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

Hamid Fathi 1, * and Tristan van Leeuwen 1, 2

1 Computational Imaging Group, Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands
2 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 need to be acquired and thereby 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 cases 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

MSC: 94A08

1. Introduction

Non-destructive testing is essential 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 widely employed across various industries for quality control [2, 3]. Conventional tomographic acquisition requires a full rotation around the object, limiting its applicability to the 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 often rely on 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 certain applications where objects are wide and flat, tomographic data can only be collected over a very limited angular range. This pushes the limits of conventional limited-angle tomography and we have to resort to another class of methods [4]. 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 [5].

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, particularly in scenarios where adjusting the source position is time-consuming, such as in robotic arm imaging systems. Bayesian Optimal Experimental Design (OED) is a mathematical approach that facilitates the design of experimental parameters while minimizing experimental costs. It utilizes prior knowledge and Bayesian inference to refine the selection of experimental settings, thereby improving data collection efficiency and decision-making effectiveness [6]. Various criteria, including D-optimality and A-optimality, are frequently employed to assess and compare the efficacy of experimental designs [6,7]. D-optimality focuses on maximizing the Kullback–Leibler divergence between the posterior and prior distributions, whereas A-optimality aims to minimize the expected error between reconstructions and actual ground truths [8,9]. In [10], the authors proposed the optimization of projection angles in X-ray imaging. They incorporated empirical A-optimal design in constrained problems based on calibration data. Specifically, they obtained a sparse design out of a predetermined set of angles and implemented it with a gradient-based optimization method. Similarly, we pose the OED problem for laminography as a bilevel optimization based on the A-optimality criterion. Additionally, we employ gradient-based methods to facilitate implementation in high-dimensional settings.

**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 3D and 2D (slice-based) image reconstruction process for laminography, present an analog of the Fourier Slice Theorem specific to 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 across a specific class of objects. Additionally, we formulate a bilevel optimization within the Bayes risk minimization framework to solve it.

The remainder of this 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 bilevel 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 bilevel optimization problem and present its pseudocode in Section 6. In Section 7, we show the results obtained by our method and compare them to optimal source designs obtained by exhaustive search. Finally, Sections 8 and 9 discuss the limitations and future directions of our research and conclude the paper, respectively.

**Notation**

Throughout this 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 a transpose of either a vector or matrix. Moreover, \( \# \) denotes the cardinality of a vector and \( \|x\| = \sqrt{\sum_{i=1}^{n} x_i^2} \) denotes the \( \ell_2 \) norm of a vector.
2. Forward and Backward Operators in the Continuous Domain

In this section, we first review the linear X-ray projection model. Subsequently, we introduce the forward and backward operators as depicted in Figure 2. Finally, we detail the depth imaging process for laminography in the continuous domain, utilizing these operators.

Briefly, the forward projection operator transforms the object’s internal features into the collected projection data, effectively modeling the imaging process. Conversely, the backward projection operator transforms the projection data back into a representation of the object’s internal features, serving as the basis for reconstructing the original image [11].

Figure 2. Two-dimensional 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:

\[
I_i = I_0 \exp \left( - \int x(t_i + \eta_i r) \, dr \right),
\]

where \( I_i \) denotes the number of photons that reach the detector from starting point \( t_i \) in the direction of \( \eta_i \), where \( i = 1, \ldots, m \). \( I_0 \) 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

\[
y_i = \int x(t_i + \eta_i r) \, dr,
\]

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 parameterise the detector position by \( r \in \mathbb{R}^2 \). The forward operator, \( A \), is now given by

\[
y(r, s) = (Ax)(r, s) = \int x(tr + (1 - t)s, (1 - t)h) \, dt,
\]

as shown in Figure 2. Here, \( x(u, z) \) denotes the attenuation coefficient of the object in terms of its 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. The Fourier Slice Theorem

