



RESEARCH PROJECT IN COMMUNICATION SYSTEMS
12 CREDITS

Multiresolution Framework for Ptychography

January 3, 2025

Author:
Efe Tarhan

Supervisors:
Jonathan Dong
Vincent P. L. Guillemet

Biomedical Imaging Group
Ecole Polytechnique Fédérale de Lausanne

Abstract

Ptychography is a computational imaging technique that enables high-resolution and quantitative phase imaging by reconstructing complex-valued images from coherent diffraction patterns. However, the reconstruction process poses significant computational challenges due to the non-convexity of the optimization problem and the high dimensionality of the data. This report presents a novel multiresolution framework inspired by Partial Differential Equations (PDEs) and wavelets to improve the efficiency and reliability of ptychographic image reconstruction. The proposed method leverages box-splines to represent the high-resolution image as a hierarchy of basis functions, allowing for iterative refinement from low to high frequencies. By incorporating redundancy and gradual resolution enhancement, the framework mitigates the risks of local minima and accelerates convergence. Experimental validation on biological microscopy data demonstrates the effectiveness of the approach, offering significant improvements in computational efficiency and reconstruction quality.

Contents

1	Introduction	3
2	Background	4
2.1	Ptychography	4
2.2	Multiresolution Algorithm	6
2.2.1	Basis Function Framework	6
2.2.2	Discretization	7
2.2.3	Optimization Objective	7
2.2.4	Regularization: HTV	8
2.2.5	Algorithm	8
2.3	A Note on the Choice of Ptychography	9
3	Methods	9
3.1	Ptychography with Multiresolution	9
3.2	Linear Problem Setup	11
3.3	Probe Parameters	12
3.3.1	Probe Type	12
3.3.2	Intersection Ratio of the Probes Between Measurements	13
3.4	Nonlinear Problem Setup	14
4	Results	15
4.1	Linear Problem Results	15
4.2	Nonlinear Problem Results with CPWL Functions	16
4.3	Nonlinear Problem Results with CPC Functions	19
5	Future Work	22
6	Conclusion	24
A	Efficient Forward Implementation	25
A.1	Forward Operator	25

1 Introduction

Ptychography is a state-of-the-art imaging technique that combines computational algorithms with optical microscopy to achieve high-resolution, complex-valued reconstructions of samples. One of its key advantages over conventional microscopy techniques is that it does not require high-quality lenses with high diffraction efficiency to produce high-quality images of large fields of view. Instead, ptychography relies on oversampling and algorithmic reconstruction [1]. By capturing a series of diffraction patterns as the sample is scanned through overlapping regions, ptychography enables the synthesis of a high-resolution image with a large field of view, while also allowing for the quantitative recovery of phase information. This capability is particularly valuable for studying biological cells. A major challenge in obtaining high-quality reconstructions arises from the *phase retrieval* problem, where the phase of light waves is lost during the imaging process and must be accurately recovered. This is achieved by leveraging the captured diffraction patterns. Consequently, ptychography can achieve spatial resolutions that surpass the diffraction limit of the objective lens [2]. The structure of the measurements used for the phase retrieval task in order to solve the reconstruction process can be described as follows:

$$y = |Ax|^2, \quad (1)$$

where y represents the measurements, A is the ptychography operator, and x is the signal to be sampled.

The reconstruction process in ptychography is computationally challenging due to the non-convex nature of the underlying optimization problem, which often leads to slow convergence and susceptibility to local minima. This project addresses these challenges by proposing a multiresolution framework inspired by Partial Differential Equations (PDEs) and wavelets, leveraging box-splines to progressively refine the reconstruction from low to high frequencies. The essence of this approach lies in the progressive refinement of the target image using a hierarchical representation based on box-splines, which serve as a basis for modeling the image. The recovery of an unknown signal from noisy measurements, an inverse problem, typically involves formulating the reconstruction as an optimization problem. Following the continuum formulation [3], the inverse problem can be expressed as:

$$\arg \min_{f \in X} (E(\nu(f), y) + \lambda R(f)), \quad (2)$$

where f is the unknown signal, X is an appropriate search space, y represents the measurements, E quantifies data fidelity, and R is a regularizer incorporating prior knowledge. The regularizer acts in the continuous domain, ensuring well-posedness and enabling representation at any resolution. To solve this problem, the multiresolution framework starts with a coarse representation of the signal and progressively incorporates finer details through a series of basis functions such as box-splines. These functions allow the image to be represented as a combination of low- and high-frequency components, facilitating efficient and accurate reconstruction.

By employing a multiresolution strategy, the framework reduces the computational burden associated with solving the non-convex optimization problem. Starting

with a coarse approximation of the image, the algorithm incrementally integrates higher-frequency details, enabling faster convergence and mitigating the risks associated with local minima. The hierarchical nature of this approach allows for a natural balance between computational efficiency and reconstruction accuracy. Moreover, the use of box-splines ensures that the solution space is well-defined, offering a structured and effective way to address the challenges of ptychographic image reconstruction.

This project brings together the domains of ptychography and multiresolution to offer an alternative and efficient solution for high-resolution image reconstruction. By integrating principles from Partial Differential Equations (PDEs), wavelets, and box-splines modeling, the project develops a cohesive framework that bridges existing tools for ptychographic simulation and multiresolution techniques. The resulting algorithm is designed to efficiently address the computational challenges inherent in ptychographic reconstruction.

The report is structured as follows: following this introduction, the theoretical background on ptychography and multiresolution frameworks is presented in Section 2. Section 3 outlines the methods used, including the formulation of the linear and nonlinear problems, the probe parameter adjustments, and the multiresolution approach. The results of the linear and nonlinear reconstructions are detailed and analyzed in Section 4. Finally, the report concludes with a discussion of recent improvements, the challenges encountered and potential future directions for further development in Appendix A.

2 Background

In this section, the concepts and tools used to establish a baseline for the project are briefly presented to facilitate an understanding of the individual components.

2.1 Ptychography

Ptychography is a computational imaging technique that enables high-resolution, quantitative phase imaging by reconstructing complex-valued images from coherent diffraction patterns. Originally introduced by Hoppe in 1969, ptychography addresses the phase problem by combining diffraction data from multiple overlapping regions of the sample. This overlapping measurement strategy enables the recovery of lost phase information, a critical requirement for reconstructing the full complex-valued sample image [1].

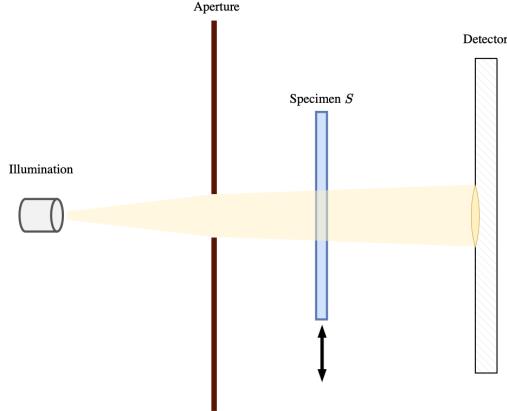


Figure 1: Ptychography measurement setup.

The forward operation in ptychography can be expressed mathematically. The pupil function $p : \mathbb{R}^2 \rightarrow \mathbb{C}$ models the aperture and is compactly supported on a square. The moving sample can be represented as a complex-valued function $x(., \mathbf{s}) : \Omega_X \times \mathbb{R}^2 \rightarrow \mathbb{C}$, where $\Omega_X \subset \mathbb{R}^2$ represents the region of interest of the sample, and \mathbf{s} is the amount of shift of the sample for each measurement.

The Fourier transform operator is denoted by \mathcal{F} . The forward model for the measured intensity can be expressed as [4]:

$$y_{\mathbf{s}}(\mathbf{r}) = |\mathcal{F}\{p(x, y) \cdot x(., \mathbf{s})\}(\mathbf{r})|^2, \quad (3)$$

where $y_n(\mathbf{r})$ represents the measured intensity at the detector, $p(x, y)$ is the pupil function, $x_n(x, y)$ is the sample function for the n -th measurement, and (s_x, s_y) is the spatial shift of the sample.

Although this measurement encodes both the amplitude and phase information of the sample, only the intensity is directly observed. Consequently, recovering the phase information becomes essential for reconstructing the sample image x .

In its discrete form, this operation can be simplified as:

$$A = F \cdot P \cdot S \cdot C, \quad (4)$$

$$y = |Ax|^2, \quad (5)$$

where:

- F : FFT (Fast Fourier Transform) matrix, representing the Fourier transform of the signal resulting from the optical properties of far-field illumination.
- P : Probe matrix, representing the geometrical aperture of the probe function, which is a square in this case.
- S : Shift matrix, representing the shift of the input. This is a space-shift-invariant feature and can also be interpreted as the entire setup being shifted.

- C : Crop matrix, used because the aperture has a finite support. It reduces the region of interest to a smaller area to optimize memory usage.

This formulation will be used throughout the report to represent the full ptychographic forward operation and its components in both continuous and discrete domains.

