

CoShaRP: A Convex Program for Single-shot Tomographic Shape Sensing

Ajinkya Kadu^{1,2}, Tristan van Leeuwen^{1,3}, K. Joost Batenburg^{1,4}

¹ Computational Imaging, Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

² Electron Microscopy for Material Science, University of Antwerp, Belgium

³ Mathematical Institute, Utrecht University, Utrecht, The Netherlands

⁴ Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands

Abstract. We introduce single-shot X-ray tomography that aims to estimate the target image from a single cone-beam projection measurement. This linear inverse problem is extremely under-determined since the measurements are far fewer than the number of unknowns. Moreover, it is more challenging than conventional tomography, where a sufficiently large number of projection angles forms the measurements, allowing for a simple inversion process. However, single-shot tomography becomes less severe if the target image is only composed of known shapes. This paper restricts analysis to target image function that can be decomposed into known compactly supported non-negative-valued functions termed shapes. Hence, the shape prior transforms a linear ill-posed image estimation problem to a non-linear problem of estimating the roto-translations of the shapes. We circumvent the non-linearity by using a dictionary of possible roto-translations of the shapes. We propose a convex program CoShaRP, to recover the dictionary coefficients successfully. CoShaRP relies on simplex-type constraints and can be solved quickly using a primal-dual algorithm. The numerical experiments show that CoShaRP recovers shape stably from moderately noisy measurements.

Keywords: computational imaging, X-ray tomography, compressed sensing, shape identification, convex optimization

Submitted to: *Inverse Problems*

1. Introduction

In tomographic imaging, the aim is to characterize the three-dimensional structure of an object from X-ray projections. In applications like medical CT, projections are gathered from all directions. This allows for a relatively straightforward reconstruction of the object using so-called filtered back-projection methods (FBP)[1, 2, 3]. When only a limited angular sampling is available, more advanced iterative reconstruction techniques that use prior information about the structure of the object have been

developed [4, 5, 6, 7, 8]. This paper considers an extreme case where we acquire only a single X-ray projection on which to base a complete three-dimensional reconstruction. We refer to this problem as *single-shot X-ray tomography*. This problem is highly relevant for many applications, including industrial quality control and high-throughput imaging [9, 10, 11, 12]. In particular, the industrial tomography in a conveyor belt configuration [13], electron tomography of extremely beam-sensitive nano-assemblies [14] and cryo-electron tomography of biological samples with repetitive structures such as amino-acid groups [15] demand single-shot X-ray projection. A typical setup consists of a fixed X-ray source and detector that collects single X-ray projections of the objects of interest. While we specifically target reconstruction from single-angle X-ray projections, the techniques we develop are also relevant for limited-angle tomography with applications including high-resolution dynamic imaging[16, 17] and cryo-electron microscopy [18, 19, 20].

The single-shot X-ray tomography problem is extremely under-determined, making it an ill-posed inverse problem. A single cone-beam projection of a volume containing n^3 voxels consists of $\mathcal{O}(n^2)$ measurements. Hence, the measurements are undersampled by a factor of n . To reduce the ill-posedness, it is a general practice to incorporate prior information via regularization. However, classic regularization methods fail on single-shot X-ray tomography, as demonstrated in Figure 1.

Since it is evident that a strong prior is necessary to recover the target image from a single projection, we consider the class of objects that are composed of a limited number of known building blocks. This is a reasonable assumption when imaging materials that are made up of basic structures, such as a 3D structure comprising interlocking bricks or a protein structure consisting of repeated amino acid groups. Hence, if such shapes and their number of repetitions are known a-priori, the image estimation problem can be recast as an estimation of roto-translation parameters of these shapes. The recovery of shape parameters (*i.e.*, the roto-translations of shapes) from the linear measurements of the image that consists of these shapes is called *shape sensing*. However, unlike the image estimation problem – which has a simple linear structure – the estimation of roto-translation parameters is a non-linear problem. This, in turn, makes the inversion process computationally intractable. To avoid such non-linearity, we use a shape dictionary approach that expresses the target image as a linear combination of shapes from the available dictionary. Due to the linear structure, we show that it is possible to recover the shapes from a single-shot X-ray projection (see Figure 1 and 2) by solving a convex problem. We regard the recovery of shapes from their tomographic projections as *tomographic shape sensing*.

1.1. Related Work

The shape sensing problem has been studied in the context of shape-based characterization, object tracking and optical character recognition. Inspired by the compressed sensing, a recent approach called Sparse Shape Composition (SSC) imposes

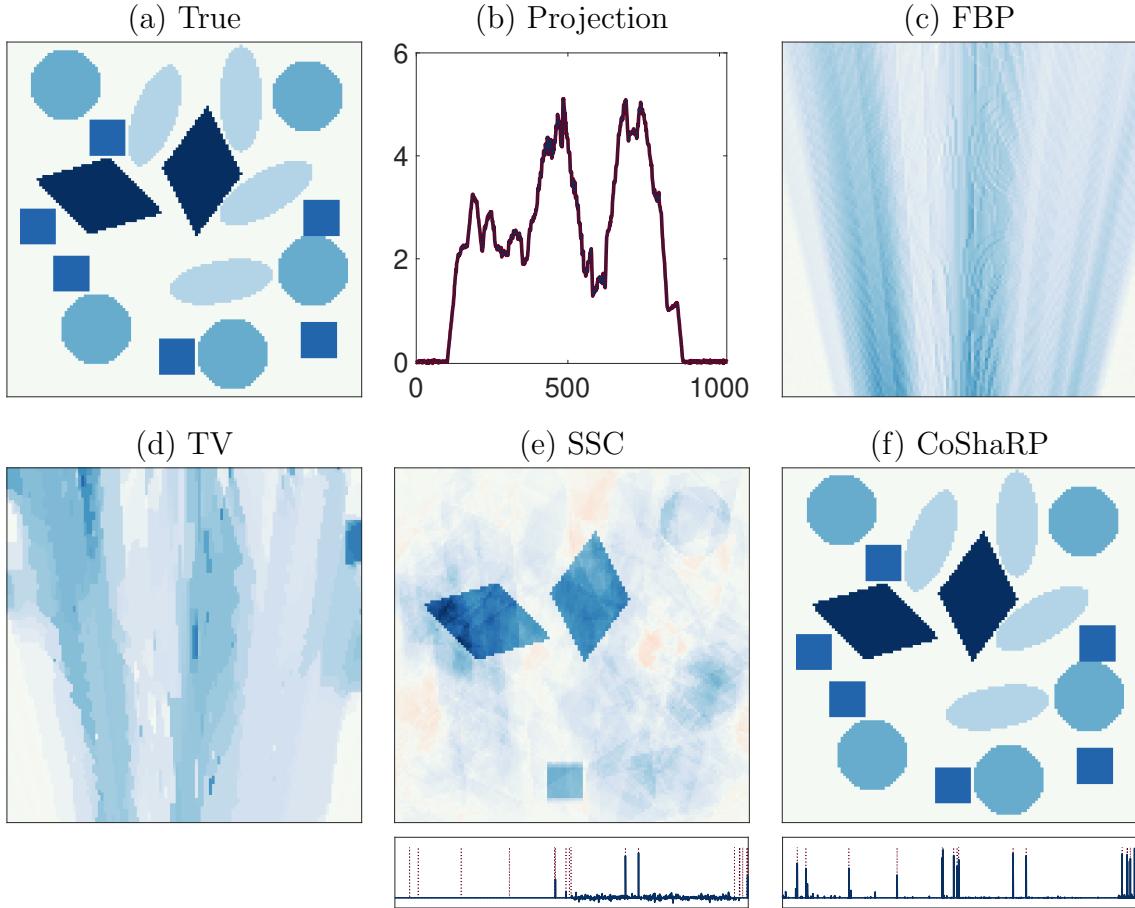


Figure 1. Demonstration of single-shot X-ray tomography on 2D numerical phantom: A target image (a) of size $1 m \times 1 m$ discretized on a 128×128 pixel grid has a fan-beam projection shown in (c). The image consists of 4 different shapes with different intensities. The equally-spaced detectors place at the top of the image collect 1024 measurements resulting from an X-ray source at the bottom of the image. We consider FBP (c), Total Variation (TV) regularization (d), and Sparse Shape Composition (e). Our proposed approach CoShaRP is given in (f). The shape coefficients for SSC and CoShaRP are given below their figures, while dotted ones denote the correct coefficients.

an ℓ_1 -norm constraint on the shape-dictionary coefficients [21, 22]. SSC has the advantage that it can form new shapes from the intersection and union of basic shapes. However, the main drawback comes from large feasible solution space inherent to the ℓ_1 -norm constraint in high dimension (see Remark 1). This large feasible space may lead to an incorrect solution. A simplistic version of SSC performs the 3D characterization of nano-particles using electron tomography [23]. This method uses a simple ℓ_1 -norm constraint to recover spherical nano-particles from their tomographic projections. However, their tomographic projections have a parallel-beam geometry and require measurements from more than one projection angle. Although SSC has been extended to tomographic inverse problem, it has never been tested for single-shot tomographic shape sensing. However, we demonstrate the failure of SSC in single-shot fan-beam tomography in Figure 1.

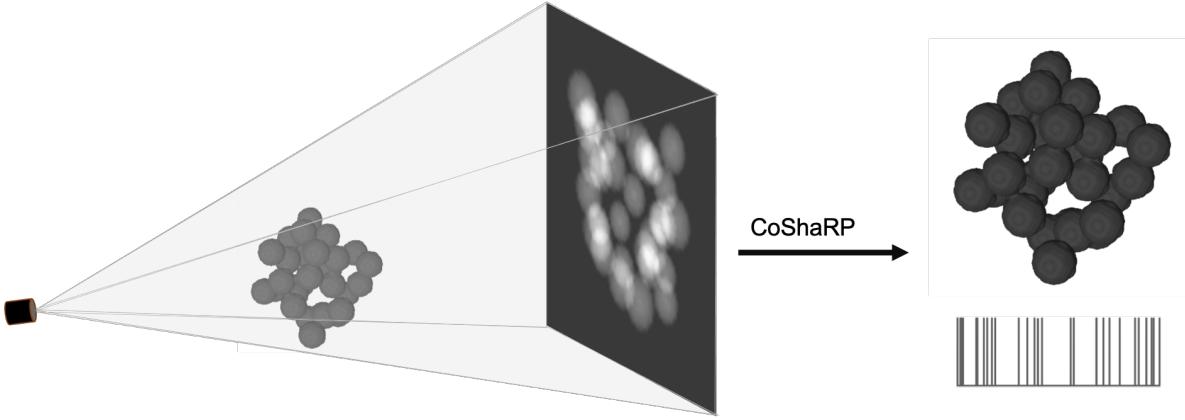


Figure 2. Demonstration of single-shot tomography on 3D numerical phantom: A target image (acquisition of which is shown on the left) consists of 30 spheres of equal radius spread around the domain. A single-shot projection is acquired on the 2D pixel grid. The reconstructed image from the CoShaRP algorithm is shown on the right, while the bottom figure shows the dictionary coefficients retrieved from CoShaRP.