The Fourier Slice Theorem illustrates the relationship between the collected measurements and the Fourier transform of the object. In traditional computed tomography, a complete sampling process enables the straightforward inversion of the Radon transform using the filtered back projection method \cite{12}. Due to the limited acquisition view in laminography, the Fourier Slice Theorem shows that some parts of the object spectrum are missing. The absence of complete data leads to the appearance of artifacts within the reconstructed image at specific depths. According to the setup shown in Figure 2, we define the measurements in receiver-offset coordinates as \( g(r, o) = y(r, r - o) \) and 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 as follows (see the derivation in Appendix A.2):

\[
\mathcal{F}_{2D}(g)(\xi, o) = \mathcal{F}_{3D}(x)(\xi, \zeta),
\]

\[
\zeta = h^{-1} \xi \top o,
\]

where \( \mathcal{F}_{2D} \) and \( \mathcal{F}_{3D} \) denote the 2D and 3D Fourier transforms, respectively, and \( o = r - s \) signifies the vector difference between the detector position \( r \) and the source position \( s \). 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 \( o \), i.e., the forward operator samples from \( \zeta = h^{-1} \xi \top o \). Since the range of \( o \) 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](image.png)

**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 \( \mathcal{F}_{2D}(g)(\xi, o) \), which are limited in \( \zeta \) by \( h^{-1} \xi \top o_{\text{max}} \). The limited aperture in \( o \) 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 $A_z$ which maps a 2D image at depth $z$, denoted by $x_z(u)$, to the corresponding projection data $y(r,s)$ (see Appendix A.3):

$$y(r,s) = (A_z x_z)(s,r) = x_z \left( 1 - \frac{1}{h} - \frac{1}{z} \right) r + \frac{1}{h} z s,$$

(5)

the corresponding adjoint operator $A_z^*$ is given by (see Appendix A.3):

$$\left( A_z^* y \right)(u) = \frac{1}{1 - \frac{1}{h} - \frac{1}{z}} \int y \left( 1 - \frac{1}{h} - \frac{1}{z} \right) r + \frac{1}{h} z s ds,$$

(6)

we then find that $A_z^* A_z$ is given by

$$\left( A_z^* A_z x_z \right)(u) = \frac{1}{1 - \frac{1}{h} - \frac{1}{z}} x_z(u),$$

(7)

so that for a limited range of $s$, $A_z^* A_z \propto I$ 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 $A_z^* A_z$ (see Appendix A.4). In particular, we expect to image layered objects, which we can model as

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

using a finite number of sources $\{s_j\}_{j=1}^{N}$, the depth image of the $k$th layer is then given by

$$\hat{x}_k(u) = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i \left( \frac{1}{1 - z_i} u + \frac{z_i - z_k}{1 - z_k} s_j \right),$$

(8)

$$= x_k(u) + N^{-1} \sum_{i \neq k} \sum_{j=1}^{N} x_i \left( \frac{1}{1 - z_i} u + \frac{z_i - z_k}{1 - z_k} s_j \right),$$

(9)

Hence, the error in the depth image of the $k$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.

So far, it is evident that imaging a specific depth of a flat object leads to artifacts, as demonstrated by the derived depth image in Equation (9). Furthermore, these artifacts are influenced by the positions of the X-ray sources. The source design problem is now to identify a configuration of source positions, $s_j$, that minimize unwanted features in the depth image. As we do not 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 \( u \) at which we evaluate the output.

3. Discrete Problem

In this section, we explore depth imaging in the discrete domain, discuss the challenges that arise from it, and present the algebraic reconstruction technique for addressing these challenges.

3.1. Forward and Back Projection

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

\[
y = Ax,
\]

where \( y \in \mathbb{R}^m \), \( x \in \mathbb{R}^n \), and \( A \in \mathbb{R}^{m \times n} \). Note that the number of rows in \( A \) 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 \( A_z \) and \( x_z \) denote the forward matrix and the slice object at depth \( z \), respectively. In general, \( A_z^\top A_z \) will not be a multiple of the identity as in the continuous setting (7), so we consider solving \( A_z^\top A_z x_z = A_z^\top y \). We can view this as an approximation of 3D reconstruction through the Schur complement [13]. To see this, let \( A_c \) and \( x_c \) represent the forward operator and the object for all pixels except the ones contained in \( x_z \). The linear equations for the forward projection (10) can now be expressed as follows:

\[
\begin{bmatrix}
A_z & A_c
\end{bmatrix}
\begin{bmatrix}
x_z \\
x_c
\end{bmatrix} = y.
\]