To address the challenges of reconstructing x in ptychography, a multiresolution algorithm will be employed. This approach progressively refines the reconstruction by starting at coarser scales and incorporating finer details iteratively, offering improved computational efficiency and stability in solving the non-convex optimization problem. The details of this algorithm and its implementation will be discussed in the next section.

2.2 Multiresolution Algorithm

The multiresolution algorithm leverages the hierarchical properties of CPWL (Continuous and Piecewise Linear) functions, parameterized through box splines, to iteratively refine the solution at multiple resolutions. This approach provides a robust framework for solving high-dimensional, non-convex optimization problems while maintaining computational efficiency [3, 5].

2.2.1 Basis Function Framework

A CPWL (Continuous and Piecewise-Linear) function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is defined on a domain that is partitioned into a set of non-overlapping polytopes $\{P_k\}$. Within each polytope, $f(x)$ behaves as a linear function. Formally, $f(x)$ can be expressed as:

$$f(x) = \sum_k (a_k^\top x + b_k) \mathbf{1}_{P_k}(x), \quad (6)$$

where:

- a_k and b_k are parameters defining the affine function within the polytope P_k ,
- $\mathbf{1}_{P_k}(x)$ is the indicator function, which is 1 if $x \in P_k$ and 0 otherwise.

Box splines are used as basis functions for parameterizing CPWL functions. These splines have several important properties:

- Compact Support: Each box spline affects only a localized region of the domain, making computations efficient.
- Refinability: Box splines support hierarchical (coarse-to-fine) representations through dyadic scaling.
- Exact Representation: Refinability ensures that functions represented at a coarse scale can be exactly represented at finer scales.

A box spline $B_{\Xi_0}(x)$ is defined recursively. For example, for the piecewise-constant box spline:

$$B_{\Xi_0}(x) = \begin{cases} \frac{1}{|\det \Xi_0|}, & x = \sum_{r=1}^d t_r \xi_r \text{ for } t_r \in [0, 1], \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Higher-order box splines are defined recursively as:

$$B_{\Xi_p}(x) = \int_0^1 B_{\Xi_{p-1}}(x - t\xi_{d+p}) dt. \quad (8)$$

The refinement of CPWL functions is based on the dyadic refinement property of box splines:

$$\phi\left(\frac{x}{2}\right) = \frac{1}{2} \sum_{k \in \{0,1\}^d} \phi(x + k). \quad (9)$$

This property ensures that a CPWL function represented at a coarse scale s can be exactly represented at a finer scale $s + 1$. The search spaces at different resolutions are nested:

$$\dots \subseteq X^{(s-1)} \subseteq X^{(s)} \subseteq X^{(s+1)} \subseteq \dots, \quad (10)$$

where $X^{(s)}$ is the space of CPWL functions spanned by the box spline basis at scale s [3].

2.2.2 Discretization

For a linear forward operator ν , the discretization at scale s expands the CPWL function in terms of its basis coefficients c_s :

$$v_m(f_s) = \langle \nu_m, f_s \rangle = \sum_{n=1}^{N_s} c_{s,n} \langle \nu_m, \phi_{s,n} \rangle = h_{s,m}^\top c_s. \quad (11)$$

The forward operator at scale s is represented as:

$$A_s = A_{s+1} U_s, \quad (12)$$

where U_s is the matrix that refines coefficients from scale s to $s + 1$ [3].

For Fourier-domain operators, the box spline basis is evaluated in the Fourier domain:

$$\langle \nu_m, \phi_{s,n} \rangle = 2^{-sd} \hat{\phi}(2^{-s} \omega_m) e^{-j2^{-s} \omega_m^\top \text{idx}_s^{-1}[n]}. \quad (13)$$

2.2.3 Optimization Objective

The multiresolution optimization framework minimizes a cost functional at each scale:

$$\mathcal{L}_s(\tilde{x}_{s_s}) = \frac{1}{2} \|y - A_s \tilde{x}_{s_s}\|_2^2 + \lambda R_s(\tilde{x}_{s_s}), \quad (14)$$

where $R_s(\tilde{x}_{s_s})$ is a regularizer applied to the input \tilde{x}_{s_s} at the discrete scale s_s , specifically using Hessian Total Variation (HTV) in this project.

1. Initializing at a coarse scale s_0 with coefficients \tilde{x}_{s_0} .
2. Solving the optimization problem at scale s_s .
3. Refining the coefficients to the next scale: $\tilde{x}_{s_{s+1}} = U_s \tilde{x}_{s_s}$.
4. Repeating until the optimization is concluded at the finest scale s_{fine} .

2.2.4 Regularization: HTV

Hessian Total Variation (HTV) is a second-order extension of total variation, characterizing the second-order variation of a function f . It is particularly effective for promoting sparsity in second-order derivatives, making it well-suited for CPWL functions, where second-order derivatives vanish almost everywhere except at boundaries.

The Hessian matrix of a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is:

$$H\{f\} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix}. \quad (15)$$

For CPWL functions, the Hessian is defined in the distributional sense. The second-order directional derivative along a direction u is given by $D_u^2 f(x) = u^\top H\{f\}(x)u$, and using the eigenvectors v_q and eigenvalues λ_q of $H\{f\}$, we have:

$$D_u^2 f(x) = \sum_{q=1}^d t_q^2 \lambda_q, \quad (16)$$

where $u = \sum_{q=1}^d t_q v_q$. This fully characterizes the second-order directional derivatives of f . The explicit equation for the HTV regularizer $R(c)$ is:

$$\text{HTV}(f) = \int_{\mathbb{R}^2} \|\mathbf{H}_f(x, y)\|_{S_1} dx dy = \int_{\mathbb{R}^2} \sum_{r=1}^2 \sigma_r(\mathbf{H}_f(x, y)) dx dy. \quad (17)$$

By using HTV as a regularizer, we promote configurations where the second-order derivatives of f vanish almost everywhere. This is particularly relevant for CPWL functions, where HTV penalizes changes in gradients across polytope boundaries, ensuring sparsity in second-order variations and favoring piecewise-affine reconstructions. This makes HTV an effective regularization tool in high-dimensional inverse problems.

2.2.5 Algorithm

The multiresolution algorithm operates iteratively, starting from a coarse grid and progressively introducing finer details. At each scale, the problem is formulated as:

$$\arg \min_{f \in X^{(s)}} (E(\nu(f), y) + \lambda R(f)), \quad (18)$$

where E quantifies data fidelity, $R(f)$ is the regularization term, and λ balances the two [3].

The specific multiresolution algorithms and their implementation details can be found in the paper "A Box-Spline Framework for Inverse Problems With Continuous-Domain Sparsity Constraints" by Pourya et al. [3]. These algorithms leverage the hierarchical structure of box splines to efficiently solve high-dimensional inverse problems by iteratively refining the representation of the solution across scales.

The combination of the ptychography problem and the multiresolution algorithm, incorporating box spline representations, will be detailed in the next section.

2.3 A Note on the Choice of Ptychography

Initially, the project was designated for the implementation of the multiresolution framework on the Fourier Ptychography method, which has the following forward operation [4]:

$$y_s(\mathbf{r}) = \left| \mathcal{F}^{-1} \left\{ \hat{p}(\mathbf{k}) \cdot \mathcal{F} [x(\cdot, \mathbf{s}) \cdot e^{j\langle \mathbf{k}, \mathbf{s} \rangle}] (\mathbf{k}) \right\} (\mathbf{r}) \right|^2. \quad (19)$$

However, since this algorithm involves a Fourier transform of the tilted sample image before its multiplication with the probe pupil, the information is compressed primarily into the low-frequency regions of the image. As the multiresolution algorithm applies averaging at coarser scales, a significant amount of information would be lost, leading to severe aliasing and the optimization approach less effective and unstable.

To address this issue, and after discussions with the project supervisors, it was decided to replace Fourier Ptychography with Ptychography. This choice better aligns with the multiresolution framework and avoids the aforementioned challenges.

3 Methods

3.1 Ptychography with Multiresolution

In this section, the integration of the two algorithms is explained in detail. The forward operation A_{s_m} , which depends on the scale s_m at which it is applied, represents the general expression for the ptychography measurement process. Measurements are generated using different diffraction patterns with a probe, taking an input x of shape (1, height, width) and producing measurements y with shape (# of measurements, height, width). This process is illustrated in Figure 2.

