# Universiteit Antwerpen

CoShaRP : a convex program for single-shot tomographic shape sensing

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

**Ajinkya Kadu[1], Tristan van Leeuwen[2], Joost Batenburg[3]**

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

[2] Mathematical Institute, Utrecht University, Utrecht, The Netherlands

[3] 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. 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. In this paper, 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 constraint and can be solved quickly using a primal-dual algorithm. The numerical experiments show that CoShaRP recovers shapes stably from moderately noisy measurements.

## 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 straight-forward 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]. In this paper, we consider an extreme case where we acquire only a single X-ray projection on which to base a full three-dimensional reconstruction. We refer to this problem as *single-shot X-ray tomography*. This problem is highly
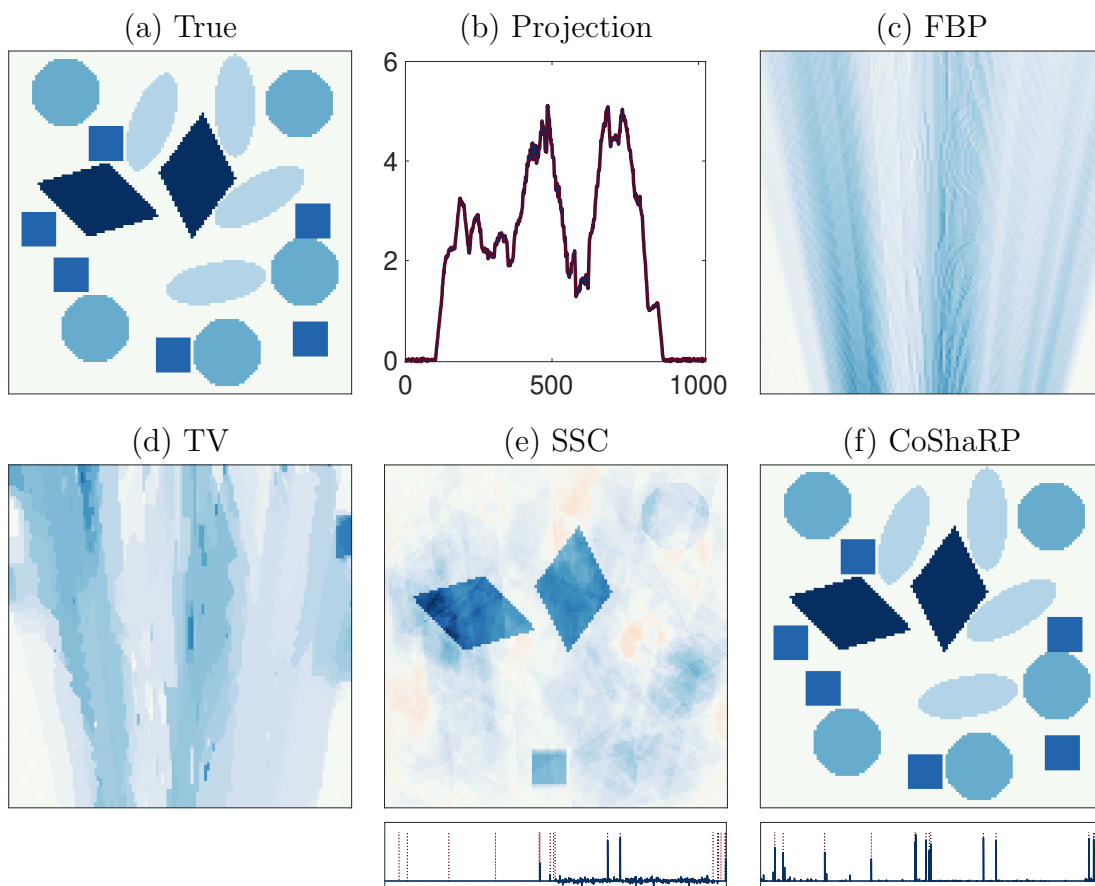
**Figure 1.** Demonstration of single-shot X-ray tomography. A target image (a) of size $1\ m \times 1\ m$ discretized on $128 \times 128$ pixel grid has a fan-beam projection shown in (c). The image consists of 4 different shapes with different intensities. The equally-spaced detectors place at the top of the image collect a total of 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.

relevant for many applications, including industrial quality control and high-throughput imaging [9, 10, 11, 12]. 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[13, 14] and cryo-electron microscopy [15, 16].