By multiplying \( A_z^\top \), it reads

\[
\begin{bmatrix}
A_z^\top A_z & A_z^\top A_c \\
A_z A_z^\top & A_z A_c^\top
\end{bmatrix}
\begin{bmatrix}
x_z \\
x_c
\end{bmatrix} =
\begin{bmatrix}
A_z^\top y \\
A_z A_c^\top y
\end{bmatrix}.
\]

By utilizing the Schur complement, one can re-write this in terms of \( x_z \) as

\[
A' x_z = y',
\]
where
\[
A' = \left( A_z^\top A_z - A_z^\top A_c \left( A_c^\top A_c \right)^{-1} A_c^\top A_z \right),
\]
\[
y' = \left( A_z^\top y - A_z^\top A_c \left( A_c^\top A_c \right)^{-1} A_c^\top y \right).
\]

However, it is not computationally feasible to compute the corresponding Schur complements, and we can interpret depth imaging as neglecting the additional terms involving \( A_c \).

### 3.2. Algebraic Reconstruction

Although the operator \( A_z \) is unitary (up to a scaling factor), the corresponding discrete operator \( A_z \) 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:

\[
\hat{x}_z = \arg\min_{x_z} \frac{1}{2} \| A_z x_z - y \|^2 + \frac{\lambda}{2} \| x_z \|^2,
\]

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 LSQR [14]. Various other regularization techniques, such as Total Variation regularization and sparse regularization, introduce different penalty functions that promote desired properties in the reconstructed image.

### 4. Efficient Source Design for \( K \) Sources

Here, we address the OED problem in laminography through a bilevel optimization framework. Bilevel optimization addresses a specific class of problems characterized by one problem being nested within another. This hierarchical structure comprises two levels of optimization tasks: the upper level (also referred to as the leader problem) and the lower level (known as the follower problem). The solution to the upper-level problem typically depends on the solution to the lower-level problem. Below, we outline the bilevel optimization strategy for the OED problem in laminography [15].

Let \( \beta \in \mathbb{R}^N \) denote 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 bilevel optimization problem whose upper level involves

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

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

\[
\hat{x}_z(\beta) = \arg\min_{x_z} \frac{1}{2} \| T_\beta (A_z x_z - y) \|^2 + \frac{1}{2} \| x_z \|^2,
\]

where \( y \) denotes the data corresponding to \( x_{gt} \) 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 \( \alpha \)) 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 [16], and the bilevel optimization is then formulated as

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

where
\[
A' = \left( A_z^\top A_z - A_z^\top A_c \left( A_c^\top A_c \right)^{-1} A_c^\top A_z \right),
\]
\[
y' = \left( A_z^\top y - A_z^\top A_c \left( A_c^\top A_c \right)^{-1} A_c^\top y \right).
\]
while the optimal value of the scalar parameter \( a \) in (15) is substituted in the above optimization with \( \hat{a}(\beta) = \frac{\hat{x}(\beta) \cdot x_{gt}}{||\hat{x}(\beta)||^2} \). Mathematically speaking, this problem seeks a \( \beta \) that minimizes the Euclidean distance of \( \hat{x}(\beta) \) and \( x_{gt} \) and lies on the intersection of a hyperplane \( \beta^\top 1 = K \) 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 one by setting the \( K \) largest elements to one and assigning zeros to the remaining elements.

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

\[
\min \beta \frac{1}{2M} \sum_{i=1}^M \|\hat{a}(\beta)x_{gt}^i(\beta) - x_{gt}^i\|^2 \quad \text{s.t.} \quad \beta^\top 1 = K, \ 0 \leq \beta \leq 1,
\]

where \( x_{gt}^i(\beta) \) is the solution of (16) for the \( i \)-th calibration data \( y^i \).

In the next section, we will present an algorithm for solving the bilevel optimization problem (17) using a gradient-based method.

5. Implementation: Projected Gradient Approach

To address the implementation of the Efficient K-Source Design (EKSD) method, it is important to note that the non-linear optimization problem specified in Equation (17) requires the use of numerical optimization algorithms that begin with an initial guess and iteratively refine the solution. This section discusses the solution of (17) employing a projected gradient method.