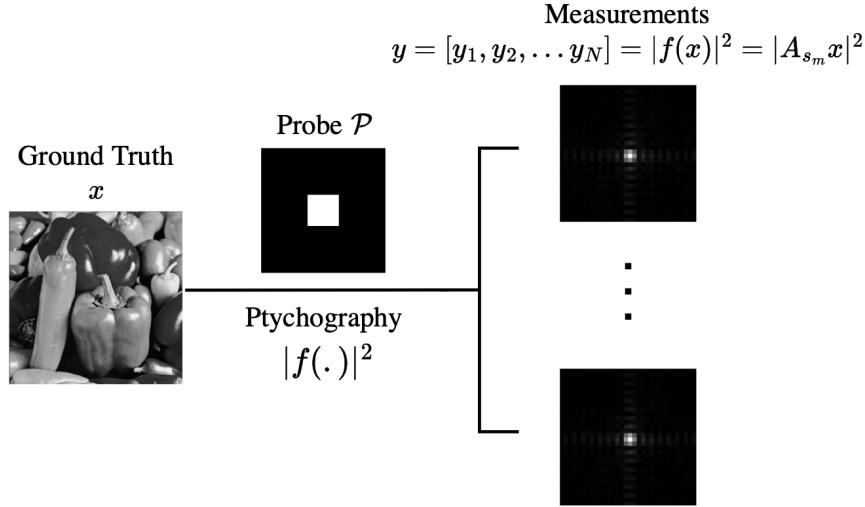
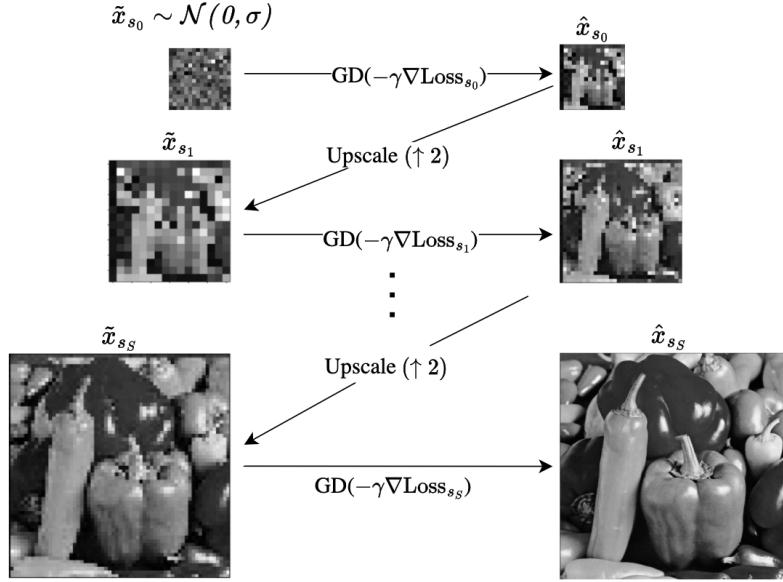


Figure 2: The figure illustrates the ptychographic forward operation. The ground truth image x is illuminated by a probe \mathcal{P} , generating a series of diffraction patterns. The forward operation A_{s_m} computes the ptychographic measurements $y = [y_1, y_2, \dots, y_N]$, where each measurement corresponds to the squared magnitude of the Fourier-transformed signal, $|A_{s_m}x|^2$.

Following the forward operation, the multiresolution optimization framework is applied to the project. As explained in Section 2.2.3, the algorithm begins with an initial random guess of the image at the coarsest scale s_0 . A gradient descent optimization is then performed at this scale using the initial guess and the measurements obtained from the forward operation described above. The optimization process incorporates regularization and additional features, such as a learning rate guided by the gradient ∇Loss_{s_0} . The calculations of these gradients will be provided in the following subsections.

Upon completing the optimization at a given scale, which is determined by a predefined number of epochs or a stopping criterion, the obtained estimate \hat{x}_{s_0} is upscaled to a finer scale by a factor of 2, resulting in the input \tilde{x}_{s_m} for the optimization at scale s_m .

This process is repeated iteratively, progressively refining the resolution until the finest possible scale of the image is reached. The optimization stops based on a predefined stopping criterion. The overall process, naming conventions, and the loss function are illustrated in Figure 3.



$$\text{Loss}_{s_m} := \frac{1}{2} \|y - A\tilde{x}_{s_m}\|_2^2 + \lambda R_{s_m}(\tilde{x}_{s_m})$$

Figure 3: The framework begins with a random initialization $\tilde{x}_{s_0} \sim \mathcal{N}(0, \sigma)$ at the coarsest scale s_0 . Gradient descent is applied iteratively using the loss function Loss_{s_m} , and the result \hat{x}_{s_m} is upscaled to the next finer scale. This process continues until the finest scale s_s , yielding the final reconstruction \hat{x}_{s_s} .

3.2 Linear Problem Setup

At the initial phase of the project, the algorithm and the integrity of its modules were tested using a simplified problem where the detector was assumed to directly capture the complex measurements, i.e.,

$$\hat{x} = \arg \min_{x \in \mathbb{R}^{N \times N}} \|y - Ax\|^2 + \lambda R(x), \quad (20)$$

where y represents the observed measurements, A is the forward operator, and $R(\hat{x})$ is the regularizer with a weighting parameter λ .

In this case, the regularization parameter is non-differentiable, whereas the data fidelity term is differentiable. The gradient of the data fidelity term is given by:

$$\nabla_x \text{Loss} = 2A^H(Ax - y), \quad (21)$$

which will be utilized in the implementation. The gradient of the data fidelity term will be handled using $\text{proj}(\cdot)$ operator which is explained in more detail in [3].

In this setup, the test image is encoded into the phase component of the input image, and the forward operator A is applied to generate the measurements y . This setup was chosen because the problem is linear and convex, making the solution relatively straightforward.

3.3 Probe Parameters

The parameters of the probe objective are among the most critical values to optimize. Various types of probes were tested during the research process, with several properties of the probe being modified and refined at different stages of the research.

3.3.1 Probe Type

The probe shown in Figure 4 was used for the linear case. This probe was generated using a sinc-based defocused pupil function with a circular geometry.

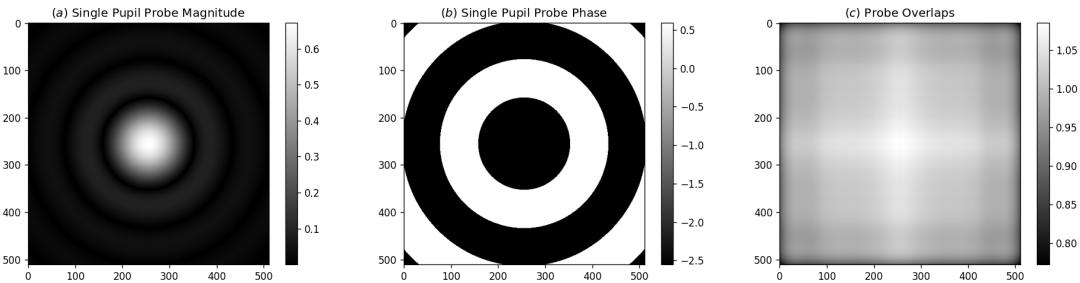


Figure 4: Probe function used with circular defocused pupil where (a) shows the amplitude of the probe, (b) shows the phase of the probe, and (c) is the profile of the overlapped probes in terms of different measurements taken.

The probe function has a circular shape, and while the overlay profile of the probes shown in (c) is sufficient for capturing the necessary information between the cross-sectional areas of the diffraction pattern, it caused problems during the initial trials of the nonlinear case. Due to the limitations of discrete sampling, a perfect representation of the continuous circular probe cannot be achieved. Since the image is initially discretized and fed into the forward process, the discretization introduces artifacts resembling a "staircase" pattern along the circumference of the circular probe. These artifacts, which become more prominent at the coarsest scale, significantly affect the optimization procedure by introducing inaccuracies that hinder the reconstruction quality.

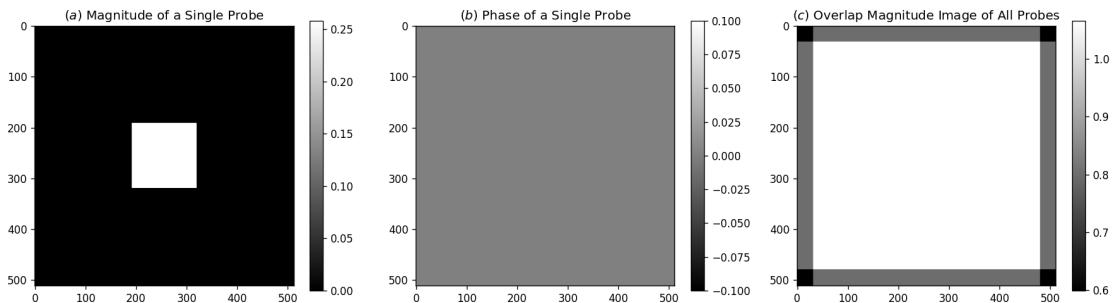
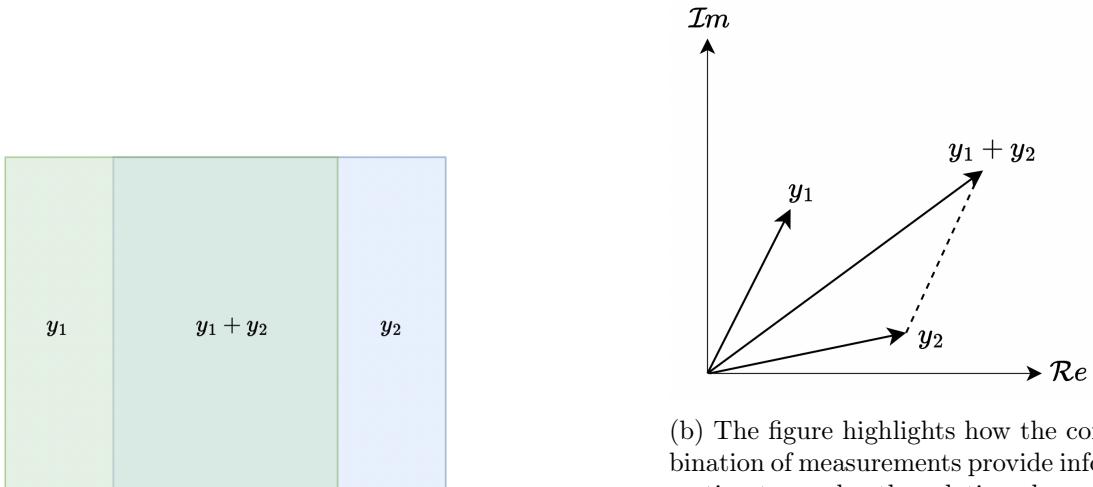


Figure 5: The square probe function is displayed, where (a) shows the amplitude of the probe, (b) illustrates the intersection ratio of two neighboring diffraction patterns, and (c) presents the profile of the overlapped probes based on different measurements taken.