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) 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 an $\ell_1$-norm constraint on the shape-dictionary coefficients [17, 18]. 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 $\ell_1$-norm constraint in high dimension (see Remark 3.2). 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 [19]. This method uses a simple $\ell_1$-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 cone-beam tomography in Figure 1.

## 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 6.

### 1.3. Notation

Throughout this paper, small boldface letters (*e.g.*, $\boldsymbol{x}, \boldsymbol{z}$) denote vectors in $\mathbb{R}^n$. The identity and zero elements are denoted by $\boldsymbol{0}$ and $\boldsymbol{1}$ respectively. The Euclidean inner product is denoted by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ with a corresponding norm $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. However, for all other norms, we use subscripts (*e.g.*, $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$, $\|\boldsymbol{x}\|_\infty = \max_i |x_i|$). To represent the matrices, we use uppercase letters (*e.g.*, $\boldsymbol{A}, \boldsymbol{\Psi}$). The elements of a matrix $\boldsymbol{A}$ 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 $f^\star$. $\mathbf{prox}_f(\boldsymbol{z})$ denotes the proximal of function $f$ evaluated at point $\boldsymbol{z}$ (for definition, please refer to [20]). We represent the optimal solution to the optimization problem using overline (*e.g.*, $\overline{\boldsymbol{x}}, \overline{\mu}$).

## 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. Let $\boldsymbol{\varphi} \in \mathbb{S}^{d-1}$ be a directional vector, and $\boldsymbol{r} \in \mathbb{R}^d$ 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 $\boldsymbol{\varphi}$ passing through $\boldsymbol{r}$. It is mathematically given by

$$(A_C x)(\boldsymbol{r}, \boldsymbol{\varphi}) = \int_0^\infty x(\boldsymbol{r} + t\boldsymbol{\varphi}) \, \mathrm{d}t.$$

In a single-shot setup, we have a source located at $\boldsymbol{r}_0$. It sends multiple X-rays through the object (compactly supported on $\Omega \subset \mathbb{R}^d$) in a cone with a vertex at $\boldsymbol{r}_0$ and spanning angles in a set $\Phi$ that determines the geometry of cone. Let these angles be $\boldsymbol{\varphi}_i \in \Phi$, $i = 1, \ldots, m$, then the measurement $y_i$ is

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

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

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

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.

## 3. Convex Shape Recovery

In this section, we discuss the tomographic shape sensing problem. Our formulation hinges 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 $u_i : \Omega \mapsto \mathbb{R}$, $i = 1, \ldots, S$, denote the compactly supported shape functions. The image is now assumed to be composed of roto-translations of these shapes

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

where $\boldsymbol{\theta}_{i,j} \in \mathbb{R}^{d(d-1)/2}$ and $\boldsymbol{s}_{i,j} \in \mathbb{R}^d$ are the angle and the shift of $j^{th}$ copy of shape $i$ respectively, and $\boldsymbol{R} \in \mathbb{R}^{d \times d}$ is a rotation matrix that depends on the angle $\boldsymbol{\theta}$. The total number of shapes in an image are $K = k_1 + \cdots + k_S$. Hence, from the knowledge of the shapes, the image estimation translates to finding the roto-translation parameters $(\boldsymbol{\theta}, \boldsymbol{s})$ 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

$$\mathbf{\Psi}(\boldsymbol{r}) = \left[\widehat{\mathbf{\Psi}}_1(\boldsymbol{r}), \ldots, \widehat{\mathbf{\Psi}}_S(\boldsymbol{r})\right],$$

$$\text{with} \quad \widehat{\mathbf{\Psi}}_i(\boldsymbol{r}) = \left[u_i\left(\boldsymbol{R}\left(\boldsymbol{\theta}_{i,j}\right)\boldsymbol{r} + \boldsymbol{s}_{i,j}\right)\right]_{j=1}^{J}, \qquad i = 1, \ldots, S,$$

