

Article

Not peer-reviewed version

---

# Source Design Optimization for Depth Image Reconstruction in X-ray Imaging

---

[Hamid Fathi](#)<sup>\*</sup> and [Tristan Van Leeuwen](#)

Posted Date: 15 April 2024

doi: 10.20944/preprints202404.0960.v1

Keywords: Laminography; Limited angle tomography; Depth image



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Source Design Optimization for Depth Image Reconstruction in X-ray Imaging

Hamid Fathi <sup>1,\*</sup>  and Tristan van Leeuwen <sup>1,2</sup> 

<sup>1</sup> Centrum Wiskunde & Informatica (CWI), NWO, 1098 XG Amsterdam, The Netherlands

<sup>2</sup> Mathematical Institute, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands

\* Correspondence: Hamid.fathi@cwi.nl

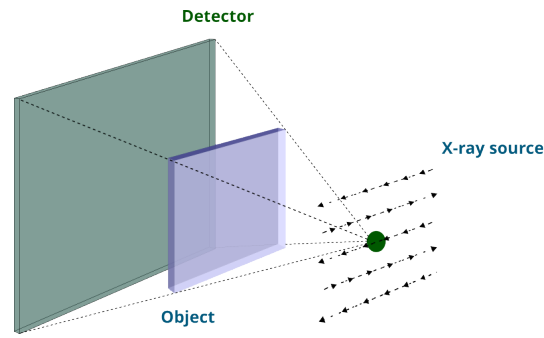
**Abstract:** X-ray tomography is an effective non-destructive testing method for industrial quality control. Limited angle tomography can be used to reduce the amount of data that needs to be acquired and thus speed up the process. In some industrial applications, however, objects are flat and layered, and laminography is preferred. It can deliver 2D images of the structure of a layered object at a particular depth from very limited data. An image at a particular depth is obtained by summing those parts of the data that contribute to that slice. This produces a sharp image of that slice while superimposing a blurred version of structures present at other depths. In this paper, we investigate an optimal experimental design (OED) problem for laminography that aims to determine the optimal source positions. Not only can this be used to mitigate imaging artifacts, it can also speed up the acquisition process in case where moving the source and detector is time consuming (e.g., in robotic arm imaging systems). We investigate the imaging artifacts in detail through a modified Fourier Slice Theorem. We address the experimental design problem within the Bayesian risk framework using empirical Bayes risk minimization. Finally, we present numerical experiments on simulated data.

**Keywords:** laminography; limited angle tomography; depth image

## 1. Introduction

Non-destructive testing plays a key role in industrial quality control [1]. X-ray imaging in particular is a highly effective non-destructive testing technique that can deliver detailed 3D images of the internal structure of products. It is therefore widely used in various industries for quality control [2–4]. Conventional tomographic acquisition requires a full rotation around the object, limiting its applicability to off-line inspection of small to medium-sized objects. To overcome this limitation, several approaches to tomographic reconstruction from limited-angle data have been developed. These methods typically assume additional prior knowledge of the object to fill in information that is missing in the measurements. A prime example is discrete tomography, where objects consisting of a few distinct materials can be reconstructed from severely under-sampled data. In some applications, objects are wide and flat, and tomographic data can only be collected for a very limited angular range. This pushes the limits of conventional limited-angle tomography and we have to resort to another class of methods [5]. Laminography offers a method to reconstruct an image of a specific slice of a flat layered object. As shown in Figure 1, it needs a flexible source and detector that can move and collect data from several directions such that during scanning the X-ray tube is located on the same side of the flat object and the detector is on the opposite side of it [6].

In this paper we consider a particular setup for laminographic imaging which requires the acquisition of X-ray projection images from  $N$  source positions on a fixed grid, while the detector is fixed on the opposite side of the object. Our aim is to be able to optimally design acquisition protocols with  $K \leq N$  sources for a given set of objects. By optimally designing the setup we can reduce imaging artifacts as well as potentially speed up the data-acquisition process in cases where moving the source is time consuming (e.g., robotic arm imaging). We pose the resulting optimal experimental design (OED) problem [7,8] as a bi-level optimization problem that employs derivative-based methods to facilitate implementation in high-dimensional settings [9].



**Figure 1.** A laminographic setup designed for scanning flat objects, featuring a flexible source that moves along a raster trajectory. Note that dashed arrows specify the source trajectory.

Our main contributions are the following.

- We analyze the corresponding 3D and 2D (slice-based) image reconstruction process, present an analog of the Fourier Slice Theorem for this setup and analyze the image artifacts present in slice-images.
- We pose an experimental design problem for selecting the  $K$  most informative source positions for laminography of a class of objects, and formulate a bi-level optimization in the Bayes risk minimization framework to solve it.

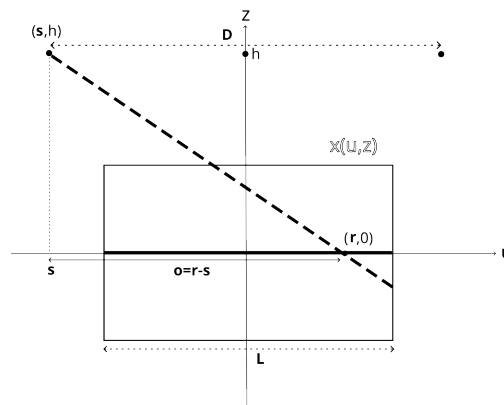
The remainder of the paper is organized as follows. In Section 2, we present the forward and backward operators in the continuous domain for the depth imaging procedure. We study depth reconstruction in the discrete domain in Section 3. In Section 4, we propose a bi-level optimization problem to find a source design with  $K$  sources chosen from a fixed array of  $N \geq K$  sources. In Section 5, we propose an algorithm to solve the proposed bi-level optimization problem. In Section 6, we show the results obtained by our method and compare them to optimal source designs, obtained by exhaustive search. In Section 8, we provide a discussion on the limitations and future work of the proposed method.

### 1.1. Notation

Throughout the paper, scalars and vector elements are denoted by lower-case letters, vectors by lower-case boldface letters, and matrices by upper-case boldface letters. The notation  $\cdot^\top$  stands for transpose of either vector or matrix. Moreover,  $\#$  denotes cardinality of a vector and  $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$  denotes the  $\ell_2$ -norm of a vector.

## 2. Forward and Backward Operators in the Continuous Domain

In this section, we will first review the linear X-ray projection model, followed by introducing the forward and backward operators for the setup shown in Figure 2.



**Figure 2.** 2D view of the acquisition setup. The source is located at lateral position  $s \in \mathbb{R}^2$  and height  $h$ ; the detector is parameterized without loss of generality by  $r \in \mathbb{R}^2$ . The object is characterized by  $x$  as a function of lateral coordinate  $u$  and vertical coordinate  $z$ . The ray from the source to the detector is shown by the dashed line.

### 2.1. Linear Projection Model

During the X-ray imaging process, an object is exposed to a beam of X-ray photons emitted by an X-ray tube. Some of these photons are absorbed by the object, while others reach the detector. The captured measurements can be modeled using the Beer-Lambert law, which describes the exponential relationship between the intensity of the X-ray beam and the distance it travels through the object

$$I_i = I_0 \exp \left( - \int x(t_i + \eta_i r) dr \right). \quad (1)$$

Here,  $I_i$  denotes the number of photons that reach the detector from a starting point  $t_i$  in the direction of  $\eta_i$ , where  $i = 1, \dots, m$ .  $I_0$  is the initial intensity of the beam (i.e., the number of photons emitted by the source), and  $x$  is the function representing the attenuation coefficient of the object. This relation is usually written as:

$$y_i = \int x(t_i + \eta_i r) dr, \quad (2)$$

with  $y_i = -\log \frac{I_i}{I_0}$  representing the measured data for a single ray. In the following, we present the forward and backward operators for the X-ray setup shown in Figure 2. To study the depth imaging procedure within this specific setup, we need to understand forward and backward operators for both 2D and 3D objects.

### 2.2. Forward and Backward Operators for the Slice Imaging Setup

The proposed imaging setup is schematically depicted in Figure 2. X-rays are emitted by a source at lateral position  $s \in \mathbb{R}^2$  and height  $h$ . Without loss of generality, we parametrize the detector position by  $r \in \mathbb{R}^2$ . The forward operator,  $\mathcal{A}$ , is now given by

$$y(r, s) = (\mathcal{A}x)(r, s) = \int x(tr + (1-t)s, (1-t)h) dt, \quad (3)$$

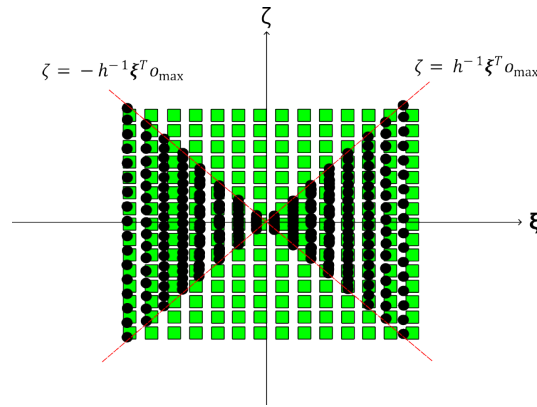
as shown in Figure 2. Here,  $x(u, z)$  denotes the attenuation coefficient of the object in terms of lateral ( $u \in \mathbb{R}^2$ ) and vertical ( $z \in \mathbb{R}$ ) coordinates. In fact, it integrates along the ray from the source at  $(s, h)$  to the detector located at  $(r, 0)$ .