1.2. Contributions and Outline

To the best of our knowledge, the single-shot X-ray tomography has never been studied, and no reconstruction method exist till date to recover back an image successfully from a single-shot. We introduce the tomographic shape sensing problem that assumes the prior knowledge about the shapes in the image. The principal contribution of this paper is the development of the convex program CoShaRP to reconstruct images composed of non-overlapping shapes from a single-shot. The convex program is novel in the sense that the simplex-type constraint enables sharp recovery results from extremely under-determined single-shot tomographic projections. Although the exact recovery problem is NP-hard, our proposed convex program CoShaRP stably recovers the shapes. Moreover, we propose a primal-dual algorithm to find the optimal solution of CoShaRP. The algorithm does not rely on the inversion of large matrices and uses a simple proximal operator to project onto a K -simplex constraint. Using numerical experiments, we answer the following questions:

- (i) What is the minimum resolution of the shape that CoShaRP can sense?
- (ii) Is CoShaRP robust to the number of repetitions and the possible rotations of the shapes in the target image?
- (iii) Can CoShaRP recover non-homogeneous as well as non-convex shapes?
- (iv) How sensitive is CoShaRP to the measurement noise?

We discuss the single-shot tomographic inverse problem in Section 2. Section 3 discusses the tomographic shape sensing problem and introduces a convex program CoShaRP. We describe an efficient iterative scheme to find an optimal solution to CoShaRP in Section 4. We illustrate the numerical experiments in Section 5 and conclude the paper in Section 7.

1.3. Notation

Throughout this paper, small boldface letters (*e.g.*, \mathbf{x}, \mathbf{z}) denote vectors in \mathbb{R}^n . The identity and zero elements are denoted by $\mathbf{0}$ and $\mathbf{1}$ respectively. The Euclidean inner product is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with a corresponding norm $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. However, for all other norms, we use subscripts (*e.g.*, $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$, $\|\mathbf{x}\|_\infty = \max_i |x_i|$). To represent the matrices, we use uppercase letters (*e.g.*, $\mathbf{A}, \mathbf{\Psi}$). The elements of a matrix \mathbf{A} are denoted by a_{ij} . All the functions are represented as $f : \mathcal{X} \mapsto \mathcal{Y}$, where \mathcal{X} and \mathcal{Y} are the domain and co-domain of f , respectively. We denote the convex conjugate of the function f by f^* . $\text{prox}_f(\mathbf{z})$ denotes the proximal of function f evaluated at point \mathbf{z} (for definition, please refer to [24]). We represent the optimal solution to the optimization problem using overline (*e.g.*, $\overline{\mathbf{x}}, \overline{\mu}$).

2. Single-shot X-ray Tomography

The acquisition geometry of single-shot X-ray tomography consists of one source and an array of regularly spaced detectors (See Figure 3 for the reference). Let $\varphi \in \mathbb{S}^{d-1}$ be a directional vector, and $\mathbf{r} \in \mathbb{R}^d$ be any position vector, where $d \in \{2, 3\}$ is the dimension of the scene. The cone-beam transform A_C of an image function $x : \mathbb{R}^d \mapsto \mathbb{R}$ is its integral along a line in the direction φ passing through \mathbf{r} . It is mathematically given by

$$(A_C x)(\mathbf{r}, \varphi) = \int_0^\infty x(\mathbf{r} + t\varphi) dt.$$

In a single-shot setup, we have a source located at \mathbf{r}_0 . It sends multiple X-rays through the object (compactly supported on $\Omega \subset \mathbb{R}^d$) in a cone with a vertex at \mathbf{r}_0 and spanning

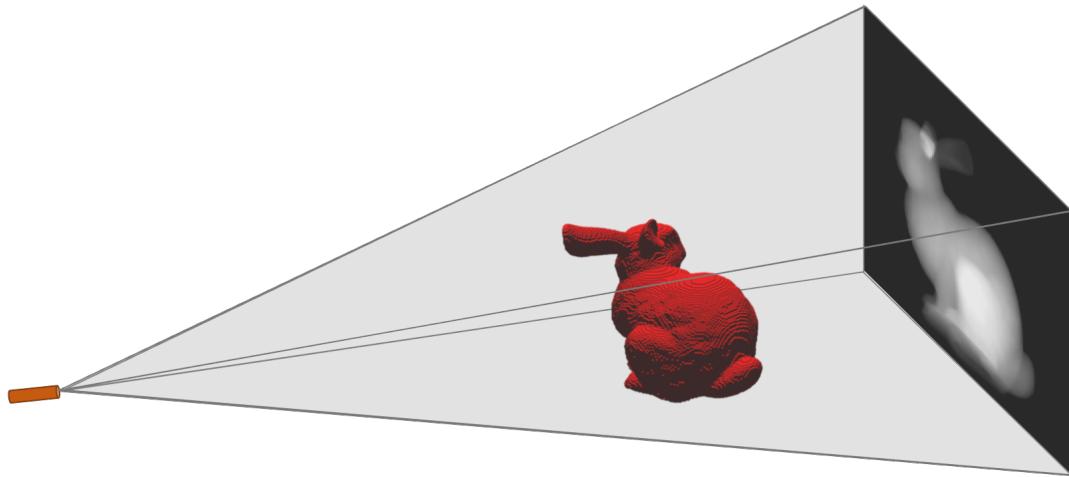


Figure 3. A single-shot cone-beam projection of the Bunny phantom: On the left, a source positioned at \mathbf{r}_0 emits X-rays in a tetrahedral cone form. The panel, shown on the right, has an array of detectors that measure X-rays' attenuated intensity profile (*i.e.*, $(A_C x)$). In a single-shot tomography, we want to estimate the Bunny phantom (shown in the middle) from the projection image we measure on the right.

directional vectors in a set Φ that determines the geometry of cone. Let these vectors be $\varphi_i \in \Phi$, $i = 1, \dots, m$, then the measurement y_i is

$$y_i = (A_C x) (\mathbf{r}_0, \varphi_i) \approx \sum_{j=1}^n a_{ij} x_j,$$

where x_j denotes the value of x in the j^{th} voxel and a_{ij} is the contribution of the j^{th} voxel to the i^{th} ray. The measurements can now be expressed as a linear system of equations

$$\mathbf{y} = \mathbf{Ax}.$$

The above linear system of equations is extremely under-determined since the number of measurements m is far smaller than the number of unknowns n . We do assume that each voxel is intersected by at least one ray, so that each column of the matrix has at least one non-zero element. Determining the image from the measurements is an ill-posed inverse problem. In general, to resolve this ill-posedness, regularization needs to be added in the inversion procedure to incorporate the prior knowledge about the target image. However, conventional regularization techniques are not sufficient to resolve the true image, as was illustrated in Figure 1.

2.1. Conventional vs Single-shot Tomography

In conventional cone-beam tomography, we acquire the measurements (on the 2D grid) for various source positions around the object. Although multiple combinations of the

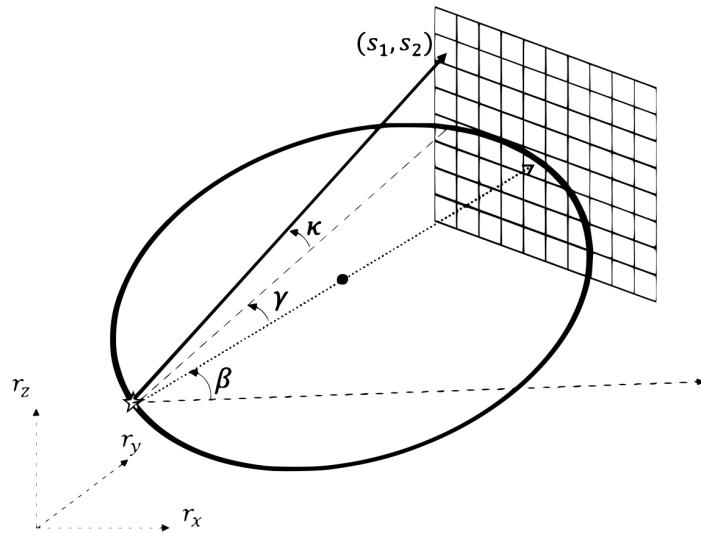


Figure 4. Cone-beam acquisition with circular trajectory. The circle has a radius R which has a trajectory parallel to the $r_x - r_y$ plane. ‘ x ’ denotes the source, and the pixel grid on the right is the measurement plane. β is the projection angle (measured as the angle made by a ray from the source to the centre of the measurement grid with the r_x axis). γ is the fan-angle, and κ is the cone-angle. (s_1, s_2) are the local coordinates on the measurement grid.

source trajectories are possible, we restrict ourselves to a circular cone-beam geometry prominent in industrial and biomedical X-ray machines. As described in Figure 4, the source moves around the object in a circle with radius R . The measurement plane, placed across the source, receives the attenuated X-rays from the object. We represent the measurements for conventional tomography in the form as follows:

$$y(\beta, \mathbf{s}) = \int_0^\infty x (\mathbf{r}(\beta) + t\boldsymbol{\varphi}(\beta, \mathbf{s})) dt,$$

where $\beta \in [-\pi, \pi]$ is the projection angle, $\mathbf{s} = (s_1, s_2)$ are the local coordinates on the measurement grid. If the source trajectory is aligned with the $r_x - r_y$ axis (refer to Figure 4), then

$$\mathbf{r}(\beta) = \begin{bmatrix} R \cos \beta \\ R \sin \beta \\ 0 \end{bmatrix}, \quad \begin{aligned} \langle \boldsymbol{\varphi}(\beta, \mathbf{s}), \mathbf{r}_x \rangle &= \cos(\beta + \gamma) \\ \langle \boldsymbol{\varphi}(\beta, \mathbf{s}), \mathbf{r}_y \rangle &= \cos(\pi/2 - \beta - \gamma), \\ \langle \boldsymbol{\varphi}(\beta, \mathbf{s}), \mathbf{r}_z \rangle &= \cos(\kappa) \end{aligned}$$

with fan-angle γ , cone-angle κ , and $\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z$ denoting the canonical vectors in \mathbb{R}^3 . For the 2D case (*i.e.*, fan-beam), we can collapse the r_z -axis and derive an equivalent formulation. We use this convention mainly to discuss practical reconstruction algorithms (refer to Section 2.2). In the conventional tomography, we sample β sufficiently large enough in $[-\pi, \pi]$ to satisfy the projection requirements. However, an extensive sampling of β translates to high acquisition time and beam damage to the sample. To improve the acquisition time and reduce the beam damage, we lower the sampling of β to just 10-20 angles. This tomography is generally termed *sparse-angle tomography*.

In Figure 5, we plot the information content available in each of these acquisitions: conventional, sparse-angle and single-shot. In particular, we have plotted the frequency content by backprojecting these measurements in the image space. As shown in Figure 5, the information about the target image (in the frequency domain) reduces drastically as we move from conventional to sparse-angle to single-shot. This is because the conventional acquisition provides low- and high-frequency content sufficient to retrieve the target image. The sparse-angle setting contains the reduced information but consists of enough low-frequency content and few high-frequency components. However, for single-shot tomography, there is almost no information available in the high-frequency region. Moreover, the low-frequency content is not enough to retrieve the complete 3D structure of the object.