In general, the projected gradient (PG) method solves

\[
\min \beta f(\beta) \quad \text{s.t.} \quad \beta \in B,
\]

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

\[
\mathcal{P}(u) = \arg\min_x \|x - u\|^2 \quad \text{s.t.} \quad x \in B.
\]

The PG algorithm is defined by

\[
\beta_{k+1} = \mathcal{P}(\beta_k - \gamma \nabla f(\beta_k))
\]

where \( \gamma \in (0, 2/L) \) is the step size, \( L \) is the Lipschitz constant of \( \nabla f \), and \( \nabla f(\beta_k) \) is the gradient of \( f \) at \( \beta_k \). The stopping criteria for the PG method is determined by the Euclidean norm of the difference between consecutive iterates, denoted as \( \|\beta_{k+1} - \beta_k\| \), being less than or equal to a specified threshold value, denoted as \( \epsilon \). This condition ensures that the optimization process is halted at an approximate solution point where further iterations do not significantly alter the solution [16]. In the following subsection, we will outline the gradient calculation for the objective function and describe a method for projecting onto a simplex.
5.1. Gradient of the Objective Function

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

$$
\nabla F(\beta) = \hat{\alpha}(\beta) \frac{\partial \hat{x}(\beta)}{\partial \beta} (\hat{\alpha}(\beta) \hat{x}(\beta) - x_{gt})
= 2\hat{\alpha}(\beta) C^\top \left( A_z^\top T_{\beta}^\top \right)_\lambda (\hat{\alpha}(\beta) \hat{x}(\beta) - x_{gt}),
$$

where \( (A_z^\top T_{\beta}^\top)_\lambda \) is required. Here, it reads

\[
\left( A_z^\top T_{\beta}^\top \right)_\lambda = \left( A_z^\top T_{\beta}^\top T_{\beta} A_z + \lambda I \right)^{-1} A_z^\top T_{\beta}^\top, \quad \text{and} \quad C = [c_1, \ldots, c_N] \text{ with } c_i = \frac{\partial T_{\beta}}{\partial \beta_i},
\]

(See the derivation in Appendix A.6).

5.2. Projection onto the Simplex Constraint

Here, we investigate the projection operator $P$ presented in (21) which maps each point onto the simplex set of $B = \{ u \in \mathbb{R}^m \mid u^\top 1 = K, \ 0 \leq u_i \leq 1 \}$. (See the derivation in Appendix A.6)

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

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

$$
\sum_i \min \{ \max \{ p_i - \mu, 0 \}, 1 \} = K,
$$

which can be solved using the Newton method. So far, we have examined the implementation of the OED method via a projected gradient approach. The following section presents the pseudocode for our proposed algorithm.

6. EKSD Algorithm

In this section, we outline the pseudocode in (Algorithm 1) for implementing the proposed bilevel optimization in (17) through the projected gradient method.

**Algorithm 1** Pseudocode for Efficient K-Source Design algorithm.

Require: $K, \beta^0, \gamma, \lambda, \epsilon$
1: \( i \leftarrow 0 \)
2: repeat
3: \( \hat{x}_i(\beta^i) \leftarrow \text{LSQR}(T_{\beta^i}(A_z), T_{\beta^i}(y), \lambda), \quad \triangleright \text{using Equation (16)} \)
4: \( \hat{\alpha}^i \leftarrow \frac{\hat{x}_i(\beta^i)^\top x_{gt}}{\| \hat{x}_i(\beta^i) \|^2}, \)
5: \( C := [c_1, \ldots, c_N], \ c_j \leftarrow \frac{\partial T_{\beta^i}}{\partial \beta_j}(y - A_z \hat{x}_i(\beta^i)), \)
6: \( r \leftarrow (\hat{\alpha}^i \hat{x}_i(\beta^i) - x_{gt}), \)
7: \( \nabla F(\beta^i) \leftarrow 2\hat{\alpha}^i C^\top \left( A_z^\top T_{\beta^i}^\top \right)_\lambda r, \)
8: \( \beta^{i+1} \leftarrow P(\beta^i - \gamma \nabla F(\beta^i)), \)
9: \( i \leftarrow i + 1, \)
10: until $\| \beta^{i+1} - \beta^i \| < \epsilon$

