Department of Precision and Microsystems Engineering

Automatic Differentiation based Multi-Mode Ptychography

Yabin Wang

Report no	: 2022.028
Coach	: Dr. Yifeng Shao (TNW)
Professor	: Prof. Dr. Wim Coene (TNW)
	Dr. Wouter Westerveld (3ME)
Specialisation	: Optics for Technology (OPT)
Type of report	: master thesis
Date	: August 12th, 2022

Automatic Differentiation based Multi-Mode Ptychography

A flexible and highly efficient lensless imaging algorithm

by

Yabin Wang

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Friday August 12, 2022 at 1:00 PM.

Student number:	5238897
Supervisors:	
Dr. Yifeng Shao (TNW)	Technische Universiteit Delft
Prof.dr. W.M.J.M. Coene	Technische Universiteit Delft & ASML Holding N.V.
Dr. Wouter Westerveld	Technische Universiteit Delft
Thesis committee:	
Prof.dr. W.M.J.M. Coene	Technische Universiteit Delft & ASML Holding N.V.
Dr. Gerard Verbiest	Technische Universiteit Delft
Dr. Wouter Westerveld	Technische Universiteit Delft
Dr. Jeroen Kalkman	Technische Universiteit Delft

TU Delft

Contents

Abstract

The scientific community recognizes the critical role played by ptychography in nanoscale imaging. Compared with the conventional imaging, which has high requirements on the manufacturing of optical elements, ptychography, as a computational imaging technique, uses a set of measured intensities of the diffraction patterns to reconstruct the image of the object and hence no imaging system is needed. This technique is especially useful in the short wavelength, e.g. EUV, regime, where manufacturing high quality optical elements such as mirrors is extremely expensive.

Most of the present ptychographic algorithms require the illumination of the object to be both spatially and temporally coherent so that the diffraction pattern can be interpreted as the intensity of the Fourier transform of the field exiting the object. However, the coherence of the sources that produce the EUV radiation often cannot be guaranteed. Therefore, it is crucial to extend the ptychography method to consider partial coherence effects. This requires the use of a flexible propagator which depends on the wavelength to deal with the temporal partial coherence and a modal representation for the spatially partially coherent field. Also, the ambiguity of the reconstructed modes of the probe will be solved by an orthogonalization approach, which could enhance the reproducibility of the results. These methods will be implemented on an existing ptychography platform based on automatic-differentiation and will be validated using both simulation data and experimental data.

Chapter 1

Introduction

As a promising high resolution imaging technique, diffractive imaging, also known as lensless imaging, is important for a wide range of scientific and industrial fields, such as the biomedical research and other fields that produce high resolution images of objects. Modern diffractive imaging techniques use diffraction patterns to reconstruct the illumination field (probe) and the sample transmission/reflection function of the object under study. Since these techniques rely on a mathematical model that is built with coherent light, diffractive imaging is also called coherent diffractive imaging (CDI) in the early research papers. A typical configuration of the CDI experimental set up is shown in Figure 1.1, in which the sample is illuminated by the probe generated by a source and the diffraction pattern is collected by the detector.

The initial but notable CDI algorithm is the *Gerchberg-Saxton algorithm* [1] (GS algorithm), which reconstructs the phase of a field by propagating it back-forth between two detection planes, one being at the image plane (real space) and the other being at the diffraction plane (Fourier space). The intensity distribution of this field must be known in both planes and the GS algorithm updates the object function iteratively by keeping the phase of the field while replacing the amplitude by the square root of the measured intensity distribution in each plane.

Later Fienup, et. al. modifies the GS algorithm such that the intensity measurement in the object plane is replaced by the support constraint (using the prior knowledge of the size of the object). The latter algorithm is referred to as the *error-reduction algorithm* (ER algorithm) [2] as it minimizes the error of the predicted and the measured diffraction pattern in the detection plane. To solve the convergence problem of the ER algorithm, Fienup, et. al. further modified the update scheme of the ER algorithm in reference [3] and the resulting algorithm is called the *Hybrid Input-output algorithm* (HIO algorithm).

Both ER and HIO algorithms require the intensity measurement only in the detection plane and only a support constraint in the object plane. This has significantly broadened the scope of the use of CDI. However, in most cases, obtaining the size of the object could already be very challenging. The reason for using the support constraint is that the some additional information is needed on top of the measured intensity of the diffraction pattern in order to reconstruct both the amplitude and the phase of the object.

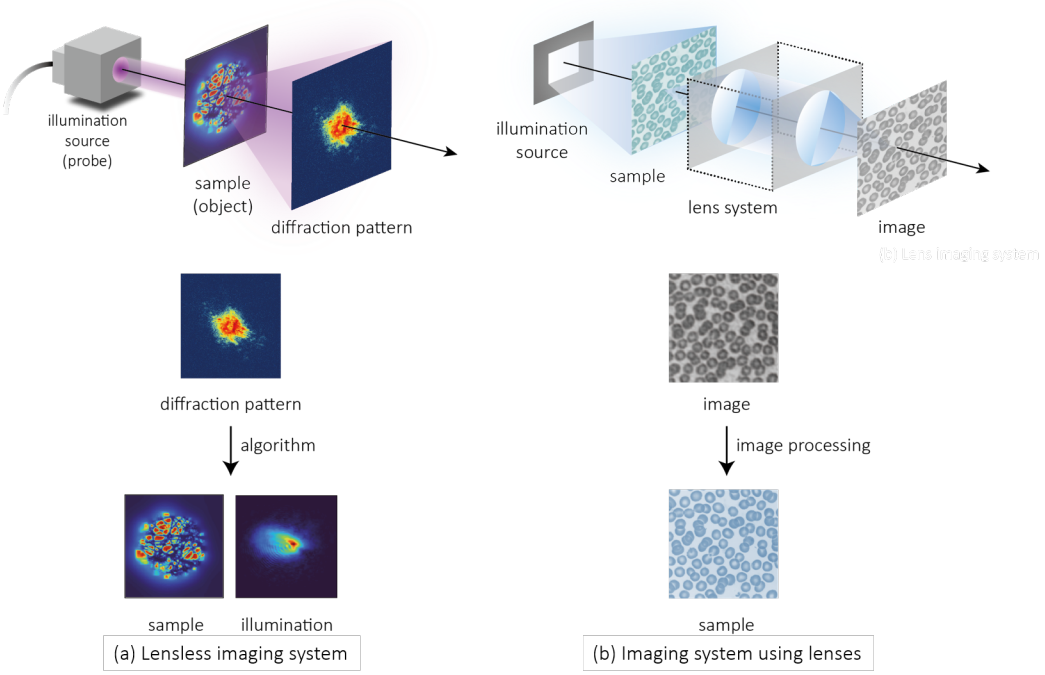


Figure 1.1: Schematic illustration of the lensless imaging system and a traditional imaging system using lenses.

The break-through of CDI took place in 2004 when Rodenburg, et. al. proposed the ptychography algorithm [4]. Compared to the above mentioned CDI methods, it still requires intensity measurement of the diffraction pattern, but the support constraint of the object is not needed any more. Instead, it needs the object to be illuminated by a probe (finite sized illumination field, e.g. the light through an aperture, or the probe as focused by illumination optics), and, most importantly, a scanning process. At each scanning position, a diffraction pattern is measured, and between adjacent scanning positions, there must be sufficiently overlap between the illuminated area of the object. The overlap provides sufficient redundancy in the measurements to overcome the lack of phase information. In the meantime, the scanning process significantly increases the field-of-view of the imaging over the original CDI approach, which uses only a single acquisition. Therefore, the ptychography algorithm is so far the most reliable and efficient CDI algorithm.

The history of ptychography dated back to 1969 when Hoppe, et. al. [5] proposed a method which is based on a series of diffraction patterns generated by a shifted probe across the sample at several scanning positions. In 1989, Rodenburg and Bates published a method based on ptychography [6] in which they consider the collected series of diffraction patterns as a four-dimensional data-set, dependent on the coordinates in the object plane, the grid of the scanning positions, and the coordinates of the spatial frequencies (related to the scattering angles), which is the grid of the coordinates in the detection plane. In this way, the collected series of diffraction patterns can be interpreted as the Wigner distribution of the field in the object plane and hence this field (the product of the object and the illumination field) can be reconstructed by deconvolution, that is,

there is no iterative update required. However, this method requires dramatically more scanning positions than conventional ptychography to achieve a high enough sampling density in the real space of the object plane.

The core to modern ptychography is an update scheme: *Ptychography iterative engine* (PIE), which was first proposed in 2004 [4] by Rodenburg, et. al. Since that pioneering paper, PIE has undergone several modifications and extensions, which enable it to reconstruct not only the object but also the probe (known as ePIE, which stands for extended PIE) [7], as well as to deal with more complicated situations in which partial coherence effects cannot be neglected.

The first of the latter kind of algorithms was introduced by Thibault and Menzel [8] to the community of ptychography in 2013. In their paper, eigen decomposition is applied to the mutual coherence function that describes a partially spatially coherent field. Each orthogonal eigen mode is referred to as a state of the system in analogy to the that in quantum physics. As a result, a fully coherent field is described as a pure state, while a partially coherent field needs to be described by a mixture of states. The most important contribution of their paper is that they demonstrate that ptychography provides enough redundancy of data (through enough overlap between the illuminated area on the sample between adjacent scanning positions) to reconstruct each eigen mode. As a result, the mixed state ptychography can deal with an extended source size for partial spatial coherence, or with vibrations during the experiment. Later on the state mixture has been demonstrated for ptychography on dynamic systems and broad-bandwidth sources [9][10][11]. The dynamic experimental system in Ref.[9] input an additional periodic signal (a sin wave and a square wave were used, respectively) to the sample positions. The dynamic of the sample is considerably shorter than the acquisition time of the experiment. It resulted in the piezo motors having a time-dependant motion, which caused periodic intraexposure sample vibration.

Inspired by the above paper, ptychography evolved into *ptychographical information multiplexing* (PIM) by Batey, et al [12], who extended ePIE of [7] to deal with plychromatic illumination. Meanwhile, multi modal ptychography starts to gain more attention [13][14].

Instead of extending the model of PIE reconstruction, Fineup et al interpreted it as a nonlinear optimization problem [15]. The *nonlinear optimization method* only requires the forward physical propagation model to establish the error metric, and the iterative updates can be performed based on a gradient descent method or other optimization algorithms. With the forward propagation model, this method can add more optimization variables besides object and probe, which can improve the performance against noise and initial errors. Nevertheless, the *nonlinear optimization method* requires the explicit expression of the gradient with respect to each variable. The tedious derivations of the gradients of the multiple gradients for optimization negatively impacts the flexibility in the design of the forward model in this approach..

Automatic differentiation (AD) provides an alternative for computing the gradients, which largely removes the above mentioned drawback. In 2014, Fienup et al. first suggested the AD framework for ptychography [16]. However, the lack of efficient computing packages limits its implementation. As a significant tool in deep learning, AD has undergone long-term development and evolution. The explosive appearance of several powerful packages, such as Tensorflow, Pytroch and Autograd, has made this framework come to fruition recently [17][18][19][20][21].

The increasing research interest in ptychography algorithm attributes not only to its extraordinary performance and flexibility, but also to the advances in the development of the light sources. As we

discussed previously, the visualization of small features down to nanoscale is crucial for industrial applications and academical research. The resolution of a traditional microscope, as shown in Figure 1.1, is restricted by Abbe's law to approximately half of the wavelength of the light source. Therefore, a straightforward method to improve the resolution is using radiation with shorter wavelengths. Some research used microscopy with the illumination of the extreme ultraviolet (10-124 nm wavelength) and soft X-ray range (1-10 nm) exhibits a nanoscale resolution. Traditionally, synchrotrons are regarded as the only source for extreme ultraviolet (EUV) and soft X-ray (SXR) imaging applications, which can produce adequate photon flux [22][23]. However, recent advances in the development of high-power ultrashort lasers drastically improved the photon-flux of laser-driven EUV sources using the process of high harmonic generation (HHG) in gases [24][25]. The ultrashort infrared laser pulses are focused into diluted gases, and a small amount of the infrared photons is coherently converted into higher harmonics of the fundamental frequency, reaching wavelengths down to a few nanometers. The emitted EUV or SXR radiation has laser-like properties such as high spatial coherence and low beam divergence. Recently, HHG, at least in the low energetic EUV range (about 30eV), reaches photon-fluxes that are comparable to large-scale facility synchrotrons [26]. Consequently, table-top laboratory-based EUV light source triggered the research activities in nanoscale imaging [27][28]. However, due to the strong absorption of the EUV light, refractive optics like lenses cannot be used. The unique advantage of ptychography that it does not require the lens system but uses algorithm to reconstruct the sample based on the diffracted field makes it well adapted to table-top EUV light source.

It should be noted that most of the existing ptychography algorithms with the AD framework require the illumination of the object to be both spatially and temporally coherent. However, in the short wavelength regime, (e.g., EUV), in which the ptychography is especially useful, the coherence of the light cannot be guaranteed. It is therefore of utmost importance to extend the ptychography algorithm to deal with partial coherence effects. To solve this problem, we have introduced a flexible propagator that depends on the wavelength to deal with the temporal partial coherence and a modal representation for the partially spatially coherent field. This new propagation method will be implemented into an AD-based ptychography algorithm, and validated using both simulation data and experimental data.

The remaining part of this thesis proceeds as follows: we first illustrate the idea of automatic differentiation framework and the structure of the algorithm (Chapter 2). We elaborate the model of multi-mode ptychography and demonstrate an orthogonal processing solution to enhance the reproducibility of the reconstructed probe modes (Chapter 3). We introduce a flexible propagation layer to deal with the initial error in propagation distance and wavelength (Chapter 4). Finally, we use the flexible propagation layer to extend the model to deal with ptychographyic data sets acquired under polychromatic illumination (Chapter 4). Finally, we propose several future research directions for ptychography with the AD framework (Chapter 5).

Chapter 2

Automatic differentiation

With the continuous development of the field of machine learning, automatic differentiation (AD) has received extensive attention as a method that can speed up gradient computation. As we discussed in the first chapter, the PIE reconstruction has been interpreted as a nonlinear optimization problem, where we can minimize the error metric through using some gradient-based optimization methods. In this chapter, we first give a limited overview of the automatic differentiation approach. Then we elaborate the framework of the automatic differentiation based ptychography algorithm.

2.1 Overview of automatic differentiation

The methods for the computing the derivatives in the computer programs can be classified into four categories: (1) manually deriving the explicit expression of the gradient and coding it; (2) *numerical differentiation*, using finite difference approximations generally; (3) *symbolic differentiation* using expression manipulation in computer algebra systems such as Mathematica, Maxima and Maple; and (4) *automatic differentiation* taking advantage of the chain rule.

The first method is the most intuitive, using basic derivative calculation rules to derive explicit expressions manually. However, for most engineering application and theoretical research problems, the objective function is usually complex and contains several different parameters. Furthermore, if we add new considerations (such as noise, new experimental variables) to the research, the derivation becomes a time-consuming and error-prone procedure.

The *numerical differentiation* can give an estimation of the derivative of a mathematical function based on the values of the function and other knowledge about the function. The simplest method is to use finite difference approximations. For example, for a function f with a single variant x, the gradient $\nabla f = \frac{df}{dx}$ can be approximated as the slope of a nearby secant line through the points $(x, f(x))$ and $(x+h, f(x+h))$:

$$\frac{df}{dx} \approx \frac{f(x+h) - f(x)}{h}, \tag{2.1}$$

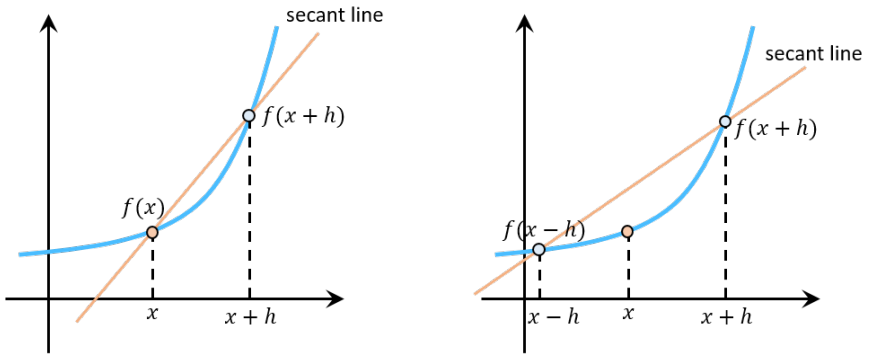
where $h > 0$ is a small step size. Different from the manual derivation, the numerical differentiation is not complicated. However, the slope of this secant line (approximate derivative value) differs

from the slope of the tangent line at the point $(x, f(x))$ (true derivative value) by an amount that is approximately proportional to the step size h. To obtain a good estimation of the gradient, the step size h needs to be selected carefully.

Various techniques have been developed to mitigate the approximation errors in numerical errors, such as using symmetric difference quotient:

$$\frac{df}{dx} = \frac{f(x+h) - f(x-h)}{2h} + O(h^2), \tag{2.2}$$

where the truncation error (error caused by approximating a mathematical process) moves to the second order, represented as $O(h^2)$. However, the approximation error is still a problem for this method, especially when the gradients with respect to a lot of parameters are needed in some complex model.



(a) first-order divided difference (b) symmetric difference quotient

Figure 2.1: Two examples of the methods for the numerical differentiation. (a) Newton's difference quotient (also known as a first-order divided difference), expressed as Equation 2.1, computes the slope of a nearby secant line through the points $(x, f(x))$ and $(x+h, f(x+h))$ to estimate the derivative. (b) Symmetric difference quotient computes the the slope of a nearby secant line through the points $(x-h, f(x-h))$ and $(x+h, f(x+h))$. Both of the numerical differentiation methods require a suitable step size h, and return to an approximate value of the derivative.

The third method, *symbolic differentiation*, resolves the disadvantages of the first two methods, but often results in complex and enigmatic expressions afflicted with the problem of "expression swell", which describes a phenomenon of exact computations in which the size of numbers and expressions involved in a calculation grows dramatically as the calculation progresses. Furthermore, both manual derivation and symbolic differentiation require a closed-form expression to describe the model.

The *automatic differentiation* is based on the chain rule for the gradient computation. A function, in general, can be interrupted as an amalgamation of a finite number of equations and some arithmetic operations. Using the chain rule, the differentiation of the functions can be described as a recursive procedure, where the same elementary function is differentiated repeatedly, only

with different input parameters. In the following explanation, we only focus on the *reverse-mode automatic differentiation* framework, which will be used in the ptychography algorithm.

To elucidate the concept of automatic differentiation, we first consider a differentiable function $f(x) = \phi_3(\phi_2(\phi_1(x)))$, and the elementary functions ϕ_1, ϕ_2 and ϕ_3 are also differentiable.

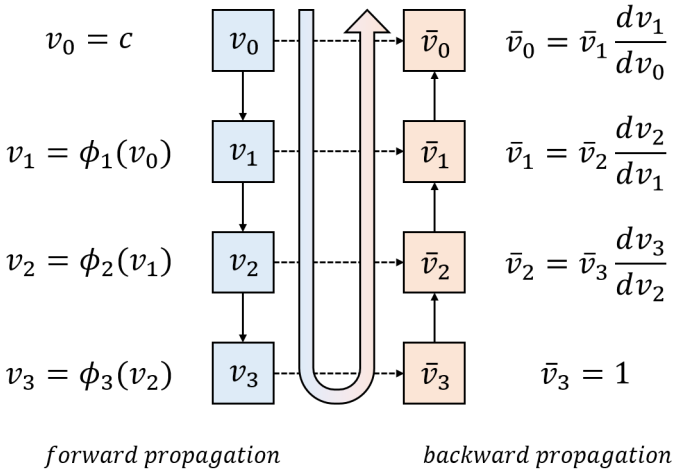


Figure 2.2: Schematic illustration of the automatic differentiation.

As shown in Figure 2.2, to calculate the derivative of $f(x)$ at $x = c$, according to the chain rule, we can compute the derivatives of the elements:

$$\frac{df}{dx}\Big|_{x=c} = \frac{d\phi_3}{dx}\Big|_{x=\phi_2(\phi_1(c))} \frac{d\phi_2}{dx}\Big|_{x=\phi_1(c)} \frac{d\phi_1}{dx}\Big|_{x=c}. \tag{2.3}$$

To evaluate the derivative, we need to perform a backward propagation to use the intermediate values obtained in the forward propagation. The backward propagation can be denoted as conduction of several derivatives:

$$\overline{v}_i = \overline{v}_{i+1} \frac{dv_{i+1}}{dv_i}. \tag{2.4}$$

From Equation 2.4, we can notice that each step in the backward propagation requires the not only the derivative values calculated in the previous step, but also the intermediate values calculated in the forward propagation. Therefore, to obtain the final value of the derivative, we need to:

1. identify the forward propagation model with several elements ϕ_1, ϕ_2 and ϕ_3,
2. perform the forward propagation and save the intermediate values,
3. perform the backward propagation to compute the final derivative.

The reverse-mode automatic differentiation procedure can be summarized in Figure 2.2. Without relying on the closed form of the model or tedious derivation, this method can still acquire the

exact value of the derivative. Thanks to the fast development of the deep learning packages, such as Tensorflow, Pytorch and Autograd, the AD framework can be implemented efficiently.

To elaborate the difference between these methods for the computation of derivatives in the computer programs, we establish a function as an example and use different methods to estimate its derivative. As shown in Figure 2.3, the elementary functions we use here are:

$$l_1 = x \tag{2.5}$$

$$l_{n+1} = 4l_n(1 - l_n) \tag{2.6}$$

where x represents the variable, l represents the elementary function and n represents the index of the elementary function. Taking $n = 4$ as the example, the established function can be expressed as:

$$f(x) = l_4 = 64x(1-x)(1-2x)^2(1-8x+8x^2)^2 \tag{2.7}$$

If we derive the explicit expression of the derivative manually, we can use the differentiation rules step by step:

$$\begin{aligned} f'(x) =& 128x(1-x)(-8+16x)(1-2x)^2(1-8x+8x^2) \\ &+ 64(1-x)(1-2x)^2(1-8x+8x^2) \\ &- 64x(1-2x)^2(1-8x+8x^2) \\ &- 256x(1-x)(1-2x)(1-8x+8x^2). \end{aligned} \tag{2.8}$$

If we manipulate the expression of the function $f(x)$ in some computer algebra systems, we can reach the same symbolic result as the manual derivation. Both of these two methods obtain the exact value of the derivative.