2.2. Practical Reconstruction Algorithms

Having discussed the conventional tomographic acquisition, we would like to briefly mention practical algorithms to reconstruct the object if enough measurements are available. In this subsection, we also demonstrate the failure of these algorithms on few 3D synthetic examples for sparse-angle and single-shot tomography.

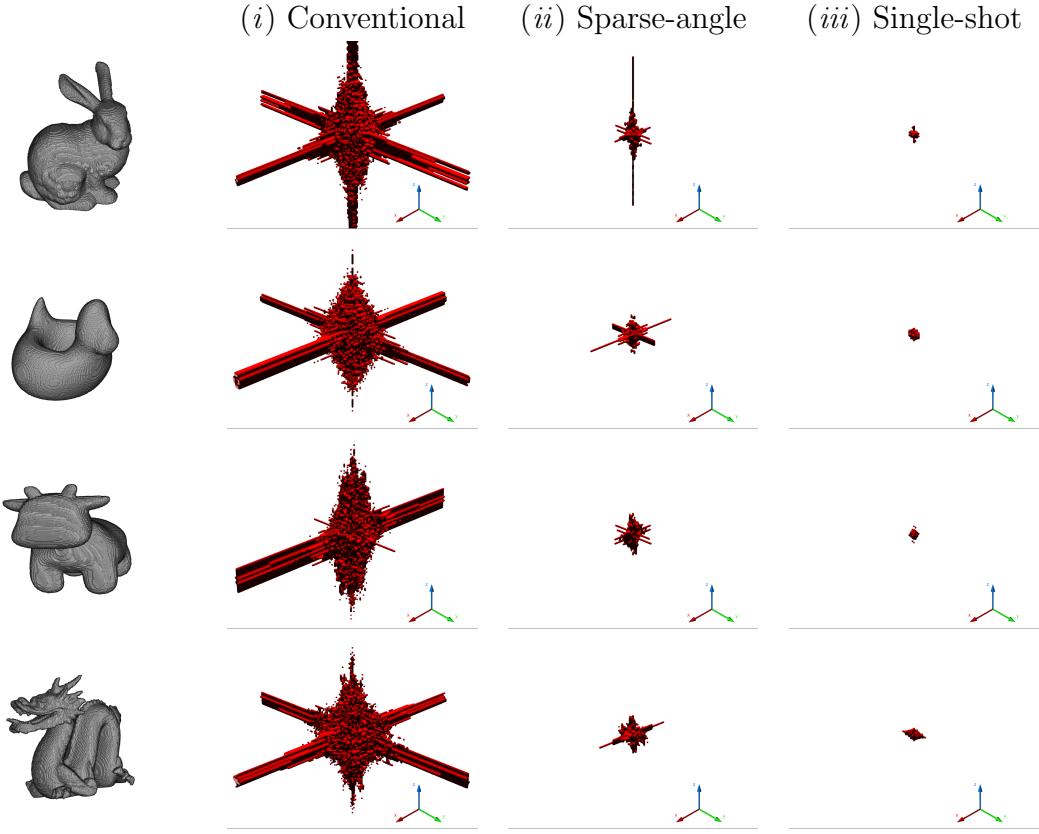


Figure 5. Frequency content information of four phantoms for three acquisitions, (i) conventional, (ii) sparse-angle, (iii) single-shot. These 3D frequency plots show the region that has strength more than the threshold. We select a threshold value of 0.1% of the maximum energy.

For conventional tomography, exact 3D reconstruction algorithms are usually based on the 3D Radon transform, where a Radon value (*i.e.*, set of tomographic measurements) is a plane integral in the object domain. From the projection-slice theorem, each plane can be represented by a unique intersection point of the plane and its normal passing the origin[25]. Radon values at these points constitute the 3D Radon space. Therefore, computing an exact reconstruction requires that all Radon values must be known. A sufficient condition for the exact reconstruction proposed by Tuy and Smith states that all planes intersecting the object must also intersect the source trajectory at least once [26, 27]. However, the circular trajectory does not satisfy this condition[28]. Hence, we resort to computationally cheap approximation algorithms, such as the FDK method[29].

For the sake of completeness, we describe the FDK method briefly. Assuming the measurements are obtained from a planar detector, the algorithm relies on filtering the measurements $y(\beta, \mathbf{s})$ to obtain

$$\hat{y}(\beta, \mathbf{s}) = (\cos \gamma \cos \kappa y(\beta, \mathbf{s})) * g(s_1)$$

where β is a projection angle, $\mathbf{s} = (s_1, s_2)$ are the local coordinates on the planar

detector, R is a source trajectory radius, γ and κ are fan-angle and cone-angle respectively, $g(t) = (1/2) \int |\rho| \exp(j2\pi\rho t) d\rho$ is often called a ramp filter due to its shape in the Fourier domain, and $*$ denotes the convolution operation. These filtered measurements produce voxel intensity at position $\mathbf{r} = (r_1, r_2, r_3)$ given by

$$x_{\text{FDK}}(\mathbf{r}) = \int_{-\pi}^{\pi} \frac{1}{U(r_1, r_2, \beta)^2} \hat{y}(\beta, s_1(r_1, r_2, \beta), s_2(\mathbf{r}, \beta)) d\beta$$

where the coordinates on the planar detector and the weighting factor U takes the following form

$$\begin{aligned} s_1(r_1, r_2, \beta) &= \frac{-r_1 \sin \beta + r_2 \cos \beta}{U(r_1, r_2, \beta)}, & s_2(\mathbf{r}, \beta) &= \frac{r_3}{U(r_1, r_2, \beta)}, \\ U(r_1, r_2, \beta) &= 1 + \left(\frac{r_1}{R}\right) \cos \beta + \left(\frac{r_2}{R}\right) \sin \beta. \end{aligned}$$

Similarly, the FDK method can be derived for the fan-beam projection in a 2D case. It is important to note that the FDK reconstruction closely approximates the true target image only when measurements are obtained for a sufficiently large number of projection angles $\beta \in [-\pi, \pi]$. If the number of projection angles is small (*i.e.*, sparse-angle tomography), the FDK method does not lead to a good solution. Figure 6 shows the reconstruction of four phantoms with the FDK algorithm for three acquisition scenario. As the number of projection angles is reduced, the FDK algorithm can no longer retrieve the target image successfully.

For sparse-angle tomography, one needs to regularize the problem by adding the appropriate prior information about the target image. The resulting regularized optimization problem can be solved using state-of-the-art iterative methods. Most common reconstruction algorithms for sparse-angle tomography solves the problem

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} \quad \|\mathbf{Ax} - \mathbf{y}\|^2 + \lambda \mathcal{R}(\mathbf{x})$$

where $\mathbf{y} \in \mathbb{R}^m$ are tomographic measurements sampled at m points, $\mathbf{A} \in \mathbb{R}^{m \times n}$ represents the matrix that approximates the cone-beam transform, $\mathbf{x} \in \mathbb{R}^n$ is the target image sampled at regular n points in the domain, and $\mathcal{R} : \mathbb{R}^n \mapsto \mathbb{R}$ is a regularization function (mostly convex) with λ balancing its importance with respect to the data-fidelity term. The most common regularization techniques are

- (i) *Tikhonov regularization* ($\mathcal{R}(\mathbf{x}) = \|\mathbf{x}\|^2$): It assumes smoothness over model parameters. The resulting problem is smooth and convex. We can obtain the solution to the regularized problem using Krylov methods[30].
- (ii) ℓ_1 *recovery* ($\mathcal{R}(\mathbf{x}) = \|\mathbf{\Gamma}\mathbf{x}\|_1$ with $\mathbf{\Gamma}$ being an orthogonal matrix, *e.g.*, Wavelet, Fourier, *etc.*): It promotes *sparsity* in either image space or transformed space. ℓ_1 recovery problems are convex but non-smooth. However, the solution can be retrieved with advanced optimization techniques, such as proximal gradient methods[31].

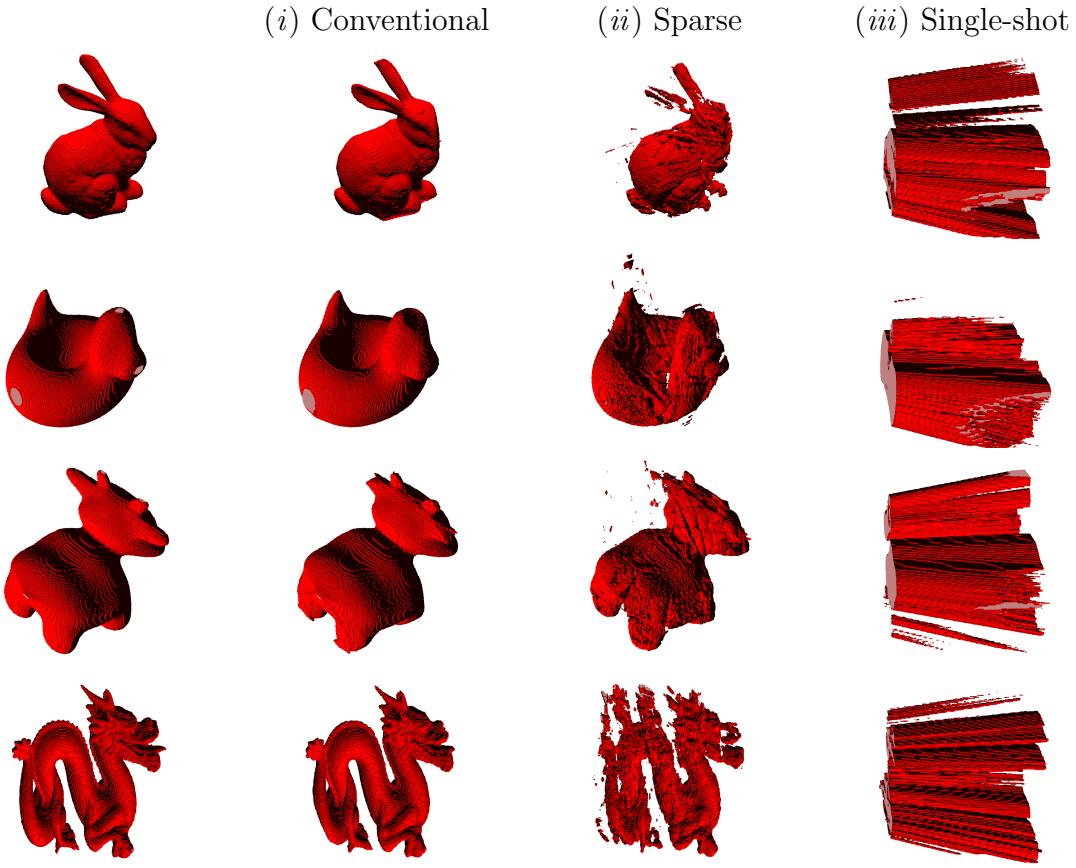


Figure 6. Demonstration of FDK algorithm on three tomographic acquisitions, (i) conventional, (ii) sparse-angle, and (iii) single-shot. These reconstructions are the segmented version of the actual FDK reconstructions.

