the first 50 objects that have the highest volume. These are saved in the *third* output of `.labels`.

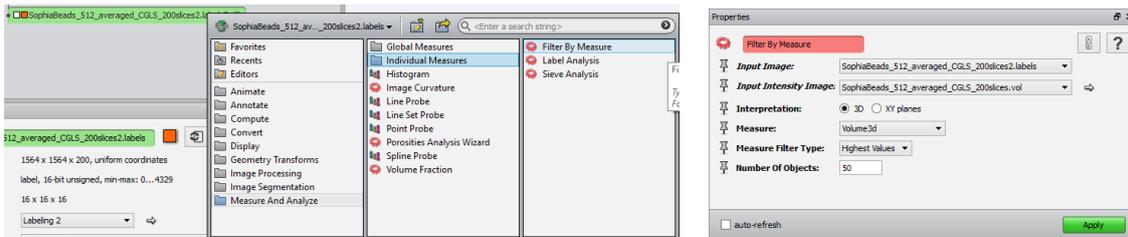
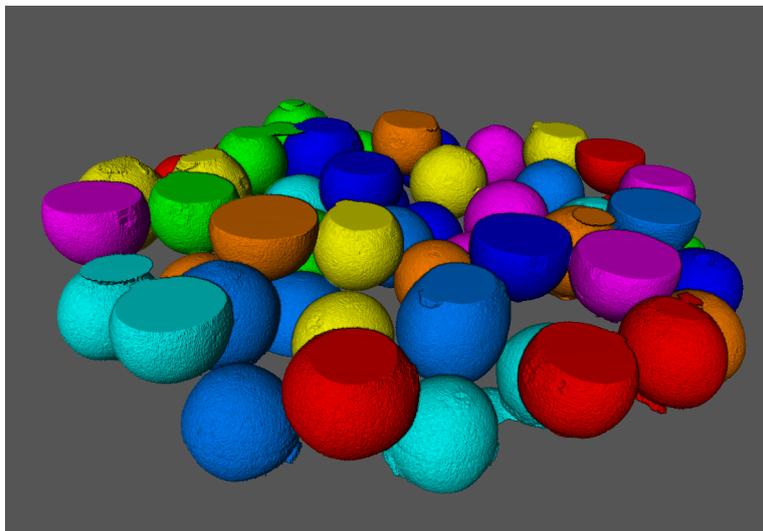
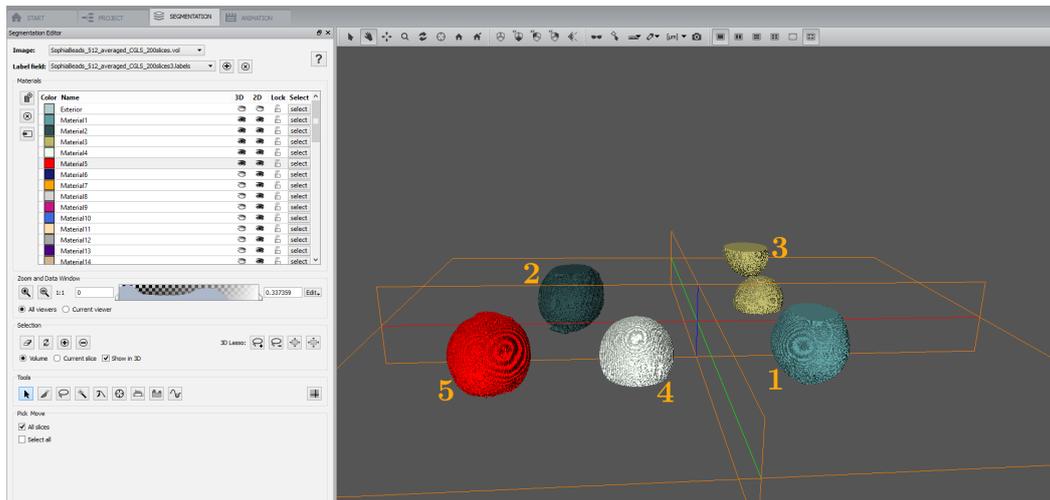
(a) Choosing *Filter by Measure*.(b) **Properties** window.(c) Volumized Rendering of the resulting `.labels`.

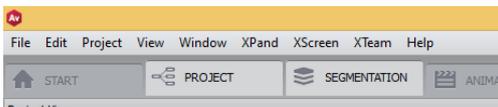
Figure 9: Filtering out the objects that are small in volume.

The next stage involves some manual work. The third output of `.labels` contains 50 objects and these are sorted by the magnitude of their volumes. What we are not taking into account here is that some objects are not separated properly, or are *mostly* within the 200 slices range so they are able to go through the filter (as seen in Figure 9(c)). We need to pick out the beads within these 50 objects that are separated and are fully within the window. To do this, we have to use the **Segmentation Editor** where all objects are listed as individual materials. In Figure 10(a), we see the first 5 objects (viewed by toggling the 3D option). From this we deduce that Material 3 is not separated properly, and Materials 1, 2 and 4 are not fully in the window. For the quantification, we can only accept Material 5 so we delete the first four materials off the list using the *Delete* key (they are simply added to the exterior and are no longer part of the third `.labels` output). Using this method, we inspect all the materials on the list. Our target is to keep around 10 materials for the *Label Analysis* stage (this can be challenging as the quality of reconstructed volume decreases, in which case we would go back to *Filter by Measure*, and increase the *Number Of Objects* option). See Figure 10(c)

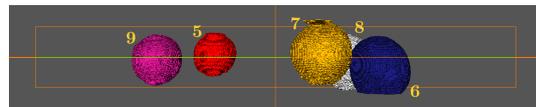
for another example, and Figure 10(d) for the resulting 10 materials we accept for the next stage. Please note that the colours may be different but by following the exact steps, the reader should be able to get the beads in the same locations.



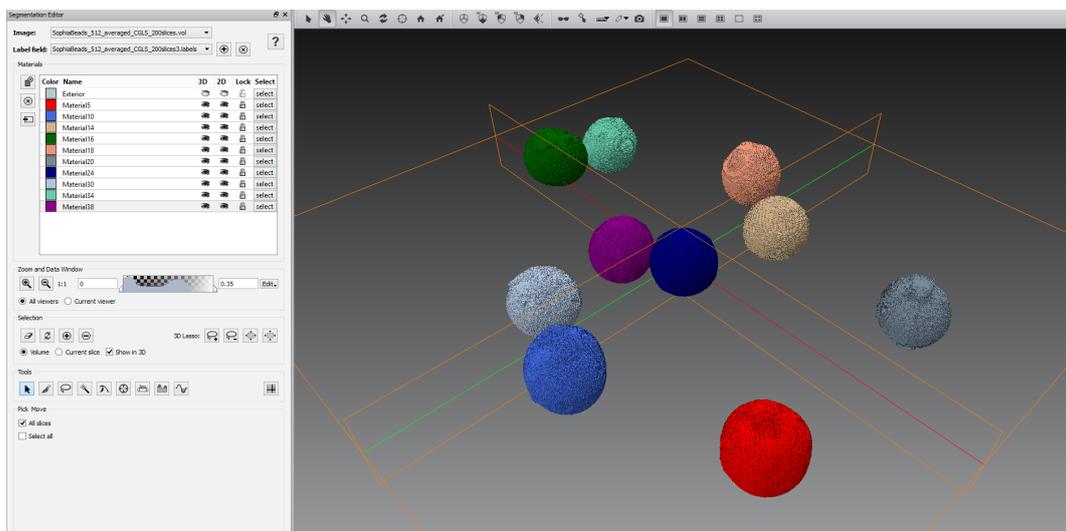
(a) The first 5 materials on the list: Materials 1 to 4 are not suitable.