If we use the numerical differentiation, we are supposed to define a suitable step size h and compute the slope of the secant line through the two points $(x, f(x))$ and $(x+h, f(x+h))$. And the result will be an approximate value.

If we implement the reverse-mode automatic differentiation framework, we should define the derivative function for each elementary function first:

$$\frac{dl_1}{dx} = 1, \tag{2.9}$$

$$\frac{dl_{n+1}}{dl_n} = 4(1 - l_n) - 4l_n. \tag{2.10}$$

With these elementary functions, we can perform the forward propagation, save the intermediate values and perform the backward propagation to obtain the exact derivative value. Please notice that the derivative functions for elementary functions are already saved in the packages, which do not need to be calculated by ourselves. Here these functions are only used to illustrate the procedure.

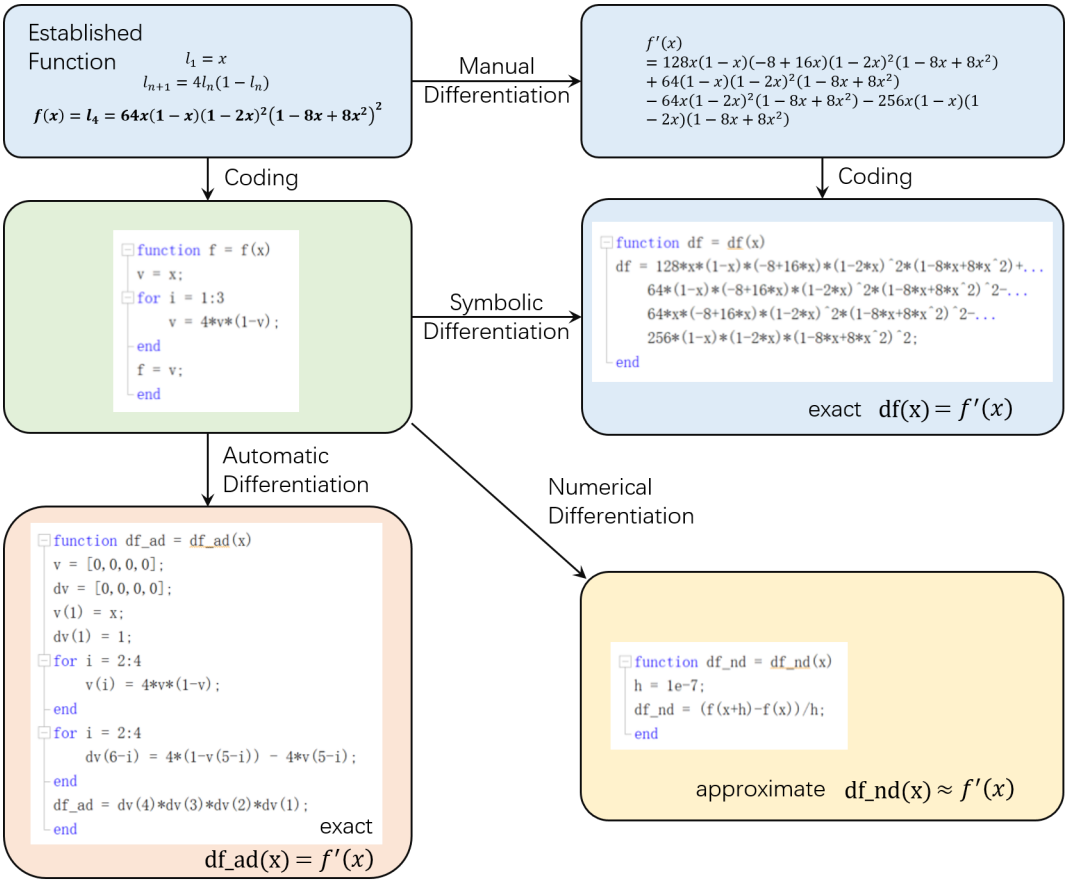


Figure 2.3: Schematic illustration of different methods of computing the derivatives. Symbolic differentiation (center right) provides exact results but requires closed-form model and suffers from expression swell; numerical differentiation (lower right) requires a suitable step size and gives an approximate result; automatic differentiation can provide the exact result and circumvent the tedious derivation.

As such, the reverse-mode AD procedure is suitable to solve optimization problems, providing gradient-based iterations efficiently. In the neural network, each training step involves a large number of operations and millions of parameters, so the closed form of the gradients is impossible to be obtained manually, which highlights the necessity of the AD framework. Compared with neural network, a phase retrieval problem is much simpler and the closed form of the error metric can also be established with several operations and functions, which has been proved by several ptychographic research. However, the flexibility of the AD framework is still impressive to reduce the laborious work of derivation. Furthermore, it also allows us to modify the forward model, which empowers us to deal with more complicated problems and to consider more parameters in the phase retrieval problem. Consequently, we expect that AD framework can yield state-of-the-art results in ptychography algorithm, which is also the quintessence of this work.

2.2 Structure of the AD based ptychography algorithm

As shown in Figure 2.6, to implement the AD framework in the ptychography algorithm, we use TensorFlow as the basic package to build the project. Also, the optimization algorithms in the TensorFlow package, such as Adam optimizer, can be used as the foundation of the optimization library in this project. We use the Numpy package to build another basic library, which can be used to do some simulations and pre/post processing for the results. Except these two basic packages, we also introduce other packages to visualize the optimization procedure, read and write data files and implement other functions.

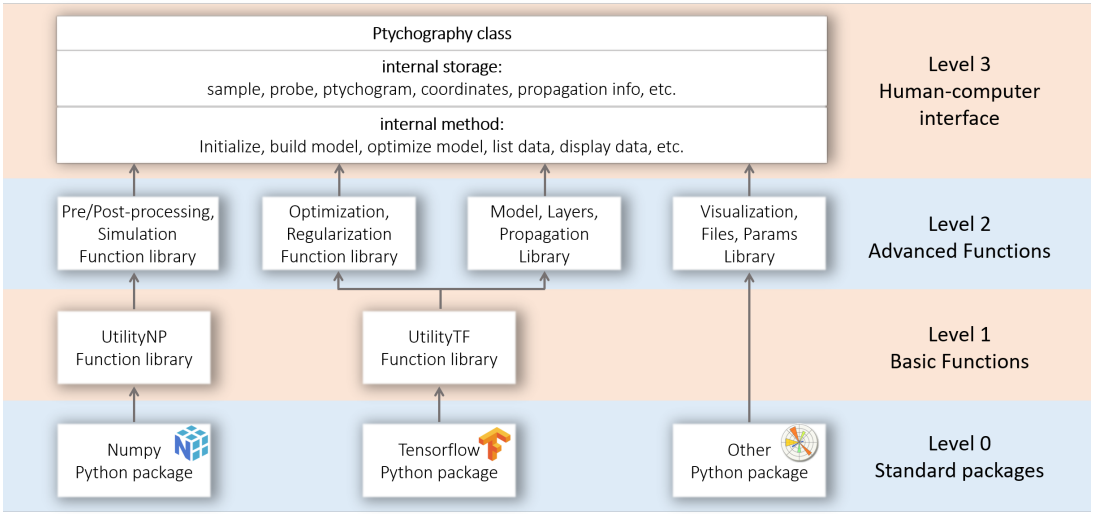


Figure 2.4: Schematic illustration of the ptychography platform.

As we explained in the previous section, the first step to implement the AD framework is to establish the forward model and perform the forward propagation. The forward model is based on the experimental set up. The configuration of the ptychography experiment can be classified into two categories: transmission ptychography and reflection ptychography, as shown in Figure 2.5. For both of these configurations, the sample is placed on a translation stage and illuminated with a light beam, called the probe, which focuses on a small region of the sample. The transmission ptychography usually utilizes a transparent or semitransparent sample. The exit field will pass through the sample and propagate to the in-line detector. However, for the highly absorbable radiation, like EUV, refractive optics like lenses can not be used. The reflection ptychography utilizes a reflective sample and mirrors in the system to avoid the strong absorption. The exit wave will be reflected to the detector.

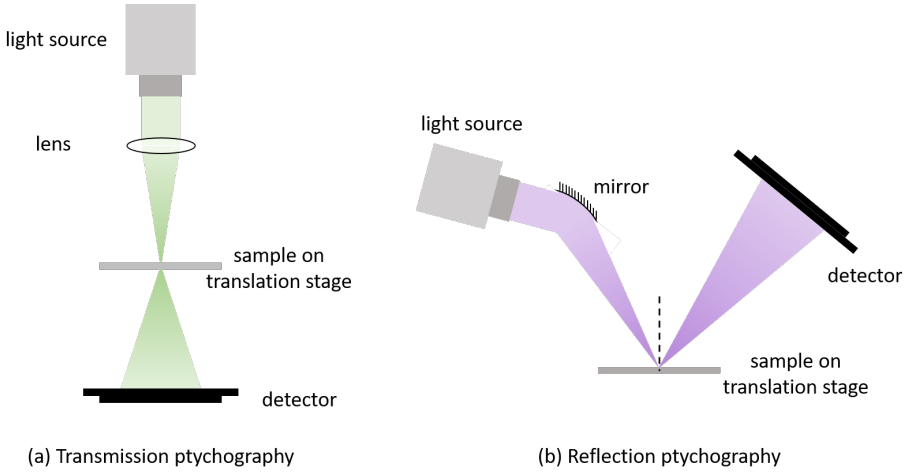


Figure 2.5: Configurations of transmission ptychography and reflection ptychography.

Here we take the experimental set up of the reflection ptychography, as shown in Figure 2.6, to elaborate the relation between the physical forward model and the code structure. In the code, we divide the forward model into several layers. Following the imaging procedure, each layer defines some elements of the experimental set up and delivers these parameters to the next layer. The *select layer* converts the recorded physical coordinates of the translation stage encoder into discrete pixel units and delivers these scanning positions to the next layer. The *illuminate layer* defines the sample function and the probe function. Generally, as the starting point for the reconstruction, the sample is initialized with uniform amplitude and zero phase; and the probe is initialized as a circular aperture with size of half of the field of view. Following the requirements of the TensorFlow package, the sample function and the probe function are programmed as some trainable tensors in the code. The exit wave is the product of the probe function and the sample function at some scanning position, which will be delivered to the *propagation layer*. The propagation layer simulates the propagated field with different propagators to deal with different sample-detector distances. After receiving the propagated field, the *camera layer* will define the mask and the bit depth of the camera based on the experimental set up to generate the diffracted field. The intensity of the simulated diffracted field will be compared with the measurement to calculate the loss function. After the forward propagation, TensorFlow can help us to compute the gradients. The internal optimization package can use these gradients to update the sample and the probe. To extend the ptychography algorithm to deal with different situations, establishing a suitable physical forward model is crucial, which is also the essence of the following chapters.

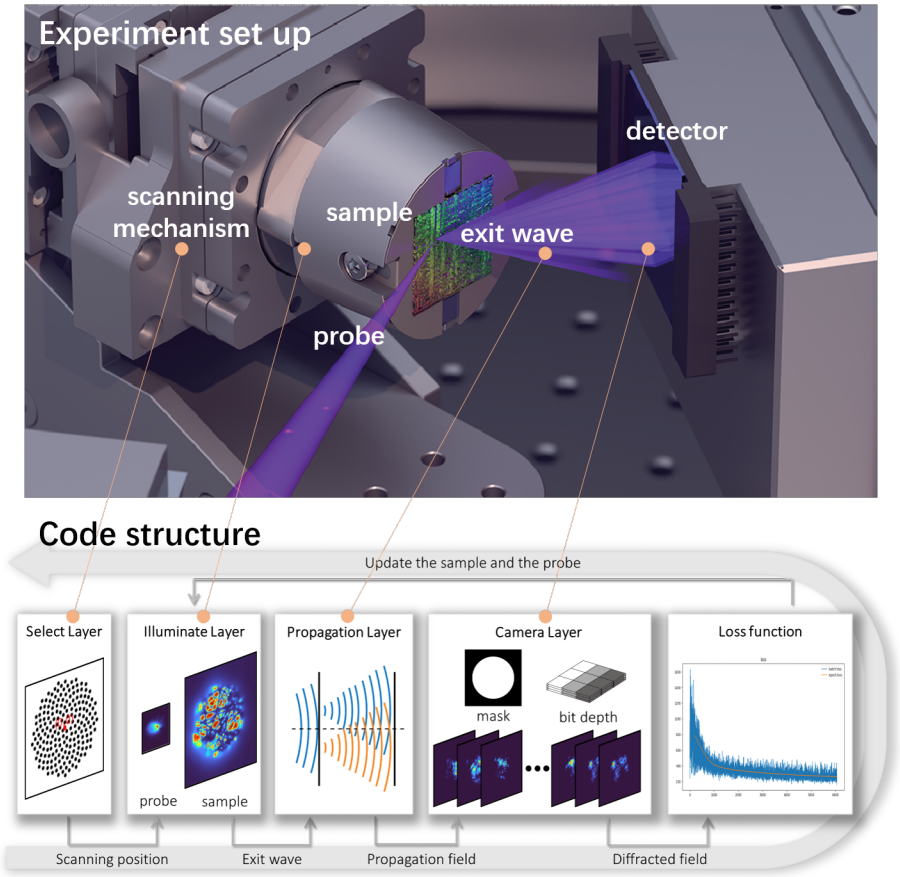


Figure 2.6: Schematic illustration of automatic differentiation ptychography experimental setup and algorithm: the select layer can match the index of the scanning positions with the coordinates and produce the scanning positions array; the illuminate layer can initialize and update the probe function and the sample function based on the loss function; the propagation layer can calculate the propagation field with different methods; the camera layer can deal with the parameters of the camera (covered with mask to resist noise or the bit depth) and produce the diffraction field to update the loss function and start the next iteration.

Chapter 3

Multi-mode ptychography

Recent developments in ptychographic imaging have significantly increased the need for multi-mode ptychography algorithm, which has a better performance when the illuminating source is partially coherent or when the object being detected is laterally vibrating [9][11][29]. The essence of the multi-mode ptychography is that the sample and the probe are now described by the mutual coherence function which is then decomposed into coherent modes. However, since the ptychography with the automatic differentiation (AD) framework cannot ensure the modes to be orthogonal during the reconstruction, the uniqueness of the result cannot be guaranteed. In this chapter, we first build the physical forward model for multi-mode ptychography. Then we discuss the origin of the ambiguity and introduce a method to make the modes mutually orthogonal. Finally, we test the performance of the introduced method on a data set of a ptychographic experiment with partially coherent illumination for its spatial coherence.

3.1 Single-mode ptychography

As mentioned in the first chapter, CDI reconstructs the object field from a multitude of diffraction patterns. To obtain the diffraction patterns, the detector is usually located at some distance from the object. Figure 3.1 demonstrates the basic model of a ptychographic setup. The distance is often long enough in order to satisfy the conditions for Fraunhofer diffraction, so that the field at the object plane and the intensity at the detector plane can be connected with the 2D Fourier transform:

$$I(u,v) = |\mathcal{F}[\Phi(x,y)]|^2 = |\iint \Phi(x,y)e^{-2\pi i(ux+vy)}dxdy|^2, \tag{3.1}$$

where Φ represents the exit wave at the object plane, (x,y) and (u,v) represent the two-dimensional coordinates in the object plane (the real space) and in the detector plane (the reciprocal space with respect to the real space of the object plane), respectively, and $\mathcal{F}$ represents the Fourier transform. The relation between these coordinates is:

$$u = \frac{x}{\lambda z}, \quad v = \frac{y}{\lambda z}, \tag{3.2}$$

where z represents the distance between these planes and λ represents the wavelength of the light source, which is assumed to be monochromatic.

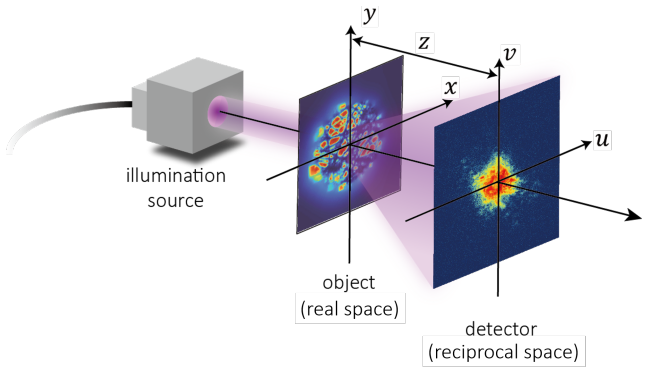


Figure 3.1: Schematic illustration of the ptychography setup in a transmission configuration. The real space (x,y) and the reciprocal space (u,v) can be connected with different propagators, depending on the propagation distance z.

In ptychography, the shift of the probe is usually realized by a piezo stage. The exit wave is the product of the object function, $o(x,y)$, and the shifted illumination (probe) function, $p(x+x_j, y+y_j)$, where the j denotes the index for the scanning position given by (x_j, y_j):

$$I_j(u,v) = |\iint o(x,y)p(x+x_j,y+y_j)e^{-2\pi i(ux+vy)}dxdy|^2 = |\mathcal{F}[o(x,y)p(x+x_j,y+y_j)]|^2. \tag{3.3}$$

There are several alternative propagators to simulate the diffracted field depending on the setup [30]. Then the general version of Equation 3.3 is:

$$I_j(u,v) = |\mathcal{D}_{\lambda,z}[o(x,y)p(x+x_j,y+y_j)]|^2, \tag{3.4}$$

where $\mathcal{D}_{\lambda,z}$ represents the operator for the free-space light propagation at wavelength λ by distance z. Figure 3.2 summarises this basic model as a flow chart.

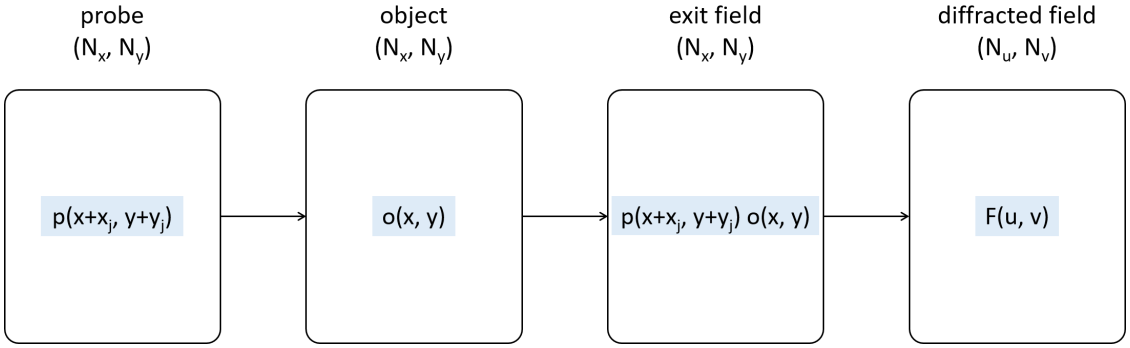


Figure 3.2: Schematic illustration of the forward model of the single modal ptychography. The exit wave can be defined as the product of the object function and the probe function. Taking the far field as an example, the diffracted field is simulated as the Fourier transform of the exit field.

3.2 Multi-mode ptychography

To introduce the multi-mode ptychography, we first analyze the imaging model of the ptychographic set up, where the exit wave in the object plane is transformed into a diffracted field in the detector plane. According to the Huygen's law, each point of the object field is itself a source. The total diffracted field is the sum of the image fields emitted by all point sources, which can be written as:

$$E_d(\boldsymbol{r_d}) = \iint H(\boldsymbol{r_d};\boldsymbol{r_o})E_o(\boldsymbol{r_o})d\boldsymbol{r_o} = \iint H(r_d;\boldsymbol{r_o})o(\boldsymbol{r_o})p(\boldsymbol{r_o}+\boldsymbol{r_j})dr_o, \tag{3.5}$$

where $\boldsymbol{r_o}$ and $\boldsymbol{r_d}$ represent the 2D coordinates in the object plane and the detector plane, respectively, $E_d(\boldsymbol{r_d})$ represents the diffracted field in the detector plane, $E_o(\boldsymbol{r_o})$ represents the exit wave in the object plane, which can be the product of the probe function and the object function on the object plane; and $H(\boldsymbol{r_d};\boldsymbol{r_o})$ represents the image field generated by the point source at $\boldsymbol{r_o}$ in the object plane. $H(\boldsymbol{r_d};\boldsymbol{r_o})$ is also known as the point-spread function (PSF) of the imaging system.

Because only the intensity of the diffracted field can be measured by the detector, the data we obtain from the detector is referred to as the intensity of the total diffracted field. The correlation between the diffracted field generated by each point in the object plane is significant in image formation. The intensity of the diffracted field $I(r_d)$can be expressed as:

$$\begin{aligned} I(\boldsymbol{r_d}) &= \langle E_d(\boldsymbol{r_d})E_d(\boldsymbol{r_d})^*\rangle \\ &= \iint\iint H(r_d;\boldsymbol{r_{o1}})H(r_d;\boldsymbol{r_{o2}})^*\langle o(\boldsymbol{r_{o1}})o(\boldsymbol{r_{o2}})^*\rangle\langle p(\boldsymbol{r_{o1}})p(\boldsymbol{r_{o2}})^*\rangle d\boldsymbol{r_{o1}}d\boldsymbol{r_{o2}}, \end{aligned} \tag{3.6}$$

where $\langle\cdot\rangle$ represents the ensemble averaging. In Equation 3.6, we define the mutual coherence function (MCF) of the probe function and that of the object function as:

$$J_p(\boldsymbol{r_{o1}},\boldsymbol{r_{o2}}) =< p(\boldsymbol{r_{o1}})p(\boldsymbol{r_{o2}})^* >, \tag{3.7}$$

and

$$J_o(\boldsymbol{r_{o1}},\boldsymbol{r_{o2}}) =< o(\boldsymbol{r_{o1}})o(\boldsymbol{r_{o2}})^* >, \tag{3.8}$$

respectively, which describe the correlation between probe functions and object functions at $\boldsymbol{r_{o1}}$ and $\boldsymbol{r_{o2}}$, respectively.

Taking the MCF of the probe function as an example, if the illumination source is spatially coherent, the MCF of the probe function is calculated by:

$$J_p(\boldsymbol{r_{o1}}, \boldsymbol{r_{o2}}) = p(\boldsymbol{r_{o1}})p(\boldsymbol{r_{o2}})^*, \tag{3.9}$$

which indicates that $p(\boldsymbol{r_{o1}})$ and $p(\boldsymbol{r_{o2}})$ are correlated for all possible combination of $\boldsymbol{r_{o1}}$ and $\boldsymbol{r_{o2}}$.