### 2.2.1. A Fourier Slice Theorem

Defining the measurements in receiver-offset coordinates as  $g(\mathbf{r}, \mathbf{o}) = y(\mathbf{r}, \mathbf{r} - \mathbf{o})$ , we find the following relation between the 2D Fourier transform of  $g$  and the 3D Fourier transform of  $x$ . Let  $\xi$  and  $\zeta$  represent the spatial frequency coordinates, then it reads: (see the derivation in Appendix. A.1)

$$\begin{aligned}\mathcal{F}_{2D}(g)(\xi, \mathbf{o}) &= \mathcal{F}_{3D}(x)(\xi, \zeta), \\ \zeta &= h^{-1}\xi^\top \mathbf{o},\end{aligned}\quad (4)$$

where  $\mathcal{F}_{2D}$  and  $\mathcal{F}_{3D}$  denote the 2D and 3D Fourier transforms, respectively, and  $\mathbf{o} = \mathbf{r} - \mathbf{s}$  signifies the vector difference between the detector position  $\mathbf{r}$  and the source position  $\mathbf{s}$ . Figure 3 depicts sampled elements  $\mathcal{F}_{2D}(g)$  (black circles) with respect to the Fourier object samples  $\mathcal{F}_{3D}(x)$  (shown by green squares). Obviously, the sampling procedure is missing a portion of the object spectrum due to the limited values of  $\mathbf{o}$ , i.e., the forward operator samples from  $\zeta = h^{-1}\xi^\top \mathbf{o}$ . Since the range of  $\mathbf{o}$  is finite, this leads to a missing wedge in the frequency domain. This missing wedge leads to an ill-posed inverse problem of recovering  $x$  from  $g$  (or equivalently,  $y$ ).



**Figure 3.** Schematic depiction of the missing wedge in the Fourier domain sampling of  $\mathcal{F}_{3D}(x)$  by the measurements  $\mathcal{F}_{2D}(g)$ . The green squares represent  $\mathcal{F}_{3D}(x)(\xi, \zeta)$  and black circles represent  $\mathcal{F}_{2D}(g)(\xi, \mathbf{o})$  which are limited in  $\zeta$  by  $h^{-1}\xi^\top \mathbf{o}_{\max}$ . The limited aperture in  $\mathbf{o}$  gives rise to the missing wedge.

### 2.2.2. The 2D Slice at Depth $z$

To study the depth-imaging procedure, we introduce the corresponding forward operator  $\mathcal{A}_z$  which maps a 2D image at depth  $z$ , denoted by  $x_z(\mathbf{u})$  to the corresponding projection data  $y(\mathbf{r}, \mathbf{s})$  (See Appendix. A.2),

$$y(\mathbf{r}, \mathbf{s}) = (\mathcal{A}_z x_z)(\mathbf{s}, \mathbf{r}) = x_z\left(\left(1 - h^{-1}z\right)\mathbf{r} + h^{-1}z\mathbf{s}\right). \quad (5)$$

The corresponding adjoint operator  $\mathcal{A}_z^*$  is given by (See Appendix. A.2),

$$(\mathcal{A}_z^* y)(\mathbf{u}) = \frac{1}{1 - h^{-1}z} \int y\left(\frac{\mathbf{u} - h^{-1}z\mathbf{s}}{1 - h^{-1}z}, \mathbf{s}\right) d\mathbf{s}. \quad (6)$$

We then find that  $\mathcal{A}_z^* \mathcal{A}_z$  is given by

$$(\mathcal{A}_z^* \mathcal{A}_z x_z)(\mathbf{u}) = \frac{\int d\mathbf{s}}{1 - h^{-1}z} x_z(\mathbf{u}), \quad (7)$$

so that for a limited range of  $s$ ,  $\mathcal{A}_z^* \mathcal{A}_z \propto \mathcal{I}$  and hence is trivially inverted. Depth imaging of an object consisting of a single thin layer can thus be achieved by a simple back-projection. When multiple layers are present we expect artifacts, which we study next.

### 2.3. Depth Image Reconstruction

Here, we are interested in studying the artifacts that arise when imaging a slice from measurements of a 3D layered object. In the remainder of this section, we let  $h = 1$  for ease of notation. As argued above, we can form a depth image by back projection, so it suffices to study the operator  $\mathcal{A}_z^* \mathcal{A}$  (See Appendix. A.3). In particular, we expect to image layered objects, which we can model as

$$x(\mathbf{u}, z) = \sum_{i=1}^M x_i(\mathbf{u}) \delta(z - z_i).$$

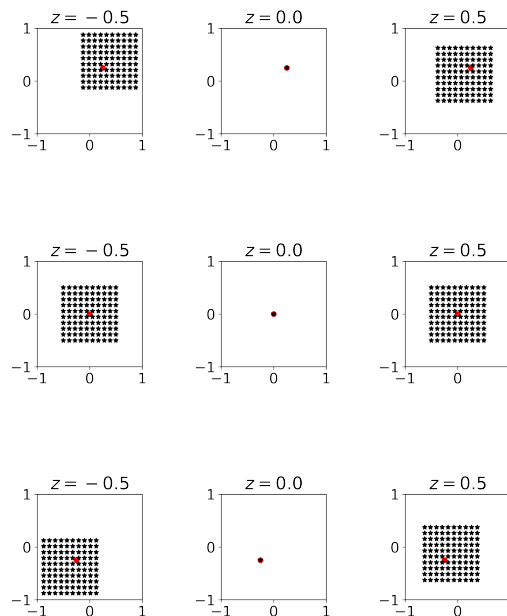
Using a finite number of sources  $\{\mathbf{s}_j\}_{j=1}^N$ , the depth image of the  $k^{\text{th}}$  layer is then given by

$$\tilde{x}_k(\mathbf{u}) = \sum_{i=1}^M \sum_{j=1}^N x_i \left( \frac{1 - z_i}{1 - z_k} \mathbf{u} + \frac{z_i - z_k}{1 - z_k} \mathbf{s}_j \right) \quad (8)$$

$$= x_k(\mathbf{u}) + N^{-1} \sum_{i \neq k} \sum_{j=1}^N x_i \left( \frac{1 - z_i}{1 - z_k} \mathbf{u} + \frac{z_i - z_k}{1 - z_k} \mathbf{s}_j \right) \quad (9)$$

Hence, the error in the depth image of the  $k^{\text{th}}$  layer is a superposition of spatially averaged versions of the other layers. An illustration of the resulting spatial sampling patterns is shown in Figure 4.

The source design problem is now to find a set of source positions  $\mathbf{s}_j$  such that unwanted parts in the depth image are minimized. As we don't expect to be able to find a closed-form solution to the source design problem in general, we discretize the problem and resort to numerical optimization.



**Figure 4.** Example of the spatial sampling patterns when imaging the middle layer (at  $z = 0$ ) of an object with  $M = 3$  layers and  $N = 100$  regularly spaced sources in  $[-1, 1]^2$  at  $z = 1$ . The red circle indicates the point  $\mathbf{u}$  at which we evaluate the output.

### 3. Discrete Problem

#### 3.1. Forward and Back Projection

We discretize the forward model in (3) and express it as

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (10)$$

where  $\mathbf{y} \in \mathbb{R}^m$ ,  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Note that the number of rows in  $\mathbf{A}$  is the product of the number of detector pixels and the number of sources. Every block of rows corresponds to a particular X-ray source.