7. Numerical Experiments

Here, we present several results related to our proposed algorithm. Our goal is to evaluate the algorithm’s performance across two distinct initialization scenarios for the design parameters, focusing on convergence behavior. The performance is assessed based on the following criteria:
i: Normalized mean square error (NMSE): it can be defined as the error between the reconstructed image and ground truth

\[
\text{NMSE} := \min_{\hat{\alpha}} \frac{\|\alpha^{\hat{\alpha}}_z - x_{\text{gt}}\|^2}{\|x_{\text{gt}}\|^2}.
\] (25)

ii: Structural similarity index measure (SSIM): Unlike the NMSE, the 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 [17].

7.1. 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 6\(L\). 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 [11]. For illustration, we present the projection images from all sources in a \(4 \times 4\) array of sources in Figure 5. It is challenging to inspect individual layers of the phantom with a single projection, as overlapping them hinders a proper investigation.

We chose 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. We included additional plots in Appendix A.1 to illustrate the sensitivity of the algorithm to variations in the parameters. 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 or with i.i.d. uniform random values in \([0, 1]\). We refer to these methods as EKSD(\(K\)) and EKSDR(\(K\)), respectively. In several 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.

7.2. 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 6) 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 6a.

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 7. We use half of this data set for source design (calibration) and half for validation.
Figure 5. Different views from sources in a $4 \times 4$ array of sources.

Figure 6. The wire designs with a size of $250 \times 250$. (a) Bottom layer of the PCB. (b) Top layer of the PCB.
7.3. Convergence

We evaluate the convergence behavior of the EKSD algorithm across various iterations. Here, we let $K = 10$, $N = 16$, and $D = 4L$. In Figure 8, we depict the NMSE at each iteration as evaluated on the calibration (left) and validation (right) data. This figure 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 in the NMSE at each iteration.

7.4. 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. Next, we will examine the results for array lengths of $2.4L$ and $4L$, which serve as examples of optimal and suboptimal array lengths, respectively.
Figure 9. The reconstruction error for various ranges of the area occupied by the source.

7.5. Exhaustive Search

Before evaluating the OED method for a large number of source positions, we aimed to assess its efficiency by comparing it with a combinatorial optimization method. We carried out an exhaustive search to assess the performance of EKSD. We considered $N = 16$ 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 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.

7.6. 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 the case using all source positions. This enhancement is largely attributed to the reduction in imaging artifacts, which likely results from the selective measurement approach inherent in EKSD and EKSDR strategies. As a comparison, we show the average NMSE and SSIM of random designs with $K = 10$ sources, which significantly underperform compared to the other methods.

![Image](image-url)

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.

<table>
<thead>
<tr>
<th>Source Design</th>
<th>$N = 16$</th>
<th>$N = 36$</th>
<th>$N = 64$</th>
<th>$N = 16$</th>
<th>$N = 36$</th>
<th>$N = 64$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL ($K = N$)</td>
<td>0.1749</td>
<td>0.1632</td>
<td>0.1603</td>
<td>0.5630</td>
<td>0.5921</td>
<td>0.6046</td>
</tr>
<tr>
<td>EKSD ($K = 10$)</td>
<td>0.1611</td>
<td>0.1508</td>
<td>0.1528</td>
<td>0.6275</td>
<td>0.7304</td>
<td>0.7397</td>
</tr>
<tr>
<td>EKSDR ($K = 10$)</td>
<td>0.1611 ± $3 \times 10^{-4}$</td>
<td>0.1508 ± $4 \times 10^{-4}$</td>
<td>0.1518 ± $1.1 \times 10^{-3}$</td>
<td>0.6275 ± $3 \times 10^{-4}$</td>
<td>0.7304 ± $4 \times 10^{-4}$</td>
<td>0.7380 ± $4.6 \times 10^{-3}$</td>
</tr>
<tr>
<td>RANDOM ($K = 10$)</td>
<td>0.1819 ± $7 \times 10^{-3}$</td>
<td>0.1781 ± $1.1 \times 10^{-2}$</td>
<td>0.1759 ± $1.4 \times 10^{-2}$</td>
<td>0.57 ± $1.5 \times 10^{-2}$</td>
<td>0.6075 ± $2.1 \times 10^{-2}$</td>
<td>0.6248 ± $3 \times 10^{-2}$</td>
</tr>
</tbody>
</table>
Figure 11 shows the corresponding depth images of the bottom layer reconstructed with different source positions. As demonstrated in Section 2.3, artifacts in these images are attributed to the superposition of features from different layers. The lack of rotational acquisition in laminography leads to these 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](image)

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

### 7.7. Source Design 2

We repeat the previous experiment with the array length optimally chosen 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.