If the illumination source is spatially incoherent, the MCF of the probe function is calculated by:

$$J_p(\boldsymbol{r_{o1}}, \boldsymbol{r_{o2}}) = p(\boldsymbol{r_{o1}})p(\boldsymbol{r_{o2}})^* \delta(\boldsymbol{r_{o1}} - \boldsymbol{r_{o2}}), \tag{3.10}$$

which indicates that $p(\boldsymbol{r_{o1}})$ and $p(\boldsymbol{r_{o2}})$ are correlated only when $p(\boldsymbol{r_{o1}}) = p(\boldsymbol{r_{o2}})$ and are uncorrelated when $p(\boldsymbol{r_{o1}}) \neq p(\boldsymbol{r_{o2}})$.

In other words, the expression of the MCF of the probe function depends on the degree of spatial coherence of the illumination source. Except the spatially coherent situation and the spatially incoherent situation, we can also reach an intermediate state when the illumination source is partially spatial coherent.

Similarly, if the sample is static, the MCF of the object function can be expressed as:

$$J_o(\boldsymbol{r_{o1}}, \boldsymbol{r_{o2}}) = o(\boldsymbol{r_{o1}})o(\boldsymbol{r_{o2}})^*, \tag{3.11}$$

which indicates that $o(\boldsymbol{r_{o1}})$ and $o(\boldsymbol{r_{o2}})$ are correlated for all possible combination of $\boldsymbol{r_{o1}}$ and $\boldsymbol{r_{o2}}$. To reach and calibrate an incoherent object field, the authors of Ref.[31] use a spatial light modulator to alternately display two phase patterns at a rate of 20Hz. We can also define an intermediate state of the object function caused by the vibrations of the sample.

Normally computing the image intensity based on Equation 3.6 is very time-consuming. To solve this problem, Emil Wolf first introduced the coherent mode representation of the mutual coherence function in 1984 [32], which is an efficient tool for calculating the field propagation. The MCF can be expanded into a diagonal representation of several modes as:

$$J_p(\boldsymbol{r_{o1}}, \boldsymbol{r_{o2}}, \omega) = \sum_p \lambda_p(\omega) p_n(\boldsymbol{r_{o1}}, \omega) p_n(\boldsymbol{r_{o2}}, \omega)^*, \tag{3.12}$$

$$J_o(\boldsymbol{r_{o1}}, \boldsymbol{r_{o2}}, \omega) = \sum_o \lambda_o(\omega) o_m(\boldsymbol{r_{o1}}, \omega) o_m(\boldsymbol{r_{o2}}, \omega)^*. \tag{3.13}$$

where $\lambda_o(\omega)$ and $\lambda_p(\omega)$ represents the weights for different modes of the object and the probe, and p_n and o_n are orthogonal modes for the probe and the object.

Since the basic model of ptychography is based on the coherent light, the experimental requirements of the previous research are usually stringent to ensure the coherence of light. However, as we mentioned before, there may be some imperfection on the experimental components which will introduce a reduction of coherence to the system. Thibault and Menzal introduced the coherent mode representation of MCF to ptychography [8]. In their paper, the probe and the object function are decomposed into several modes, and each mode is propagated coherently but

added up incoherently to calculate the diffraction pattern. They have shown that ptychography can reconstruct each mode of the probe and the object. Their approach is called the multi-mode ptychography in the following research.

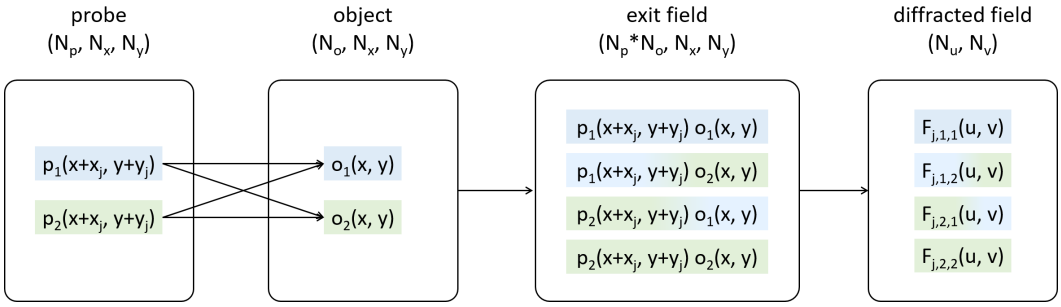


Figure 3.3: Schematic illustration of the forward model of multi-mode ptychography. The object function and the probe function are decomposed into several modes. The definitions of the exit field and the diffracted field are similar to Figure 3.2

.

3.3 Ambiguity in multi-mode ptychography

As we discussed in the previous session, the MCF can be decomposed into several modes to simulate the intensity of the diffraction pattern, which was introduced to ptychography to deal with the partially coherent illumination and the vibration of the sample (object). For a given probe and object, the sets of orthogonal modes are uniquely determined. However, since there is no restriction implemented to guarantee the orthogonality between modes, the reconstructed modes may be any linear combination of the uniquely determined orthogonal modes depending on the measurement noises and the initial conditions. In this session, we will explain the origin of the ambiguity in the reconstructed probe and object modes following the mathematical analysis in Ref. [31].

Taking the notation in Figure 3.3, we assume that the object and the probe can be sufficiently decomposed into N_o and N_p orthogonal modes, respectively. Then we can define unique sets of orthogonal probe modes $P^{n_p}(x,y)$ and orthogonal object modes $O^{n_o}(x,y)$. The exit wave at each scanning position (denoted as index j) can be modeled as:

$$F_j^{(n_p,n_o)}(u,v) = \mathcal{D}[P^{n_p}(x,y)O^{n_o}(x,y)]. \tag{3.14}$$

If we assemble the probe modes as a row vector $\boldsymbol{P} = [P^1(x,y), P^2(x,y), ..., P^{N_p}(x,y)]$, the object modes as a row vector $\boldsymbol{O} = [O^1(x,y), O^2(x,y), ..., O^{N_o}(x,y)]$ and the diffracted field as a row vector $\boldsymbol{F}_j = [F_j^{(1,1)}(u,v), F_j^{(j,1,2)}(u,v), ..., F_j^{(j,N_p,N_o)}(u,v)]$, the relation between the three row vectors can be denoted as a series of a state-wise Kronecker product:

$$\boldsymbol{F}_j = [\mathcal{D}(\boldsymbol{P} \otimes \boldsymbol{O})]. \tag{3.15}$$

In Figure 3.4, we give an example of object with two states $\boldsymbol{O} = [O^1(x,y), O^2(x,y)]$ and probe with two states $\boldsymbol{P} = [P^1(x,y), P^2(x,y)]$ to illustrate the state-wise Kronecker product. The

state-wise Kronecker product of these two row vectors is a row vector of the exit waves $\boldsymbol{U} = [U^{(1,1)}(x,y), U^{(1,2)}(x,y), U^{(2,1)}(x,y), U^{(2,2)}(x,y)]$, where $U^{(n_p,n_o)} = P^{n_p} O^{n_o}$. Each elements in $\boldsymbol{U}$ is a pixel-wise product between the probe and the object.

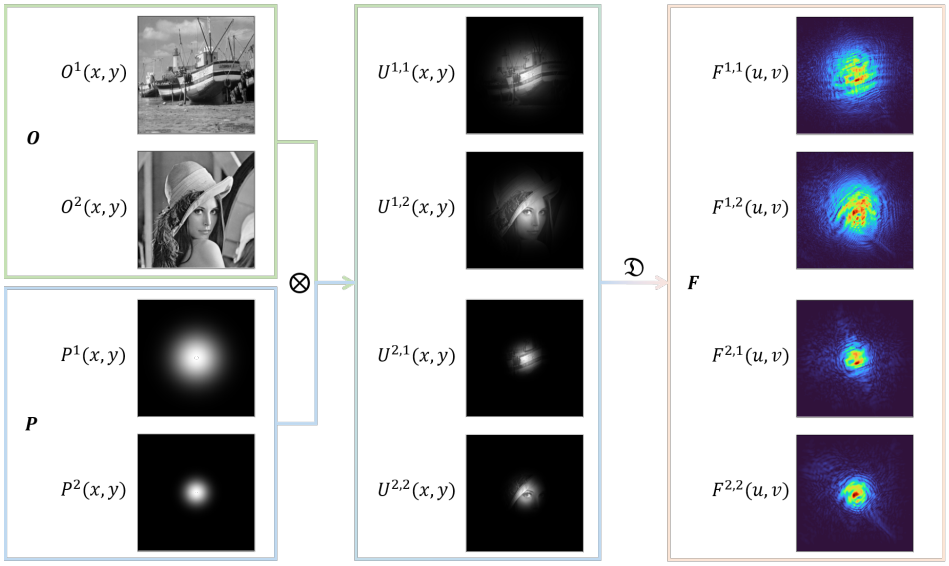


Figure 3.4: Schematic illustration of state-wise Kronecker product, denoted as $\otimes$ in the figure. The $\mathcal{D}$ denotes the propagator. Here the simulation is based on the Fraunhofer diffraction in the far field region.

To mathematically prove that the modes reconstructed by the ptychographic algorithm may be a linear combination of the orthogonal modes, here we denote the result from the ptychography algorithm as $\widetilde{\boldsymbol{P}} = [\widetilde{P}^1(x,y), \widetilde{P}^2(x,y), ..., \widetilde{P}^{N_p}(x,y)]$ and $\widetilde{\boldsymbol{O}} = [\widetilde{O}^1(x,y), \widetilde{O}^2(x,y), ..., \widetilde{O}^{N_o}(x,y)]$, respectively. The retrieved diffracted field can be expressed as the propagated state-wise Kronecker product of $\widetilde{\boldsymbol{P}}$ and $\widetilde{\boldsymbol{O}}$ as follows:

$$\widetilde{\boldsymbol{F}}_j = [\mathcal{D}(\widetilde{\boldsymbol{P}} \otimes \widetilde{\boldsymbol{O}})]. \tag{3.16}$$

The relation between the reconstructed (non-orthogonal) $\widetilde{\boldsymbol{P}}$ and $\widetilde{\boldsymbol{O}}$ and the original (orthogonal) $\boldsymbol{P}$ and $\boldsymbol{O}$ can be expressed as:

$$\widetilde{\boldsymbol{P}} = \boldsymbol{P}\boldsymbol{K}_{\boldsymbol{p}}, \tag{3.17}$$

$$\widetilde{\boldsymbol{O}} = \boldsymbol{O}\boldsymbol{K}_{\boldsymbol{o}}, \tag{3.18}$$

where $\boldsymbol{K_p}$ and $\boldsymbol{K_o}$ are two complex-valued transformation matrices with row- and column-dimensions of $N_p \times \widetilde{N_p}$ and $N_o \times \widetilde{N_o}$, respectively. Since both the reconstructed and the original sets satisfy the modulus constraints of the diffracted field, the summed intensity of the diffracted field generated by these two sets should be equal to each other:

$$\boldsymbol{F}_j \boldsymbol{F}_j^{\dagger} = \widetilde{\boldsymbol{F}}_j \widetilde{\boldsymbol{F}}_j^{\dagger}, \tag{3.19}$$

where † represents the Hermitian conjugate. The modulus of the original orthogonal sets can be expressed as:

$$\boldsymbol{F}_j\boldsymbol{F}_j^{\dagger} = \sum_{n_p=1}^{N_p}\sum_{n_o=1}^{N_o} F^{(n_p,n_o)}F^{(n_p,n_o)\dagger} = \sum_{n_p=1}^{N_p}\sum_{n_o=1}^{N_o}\left|\mathcal{D}[P^{n_p}O^{n_o}]\right|^2. \tag{3.20}$$

Considering Equation 3.17 and Equation 3.18, the modulus of the reconstructed non-orthogonal sets can be expressed as:

$$\widetilde{\boldsymbol{F}}_j\widetilde{\boldsymbol{F}}_j^{\dagger} = [\mathcal{D}(\widetilde{\boldsymbol{P}}\otimes\widetilde{\boldsymbol{O}})][\mathcal{D}(\widetilde{\boldsymbol{P}}\otimes\widetilde{\boldsymbol{O}})]^{\dagger} = [\mathcal{D}(\boldsymbol{P}\boldsymbol{K_p}\otimes\boldsymbol{O}\boldsymbol{K_o})][\mathcal{D}(\boldsymbol{P}\boldsymbol{K_p}\otimes\boldsymbol{O}\boldsymbol{K_o})]^{\dagger}. \tag{3.21}$$

Equation 3.17 and Equation 3.18 can be defined state-wisely as:

$$\widetilde{P}^{n_p} = \sum_{m_p} P^{m_p}k_{m_p,n_p}, \tag{3.22}$$

$$\widetilde{O}^{n_o} = \sum_{m_o} O^{m_o}k_{m_o,n_o}, \tag{3.23}$$

where k_{m_p,n_p} is the element of K_p and k_{m_o,n_o} is the element of K_o. Then the propagated field of each state is:

$$\mathcal{D}[\widetilde{P}^{n_p}\widetilde{O}^{n_o}] = \sum_{m_p}\sum_{m_o}\mathcal{D}[P^{m_p}k_{n_p}O^{m_o}k_{n_o}]k_{m_p,n_p}k_{m_o,n_o}. \tag{3.24}$$

We can rewrite Equation 3.24 into matrix containing all the modes:

$$\mathcal{D}(\widetilde{\boldsymbol{P}}\otimes\widetilde{\boldsymbol{O}}) = \mathcal{D}(\boldsymbol{P}\boldsymbol{K_p}\otimes\boldsymbol{O}\boldsymbol{K_o})\boldsymbol{K_p}\otimes\boldsymbol{K_o}, \tag{3.25}$$

then Equation 3.21 can be rewritten as:

$$\widetilde{\boldsymbol{F}}_j\widetilde{\boldsymbol{F}}_j^{\dagger} = [\mathcal{D}(\boldsymbol{P}\otimes\boldsymbol{O})][\boldsymbol{K_p}\otimes\boldsymbol{K_o}][\boldsymbol{K_p}\otimes\boldsymbol{K_o}]^{\dagger}[\mathcal{D}(\boldsymbol{P}\otimes\boldsymbol{O})]^{\dagger} = [\mathcal{D}(\boldsymbol{P}\otimes\boldsymbol{O})][\mathcal{D}(\boldsymbol{P}\otimes\boldsymbol{O})]^{\dagger}. \tag{3.26}$$

Subsequently, it requires:

$$[\boldsymbol{K_p}\otimes\boldsymbol{K_o}][\boldsymbol{K_p}\otimes\boldsymbol{K_o}]^{\dagger} = \boldsymbol{I}_{N_pN_o}, \tag{3.27}$$

where $\boldsymbol{I}_{N_pN_o}$ denotes the identity matrix with dimension $N_p \times N_o$. This means $\boldsymbol{K_p}$ and $\boldsymbol{K_o}$ have to satisfy:

$$\boldsymbol{K_p}\boldsymbol{K_p^{\dagger}} = c\boldsymbol{I_{Np}}, \tag{3.28}$$

$$\boldsymbol{K_o}\boldsymbol{K_o^{\dagger}} = \boldsymbol{I_{No}}/c, \tag{3.29}$$

where c is the positive real number and $\boldsymbol{I}$ is a identity matrix. This further implies that:

$$\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{P}}^{\dagger} = c\boldsymbol{P}\boldsymbol{P}^{\dagger}, \tag{3.30}$$

$$\widetilde{\boldsymbol{O}}\widetilde{\boldsymbol{O}}^{\dagger} = \boldsymbol{O}\boldsymbol{O}^{\dagger}/c. \tag{3.31}$$

Equation 3.30 and Equation 3.31 indicate that any linear combination of the orthogonal solutions can also satisfy the modulus constraints, which leads to the ambiguity of the reconstruction result.

3.4 Orthogonalization of the reconstructed probe

According to the concept of multi-mode ptychography, the partially coherent illumination probe can be decomposed into several orthogonal modes. Here we follow the analysis in Ref.[31], demonstrating that imposing the constraint on the reconstructed modes can break the ambiguity in the ptychographic result.

For the convenience of analysis, each probe mode $P^{n_p}(x,y)$ is reformed to a column vector $\boldsymbol{s}^{n_p} = [p^{n_p}(x_1,y_1), p^{n_p}(x_1,y_2), ..., p^{n_p}(x_n,y_n)]^T$, where T represents the transpose operator, $p^{n_p}(x_n,y_n)$ represents the pixel value of the n_p mode at the position of (x_n,y_n), n is the pixel number. Based on the orthogonal assumption, we have:

$$\boldsymbol{s}^{n_{p1}\dagger}\boldsymbol{s}^{n_{p2}} = \begin{cases} 0, & n_{p1} \neq n_{p2} \\ \\ \eta, & n_{p1} = n_{p2} \end{cases} \tag{3.32}$$

where the η is a positive real number representing the power of the mode. All the column vectors of modes can be assembled to form a matrix:

$$\boldsymbol{S} = [\boldsymbol{s}^1, \boldsymbol{s}^2, \cdots, \boldsymbol{s}^{n_p}] = \begin{bmatrix} p^1(x_1,y_1) & p^2(x_1,y_1) & \cdots & p^{n_p}(x_1,y_1) \\ p^1(x_1,y_2) & p^2(x_1,y_2) & \cdots & p^{n_p}(x_1,y_2) \\ \vdots & \vdots & \ddots & \vdots \\ p^1(x_n,y_n) & p^2(x_n,y_n) & \cdots & p^{n_p}(x_n,y_n) \end{bmatrix}. \tag{3.33}$$

Then we can form a $\boldsymbol{T}$ matrix containing all the dot products of the probe states:

$$\boldsymbol{T} = \boldsymbol{S}^\dagger\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}^1\dagger\boldsymbol{s}^1 & \boldsymbol{s}^1\dagger\boldsymbol{s}^2 & \cdots & \boldsymbol{s}^1\dagger\boldsymbol{s}^{N_p} \\ \boldsymbol{s}^2\dagger\boldsymbol{s}^1 & \boldsymbol{s}^2\dagger\boldsymbol{s}^2 & \cdots & \boldsymbol{s}^2\dagger\boldsymbol{s}^{N_p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{s}^{N_p}\dagger\boldsymbol{s}^1 & \boldsymbol{s}^{N_p}\dagger\boldsymbol{s}^2 & \cdots & \boldsymbol{s}^{N_p}\dagger\boldsymbol{s}^{N_p} \end{bmatrix} = \begin{bmatrix} \eta_1 & 0 & \cdots & 0 \\ 0 & \eta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_{N_p} \end{bmatrix}. \tag{3.34}$$

Recalling the definition of K_p in Equation 3.17, we can also form a matrix $\widetilde{\boldsymbol{T}}$ based on the reconstructed non-orthogonal modes:

$$\widetilde{\boldsymbol{T}} = \widetilde{\boldsymbol{S}}^\dagger\widetilde{\boldsymbol{S}} = \boldsymbol{K_p^\dagger P^\dagger P K_p} = \boldsymbol{K_p^\dagger T K_p}, \tag{3.35}$$

which is actually an eigen decomposition of the matrix $\widetilde{\boldsymbol{T}}$, with the column vectors in $\boldsymbol{K_p^\dagger}$ the eigen vectors and the diagonal elements in the matrix $\boldsymbol{T}$ the eigen values. Starting from this point, we can apply an eigen decomposition to the matrix $\widetilde{\boldsymbol{T}}$, which is formed using the reconstructed modes, to get the matrix $\boldsymbol{K_p}$, and then using Equation 3.17 to obtain the orthogonal modes.

Here we give a schematic illustration of the orthogonalization procedure for three reconstructed probe modes (Figure 3.5). The reconstructed modes first are reshaped and assembled to form the matrix $\widetilde{\boldsymbol{S}}$. Then we multiply $\widetilde{\boldsymbol{S}}$ and its Hermittian transpose to form matrix $\widetilde{\boldsymbol{T}}$. Eigen decomposition is applied to $\widetilde{\boldsymbol{T}}$ to obtain K_p for reallocating the probe modes. Finally, $\boldsymbol{S}$ is split and reshaped to the original shape of the probe modes.

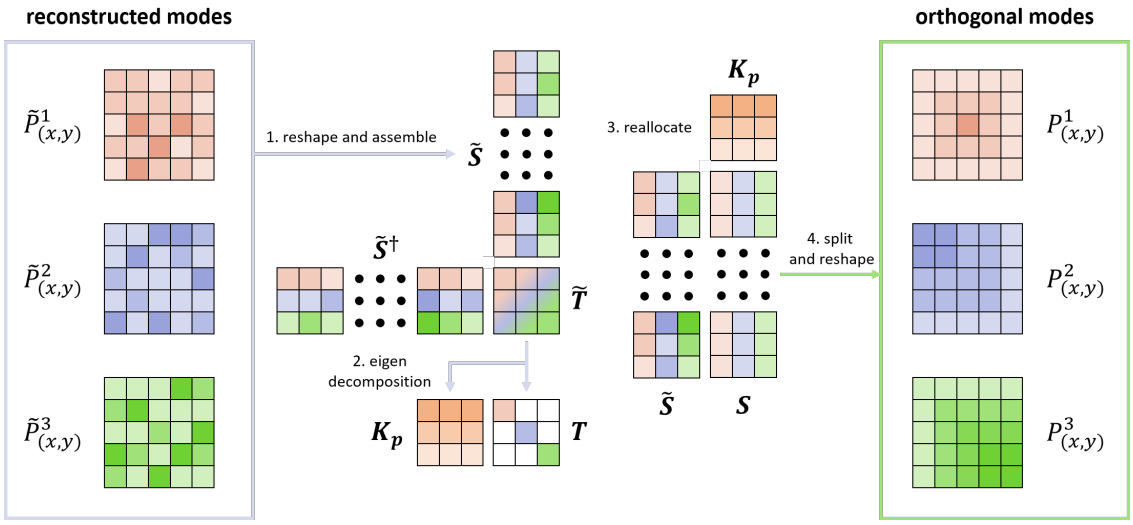


Figure 3.5: Schematic illustration of orthogonalization of the reconstructed probe. Step 1: Reshape the reconstructed modes to column vectors and stack them together to construct the $\widetilde{S}$ matrix. Step 2: Implement eigen decomposition on the matrix product of the $\widetilde{S}$ and the $\widetilde{S}^{\dagger}$. Step 3: Multiply the eigenvectors with the $\widetilde{S}$ matrix to reallocate the pixel values of each modes. Step 4: Split the result of step 3 and reshape each mode to the initial shape of the probe modes.