(b) Segmentation Editor window.



(c) A second example of beads: Materials 5 and 9 are acceptable for the analysis.

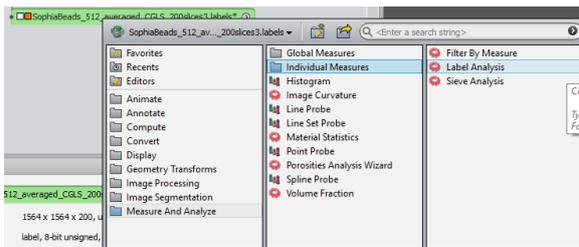
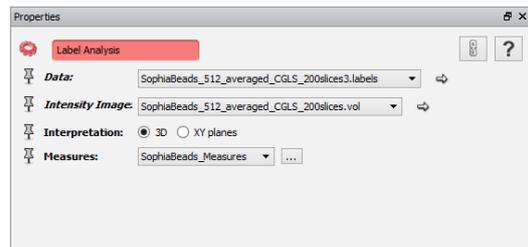
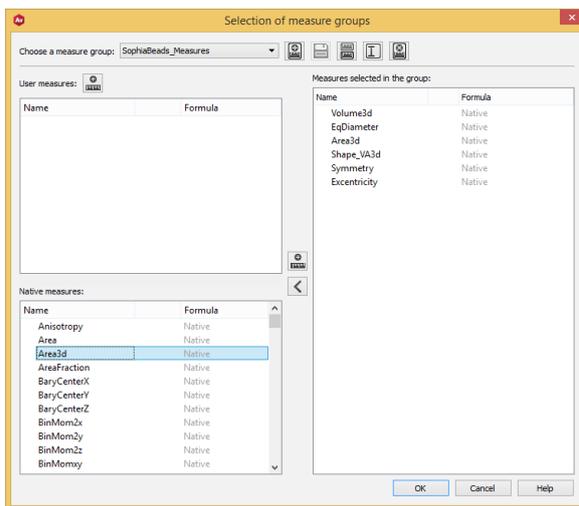


(d) Final look of the appropriate beads in the **Segmentation Editor**.

Figure 10: Manually choosing beads that are separated and fully within the 200 slices.

The final stage is running the *Label Analysis*, which concludes the quantification. To perform the *Label Analysis*, right click on the third `.labels` output → *Measure and Analyze* →

Individual Measures → Label Analysis → Create (Figure 11(a)). In the **Properties** window, choose `.vol` for the *Intensity Image*. For convenience, we created a measure group for the SophiaBeads experiments. This is done by clicking  next to *Measures*. This prompts the window shown in Figure 11(c). To create a new measure group, press  on top of the page, next to the “Choose a measure group option”, and enter a name for a new group. Then, browse the “Native measures” section and double click on the preferred measures to include in the group. The measures picked should appear on the right, under the section “Measures selected in the group”. Press OK when finished, and make sure the newly created measure group is chosen as the *Measures* option in the **Properties** window (compare with Figure 11(b)). Clicking **Apply** will prompt Avizo to measure the remaining objects in `.labels`, and output results in a table for all the measures in the group. The resulting table for our run is given in Figure 11(d). By inspecting these results, we can quantify how close the reconstructed volume is to a perfect solution, and thus compare with different methods applied to the same volume (see e.g. [6, Fig. 7]).

(a) Choosing *Label Analysis*.(b) The final **Properties** window of *Label Analysis*.

(c) Creating measure groups in Avizo.

	Volume3d	EqDiameter	Area3d	Shape_VA3d	index	Materials
Mean	8.79831e+09	2557.25	2.44956e+07	1.70264	5.5	..
Min	7.40989e-09	2418.83	1.93694e+07	1.10405	1	..
Max	1.0982e+10	2757.79	2.88112e+07	2.12112	10	..
Median	8.83652e+09	2564.84	2.47988e+07	1.75334	5	..
Variance	1.14513e+18	10414.7	9.04389e+12	0.120246	8.25	..
Kurtosis	..	-0.602641	-0.968621	-1.04148	-1.22425	..
Skewness	0.593039	0.468217	-0.274873	-0.516507	0	..
	Volume3d	EqDiameter	Area3d	Shape_VA3d	index	Materials
1	1.0982e+10	2757.79	2.88112e+07	1.75334	1	Material5
2	1.00455e+10	2677.06	2.83198e+07	1.99011	2	Material...
3	9.29738e+09	2608.88	2.39781e+07	1.41017	3	Material...
4	9.2375e+09	2603.27	2.55478e+07	1.72781	4	Material...
5	8.9129e+09	2572.41	2.67107e+07	2.12112	5	Material...
6	8.47617e+09	2529.69	2.57804e+07	2.10872	6	Material...
7	8.0601e+09	2487.6	2.00928e+07	1.10405	7	Material...
8	7.89253e+09	2470.24	2.29538e+07	1.71663	8	Material...
9	7.66908e+09	2446.7	2.33918e+07	1.92421	9	Material...
10	7.40989e+09	2418.83	1.93694e+07	1.17024	10	Material...

(d) The final table with label analysis results listed.

Figure 11: Final steps of quantifying the quality of a reconstructed volume using Avizo.

It is possible to visualize the results in 3D by applying *Volumized Rendering* to the final `.labels` output, or by applying *Generate Surface* to the same output, followed by *Surface View* (the outputs will be similar to Figures 9(c) or 8(d)). The tree flow in the **Project**

View window should match the one given in Figure 2. The reader can save the results on the analysis table in a csv or xml format.

6 Additional Notes and Licensing Information

Due to the nature of the sample, the datasets are also suitable for developing and/or testing

- segmentation methods,
- image or data correction techniques,
- forward and back project implementations, and
- benchmarking own codes or method.

We include mex files suitable for Windows, Mac OSX and Linux operating systems. We also include the Linux template of the script we have written to create these mex files (`setup.m`). Please note that as this is a template script, it may not work with Windows, and may need some tweaking to work for Mac OSX systems (this depends on the installed versions of MATLAB and Xcode).

If the reader runs `setup.m`, the existing projector and back projector mex files will be overwritten. To avoid this, modify the names after the **-output** option in the `setup.m`:

```
Line 22: -output mex/CBbackproject_c
Line 23: -output mex/CBproject_c
```

Please note that in doing so, the reader will also have to modify the interface scripts for the mex files (`CBproject.m` and `CBbackproject.m`) to update the lines with the new projector and back projector mex file names.

Finally, the project codes are tested using the following MATLAB versions:

- **Scientific Linux 6, Ubuntu 12.04 and 14.04:** MATLAB R2010a, R2013a, 2014b.
- **Mac OSX 10.9 (Mavericks) and 10.10 (Yosemite):** MATLAB R2009b, R2013a, R2014b.
- **Windows 7 and 8.1:** MATLAB R2012b, R2013a, R2014a, R2014b.

Licensing Information

The SophiaBeads Datasets are distributed under the CREATIVE COMMONS ATTRIBUTION SHARE-ALIKE license. The project codes are distributed under the GPL v2 and MIT licenses. Please read the documents `LICENSE` and `LICENSE2` to learn more about using the project codes. The reader may also find it useful to visit www.creativecommons.org/licenses/ and www.choosealicense.com to learn about these licenses.