Let  $\mathbf{A}_z$ ,  $\mathbf{x}_z$  denote the forward matrix and the slice object at depth  $z$ , respectively. In general,  $\mathbf{A}_z^\top \mathbf{A}_z$  will not be a multiple of the identity as in the continuous setting (7), so we consider solving  $\mathbf{A}_z^\top \mathbf{A}_z \mathbf{x}_z = \mathbf{A}_z^\top \mathbf{y}$ . We can view this as an approximation of 3D reconstruction through the Schur complement [10]. To see this, let  $\mathbf{A}_c$  and  $\mathbf{x}_c$  represent the forward operator and the object for all pixels except the ones contained in  $\mathbf{x}_z$ . The linear equations for the forward projection (10) can now be expressed as follows:

$$\begin{bmatrix} \mathbf{A}_z & \mathbf{A}_c \end{bmatrix} \begin{bmatrix} \mathbf{x}_z \\ \mathbf{x}_c \end{bmatrix} = \mathbf{y}. \quad (11)$$

By multiplying  $\mathbf{A}^\top$ , it reads

$$\begin{bmatrix} \mathbf{A}_z^\top \mathbf{A}_z & \mathbf{A}_z^\top \mathbf{A}_c \\ \mathbf{A}_c^\top \mathbf{A}_z & \mathbf{A}_c^\top \mathbf{A}_c \end{bmatrix} \begin{bmatrix} \mathbf{x}_z \\ \mathbf{x}_c \end{bmatrix} = \begin{bmatrix} \mathbf{A}_z^\top \mathbf{y} \\ \mathbf{A}_c^\top \mathbf{y} \end{bmatrix}. \quad (12)$$

By utilizing Schur complement, one can re-write this in terms of  $\mathbf{x}_z$  as

$$\mathbf{A}' \mathbf{x}_z = \mathbf{y}', \quad (13)$$

where

$$\begin{aligned} \mathbf{A}' &= \left( \mathbf{A}_z^\top \mathbf{A}_z - \mathbf{A}_z^\top \mathbf{A}_c \left( \mathbf{A}_c^\top \mathbf{A}_c \right)^{-1} \mathbf{A}_c^\top \mathbf{A}_z \right), \\ \mathbf{y}' &= \left( \mathbf{A}_z^\top \mathbf{y} - \mathbf{A}_z^\top \mathbf{A}_c \left( \mathbf{A}_c^\top \mathbf{A}_c \right)^{-1} \mathbf{A}_c^\top \mathbf{y} \right). \end{aligned}$$

However, it is computationally not feasible to compute the corresponding Schur complements, and we can interpret depth imaging as neglecting the additional terms involving  $\mathbf{A}_c$ .

#### 3.2. Algebraic Reconstruction

Although the operator  $\mathcal{A}_z$  is unitary (up to a scaling factor), the corresponding discrete operator  $\mathbf{A}_z$  is generally not. Moreover, it is desirable to stabilize the reconstruction with additional regularization. We therefore pose depth imaging as a regularized least-squares problem:

$$\hat{\mathbf{x}}_z = \underset{\mathbf{x}_z}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{A}_z \mathbf{x}_z - \mathbf{y}\|^2 + \frac{\lambda}{2} \|\mathbf{x}_z\|^2, \quad (14)$$

where  $\lambda$  controls the trade-off between the two terms. While the problem is not extremely ill-posed, we utilize  $\lambda$  as a safety factor to ensure the uniqueness of the reconstructed depth image. The above optimization can be solved using iterative algorithms such as the LSQR [11]. Various other regularization techniques, such as Total Variation regularization, or sparse regularization, introduce different penalty functions that promote desired properties in the reconstructed image.



#### 4. Efficient Source Design for $K$ Sources

Let  $\beta \in \mathbb{R}^N$  be the design parameter, where the components of  $\beta$  determine the influence of each source in the reconstruction process. The problem of finding the optimal design parameter  $\beta$  for a fixed number of sources in the reconstruction of a specific depth plane can be formulated as a bi-level optimization problem as below,

$$\min_{\beta} \min_{\alpha(\beta)} \frac{1}{2} \|\alpha(\beta) \hat{x}_z(\beta) - x_{\text{gt}}\|^2 \quad \text{s.t. } \#\beta = K, \beta \in \{0, 1\}^N, \quad (15)$$

where  $x_{\text{gt}} \in \mathbb{R}^n$  represents the ground truth and the hyper-parameter  $\alpha(\beta) \in \mathbb{R}$  serves as scaling factor, ensuring that  $\alpha(\beta) \hat{x}_z(\beta)$  is on the same scale as  $x_{\text{gt}}$ . Also,  $\hat{x}_z(\beta)$  is obtained by solving

$$\hat{x}_z(\beta) = \operatorname{argmin}_{x_z} \frac{1}{2} \|T_{\beta}(A_z x_z - y)\|^2 + \frac{\lambda}{2} \|x_z\|^2, \quad (16)$$

where  $y$  denotes the data corresponding to  $x_{\text{gt}}$  and  $T_{\beta}$  is a linear operator which applies the weights  $\beta$  to each source's corresponding measurements. The optimization problem (15) minimizes the Euclidean distance between the reconstructed image (scaled by  $\alpha(\beta)$ ) and the ground truth subject to certain constraints on  $\beta$  to enforce its binary nature with  $K$  elements.

The binary constraint on  $\beta$  in (15) can be relaxed using bound constraints [12], and the bi-level optimization is then formulated as

$$\min_{\beta} \frac{1}{2} \|\hat{\alpha}(\beta) \hat{x}_z(\beta) - x_{\text{gt}}\|^2 \quad \text{s.t. } \beta^{\top} \mathbf{1} = K, 0 \leq \beta \leq \mathbf{1}, \quad (17)$$

while the optimal value of the scalar parameter  $\alpha(\beta)$  in (15) is substituted in above optimization with  $\hat{\alpha}(\beta) = \frac{\hat{x}_z(\beta)^{\top} x_{\text{gt}}}{\|\hat{x}_z(\beta)\|^2}$ . Mathematically speaking, this problem seeks a  $\beta$  that minimizes the Euclidean distance of  $\hat{x}(\beta)$  and  $x_{\text{gt}}$  and lies on the intersection of a hyperplane  $\beta^{\top} \mathbf{1} = K$  and the hyperbox  $0 \leq \beta \leq 1$ . Note that we cannot guarantee that this will produce binary solutions. In such cases, we convert the solution to a binary form by setting the  $K$  largest elements to one and assigning zeros to the remaining elements.

In practice the ground truth of  $x_{\text{gt}}$  is unknown, however, a set of calibration data  $\{x_{\text{gt}}^i\}_{i=1}^M$  of representative images may be available. This can then be utilized to compute an efficient source design such that minimizes the empirical Bayes risk. More precisely, for a given set of calibration data  $\{x_{\text{gt}}^i\}_{i=1}^M$  with corresponding measurements  $\{y^i\}_{i=1}^M$ , we can generalize (17) to

$$\min_{\beta} \frac{1}{2M} \sum_{i=1}^M \|\hat{\alpha}^i(\beta) \hat{x}_z^i(\beta) - x_{\text{gt}}^i\|^2 \quad \text{s.t. } \beta^{\top} \mathbf{1} = K, 0 \leq \beta \leq \mathbf{1}, \quad (18)$$

where  $\hat{x}_z^i(\beta)$  is the solution of (16) for the  $i$ -th calibration data  $y^i$ .

#### 5. Implementation

To address the implementation of the Efficient  $K$ -Source Design (EKSD) method, it is worth mentioning that the non-linear optimization problem of (17) requires numerical optimization algorithms to start from an initial guess and refine the solution iteratively. This section discusses the solution of (17) using a projected gradient method.

##### 5.1. Projected Gradient Method

In general, the projected gradient (PG) method solves,

$$\min_{\beta} f(\beta) \quad \text{s.t. } \beta \in \mathcal{B}, \quad (19)$$



where  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  is a smooth function and  $\mathcal{B}$  is a nonempty closed convex set. The projection onto  $\mathcal{B}$  is the mapping  $\mathcal{P} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  defined by:

$$\mathcal{P}(\mathbf{u}) = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{u}\|^2 \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{B}. \quad (20)$$

The PG algorithm is defined by,

$$\boldsymbol{\beta}_{k+1} = \mathcal{P}(\boldsymbol{\beta}_k - \gamma \nabla f(\boldsymbol{\beta}_k)) \quad (21)$$

where  $\gamma \in (0, 2/L)$  is the step size,  $L$  is the Lipschitz constant of  $\nabla f$ , and  $\nabla f(\boldsymbol{\beta}_k)$  is the gradient of  $f$  at  $\boldsymbol{\beta}_k$ . The stopping criteria is defined as the condition where the Euclidean norm of the difference between consecutive design parameters, denoted as  $\|\boldsymbol{\beta}_{k+1} - \boldsymbol{\beta}_k\|$ , is less than or equal to a specified threshold value, denoted as  $\epsilon$ .