To tackle this problem a square shaped probe with constant phase is decided to be used, shown in Figure 5. Since discrete sampling of the square probe also resembles a perfect square probe it considered to

3.3.2 Intersection Ratio of the Probes Between Measurements

The ptychography problem, as defined in Section 2.1, involves a phase retrieval process that must be addressed. In the typical case, the phases of different measurements y_1, y_2, \dots, y_M are captured as power measurements, resulting in the loss of phase information for all measurements. Since a complete reconstruction requires both magnitude and phase information, additional information must be implicitly or explicitly provided to the system. In ptychography, the diffraction patterns of the measurements generated on the detector plane contain overlap regions, as illustrated in Figure 6a, which serve as a source of the necessary additional information [6].



(a) Overlap region of two probes of consecutive measurements, $|y_1|$, $|y_2|$ and $|y_1 + y_2|$

(b) The figure highlights how the combination of measurements provide information to resolve the relative phases of components. The overlap region introduces additional constraints from combined measurements.

As can be seen, the information from the cross-sectional area can be utilized to address the phase issue in the problem, aiding in the retrieval of the relative phase with respect to the base measurement. To achieve this, it was decided that the overlap ratio would be set to 75%, with each probe shifted by 1/4 of its length. Consequently, the probe structure shown in Figure 7 was selected for the remainder of the research.

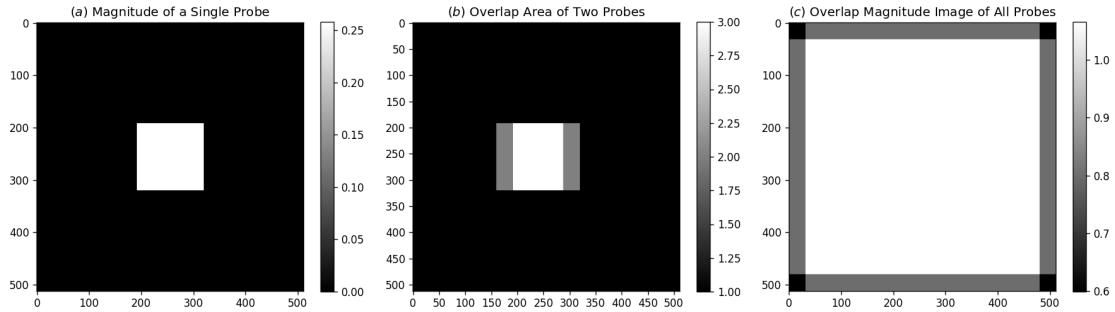


Figure 7: The square probe function is displayed, where (a) shows the amplitude of the probe, (b) illustrates the intersection ratio of two neighboring diffraction patterns, and (c) presents the profile of the overlapped probes based on different measurements taken.

3.4 Nonlinear Problem Setup

After verifying the algorithm's integrity for the linear problem, the research shifted focus to the nonlinear phase retrieval problem. The formulation of the nonlinear problem is as follows:

$$\hat{x} = \arg \min_{x \in \mathbb{R}^{N \times N}} ||y - |Ax|^2||^2 + \lambda R(x), \quad (22)$$

where y represents the observed power measurements, A is the forward operator representing the ptychography process, and $R(\hat{x})$ is the HTV regularizer weighted by the parameter λ .

In this setup, the regularization parameter is non-differentiable, whereas the data fidelity term is differentiable. The gradient of the data fidelity term is given by:

$$\nabla_x \text{Loss} = 2A^H \left(Ax * (|Ax|^2 - y) \right), \quad (23)$$

which will be utilized in the implementation. The gradient of the data fidelity term will be handled using $\text{proj}(\cdot)$ operator which is explained in more detail in [3].

The introduction of the regularization term $R(\hat{x})$ stabilizes the optimization by promoting continuous piecewise linear (CPWL) solutions. However, the problem is non-convex and challenging due to the loss of phase information, necessitating the use of a multiresolution algorithm based on box splines.

As explained in Section 1, the multiresolution algorithm initializes the solution at a coarser scale and incrementally refines it at finer scales, using the results from previous scales to guide the optimization. This approach reduces the risk of local minima and accelerates convergence, especially for high-dimensional, nonlinear problems.

In the next section, the results of the linear problem and the nonlinear problem, after applying the adjustments to the probe as described above, will be presented and analyzed. The adjustments aim to improve the overall reconstruction quality and algorithmic efficiency, particularly in addressing the artifacts caused by the probe structure.

4 Results

4.1 Linear Problem Results

The multiresolution algorithm was applied to solve the linear problem described in Section 3.2. Examples of the resulting samples from this problem are shown in Figure 8.

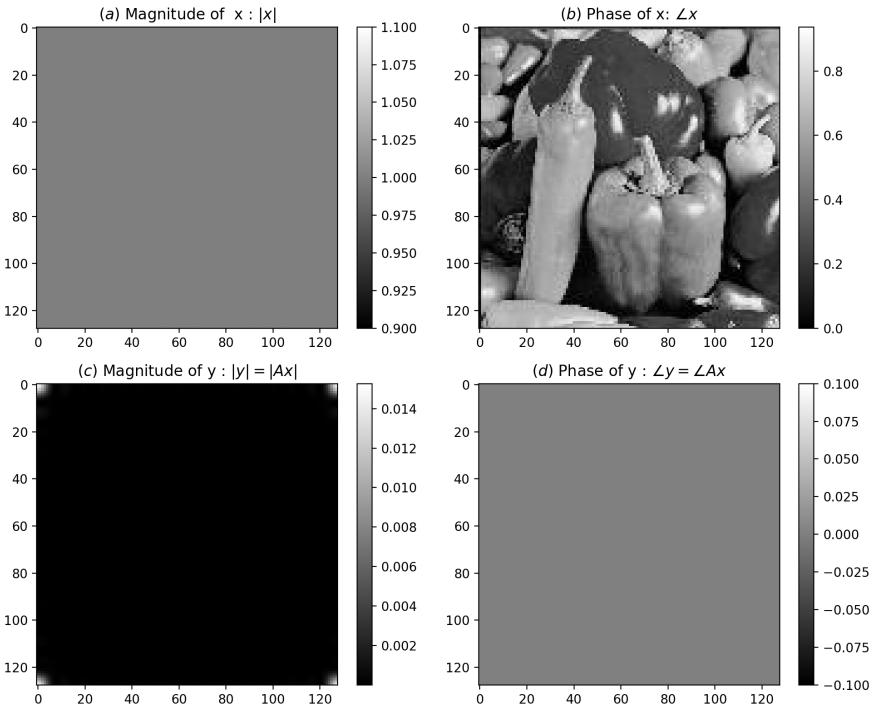


Figure 8: Magnitude (a) and phase (b) of the input (x) and the magnitude (c) and phase (d) of an example measurement ($y = Ax$).

Running the multiresolution algorithm produced good reconstruction results, as expected. These results are shown in Figure 9. The optimization successfully converged under various settings, with different numbers of iterations for each scale and varying learning rates. This is because the linear problem is convex, making the solution relatively straightforward to achieve, whether using a classical gradient descent approach or the multiresolution algorithm.

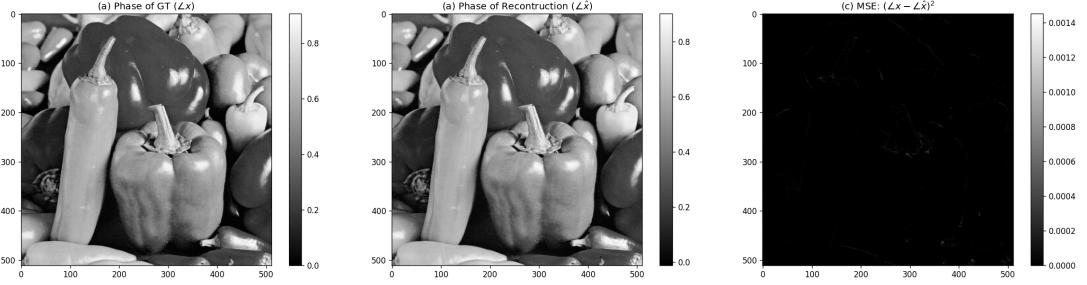


Figure 9: Result of the optimization: the ground truth image is embedded to the $\angle x$ shown in (a), phase of the reconstruction $\angle \hat{x}$ shown in (b), and finally the loss between the ground truth and reconstruction.