Acknowledgments

This project is funded by the School of Mathematics, EPSRC CCPi (EP/J010456/1) and BP through the ICAM framework.

The author is grateful to the MXIF staff Dr. Sam McDonald and Dr. Julia Behnsen for their help with acquiring the datasets and with visualizing the results. The author is also grateful to the PhD supervisors Prof. Bill Lionheart and Prof. Phil Withers for their helpful comments and guidance.

References

- [1] Avizo 8. *Avizo User's Guide*. English. FEI Visualization Sciences Group.
- [2] Å. Björck, T. Elfving, and Z. Strakos. "Stability of conjugate gradient and Lanczos methods for linear least squares problems". In: *SIAM J. Matrix Anal. Appl.* 19(3) (1998), pp. 720–736.
- [3] S.B. Coban. *SophiaBeads Datasets Project Codes*. Apr. 2015. DOI: 10.5281/zenodo.16539. URL: <http://dx.doi.org/10.5281/zenodo.16539>.
- [4] S.B. Coban and S.A. McDonald. *SophiaBeads Datasets*. Mar. 2015. DOI: 10.5281/zenodo.16474. URL: <http://dx.doi.org/10.5281/zenodo.16474>.
- [5] S.B. Coban et al. "A framework for exploring the limits of reconstruction algorithms". In: (*to appear*) (2015).
- [6] S.B. Coban et al. "When do iterative reconstruction methods become worth the effort?" In: *the proceedings of The 13th International Meeting of Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. 2015.
- [7] T. Liu. "Direct central ray determination in computed microtomography". In: *Optical Engineering* 48.4 (2009), pp. 046501–046501.
- [8] *Manchester X-Ray Imaging Facility – Nikon Custom Bay*. <http://www.mxif.manchester.ac.uk/resources/imaging-systems/nikon-custom-bay>. Accessed: 20-MAR-2015.
- [9] *SophiaBeads Datasets Project Codes – GitHub Page*. <http://sophilyplum.github.io/sophiabeads-datasets/>. Accessed: 01-APR-2015.
- [10] E. Sundermann et al. "A fast algorithm to calculate the exact radiological path through a pixel or voxel space". In: (1998).
- [11] D. Szotten. "Limited Data Problems in X-ray and Polarized Light Tomography". PhD thesis. School of Mathematics, University of Manchester, 2011.
- [12] W.M. Thompson. "Source firing patterns and reconstruction algorithms for a switched source, offset detector CT machine". PhD thesis. School of Mathematics, University of Manchester, 2011.
- [13] N. Wadson. "Aluminium foam data reconstruction using CGLS and TV Regularisation - 100 and 200 projection data." MIMS Preprint. 2012.

Appendix B

When do the iterative reconstruction methods become worth the effort?

When do the iterative reconstruction methods become worth the effort?

S.B. Coban*, P.J. Withers[†], W.R.B. Lionheart*, S.A. McDonald[†]

Abstract—A driving force for the development of new reconstruction algorithms is to achieve better quality images using less information (lower dose, fewer projections, in less time), but under what circumstances do iterative methods become worth the effort? In this paper we propose a framework that enables the performance of reconstruction algorithms to be mapped. Such a framework allows fair comparisons to be made, providing insights into experimental acquisition strategies and methods of quantifying the quality of reconstructions, and identifying the sweet spot for different algorithms.

In the CT imaging community, the challenge is to be able to produce the best quality images with the least amount of information. Depending on the application, this *information* could be a series of quickly acquired projections if we wish to capture rapid changes in the sample, or low dose exposures given to a patient during the scans; or it could be a limited angles of illumination due to the constraints of the hardware, the sample, or restrictions to minimize the computational memory requirements.

The science of reconstructing a 3D volume from 2D projections is a problem with many possible solutions. Even in the case of sufficient data, the solutions can be unstable. Additionally, these solutions are sensitive to small changes in the measured data (noise due to modeling or experimental errors), which means that even a noisy image can qualify as a feasible solution. This solution can be an image of the location of a landmine in the ground, or detection of a weapon in a bag, or used to diagnose cancerous cells. This becomes an issue when the reconstruction techniques do not converge to a better solution than a noisy image. Therefore it is natural that we consider the data to be measured, any prior information about the problem we can make use of, and the methods for reconstructing high quality images. However before thinking about the reconstruction algorithms, we must first consider whether the information we have is useful. This can be done by characterizing the *cost of measure*. This is simply a variable specific to the application. In our case, this could be the available acquisition time, or perhaps the level of dose we use to scan a patient. It is clear that, whatever the application is, we do not want a dataset comprising just one projection acquired over a long period of time (strong signal), or many projections in very short periods (weak

signal). We want to be able to determine what we need to know, and plan experiments accordingly. Taking into account the important variables in CT, we propose a map to guide the reader when planning experiments. This map displays aspects of the information content of the dataset with the abscissa quantifying the number of projections and the ordinate the number of photons collected (proportional to the number of frames acquired per projection times the number of projections). In this space a given experimental strategy is a point on the map. If the quality of the reconstructed image can be expressed as a metric, then it can be used to map the capability of a given algorithm highlight acquisition regimes over which it maybe deemed acceptable. Similarly regimes over which different algorithms are beneficial can be identified. Conversely, acquisition strategies can be identified to achieve a given performance in the least time or dose.

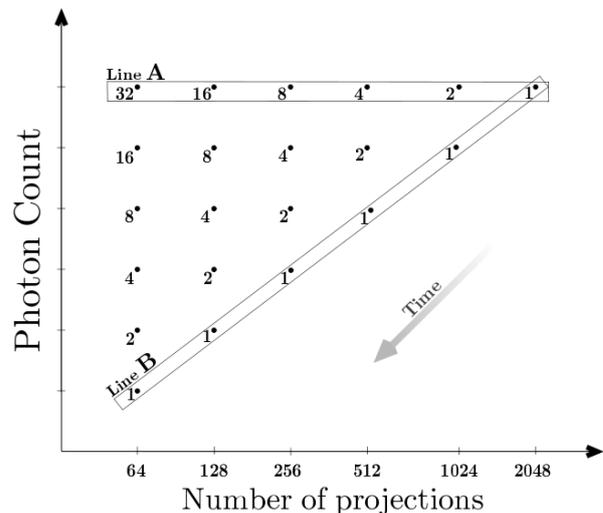


Fig. 1. The log-plot of number of projections vs photon count (number of projections \times number of frames). The numbers on the chart refer to the number of frames for scans lying on LINES A and B.

From this map it is clear that if we want to identify the performance of algorithms as a function of the number of projections, it is fairer to keep the number of photons collected constant while the number of projections is varied (LINE A in Fig. 1). By contrast it is often the case that in practice the comparison is made using one frame for each projection (LINE B) which convolves the decreasing number of projections with the decreasing signal, influencing our conclusions. The

*School of Mathematics, [†]School of Materials,
The University of Manchester, Manchester, M13 9PL, United Kingdom.
sophia.coban@manchester.ac.uk, www.maths.manchester.ac.uk/~scoban;
sam.mcdonald@manchester.ac.uk;
p.j.withers@manchester.ac.uk;
bill.lionheart@manchester.ac.uk, www.maths.manchester.ac.uk/~bl.

difference in reconstructed images for both LINES A and B is illustrated in Fig. 2(c), where we have used the SophiaBeads 128 dataset with 1 frame (lying on LINE B) and 16 frames (on LINE A). We introduce the SophiaBeads Datasets in the next section.