### 5.2. Gradient of the Objective Function

To implement the non-linear optimization of (17) using the PG algorithm, the gradient calculation of  $F(\hat{\mathbf{x}}(\boldsymbol{\beta}), \hat{\mathbf{a}}(\boldsymbol{\beta})) := \frac{1}{2} \|\hat{\mathbf{a}}(\boldsymbol{\beta})\hat{\mathbf{x}}(\boldsymbol{\beta}) - \mathbf{x}_{\text{gt}}\|^2$  is required. Here, it reads,

$$\begin{aligned} \nabla F(\boldsymbol{\beta}) &= \hat{\mathbf{a}} \frac{\partial \hat{\mathbf{x}}_z(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}^\top (\hat{\mathbf{a}}\hat{\mathbf{x}}_z(\boldsymbol{\beta}) - \mathbf{x}_{\text{gt}}) \\ &= 2\hat{\mathbf{a}}\mathbf{C}^\top \left( \mathbf{A}^\top \mathbf{T}_\beta^\top \right)_\lambda^\dagger (\hat{\mathbf{a}}\hat{\mathbf{x}}_z(\boldsymbol{\beta}) - \mathbf{x}_{\text{gt}}), \end{aligned} \quad (22)$$

where  $\left( \mathbf{A}^\top \mathbf{T}_\beta^\top \right)_\lambda^\dagger = \left( \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{A} + \lambda \mathbf{I} \right)^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top$ , and  $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_N]$  with  $\mathbf{c}_i = \frac{\partial \mathbf{T}_\beta}{\partial \beta_i} (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_z(\boldsymbol{\beta}))$ . (See the derivation in Appendix. A.4)

### 5.3. Projection Onto the Simplex Constraint

Here, we investigate the projection operator  $\mathcal{P}$  presented in (21) which maps each point onto the simplex set of  $\mathcal{B} = \{\mathbf{u} \in \mathbb{R}^m \mid \mathbf{u}^\top \mathbf{1} = K, 0 \leq u_i \leq 1\}$ . (See the derivation in Appendix. A.5)

$$\mathcal{P}(\mathbf{p}) = \min\{\max\{\mathbf{p} - \mu\mathbf{1}, \mathbf{0}\}, \mathbf{1}\}, \quad (23)$$

where  $\mathbf{p}, \mathbf{1} \in \mathbb{R}^m$  and  $\mu$  can be obtained by solving the following optimization problem:

$$\sum_i \min\{\max\{p_i - \mu, 0\}, 1\} = K, \quad (24)$$

which can be solved using the Newton method.

## 6. Numerical Experiments

Here, we present some results related to our proposed method (EKSD). We evaluate the performance of EKSD using the following criteria:

i: Normalized mean square error (NMSE) can be defined as the error between the reconstructed image and ground truth

$$\text{NMSE} := \frac{\|\hat{\mathbf{a}}\hat{\mathbf{x}}_z - \mathbf{x}_{\text{gt}}\|^2}{\|\mathbf{x}_{\text{gt}}\|^2}. \quad (25)$$

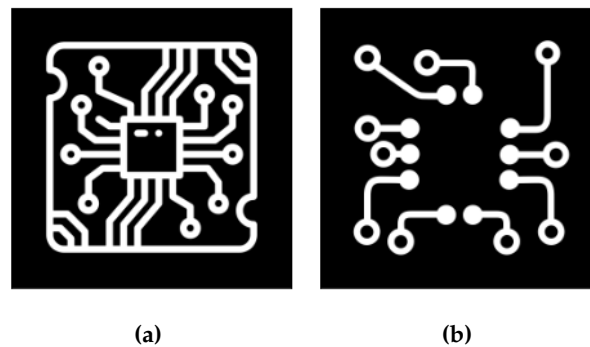
ii: Structural similarity index measure (SSIM): Unlike NMSE, SSIM takes into account the structural information and perceived quality of the images. It compares three aspects of the images:

luminance, contrast, and structure, and computes a similarity index between 0 and 1, where 1 indicates perfect similarity [13].

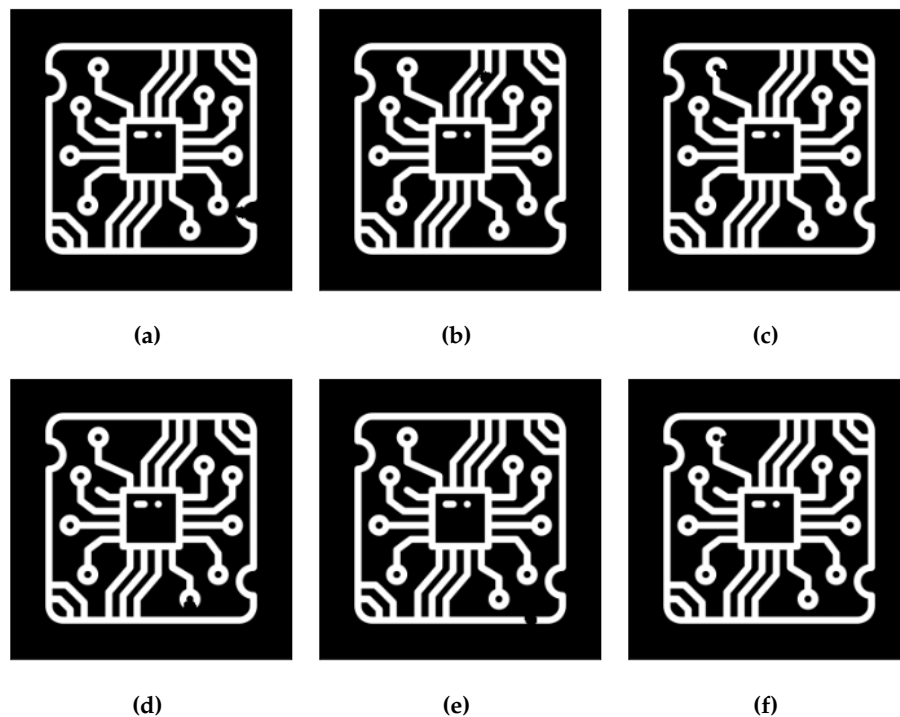
### 6.1. Data Set

As a benchmark example we consider a two-layer PCB of size  $80 \times 250 \times 250$  pixels. It consists of two distinct wire designs (as illustrated in Figure 5) with a thickness of 25 pixels each. These designs are separated by a gap of 20 pixels. The layer of interest is the one depicted in Figure 5a.

We create 200 defect shapes in various locations, forming a data set of defective wire design for the bottom layer. A few samples of this data set are depicted in Figure 6. We use half of this data set for source design (calibration) and half for validation.



**Figure 5.** The wire designs with a size of  $250 \times 250$  are located at the bottom and top layers within the PCB, respectively.



**Figure 6.** Various defects in different locations on the wire design.

### 6.2. Implementation Details

We define the X-ray setup as shown in Figure 2. The distance  $h$ , between the center of the array and the detector, is equal to  $6L$ . Here,  $L$  represents the width of the phantom. Moreover, we consider  $r \in [-L/2, L/2]^2$  (discretized uniformly with  $250 \times 250$  pixels) and  $s \in [-D/2, D/2]^2$  (discretized

uniformly with  $\sqrt{N} \times \sqrt{N}$  source positions), where  $D$  is the fixed array length. The projection data are generated using the ASTRA toolbox [14]. For illustration, we present the projection images from all sources in a  $4 \times 4$  array of sources in Figure 7. It is challenging to inspect individual layers of a the phantom with a single projection, as overlapping them hinders a proper investigation.

We choose a small regularization parameter  $\lambda = 0.1$  to ensure the solution's stability and uniqueness without significantly biasing the reconstruction. We carefully tuned the step size,  $\gamma = 10^{-9}$ , to be sufficiently small, reducing the risk of algorithmic overshooting during optimization. Additionally, we manually adjusted the stopping threshold,  $\epsilon = 0.001$ , to ensure that the algorithm stops at approximately consistent convergence points. The initial value of the parameter vector  $\beta$  is initialized either with a vector ones or with i.i.d. uniform random values in  $[0, 1]$ . We refer to these methods as EKSD(K) and EKSDR(K), respectively. In some experiments, we will compare the results to completely random designs of  $K$  out of  $N$  source positions. We refer to this as the RANDOM method.