3.5 Experimental results

To validate the performance of the orthogonalization, we perform a ptychographic reconstruction on an experimental data set. The illumination probe is generated by focusing a laser through a diffuser, which can be regarded as a partially coherent illumination source. The wavelength of the laser is 532nm. The sample is a part of a binary Siemens star.

To define the starting point for the ptychographic reconstruction, the sample is initialized with uniform amplitude and zero phase; the probe is initialized as a circular aperture with the size of half of the field of view. This initialization will also be implemented in the following simulations. The weights of different modes are randomly defined, but are guaranteed to sum to 1. Here we define the batch size as the number of the diffraction patterns to work through before updating the internal modal parameters, and the epoch size as the number that the learning algorithm will work through the entire training data set. In other words, at the end of the batch, the simulated intensity of the diffracted field is compared to the measured intensity of the diffracted field and an error is calculated. In one epoch, each diffraction pattern in the data set has an opportunity to update the internal modal parameters. In this reconstruction, we set the batch size to be 1, and the epoch size to be 40. Every time when the algorithm runs through all the diffraction patterns and is ready to start a new epoch, the order of the input scanning position for the new epoch is randomized.

Since the degree of coherence of the illumination is unknown, a series of reconstruction of this data set is implemented with different number of probe modes, and the different reconstructed samples are shown in Figure 3.6. It can be clearly observed that the reconstruction with a single mode is considerably worse than that with multiple modes. The failure is attributed to the fact that

a single mode cannot incorporate the partial coherence effect of the illumination. As can be seen in Figure 3.6, the amount of details in the reconstructed sample does not increases as we increase in the number of modes when using the number of modes exceeding 6, implying that it is sufficient to use 6 modes to characterize the spatial coherence of the illumination beam in the experiment.

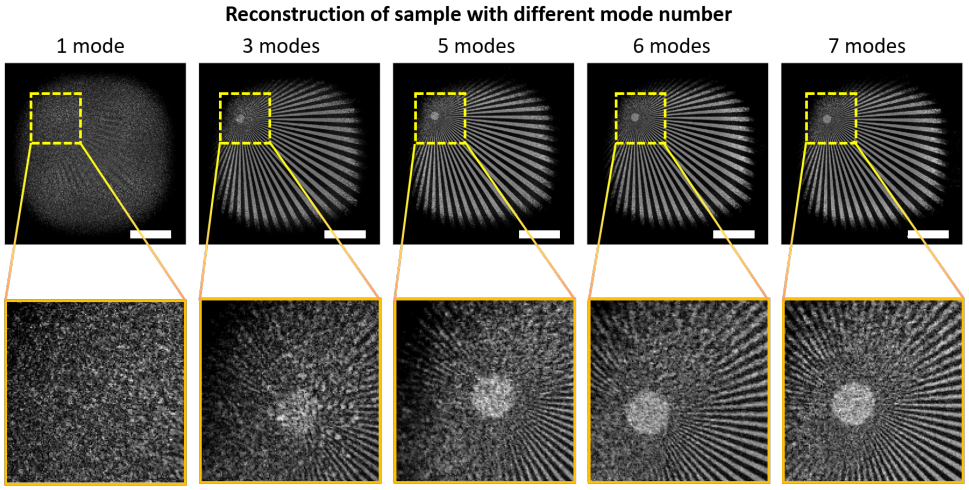


Figure 3.6: Reconstruction of the sample with different probe mode numbers: from left to right the amount of included probe modes is increased. The reconstructed details imply that there are around 6 modes in the illumination. The scale bars indicate a length of 1mm.

Due to the ambiguity of the reconstructed probe modes and the stochastic order of the input scanning positions, we cannot get a reproducible result of the probe modes. Figure 3.7 demonstrates the ptychographic reconstructions with the same starting point. Due to the strong phase curvature, the structure of the probe modes are difficult to be observed. So the first mode with highest spectral weight is used as the reference to remove all the phase curvatures of the other modes. For convenience of observing, the central part of the other probe modes are also enlarged. Due to the randomness of the scanning positions, although all these three ptychographic reconstructions share the same starting point and hyperparameters, neither the reconstructed profile nor the spectral weight is consistent.

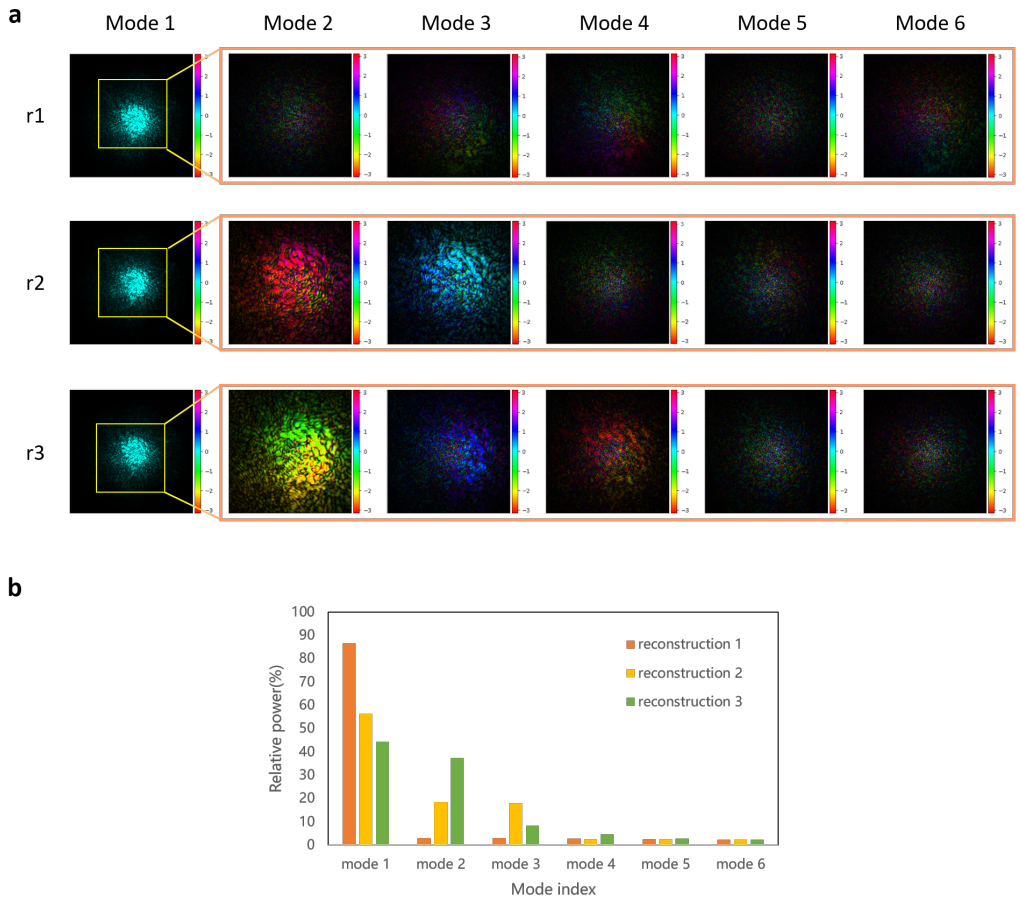


Figure 3.7: Ptychographic reconstructions of six probe modes without orthogonalization. The three reconstructions here share the same starting point and hyperparameters. The probe mode with the highest spectral weight is used as the reference to remove all the phase curvatures. The profiles of the probe modes (a) and the spectral weights (b) are not consistent.

After implementing the orthogonalization on the ptychographic algorithm, Figure 3.8 demonstrates the ptychographic reconstructions with a better reproducibility. Both the profiles and the modal weights are consistent, despite of the stochastic order of the input scanning positions.

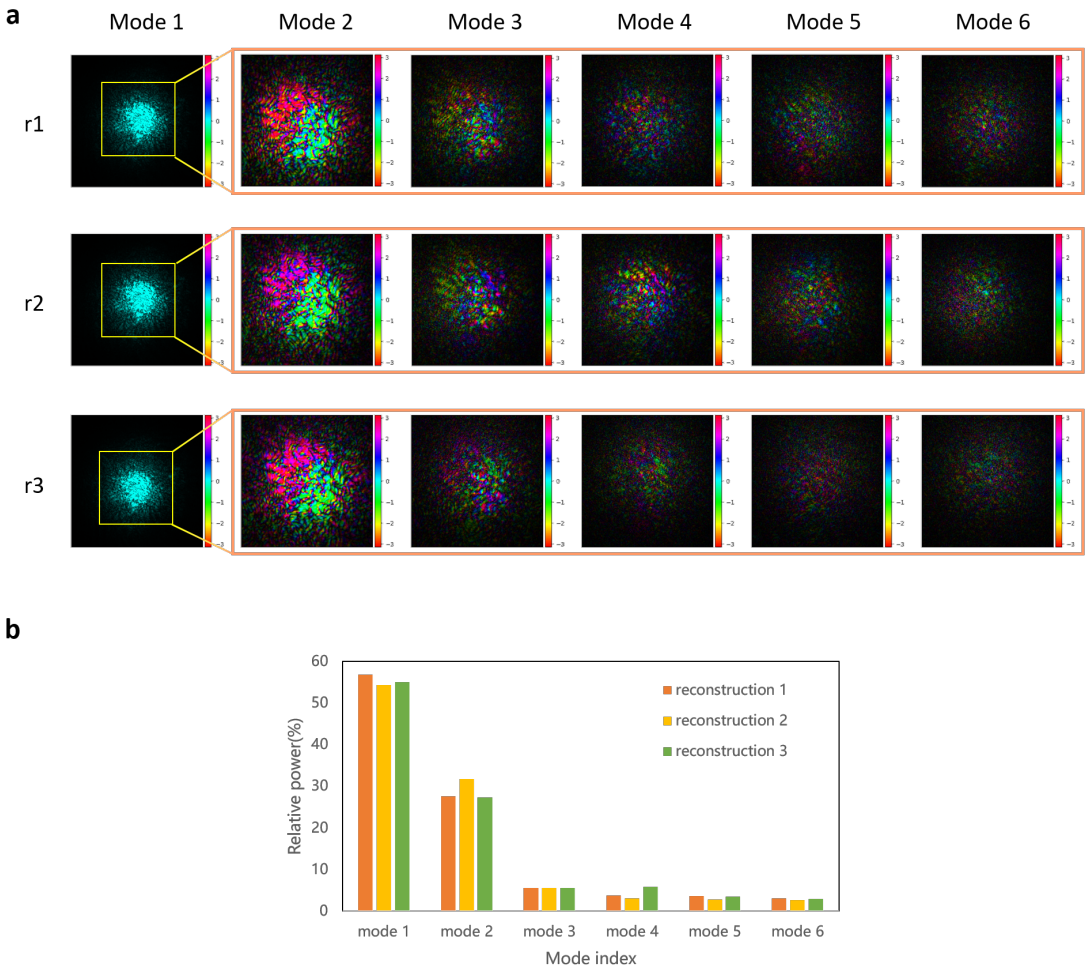


Figure 3.8: Ptychographic reconstructions of six probe modes with orthogonalization. The three reconstructions here share the same starting point and hyperparameters. The probe mode with the highest spectral weight is used as the reference to remove all the phase curvatures. The profiles of the probe modes (a) and the spectral weights (b) are consistent.

Except for the reconstructions with the same number of modes, we also test its performance on different numbers of modes. As shown in Figure 3.9, the ptychographic reconstructions are similar to the previous results, despite of the different mode numbers. The first mode, containing most of the power in the illumination, looked similar for all the reconstructions. And the structures of first three modes containing most of the energy are clear. Although the other modes are not that strong, the reconstructed sample implies their necessity. Although the mode number is not correct, the orthogonalization can ensure the structure of the reconstructed modes similar to the correct solutions. If the mode number exceeds the correct value, orthogonalization can help us to identify the redundant modes, of which the profile and the weight are chaos and unclear. This can be useful to calibrate the illumination probes for ptychography.

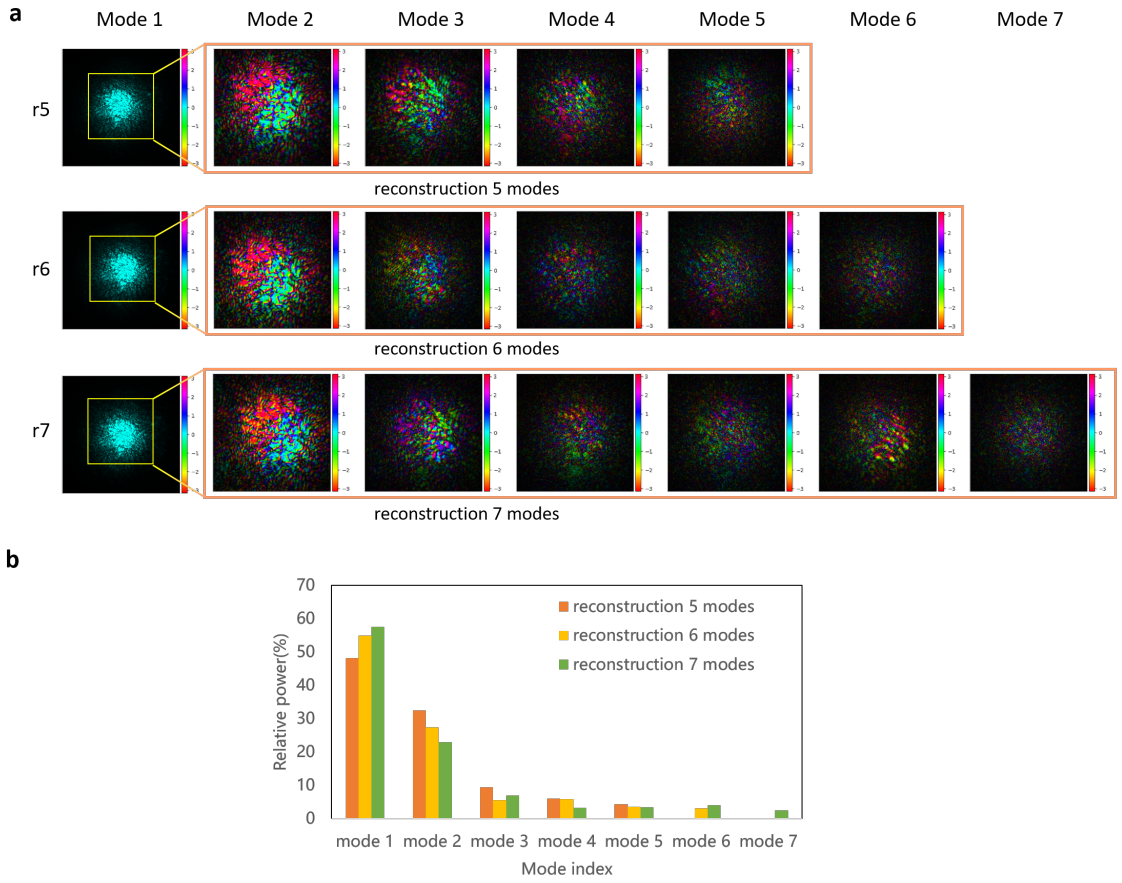


Figure 3.9: Reconstruction of probe modes after implementing orthogonalization. The brightness of the image represents the amplitude, and the color represents the phase (same below). In (a), from left to right the relative power of the modes decreases. When the mode number is increased to seven, the relative power and the structure of the mode seven and mode six are randomly distributed and the relative power is still very small.

These results validate that the orthogonalization can break the probe reconstruction ambiguity. We also have test multiple modes of the sample on this data set, however, since the sample is static and the structure is not designed to contain several independent shapes like in Ref.[31], the relative power of the other modes of the sample are extremely small.

Chapter 4

Polychromatic ptychography

Following the inspiration of the multimodal ptychography, another new dimension, wavelength, can also be introduced into the ptychographic algorithm. The polychromatic illumination guarantees a deeper understanding of underlying physics in the object's absorption and phase response at different wavelengths. To date, several studies have investigated ptychography with polychromatic illumination. To solve the sampling problem due to the multi wavelengths, some of them use a scale factor to scale the diffraction pattern [12][33][34], some introduce a new propagator to simulate the diffraction pattern [35]. To make it suitable for ptychography with the AD framework, in this session, we first introduce the basic model model of the polychromatic ptychography and the related sampling problem. Then we discuss and test two different methods for simulating the polychromatic diffraction. Based on these new methods, an optimization of propagation distance and wavelength with an experimental data set is discussed. Finally, an experimental ptychographic data set with an illumination comprising 9 wavelengths is used as a validation.

4.1 Model of multi-mode ptychography with polychromatic illumination

Conventionally, the light source for ptychography is almost monochromatic, but sometimes it will suffer from some finite bandwidth. However, what we discuss here is a more ambitious model with more discrete wavelengths, which means the intensity of the diffraction pattern can be the summation of the diffraction patterns of these wavelengths. Then the physical model of the polychromatic ptychography can be expressed as:

$$I_j(u,v) = \sum_{\lambda} \sum_{m,n} |\mathcal{D}_{\lambda,z}[o_m(x,y)p_n(x+x_j, y+y_j)]|^2. \tag{4.1}$$

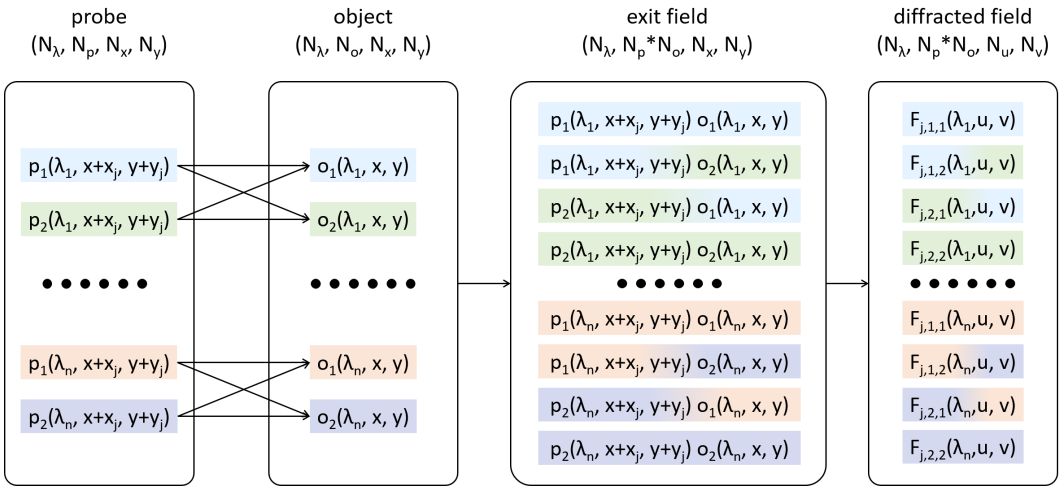


Figure 4.1: Schematic illustration of forward model of multi-mode ptychography with polychromatic illumination.

4.2 Simulation of the free space propagation

In the previous discussion, the detector in the ptychography experiment collects the diffraction pattern of an illuminated sample. To simulate this procedure, the propagation is the key to connect the exit wave and the propagation field. To simulate the propagation, three basic methods (also referred to as propagator, the operator $\mathcal{D}$ in Equation 4.1) are introduced[36]: Fraunhofer diffraction aims to deal with the far field diffraction:

$$U(x_2, y_2, z) = e^{ik\frac{x_2^2+y_2^2}{2z}}\mathcal{F}\{U(x_1, y_1, 0)\}, \tag{4.2}$$

the Fresnel method aims to deal with the near field diffraction where its condition is satisfied:

$$U(x_2, y_2, z) = \frac{e^{ikz}}{i\lambda z}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(x_1, y_1, 0)e^{i\frac{k}{2z}[(x_2-x_1)^2+(y_2-y_1)^2]}dx_1dy_1, \tag{4.3}$$

and the angular spectrum method aims to deal with the near field diffraction as:

$$U(x_2, y_2, z) = \mathcal{F}^{-1}\{\mathcal{F}\{U(x_1, y_1, 0)\}e^{i2\pi z\sqrt{\frac{1}{\lambda^2}-f_x^2-f_y^2}}\}, \tag{4.4}$$

where $U(x_1, y_1, 0)$ and $U(x_2, y_2, z)$ represent the wave field before and after propagation over distance of z; f_x and f_y represent the spatial frequencies; $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier transform and its inverse, respectively, and k is the wave number.

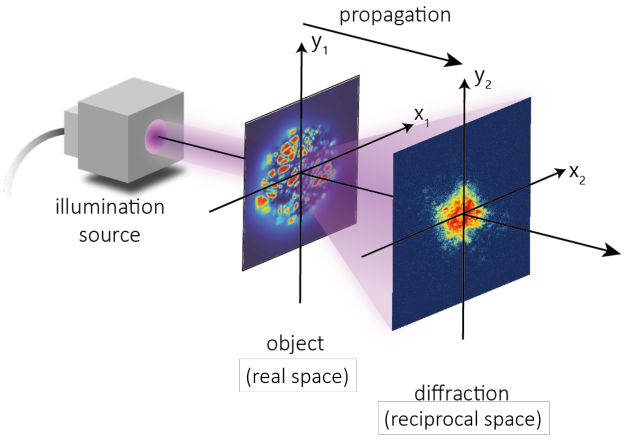


Figure 4.2: Schematic illustration of the geometry used in the propagation model.

Conventionally, the methods we mentioned are computed with the fast Fourier transform algorithm (FFT). However, there are certain drawbacks associated with the use of the FFT algorithm, which requires the pixel size and the number of pixels in the object domain and the detector domain to follow some specific conditions.

To demonstrate the sampling problem in the FFT, we take the 1D discrete Fourier transform (DFT) as an example:

$$F_m = \sum_{n=0}^{N-1} f_n e^{-i2\pi x_n u_m}, \tag{4.5}$$

where x_n and u_m are discrete sampling points in the spatial domain and the spatial frequency domain, respectively, which are given by:

$$x_n = x_0 + n\Delta x, \tag{4.6}$$

$$u_m = u_0 + m\Delta u, \tag{4.7}$$

where x_0 ad u_0 are the starting locations, n and m are the integer indices (with number of pixels being equal to N and M, respectively), and Δx and Δu are the sampling intervals of the sampling in the spatial and the spatial frequency domain, respectively. The coordinates of the reciprocal plane x_2 and the spatial frequency u satisfy the following relation:

$$u = \frac{x_2}{\lambda z}, \tag{4.8}$$

where λ is the wavelength of the source and z is the propagation distance.