4.2 Nonlinear Problem Results with CPWL Functions

For the nonlinear phase retrieval problem described in Section 3.4, the multiresolution algorithm was used to address the optimization challenges. Results of the optimization with $\lambda = 0$ are shown in Figure 10.

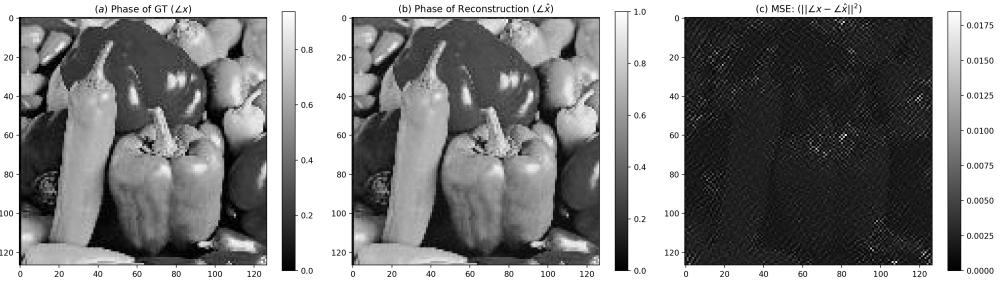


Figure 10: Result of the optimization: the ground truth image is embedded to the $\angle x$ shown in (a), phase of the reconstruction $\angle \hat{x}$ shown in (b), and finally the loss between the ground truth and reconstruction.

To accelerate the algorithm, an adaptive learning rate strategy was implemented, with two multipliers, α_u and α_d , used to dynamically adjust the learning rate based on the loss function. The evolution of the loss and learning rate over iterations is shown in Figure 11.

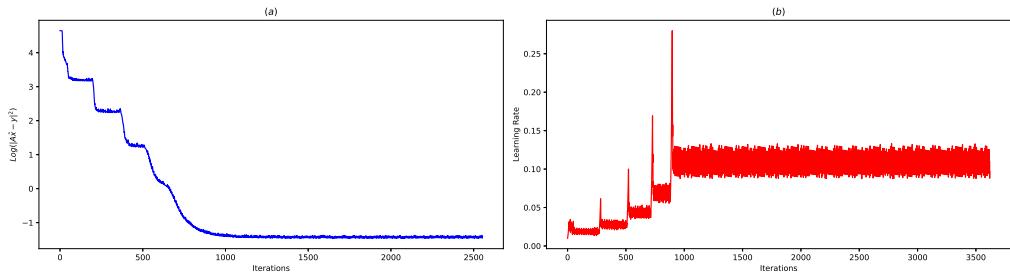


Figure 11: Log-Loss (a) and step size (b) through iterations.

The effect of HTV regularization was evaluated by varying the regularization parameter λ over $\{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$. The reconstructed images and their corresponding loss values are shown in Figure 12.

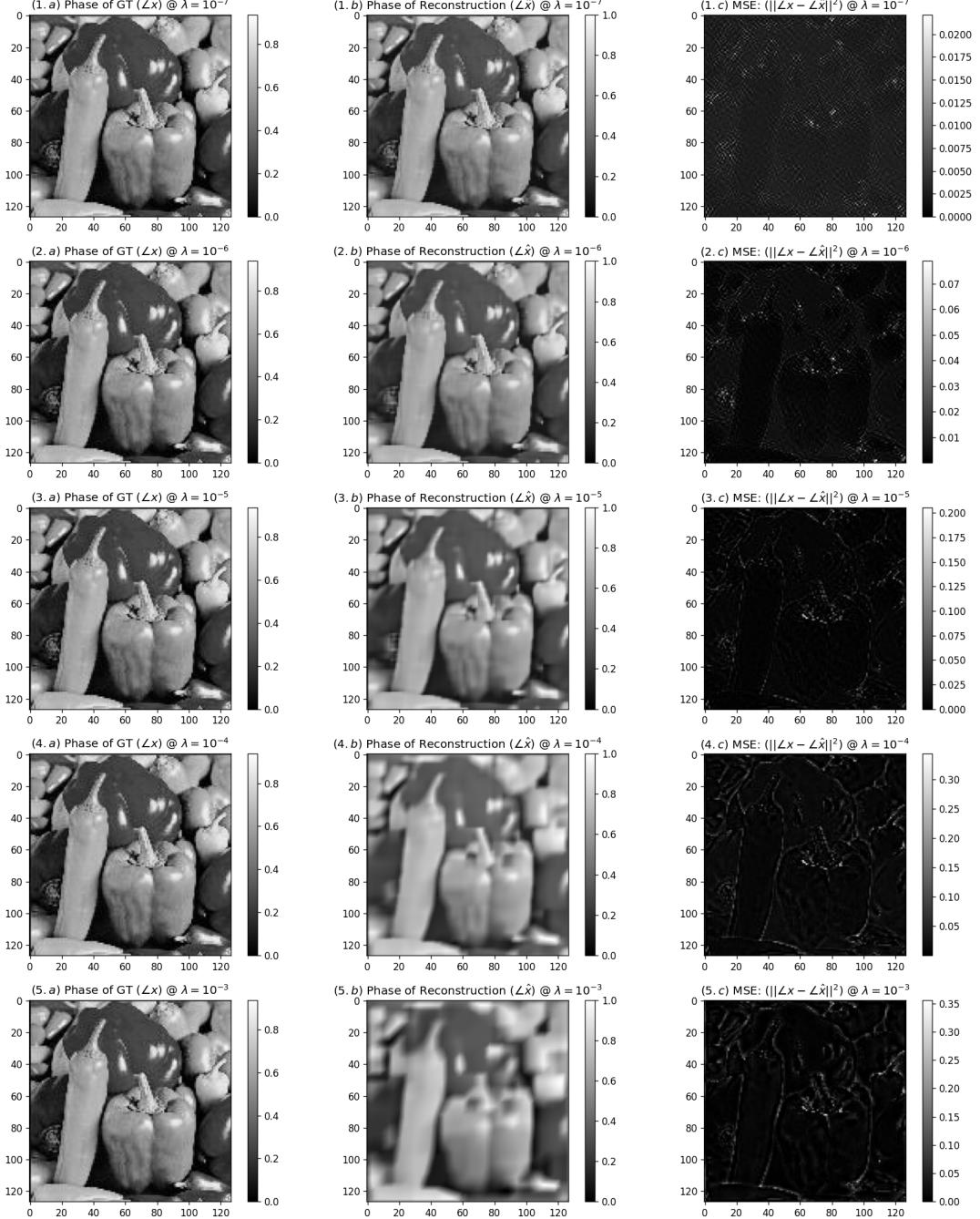


Figure 12: Ground truth (GT) images (a), reconstructed images (b), and corresponding loss results (c) for $\lambda = \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$.

The learning rate and loss curves for each value of λ are shown in Figures 13 and 14, respectively.

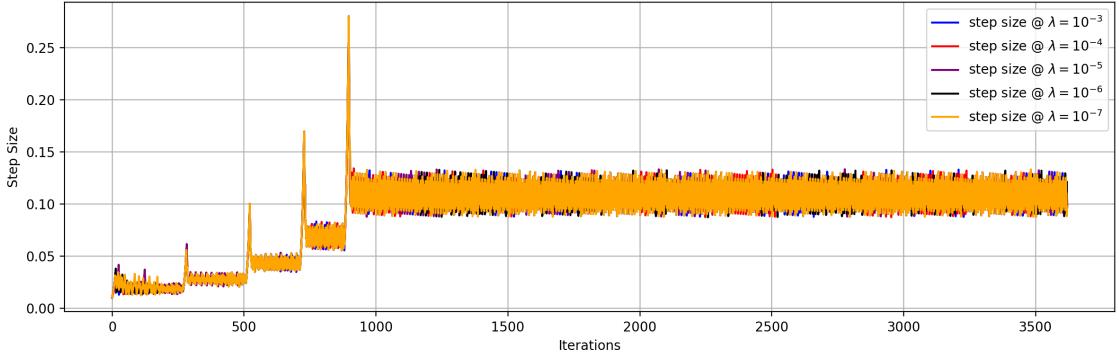


Figure 13: Step size values for different values of regularization parameter λ .