<table>
<thead>
<tr>
<th>Source Design</th>
<th>( N = 16 )</th>
<th>( N = 36 )</th>
<th>( N = 64 )</th>
<th>( N = 16 )</th>
<th>( N = 36 )</th>
<th>( N = 64 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL (( K = N ))</td>
<td>0.1467</td>
<td>0.1454</td>
<td>0.1459</td>
<td>0.7030</td>
<td>0.7174</td>
<td>0.7233</td>
</tr>
<tr>
<td>EKSD (( K = 10 ))</td>
<td>0.1488 ± 5 \times 10^{-4}</td>
<td>0.1484 ± 9 \times 10^{-4}</td>
<td>0.1530 ± 2.5 \times 10^{-3}</td>
<td>0.7230 ± 5 \times 10^{-4}</td>
<td>0.7259 ± 2.7 \times 10^{-3}</td>
<td>0.7240 ± 4.7 \times 10^{-3}</td>
</tr>
<tr>
<td>EKSDR (( K = 10 ))</td>
<td>0.1537 ± 4 \times 10^{-3}</td>
<td>0.1580 ± 5 \times 10^{-3}</td>
<td>0.1614 ± 8 \times 10^{-3}</td>
<td>0.7040 ± 1 \times 10^{-2}</td>
<td>0.7178 ± 1.1 \times 10^{-2}</td>
<td>0.7237 ± 1.5 \times 10^{-2}</td>
</tr>
<tr>
<td>RANDOM (( K = 10 ))</td>
<td>0.1537 ± 4 \times 10^{-3}</td>
<td>0.1580 ± 5 \times 10^{-3}</td>
<td>0.1614 ± 8 \times 10^{-3}</td>
<td>0.7040 ± 1 \times 10^{-2}</td>
<td>0.7178 ± 1.1 \times 10^{-2}</td>
<td>0.7237 ± 1.5 \times 10^{-2}</td>
</tr>
</tbody>
</table>
8. Discussion

In this research, we explore the methodology of experimental design for a lamino-graphic setup, acknowledging potential challenges that may arise in real-world applications. Here, we outline several significant challenges with our proposed method: (i) As observed in the differences in EKSD and EKSDR results, the non-convex nature of the objective function may lead to entrapment in local minima. To mitigate this issue, one could explore the use of optimization methods that thoroughly explore the entire search space, such as the Genetic Algorithm [18], which can potentially provide more robust solutions by avoiding suboptimal local minima. However, it is worth mentioning that these methods are restricted to low-dimensional problems. (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 current investigation of the experimental design problem uses 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.

9. Conclusions

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 bilevel optimization problem. To solve this problem, we employed bilevel optimization and implemented it using gradient-based methods. The suggested framework can be tailored more specifically to a particular application by incorporating different regularization terms in the lower-level optimization problem or by adopting the alternate loss function in the upper-level optimization problem to more effectively capture the characteristics of the image.
Author Contributions: Conceptualization, T.v.L.; methodology, H.F. and T.v.L.; software, H.F.; formal analysis, H.F.; writing—original draft, H.F.; writing—review & editing, T.v.L.; visualization, H.F.; supervision, T.v.L. All authors have read and agreed to the published version of the manuscript.

Funding: Our work was supported by funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 956172 (xCTing).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Parameter Tuning Plots

In Figure A1, we provide the convergence behavior of our method for distinct values of $\lambda$ and $\gamma$, presenting both the average values (represented by solid points) and the standard deviations (illustrated as shaded areas), with respect to the normalized mean square error (NMSE) in each iteration.