To implement the FFT algorithm to calculate the Fourier transform, the following specific conditions have to be satisfied:

$$x_0 = u_0 = 0, \tag{4.9}$$

$$\frac{1}{\Delta x_1 \Delta u} = \frac{\lambda z}{\Delta x_1 \Delta x_2} = M = N. \tag{4.10}$$

If we use the FFT algorithm to simulate the diffracted field, once we define the pixel size Δx_1, Δx_2, and pixel number M, N on the spatial domain of the object and the detector, respectively, to satisfy Equation 4.10, the product of the propagation distance and the wavelength are determined. However, in the ptychographic experiment with polychromatic illumination, the sample is irradiated by different wavelengths simultaneously, and the distance between the sample and the detector is fixed for different wavelengths. Then the product of the wavelengths and the propagation distance can take several numerical values, which means Equation 4.10 will be violated.

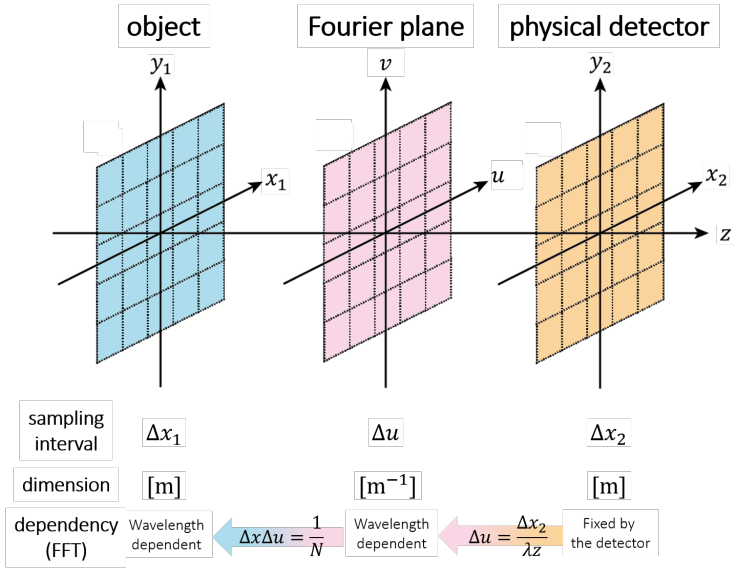


Figure 4.3: Schematic illustration of the dependency between the sampling intervals on different planes based on the FFT algorithm.

In a ptychographic research paper with polychromatic illumination, this violation was resolved by scaling the object field with a wavelength-dependent factor [34]. Since the pixel size Δx_2 of the detector is fixed in the experiment, the pixel sizes Δx_1 of the object for different wavelengths can be determined based on Equation 4.10. And to avoid the problem of aliasing, the object pixel size Δx, calculated with the shortest wavelength, will be used as a reference. Then the pixel sizes of the object calculated with the other wavelengths will be scaled to the same values as used for the reference.

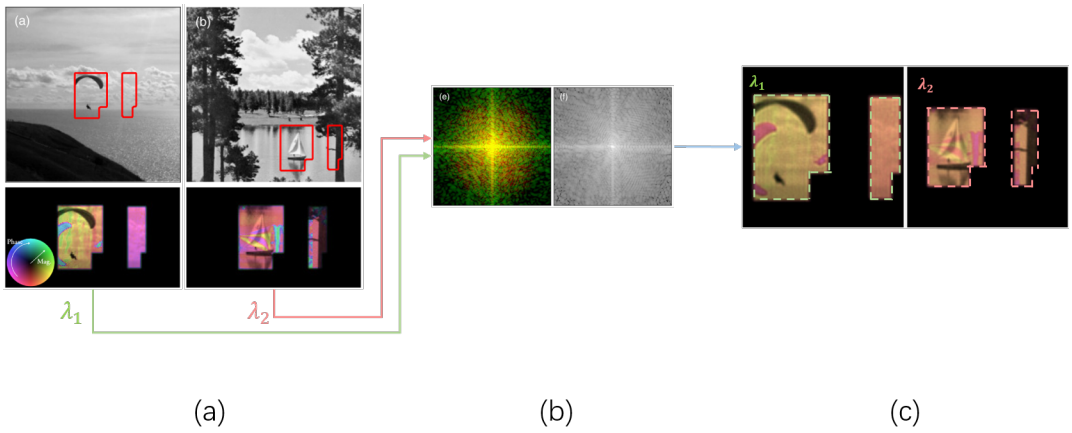


Figure 4.4: A reconstruction result with illumination of two wavelengths [34]. (a) Paraglider and sailboat images were used to simulate the objects, respectively. These two objects were radiated with different wavelengths, λ_1 and λ_2, respectively. (b) Polychromatic diffraction pattern shows the individual contributions from λ_1 and λ_2 in red and green, respectively. The grey pattern denotes the polychromatic data used by the algorithms. (c) Reconstructions of the two objects radiated with two different wavelengths. In (a), the selected regions of the two images are the same size. Due to the scaling factor used in this research paper to satisfy the requirement of the FFT algorithm, the scaling effect can be seen clearly in the reconstructions (c).

Another ptychographic research paper with polychromatic illumination used a two step propagator to solve the sampling problem [35]. Because of the double Fourier transform (one forward FT and one inverse FT) of the angular spectrum approach, its sampling intervals in the two respective planes does not depend on the wavelength. Using this property, the exit wave first propagates to the intermediate plane with the angular spectrum method with a fixed pixel size, and subsequently propagates over the remaining distance $z - z_{as}$ to the detector plane with a Fresnel propagator (with z_{as} the distance over which the angular spectrum propagation has been computed starting from the object plane). After introducing the virtual intermediate plane, The sampling intervals of the object plane and the intermediate plane are equal to each other. The sampling intervals of the intermediate plane and the detector plane have to satisfy Equation 4.10. By changing the position of the intermediate plane, the ratio between the pixel sizes Δx_1 and Δx_2 in the intermediate plane and the detector plane, respectively, can be selected as we need.

In the next section, we will introduce a new algorithm, Chirp Z-transform (CZT), to replace the FFT algorithm, which can be implemented in all the three propagators.

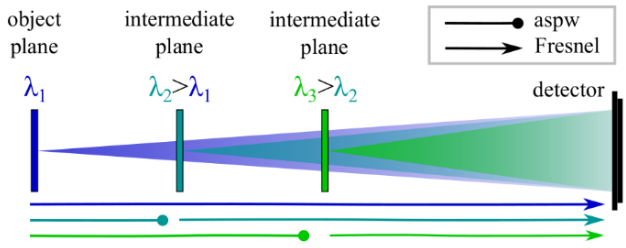


Figure 4.5: A two-step propagation scheme is used to scale the pixel size of each spectral reconstruction to be the same [35]. The spectral wave emanating from the sample at the lowest wavelength directly propagated into the detector plane via Fresnel propagation. All other spectral waves are first propagated to intermediate planes using an angular spectrum propagator. In a second step each spectral exit waves is propagated from its respective intermediate plane to the detector.

4.3 Chirp Z-transform

4.3.1 Bluestein's algorithm

CZT is expressed as a convolution and implemented with the FFT algorithm and its inverse [37]. Specifically, the CZT calculates the Z-transform at a finite number of points z_k along a logarithmic spiral contour, defined as:

$$F_m = \sum_{n=0}^{N-1} f_n z_m^{-n}, \tag{4.11}$$

$$z_m = AW^{-m}, \quad m = 0, 1, \ldots, M-1, \tag{4.12}$$

where A and W are both complex numbers. If we set $A = 1$ and $W = \exp(-i2\pi\Delta u\Delta x)$, then again Equation 4.9 and Equation 4.10 are satisfied, and hence the CZT algorithm is equivalent to FFT algorithm.

The exponential term in Equation 4.5 can be expanded as:

$$\exp(-i2\pi x_n u_m) = \exp[-i2\pi(u_0 x_0 + u_0 n\Delta x + x_0 m\Delta u + mn\Delta u\Delta x)]. \tag{4.13}$$

Here we assume that in the case of the starting position $x_0 = 0$, we can make an analogy with $z_m = AW^{-m}$ to obtain:

$$A = e^{i2\pi u_0 \Delta x}, \tag{4.14}$$

$$W = e^{-i2\pi\Delta u\Delta x}. \tag{4.15}$$

However, for practical optical simulation, the starting location x_0 is usually not 0.

We can observe that the exponential term containing x_0 in Equation 4.13 does not contribute to the summation in Equation 4.11. Therefore, the exponential term can be multiplied with the result in the case of $x_0 = 0$ to obtain the result in the case of $x_0 \neq 0$.

Neglecting the exponential term $e^{-i2\pi(x_0 m\Delta u + u_0 x_0)}$, Equation 4.11 can be expressed as:

$$F_m = \sum_{n=0}^{N-1} f_n (AW^{-m})^{-n} = \sum_{n=0}^{N-1} f_n A^{-n} W^{mn}, \quad m = 0, 1, \ldots, M-1. \tag{4.16}$$

Replacing the product mn with the identity $mn = \frac{m^2+n^2-(m-n)^2}{2}$, we obtain:

$$F_m = W^{\frac{m^2}{2}} \sum_{n=0}^{N-1} (f_n A^{-n} W^{\frac{n^2}{2}}) W^{-\frac{(m-n)^2}{2}}, \quad m = 0, 1, \ldots, M-1. \tag{4.17}$$

This summation can be calculated by a convolution operation, in which the FFT/IFFT pair can be used here as:

$$F_m = W^{\frac{m^2}{2}} \text{iFFT}[\text{FFT}[f_n A^{-n} W^{\frac{n^2}{2}}] \text{FFT}[W^{-\frac{k^2}{2}}]], \quad m = 0, 1, \ldots, M-1. \tag{4.18}$$

The inverse CZT can be derived following above derivation straightforwardly.

We can notice that no restrictions on the relation between the sampling interval of the spatial domain Δx and that of the spatial frequency domain Δu are imposed during the derivation of the CZT algorithm. The sampling number in the spatial domain and that in the spatial frequency domain, N and M, are also not required to be equal. The latter observation yields the very significant result that we can use the CZT algorithm to simulate the polychromatic diffracted field with fixed sampling intervals in both the sample plane and the detector plane when the wavelength varies.

4.3.2 Fraunhofer method based on CZT

To test the performance of CZT algorithm, we perform a simulation using the CZT algorithm, instead of the FFT algorithm, to compute the Fraunhofer propagation in Equation 4.2. In this simulation, we use a detector with number of pixels equal to 2048×2048 and pixel size $15\mu m \times 15\mu m$. The wavelength of the light source is 13 nm. The propagation distance between the aperture and the detector is 50 mm. Here we make the pixel size and the pixel number of the sampling in the object plane to satisfy the requirements of FFT algorithm (Equation 4.9 and Equation 4.10). The diameter of the circular aperture in the object is 10% of the field of view (FOV) of the object plane.

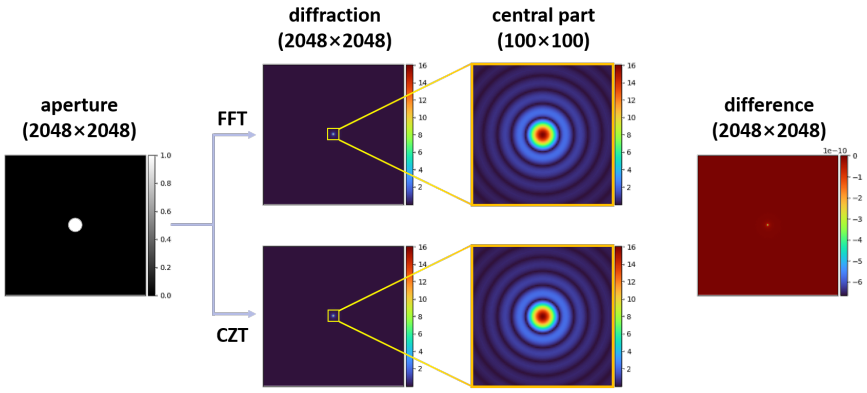


Figure 4.6: Simulated amplitudes of the diffracted fields with Fraunhofer method based on the FFT algorithm (top row) and the CZT algorithm (bottom row), respectively. The aperture in the object plane and the difference between the two simulated results are shown on the left and on the right, respectively.

Figure 4.6 shows the simulated results based on the two algorithms. In the ptychography algorithm, only the intensity is used in the loss function, therefore, here only the amplitude is shown. We can notice that the far field diffraction pattern based on the CZT algorithm is reasonable and the largest difference between the results of the two algorithms is smaller than 10×10^{-11}, which is extremely small relative to the amplitude, whose largest value is around 16.

As we discussed in the previous session, the CZT algorithm relaxes the stringent requirement on the sampling relation between the spatial domain and spatial frequency domain. To validate its performance, we use the same object, wavelength and the propagation distance in the first simulation, but scale the field of view of the detector by a factor of $\frac{1}{5}$, $\frac{1}{10}$ and $\frac{1}{20}$. It should be noticed that in the experiment, the range of the wavelengths of the illumination for polychromatic ptychography will not be so large. The factors we used here are only for validating the algorithm in terms of relaxing the sampling relation.

In Figure 4.7, we can notice that the results from the CZT algorithm with different pixel sizes in the detector plane are still reasonable. For a fair comparison, we plot the cross-sections of different diffraction patterns in the same figure. The simulation results with different scaling factors match well. We can observe that these cross-sections match well with each other in the region where they overlap. Therefore, we can conclude that the CZT algorithm can effectively scale the sampling interval and hence can relax the stringent requirement of the FFT algorithm.

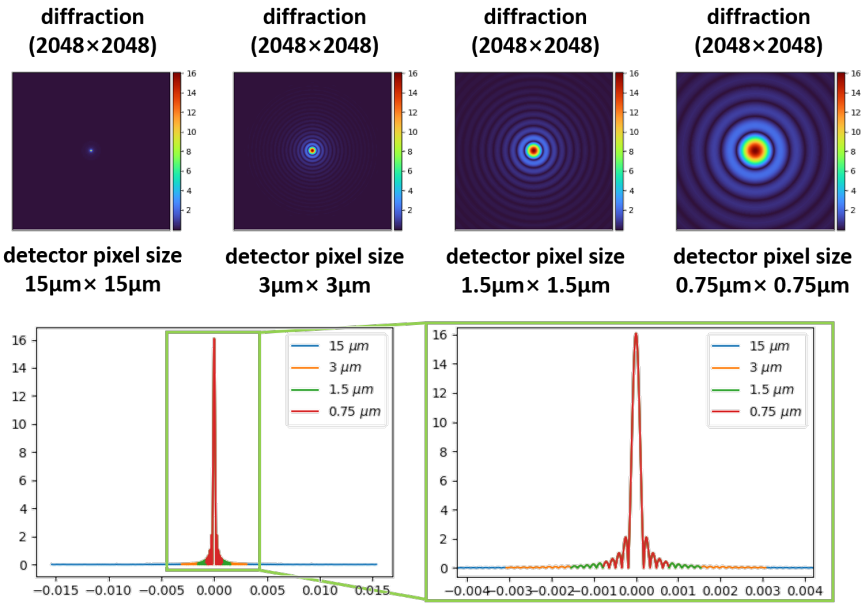


Figure 4.7: Simulation amplitudes of the diffracted fields with Fraunhofer method based on CZT algorithm (bottom row). Taking $15\mu m\times15\mu m$ as the reference, from left to right in the top row, the sampling intervals of the detector are scaled by factor of 1, $\frac{1}{5}$, $\frac{1}{10}$ and $\frac{1}{20}$. For a fair comparison, we plot the cross-sections of different diffraction patterns in the same figure.

4.3.3 Fresnel method based on CZT

To test the performance of the CZT algorithm, we perform a simulation using the CZT algorithm, instead of the FFT algorithm, to compute the Fresnel propagation in Equation 4.3. In this simulation, we use a detector with pixel number 2048×2048 and pixel size $5\mu m \times 5\mu m$. The wavelength of the light source is 750 nm. The propagation distance between the aperture and the detector is 50 mm. Here we make the pixel size and the number of pixels of the sampling in the object plane to satisfy the requirements of the FFT algorithm (Equation 4.9 and Equation 4.10). During the test based on the Fresnel method, we notice that a wrong quadratic phase term will flip the diffraction pattern, which is not observable on the diffracted field of the circular aperture. To break the symmetry of the circular aperture, here an equilateral triangular aperture is used as the object. The side length of the triangular aperture is 20% of the FOV of the object plane.

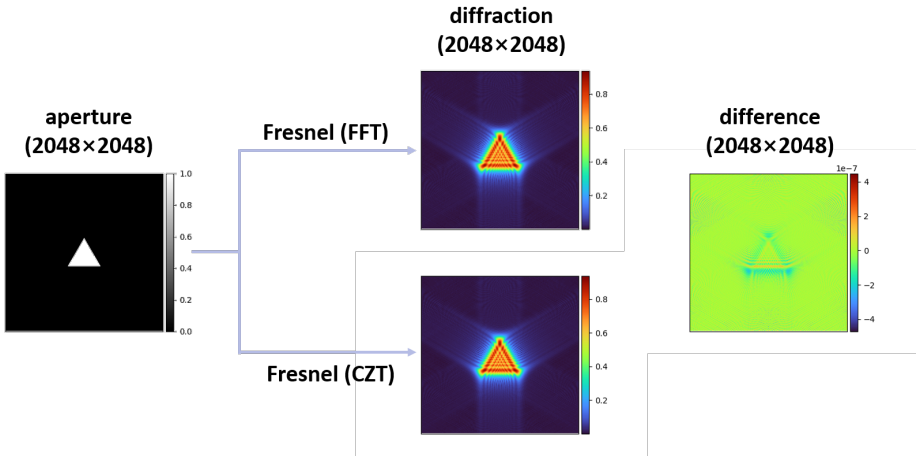


Figure 4.8: Simulated amplitudes of the diffracted fields with Fresnel method based on the FFT algorithm (top row) and the CZT algorithm (bottom row), respectively. The aperture in the object plane and the difference between the two simulated results are shown on the left and on the right, respectively.

Figure 4.8 shows the simulated results based on the two algorithms. We can notice that the near field diffraction pattern based on the CZT algorithm is reasonable and the largest difference between the results of the two algorithms is smaller than 10×10^{-7}, which is extremely small relative to the amplitude, whose largest value is around 1.3.

To validate the performance of the CZT algorithm in relaxing the stringent requirement on the sampling relation between the spatial domain and spatial frequency domain in the Fresnel method, we use the same object, wavelength and the propagation distance as in the first simulation, but scale the FOV of the detector by a factor of 2 and $\frac{1}{2}$.

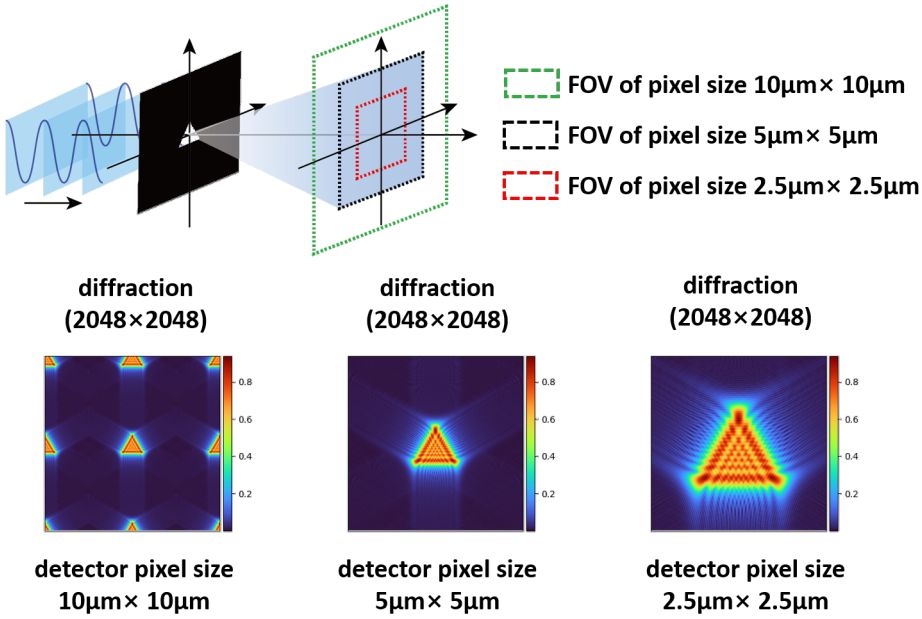


Figure 4.9: Schematic illustration of the scaling test for the Fresnel method (top row). The simulated amplitudes of the diffracted fields with different detector sizes are shown. The aliasing result of the pixel size $10\mu m \times 10\mu m$ attributes to the violation of the Nyquist critical frequency.

Figure 4.9 presents the schematic illustration of the scaling operation on the FOV. The simulated diffracted field of scaling by $\frac{1}{2}$ is reasonable. However, an aliasing result is shown when the FOV is expanded by a factor of 2.

To explain this phenomenon, we should refer to the range of the sampling interval in the frequency domain:

$$\Delta u \leq \frac{1}{L_x} = \frac{1}{M\Delta x}, \tag{4.19}$$

which uses the same notation with Equation 4.9 and Equation 4.10, and L_x denotes the range of the spatial domain. This relation comes from that for a bandlimited function, the region over which the spectrum of the sampled function is nonzero is limited. The maximum width for this region is $1/\Delta x$. The aliasing result with the scaling by 2 attributes to the violation of this theorem. We can imply that to simulate the polychromatic diffraction in ptychography, we have to use the minimum wavelength to define the object pixel size to satisfy the sampling theorem.

4.3.4 Angular spectrum method based on CZT

For the convenience of reading, we write the angular spectrum method here again:

$$U(x_2, y_2, z) = \iint_{-\infty}^{\infty} U(f_x, f_y, 0) \exp\left(iz2\pi\sqrt{\frac{1}{\lambda^2} - f_x^2 - f_y^2}\right) df_x df_y, \tag{4.20}$$

with

$$U(f_x, f_y, 0) = \iint_{-\infty}^{\infty} U(x_1, y_1, 0) \exp[-i2\pi(f_x x_1 + f_y y_1)] dx_1 dy_1 = \mathcal{F}\{U(x_1, y_1, 0)\}, \tag{4.21}$$

where (f_x, f_y) represents the coordinate in the spatial frequency domain, and every spatial frequency is related to an angle.