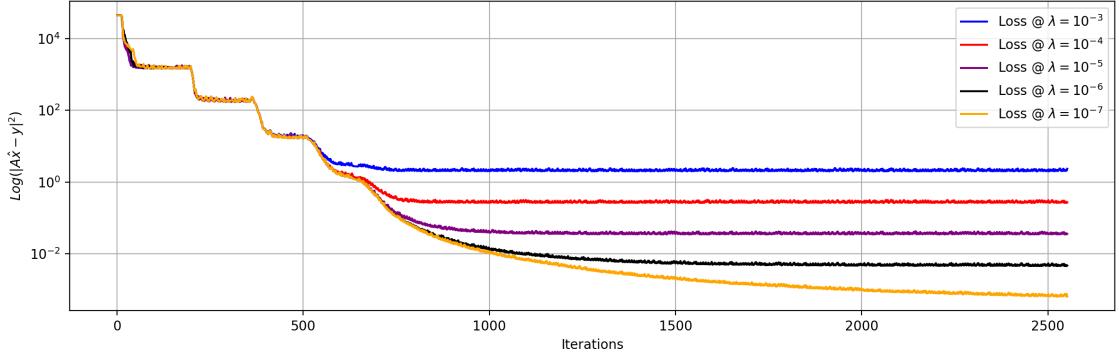


Figure 14: Log-Loss values for different values of regularization parameter λ .

The results demonstrate that while HTV regularization stabilizes the solution by enforcing CPWL behavior, excessive regularization introduces artifacts and reduces detail fidelity. These findings highlight the need for careful parameter selection to balance reconstruction accuracy and artifact suppression.

Comparing the results of the CPWL-based multiresolution algorithm with the classical gradient descent (GD) approach provides valuable insights. The GD method can be viewed as a special case of the multiresolution algorithm, where optimization is restricted to the finest scale. Figure 15 illustrates the comparison of the loss values between the two approaches over iterations.

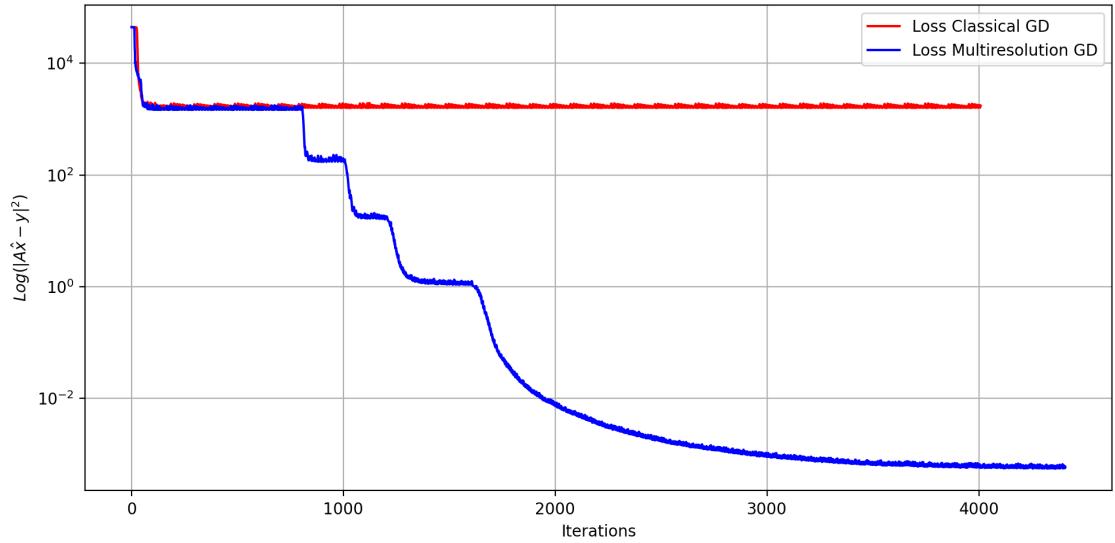


Figure 15: Log-Loss values of the optimization process over iterations for the CPWL-based multiresolution and gradient descent algorithms.

The results indicate that the CPWL-based multiresolution algorithm offers a more efficient optimization path, achieving faster convergence compared to the classical gradient descent approach.

4.3 Nonlinear Problem Results with CPC Functions

In the latter stages of the project, a bottleneck was identified in the implementation of the multiresolution algorithm. Until now, the code operated on all tensors and images at the finest discrete scale, resulting in suboptimal algorithmic performance. This inefficiency arises because the forward and backward operations do not necessarily require computations to be performed at the finest scale throughout.

To address the computational challenges, operations at coarser scales were implemented with reduced memory and tensor sizes, as described in the ptychography forward operation (Section 2.1) and the corresponding discretization (Section 2.2.2). This improvement significantly accelerates computationally expensive tasks, such as the FFT in the forward pass and the IFFT in transpose operations.

Initially, CPWL (Continuous Piecewise Linear) basis functions were used to enable HTV regularization, as the implementation of Pourya et al. [3]. CPWL box splines provide a robust framework for multiscale regularization. However, it was observed that the initial ptychography implementation operates on a pixel-based representation of images, which makes the multiscale relationships of CPWL functions difficult to implement efficiently within this framework. This limitation led to sub-optimal implementations and necessitated a reevaluation of the basis functions.

To simplify the implementation at this stage, CPWL basis functions were replaced with CPC (Continuous Piecewise Constant) functions, a different class of box splines. While regularization was temporarily removed to simplify the implementation, it is planned to reintroduce a suitable regularization approach in future

refinements. The derivation of the new implementation is provided in Section A, and the updated forward operator is expressed in Equation 24:

$$A_{s_\zeta}x_{s_\zeta} = 2^{-2\zeta}\text{DFT}_{2^\zeta}(x_{s_\zeta}) \odot \hat{\phi}(2^{-\zeta}), \quad (24)$$

where $A_{s_\zeta}x_{s_\zeta}$ represents the forward operation and input at scale s_ζ , and $\hat{\phi}(2^{-\zeta})$ incorporates the coarser scale transformation.

The method processes images at respective coarser scales, applies spatial copies, and uses fixed multipliers at each scale. This process is illustrated in Figure 16, which includes the input x , the ground truth (GT) image embedded into the phase, the result of $|Ax|$, the multiplier at a given scale, and the phase of $A^T Ax$.

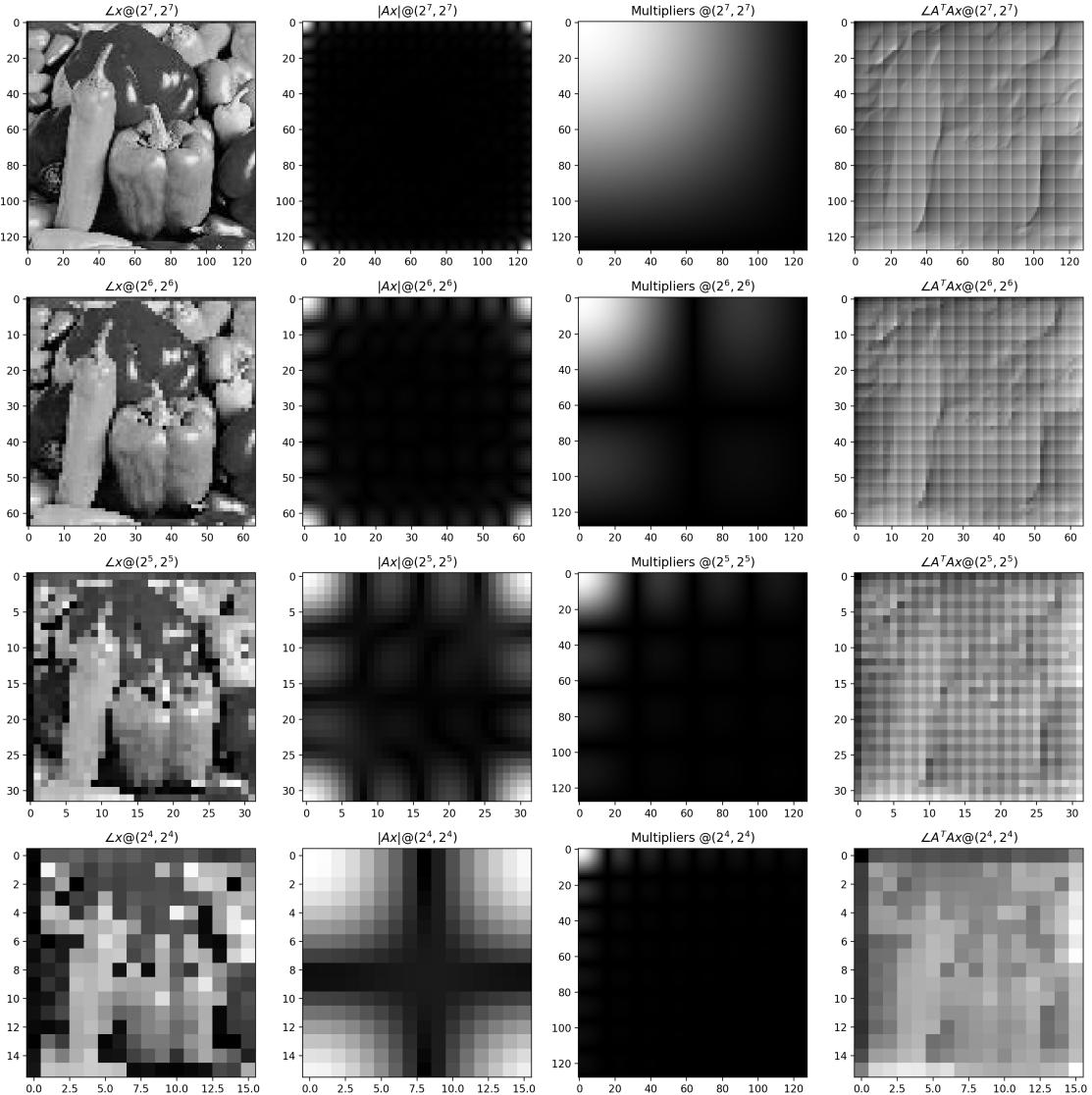


Figure 16: Illustration of the forward operation and its components for various scales.