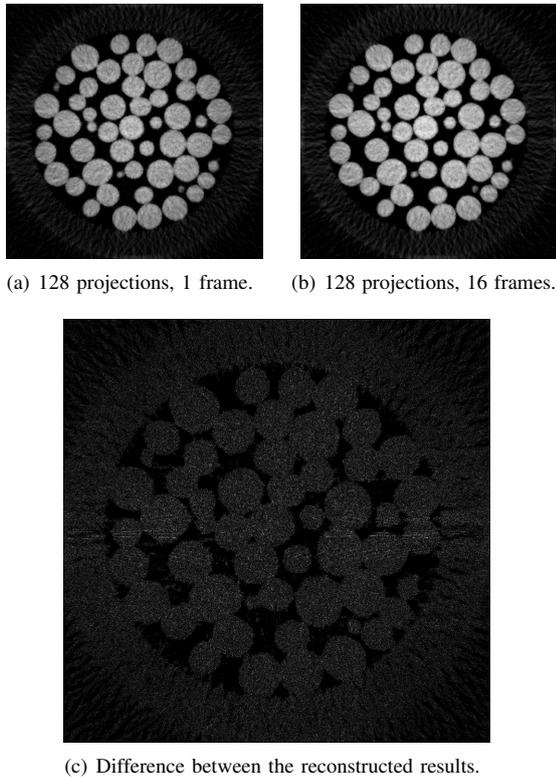


Fig. 2. Highlighting the difference in reconstructions with the traditional approach following LINE B (top left), and keeping the cost of measure constant by following LINE A (top right). The eliminated noise is shown in the image (bottom). Reconstructions are obtained using CGLS.

Our ultimate aim is to introduce a framework for designing experiments and choosing appropriate reconstruction methods via fair comparisons. In this paper, we wish to discuss this aspect of our work, and support our logic with reconstructed results. Before concluding the paper, we explore ways of quantifying results using a real dataset.

I. EXPERIMENTAL DESIGN AND QUANTIFICATION METHODS

We have established an experimental glass bead pack dataset, [5], based on the above framework acquiring 1 frame for each of 2048 projections; 2 frames at 1024, 4 frames at 512, 8 frames at 256, 16 frames at 128 and 32 frames for 64 projections (see points lying on LINE A in Fig. 1). This enables a wide range of algorithm comparisons and information content optimizations to be examined. In this paper we examine the performance of algorithms along LINE A, namely we compare the performance of algorithms using different numbers of frames but at a constant signal.

The experiment dataset is called SophiaBeads, available as part of the SophiaBeads Datasets Project. More information

on the sample, data acquisition and quantitative analysis of the reconstructions can be found in [6, 4]. We have chosen a beads problem because it is easy to make the dimensions of the solid spheres precisely known and the problem is representative of many X-ray imaging problems [7]. A key element of the beads problems is that the samples often consist of only beads and air, making them suitable for studying ‘porous channels’ (bottlenecks) and ‘touching of beads’. These are important characterizations for studying segmentation techniques or discrete CT algorithms.

Our motivation for using SophiaBeads Datasets in particular is that we know what the reconstructions should look like. We know that the beads are of one size¹, and thus the following items can be considered when quantifying our results:

- I. **Volume of the beads²**: The range of expected volume of each bead is known.
- II. **Shape of the beads²**: We can parameterize how close (in shape) a reconstructed bead is to a perfect sphere.
- III. **Circular cross-section of the beads³**: Because the beads are (nearly) perfectly spherical, each bead should have the same¹ diameters or radii in all axes (i.e. a cross-section of each bead should be a perfect circle).
- VI. **Smoothness of the beads³**: boundary of each bead should be smooth.

II. COMPARISON OF RECONSTRUCTION METHODS

In this section, we present the 2D reconstructions of the SophiaBeads Datasets, using FDK [8], CGLS [3] and SART [2]. To examine the reconstructions in detail, we focus on a central window of the reconstructed slices (see Fig. 3).

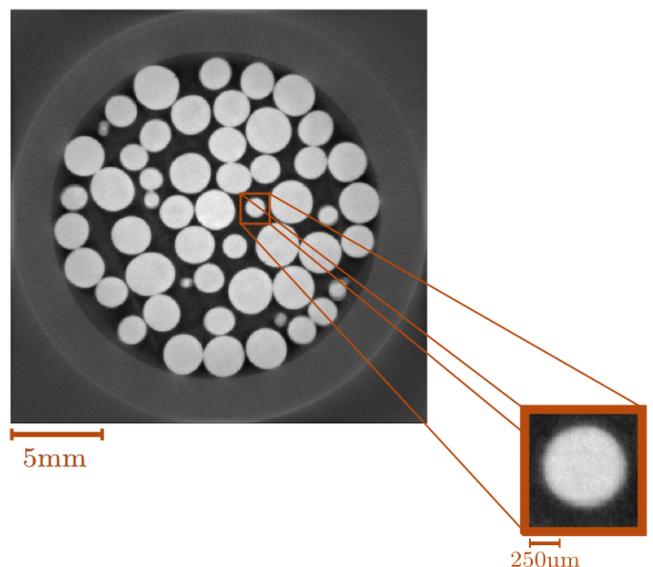


Fig. 3. 2D CGLS reconstruction of 2048 projections, and its center window.

¹The size of the beads is normally distributed with a mean of 2.5mm (in diameter), and a standard deviation of 0.01mm (or 100µm). This means that even though most beads in 2D will look like perfect circles, there will be a proportion of them that are egg-shaped.

³Suitable for 2D and 3D reconstructions.

²Requires 3D reconstructions.

FDK Results

FDK [8] is the standard approach employed by most commercial scanners. Results below are obtained using the in-house implementation of FDK.

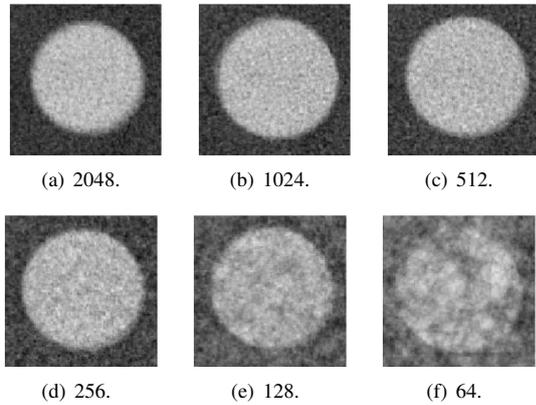


Fig. 4. 2D FDK reconstructions of each SophiaBeads dataset.

One can observe line artefacts and loss of contrast in the reconstructed images in scans with fewer projections. In particular in Fig. 4(f), the bead is almost unidentifiable due to loss of definition in shape.

CGLS Results

This is the Conjugate Gradient method modified for non-square systems such as the CT problem, as explained in [3].

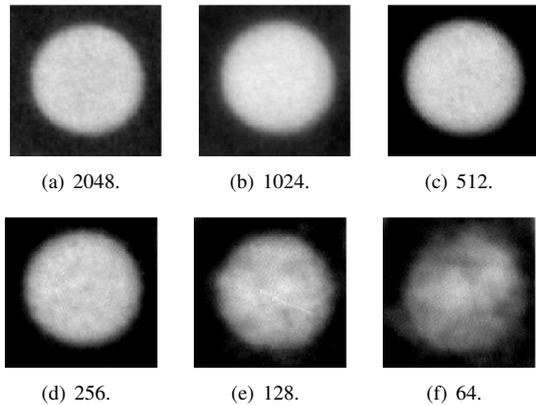


Fig. 5. CGLS reconstructions at iteration 12.