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

$$x(\boldsymbol{r}) = \sum_{i=1}^{p} z_i \psi_i(\mathbf{r})$$

$$\text{with} \quad z_i \in \{0, 1\}, \quad i = 1, \ldots, p,$$

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

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

Here, $\boldsymbol{A}\mathbf{\Psi}$ contains the projections of the individual dictionary elements, sampled at the appropriate points. The binary constraints on $\boldsymbol{z}$ make the recovery problem an integer program, and hence, NP-hard in general [21].

### 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

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

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 $\boldsymbol{z}$ closest to the hyperplane $\boldsymbol{A}\mathbf{\Psi}\boldsymbol{z} = \boldsymbol{y}$ in a Euclidean sense that lies on the intersection of a hyperplane $\boldsymbol{z}^T \mathbf{1} = K$ and the hyperbox $\mathbf{0} \leq \boldsymbol{z} \leq \mathbf{1}$. In Figure 2, we show the geometry for a shape-sensing problem with two possible shapes $(p = 2)$. In Figure 2(a), the hyperplane corresponding to tomographic measurements intersects the hyperplane corresponds to equality constraints $(\boldsymbol{z}^T \mathbf{1} = K)$ 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

(a)           (b)
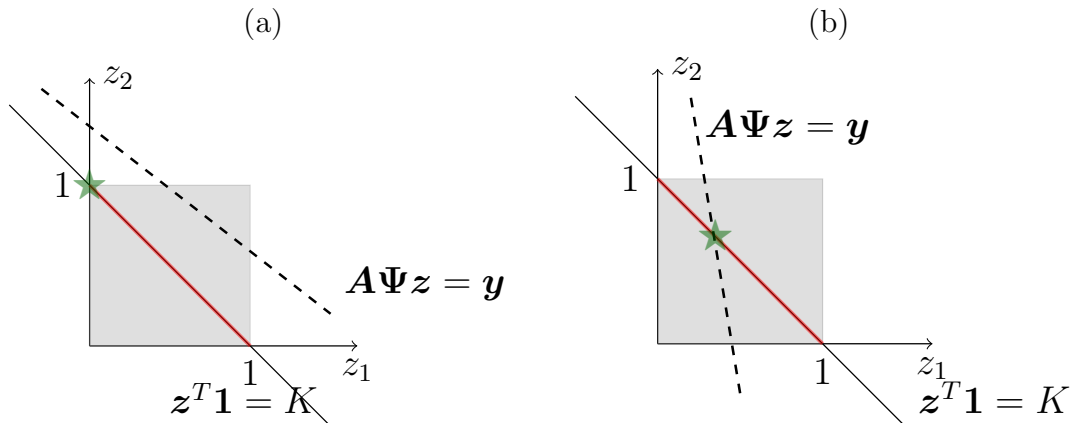


**Figure 2.** Geometry of CoShaRP. The grey region denotes the hyperbox that corresponds to bounds constraints($\mathbf{0} \leq \mathbf{z} \leq \mathbf{1}$). Solid line denotes the hyperplane for constraint on the number of shapes in the image, while dotted line denotes the hyperplane for measurement $\mathbf{A\Psi z} = \mathbf{y}$. 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.

hyperbox (cf. Figure 2(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

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

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 3. 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.** *CoShaRP differs significantly from the Sparse Shape Composition (SSC) [17, 19]. SSC formulates the shape-sensing problem as*

$$minimize \quad \tfrac{1}{2} \left\| \mathbf{A\Psi z} - \mathbf{y} \right\|^2 \quad subject\ to \quad \left\| \mathbf{z} \right\|_1 \leq K. \tag{2}$$