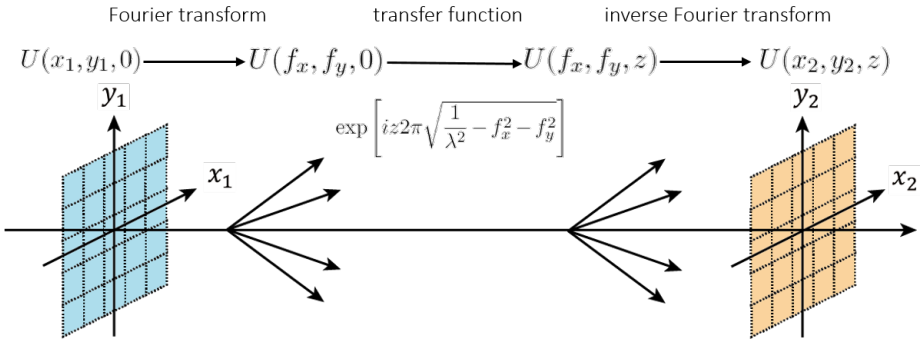


Figure 4.10: Schematic illumination of the angular spectrum method.

If we use the FFT algorithm to calculate the Fourier transform and its inverse, the sampling intervals in the three domains should satisfy:

$$\Delta f_x = \frac{1}{N_x \Delta x_1} = \frac{1}{N_x \Delta x_2}, \quad \Delta f_y = \frac{1}{N_y \Delta y_1} = \frac{1}{N_y \Delta y_2}. \tag{4.22}$$

To break the requirements on the sampling relation between the spatial domain and spatial frequency domain from the FFT/IFFT pair in the angular spectrum method, we have two options to combine the CZT algorithm and the FFT algorithm: CZT/IFFT pair or FFT/ICZT pair.

In fact, only FFT/ICZT can be used in the angular spectrum method. Here we scale the pixel size of the detector (calculated with Equation 4.10) by a factor of 2 to explain the reason. For the CZT/IFFT pair, the sampling intervals in the three domains are obtained as:

$$\Delta x_2 = \frac{1}{N_x \Delta f_x'} = \frac{\Delta x_1}{2}, \quad \Delta y_2 = \frac{1}{N_y \Delta f_y'} = \frac{\Delta y_1}{2}. \tag{4.23}$$

$$\Delta f_x' = \frac{2}{N_x \Delta x_1}, \quad \Delta f_y' = \frac{2}{N_y \Delta y_1}. \tag{4.24}$$

However, the sampling theorem requires the sampling interval of the frequency domain and that of the object domain to satisfy:

$$\Delta f_x' \leq \frac{1}{N_x \Delta x_1}, \tag{4.25}$$

which is violated by the CZT/IFFT pair. So only the FFT/ICZT pair can be used in the angular spectrum method.

To test the performance of the CZT algorithm, we perform a simulation using the FFT/ICZT pair, instead of the FFT/IFFT pair, to compute the angular spectrum method. In this simulation, we use a detector with a number of pixels equal to 2048×2048 and a pixel size $5\mu m \times 5\mu m$. The wavelength of the light source is 750 nm. The propagation distance between the aperture and the detector is 50 mm. In order to compare the result of the FFT/IFFT pair with that of FFT/ICZT pair, here we make the pixel size and the number of the pixels of the sampling in the object plane to satisfy the requirements of the FFT algorithm (Equation 4.9 and Equation 4.10). Here an equilateral triangular aperture is used as the object. The side length of the triangular aperture is 30% of the FOV of the object plane.

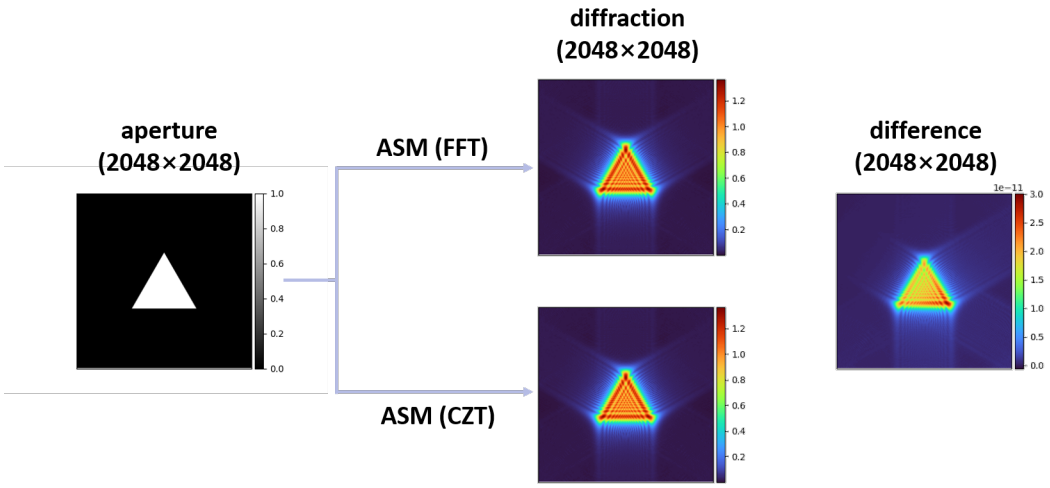


Figure 4.11: Simulated amplitudes of the diffracted fields based on the angular spectrum method with the FFT/IFFT pair (top row) and the FFT/ICZT pair (bottom row). The aperture in the object plane and the difference between the two simulated results are shown on the left and on the right, respectively.

Figure 4.11 shows the simulated results based on the two different pairs. We can notice that the near field diffraction pattern based on the the FFT/ICZT pair is reasonable and the largest difference between the results of the two algorithms is smaller than 10×10^{-11}, which is extremely small relative to the amplitude, whose largest value is around 1.3.

To validate the performance of the CZT algorithm in relaxing the stringent requirement on the sampling relation between the spatial domain and spatial frequency domain in the angular spectrum method, we use the same object, wavelength and the propagation distance in the first simulation, but scale the FOV of the detector by a factor of 2 and $\frac{1}{2}$. Figure 4.12 demonstrates the result where we can notice a similar aliasing result when the sampling interval of the detector violates the sampling theorem.

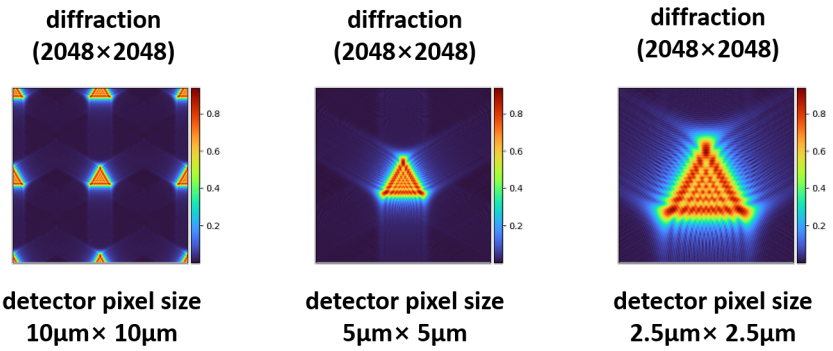


Figure 4.12: The simulated amplitudes of the diffracted fields with different detector sizes based on the angular spectrum method. The aliasing result of the detector pixel size $10\mu m \times 10\mu m$ attributes to the violation of the sampling theorem.

4.4 Fresnel two step propagation

In addition to the CZT algorithm, another method called Fresnel two step propagation is also introduced. Considering the Rayleigh-Sommerfeld diffraction:

$$U(x_2, y_2) = \frac{\Delta z}{i\lambda} \iint U(x_1, y_1) \frac{\exp(ikr)}{r^2} dx_1 dy_1, \tag{4.26}$$

with

$$r = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + \Delta z^2}, \tag{4.27}$$

where (x_1, y_1) and (x_2, y_2) represents the coordinates of the object plane and the diffraction plane, respectively; r represents the distance between two points on the two planes; $k = \frac{2\pi}{\lambda}$ represents the wave number; and Δz represents the vertical distance between the two planes.

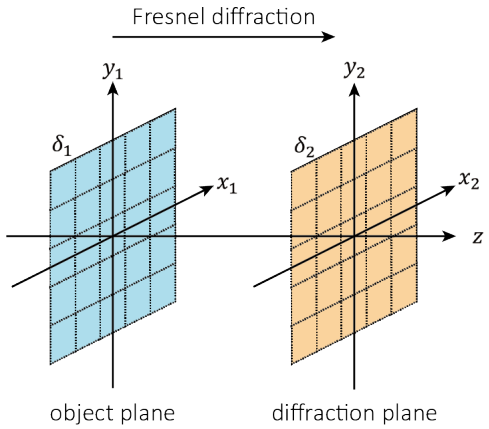


Figure 4.13: Schematic illustration of Fresnel one step propagation (Fresnel method).

In the near-field, we could simplify the distance r as:

$$r \approx \Delta z + \frac{(x_2 - x_1)^2}{2\Delta z} + \frac{(y_2 - y_1)^2}{2\Delta z}, \tag{4.28}$$

with this simplification, we can get the Fresnel diffraction:

$$U(x_2, y_2) = \frac{e^{ik\Delta z}}{i\lambda\Delta z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_1, y_1) e^{i\frac{k}{2\Delta z}[(x_2 - x_1)^2 + (y_2 - y_1)^2]} dx_1 dy_1 \tag{4.29}$$

Expanding the squared terms in the exponential part yields:

$$\begin{aligned} U(x_2, y_2) = & \frac{e^{ik\Delta z}}{i\lambda\Delta z} e^{i\frac{k}{2\Delta z}(x_2^2 + y_2^2)} \\ & \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_1, y_1) e^{i\frac{k}{2\Delta z}(x_1^2 + y_1^2)} e^{-i\frac{2\pi}{\lambda\Delta z}(x_2 x_1 + y_2 y_1)} dx_1 dy_1. \end{aligned} \tag{4.30}$$

Based on Equation 4.30, the diffracted field is computed by multiplying the input field and a quadratic phase ($\exp[i\frac{k}{2\Delta z}(x_1^2 + y_1^2)]$), performing Fourier transform on the product, scaling it by a constant, $(1/(\lambda\Delta z))$, and multiplying the result with another quadratic phase ($e^{i\frac{k}{2\Delta z}(x_2^2 + y_2^2)}$). To write it more compactly, here we use the $\mathcal{R}[\ ]$ to represent these steps of the Fresnel method, and r_1, r_2 represent (x_1, y_1) and (x_2, y_2), respectively, then we have:

$$U(r_2) = \mathcal{R}[\Delta z, r_1, r_2]\{U(r_1)\}. \tag{4.31}$$

The sampling sizes in the object plane (δ_1) and the diffraction plane (δ_2) satisfy:

$$\delta_2 = \frac{\lambda\Delta z}{N\delta_1}. \tag{4.32}$$

However, in the simulation with the fixed sampling at the detector plane as in the real experimental setup, the sampling interval of the object and the sampling interval of the detector may not satisfy this relation. To choose the sampling interval freely, a scaling number m is introduced:

$$m = \frac{\delta_2}{\delta_1}. \tag{4.33}$$

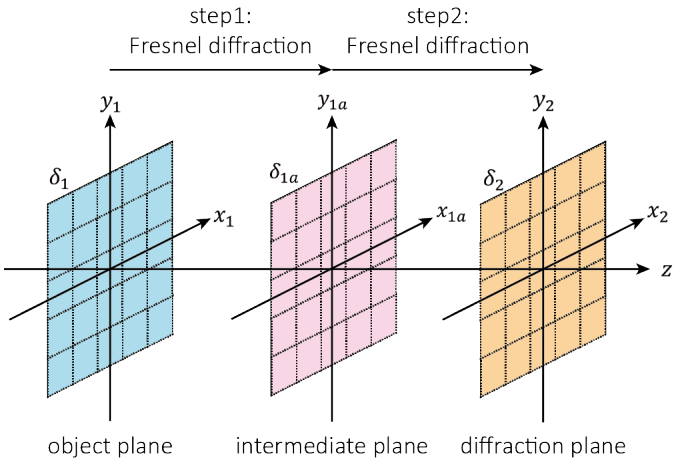


Figure 4.14: Schematic illustration of Fresnel two step propagation. Note that the intermediate plane does not have to be between the object plane and the diffraction plane.

As shown in Figure 4.14, an intermediate plane (x_{1a}, y_{1a}) is introduced into the system. The propagation distance is divided into two parts, the first propagation distance is $\Delta z_1 = z_{1a} - z_1$ and the second propagation distance is $\Delta z_2 = z_2 - z_{1a}$. Using the notation of Equation 4.31, the two step propagation can be expressed as:

$$U(r_2) = \mathcal{R}[\Delta z_2, r_{1a}, r_2]\mathcal{R}[\Delta z_1, r_1, r_{1a}]\{U(r_1)\}. \tag{4.34}$$

Considering the sampling interval of these three planes, we have that:

$$\delta_{1a} = \frac{\lambda|\Delta z_1|}{N\delta_1}, \tag{4.35}$$

$$\delta_2 = \frac{\lambda|\Delta z_2|}{N\delta_{1a}}. \tag{4.36}$$

By introducing Equation 4.35 into Equation 4.36, we can derive the relation between the sampling intervals of the object plane and the diffraction plane:

$$\delta_2 = \frac{\lambda|\Delta z_2|}{N\frac{\lambda|\Delta z_1|}{N\delta_1}} = \left|\frac{\Delta z_2}{\Delta z_1}\right|\delta_1 \tag{4.37}$$

which means the introduced scaled parameter m can be used to determine the location of the intermediate plane:

$$m = \left|\frac{\Delta z_2}{\Delta z_1}\right| = \left|\frac{z_2 - z_{1a}}{z_{1a} - z_1}\right| \tag{4.38}$$

In the test for Fresnel two step propagation, we use a detector with a number of pixels equal to 2048×2048 and pixel size $5\mu m \times 5\mu m$. The wavelength of the light source is 750 nm. The propagation

distance between the aperture and the detector is 50 mm. In order to compare the result of the Fresnel two step propagation with that of the Fresnel method, here we make the pixel size and the number of pixels of the sampling in the object plane to satisfy the requirements of the FFT algorithm (Equation 4.9 and Equation 4.10). An equilateral triangular aperture is used as the object. The side length of the triangular aperture is 30% of the FOV of the object plane.

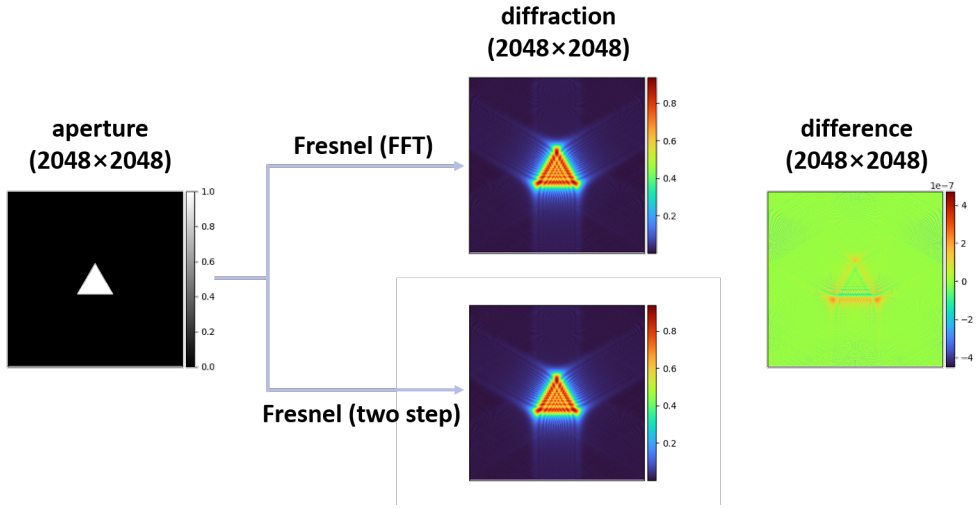


Figure 4.15: Simulated intensities of the diffractive fields based on the angular spectrum method (FFT/IFFT pair) and Fresnel two step propagation. For the convenience of comparison, the difference between the two results is also shown in the last figure.

Figure 4.15 presents the simulation results based on the two different algorithms. We can notice that the diffracted field based on the Fresnel two step propagation is reasonable and the highest difference between the results of the two algorithms is about 4×10^{-7}, which is extremely small relative to the amplitude.

Similar to the simulation in the Fresnel method, we scale the sampling size of the detector by a factor of 2 and $\frac{1}{2}$, respectively, see the results in Figure 4.16.

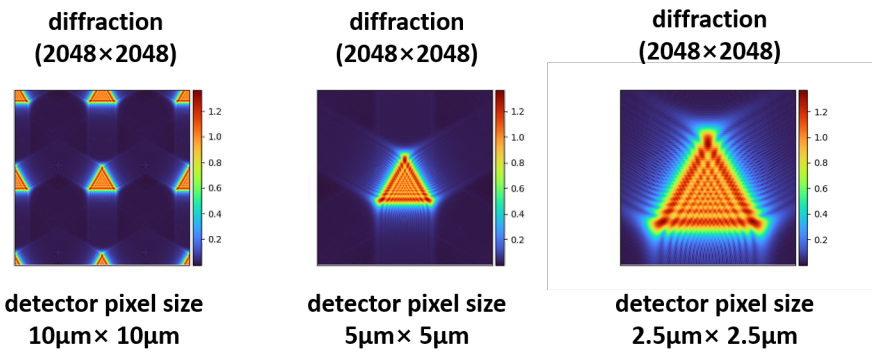


Figure 4.16: The simulated amplitudes of the diffracted fields with different detector sizes based on the Fresnel two step propagation. The aliasing result of the sampling interval $10\mu m \times 10\mu m$ attributes to the violation of the sampling theorem.

4.5 Computational speed test

To select an appropriate algorithm for further reconstruction of a polychromatic data set, we first compare the computational speed of the Fraunhofer method based on the FFT algorithm and the CZT algorithm, respectively. The number of pixels of the input matrices are 500×500, 512×512, 1000×1000, 1024×1024, 2000×2000 and 2048×2048. For the CZT algorithm, similar to the previous simulation, we set the scale factor for interval of the spatial frequency domain to be 0.5, 1 and 2. For each combination of a scale factor and a number of pixels, the test is implemented for 100 times, and the average of the running times will be recorded. The experiment is performed on an Intel i7-8650U CPU and a GeForce RTX 3080 GPU, respectively.

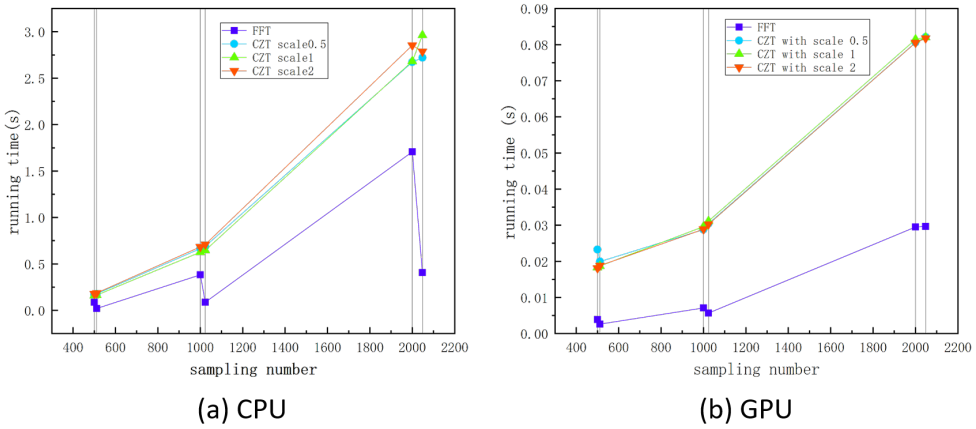


Figure 4.17: Computational speed test for the Fraunhofer method based on the FFT algorithm and the CZT algorithm.

Figure 4.17 demonstrates the runtimes for different combinations of the algorithms and scale

factors on CPU and GPU. It shows that the scale factor has no influence on the computational speed for the same number of pixels on both the CPU and the GPU, which means that the simulations for the same object with different sampling intervals of the detector are equivalently time-consuming. On the CPU, when the number of pixels of the input matrix is power of 2, the computation is accelerated due to the fact that the FFT algorithm can be run efficiently on the combinations of powers of prime numbers (e.g. 2^m, 3^n, 5^k, ...). The CZT algorithm needs two steps of FFT and one step of IFFT, and the padding size is designed to be the exponent of next higher power of 2 compared with the number of pixels of the input matrix. Therefore, when the number of pixels is power-of-two, the computational time of the CZT algorithm is $2^3 = 8$ times of that of the FFT algorithm; when the number of pixels is not power-of-two, the computational speed of the FFT algorithm will be decelerated, even if the number of pixels is close to each other, but the speed of the CZT algorithm is still linear to the number of pixels. Furthermore, due to the architecture, a CPU is ideal for serial instruction processing, while a GPU is designed for parallel instruction processing. The computational test on the GPU is much faster than it is on the CPU. Also, the deterioration due to the non-power-of-two number of pixels is alleviated.

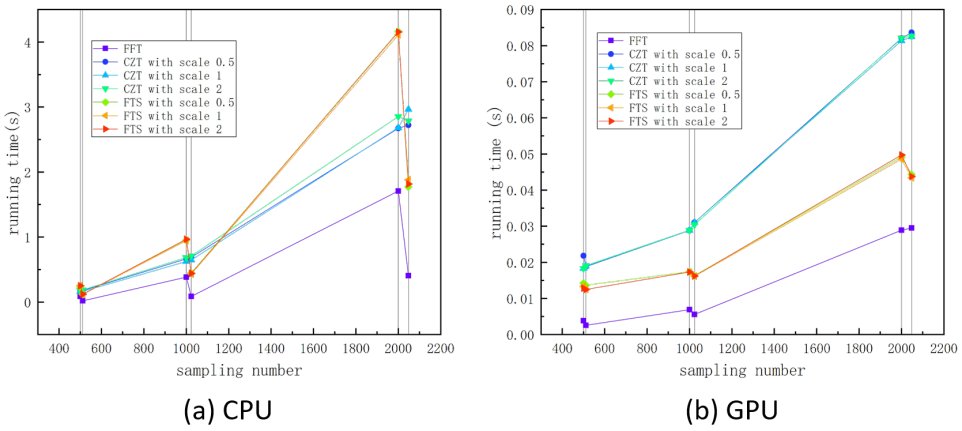


Figure 4.18: Computational speed test for near field diffraction based on the Fresnel method with the FFT algorithm, the CZT algorithm and the Fresnel two step propagation.