The method is implemented in MATLAB R2014b, with the forward and back projector codes written in C. This method is also used in the SophiaBeads tutorials [4]. The number of iterations is fixed at 12. We observe increase in blur and loss of definition of bead shape in scans with fewer projections.

SART Results

As the third example, we present results using a popular method from the family of algebraic reconstruction techniques [2]. For these runs, we performed 200 sweeps with a relaxation factor chosen as 0.8. Just as with CGLS, we have implemented

and plotted results in MATLAB with forward and back projectors implemented in C.

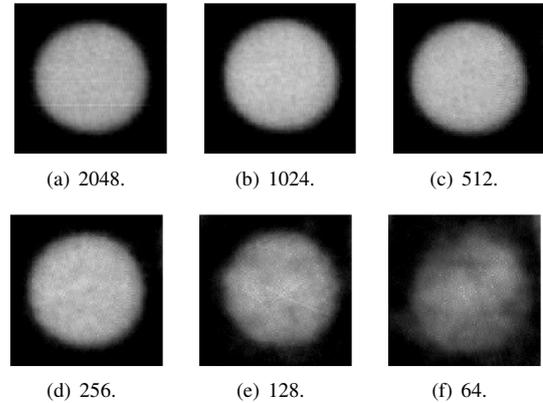


Fig. 6. SART reconstructions with 200 sweeps, and relaxation factor $\omega = 0.8$.

We observe relatively greater loss of definition in the shape and contrast (compared to CGLS), as the number of projections decrease.

Quantifying the SophiaBeads Reconstructions

To evaluate our results we use the quantification item II (henceforth referred to as SHAPE3D). For the analysis, we have used built-in image-measuring techniques in Avizo Fire 8, where the reconstructed volume is read by Avizo (see the quantification tutorial in [4]).

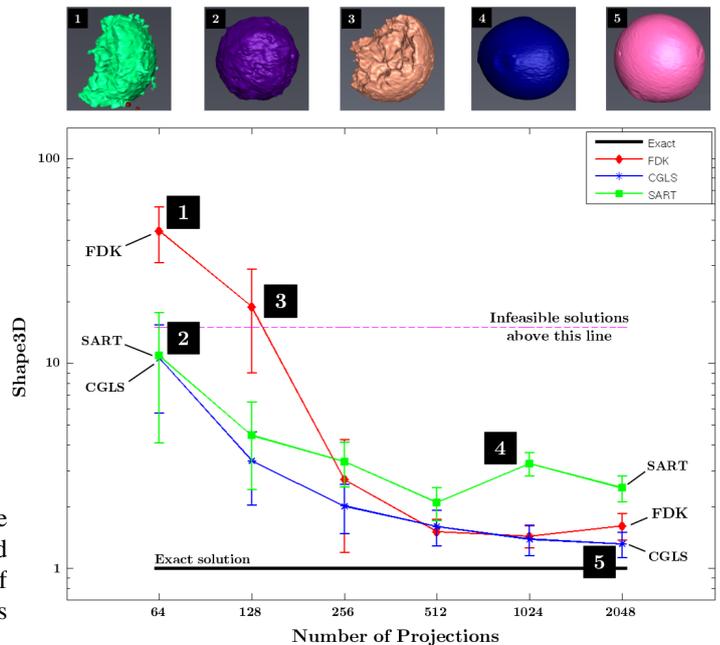


Fig. 7. Results of the SHAPE3D analysis plotted with errorbars in MATLAB.

Fig. 7 is the log-plot of the mean and the standard deviation of each reconstructed volume. The image-measuring technique in Avizo attempts to fit the each bead to a unit sphere, and

parameterizes how close that bead is to a perfect fit (more details in the Avizo user manual [1]). If the reconstructed beads fit the model sphere perfectly, then Avizo outputs 1, so this is taken as the exact answer. Anything above the dashed line can be dismissed as an infeasible solution. From this it is clear that FDK at 64 iterations is a poor choice as a reconstruction method. For datasets with 256 projections or more, we see all three methods giving similar results with small standard deviations.

III. CONCLUSIONS

As the CT community, we welcome novel ideas for iterative reconstruction methods for better quality reconstructions. We spend time on defining our problems, creating and testing ideas, and developing algorithms. Yet we still struggle to answer this simple question: When should we be using iterative methods?

In this paper, we offered a strategy to help us answer this question by introducing a map to plan trials and across which the performance of various algorithms can be charted. Here we examined the effect of altering the number of projections whilst keeping the photon count constant. This has shown that iterative methods deal better with datasets with fewer projections, whereas the FDK method is adequate for scans with 256 projections or higher.

The SophiaBeads Datasets [5] were acquired in such a way that it allows a multi-faceted exploration of the effect of decreasing the information content on the performance of reconstruction algorithms as outlined in Fig. 1. Another key advantage of the SophiaBeads Datasets was that many aspects of the actual 3D object are precisely known, enabling us to quantify algorithm performance. It is noteworthy that the framework in Fig. 1 and the SophiaBeads Datasets allow a wide range of experimental strategies to be simulated (minimum time, dose, number of projections) and the limits of the algorithms delineated or the most appropriate one identified. We would like to note here that there was no prior information used in these reconstructions, which is outside the scope of this discussion. However, improvements in images and changes in fewer projection artefacts when prior information is used are interesting topics that deserve further discussion. In addition, because the beads problem is amenable to discrete tomography, an algorithm (with a suitable prior information) may outperform the current algorithms in quality and speed.

ACKNOWLEDGMENT

This project is funded by the School of Mathematics, EPSRC CCPi (EP/J010456/1) and BP through the ICAM framework.

REFERENCES

- [1] Avizo 8. *Avizo User's Guide*. English. FEI Visualization Sciences Group.
- [2] A.H. Andersen and A.C. Kak. "Simultaneous Algebraic Reconstruction Technique (SART): a superior implementation of the ART algorithm". In: *Ultrason. Img.* 6 (1984), pp. 81–94.
- [3] Å. Björck, T. Elfving, and Z. Strakos. "Stability of conjugate gradient and Lanczos methods for linear least squares problems". In: *SIAM J. Matrix Anal. Appl.* 19(3) (1998), pp. 720–736.
- [4] S.B. Coban. *SophiaBeads Datasets Project Documentation and Tutorials*. MIMS EPrint 2015.26. UK: Manchester Institute for Mathematical Sciences, The University of Manchester, Apr. 2015, p. 22.
- [5] S.B. Coban and S.A. McDonald. *SophiaBeads Datasets*. Mar. 2015. DOI: 10.5281/zenodo.16474. URL: <http://dx.doi.org/10.5281/zenodo.16474>.
- [6] S.B. Coban et al. "A framework for exploring the limits of reconstruction algorithms". In: (*to appear*) (2015).
- [7] S.J. Cooper et al. "Image based modelling of microstructural heterogeneity in LiFePO₄ electrodes for Li-ion batteries". In: *Journal of Power Sources* 247.0 (2014), pp. 1033–1039.
- [8] L.A. Feldkamp, L.C. Davis, and J.W. Kress. "Practical cone-beam algorithm". In: *JOSA A* 1.6 (1984), pp. 612–619.