*The $\ell_1$-norm ball is bigger in size than the K-simplex constraint set used in CoShaRP. In particular, the K-simplex constraint represents the strict $\ell_1$ ball, i.e., $\left\| \mathbf{z} \right\|_1 = K$, 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).*

## 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:

$$\text{minimize} \quad f(\boldsymbol{A\Psi z}) + g(\boldsymbol{z}),$$

$$\text{where} \quad f(\boldsymbol{z}) = \|\boldsymbol{z} - \boldsymbol{y}\|, \quad g(\boldsymbol{z}) = \delta_{\mathcal{C}}(\boldsymbol{z}),$$

$$\text{with} \quad \mathcal{C}(\boldsymbol{z}) = \left\{ \boldsymbol{z} \in \mathbb{R}^n \,|\, \boldsymbol{z}^T \boldsymbol{1} = K, \boldsymbol{0} \le \boldsymbol{z} \le \boldsymbol{1} \right\},$$

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 [22, 23]. The iterates of this primal-dual algorithm takes the following form:

$$\boldsymbol{z}_{t+1} = \mathbf{prox}_{\gamma g}\left( \boldsymbol{z}_t - \gamma \boldsymbol{\Psi}^T \boldsymbol{A}^T \boldsymbol{u}_t \right),$$

$$\boldsymbol{u}_{t+1} = \mathbf{prox}_{\tau f^\star}\left( \boldsymbol{u}_t - \tau \boldsymbol{A\Psi} \left( \boldsymbol{z}_t - 2\boldsymbol{z}_{t+1} \right) \right),$$

for $t = [0, 1, \ldots, T]$, where $\gamma, \tau > 0$, with $\gamma\tau \le \|\boldsymbol{A\Psi}\|^{-1}$, are parameters that controls the speed of convergence. The main characteristic of this algorithm is that we avoid an inversion of a large matrix $\boldsymbol{\Psi}$ which often occurs in other splitting methods such as alternating direction method of multipliers [24]. Moreover, the proximal of both functions are easy to compute.

**Remark.** *The proposed primal-dual algorithm does not require user to store a large dictionary matrix $\boldsymbol{\Psi}$ as well as tomography matrix $\boldsymbol{A}$. Hence, the algorithm can utilize*



(a) 1-simplex  (b) 2-simplex  (c) 3-simplex
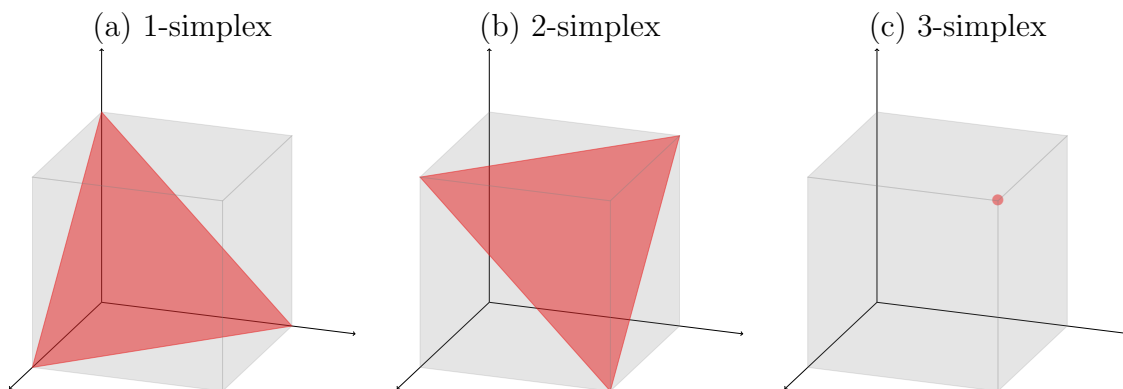
**Figure 3.** K-simplex in 3D. Barring $K = 3$ case where the K-simplex reduces to a point, the simplex are equilateral triangles (denoted by red color). The gray box denotes the bounds constraints.