A similar test is implemented on the Fresnel method with the FFT algorithm, the CZT algorithm and the Fresnel two step propagation approach. As shown in Figure 4.18, the scale factor still has no influence on the computational speed for the same number of pixels on both the CPU and the GPU, which means the simulations for the same object with different wavelengths are equally time-consuming. Theoretically, the Fresnel two step propagation consists of two steps of the FFT algorithm. The result on the CPU is consistent with this statement, and we can still notice the deterioration due to the non-power-of-two number of pixels on the Fresnel two step propagation. However, due to the improvement of architecture, the deterioration is alleviated on the GPU. Since the CZT algorithm contains three Fourier transform steps, we can notice that Fresnel two step propagation is faster than the CZT algorithm.

Although it seems that the CZT algorithm is theoretically slower than the Fresnel two step propagation, there is some trick to accelerate it during the simulation. For convenience of reading, we write the equation of the CZT here again:

$$F_m = W^{\frac{m^2}{2}} \text{iFFT}[\text{FFT}[f_n A^{-n} W^{\frac{n^2}{2}}]\text{FFT}[W^{-\frac{k^2}{2}}]], \quad m = 0, 1, ..., M-1, \tag{4.39}$$

where

$$\begin{cases} A = e^{i2\pi u_0 \Delta x} \\ W = e^{-i2\pi \Delta u \Delta x}. \end{cases} \tag{4.40}$$

Based on the formula, the CZT algorithm requires 3 FFT steps. However, during the ptychographic iterations, the A and W are fixed for one combination of the wavelength and the propagation distance, and f_n (exit wave) is the only one that will be updated. Therefore, the Fourier transform of the W, FFT$[W^{-\frac{k^2}{2}}]$, can be saved in memory as a fixed term after the first iteration. The following calculation can use this term without redundant calculations. Then the calculation of the CZT algorithm only requires 2 FFT steps, which is equal to that of the Fresnel two step propagation.

In conclusion, both of the CZT algorithm and the Fresnel two step propagation can be used to relax the stringent requirement on the sampling relation between the spatial domain and spatial frequency domain. By saving the term FFT$[W^{-\frac{(m-n)^2}{2}}]$ in the CZT algorithm as a fixed term to be reused in the iterative reconstructions, these two algorithms are compatibly fast.

4.6 Correction of initial error

Ptychography has been proven to be able to correct misalignment and calibrate experimental setups. Lateral position correction algorithms have been reported to mitigate the effect of sample drift and errors in the encoders of xy translation stages. Except the encoders, axial uncertainty in the sample-detector distance or a wrong wavelength results in scaling of the reconstruction pixel size and as such introduces spatially dependent artifacts.

Most existing solutions for correcting these initial errors rely on using a sharpness metric to evaluate the diffraction pattern and make them focusing on the correct axial position. However, in ptychography with the AD framework, these correction can be implemented easily by setting these values as trainable variables, only if the gradients of these variables can be calculated. Before we developed the new propagators, once the pixel size of the sample and that of the detector are determined, the product of the propagation distance and the wavelength is fixed, which means these parameters cannot be updated or corrected. Since the newly developed propagators can relax the stringent requirement on the sampling relation, we can implement the optimization for propagation distance and the wavelength. In other words, we can make the object autofousing on the correct plane. It is worth to notice that, according to Equation 4.8, the propagation distance and the wavelength are multiplied in the physical model, making them inseparable during the optimization. In other words, if both of the propagation distance and the wavelength are incorrect, both of them may not converge to exact correct values, but the product will be the correct value.

Generally, the range of the amplitude of the object and the probe is two digits, but that of the propagation distance and the wavelength is relatively small, which means that the learning rates

for different parameters should be different. To select a proper learning rate, we implemented a reconstruction on an experimental data set. In the first 20 epochs (red rectangle in Figure 4.19), the sample and the probe are reconstructed with the correct detector-sample distance. The learning rate for the sample and the probe are both 0.1. Then we take the reconstructed sample, the reconstructed probe and a wrong distance as the initial point for the next 20 epochs of the reconstruction with different learning rates.

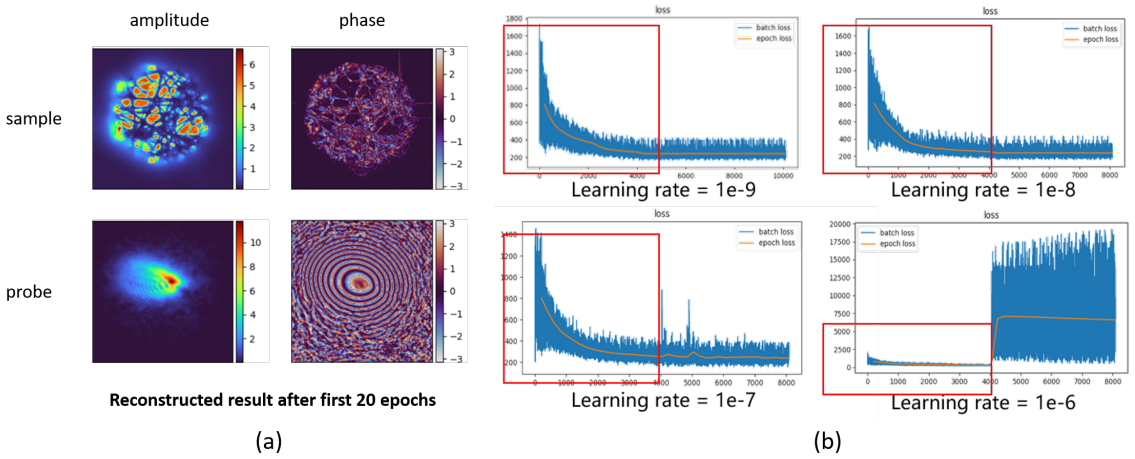


Figure 4.19: Test for different learning rates. (a) Reconstructed sample and probe after first 20 epochs. (b) Subsequent optimization of distance with different learning rate.

As shown in Figure 4.19, if the learning rate for the correction of the propagation distance is too large, the loss function will extremely increase and the reconstructed distance will deviate from the correct value. From this test, we can imply that the proper learning rate for propagation distance is relatively small ($< 1 \times 10^{-6}$) compared to the ones for the sample and the probe (0.1).

With a suitable learning rate for the distance and the wavelength, another reconstruction test is implemented on another experimental data set, which uses an etched number three with some point stains as the sample. The wavelength of the light source is 561 nm.

To validate the performance of the algorithm, we implement two reconstructions on this data set. The initialization of the sample and the probe for the two reconstructions are similar to the previous reconstructions, but the input propagation distance is wrong. One of these two reconstructions uses the propagator based on the FFT algorithm (without autofocusing) and the other one uses the the propagator based on the CZT algorithm (with autofocusing).

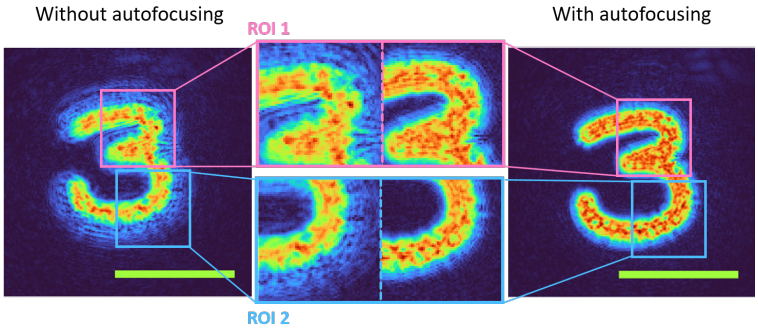


Figure 4.20: Ptychographic reconstruction without correction of propagation distance (FFT algorithm) and with correction based on the new propagators (CZT algorithm). The reconstructed sample based on the CZT algorithm is more clear than that based on the FFT algorithm.

Figure 4.20 demonstrates the reconstructed sample with and without correction. The reconstructed sample without distance correction suffers from an out of focus effect, demonstrating a blurred edge. But the sample with the correction of distance shows a clear edge, and the propagation distance converges to the correct value after several iterations. In order to make the improvement of the reconstructed quality quantifiable, we perform the reconstruction on the data set with a correct propagation distance. We set the reconstructed sample from this reconstruction as reference, and introduce Fourier ring correlation (FRC) to evaluate the similarity between the reference and the reconstruction of the same object with/without correction in propagation distance.

FRC measures the degree of correlation of the two images at different spatial frequencies [38]. As shown in Figure 4.21, in order to calculate the Fourier ring correlation between two images, these two images are Fourier transformed and multiplied. The normalized average correlation is computed for several concentric rings of increasing radius (corresponding to increasing spatial frequencies) centered around the (0,0) spatial frequency. The higher the similarity between the two images, the greater the spatial frequency corresponding to the point that falls below the threshold.

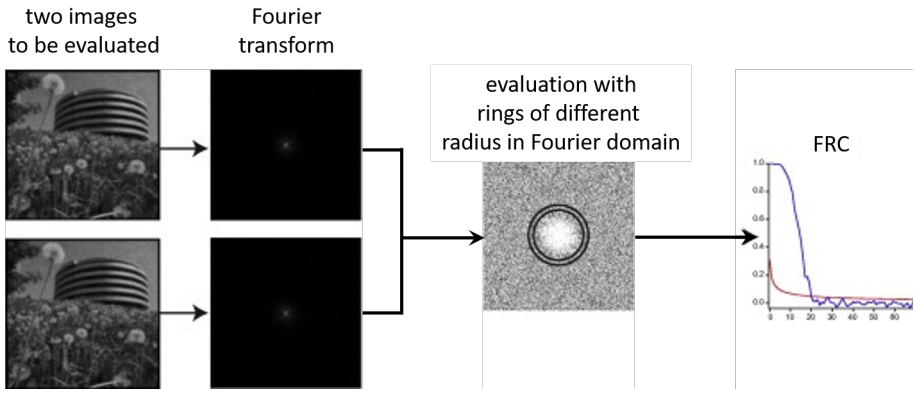


Figure 4.21: Workflow used for evaluating the similarity between two images through FRC.

As shown in Figure 4.22, the reference sample is more similar to the autofocusing result (crossing the threshold at 6.8539e-05m in spatial domain) than it is to the result without autofocusing(crossing the threshold at 2.6522e-04m in spatial domain).

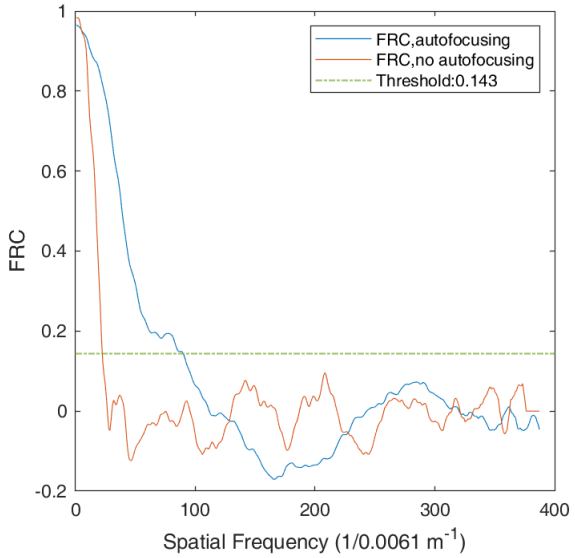


Figure 4.22: FRC plot of the reconstructed sample with/without initial errors. As shown in the figure, the reference sample is more similar to the autofocusing result than it is to the result without autofocusing.

Considering the previous analysis, the wavelength and the propagation distance are inseparable in the physical model, implying that the product of these two parameters should converge to the correct value, independent of the initial error. To further validate this statement, we execute another two separate reconstructions on this data set. Different from the previous reconstruction, we set the input propagation distance and the input wavelength to be wrong for these two reconstructions, respectively.

Figure 4.23 demonstrates the reconstruction results for the two different simulations. Both of the results show a clear reconstruction for the sample and the probe. The products of the propagation distance and the wavelength in these two tests converge to $3.877769e-8m^2$ and $3.877288e-8m^2$, respectively. The percentage difference of these two values is:

$$diff = \frac{(\lambda z)_1 - (\lambda z)_2}{\frac{(\lambda z)_1 + (\lambda z)_2}{2}} \times 100\% = 0.0124\% \tag{4.41}$$

which is small enough to be neglected. Therefore, we think the correction is successful.

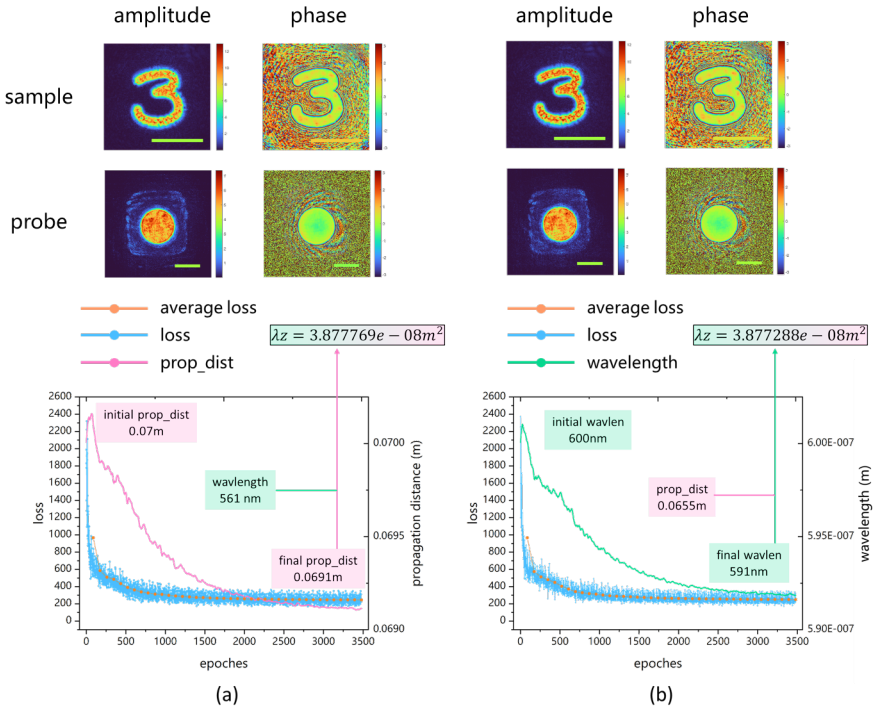


Figure 4.23: Reconstruction result for the correction of the propagation distance and wavelength. (a) Correction of propagation distance. (b) Correction of wavelength. The scale bar indicates a length of 1mm. Both of them can reconstruct a clear sample and probe. The product of the propagation distance and the wavelength can converge to the same value for the two different initial error, which means our correction is successful.

The data set used before is collected in the near field. To test the algorithm in the far field, we utilized another experimental data set. The sample here is a piece of some lignocellulosic fibrous material (pulp, the material for making paper). We perform the reconstruction three times: (a) no initial error in propagation distance (b) correct the initial error with the CZT algorithm (c) no correction for the initial error. The reconstructed results are shown in Figure 4.24. All the three reconstruction can converge to a reasonable result. The shifts of the sample and the probe have no influences on the resolution, since only the relative positions of the sample and the probe are important for ptychography. The scaling effect on the result (c) attributes to the wrong discrete pixel size in the object domain.

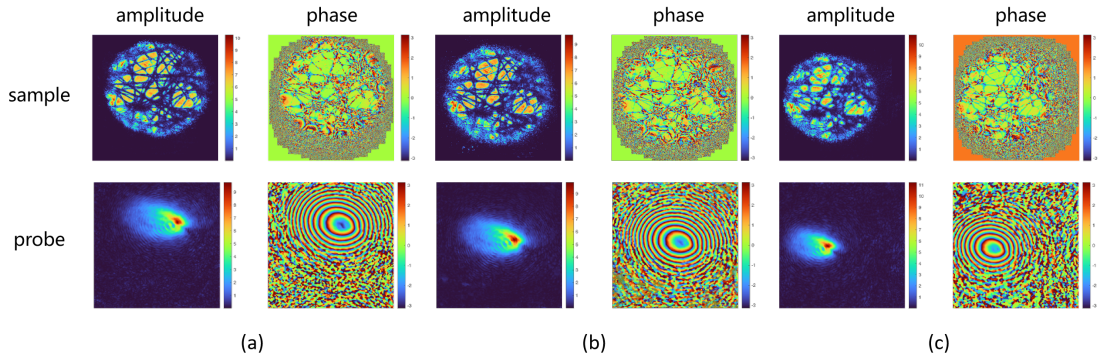


Figure 4.24: Ptychographic reconstruction results: (a) no initial errors in propagation distance; (b) correct the initial errors with the propagator based on the CZT algorithm; (c) no correction for the initial errors.

Figure 4.25 shows three selected ptychographic reconstructed regions of interest (ROI) across the FOV. Similar to the previous comparison, we take the reconstructed sample with the correct input propagation distance as the reference and use FRC to evaluate the similarity between the reference and the other reconstructed samples. The blurring effect on the reconstructed sample due to the wrong propagation distance is significantly alleviated, which is clearly shown in both the enlarged samples and the FRC evaluation.

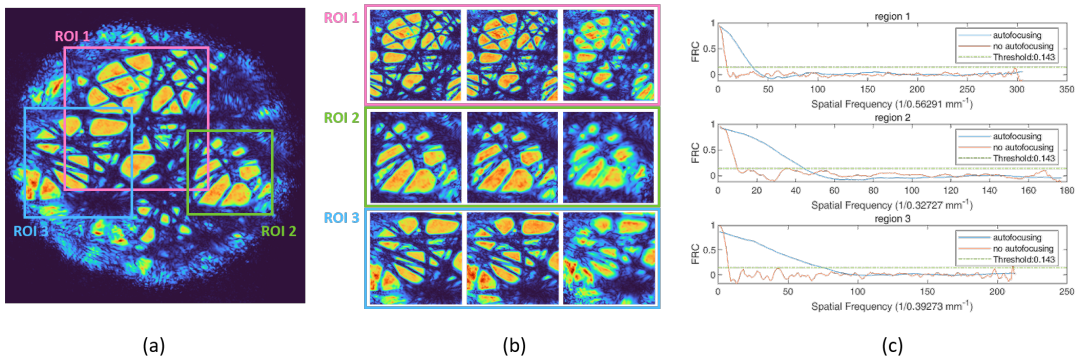


Figure 4.25: (a) Reconstructed sample without initial errors. (b) Comparison of the reconstructed sample of three different situations, from left to right are the three situations of a,b,c mentioned in Figure 4.24. (c) Evaluation of the similarity between the autofocusing/no-autofocusing images and the correct reconstructed image of different regions. The reconstruction results without the initial error are much more similar to the autofocusing results with the initial errors than they are to the results without autofocusing.

As we mentioned previously, the product of the sample-detector distance and the wavelength can converge to the same value, no matter where we introduce the initial errors.

Similar to the previous test in the near field region, initial errors are introduced to the propagation distance and the wavelength, respectively. Figure 4.26 shows the results. The percentage difference of these two products is 0.0135%, which means that our algorithm also works in the far field region.

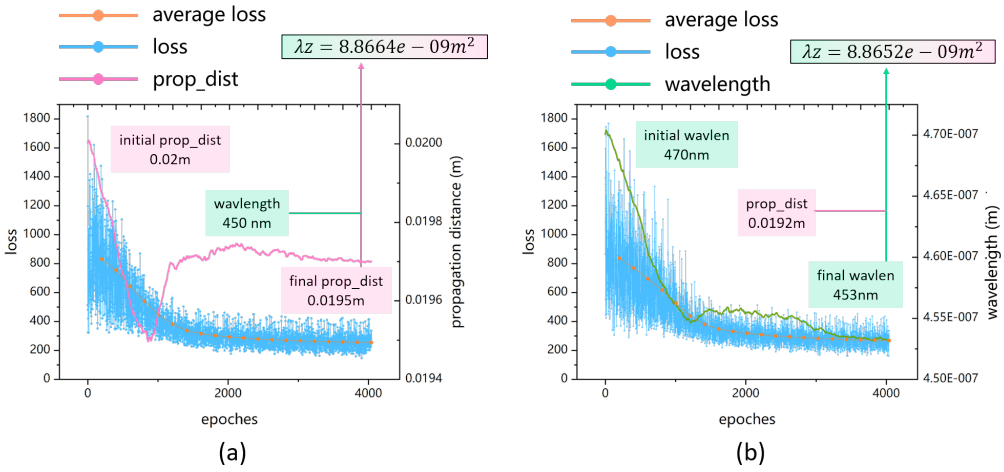


Figure 4.26: Plot of the loss function, sample-detector distance (a) and the wavelength (b) during the optimization. The products of the distance and the wavelength starting from different initial errors can converge to the same value, proving that our algorithm is successful.

4.7 Polychromatic ptychography

To test performance of the newly developed propagators based on the CZT algorithm in dealing with polychromatic illumination, we first execute the ptychographic simulation, where we use the physical forward model to emulate the data acquisition of the diffraction patterns with some known object and probe first; and next, we take the scanning positions and the corresponding simulated diffracted fields as the input for the ptychographic algorithm to execute the reconstruction. Here we used the image of the boat and Lenna as the amplitude and the phase of the sample, respectively. The probes with the two wavelengths are two Gaussian beams with different sizes. And the scanning positions are evenly spread over the rectangular FOV.