Appendix C

Effect of sparsity and exposure on
total variation regularized X-ray
tomography from few projections

Effect of sparsity and exposure on total variation regularized X-ray tomography from few projections

Jakob S. Jørgensen*, Sophia B. Coban†, William R.B. Lionheart† and Philip J. Withers†

Abstract—We address effects of exposure and image gradient sparsity for total variation-regularized reconstruction: is it better to collect many low-quality or few high-quality projections, and can gradient sparsity predict how many projections are necessary? Preliminary results suggest collecting many low-quality projections is favorable, and that a link may exist between gradient sparsity level and successful reconstruction.

I. INTRODUCTION

Sparsity regularization for X-ray computed tomography (CT) image reconstruction, for example total variation (TV) regularization [1] for gradient-sparse images, has been seen to allow drastically reduced numbers of projections compared to conventional analytical methods, see, e.g. [2]. In medical imaging and non-destructive testing this may allow reduced X-ray exposure or data acquisition time. In today’s literature, there is little quantitative guidance on how much TV-regularization allows us to reduce the number of projections. In order for TV and other forms of sparsity regularization to become appropriately used this lack of knowledge must be filled. Our recent work [3] has indicated in simulations inspired by compressed sensing [4] that sparsity of the image gradient can predict how few projections will suffice for accurate TV-regularized reconstruction. A main goal of the present work is to investigate, for the first time, if the same argument holds using real X-ray CT data.

In the present study we consider exposure as the measurement cost, and – given a fixed total exposure – look at the trade off between more information obtained at lower quality (more projections at low exposure) and less information at higher quality (fewer projections at high exposure). The sparsity-regularization literature often takes number of projections as the primary variable, however the total exposure of a CT scan depends both on the number of projections and the exposure-per-projection. [2], [4]. So one could also reduce the total exposure by keeping the usual high number of projections but decreasing the exposure time of each. It is not immediately clear in which scenario TV-regularized reconstruction will perform better. Addressing this issue is the other main goal of this work.

In the present work we systematically study TV-regularized reconstruction quality at reduced numbers of projections as a function of both exposure time and gradient sparsity using real CT data. Specifically we address:

Q1: Does TV-regularized reconstruction compensate better for reduced information from few high-exposure or many low-exposure projections?

Q2: Is there a connection between gradient sparsity and how few projections provide enough information that TV-regularized reconstruction succeeds?

While Q1 considers a fixed total exposure with exposure-per-projection inversely proportional to the number of projections, Q2 considers a constant exposure-per-projection and a total exposure proportional to the number of projections.

We will use the recently published SophiaBeads data set, which has been designed specifically for systematic studies of advanced reconstruction algorithms. In addition to using this data set to address the stated questions, we apply the present work to examine how appropriately the SophiaBeads data set can serve the purpose of testing sparsity-regularized reconstruction methods.

An important note needs to be made here about our definition of an ‘adequate reconstruction’. In our earlier work [3], we used a relative 2-norm measure to assess if reconstructions *perfectly* recovered the ground truth. This was appropriate for the idealized scenario and to stay consistent, we also report 2-norm errors in this work. However with real data, we wish to assess how well important features can be quantified; in this case known to be disk-shaped and we employ an aspect-ratio quality measure as explained in §IV-B.

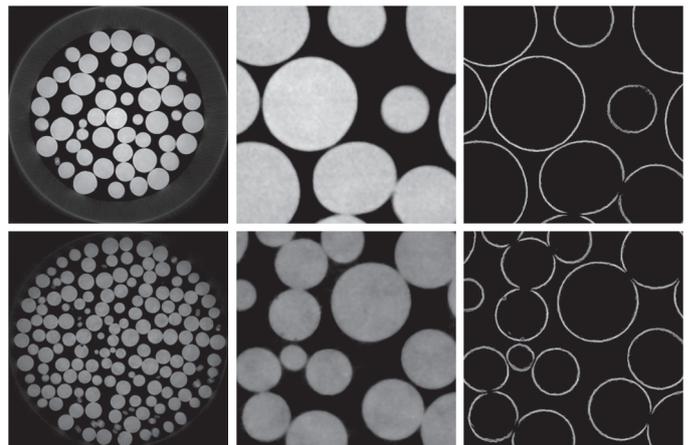


Fig. 1. Ground truths for SophiaBeads data sets S1 (top row) and S2 (bottom row) obtained by 30 CGLS iterations from pooled projections, followed by median filtering. Full 1564×1564 images (left), 350×350 region of interest around the centre (centre), and sparse thresholded gradient magnitude region of interests (right).

*Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

†The University of Manchester, Manchester, M13 9PL, United Kingdom.
Corresponding author contact: jakj@dtu.dk.

II. TEST DATA

A. The SophiaBeads test data set

The SophiaBeads Dataset Project [5] is a collection of cone-beam X-ray CT data sets where the number of projections are varied while the total photon count (or the total exposure time) is kept constant, i.e. the exposure-per-projection is inversely proportional to the number of projections, as in Q1. This enables a wide range of algorithm comparisons and information content optimizations to be examined. For more detailed information on this experimental framework and the examples of such scenarios, we refer the reader to [6].

The SophiaBeads data set were collected using the 320/225 kV Nikon XTEK Bay at the Manchester X-ray Imaging Facility (MXIF), the University of Manchester. The apparatus consists of a cone-beam microfocus X-ray source that projects polychromatic X-rays onto a 2000×2000 pixel-length and width, 16-bit flat detector panel. The optimal window size for the SophiaBeads reconstructions is 1564×1564 , see [7].

There are two samples (henceforth referred to as S1 and S2) that were scanned using the framework described in [6], and both samples comprised a plastic tube with a diameter of 25mm, filled with uniform Soda-Lime Glass ($\text{SiO}_2\text{-Na}_2\text{O}$) beads of diameters 2.5mm (S1) and 1.0mm (S2). S1 is publicly available; S2 on request. Here, we use S1 and S2 to represent different sparsity levels: the smaller beads of S2 have relatively more boundary pixels, which equates to more non-zero pixels in the gradient, and hence is less sparse than S1.

The present study uses a single central row of the 3D cone-beam data, and a 2D fan-beam geometry. For the constant-exposure series, the available data sets labelled 64-, 128-, 256-, 512- and 1024-projection are used. For the reduced-exposure series the 1024-projection data set is downsampled by repeatedly halving the number of projections while keeping every other one, thereby preserving the equiangular distribution.

B. Determining a ground truth image and its sparsity

The SophiaBeads data set is designed with fixed total exposure ranging from few high-exposure projections to many low-exposure projections. No high-quality data set (many high-exposure projections) is provided for the construction of a ground truth. However, each data set is recorded at slightly offset angular positions and we obtain a ground truth by pooling all projections for each of S1 and S2 and reconstruct using 30 iterations of the Conjugate Gradient Least Squares (CGLS) algorithm, followed by median filtering with a 5×5 filter to reduce noise. The resulting S1 and S2 ground truths are shown in full and close-up in Fig. 1.