---

**Algorithm 1** Primal-Dual Algorithm for CoShaRP

---

**Input:** $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{\Psi} \in \mathbb{R}^{n \times p}, \boldsymbol{y} \in \mathbb{R}^m, \gamma, \tau, T, \epsilon$

**Output:** $\overline{\boldsymbol{z}}$

$\quad$ *Initialisation* : $\boldsymbol{z}_0, \boldsymbol{u}_0$

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

---

*the functional forms of the dictionary as well as tomography operator since it only requires the forward and the adjoint operation with the operator.*

*4.2. Proximal operators*

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

$$f^\star(\boldsymbol{w}) = -\inf_{\boldsymbol{x}} \left\{ f(\boldsymbol{x}) - \boldsymbol{x}^T \boldsymbol{w} \right\}$$

$$= \begin{cases} \boldsymbol{w}^T \boldsymbol{y} & \text{if} \quad \|\boldsymbol{w}\| \leq 1 \\ +\infty & \text{otherwise} \end{cases}$$

We refer to [25, Example 3.26] for the derivation of the conjugate function. The conjugate function is linear inside the Euclidean norm ball of size 1 and $\infty$ outside. Hence, the conjugate function is convex. Its proximal operator is given in the following theorem.

**Theorem 1.** *The proximal operator of function*

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

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

$$\mathbf{prox}_{\gamma f^\star}(\boldsymbol{z}) = \frac{\boldsymbol{z} - \gamma \boldsymbol{y}}{\max\left(1, \|\boldsymbol{z} - \gamma \boldsymbol{y}\|\right)}. \tag{3}$$

*Proof.* The proximal operator for function $f^\star$ reads

$$\mathbf{prox}_{\gamma f}(\boldsymbol{z}) = \operatorname*{argmin}_{\boldsymbol{x}} \left\{ \frac{1}{2\gamma} \|\boldsymbol{x} - \boldsymbol{z}\|^2 + \boldsymbol{x}^T \boldsymbol{y} : \|\boldsymbol{x}\| \leq 1 \right\},$$

$$= \operatorname*{argmin}_{\|\boldsymbol{x}\| \leq 1} \left\{ \|\boldsymbol{x} - (\boldsymbol{z} - \gamma \boldsymbol{y})\|^2 - \gamma^2 \|\boldsymbol{y}\|^2 + 2\gamma \boldsymbol{z}^T \boldsymbol{y} \right\}.$$

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

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

This concludes the proof. □

We use the following theorem to compute the proximal of $g(\boldsymbol{z}) = \delta_{\mathcal{C}}(\boldsymbol{z})$, adapted from [26, Theorem 6.27].

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

$$\mathbf{prox}_{\gamma\delta_{\mathcal{C}}}(\boldsymbol{x}) = \mathcal{P}_{[\mathbf{0},\mathbf{1}]}(\boldsymbol{x} - \overline{\mu}\mathbf{1}) \tag{4}$$

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

*Proof.* The orthogonal projection of $\boldsymbol{x}$ is the unique solution of

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \left\{ \tfrac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|^2 \; : \; \mathbf{1}^T\boldsymbol{z} = K, \mathbf{0} \leq \boldsymbol{z} \leq \mathbf{1} \right\}. \tag{5}$$

A Lagrangian of this minimization problem reads

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

where $\mu \in \mathbb{R}$ is a Lagrange multiplier. It follows from the strong duality that $\overline{\boldsymbol{z}}$ is an optimal solution of problem (5) if and only if there exists a dual variable $\mu \in \mathbb{R}$ for which

$$\boldsymbol{y}^\star \in \operatorname*{argmin}_{\mathbf{0}\leq\boldsymbol{z}\leq\mathbf{1}} \; \mathcal{L}(\boldsymbol{z}, \overline{\mu}), \tag{6}$$

$$\mathbf{1}^T\boldsymbol{z} = K. \tag{7}$$

Using the expression for the Lagrangian, the relation (6) can be equivalently written as $\overline{\boldsymbol{z}} = \mathcal{P}_{[\mathbf{0},\mathbf{1}]}(\boldsymbol{x} - \overline{\mu}\mathbf{1})$. The feasibility condition (7) takes a form $\mathbf{1}^T\mathcal{P}_{[\mathbf{0},\mathbf{1}]}(\boldsymbol{x} - \overline{\mu}\mathbf{1}) = K$. □