- (iii) *Total-variation* ($\mathcal{R}(\mathbf{x}) = \|\nabla \mathbf{x}\|_1$ with ∇ denoting the spatial gradient): It retrieves a piecewise-constant approximation of the target image. The regularized problem is non-smooth but convex. However, Douglas-Rachford-type splitting techniques can obtain an optimal solution to this problem[32, 33].
- (iv) *Discrete Tomography* ($\mathcal{R}(\mathbf{x}) = \mathcal{I}_{\{0,1\}}(\mathbf{x})$ with \mathcal{I} being an indicator function that is 0 when x_i is either 0 or 1, and ∞ otherwise). This discrete prior is useful when the image consists of only two grey values that are already known. The resulting problem is NP-hard and can be solved either using heuristic techniques such as DART [7], TVR-DART[34], or with convex programs [35].

Although these methods are successful for sparse-angle tomography, they fail to recover the target image with a single shot (see Figure 1 for instance). The primary reason for the failure of these advanced reconstruction techniques is that the information contained in a single shot is not enough to fulfil the complete information required by the regularized inverse problem to produce a complete 3D image.

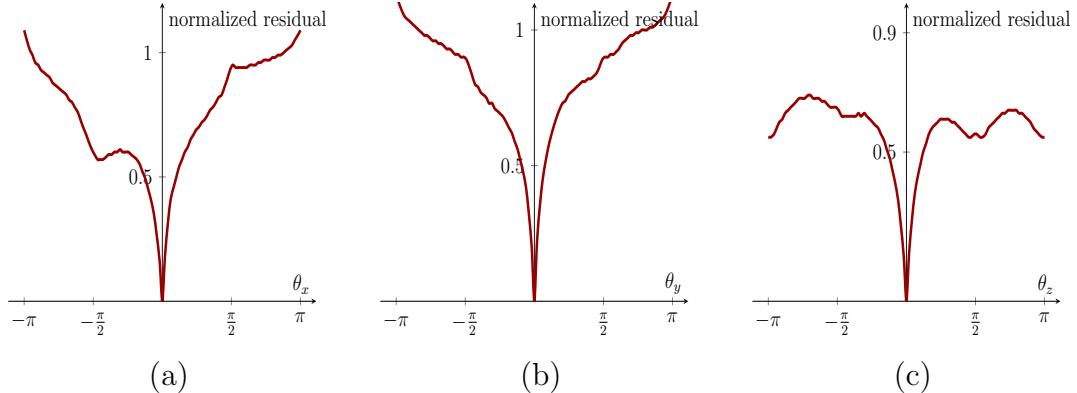


Figure 7. Misfit functional landscape for Bunny phantom with single-shot tomography for three rotation angles.

2.3. Tomographic Shape Sensing

Having noted that the failure of advanced regularization techniques on single-shot tomography, we discuss a special prior where the basic shapes in the target image are known to the user. The knowledge of shape (*i.e.*, its 3D structure and spatial intensity profile) will transform the problem of estimating a target image from single-shot projections to finding the rotation and translation parameters of the shape from its single projection. We refer to this problem as *Tomographic Shape Sensing*. Such advanced prior will be available when the target image is made up of building blocks. The information about the building blocks is available either through previous experiments or described by physical/material sciences.

Let us consider a target image composed of a single shape denoted by a function $u : \mathbb{R}^d \mapsto \mathbb{R}_+$. The rotated and translated version of this shape mathematically reads

$$\mathcal{T}_{\mathbf{R}, \mathbf{s}}(u) = \left\{ \int u(\mathbf{r}') \delta(\mathbf{r}_i - \mathbf{R}\mathbf{r}' - \mathbf{s}) d\mathbf{r}' \mid \mathbf{r}_i \in \Omega \forall i = 1, \dots, n \right\}, \quad (1)$$

where $\mathbf{R} \in \mathbb{R}^{d \times d}$ is a rotation matrix, $\mathbf{s} \in \mathbb{R}^d$ is a translation vector, and $\delta : \mathbb{R}^d \mapsto \mathbb{R}$ is a Dirac-delta function. Here, we have discretized the image on the n -point grid in the image space $\Omega \subset \mathbb{R}^d$. Hence, the tomography problem, where we were estimating the image function $x : \mathbb{R}^d \mapsto \mathbb{R}$ from m tomographic measurements $\mathbf{y} \in \mathbb{R}^m$, can now be rephrased as

$$\text{find } \mathbf{R}, \mathbf{s} \quad \text{subject to} \quad A\mathcal{T}_{\mathbf{R}, \mathbf{s}}(\mathbf{u}) \approx \mathbf{y}, \quad (2)$$

where \mathbf{A} is a tomographic matrix with multiple source positions, and \mathbf{y} contains the corresponding tomographic measurements. This is a mathematical formulation of *tomographic shape sensing* with a single shape. However, it is relatively straightforward to derive the formulation for multiple shapes with their multiple roto-translated copies.

Equation (2) is inherently a non-linear problem due to the implicit relationship between the image and the roto-translation parameters (refer to eq. (1)). Moreover, this non-linearity also leads to a non-convex program. To see this, we perform a numerical

experiment on a bunny phantom. Assuming the translation vector $\mathbf{s} \in \mathbb{R}^3$ is known, we try to estimate the rotation matrix $\mathbf{R} \in \mathbb{R}^{3 \times 3}$. The rotation matrices in 3D take the following form:

$$\mathbf{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}, \mathbf{R}_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}, \mathbf{R}_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

with a general rotation matrix represented by $\mathbf{R}(\alpha, \beta, \gamma) = \mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\gamma)$ in terms of Eulers angles $\alpha, \beta, \gamma \in [-\pi, \pi]$. In Figure 7, we plot the misfit landscape for three euclidean rotation angles. For every plot, we assume that the other two rotation angles are fixed and correct. From these plots, it is evident that the misfit function $\|\mathbf{A}\mathcal{T}_{\mathbf{R}, \mathbf{s}}(\mathbf{u}) - \mathbf{y}\|$ is highly non-convex. However, the rotational angles are approximately known, then the actual rotation angles can be recovered using a steepest-descent algorithm. Although the formulation (2) for tomographic shape sensing takes shape prior into account, it leads to a computational non-tractable problem due to the non-convex misfit landscape. In the next section, we discuss an efficient modelling approach to make the tomographic shape sensing a linear inverse problem.

3. Convex Shape Recovery

In this section, we discuss the tomographic shape sensing problem. Our formulation relies on the formation of a dictionary that consists of possible roto-translations of the known shapes and the representation of target image as a convex combination of dictionary elements.

3.1. Image model and Dictionary

Let the functions $u_i : \Omega \mapsto \mathbb{R}$, $i = 1, \dots, S$, denote the compactly supported shape functions. The image is now assumed to be composed of roto-translations of these shapes

$$x(\mathbf{r}) = \sum_{i=1}^S \sum_{j=1}^{k_i} u_i(\mathbf{R}(\boldsymbol{\theta}_{i,j}) \mathbf{r} + \mathbf{s}_{i,j})$$

where $\boldsymbol{\theta}_{i,j} \in \mathbb{R}^{d(d-1)/2}$ and $\mathbf{s}_{i,j} \in \mathbb{R}^d$ are the angle and the shift of j^{th} copy of shape i respectively, and $\mathbf{R} \in \mathbb{R}^{d \times d}$ is a rotation matrix that depends on the angle vector $\boldsymbol{\theta}$. The total number of shapes in an image are $K = k_1 + \dots + k_S$. Hence, from the knowledge of the shapes, the image estimation translates to finding the roto-translation parameters $(\boldsymbol{\theta}, \mathbf{s})$ of the shapes. However, the image is a non-linear function of these parameters. Hence, the recovery of these parameters becomes a computationally intractable problem due to the non-convex structure of the cost function.

To mitigate the non-linearity associated with the roto-translation parameters, we create a *shape* dictionary that consists of roto-translations of the shapes. Let the

dictionary

$$\begin{aligned} \Psi(\mathbf{r}) &= \left[\widehat{\Psi}_1(\mathbf{r}), \dots, \widehat{\Psi}_S(\mathbf{r}) \right], \\ \text{with } \widehat{\Psi}_i(\mathbf{r}) &= \left[u_i(\mathbf{R}(\boldsymbol{\theta}_{i,j})\mathbf{r} + \mathbf{s}_{i,j}) \right]_{j=1}^J, \quad i = 1, \dots, S, \end{aligned} \tag{3}$$

where $j = 1, \dots, J$ covers possible roto-translations of the shapes. Hence, the target image can be represented as a linear combination of the elements of this dictionary,

$$\begin{aligned} x(\mathbf{r}) &= \sum_{i=1}^p z_i \psi_i(\mathbf{r}) \\ \text{with } z_i &\in \{0, 1\}, \quad i = 1, \dots, p, \end{aligned}$$

where $\mathbf{z} = [z_1, \dots, z_p]^T$ is a coefficient vector, with $p = JS$. Hence, the shape recovery problem is to find a high-dimensional binary vector \mathbf{z} from its linear measurements

$$\mathbf{y} = \mathbf{A}\Psi\mathbf{z}.$$

Here, $\mathbf{A}\Psi$ contains the projections of the individual dictionary elements, sampled at the appropriate points. The binary constraints on \mathbf{z} make the recovery problem an integer program, and hence, NP-hard in general [36].

3.2. Convex Shape Recovery Program (CoShaRP)

The binary constraints on the coefficients can be relaxed using the bounds constraints. Moreover, a Gaussian assumption on the noise leads to a least-squares formulation for the data misfit. Hence, the resulting convex program, which we refer to as the *Convex Shape Recovery Program* (CoShaRP), reads

$$\begin{aligned} \underset{\mathbf{z} \in \mathbb{R}^p}{\text{minimize}} \quad & \|\mathbf{A}\Psi\mathbf{z} - \mathbf{y}\| \\ \text{subject to} \quad & \mathbf{z}^T \mathbf{1} = K, \quad \mathbf{0} \leq \mathbf{z} \leq \mathbf{1}. \end{aligned} \tag{4}$$

Here, the inequality between vectors is imposed elementwise. Note that we have used the Euclidean norm instead of its square to measure the misfit.