To determine gradient sparsity of the ground truth images we count only nonzero gradient magnitude values greater than a threshold chosen empirically to preserve only bead edges and not noise. Thresholded gradient magnitude images are shown in Fig. 1. The S1 ground truth has 54543 nonzero values in its gradient, corresponding to a sparsity level (relative to the total number of pixels) of $54543/1564^2 = 2.2\%$. The same numbers for S2 are 123870 and 5.1%. This quantifies the

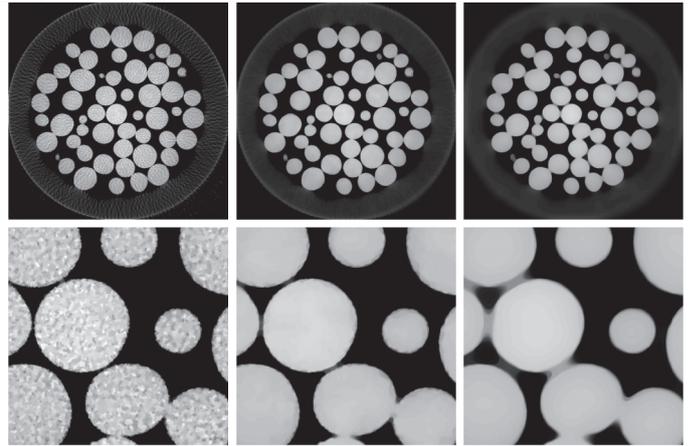


Fig. 2. TV-regularized reconstructions using regularization parameters 10^{-3} , 10^{-2} and 10^{-1} (left to right). Full 1564×1564 images (top row) and 350×350 region of interest around the center (bottom row).

intuition that S1 is more gradient-sparse than S2, though exact numbers may vary depending on thresholds chosen.

III. RECONSTRUCTION PROBLEM AND ALGORITHM

A. Total variation optimization problem

We denote the log-transformed projection data by b , the 2D fan-beam system matrix by A , an image such as a reconstruction by u , in particular a TV-regularized solution by u_{TV} , and the number of projections by N_θ . To determine a TV-regularized reconstruction (which can be seen as the maximum a posteriori estimate in a Bayesian formulation) of the discrete imaging model $Au = b$ we solve the optimization problem

$$u_{\text{TV}} = \arg \min_u \frac{1}{2N_\theta} \|Au - b\|_2^2 + \alpha T_\tau(u), \quad u \geq 0, \quad (1)$$

where we employ a standard Huber-smoothed TV defined as

$$T_\tau(u) = \sum_j \Phi_\tau(\|D_j u\|_2), \quad \text{where} \quad (2)$$

$$\Phi_\tau(z) = \begin{cases} |z| - \frac{1}{2}\tau & \text{if } |z| \geq \tau, \\ \frac{1}{2\tau}z^2 & \text{else.} \end{cases} \quad (3)$$

Here, α is the TV regularization parameter, D_j is a finite difference approximation to the gradient at pixel j and $\|\cdot\|_2$ denotes the vector 2-norm (or Euclidian norm).

Smoothing is used to make the problem solvable by smooth optimization techniques which are generally faster than their non-smooth counterparts. Depending on the choice of smoothing parameter, τ , this might modify the reconstruction; however here we use a sufficiently small value of $\tau = 10^{-5}$ relative to the image values that smoothing effects are negligible.

Non-negativity is enforced as the object's attenuation coefficients are known to be non-negative and in general non-negativity can lead to substantial reconstruction improvement.

The normalization by N_θ helps to compare reconstructions obtained at different N_θ by compensating the magnitude of the first term which is otherwise proportional to N_θ . As a result, a fixed α value yields the same balance between the two terms

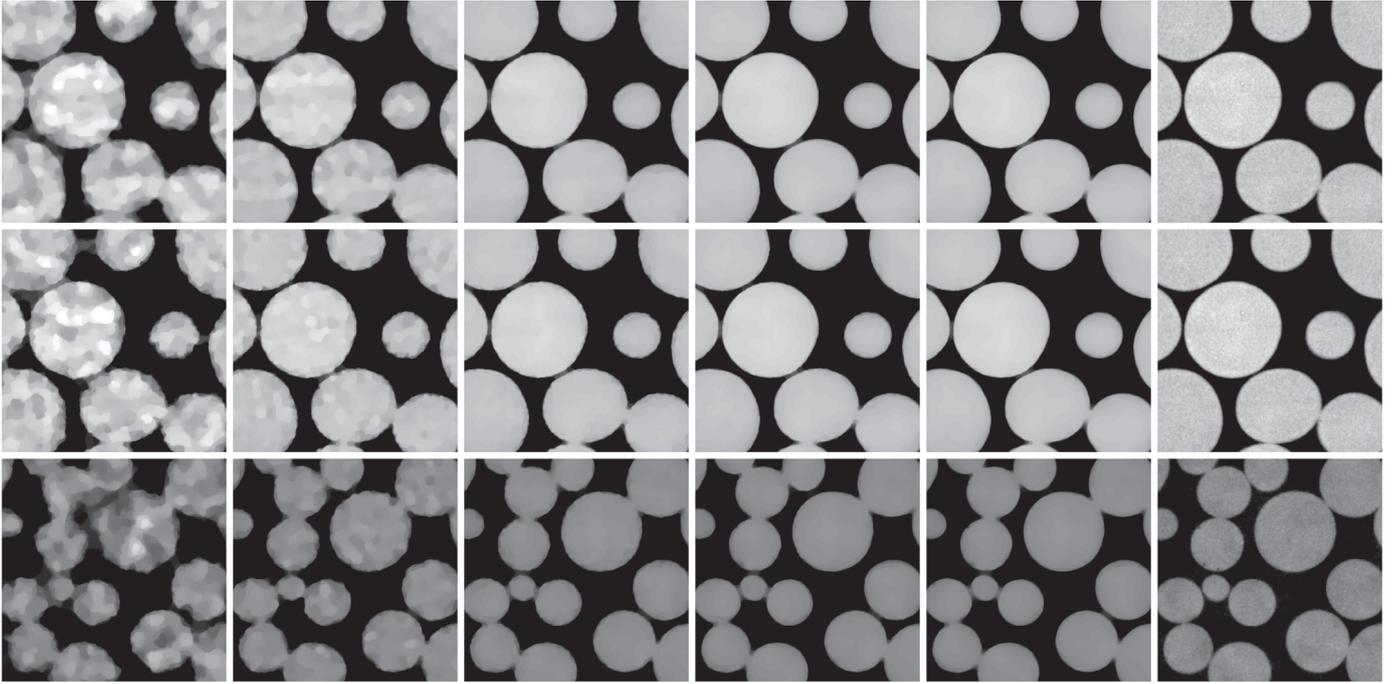


Fig. 3. TV-regularized reconstructions of 64-, 128-, 256-, 512-, 1024-projection data sets and pooled-data ground truths (left to right), showing a 350×350 -pixel region of interest. S1 data set of fixed total exposure (top row), S1 and S2 data sets with fixed per-projection exposure (middle and bottom rows).

at different N_θ . This reduces the search for the optimal α to a single initial sweep, the resulting α of which can be reused. In practice we verified this through α sweeps at different N_θ but for brevity have not included results here.

B. High-accuracy optimization algorithm

To solve (1) we used the toolbox TVReg [8], which offers implementations (written in C with MATLAB interface) of accelerated gradient projection methods; specifically we used the provided GPBB (Gradient Projection Barzilai-Borwein) method which among other techniques employ acceleration in form of the Barzilai-Borwein step-size selection. To further accelerate the reconstruction, we employed the ASTRA Tomography Toolbox [9] for GPU-acceleration of the computationally expensive forward and back-projection operators.

We emphasize that our goal here is not necessarily to use the fastest algorithm but one that can reliably solve (1) to high accuracy in reasonable time in order that we indeed assess the quality of the TV-regularized reconstruction and not of an arbitrary early-termination result. TVReg is capable of this through a non-heuristic termination criterion based on the gradient norm magnitude, in contrast to, for example, running a pre-set fixed number of iterations or terminating when a small difference between iterates is encountered.