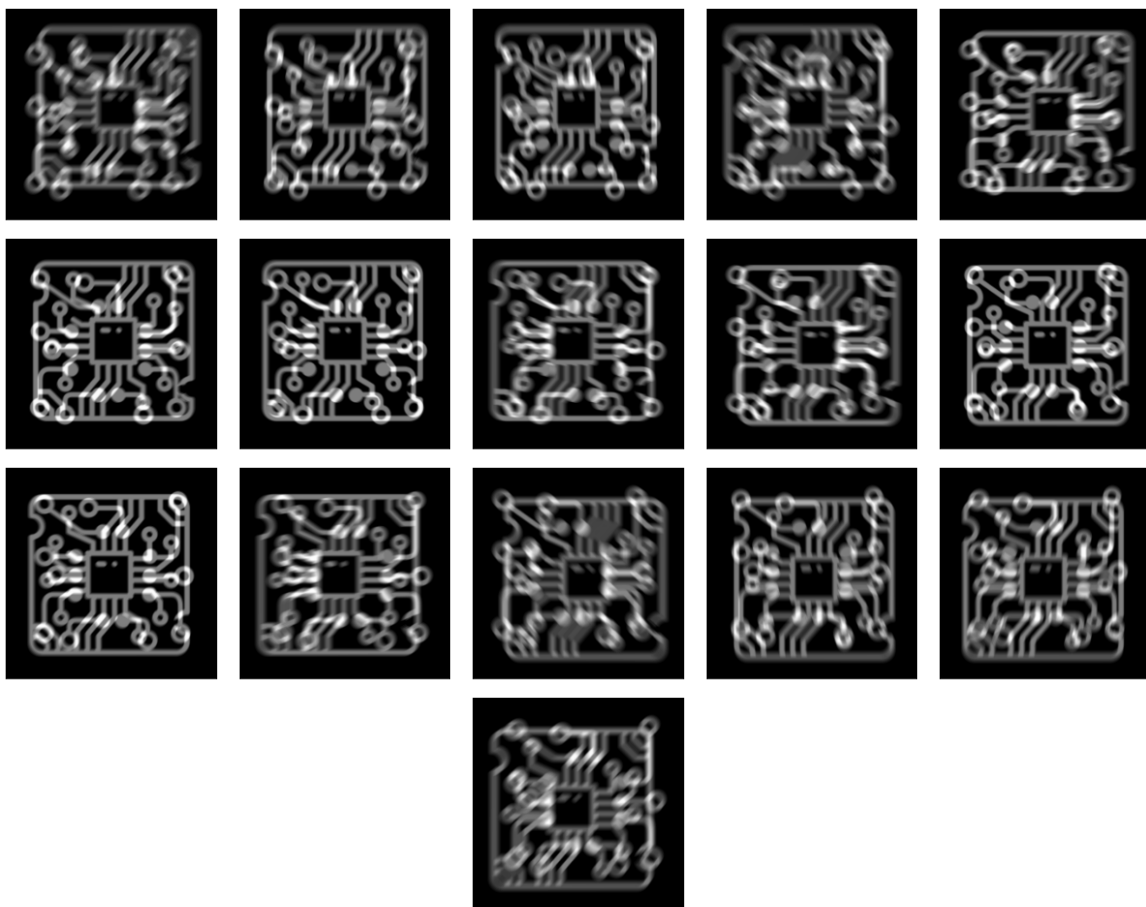
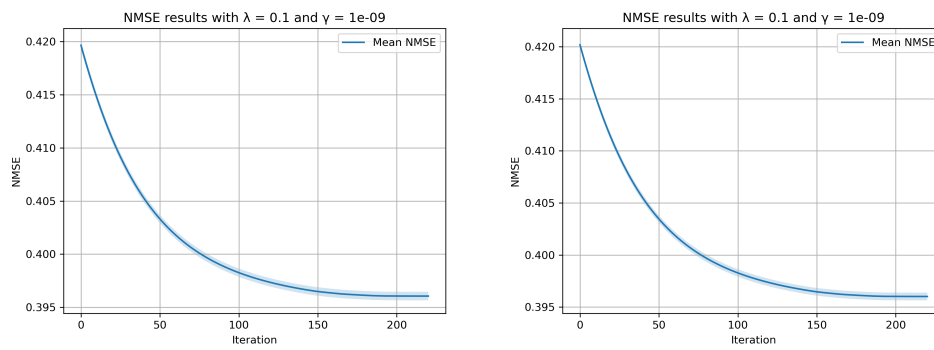


Figure 7. Different views from sources in a  $4 \times 4$  array of sources.

### 6.3. Convergence

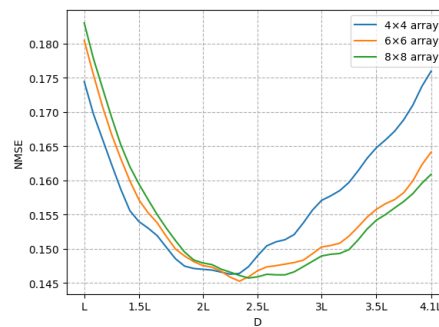
We evaluate the convergence behavior of the EKSD algorithm across various iterations. Here, we let  $K = 10$ ,  $N = 16$ ,  $D = 4L$ . In Figure 8, we depict the NMSE at each iteration as evaluated on the calibration (left) and validation (right) data. This figures shows that the design is not overfitted on the calibration data.



**Figure 8.** The EKSD algorithm's convergence as applied to a calibration data set, is depicted. The blue line indicates the mean values, while the black dots highlight the variance of the NMSE at each iteration.

#### 6.4. Exploring the Effects of Array Length

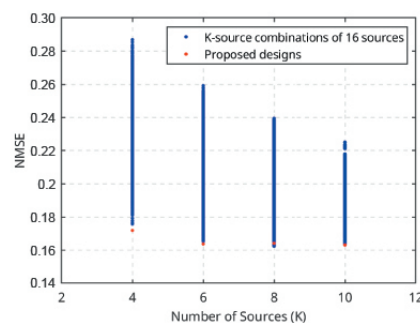
Here, we study the influence of the array length  $D$  on the reconstruction error for  $N = 16, 36, 64$ . Hence, we compare the average NMSE over the validation data across various array lengths, as depicted in Figure 9. The heightened errors observed for both smaller and larger array lengths indicate that certain source positions within the array contribute to artifacts in the depth image, leading to an elevated NMSE.



**Figure 9.** The reconstruction error for various ranges of the area occupied by the source.

#### 6.5. Exhaustive Search

We carried out an exhaustive search to assess the performance of EKSD. We consider  $N = 16$  source positions with  $D = 4L$ . Figure 10 shows the average NMSE over the validation data for all possible  $K$ -source designs for each  $K$ , as well as the average NMSE achieved by the EKSD-design. We see the the designs found by EKSD are nearly optimal.



**Figure 10.** The error range applies to any  $K$  source positions within the array of 16 positions, where  $K$  can be 4, 6, 8, or 10.

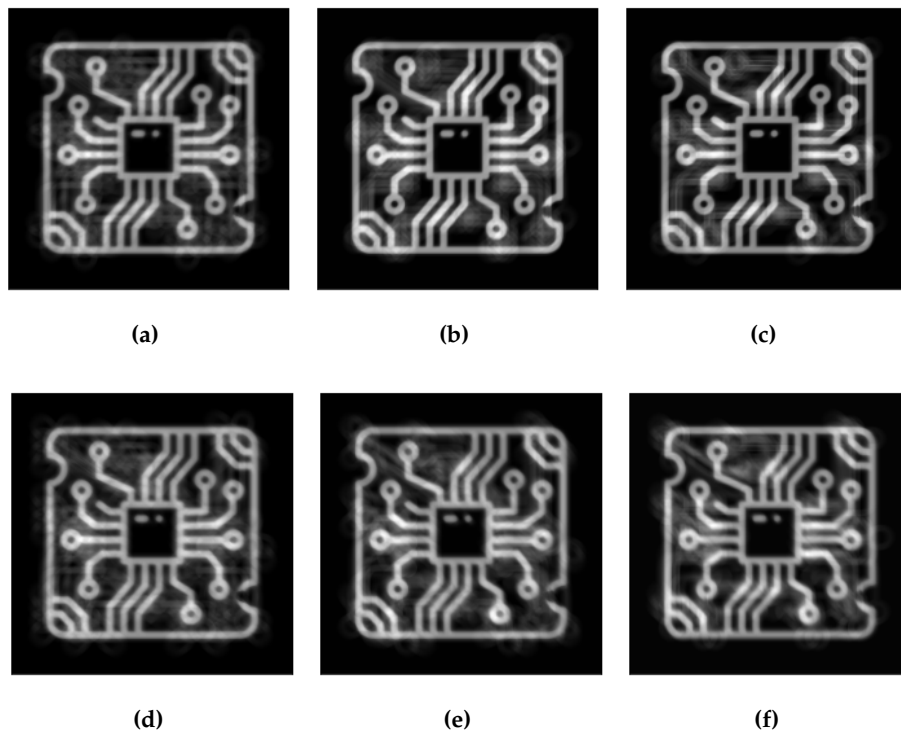
### 6.6. Source Design 1

We consider three configurations with  $N = 16, 36, 64$  at  $D = 4L$ , and we show the NMSE and SSIM results in Table 1. Initially, we display results for depth-reconstructed images utilizing all source positions within the array. Subsequently, we derive  $K = 10$  efficient source designs with EKSD and EKSDR and present the corresponding NMSE and SSIM results. For each value of  $N$ , we observe that EKSD and EKSDR outperform than the case using all source positions. This improvement can be attributed to the reduction of artifacts in the images, which is likely due to the selective nature of measurements in the EKSD and EKSDR approaches. As a comparison we show the average NMSE and SSIM of random designs with  $K = 10$  sources, and these are significantly worse than all other methods.

**Table 1.** NMSE and SSIM results with the occupied array length of  $4L$ . For randomized methods, we present the average and standard deviation over 50 runs.