![Parameter Tuning Plots](image)

Figure A1. Various $\gamma$ and $\lambda$ applied to the Algorithm 1 and results in different convergence speed.

Appendix A.2. Forward Operator for the 3D Object

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

$$y(r, s) = \int x(s + (r - s)t, h - ht) \, dt$$  \hspace{1cm} (A1)

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

$$x(u, z) = \int \int F_{3D}(x)(\xi, \zeta)e^{j2\pi(\xi^\top u + \zeta^\top z)} \, d\xi \, d\zeta,$$  \hspace{1cm} (A2)
We then define the forward operator \( A \) as projecting a 2D object to the detector at \( z = 0 \) via
\[
y(r, s) = x_z \left( s + \left( 1 - h^{-1}z \right) (r - s) \right)
\] (A8)
or, equivalently, as
\[
y(r, s) = x_z \left( r \left( 1 - h^{-1}z \right) + h^{-1}zs \right).
\] (A9)

The adjoint operator \( A_z^* \) can be obtained by considering the definition \( \langle A_z x_z, y \rangle_{(s,r)} = \langle x_z, A_z^* y \rangle_{(v)} \). First, we have
\[
\langle A_z x_z, y \rangle_{(s,r)} = \int \int x_z \left( r \left( 1 - h^{-1}z \right) + h^{-1}zs \right) y(r, s) dr ds.
\] (A10)

Now, let \( v = r \left( 1 - h^{-1}z \right) + h^{-1}zs \); then, \( r = \frac{v - h^{-1}zs}{1 - h^{-1}z} \) and \( dr = \frac{1}{1 - h^{-1}z} dv \); and it follows that
\[
\langle A_z x_z, y \rangle_{(s,r)} = \frac{1}{1 - h^{-1}z} \int x_z(v) y \left( \frac{v - h^{-1}zs}{1 - h^{-1}z}, s \right) dv ds = \langle x_z, A_z^* y \rangle_{(v)}.
\] (A11)

We thus find that \( A_z^* f \) is defined as
\[
(A_z^* y)(u) = \frac{1}{1 - h^{-1}z} \int y \left( u - \frac{h^{-1}zs}{1 - h^{-1}z}, s \right) ds.
\] (A12)
By using (A9) in (A12), the composition $A_z^*A_z$ can be written as below:

$$\left(A_z^*A_z \circ x_z\right)(u) = \frac{x_z(u)}{1 - h^{-1}z} \int ds. \quad (A13)$$

**Appendix A.4. Back Projection for a Depth Image**

We use $y(r, s)$ in (A1) to express $A_z^*A$ using (A12):

$$\left(A_z^*A\right)(u) = \frac{1}{1 - h^{-1}z} \int x \left(\frac{u - h^{-1}zs}{1 - h^{-1}z} + (1 - t)s, h(1 - t)\right) dt ds. \quad (A14)$$

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

$$\left(A_z^*A\right)(u) = \frac{1}{h - z} \int x \left(\frac{h - z'}{h - z}u + \frac{z' - z}{h - z} s, z'\right) dz' ds \quad (A15)$$

**Appendix A.5. Gradient Calculation of the Objective Function**

To find the gradient of $F(\beta) := \frac{1}{2} \| \hat{h}(\beta) \hat{x}(\beta) - x_{gt} \|^2$ with respect to $\beta$, by the chain roll it reads

$$\frac{\partial F}{\partial \beta} (\beta) = \frac{\partial F}{\partial \hat{x}} \cdot \frac{\partial \hat{x}}{\partial \beta} + \frac{\partial F}{\partial \hat{h}} \cdot \frac{\partial \hat{h}}{\partial \beta} \quad (A16)$$

where the second term is equal to zero because $\hat{h}(\beta)$ is the optimal point of $F$. Then, the first term can be written as

$$\frac{\partial F}{\partial \beta} (\beta) = \left(\frac{\partial \hat{x}}{\partial \beta}\right)^\top \frac{\partial F}{\partial \hat{x}} = \hat{h}(\beta) \left(\frac{\partial \hat{x}}{\partial \beta}\right)^\top \left(\hat{h}(\beta) \hat{x}(\beta) - x_{gt}\right). \quad (A17)$$