IV. RESULTS

A. Choosing the regularization parameter

Fig. 2 shows reconstructions for $\alpha = 10^{-3}$, 10^{-2} and 10^{-1} showing the well-known transition from an under-regularized noisy/patchy TV-regularized reconstruction, through to an

over-regularized solution where separated beads appear connected due to excessive smoothing. Among a range of values we found $\alpha = 10^{-2}$ to provide the best trade-off and this fixed value was reused in the remaining reconstructions.

B. Assessment of reconstruction image quality

We assess the reconstructions qualitatively through visual inspection. For quantitative assessment we use two error measures with respect to the constructed ground truth u_{GT} : First, the standard relative 2-norm of pixelwise differences: $E_1(u_{TV}) = \|u_{TV} - u_{GT}\|_2 / \|u_{GT}\|_2$, where $\|\cdot\|_2$ denotes the (Euclidian) 2-norm. The relative 2-norm provides a standardized comparison but is not necessarily the most informative about whether important features have been reliably reconstructed. For the second error measure $E_2(u_{TV})$, we evaluate the aspect ratio (width:height) of 25 reconstructed beads and report the mean relative error with respect to determined aspect ratios of the ground truth beads. This measure describes how well bead reconstructions reproduce the known bead shapes.

C. Q1: Is it better to collect few high-exposure or many low-exposure projections?

We first address Q1 by determining the TV-regularized reconstruction of the fixed total exposure S1 data set for 64, 128, 256, 512 and 1024 projections. We visually compare a 350×350 -pixel region of interest of all reconstructions with the constructed ground truth in the top row of Fig. 3. Visual quality clearly improves with increasing number of projections.

The error measures E_1 and E_2 are plotted in Fig. 4 using a full line. Both error measures agree with visual assessment that

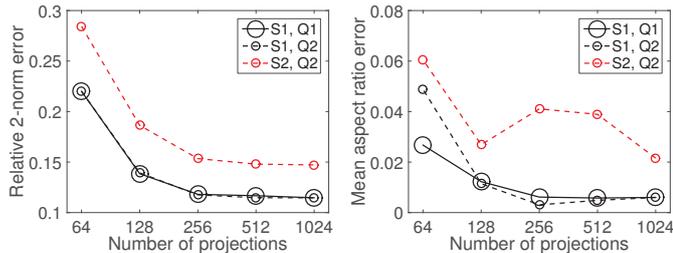


Fig. 4. Relative 2-norm error (left) and mean aspect ratio error (right) for data sets S1 and S2 with fixed total or per-projection exposure, Q1 and Q2.

the 64-projection reconstruction is substantially worse than the others and that error decays with more projections.

D. Q2: What is the dependence on sparsity?

To address Q2 the middle and bottom rows of Fig. 3 show reconstructions for the case of fixed exposure-per-projection for data sets S1 and S2. Corresponding error measures are plotted in Fig. 4 using dashed lines. First, again a clear trend of improved TV-regularized reconstruction quality with increasing number of projections is observed. This is less surprising than in the previous case, since more projections correspond to a higher X-ray exposure. However, for S2 the mean aspect ratio error for 256 and 512 projections is larger than the general trend. We also note that E_1 for S1 in this case almost coincides with the fixed total exposure case.

In case the gradient sparsity does in fact affect the number of projections sufficient for accurate reconstruction, we would expect to see clear differences between S1 and the more gradient-sparse S2 data set. However, visually the S1 and S2 reconstructions show no clear difference in their dependence on the number of projections. The error plots also do not reveal clearly different behavior of S1 and S2 as function of numbers of projections, apart from the previously mentioned E_2 values for S2 at 256 and 512 projections.

V. DISCUSSION AND CONCLUSIONS

In all considered cases the 64-projection reconstructions stand out from the rest as substantially poorer. It seems that artifacts caused by having only 64 projections cannot be effectively removed by TV-regularized reconstruction, no matter whether high- or low-exposure projections are used. This is particularly interesting considering the highly gradient-sparse and round, piece-wise flat regions, for which TV-regularized reconstruction could be expected to excel.

For Q1, we conclude that given a fixed total exposure it appears beneficial to distribute across the highest possible number of projections. Even though each projection is of low quality it appears intuitively sensible to aim for obtaining in a loose sense more independent information about the scanned sample through more projection angles, rather than few high-quality ones. This is however in contrast to the typical message from the sparsity-regularization literature, namely that reconstruction from few projections is possible.

Regarding a possible connection to sparsity in Q2, present results are inconclusive since no clear difference is observed between S1 and S2. However for both S1 and S2 results, there is a large error reduction between 64 and 128 projections. This may hint that there is a number of projections, possibly different for each of S1 and S2, below which TV-regularized reconstruction will not be successful. The SophiaBeads data set only allows subsampling by factors of 2 to preserve equiangular projections. Relevant future work includes the acquisition and analysis of data sets with finer increments of numbers of projections, as well as more sparsity levels.

It should be mentioned that the presented preliminary conclusions may depend on several aspects of the study. For example it is unclear if the pooling approach produces a reliable enough ground truth, and in potential future work, extra care should be taken to acquire ground truth data. Also it is not certain that the error measures used here are the most informative and other options could be considered.

Lastly, regarding how SophiaBeads data sets serve as sparsity-regularization test data, we found TV-regularized reconstruction to work well on the piecewise constant bead images. In that sense, SophiaBeads is quite useful. However for assessing the influence of gradient sparsity we faced shortcomings which we have offered suggestions to address in future work.

ACKNOWLEDGMENTS

The work by JSJ was supported by the project ‘‘High-Definition Tomography’’ funded by Advanced Grant No. 291405 from the European Research Council. The work by SBC was supported by the School of Mathematics (UoM), EPSRC CCPi (EP/J010456/1), and BP through the BP International Centre for Advanced Materials (BP-ICAM). Authors would also like to acknowledge the support of Samuel A. McDonald (MXIF) and EU COST action MP1207 (EXTREMA).

REFERENCES

- [1] L. I. Rudin, S. Osher, and E. Fatemi, ‘‘Nonlinear total variation based noise removal algorithms,’’ *Physica D*, vol. 60, pp. 259–268, 1992.
- [2] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, ‘‘Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT,’’ *Phys Med Biol*, vol. 55, pp. 6575–6599, 2010.
- [3] J. S. Jørgensen and E. Y. Sidky, ‘‘How little data is enough? Phase-diagram analysis of sparsity-regularized X-ray computed tomography,’’ *Philos Trans R Soc Lond Ser A*, vol. 373, p. 20140387, 2015.
- [4] E. J. Candès, J. Romberg, and T. Tao, ‘‘Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,’’ *IEEE Trans Inf Theory*, vol. 52, pp. 489–509, 2006.
- [5] S. B. Coban and S. A. McDonald, ‘‘SophiaBeads Dataset Project,’’ Tech. Rep., 2015. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.16474>
- [6] S. B. Coban, P. J. Withers, W. R. B. Lionheart, and S. A. McDonald, ‘‘When do iterative reconstruction methods become worth the effort?’’ in *Proc Fully3D*, Newport, RI, USA, 2015.
- [7] S. B. Coban, ‘‘SophiaBeads Dataset Project Codes,’’ 2015. [Online]. Available: <http://dx.doi.org/10.5281/zenodo.16539>
- [8] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, ‘‘Implementation of an optimal first-order method for strongly convex total variation regularization,’’ *BIT*, vol. 52, pp. 329–356, 2012.
- [9] 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*, vol. 157, pp. 35–47, 2015.