Source design	NMSE ( $N$ )			SSIM ( $N$ )		
	$N = 16$	$N = 36$	$N = 64$	$N = 16$	$N = 36$	$N = 64$
ALL( $K = N$ )	0.1749	0.1632	0.1603	0.5630	0.5921	0.6046
EKSD ( $K = 10$ )	0.1611	0.1508	0.1528	0.6275	0.7304	0.7397
EKSDR ( $K = 10$ )	$0.1611 \pm 3 \times 10^{-4}$	$0.1508 \pm 4 \times 10^{-4}$	$0.1518 \pm 1.1 \times 10^{-3}$	$0.6275 \pm 3 \times 10^{-4}$	$0.7304 \pm 4 \times 10^{-4}$	$0.7380 \pm 4.6 \times 10^{-3}$
RANDOM ( $K = 10$ )	$0.1819 \pm 7 \times 10^{-3}$	$0.1781 \pm 1.1 \times 10^{-2}$	$0.1759 \pm 1.4 \times 10^{-2}$	$0.57 \pm 1.5 \times 10^{-2}$	$0.6075 \pm 2.1 \times 10^{-2}$	$0.6248 \pm 3 \times 10^{-2}$

Figure 11 shows the corresponding depth images of the bottom layer reconstructed with source positions suggested by EKSD. As demonstrated in Section 2.3, artifacts in these images attributed to the superposition of features from different layers. The lack of rotational acquisition in laminography leads to this artifacts. Clearly, the designs proposed by EKSD minimize artifacts in depth images compared to random designs, facilitating accurate defect detection within the depth images.



**Figure 11.** Depth images in [a](#), [b](#), [c](#) are reconstructed using 10 proposed source positions from the total of 16, 36, and 64 positions, respectively. Conversely, the depth images in [d](#), [e](#), [f](#) are derived from 10 random source positions out of the same total positions.

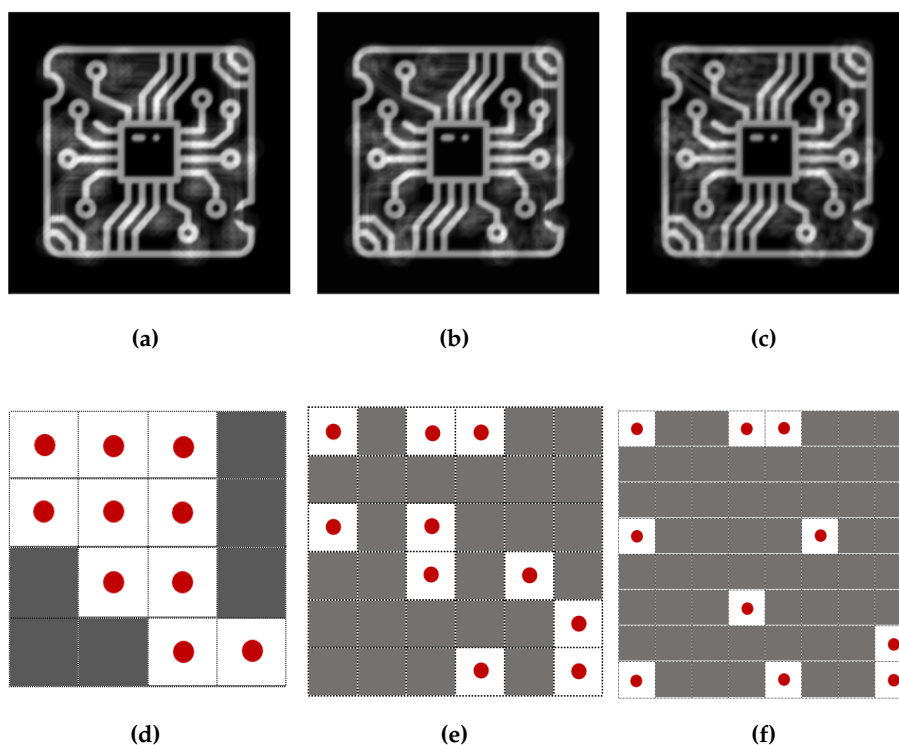
### 6.7. Source Design 2

We repeat the previous experiment with the array length optimally according to Figure 9;  $D = 2.4L$  for different  $N$  values. The results are summarized in Table 2. As the full designs are already optimal

in a sense, there is less dramatic improvement when optimizing the design with fewer source positions. However, the results for  $N = 64$  suggest that increasing the number of positions while maintaining a fixed array length, thereby reducing the distance between sources, may result in a non-binary solution. Then, converting the solution into a sparse vector with  $K = 10$  elements, results in a less optimal outcome. The proposed source designs and corresponding depth images are shown in Figure 12.

**Table 2.** NMSE and SSIM results with the occupied array length of  $2.4L$ . For randomized methods, we present the average and standard deviation over 50 runs.

Source design	NMSE ( $N$ )			SSIM ( $N$ )		
	$N = 16$	$N = 36$	$N = 64$	$N = 16$	$N = 36$	$N = 64$
ALL( $K = N$ )	0.1467	0.1454	0.1459	0.7030	0.7174	0.7233
EKSD ( $K = 10$ )	0.1488	0.1478	0.1548	0.7230	0.7271	0.7011
EKSDR ( $K = 10$ )	$0.1488 \pm 5 \times 10^{-4}$	$0.1484 \pm 9 \times 10^{-4}$	$0.1530 \pm 2.5 \times 10^{-3}$	$0.7230 \pm 5 \times 10^{-4}$	$0.7259 \pm 2.7 \times 10^{-3}$	$0.7040 \pm 4.7 \times 10^{-3}$
RANDOM ( $K = 10$ )	$0.1537 \pm 4 \times 10^{-3}$	$0.1580 \pm 5 \times 10^{-3}$	$0.1614 \pm 8 \times 10^{-3}$	$0.7040 \pm 1 \times 10^{-2}$	$0.7178 \pm 1.1 \times 10^{-2}$	$0.7237 \pm 1.5 \times 10^{-2}$



**Figure 12.** Depth images shown in [a](#), [b](#), [c](#) are reconstructed by 10 proposed source positions shown in [d](#), [e](#), [f](#), respectively.

## 7. Discussion

In this research, we explore the methodology of experimental design for a laminographic setup, acknowledging potential challenges that may arise in real-world applications. Here, we outline several significant challenges with our proposed method: (i) The non-convex nature of the objective function may lead to entrapment in local minima. Addressing this, diversifying initial conditions enhances the likelihood of identifying more optimal solutions or the global minimum. (ii) Our method optimizes the positions of sources within a specified grid. However, this grid may not represent the most efficient set of source positions for optimization. An alternative approach could involve optimizing within a continuous domain of positions, leveraging the closed-form formulation for depth image reconstruction presented in this paper. (iii) While our investigation of the experimental design problem assumes noise-free measurements, this assumption may not hold in real-world scenarios characterized by noisy data. This can be addressed by incorporating noise directly into the optimization process to enhance real-world applicability.

## 8. Conclusion

In real world scenarios, imaging a particular depth of a planar object can yield valuable information with less effort than a full 3D reconstruction. In a depth image, the information corresponding to the chosen layer appears sharp, while the details of other layers appear blurred. We argue that source design plays a role in mitigating these unwanted artifacts and pose the experimental design problem as a bi-level optimization problem. To solve this problem, we employed a bilvel optimization and implemented it by gradient-based methods. The suggested framework can be tailored more specifically to a particular application by incorporating different regularization terms in lower level optimization or by adopting alternate loss function in the upper level optimization to more effectively capture the characteristics of the image.

## Appendix A

### Appendix A.1 Forward Operator for the 3D Object

Figure 2 presents the 2D view of acquisition setup for a planar object  $x(\mathbf{u}, z)$  where  $\mathbf{u} \in \mathbb{R}^2$  is the lateral coordinate and  $z \in \mathbb{R}$  the depth coordinate. The line integral of  $x(\mathbf{u}, z)$  from  $(s, h)$  to  $(r, 0)$  for a fixed  $h$  can be expressed as follows:

$$y(\mathbf{r}, s) = \int x(\mathbf{s} + (\mathbf{r} - \mathbf{s})t, h - ht) dt \quad (\text{A1})$$

where  $\mathbf{s} \in \mathbb{R}^2$ ,  $\mathbf{r} \in \mathbb{R}^2$  denote the source and detector positions, respectively. Next, we substitute  $x(\mathbf{u}, z)$  by its representation in the Fourier domain  $\mathcal{F}_{3D}(x)(\boldsymbol{\xi}, \zeta)$ :

$$x(\mathbf{u}, z) = \iint \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, \zeta) e^{j2\pi(\boldsymbol{\xi}^\top \mathbf{u} + \zeta z)} d\boldsymbol{\xi} d\zeta, \quad (\text{A2})$$

using (A2) in (A1), it yields:

$$y(\mathbf{r}, s) = \int \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, \zeta) e^{j2\pi(\boldsymbol{\xi}^\top (\mathbf{s} + (\mathbf{r} - \mathbf{s})t) + \zeta(h - ht))} dt d\boldsymbol{\xi} d\zeta, \quad (\text{A3})$$

using  $\int e^{j2\pi(\boldsymbol{\xi}^\top (\mathbf{r} - \mathbf{s}) - \zeta h)t} dt = \delta(\boldsymbol{\xi}^\top (\mathbf{r} - \mathbf{s}) - \zeta h)$ , it follows,

$$y(\mathbf{r}, s) = \int \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, \zeta) e^{j2\pi(\boldsymbol{\xi}^\top \mathbf{s} + \zeta h)} \delta(\boldsymbol{\xi}^\top (\mathbf{r} - \mathbf{s}) - \zeta h) d\boldsymbol{\xi} d\zeta, \quad (\text{A4})$$

and then,

$$y(\mathbf{r}, s) = \int \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, h^{-1}\boldsymbol{\xi}^\top (\mathbf{r} - \mathbf{s})) e^{j2\pi\boldsymbol{\xi}^\top \mathbf{r}} d\boldsymbol{\xi} \quad (\text{A5})$$

using  $\mathbf{o} = \mathbf{r} - \mathbf{s}$ , it can be written,

$$g(\mathbf{r}, \mathbf{o}) = \int \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, h^{-1}\boldsymbol{\xi}^\top \mathbf{o}) e^{j2\pi\boldsymbol{\xi}^\top \mathbf{r}} d\boldsymbol{\xi} \quad (\text{A6})$$

which means we can link the 3D Fourier transform of  $x$  to the 2D Fourier transform of the data as,

$$\mathcal{F}_{2D}(g)(\boldsymbol{\xi}, \mathbf{o}) = \mathcal{F}_{3D}(x)(\boldsymbol{\xi}, h^{-1}\boldsymbol{\xi}^\top \mathbf{o}). \quad (\text{A7})$$

### Appendix A.2 Forward Operator for a 2D Slice on the Depth $z$

We introduce  $x_z$  as a slice of the object  $x$  at depth  $z$ , i.e.,

$$x(\mathbf{u}, z') = x_z(\mathbf{u})\delta(z' - z).$$



We then define the forward operator  $\mathcal{A}_z$  as projecting a 2D object to the detector at  $z = 0$  via,

$$y(\mathbf{r}, \mathbf{s}) = x_z \left( \mathbf{s} + \left( 1 - h^{-1}z \right) (\mathbf{r} - \mathbf{s}) \right) \quad (\text{A8})$$

or equivalently as,

$$y(\mathbf{r}, \mathbf{s}) = x_z \left( \mathbf{r} \left( 1 - h^{-1}z \right) + h^{-1}z\mathbf{s} \right). \quad (\text{A9})$$

The adjoint operator  $\mathcal{A}_z^*$  can be obtained by considering the definition  $\langle \mathcal{A}_z x_z, y \rangle_{(\mathbf{s}, \mathbf{r})} = \langle x_z, \mathcal{A}_z^* y \rangle_{\mathbf{v}}$ . First, we have

$$\langle \mathcal{A}_z x_z, y \rangle_{(\mathbf{s}, \mathbf{r})} = \iint x_z \left( \mathbf{r} \left( 1 - h^{-1}z \right) + h^{-1}z\mathbf{s} \right) y(\mathbf{r}, \mathbf{s}) d\mathbf{r} d\mathbf{s}.$$

Now let  $\mathbf{v} = \mathbf{r} \left( 1 - h^{-1}z \right) + h^{-1}z\mathbf{s}$ , then  $\mathbf{r} = \frac{\mathbf{v} - h^{-1}z\mathbf{s}}{1 - h^{-1}z}$  and  $d\mathbf{r} = \frac{1}{1 - h^{-1}z} d\mathbf{v}$ , and it follows,

$$\begin{aligned} \langle \mathcal{A}_z x_z, y \rangle_{(\mathbf{s}, \mathbf{r})} &= \\ \frac{1}{1 - h^{-1}z} \iint x_z(\mathbf{v}) y \left( \frac{\mathbf{v} - h^{-1}z\mathbf{s}}{1 - h^{-1}z}, \mathbf{s} \right) d\mathbf{v} d\mathbf{s} \\ &= \langle x_z, \mathcal{A}_z^* y \rangle. \end{aligned} \quad (\text{A10})$$

We thus find that  $\mathcal{A}_z^* f$  is defined as

$$(\mathcal{A}_z^* y)(\mathbf{u}) = \frac{1}{1 - h^{-1}z} \int y \left( \frac{\mathbf{u} - h^{-1}z\mathbf{s}}{1 - h^{-1}z}, \mathbf{s} \right) d\mathbf{s}. \quad (\text{A11})$$

By using (A9) in (A11), the composition  $\mathcal{A}_z^* \mathcal{A}_z$  can be written as below

$$(\mathcal{A}_z^* \mathcal{A}_z x_z)(\mathbf{u}) = \frac{x_z(\mathbf{u})}{1 - h^{-1}z} \int d\mathbf{s}. \quad (\text{A12})$$

### Appendix A.3 Back Projection for a Depth Image

We use  $y(\mathbf{r}, \mathbf{s})$  in (A1) to express  $\mathcal{A}_z^* \mathcal{A}$  using (A11),

$$\begin{aligned} (\mathcal{A}_z^* \mathcal{A} x)(\mathbf{u}) &= \\ \frac{1}{1 - h^{-1}z} \iint x \left( \frac{\mathbf{u} - h^{-1}z\mathbf{s}}{1 - h^{-1}z} t + (1 - t) \mathbf{s}, h(1 - t) \right) dt d\mathbf{s}. \end{aligned} \quad (\text{A13})$$

Now let  $z' = h(1 - t)$ , then it can be rewritten as

$$\begin{aligned} (\mathcal{A}_z^* \mathcal{A} x)(\mathbf{u}) &= \\ \frac{1}{h - z} \iint x \left( \frac{h - z'}{h - z} \mathbf{u} + \frac{z' - z}{h - z} \mathbf{s}, z' \right) dz' d\mathbf{s} \end{aligned} \quad (\text{A14})$$

### Appendix A.4 Gradient Calculation of the Objective Function

To find the gradient of  $F(\hat{\mathbf{x}}(\boldsymbol{\beta}), \hat{\mathbf{a}}(\boldsymbol{\beta})) := \frac{1}{2} \|\hat{\mathbf{a}}(\boldsymbol{\beta}) \hat{\mathbf{x}}(\boldsymbol{\beta}) - \mathbf{x}_{\text{gt}}\|^2$  with respect to  $\boldsymbol{\beta}$ , by the chain rule it reads:

$$\frac{\partial F}{\partial \boldsymbol{\beta}}(\hat{\mathbf{x}}_z(\boldsymbol{\beta}), \hat{\mathbf{a}}(\boldsymbol{\beta})) = \frac{\partial F}{\partial \hat{\mathbf{x}}_z} \cdot \frac{\partial \hat{\mathbf{x}}_z}{\partial \boldsymbol{\beta}} + \frac{\partial F}{\partial \hat{\mathbf{a}}} \cdot \frac{\partial \hat{\mathbf{a}}}{\partial \boldsymbol{\beta}} \quad (\text{A15})$$

where the second term is equal to zero because  $\hat{\mathbf{x}}$  is the optimal point of  $F$ . Then, the first term can be written as:

$$\frac{\partial F}{\partial \beta} = \left( \frac{\partial \hat{\mathbf{x}}_z}{\partial \beta} \right)^\top \frac{\partial F}{\partial \hat{\mathbf{x}}_z} = \hat{\mathbf{a}} \left( \frac{\partial \hat{\mathbf{x}}_z}{\partial \beta} \right)^\top (\hat{\mathbf{a}} \hat{\mathbf{x}}_z(\beta) - \mathbf{x}_{\text{gt}}). \quad (\text{A16})$$

To obtain an expression for the derivative of  $\hat{\mathbf{x}}$  with respect to  $\beta$ , we exploit the fact that the lower-level optimization problem (16) has a closed-form solution:

$$\hat{\mathbf{x}}_z(\beta) = \left( \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{A} + \lambda \mathbf{I} \right)^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{y}. \quad (\text{A17})$$