The geometric interpretation of CoShaRP is as follows: We are trying to find a high-dimensional vector \mathbf{z} closest to the hyperplane $\mathbf{A}\Psi\mathbf{z} = \mathbf{y}$ in a Euclidean sense that lies on the intersection of a hyperplane $\mathbf{z}^T \mathbf{1} = K$ and the hyperbox $\mathbf{0} \leq \mathbf{z} \leq \mathbf{1}$. In Figure 8, we show the geometry for a shape-sensing problem with two possible shapes ($p = 2$). In Figure 8(a), the hyperplane corresponding to tomographic measurements intersects the hyperplane corresponds to equality constraints ($\mathbf{z}^T \mathbf{1} = K$) outside the hyperbox. Hence, the solution to CoShaRP in this case is binary. However, a binary solution can not always be guaranteed as these hyperplanes may intersect inside the hyperbox (cf. Figure 8(b)). In such cases, further post-processing is required to retrieve the target image. For more, refer to Section 4.3.

The CoShaRP consists of constraints that are defined by K-simplex. The K-simplex, defined as

$$\Delta_p^K = \left\{ \mathbf{z} \in \mathbb{R}^p \mid \sum_{i=1}^p z_i = K, \mathbf{0} \leq \mathbf{z} \leq \mathbf{1} \right\},$$

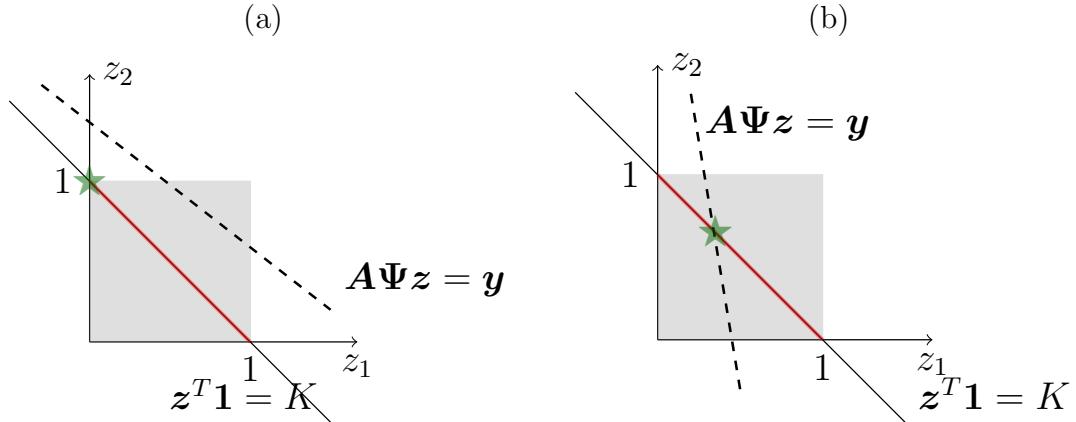


Figure 8. Geometry of CoShaRP. The grey region denotes the hyperbox that corresponds to bounds constraints($\mathbf{0} \leq \mathbf{z} \leq \mathbf{1}$). Solid line denotes the hyperplane for constraint on the number of shapes in the image, while dotted line denotes the hyperplane for measurement $\mathbf{A}\Psi\mathbf{z} = \mathbf{y}$. Note that the measurement hyperplanes do not pass through point $(0, 1)$ or $(1, 0)$ due to noise in the measurements. The star denotes the solution of CoShaRP. The figure (a) denotes the setup where coefficient is binary, while (b) with a non-binary solution.

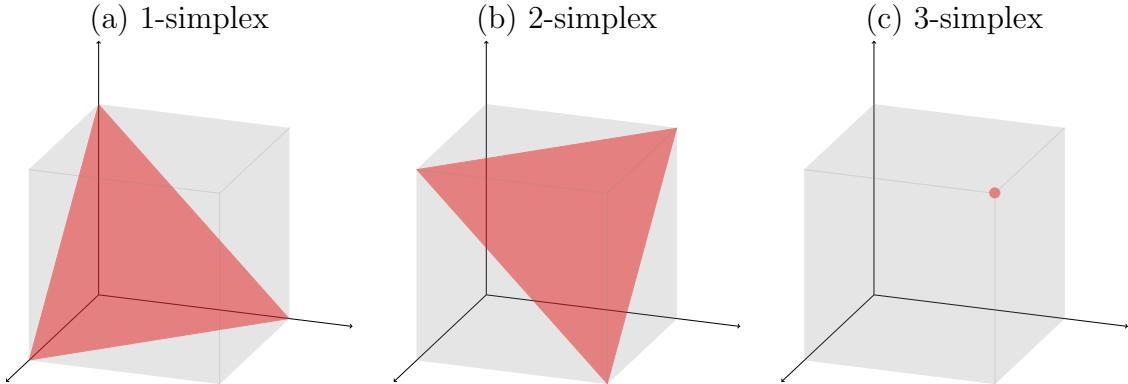


Figure 9. K-simplex in 3D. Barring $K = 3$ case where the K-simplex reduces to a point, the simplex are equilateral triangles (denoted by red color). The gray box denotes the bounds constraints.

is a generalized version simplex (simplex has $K = 1$). K-simplex represents a polytope in p -dimension with $\binom{p}{K}$ vertices. Moreover, these polytopes are *regular*, *i.e.*, they posses highest level of symmetry. We plot K -simplex in three dimension in Figure 9. These simplices are equilateral triangle except for $K = 3$. However, it is important to note that the number of shapes in the target image will be much smaller than the number of dictionary elements (*i.e.*, $K \ll p$). Hence, we will frequently encounter feasible regions to be extremely low-dimensional polytope embedded in a high-dimensional space.

Remark 1. *CoShaRP* differs significantly from the Sparse Shape Composition (SSC) [21, 23]. *SSC* formulates the shape-sensing problem as

$$\text{minimize } \frac{1}{2} \|\mathbf{A}\Psi\mathbf{z} - \mathbf{y}\|^2 \quad \text{subject to } \|\mathbf{z}\|_1 \leq K. \quad (5)$$

The ℓ_1 -norm ball is bigger in size than the K -simplex constraint set used in CoShaRP. In particular, the K -simplex constraint represents the strict ℓ_1 ball, i.e., $\|\mathbf{z}\|_1 = K$, in the non-negative region. Since the solution lies on the corners of the K -simplex constraint, the recovery with CoShaRP is stronger than that of SSC (see, e.g., Figure 1).

Remark 2. CoShaRP formulation is general, i.e., it extends beyond the single-shot tomography. Hence, CoShaRP is readily applicable for sparse-angle tomography as well as conventional tomography with noisy measurements or missing wedge.

Remark 3. Suppose the tomographic measurements are corrupted with Poisson noise of small strength. In that case, we can express CoShaRP in the following form:

$$\begin{aligned} & \underset{\mathbf{z} \in \mathbb{R}^p}{\text{minimize}} \quad \|\mathbf{A}\Psi\mathbf{z} - \mathbf{y}\|_{\mathbf{W}} \\ & \text{subject to} \quad \mathbf{z}^T \mathbf{1} = K, \quad \mathbf{0} \leq \mathbf{z} \leq \mathbf{1}, \end{aligned} \tag{6}$$

where \mathbf{W} is a diagonal matrix with elements $\{y_1, y_2, \dots, y_m\}$, and $\|\mathbf{x}\|_{\mathbf{W}} = \sqrt{\mathbf{x}^T \mathbf{W} \mathbf{x}}$ is a weighted ℓ_2 -norm. This formulation of CoShaRP is also a convex program since \mathbf{W} is a positive semi-definite matrix.

Remark 4. Suppose the target image consists of shapes that are a linear combination of roto-translated versions of other shapes present in the dictionary. In that case, the number of elements K for CoShaRP must be chosen such that the target image can be composed with the minimum number of elements from the shape dictionary Ψ .

4. Optimization

We discuss a fast iterative scheme to find an approximate solution of CoShaRP numerically. The iterative scheme is based on splitting strategy that separates the non-smooth part from the smooth. We also introduce a thresholding method to recover the image from the coefficient vector, in case the solution is not binary.

4.1. Primal-Dual Algorithm

For simplicity, we express CoShaRP in the following form:

$$\begin{aligned} & \underset{\mathbf{z}}{\text{minimize}} \quad f(\mathbf{A}\Psi\mathbf{z}) + g(\mathbf{z}), \\ & \text{where} \quad f(\mathbf{z}) = \|\mathbf{z} - \mathbf{y}\|, \quad g(\mathbf{z}) = \delta_{\mathcal{C}}(\mathbf{z}), \\ & \text{with} \quad \mathcal{C}(\mathbf{z}) = \left\{ \mathbf{z} \in \mathbb{R}^n \mid \mathbf{z}^T \mathbf{1} = K, \mathbf{0} \leq \mathbf{z} \leq \mathbf{1} \right\}, \end{aligned}$$

and $\delta_{\mathcal{C}}$ is the indicator function of the set \mathcal{C} . To solve this optimization problem, we use a primal-dual splitting algorithm [37, 38]. The iterates of this primal-dual algorithm takes the following form:

$$\begin{aligned} \mathbf{z}_{t+1} &= \text{prox}_{\gamma g} \left(\mathbf{z}_t - \gamma \Psi^T \mathbf{A}^T \mathbf{u}_t \right), \\ \mathbf{u}_{t+1} &= \text{prox}_{\tau f^*} \left(\mathbf{u}_t - \tau \mathbf{A}\Psi (\mathbf{z}_t - 2\mathbf{z}_{t+1}) \right), \end{aligned}$$

Algorithm 1 Primal-Dual Algorithm for CoShaRP**Input:** $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\Psi \in \mathbb{R}^{n \times p}$, $\mathbf{y} \in \mathbb{R}^m$, $\gamma, \tau, T, \epsilon$ **Output:** $\bar{\mathbf{z}}$

Initialisation : $\mathbf{z}_0, \mathbf{u}_0$

- 1: **for** $t = 0$ to T **do**
- 2: compute $\mathbf{z}_{t+1} = \text{prox}_{\gamma g}(\mathbf{z}_t - \gamma \Psi^T \mathbf{A}^T \mathbf{u}_t)$ using eq. (8)
- 3: compute $\mathbf{u}_{t+1} = \text{prox}_{\tau f^*}(\mathbf{u}_t - \tau \mathbf{A} \Psi(\mathbf{z}_t - 2\mathbf{z}_{t+1}))$ using eq. (7)
- 4: **if** $\|\mathbf{A} \Psi \mathbf{z}_t - \mathbf{y}\| \leq \epsilon$ **then**
- 5: **return** $\bar{\mathbf{z}} = \mathbf{z}_{t+1}$
- 6: **end if**
- 7: **end for**
- 8: **return** $\bar{\mathbf{z}} = \mathbf{z}_T$

for $t = [0, 1, \dots, T]$, where $\gamma, \tau > 0$, with $\gamma\tau \leq \|\mathbf{A}\Psi\|^{-1}$, are parameters that controls the speed of convergence. The main characteristic of this algorithm is that we avoid an inversion of a large matrix Ψ which often occurs in other splitting methods such as alternating direction method of multipliers [33]. Moreover, the proximal of both functions are easy to compute.