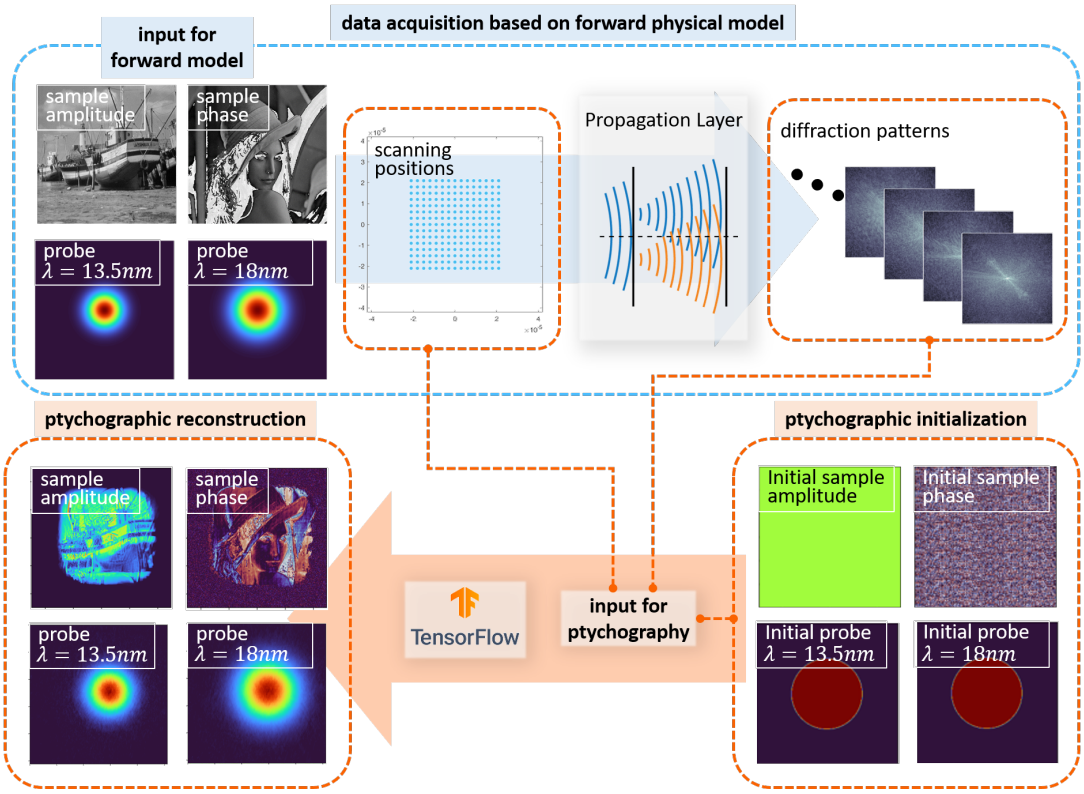


Figure 4.27: Simulation of ptychographic reconstruction with two wavelengths.

Figure 4.27 demonstrates the workflow for the ptychographic simulation of polychromatic illumination. In this simulation, both the sample and the probes are reconstructed clearly.

To further test the performance of our algorithm, we execute the reconstruction on an experimental data set. The experiment uses a designed spiral as the sample, and the illumination probe contains 9 wavelengths: 52.0nm, 46.7nm, 42.0nm, 38.2nm, 35.0nm, 32.0nm, 29.5nm, 27.4nm and 25.4nm. The ptychographic scan consists of 252 scan positions covering a FOV with a diameter of 50 μm. The average scan step is 2.5μm and the diameter of the binary spiral is 25μm, resulting in an overlap of approximately 90%. The initialization for the sample and the probe is similar to the previous test in the session about the multimodal ptychography. The initial spectral weight of the probe is initialized randomly, but ensured to sum up to 1.

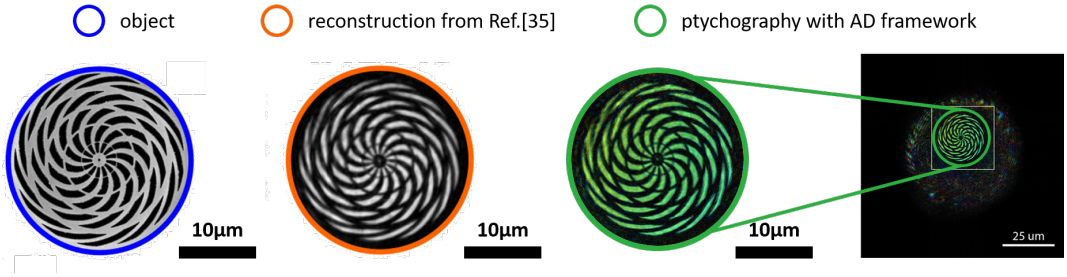


Figure 4.28: A spiral with diameter of 25 μm is used as the sample (blue circle). The reconstructed sample from Ref.[35] (orange circle) and our reconstructed sample (green circle) are compatibly clear to show the details of the object.

As shown in Figure 4.28, our reconstructed sample (green circle) is compatible with the result of Ref.[35] (orange circle). Also, in terms of the probe in Figure 4.29, good qualitative agreement between our reconstruction and the measurements is observed in the spectral weights, especially in the central harmonics. Compared with the result from Ref.[35], we have smaller discrepancy of the weight at the extremal spectral lines. As explained in Ref.[35], the differences between the simulated profiles of the simulated probes and the reconstructed probes are attributed to the fabrication errors of the experimental elements. Furthermore, to converge to such a compatible result, our algorithm only takes about 40 minutes, which is much faster than the algorithm in Ref.[35] (usually takes several hours).

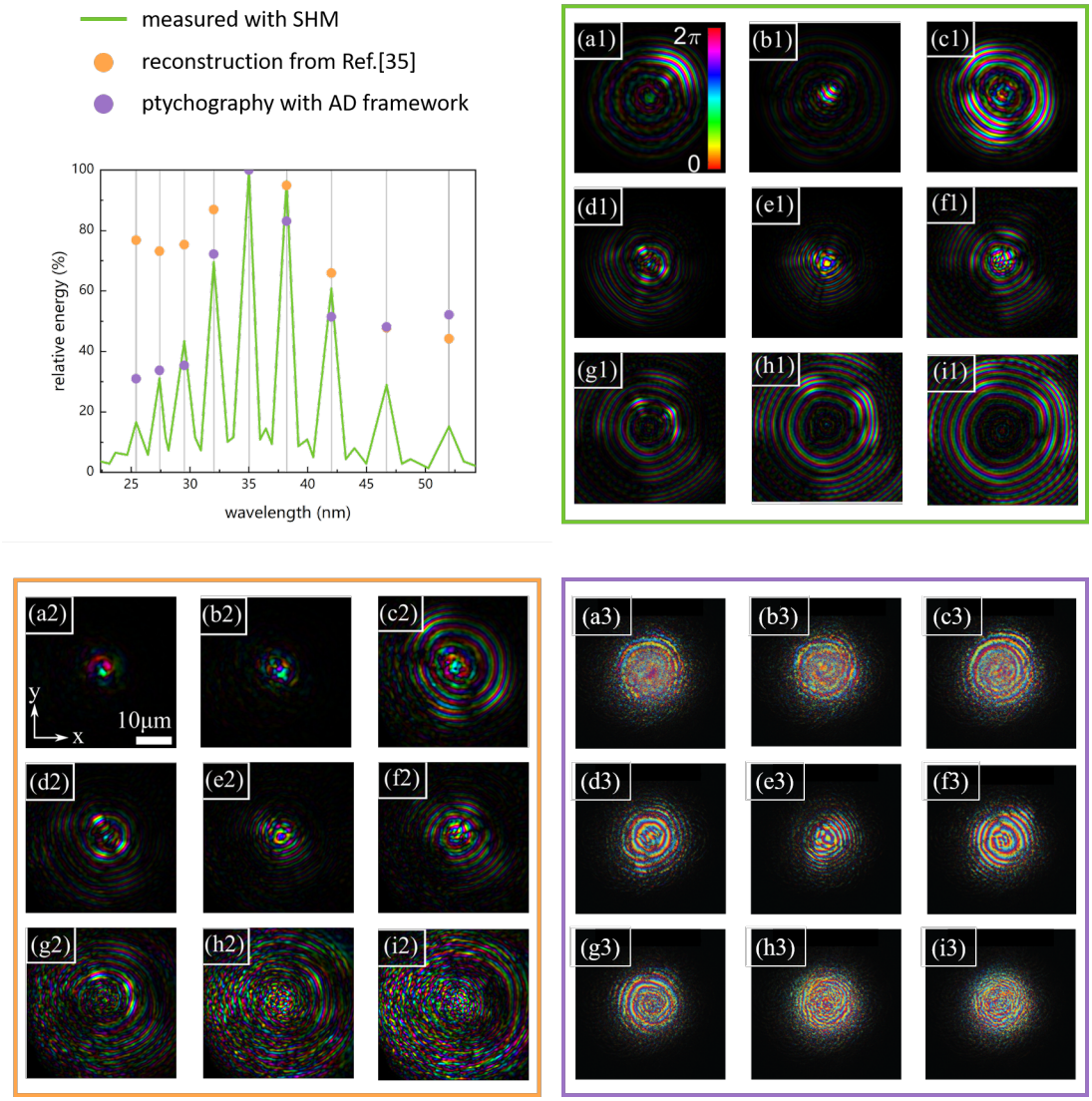


Figure 4.29: The spectrum of the source is extracted with the spectral Hartmann mask (SHM) during the experiment. Although at the extreme spectral lines, both of reconstructed weights from Ref.[35] and our result show some discrepancy, our result is still closer to the measured weights. As discussed in Ref.[35], the difference between the simulated beam profiles (a1-i1) and the reconstructed beam profiles (a2-i2) attributes to the fabrication errors of the experimental elements. Our reconstructed probes (a3-i3) shares the same size due to the use of the CZT algorithm. In (a2)-(i2), the spot size is not steadily increasing due to some discrepancy in the algorithm of Ref.[35].

Chapter 5

Conclusion and outlook

5.1 Conclusion

In this thesis, several different subjects about ptychography with the automatic differentiation (AD) framework were discussed. The main goals of this master thesis were the following:

- Solving the ambiguity in the multi-mode ptychography to deal with partially coherent illumination.
- Rewriting the structure of the platform to correct for initial errors in the propagation distance and wavelength.
- Introduce flexible propagators to deal with ptychography with polychromatic illumination.

It can be concluded that:

Firstly, to relax the stringent requirement on the illumination source and deal with a dynamic sample, multimodal ptychography decomposes the mutual coherence function (MCF) into several modes, which are incoherent with each other. Each mode can be propagated coherently and added up incoherently to obtain the final diffracted field. In Chapter 3, we have introduced the forward physical model of multimodal ptychography and discussed the origin of the ambiguity, which attributes to the modulus constraints where the only restrictions are about the intensity of the diffraction field. Consequently, the reconstructed modes can vary with different initialization. To break this ambiguity, an orthogonalization solution is introduced into the algorithm, which could obtain the unique orthogonal solution of the modes. This proposed solution has been validated on an experimental data set with a partially coherent illumination source, which is consistent with our expectation.

Secondly, to correct the misalighment and calibrate the experimental setup, several algorithms have been proposed based on the sharpness metric evaluating the simulation of the diffracted field. In Chapter 4, the sampling requirement due to the FFT algorithm has been discussed, which makes it impossible to train the propagation distance and the wavelength during the optimization. Then we introduced the CZT algorithm to replace the FFT algorithm relaxing the sampling relation

between the object plane and the detector plane. Inspired by the other research papers in the literature about optical simulation, a Fresnel two step propagator has also been introduced. The proposed correction algorithm is validated on an experimental data set. And the final results with different initialization can converge to the same solution. The sample and the probe function can also be reconstructed clearly.

Finally, the physical forward model of ptychography has been evolved to multimodal ptychography with polychromatic illumination. Different from the existing solutions, such as scaling the object plane during the iterations, an algorithm with some flexible propagators has been used in our platform. The algorithm has been validated on an experimental data set with a source comprising 9 individual wavelengths.

5.2 Outlook

For the further research about our algorithm, I have several recommendations:

- Introducing a rotation matrix into the forward physical model to deal with a correction for sample tilt in reflective ptychography.
- Using better initial guess for some data set, when it is difficult to converge.
- Introducing some neural network elements into the platform to improve converge.
- Using polychromatic ptychography to execute nondestructive fault diagnosis of multi-layer samples.

Firstly, to deal with the initial error of the product of propagation distance and wavelength, the flexible propagators make them trainable (or differentiable) with fixed sampling conditions of the sample and the detector. One of the significant benefit of our platform is that many different parameters can be introduced into the forward physical model, which makes it a powerful correction approach. To use some highly absorbable illumination source, such as EUV light, the reflective ptychography is significant. In the experiment with a reflective sample, tilting the sample introduces a nonlinear coordinate warping in the observed diffraction data, parameterized by the relative angle between the specimen and the detector. To calibrate this tilt angle, an algorithm called aPIE has been proposed recently [39]. However, we can also include the rotation matrix into the physical model, and make the tilt angle trainable during the iterations to find the correct angle.

Secondly, for some complex experimental situations, such as polychromatic illumination, it is difficult to converge to the global minimum or an acceptable local minimum. To avoid the over-fitting problem, several regularization methods have been introduced into the platform to improve the performance, such as L1 regularization to reduce the background noise and total variance to smooth the reconstructed image and emphasize edges [40]. However, for some extreme situations, it is necessary to relax our requirement. For the reconstruction in Ref.[35], the authors have used use a perfect initial guess of the sample as the starting point, which makes their profile of the probes smooth in the final result, which suggests us, that if the profile of the probe is important for some reconstruction, a better initial guess of the sample can be used in the initialization.

Furthermore, the essence of ptychography with the AD framework is building the forward physical model and obtaining the gradient based on the principle of automatic differentiation. The

procedure is taking advantage of the deep learning packages, but is still based on the physical model. With the rapid development of the neural network research, some ptychography algorithms are totally based on some existing deep learning model. However, since its internal parameters often do not have clear physical meanings, this method is questionable in terms of generality and accuracy. To combine the advantages of the neural networks and the physical essence of our model, we can introduce some layers based on neural network research into our model. For example, an auto-encoder can reduce the complexity of the input data, which can be used to reduce the complexity of the sample and the probe. Then we may be able to save the computational resource and accelerate the convergence.

In terms of the application of the ptychography algorithm, it can reduce the high requirements on the experimental set up and facilitate the flux efficiency as compared to lens-based microscopy setups, especially with the automatic differentiation framework. With the rapid development in the tabletop EUV sources, the solution we proposed here to deal with the multi-mode ambiguity and polychromatic illumination can enhance its robustness and flexibility. A useful application for polychromatic ptychography is silicon chip inspection. The chip can be regarded as a multi-layer sample, containing different materials (e.g. mental layer, silicon layer and adhesive). Different materials have different transmission/reflection coefficients at different wavelengths, for example, silicon is opaque in the visible spectrum, but the adhesive is semitransparent. We can utilize polychromatic illumination in ptychography to obtain different information under different wavelengths. Also, we can use polychromatic ptychography to detect the defects in multi-layer mirror and extract the amplitude and the phase damage of the defect. With these information, we can execute nondestructive fault diagnosis of samples such as microelectromechanical systems (MEMS) devices, heavily doped silicon samples, wafer bonding, and 3D chip stacks.

References

[1] R. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

[2] J. Fienup, "Reconstruction of an object from modulus of its fourier transform," *Optics letters*, vol. 3, pp. 27–9, Aug. 1978.

[3] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug. 1982.

[4] H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm," *Phys. Rev. Lett.*, vol. 93, p. 023 903, 2 Jul. 2004.

[5] W. Hoppe, "Beugung im inhomogenen primärstrahlwellenfeld. i. prinzip einer phasenmessung von elektronenbeungungsinterferenzen," *Acta Crystallographica Section A*, vol. 25, pp. 495–501, 1969.

[6] R. H. T. Bates and J. M. Rodenburg, "Sub-ångström transmission microscopy: A fourier transform algorithm for microdiffraction plane intensity information," *Ultramicroscopy*, vol. 31, pp. 303–307, 1989.

[7] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009, ISSN: 0304-3991.

[8] P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," *Nature*, vol. 7435, pp. 68–71, 2013.

[9] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson, "Dynamic imaging using ptychography," *Phys. Rev. Lett.*, vol. 112, p. 113 901, 11 Mar. 2014.

[10] P. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel, "On-the-fly scans for x-ray ptychography," *Applied Physics Letters*, vol. 105, p. 251 101, Dec. 2014.

[11] B. Enders, M. Dierolf, P. Cloetens, M. Stockmar, F. Pfeiffer, and P. Thibault, "Ptychography with broad-bandwidth radiation," *Applied Physics Letters*, vol. 104, no. 17, p. 171 104, 2014.

[12] D. J. Batey, D. Claus, and J. M. Rodenburg, "Information multiplexing in ptychography," *Ultramicroscopy*, vol. 138, pp. 13–21, 2014, ISSN: 0304-3991.

[13] X. Shi, N. Burdet, D. Batey, and I. Robinson, "Multi-modal ptychography: Recent developments and applications," *Applied Sciences*, vol. 8, no. 7, 2018, ISSN: 2076-3417.

[14] S. Cao, P. Kok, P. Li, A. M. Maiden, and J. M. Rodenburg, "Modal decomposition of a propagating matter wave via electron ptychography," *Phys. Rev. A*, vol. 94, p. 063 621, 6 Dec. 2016.

[15] M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with transverse translation diversity: A nonlinear optimization approach," *Opt. Express*, vol. 16, no. 10, pp. 7264–7278, May 2008.

[16] A. S. Jurling and J. R. Fienup, "Applications of algorithmic differentiation to phase retrieval algorithms," *J. Opt. Soc. Am. A*, vol. 31, no. 7, pp. 1348–1359, Jul. 2014.

[17] Y. S. Nashed, T. Peterka, J. Deng, and C. Jacobsen, "Distributed automatic differentiation for ptychography," *Procedia Computer Science*, vol. 108, pp. 404–414, 2017, International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland, ISSN: 1877-0509.

[18] S. Ghosh, Y. S. G. Nashed, O. Cossairt, and A. Katsaggelos, "Adp: Automatic differentiation ptychography," in *2018 IEEE International Conference on Computational Photography (ICCP)*, 2018, pp. 1–10.

[19] S. Kandel, S. Maddali, M. Allain, S. O. Hruszkewycz, C. Jacobsen, and Y. S. G. Nashed, "Using automatic differentiation as a general framework for ptychographic reconstruction," *Opt. Express*, vol. 27, no. 13, pp. 18 653–18 672, Jun. 2019.

[20] K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, *et al.*, "Flexible ptychography platform to expand the potential of imaging at free electron lasers," *Opt. Express*, vol. 29, no. 14, pp. 22 345–22 365, Jul. 2021.

[21] J. Seifert, D. Bouchet, L. Loetgering, and A. P. Mosk, "Efficient and flexible approach to ptychography using an optimization framework based on automatic differentiation," *OSA Continuum*, vol. 4, no. 1, pp. 121–128, Jan. 2021.

[22] A. Jucha, D. Bonin, E. Dartyge, A. Flank, A. Fontaine, and D. Raoux, "Photodiode array for position-sensitive detection using high x-ray flux provided by synchrotron radiation," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 226, no. 1, pp. 40–44, 1984, ISSN: 0168-9002.

[23] M. Olbinado, J. Grenzer, P. Pradel, *et al.*, "Advances in indirect detector systems for ultra high-speed hard x-ray imaging with synchrotron light," *Journal of Instrumentation*, vol. 13, no. 04, pp. C04004–C04004, Apr. 2018.

[24] H. Carstens, M. Högner, T. Saule, *et al.*, "High-harmonic generation at 250mhz with photon energies exceeding 100ev," *Optica*, vol. 3, no. 4, pp. 366–369, Apr. 2016.

[25] D. E. Rivas, B. Major, M. Weidman, *et al.*, "Propagation-enhanced generation of intense high-harmonic continua in the 100-ev spectral region," *Optica*, vol. 5, no. 10, pp. 1283–1289, Oct. 2018.

[26] R. Klas, S. Demmler, M. Tschernajew, *et al.*, "Table-top milliwatt-class extreme ultraviolet high harmonic light source," *Optica*, vol. 3, no. 11, pp. 1167–1170, Nov. 2016.

[27] M. D. Seaberg, B. Zhang, D. F. Gardner, *et al.*, "Tabletop nanometer extreme ultraviolet imaging in an extended reflection mode using coherent fresnel ptychography," *Optica*, vol. 1, no. 1, pp. 39–44, Jul. 2014.

[28] L. Loetgering, S. Witte, and J. Rothhardt, "Advances in laboratory-scale ptychography using high harmonic sources," *Opt. Express*, vol. 30, no. 3, pp. 4133–4164, Jan. 2022.

[29] B. Enders, M. Dierolf, P. Cloetens, M. Stockmar, F. Pfeiffer, and P. Thibault, "Ptychography with broad-bandwidth radiation," *Applied Physics Letters*, vol. 104, no. 17, p. 171 104, 2014.

[30] M.Stockmar, P. Cloetens, I. Zanette, *et al.*, "Near-field ptychography: Phase retrieval for inline holography using a structured illumination," *Scientific Reports*, vol. 3, 1 May 2013.

[31] P. Li, T. Edo, D. Batey, J. Rodenburg, and A. Maiden, "Breaking ambiguities in mixed state ptychography," *Opt. Express*, vol. 24, no. 8, pp. 9038–9052, Apr. 2016.

[32] E. Wolf, "New theory of partial coherence in the space–frequency domain. part i: Spectra and cross spectra of steady-state sources," *J. Opt. Soc. Am.*, vol. 72, no. 3, pp. 343–351, Mar. 1982.

[33] J. Barolak, D. Goldberger, J. Squier, Y. Bellouard, C. Durfee, and D. Adams, "Wavelength-multiplexed single-shot ptychography," *Ultramicroscopy*, vol. 233, p. 113 418, 2022, ISSN: 0304-3991.

[34] E. Malm, E. Fohtung, and A. Mikkelsen, "Multi-wavelength phase retrieval for coherent diffractive imaging," *Opt. Lett.*, vol. 46, no. 1, pp. 13–16, Jan. 2021.

[35] L. Loetgering, X. Liu, A. C. C. D. Beurs, *et al.*, "Tailoring spatial entropy in extreme ultraviolet focused beams for multispectral ptychography," *Optica*, vol. 8, no. 2, pp. 130–138, Feb. 2021.

[36] J. W. Goodman, *Introduction to Fourier optics*, 6th ed. New York: McGraw-Hill, 1996.

[37] L. R. Rabiner, R. W. Schafer, and C. M. Rader, "The chirp z-transform algorithm and its application," *Bell System Technical Journal*, vol. 48, pp. 1249–1292, 1969.

[38] W. O. Saxton and W. Baumeister, "The correlation averaging of a regularly arranged bacterial cell envelope protein," *Journal of Microscopy*, vol. 127, no. 2, pp. 127–138, 1982.

[39] A. de Beurs, L. Loetgering, M. Herczog, M. Du, K. S. E. Eikema, and S. Witte, "Apie: An angle calibration algorithm for reflection ptychography," *Opt. Lett.*, vol. 47, no. 8, pp. 1949–1952, Apr. 2022.

[40] P. Rodríguez, "Total variation regularization algorithms for images corrupted with different noise models: A review," *JECE*, vol. 2013, Jan. 2013, ISSN: 2090-0147.