We define  $\mathbf{H}_\beta := \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{A} + \lambda \mathbf{I}$ , and then the derivative follows,

$$\frac{\partial \hat{\mathbf{x}}(\beta)}{\partial \beta_i} = \frac{\partial \mathbf{H}_\beta^{-1}}{\partial \beta_i} \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{y} + \mathbf{H}_\beta^{-1} \frac{\partial \left( \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{y} \right)}{\partial \beta_i}, \quad (\text{A18})$$

recall that  $\frac{\partial \mathbf{H}_\beta^{-1}}{\partial \beta_i} = -\mathbf{H}_\beta^{-1} \frac{\partial \mathbf{H}_\beta}{\partial \beta_i} \mathbf{H}_\beta^{-1}$ , therefore,

$$\frac{\partial \mathbf{H}_\beta^{-1}}{\partial \beta_i} = -2\mathbf{H}_\beta^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top \frac{\partial \mathbf{T}_\beta}{\partial \beta_i} \mathbf{A} \mathbf{H}_\beta^{-1} \quad (\text{A19})$$

then (A18) can be rewritten as,

$$\frac{\partial \hat{\mathbf{x}}(\beta)}{\partial \beta_i} = -2\mathbf{H}_\beta^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top \frac{\partial \mathbf{T}_\beta}{\partial \beta_i} \left( \mathbf{A} \mathbf{H}_\beta^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{y} - \mathbf{y} \right), \quad (\text{A20})$$

where  $\mathbf{H}_\beta^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top \mathbf{T}_\beta \mathbf{y}$  is equal to  $\hat{\mathbf{x}}(\beta)$  in (A17) and  $\mathbf{H}_\beta^{-1} \mathbf{A}^\top \mathbf{T}_\beta^\top$  is denoted by  $(\mathbf{T}_\beta \mathbf{A})_\lambda^\dagger$ , then it can be rewritten,

$$\frac{\partial \hat{\mathbf{x}}(\beta)}{\partial \beta_i} = 2(\mathbf{T}_\beta \mathbf{A})_\lambda^\dagger \left( \frac{\partial \mathbf{T}_\beta}{\partial \beta_i} (\mathbf{y} - \mathbf{A} \hat{\mathbf{x}}(\beta)) \right) = 2(\mathbf{T}_\beta \mathbf{A})_\lambda^\dagger \mathbf{c}_i, \quad (\text{A21})$$

for simplicity, we denote  $\frac{\partial \mathbf{T}_\beta}{\partial \beta_i} (\mathbf{y} - \mathbf{A} \hat{\mathbf{x}}(\beta))$  as  $\mathbf{c}_i \in \mathbb{R}^m$  and it follows,

$$\frac{\partial \hat{\mathbf{x}}(\beta)}{\partial \beta} = 2(\mathbf{T}_\beta \mathbf{A})_\lambda^\dagger \mathbf{C} \quad (\text{A22})$$

where  $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_N]$ . By replacement of (A22) into (22), yields

$$\nabla F = 2\hat{\mathbf{a}} \mathbf{C}^\top \left( \mathbf{A}^\top \mathbf{T}_\beta^\top \right)_\lambda^\dagger (\hat{\mathbf{a}} \hat{\mathbf{x}}_z(\beta) - \mathbf{x}_{\text{gt}}). \quad (\text{A23})$$

#### Appendix A.5 Projection Onto the Simplex Constraint

Let  $\mathcal{P}$  be an operator that maps each point onto the simplex set of  $\mathcal{B} = \{\mathbf{u} \in \mathbb{R}^m \mid \mathbf{u}^\top \mathbf{1} = K, 0 \leq u_i \leq 1\}$ . In fact, the orthogonal projection of  $\mathbf{p}$  is the unique solution of:

$$\underset{\mathbf{t}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{t} - \mathbf{p}\|^2 \quad \text{s.t.} \quad \mathbf{t}^\top \mathbf{1} = K, 0 \leq t_i \leq 1, \quad (\text{A24})$$

to solve the above optimization via its dual problem, we introduce the Lagrangian [15],

$$\mathcal{L}(\mathbf{t}, \mu) := \frac{1}{2} \|\mathbf{t} - \mathbf{p}\|^2 + \mu \left( \mathbf{t}^\top \mathbf{1} - K \right), \quad (\text{A25})$$

where  $\mu \in \mathbb{R}$  is a Lagrangian multiplier. The dual problem can now be written as

$$\mathcal{J}(\mu) = \inf_{0 \leq t_i \leq 1} \mathcal{L}(\mathbf{t}, \mu), \quad (\text{A26})$$

where  $\mathcal{J}(\mu)$  is the dual function as the minimum value of Lagrangian over  $\mathbf{t}$ . Regarding strong duality the optimal solution of (A24) and (A26) are equal if there exists a dual variable of  $\mu \in \mathbb{R}$  that makes (A26) as a tight lower bound for (A24). The solution of the dual problem of (A26) can be written as [16],

$$\mathbf{t}^*(\mu) = \min\{\max\{\mathbf{p} - \mu \mathbf{1}, \mathbf{0}\}, \mathbf{1}\}$$

where  $\mu \in \mathbb{R}$  is obtained by solving  $\mathbf{t}^*(\mu)^\top \mathbf{1} = K$ . Let's define the function  $w(\mu) = \sum_i \min\{\max\{p_i - \mu, 0\}, 1\} - K$  that is non-increasing and can be solved using the Newton method. The derivative is given by,

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

## References

- Medici, V.; Martarelli, M.; Paone, N.; Pandarese, G.; van de Kamp, W.; Verhoef, B.; Sipsas, K.; Broechler, R.; Besada, L.B.; Alexopoulos, K.; Nikolakis, N. Integration of Non-Destructive Inspection (NDI) systems for Zero-Defect Manufacturing in the Industry 4.0 era. 2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), 2023, pp. 439–444. doi:10.1109/MetroInd4.0IoT57462.2023.10180016.
- Wang, B.; Zhong, S.; Lee, T.L.; Fancey, K.S.; Mi, J. Non-destructive testing and evaluation of composite materials/structures: A state-of-the-art review. *Advances in mechanical engineering* **2020**, *12*, 1687814020913761.
- Schimleck, L.; Dahlen, J.; Apiolaza, L.A.; Downes, G.; Emms, G.; Evans, R.; Moore, J.; Pâques, L.; Van den Bulcke, J.; Wang, X. Non-Destructive Evaluation Techniques and What They Tell Us about Wood Property Variation. *Forests* **2019**, *10*. doi:10.3390/f10090728.
- Hassler, U.; Oeckl, S.; Bauscher, I.; others. Inline ct methods for industrial production. International Symposium on NDT in Aerospace, 2008, pp. 123–131.
- Gondrom, S.; Zhou, J.; Maisl, M.; Reiter, H.; Kröning, M.; Arnold, W. X-ray computed laminography: an approach of computed tomography for applications with limited access. *Nuclear Engineering and Design* **1999**, *190*, 141–147. doi:https://doi.org/10.1016/S0029-5493(98)00319-7.
- O'Brien, N.; Mavrogordato, M.; Boardman, R.; Sinclair, I.; Hawker, S.; Blumensath, T. Comparing cone beam laminographic system trajectories for composite NDT. *Case Studies in Nondestructive Testing and Evaluation* **2016**, *6*, 56–61. Special Issue: Industrial computed tomography, doi:https://doi.org/10.1016/j.csndt.2016.05.004.
- Haber, E.; Horeh, L.; Tenorio, L. Numerical methods for experimental design of large-scale linear ill-posed inverse problems. *Inverse Problems* **2008**, *24*, 055012. doi:10.1088/0266-5611/24/5/055012.
- Haber, E.; Magnant, Z.; Lucero, C.; Tenorio, L. Numerical methods for A-optimal designs with a sparsity constraint for ill-posed inverse problems. *Computational Optimization and Applications* **2012**, *52*, 293–314.
- Ruthotto, L.; Chung, J.; Chung, M. Optimal Experimental Design for Inverse Problems with State Constraints. *SIAM Journal on Scientific Computing* **2018**, *40*, B1080–B1100.
- Zhang, F. *The Schur complement and its applications*; Vol. 4, Springer Science & Business Media, 2006.
- Paige, C.C.; Saunders, M.A. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software (TOMS)* **1982**, *8*, 43–71.
- Nemhauser, G.; Wolsey, L., Linear Programming. In *Integer and Combinatorial Optimization*; John Wiley & Sons, Ltd, 1988; chapter 1.2, pp. 27–49, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118627372.ch2]. doi:https://doi.org/10.1002/9781118627372.ch2.
- Wang, Z.; Bovik, A.; Sheikh, H.; Simoncelli, E. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **2004**, *13*, 600–612. doi:10.1109/TIP.2003.819861.
- van Aarle, W.; Palenstijn, W.J.; Cant, J.; Janssens, E.; Bleichrodt, F.; Dabrovolski, A.; Beenhouwer, J.D.; Batenburg, K.J.; Sijbers, J. Fast and flexible X-ray tomography using the ASTRA toolbox. *Opt. Express* **2016**, *24*, 25129–25147. doi:10.1364/OE.24.025129.

15. Boyd, S.; Vandenberghe, L. *Convex optimization*; Cambridge university press, 2004.
16. Kadu, A.; van Leeuwen, T.; Batenburg, K.J. CoShaRP: a convex program for single-shot tomographic shape sensing. *Inverse Problems* **2021**, *37*, 105005. doi:10.1088/1361-6420/ac1776.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.