Remark 5. *The proposed primal-dual algorithm does not require the user to store a large dictionary matrix Ψ and tomography matrix \mathbf{A} . Hence, the algorithm can utilize the functional forms of the dictionary as well as tomography operator since it only requires the forward and the adjoint operation with the operator. However, when the problem size is small, or the dictionary consists of a small number of components compared to the image size, it is advisable to pre-multiply the tomography matrix \mathbf{A} with dictionary matrix Ψ and use the resulting sparse matrix $\mathbf{A}\Psi$ for faster computations.*

4.2. Proximal operators

The conjugate function of $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\|$ is

$$\begin{aligned} f^*(\mathbf{w}) &= -\inf_{\mathbf{x}} \{f(\mathbf{x}) - \mathbf{x}^T \mathbf{w}\} \\ &= \begin{cases} \mathbf{w}^T \mathbf{y} & \text{if } \|\mathbf{w}\| \leq 1 \\ +\infty & \text{otherwise} \end{cases} \end{aligned}$$

We refer to [39, Example 3.26] for the derivation of the conjugate function. The conjugate function is linear inside the Euclidean norm ball of size 1 and ∞ outside. Hence, the conjugate function is convex. Its proximal operator is given in the following theorem.

Theorem 1. *The proximal operator of function*

$$f^*(\mathbf{x}) = \begin{cases} \mathbf{y}^T \mathbf{x} & \|\mathbf{x}\| \leq 1 \\ +\infty & \text{otherwise} \end{cases},$$

where $\mathbf{y} \in \mathbb{R}^n$ is a known vector, is given by

$$\text{prox}_{\gamma f^*}(\mathbf{z}) = \frac{\mathbf{z} - \gamma \mathbf{y}}{\max(1, \|\mathbf{z} - \gamma \mathbf{y}\|)}. \quad (7)$$

Proof. The proximal operator for function f^* reads

$$\begin{aligned} \text{prox}_{\gamma f^*}(\mathbf{z}) &= \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{z}\|^2 + \mathbf{x}^T \mathbf{y} : \|\mathbf{x}\| \leq 1 \right\}, \\ &= \underset{\|\mathbf{x}\| \leq 1}{\operatorname{argmin}} \left\{ \|\mathbf{x} - (\mathbf{z} - \gamma \mathbf{y})\|^2 - \gamma^2 \|\mathbf{y}\|^2 + 2\gamma \mathbf{z}^T \mathbf{y} \right\}. \end{aligned}$$

The Euclidean norm constraints enforces two cases: (i) The optimal point without the constraints is $\bar{\mathbf{x}} = \mathbf{z} - \gamma \mathbf{y}$. This optimal solution holds when $\|\bar{\mathbf{x}}\| \leq 1$. (ii) When $\|\mathbf{z} - \gamma \mathbf{y}\| > 1$, the optimal solution lies on the surface of the Euclidean norm ball with size 1. Moreover, the optimal solution is in the direction of $\mathbf{z} - \gamma \mathbf{y}$. Hence, the proximal point is

$$\text{prox}_{\gamma f^*}(\mathbf{z}) = \begin{cases} \mathbf{z} - \gamma \mathbf{y} & \|\mathbf{z} - \gamma \mathbf{y}\| \leq 1 \\ \frac{\mathbf{z} - \gamma \mathbf{y}}{\|\mathbf{z} - \gamma \mathbf{y}\|} & \text{otherwise} \end{cases}.$$

This concludes the proof. \square

We use the following theorem to compute the proximal of $g(\mathbf{z}) = \delta_C(\mathbf{z})$, adapted from [40, Theorem 6.27].

Theorem 2 (projection onto the intersection of a hyperplane and a box). *Let $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{1} = K, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$ be a set. The proximal operator of an indicator function to the set, δ_C , is given by*

$$\text{prox}_{\gamma \delta_C}(\mathbf{x}) = \mathcal{P}_{[\mathbf{0}, \mathbf{1}]}(\mathbf{x} - \bar{\mu} \mathbf{1}) \quad (8)$$

where $\mathcal{P}_{[\mathbf{0}, \mathbf{1}]}$ is a projection onto the box $[0, 1]^n$ and $\bar{\mu}$ is a solution of the equation $\mathbf{1}^T \mathcal{P}_{[\mathbf{0}, \mathbf{1}]}(\mathbf{x} - \mu \mathbf{1}) = K$.

Proof. The orthogonal projection of \mathbf{x} is the unique solution of

$$\min_{\mathbf{z} \in \mathbb{R}^n} \left\{ \frac{1}{2} \|\mathbf{z} - \mathbf{x}\|^2 : \mathbf{1}^T \mathbf{z} = K, \mathbf{0} \leq \mathbf{z} \leq \mathbf{1} \right\}. \quad (9)$$

A Lagrangian of this minimization problem reads

$$\mathcal{L}(\mathbf{z}, \mu) = \frac{1}{2} \|\mathbf{z} - \mathbf{x}\|^2 + \mu (\mathbf{1}^T \mathbf{z} - K),$$

where $\mu \in \mathbb{R}$ is a Lagrange multiplier. It follows from the strong duality that $\bar{\mathbf{z}}$ is an optimal solution of problem (9) if and only if there exists a dual variable $\mu \in \mathbb{R}$ for which

$$\mathbf{y}^* \in \underset{0 \leq \mathbf{z} \leq \mathbf{1}}{\operatorname{argmin}} \quad \mathcal{L}(\mathbf{z}, \bar{\mu}), \quad (10)$$

$$\mathbf{1}^T \mathbf{z} = K. \quad (11)$$

Using the expression for the Lagrangian, the relation (10) can be equivalently written as $\bar{\mathbf{z}} = \mathcal{P}_{[\mathbf{0}, \mathbf{1}]}(\mathbf{x} - \bar{\mu} \mathbf{1})$. The feasibility condition (11) takes a form $\mathbf{1}^T \mathcal{P}_{[\mathbf{0}, \mathbf{1}]}(\mathbf{x} - \bar{\mu} \mathbf{1}) = K$. \square

Algorithm 2 Image formation from shape coefficients**Input:** $\mathbf{z} \in \mathbb{R}^p$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\Psi \in \mathbb{R}^{n \times p}$, $\mathbf{y} \in \mathbb{R}^m$, K **Output:** $\bar{\mathbf{x}}$

Initialisation : $\bar{\mathbf{x}} = \mathbf{0}$, $i = 0$, $s = 0$

- 1: sort the elements of \mathbf{z} in the descending order, and store its indices as a list T
- 2: **while** ($s \leq K$) or ($i \leq p$) **do**
- 3: $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \Psi e_{T(i)}$
- 4: **if** $\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\| \leq \|\mathbf{A}\bar{\mathbf{x}} - \mathbf{y}\|$ and $\bar{\mathbf{x}}^T \Psi e_{T(i)} \leq 0$ **then**
- 5: $\bar{\mathbf{x}} = \hat{\mathbf{x}}$, $s = s + 1$
- 6: **end if**
- 7: $i = i + 1$
- 8: **end while**
- 9: **return** $\bar{\mathbf{x}}$

Remark 6. The projection onto the box $[0, 1]$ is simple. It is done component-wise as $(\min\{\max\{x_i, 0\}, 1\})_{i=1}^n$. However, equation (8) consists of finding a root of the non-increasing function $\phi(\mu) = \sum_{i=1}^n \min\{\max\{x_i - \mu, 0\}, 1\} - K$. Since $\mu \mapsto \min\{\max\{x_i - \mu, 0\}, 1\}$ is a non-increasing function for any i , ϕ is a non-increasing function. Its root can be found using the Newton procedure, where derivative is

$$\phi'(\mu) = |\mathcal{I}|, \quad \text{with } \mathcal{I} = \{i \in \{1, \dots, n\} : 0 \leq x_i - \mu \leq 1\}.$$

4.3. Image Formation

The convex program CoShaRP does not always lead to a binary solution (refer to Figure 8). Moreover, if the optimization procedure is terminated early, we may not have a binary solution. Hence, an accurate image formation process is essential to retrieve the target image from the non-binary shape coefficient vector resulting from CoShaRP. We propose the image formation procedure based on sorting of the coefficients. We first sort the coefficients in descending order, and selectively form the image consistent with the measurements. Algorithm 2 enumerates the steps in the image formation process. Here, e_j is a natural basis vector with non-zero element located at j^{th} position. To make sure the shapes do not overlap, we also add necessary conditions (see step 4 in Algorithm 2).

5. 2D Numerical Experiments

In this section, we try to answer questions regarding resolution, sparsity, rotations and the performance under noise using 2D numerical experiments. For all the experiments, we have images of size $1 m \times 1 m$ discretized on 128×128 pixels, and the tomography matrix has at least 1024 measurements. The typical tomography setup is shown in Figure 10. In the CoShaRP performance plots, we generate 100 different realizations of

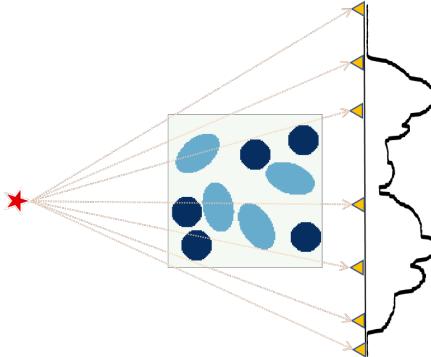


Figure 10. Single-shot tomography setup. The source (denoted by \star) is located at the left of the target image. The detector array consisting of 1024 detectors is located at the right of the image. For reference, we only show seven detectors (denoted by \triangledown). The measurement profile is shown on the right of the detector array.

the target image with given constraints (for examples, size, rotations and repetitions of the shape), and the success rate is measured from the average over all instances. We say an instance is successful if the recovered image is close to the target image (in Euclidean norm).

For all the experiments, we use ASTRA toolbox [41], and run Algorithm 1 with $\gamma = 1.2\sigma$, $\tau = 0.8\sigma$ with $\sigma = \|\mathbf{A}\Psi\|^{-1}$. Moreover, we set $T = 4p^2$ and $\epsilon = 10^{-6}$. Once we obtain the vector $\bar{\mathbf{z}}$, we run Algorithm 2 to form the image.

5.1. Resolution Analysis

In this experiment, we estimate the required minimum width of the shape sensed by a single-shot. For simplicity, we consider circular disc of constant intensity with size varying from 1 to 3168 pixels. Figure 11(a) shows the performance of CoShaRP against varying sizes of the disc. As the number of measurements (*i.e.*, detector pixels) is increased, the success rate increases implying that the recovery of even single-pixel shapes is possible with CoShaRP.

5.2. Invariance with respect to Density and Rotation

We first look at the success of CoShaRP with multiple repetitions of the shape. We consider a circular disc with size 256 pixels and the number of repetitions in the image from 1 to 20. The top figure in Figure 11(b) shows the performance of CoShaRP with density. This shows that the CoShaRP is insensitive to the number of repetitions of the shapes in the image. Next, we take an ellipsoidal disc with semi-axes 0.2 m and 0.08 m. This image is rotated for 30 angles making sure that each angle represents a different shape on the 128×128 pixels. The bottom figure in Figure 11(b) demonstrates the performance of CoShaRP with the number of possible rotations. This implies that the CoShaRP is insensitive to the number of possible rotations of the shapes.