**Remark.** *The projection onto the box $[\mathbf{0},\mathbf{1}]$ is simple. It is done component-wise as $\left(\min\{\max\{x_i, 0\}, 1\}\right)_{i=1}^n$. However, equation (4) 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 $i$, $\phi$ is a non-increasing function. Its root can be found using the Newton procedure, where derivative is*

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

---

**Algorithm 2** Image formation from shape coefficients

---

**Input:** $\boldsymbol{z} \in \mathbb{R}^p, \boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{\Psi} \in \mathbb{R}^{n \times p}, \boldsymbol{y} \in \mathbb{R}^m, K$

**Output:** $\overline{\boldsymbol{x}}$

    *Initialisation* : $\overline{\boldsymbol{x}} = \boldsymbol{0}, i = 0, s = 0$

  1: sort the elements of $\boldsymbol{z}$ in the descending order, and store its indices as a list $T$

  2: **while** $(s \leq K)$ or $(i \leq p)$ **do**

  3:     $\hat{\boldsymbol{x}} = \overline{\boldsymbol{x}} + \boldsymbol{\Psi} \boldsymbol{e}_{T(i)}$

  4:     **if** $\|\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{y}\| \leq \|\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{y}\|$ and $\overline{\boldsymbol{x}}^T \boldsymbol{\Psi} \boldsymbol{e}_{T(i)} \leq 0$ **then**

  5:         $\overline{\boldsymbol{x}} = \hat{\boldsymbol{x}}, \quad s = s + 1$

  6:     **end if**

  7:     $i = i + 1$

  8: **end while**

  9: **return** $\overline{\boldsymbol{x}}$

---

### 4.3. Image Formation

The convex program CoShaRP does not always lead to a binary solution (refer to Figure 2). 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, $\boldsymbol{e}_j$ is a natural basis vector with non-zero element located at $j^{\text{th}}$ position. To make sure the shapes do not overlap, we also add necessary conditions (see step 4 in Algorithm 2).

## 5. 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 $128 \times 128$ pixels, and the tomography matrix has at least 1024 measurements. The typical tomography setup is shown in Figure 4. In the CoShaRP performance plots, we generate 100 different realizations of 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 run Algorithm 1 with $\gamma = 1.2\sigma$, $\tau = 0.8\sigma$ with $\sigma = \|\boldsymbol{A}\boldsymbol{\Psi}\|^{-1}$. Moreover, we set $T = 4p^2$ and $\epsilon = 1e - 6$. Once we obtain the vector $\overline{\boldsymbol{z}}$, we run Algorithm 2 to form the image.
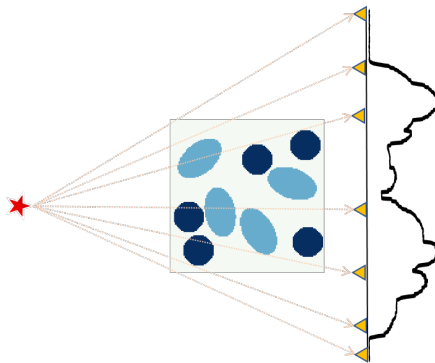
**Figure 4.** Single-shot tomography setup. The source (denoted by ⋆) is located at the left of the target image. The detector array consisting of 1024 detectors is located at the right of the image. For reference, we only show 7 detectors (denoted by ▽). Measurement profile is shown on the right of the detector array.

## 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 5(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 5(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 $128 \times 128$ pixels. The bottom figure in Figure 5(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.