Verification of the forward operation was conducted through inner product equal-

ity:

$$\langle Ax_1, y_2 \rangle = \langle x_1, A^H y_2 \rangle. \quad (25)$$

The optimization results using the proposed forward operator and scaling methods are presented in Figure 17.

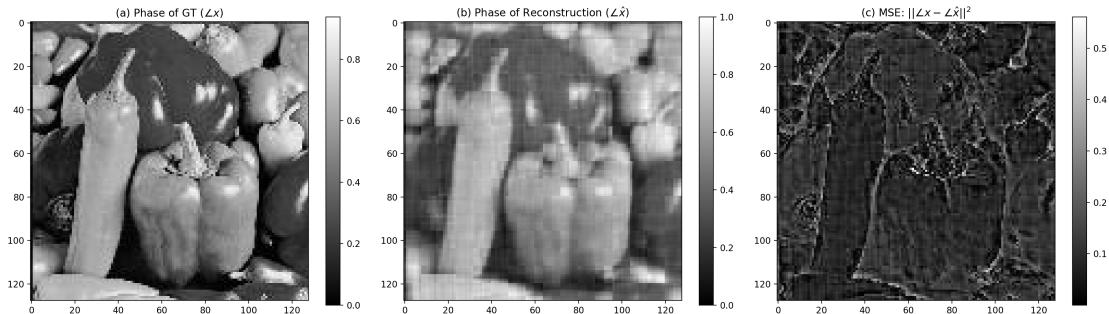


Figure 17: Illustration of the forward operation and its components for various scales with symmetric multipliers.

The reconstructed image demonstrates high quality, validating the effectiveness of the proposed approach as a reconstruction technique. Additionally, Figure 18 shows the loss values during optimization, revealing that more information was extracted at coarser scales compared to the initial algorithm, thus improving the optimization process.

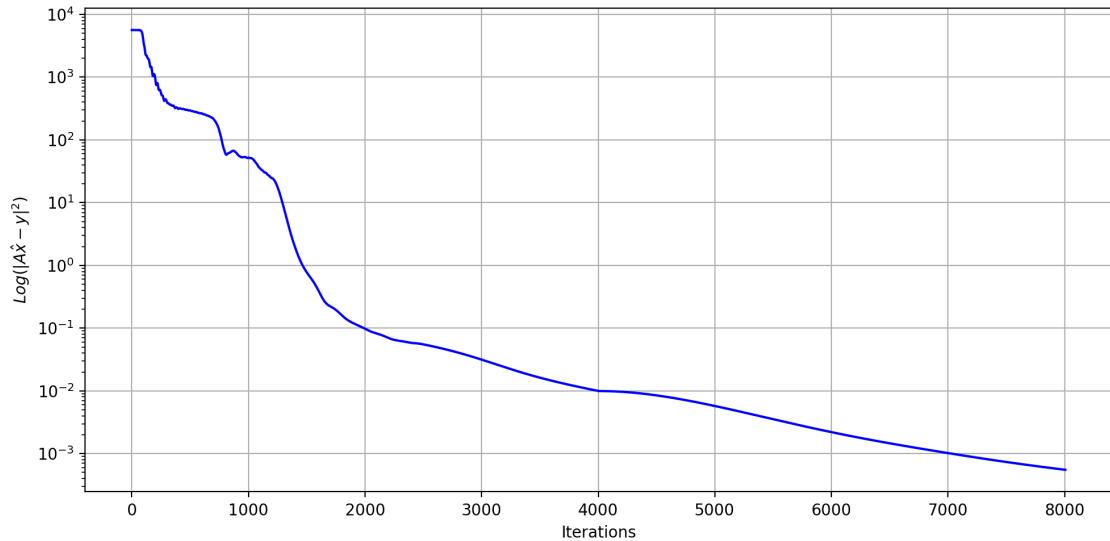


Figure 18: Log-Loss values of the optimization process over iterations.

It is also important to compare the results of utilizing the CPC-based multiresolution algorithm with the classical gradient descent (GD) approach, which can be considered a special case of the multiresolution algorithm where the optimization is performed only at the finest scale. The comparison of the losses between the two algorithms is shown in Figure 19.

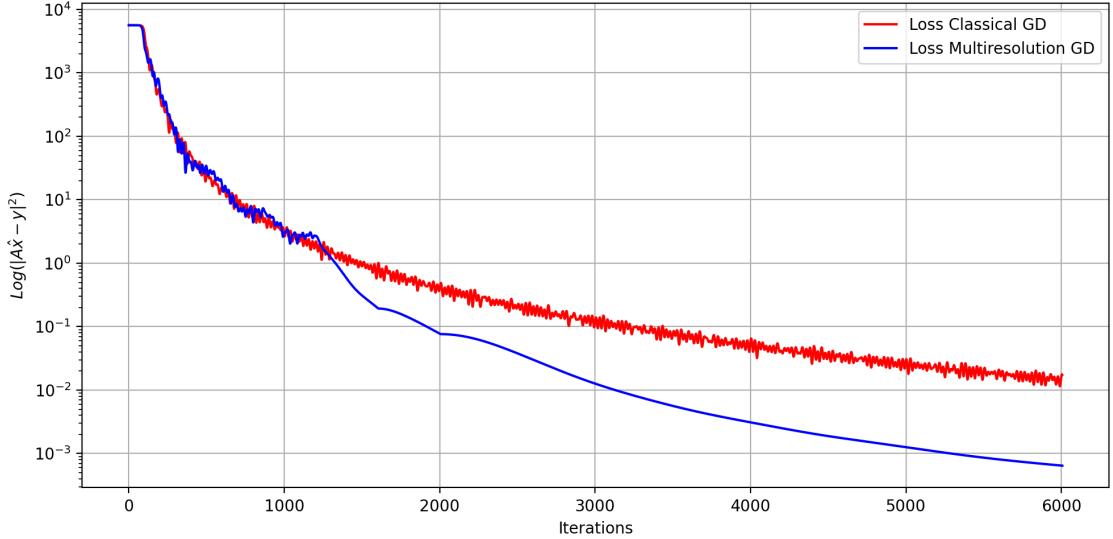


Figure 19: Log-Loss values of the optimization process over iterations for the two optimization algorithms.

The results demonstrate that the multiresolution algorithm achieves a better optimization path, with an increased convergence speed compared to the classical gradient descent approach.

5 Future Work

While the CPC-based method has shown promising results, certain challenges remain. By editing the multipliers such that the matrix becomes circularly symmetric, better definitions for the linear operator and its corresponding transpose can be achieved. This improvement addresses a potential issue arising from the implementation, which combines continuous-time Fourier transform (CTFT) and discrete-time Fourier transform (DTFT) operations. The results obtained using these symmetric multipliers are shown in Figures 20 and 21.