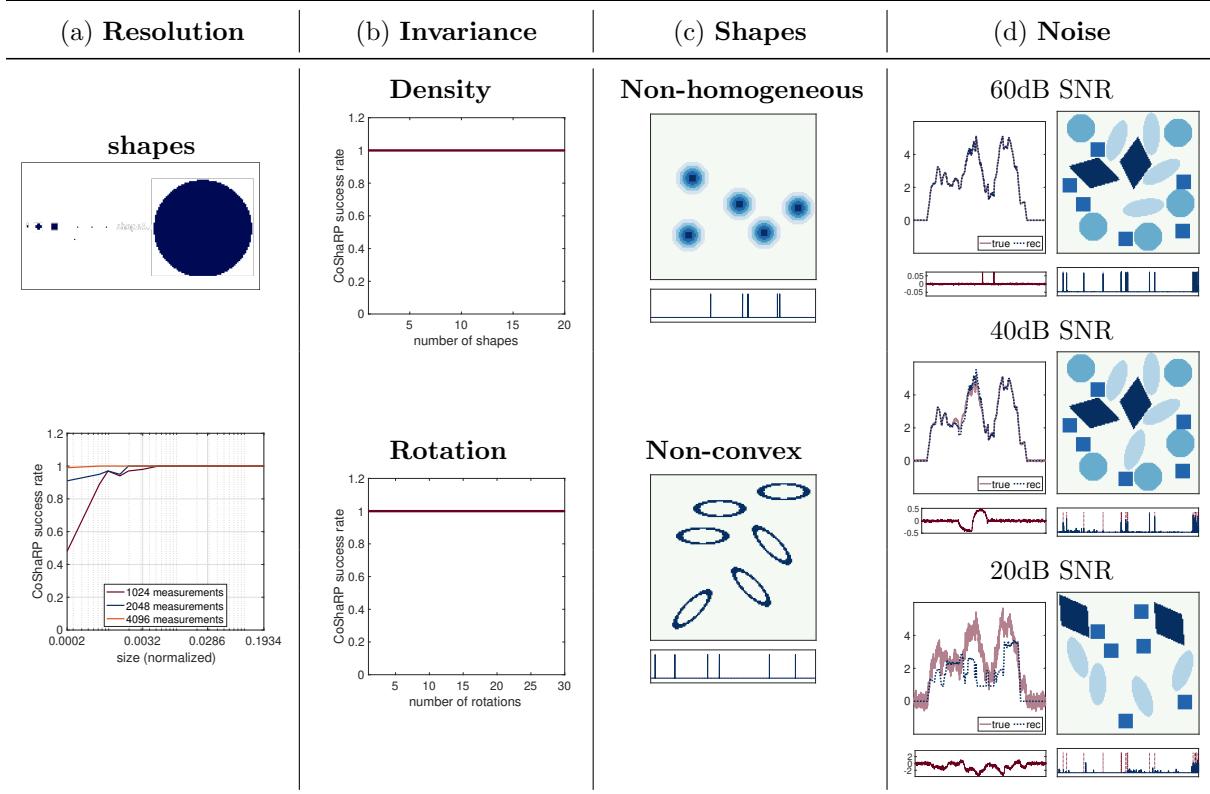


Figure 11. Numerical results of CoShaRP for single-shot tomography.

5.3. Non-homogeneous and Non-convex Shapes

The top figure in Figure 11(c) provides the reconstruction from CoShaRP for non-homogeneous shape. We consider a circular shape with four different intensities varying radially. The true image consists of 5 repetitions of this shape. The CoShaRP recovers these 5 copies successfully as shown from the shape coefficients (given below the figure). For non-convex shape, we consider ellipsoidal shell with outer axes 0.2 m and 0.05 m , and inner axes 0.15 m and 0.03 m . The bottom figure in Figure 11(c) gives the reconstruction with CoShaRP for a true image with 6 repetitions of the above-mentioned non-convex shape. From these two experiments, we conclude that the CoShaRP can recover the non-homogeneous and non-convex shapes.

5.4. Measurement Noise

For the true image shown in Figure 1, we consider three noisy scenarios where the Gaussian noise of strength 0.1%, 1% and 10% is added to 1024 measurements, resulting in signal-to-noise ratio (SNR) of 60dB, 40dB and 20dB respectively. In Figure 11(d), we plot the measurements (on the left) and the reconstructed image (on the right) for the above-mentioned noise values. In the measurements plots, the true noisy data is mentioned as ‘true’, while the forward projected data from the reconstructed image is denoted by ‘rec’. The plots below them show the difference between the two. From

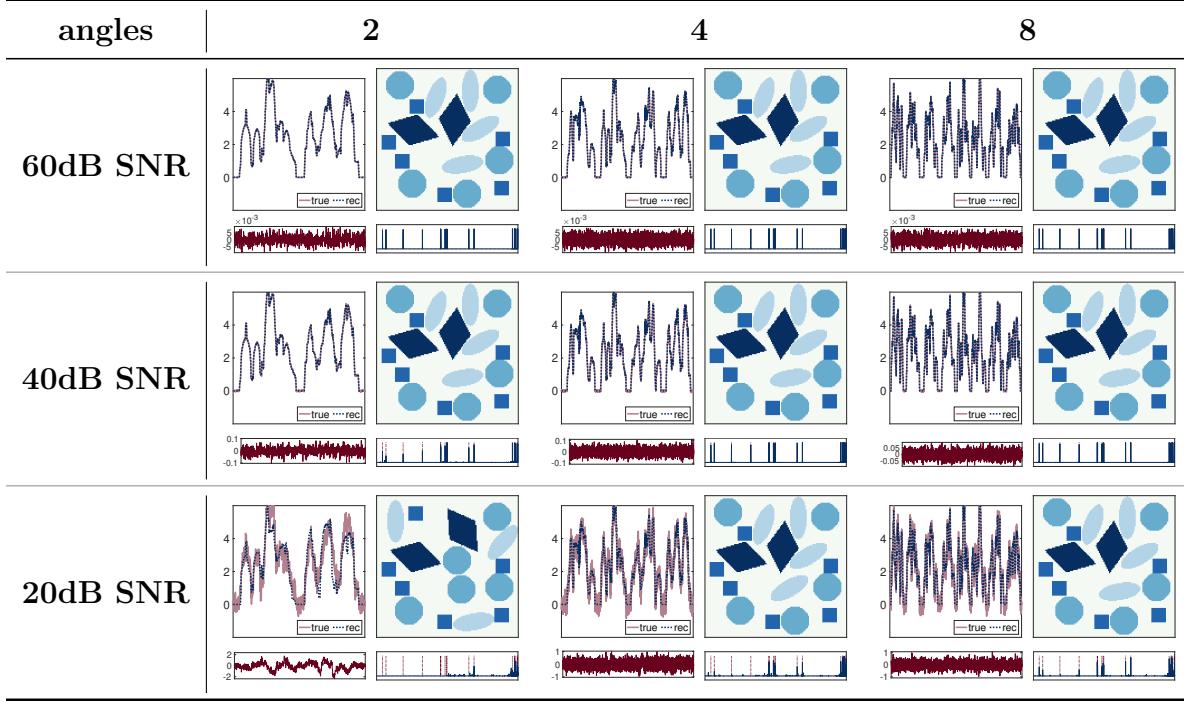


Figure 12. Demonstration of CoShaRP against Noise for sparse-angle tomography with 2-angles, 4-angles and 8-angles: Each element of the table consists of four figures: The top-left corresponds to the true projections and the forward projections from the CoShaRP image, the top-right corresponds to the reconstructed image from CoShaRP. The bottom-left shows the noise in the projections, and the bottom-right plots the Dictionary coefficients.

Figure 11(d), it is evident that the CoShaRP is stable till 1% noise in single-shot X-ray projection, while fails for extremely noisy measurements.

5.5. Stability against Noise for Sparse-Angle Tomography

To demonstrate the increment in the stability of CoShaRP against high noise with the increasing number of views, we consider the image shown in Figure 1. We mainly consider sparse-angle tomography with 2-angles ($\{0^\circ, 90^\circ\}$), 4-angles ($\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$), and 8-angles ($\{0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ\}$). We test CoShaRP with additive Gaussian noise of strength 0.1%, 1% and 10%, resulting in SNR of 60dB, 40dB and 20dB respectively. Figure 12 shows the results for these three sparse-angle settings with three different noise levels. For 60dB and 40dB SNR values, the results of CoShaRP for all three sparse-angle configurations are stable. For 20dB SNR, CoShaRP fails to reconstruct an appropriate target image for a 2-angles setting. However, for 4-angles and 8-angles setting, CoShaRP finds the true image. Hence, we can conclude from these numerical experiments that the CoShaRP becomes stable if projections for more than 4-angles are available.

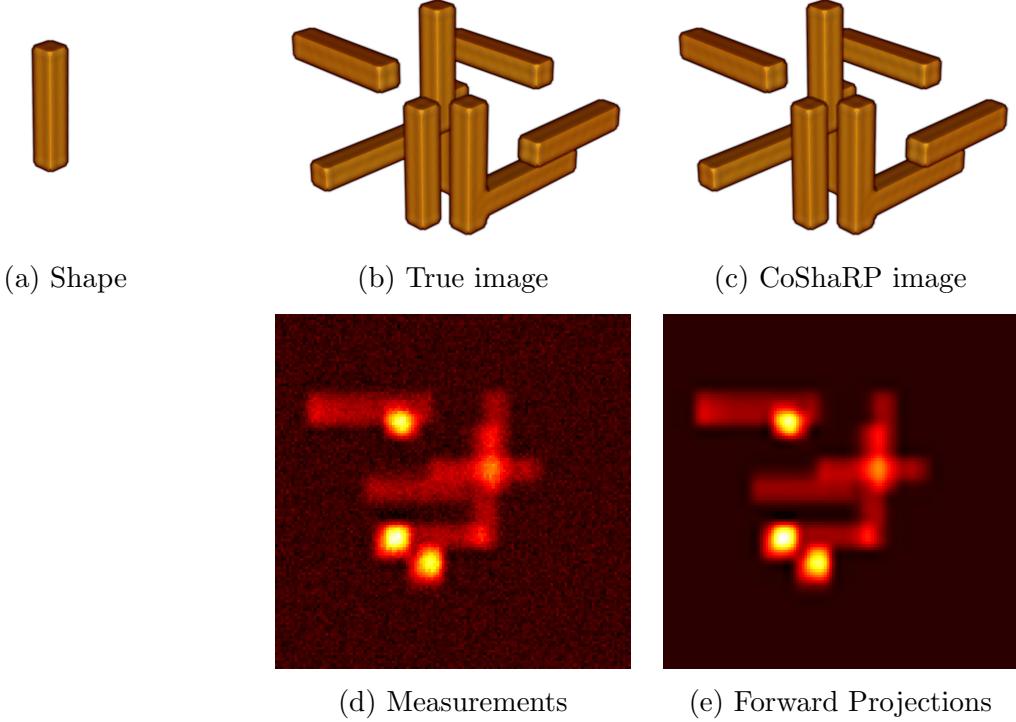


Figure 13. Numerical demonstration of CoShaRP on 3D Phantom composed of multiple roto-translated copies of a rectangular cylinder.