## 5.3. Non-homogeneous and Non-convex Shapes

The top figure in Figure 5(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 5(c) gives the reconstruction with CoShaRP for a true image with 6 repetitions of the above-mentioned non-convex
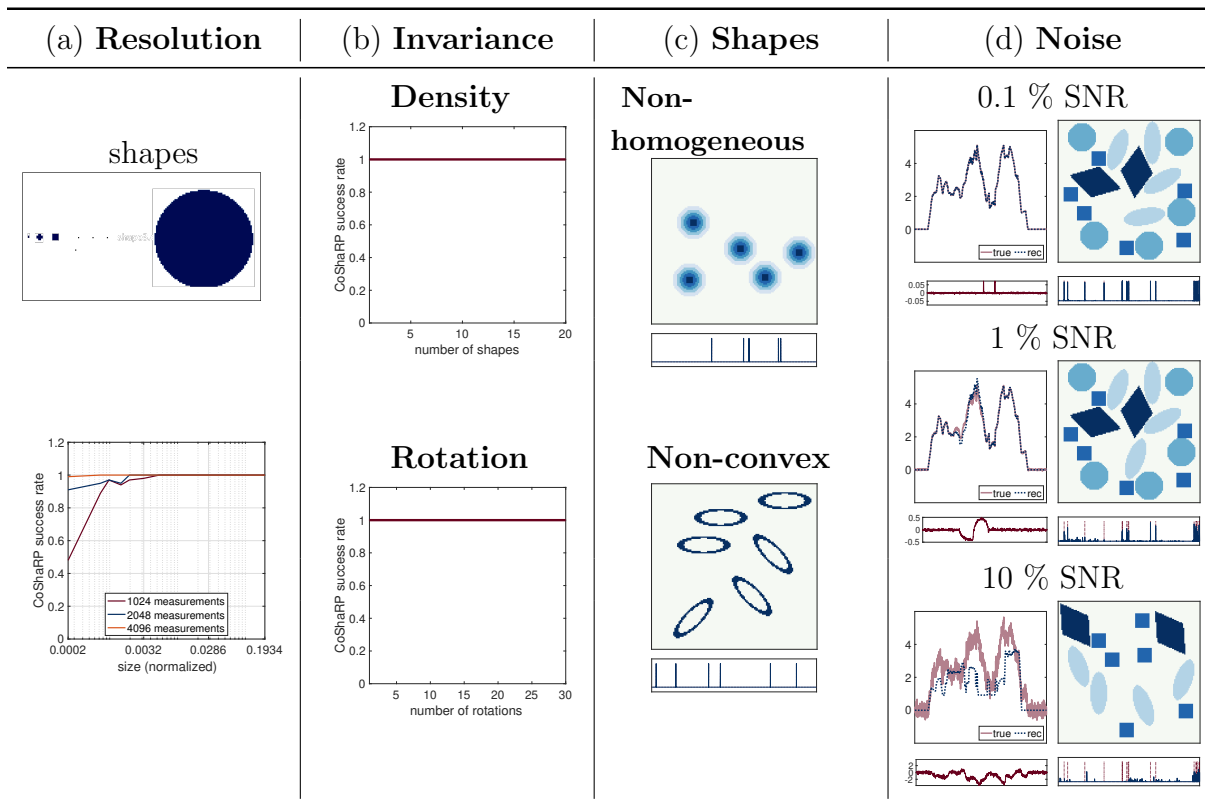
**Figure 5.** Numerical results for CoShaRP.

shape. From these two experiments, we conclude that the CoShaRP can recover the non-homogeneous and non-convex shapes.

## 5.4. Measurement Noise

We consider the true image shown in Figure 1 and consider three noisy scenarios where the Gaussian noise of strength 0.1%, 1% and 10% is added to 1024 measurements. In Figure 5(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 5(d), it is evident that the CoShaRP is stable till 1% noise, while fails for extremely noisy measurements.

## 6. 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. The numerical results 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.