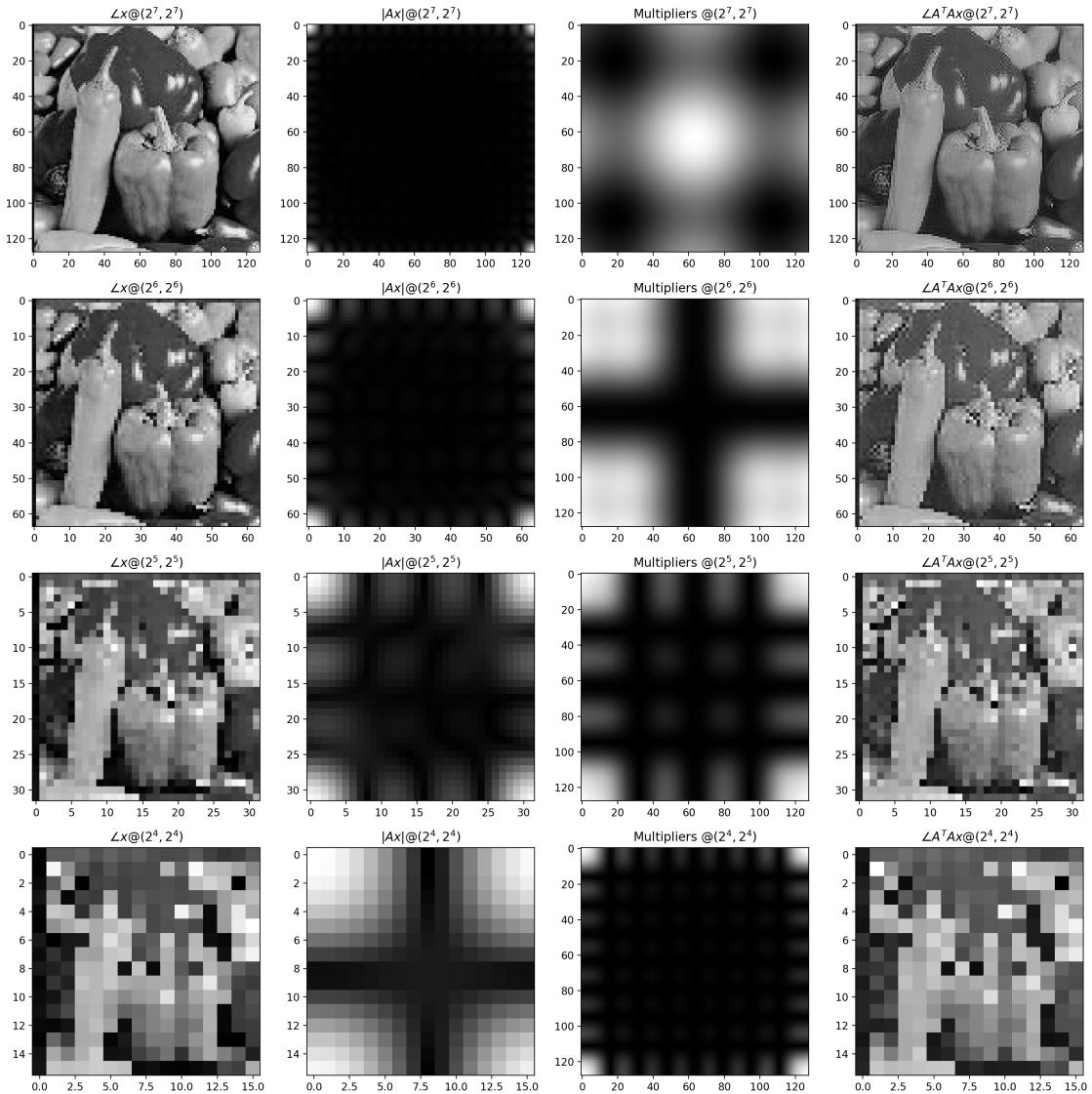


Figure 20: Log-Loss values of the optimization process over iterations using symmetric multipliers.

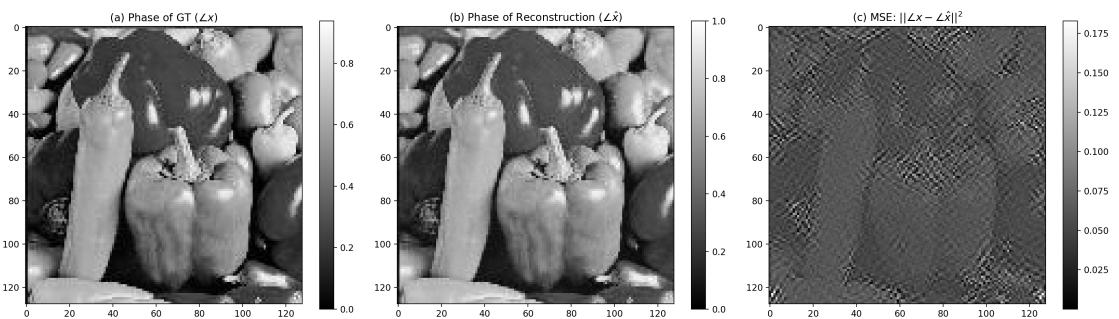


Figure 21: Reconstructed image using symmetric multipliers.

Details of the implementation of the multipliers will be thoroughly investigated as part of the future work of the project, and the mathematically correct implementation will be provided. Notably, the algorithm has yet to be fully optimized for time efficiency due to technical issues such as memory management and the selection of appropriate functions during coding. Future work could focus on developing more robust strategies for handling large-scale computations to further reduce runtime and memory footprint. Additionally, methods to integrate HTV regularization could be explored while maintaining the simplicity of CPC functions. Direct comparisons with the initial implementation are also needed to quantify improvements in time efficiency and reconstruction quality. These directions offer promising avenues to enhance the robustness, scalability, and efficiency of the proposed framework. Final remarks and conclusions for the project are provided in the following chapter.

6 Conclusion

In this study, we presented a novel multiresolution framework for ptychographic image reconstruction, leveraging the computational efficiency of box splines and the hierarchical structure of multiresolution methods. By addressing the inherent challenges of ptychography, including the non-convex nature of the optimization problem and the high dimensionality of the data, the proposed approach demonstrated significant promise in improving reconstruction accuracy and computational performance. The results achieved through the implementation of the multiresolution algorithm highlight its potential in addressing the phase retrieval problem effectively. By starting with coarse-scale representations and progressively incorporating finer details, the framework successfully balances computational efficiency with reconstruction fidelity. However, certain limitations remain. The current implementation has yet to fully optimize memory management and computational resource utilization, particularly at the finer scales. Looking ahead, the framework can be extended to incorporate additional regularization techniques tailored to CPC functions and explore advanced optimization strategies to further accelerate convergence. In conclusion, the proposed multiresolution approach provides a robust and efficient solution to the challenges of ptychographic reconstruction, with the potential for significant impact in the fields of biological imaging, materials science, and beyond. Continued advancements in this direction are expected to further broaden the applicability and effectiveness of computational imaging techniques.

References

- [1] J. Rodenburg, “Ptychography and related diffractive imaging methods,” ser. Advances in Imaging and Electron Physics, Hawkes, Ed. Elsevier, 2008, vol. 150, pp. 87–184. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1076567007000031>
- [2] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, “Experimental robustness of fourier ptychography phase retrieval algorithms,” *Optics Express*, vol. 23, no. 26, p. 33214, Dec. 2015. [Online]. Available: <http://dx.doi.org/10.1364/OE.23.033214>
- [3] M. Pourya, A. Boquet-Pujadas, and M. Unser, “A box-spline framework for inverse problems with continuous-domain sparsity constraints,” *IEEE Transactions on Computational Imaging*, vol. 10, pp. 790–805, 2024.
- [4] P. Bohra, T. an Pham, Y. Long, J. Yoo, and M. Unser, “Dynamic fourier ptychography with deep spatiotemporal priors,” *Inverse Problems*, vol. 39, no. 6, p. 064005, may 2023. [Online]. Available: <https://dx.doi.org/10.1088/1361-6420/acca72>
- [5] M. Pourya, A. Goujon, and M. Unser, “Delaunay-triangulation-based learning with hessian total-variation regularization,” *IEEE Open Journal of Signal Processing*, vol. 4, pp. 167–178, 2023.
- [6] J. Rodenburg and A. Maiden, “Ptychography,” in *Springer Handbook of Microscopy*, ser. Springer Handbooks, P. Hawkes and J. Spence, Eds. Springer, January 2019. [Online]. Available: <https://eprints.whiterose.ac.uk/127795/>

A Efficient Forward Implementation

A.1 Forward Operator

$$p(x) := \mathbf{1}_{[-0.5,0.5]^2}(x), \quad (\text{A.1.1})$$

$$\phi(x) := \mathbf{1}_{[0,1]^2}(x), \quad (\text{A.1.2})$$

$$\hat{\phi}(f) := \text{sinc}(f_1) \cdot \text{sinc}(f_2) \cdot e^{-j\pi(f_1+f_2)}, \quad (\text{A.1.3})$$

$$|FDx|^2 = \mathcal{F} \left(p(2w \cdot -k) \sum_{n \in [0, 2^\zeta - 1]^2} c_n \cdot \phi(2^\zeta \cdot -n) \right) \Big|_s \quad \text{where } s, k \in \mathbb{R}^2, \quad (\text{A.1.4})$$

$$\hat{c}_n = c_n \cdot p(2w \cdot -k) \quad \text{for } w > s, \quad (\text{A.1.5})$$

$$|FDx|^2 = \mathcal{F} \left(\sum_{n \in [0, 2^\zeta - 1]^2} \hat{c}_n \cdot \phi(2^\zeta \cdot -n) \right) \Big|_s = \sum_{n \in [0, 2^\zeta - 1]^2} \hat{c}_n \cdot \mathcal{F}(\phi(2^\zeta \cdot -n)) \Big|_s \quad (\text{A.1.6})$$

$$= \sum_{n \in [0, 2^\zeta - 1]^2} \hat{c}_n \cdot 2^{-2\zeta} \cdot \hat{\phi}(2^{-\zeta} s) \cdot e^{-j2\pi \langle n, s \rangle}. \quad (\text{A.1.7})$$

$$\sum_{n \in [0, 2^\zeta - 1]^2} \hat{c}_n \cdot e^{-j2\pi \langle n, s \rangle} \Big|_{s=m} = \text{DFT}(\hat{c}_n) \Big|_s \quad \text{for } m \in [0, 2^\zeta - 1]^2, \quad (\text{A.1.8})$$

$$\sum_{n \in [0, 2^\zeta - 1]^2} \hat{c}_n \cdot 2^{-2\zeta} \cdot \hat{\phi}(2^{-\zeta} s) \cdot e^{-j2\pi \langle n, s \rangle} \Big|_{s=m} = 2^{-2\zeta} \text{DFT}_{2^\zeta}(\hat{c}_n) \odot \hat{\phi}(2^{-\zeta}), \quad (\text{A.1.9})$$