6. 3D Experiment

We quickly demonstrate the CoShaRP framework on a 3D image. Figure 13(a) shows the shape (a rectangular rod with thickness of $32 \times 8 \times 8$ voxels) used to create a target image of size $128 \times 128 \times 128$ shown in Figure 13(b). The shape has three possible rotations, and the target image consists of 9 copies of this shape. Figure 13(d) shows a cone-beam projection of this 3D image. These projection measurements of size 128×128 are corrupted with Gaussian noise of strength 0.1%, resulting in an SNR of 40dB. We run the CoShaRP with $K = 8$, and the resulting image is plotted in Figure 13(c). The results from CoShaRP matches precisely with the true image. Moreover, the forward projections from the CoShaRP image, shown in Figure 13(e), approximately resembles the true measurements.

7. Conclusions

We introduced a *single-shot tomographic shape sensing* problem that aims to recover shapes from a single cone-beam projection. To solve this problem, we develop a convex program CoShaRP. CoShaRP is novel in the sense that the simplex-type constraint enables sharp recovery results from extremely under-determined single-shot tomographic projections. Moreover, we propose a primal-dual algorithm to find an approximately optimal solution to CoShaRP quickly. Through numerical experiments, we demonstrate

that (*i*) the resolution limit to sense the shape depends on the number of measurements, (*ii*) CoShaRP is insensitive to the number of repetitions of the shape and the number of possible rotations of the shape, (*iii*) CoShaRP can sense the non-homogeneous as well as non-convex shapes, (*iv*) CoShaRP tolerates only a moderate amount of measurement noise.

The limitations of CoShaRP are as follows: (*i*) The roto-translations of the shapes must be included in the dictionary for the exact recovery of the target image. This inclusion requirement makes CoShaRP a computationally expensive approach due to the large dictionary size. (*ii*) CoShaRP also requires the correct knowledge of shapes and their intensity. If the shape is not known accurately, the CoShaRP may fail. (*iii*) CoShaRP relies on the knowledge of total number of shapes in the target image. If unknown, its estimation may be a costly procedure due to repeated solving of CoShaRP for various estimates.

Acknowledgments

The authors thank Nick Luiken for stimulating discussions. Moreover, we thank reviewers for substantially improving the manuscript. This work was supported by the Dutch Research Council (grant no. OCENW.XS.039).

8. References

- [1] M. Defrise and R. Clack. A cone-beam reconstruction algorithm using shift-variant filtering and cone-beam backprojection. *IEEE Transactions on Medical Imaging*, 13(1):186–195, March 1994.
- [2] Hiroyuki Kudo, Frédéric Noo, and Michel Defrise. Cone-beam filtered-backprojection algorithm for truncated helical data. *Physics in Medicine and Biology*, 43(10):2885–2909, October 1998.
- [3] Yu Zou and Xiaochuan Pan. Image reconstruction on PI-lines by use of filtered backprojection in helical cone-beam CT. *Physics in Medicine and Biology*, 49(12):2717–2731, June 2004.
- [4] A.H. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Transactions on Image Processing*, 7(2):204–221, 1998.
- [5] M Persson, D Bone, and H Elmqvist. Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Physics in Medicine and Biology*, 46(3):853–866, February 2001.
- [6] Emil Y Sidky and Xiaochuan Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine and Biology*, 53(17):4777–4807, August 2008.
- [7] K. J. Batenburg and J. Sijbers. DART: A practical reconstruction algorithm for discrete tomography. *IEEE Transactions on Image Processing*, 20(9):2542–2553, September 2011.
- [8] Jürgen Frikel. Sparse regularization in limited angle tomography. *Applied and Computational Harmonic Analysis*, 34(1):117–141, January 2013.
- [9] Max F. Hantke, Dirk Hasse, Filipe R. N. C. Maia, Tomas Ekeberg, Katja John, Martin Svenda, N. Duane Loh, Andrew V. Martin, Nicusor Timneanu, Daniel S. D. Larsson, Gijs van der Schot, Gunilla H. Carlsson, Margareta Ingelman, Jakob Andreasson, Daniel Westphal, Mengning Liang, Francesco Stellato, Daniel P. DePonte, Robert Hartmann, Nils Kimmel, Richard A. Kirian, M. Marvin Seibert, Kerstin Mühlig, Sebastian Schorb, Ken Ferguson, Christoph Bostedt, Sebastian Carron, John D. Bozek, Daniel Rolles, Artem Rudenko, Sascha Epp, Henry N.

- Chapman, Anton Barty, Janos Hajdu, and Inger Andersson. High-throughput imaging of heterogeneous cell organelles with an x-ray laser. *Nature Photonics*, 8(12):943–949, November 2014.
- [10] Daniël M. Pelt and Vincent De Andrade. Improved tomographic reconstruction of large-scale real-world data by filter optimization. *Advanced Structural and Chemical Imaging*, 2(1), December 2016.
 - [11] Nouamane Laanait, Wittawat Saenrang, Hua Zhou, Chang-Beom Eom, and Zhan Zhang. Dynamic x-ray diffraction imaging of the ferroelectric response in bismuth ferrite. *Advanced Structural and Chemical Imaging*, 3(1), March 2017.
 - [12] Tekin Bicer, Doga Gürsoy, Vincent De Andrade, Rajkumar Kettimuthu, William Scullin, Francesco De Carlo, and Ian T. Foster. Trace: a high-throughput tomographic reconstruction engine for large-scale datasets. *Advanced Structural and Chemical Imaging*, 3(1), January 2017.
 - [13] Luis F Alves Pereira, Eline Janssens, George DC Cavalcanti, Ren Tsang, Mattias Van Dael, Pieter Verboven, Bart Nicolai, and Jan Sijbers. Inline discrete tomography system: Application to agricultural product inspection. *Computers and electronics in agriculture*, 138:117–126, 2017.
 - [14] T Altantzis, D Zanaga, and S Bals. Advanced electron tomography of nanoparticle assemblies. *EPL (Europhysics Letters)*, 119(3):38001, 2017.
 - [15] Friedrich Förster, Ohad Medalia, Nathan Zauberman, Wolfgang Baumeister, and Deborah Fass. Retrovirus envelope protein complex structure in situ studied by cryo-electron tomography. *Proceedings of the National Academy of Sciences*, 102(13):4729–4734, 2005.
 - [16] J C H Spence, U Weierstall, and H N Chapman. X-ray lasers for structural and dynamic biology. *Reports on Progress in Physics*, 75(10):102601, September 2012.
 - [17] Wenjuan Zhu, Wenbo Ma, Yirong Su, Zeng Chen, Xinya Chen, Yaoguang Ma, Lizhong Bai, Wenge Xiao, Tianyu Liu, Haiming Zhu, Xiaofeng Liu, Huafeng Liu, Xu Liu, and Yang (Michael) Yang. Low-dose real-time x-ray imaging with nontoxic double perovskite scintillators. *Light: Science & Applications*, 9(1), June 2020.
 - [18] X. Zhang, E. Settembre, C. Xu, P. R. Dormitzer, R. Bellamy, S. C. Harrison, and N. Grigorieff. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proceedings of the National Academy of Sciences*, 105(6):1867–1872, January 2008.
 - [19] Sjors H.W. Scheres. RELION: Implementation of a bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, 180(3):519–530, December 2012.
 - [20] Stefan Pfeffer, Maryam Khoshouei, Radostin Danev, and Friedrich Förster. Towards high resolution in cryo-electron tomography subtomogram analysis. *Microscopy and Microanalysis*, 23(S1):812–813, 2017.
 - [21] Alireza Aghasi and Justin Romberg. Sparse shape reconstruction. *SIAM Journal on Imaging Sciences*, 6(4):2075–2108, January 2013.
 - [22] Alireza Aghasi and Justin Romberg. Convex cardinal shape composition. *SIAM Journal on Imaging Sciences*, 8(4):2887–2950, January 2015.
 - [23] Daniele Zanaga, Folkert Bleichrodt, Thomas Altantzis, Naomi Winckelmans, Willem Jan Palenstijn, Jan Sijbers, Bart de Nijs, Marijn A. van Huis, Ana Sánchez-Iglesias, Luis M. Liz-Marzán, Alfons van Blaaderen, K. Joost Batenburg, Sara Bals, and Gustaaf Van Tendeloo. Quantitative 3d analysis of huge nanoparticle assemblies. *Nanoscale*, 8(1):292–299, 2016.
 - [24] Patrick L. Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. In *Springer Optimization and Its Applications*, pages 185–212. Springer New York, 2011.
 - [25] Frank Natterer. *The mathematics of computerized tomography*. SIAM, 2001.
 - [26] Heang K Tuy. An inversion formula for cone-beam reconstruction. *SIAM Journal on Applied Mathematics*, 43(3):546–552, 1983.
 - [27] Bruce D Smith. Image reconstruction from cone-beam projections: necessary and sufficient conditions and reconstruction methods. *IEEE transactions on medical imaging*, 4(1):14–25, 1985.

- [28] Eric Todd Quinto. An introduction to x-ray tomography and radon transforms. In *Proceedings of symposia in Applied Mathematics*, volume 63, page 1, 2006.
- [29] Lee A Feldkamp, Lloyd C Davis, and James W Kress. Practical cone-beam algorithm. *Josa a*, 1(6):612–619, 1984.
- [30] Henk A Van der Vorst. *Iterative Krylov methods for large linear systems*. Number 13. Cambridge University Press, 2003.
- [31] Neal Parikh and Stephen Boyd. Proximal algorithms. *Foundations and Trends in optimization*, 1(3):127–239, 2014.
- [32] Jonathan Eckstein and Dimitri P Bertsekas. On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1):293–318, 1992.
- [33] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2010.
- [34] Xiaodong Zhuge, Willem Jan Palenstijn, and Kees Joost Batenburg. Tvr-dart: A more robust algorithm for discrete tomography from limited projection data with automated gray value estimation. *IEEE Transactions on Image Processing*, 25(1):455–468, 2015.
- [35] Ajinkya Kadu and Tristan van Leeuwen. A convex formulation for binary tomography. *IEEE Transactions on Computational Imaging*, 6:1–11, 2019.
- [36] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer US, 1972.
- [37] Xiaoqun Zhang, Martin Burger, and Stanley Osher. A unified primal-dual algorithm framework based on bregman iteration. *Journal of Scientific Computing*, 46(1):20–46, August 2010.
- [38] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, December 2010.
- [39] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [40] Amir Beck. *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, October 2017.
- [41] Folkert Bleichrodt, Tristan van Leeuwen, Willem Jan Palenstijn, Wim van Aarle, Jan Sijbers, and K Joost Batenburg. Easy implementation of advanced tomography algorithms using the astra toolbox with spot operators. *Numerical algorithms*, 71(3):673–697, 2016.