## 7. References

[1] Defrise M and Clack R 1994 *IEEE Transactions on Medical Imaging* **13** 186–195
[2] Kudo H, Noo F and Defrise M 1998 *Physics in Medicine and Biology* **43** 2885–2909
[3] Zou Y and Pan X 2004 *Physics in Medicine and Biology* **49** 2717–2731
[4] Delaney A and Bresler Y 1998 *IEEE Transactions on Image Processing* **7** 204–221
[5] Persson M, Bone D and Elmqvist H 2001 *Physics in Medicine and Biology* **46** 853–866
[6] Sidky E Y and Pan X 2008 *Physics in Medicine and Biology* **53** 4777–4807
[7] Batenburg K J and Sijbers J 2011 *IEEE Transactions on Image Processing* **20** 2542–2553
[8] Frikel J 2013 *Applied and Computational Harmonic Analysis* **34** 117–141
[9] Hantke M F, Hasse D, Maia F R N C, Ekeberg T, John K, Svenda M, Loh N D, Martin A V, Timneanu N, Larsson D S D, van der Schot G, Carlsson G H, Ingelman M, Andreasson J, Westphal D, Liang M, Stellato F, DePonte D P, Hartmann R, Kimmel N, Kirian R A, Seibert M M, Mühlig K, Schorb S, Ferguson K, Bostedt C, Carron S, Bozek J D, Rolles D, Rudenko A, Epp S, Chapman H N, Barty A, Hajdu J and Andersson I 2014 *Nature Photonics* **8** 943–949
[10] Pelt D M and Andrade V D 2016 *Advanced Structural and Chemical Imaging* **2**
[11] Laanait N, Saenrang W, Zhou H, Eom C B and Zhang Z 2017 *Advanced Structural and Chemical Imaging* **3**
[12] Bicer T, Gürsoy D, Andrade V D, Kettimuthu R, Scullin W, Carlo F D and Foster I T 2017 *Advanced Structural and Chemical Imaging* **3**
[13] Spence J C H, Weierstall U and Chapman H N 2012 *Reports on Progress in Physics* **75** 102601
[14] Zhu W, Ma W, Su Y, Chen Z, Chen X, Ma Y, Bai L, Xiao W, Liu T, Zhu H, Liu X, Liu H, Liu X and Yang Y M 2020 *Light: Science & Applications* **9**
[15] Zhang X, Settembre E, Xu C, Dormitzer P R, Bellamy R, Harrison S C and Grigorieff N 2008 *Proceedings of the National Academy of Sciences* **105** 1867–1872
[16] Scheres S H 2012 *Journal of Structural Biology* **180** 519–530
[17] Aghasi A and Romberg J 2013 *SIAM Journal on Imaging Sciences* **6** 2075–2108

[18] Aghasi A and Romberg J 2015 *SIAM Journal on Imaging Sciences* **8** 2887–2950
[19] Zanaga D, Bleichrodt F, Altantzis T, Winckelmans N, Palenstijn W J, Sijbers J, de Nijs B, van Huis M A, Sánchez-Iglesias A, Liz-Marzán L M, van Blaaderen A, Batenburg K J, Bals S and Tendeloo G V 2016 *Nanoscale* **8** 292–299
[20] Combettes P L and Pesquet J C 2011 Proximal splitting methods in signal processing *Springer Optimization and Its Applications* (Springer New York) pp 185–212
[21] Karp R M 1972 Reducibility among combinatorial problems *Complexity of Computer Computations* (Springer US) pp 85–103
[22] Zhang X, Burger M and Osher S 2010 *Journal of Scientific Computing* **46** 20–46
[23] Chambolle A and Pock T 2010 *Journal of Mathematical Imaging and Vision* **40** 120–145
[24] Boyd S, Parikh N, Chu E, Peleato B and Eckstein J 2010 *Foundations and Trends® in Machine Learning* **3** 1–122
[25] Boyd S and Vandenberghe L 2004 *Convex optimization* (Cambridge university press)
[26] Beck A 2017 *First-Order Methods in Optimization* (Society for Industrial and Applied Mathematics)