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

$$\hat{x}(\beta) = \left(A_z^\top T_\beta^\top T_\beta A_z + \lambda I\right)^{-1} A_z^\top T_\beta^\top T_\beta y. \quad (A18)$$

We define $H_\beta := A_z^\top T_\beta^\top T_\beta A_z + \lambda I$, and then the derivative follows:

$$\frac{\partial \hat{x}(\beta)}{\partial \beta_i} = \frac{\partial H_\beta^{-1}}{\partial \beta_i} A_z^\top T_\beta^\top y + H_\beta^{-1} \frac{\partial \left(A_z^\top T_\beta^\top T_\beta y\right)}{\partial \beta_i}, \quad (A19)$$

Recall that $\frac{\partial H_\beta^{-1}}{\partial \beta_i} = -H_\beta^{-1} \frac{\partial H_\beta}{\partial \beta_i} H_\beta^{-1}$; therefore,

$$\frac{\partial H_\beta^{-1}}{\partial \beta_i} = -2H_\beta^{-1} A_z^\top T_\beta^\top \frac{\partial T_\beta}{\partial \beta_i} A_z H_\beta^{-1} \quad (A20)$$

Then, (A19) can be rewritten as

$$\frac{\partial \hat{x}(\beta)}{\partial \beta_i} = -2H_\beta^{-1} A_z^\top T_\beta^\top \frac{\partial T_\beta}{\partial \beta_i} \left(A_z H_\beta^{-1} A_z^\top T_\beta^\top y - y\right), \quad (A21)$$
where $H_{\beta}^{-1}A_{\beta}^T T_{\beta}^T T_{\beta}y$ is equal to $\hat{x}(\beta)$ in (A18) and $H_{\beta}^{-1}A_{\beta}^T T_{\beta}^T$ is denoted by $(T_{\beta}A_{\beta})^\dagger$. Then, it can be rewritten:

$$\frac{\partial \hat{x}_z(\beta)}{\partial \beta_i} = 2(T_{\beta}A_{\beta})^\dagger(\frac{\partial T_{\beta}}{\partial \beta_i}(y - A_{\beta}\hat{x}(\beta))) = 2(T_{\beta}A_{\beta})^\dagger c_i, \quad (A22)$$

For simplicity, we denote $\frac{\partial T_{\beta}}{\partial \beta_i}(y - A_{\beta}\hat{x}(\beta))$ as $c_i \in \mathbb{R}^m$ and it follows that

$$\frac{\partial \hat{x}_z(\beta)}{\partial \beta} = 2(T_{\beta}A_{\beta})^\dagger C \quad (A23)$$

where $C = [c_1, \ldots, c_N]$. The replacement of (A23) into (22) yields

$$\nabla F = 2\hat{h}(\beta)C^T \left(A_{\beta}^T T_{\beta}^T\right)^\dagger (\hat{h}(\beta)\hat{x}_z(\beta) - x_{gl}). \quad (A24)$$

Appendix A.6. Projection onto the Simplex Constraint

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

$$\arg\min_t \frac{1}{2}\|t - p\|^2 \text{ s.t } t^T 1 = K, \ 0 \leq t_i \leq 1, \quad (A25)$$

to solve the above optimization via its dual problem, we introduce the Lagrangian [16]:

$$\mathcal{L}(t, \mu) := \frac{1}{2}\|t - p\|^2 + \mu \left(t^T 1 - K\right), \quad (A26)$$

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

$$J(\mu) = \inf_{0 \leq t_i \leq 1} \mathcal{L}(t, \mu), \quad (A27)$$

where $J(\mu)$ is the dual function as the minimum value of the Lagrangian over $t$. Regarding strong duality, the optimal solutions of (A25) and (A27) are equal if there exists a dual variable of $\mu \in \mathbb{R}$ that makes (A27) a tight lower bound for (A25). The solution of the dual problem of (A27) can be written as [19]

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

where $\mu \in \mathbb{R}$ is obtained by solving $t^*(\mu) 1 = K$. Let us define the function $w(\mu) = \sum \min \{p_i - \mu, 0\}, 1\} - K$, which is non-increasing and can be solved using the Newton method. The derivative is given by

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

References


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.